Comparison of Estimation Techniques for a Forest Inventory in which Double Sampling for Stratification Is Used

Michael S. Williams

ABSTRACT. A number of different estimators can be used when forest inventory plots cover two or more distinctly different condition classes. In this article the properties of two approximate Horvitz-Thompson (HT) estimators, a ratio of means (RM), and a mean of ratios (MR) estimator are explored in the framework of double sampling for stratification. Relevant theoretical properties are given, and a simulation study is used to compare the properties of these estimators. It is concluded that the variance of the RM estimator is smaller than that of the approximate HT for the populations studied, but that of the MR estimator is larger. While ratio estimators are known to be biased for small sample sizes, an additional bias also occurs when the population of interest has an edge-effect. The RM classical variance estimator underestimated the true variance in the simulation study. The approximate HT estimators exhibited biases that were generally smaller than those of the RM and MR estimators for populations with an edge-effect. Its classical variance estimator was essentially unbiased. FOR. SCI. 47(4):563–576.

Key Words: Ratio of means estimator, mean of ratios estimator, Horvitz-Thompson estimator.

MANY OF THE LARGE-SCALE FOREST INVENTORIES conducted in the United States and other countries use some form of double sampling. One example is the Forest Inventory and Analysis (FIA) program in the United States, which has monitored the condition of forest resources of the nation for the last 70 yr. This program has taken measurements on approximately 2,500,000 photo-interpreted points and over 120,000 fixed-area cluster plots across all forested land. These data are used to periodically generate estimates of forest resources for attributes such as area, volume, and number of trees. Changes in the forest since the previous inventory are also estimated (e.g., growth, mortality, removals, and area converted to other uses). The estimates of all attributes are generated for all forested areas. However, of equal importance are estimates for various subpopulations, which can be based on either individual tree or land characteristics. The term *domain* will be used to refer to subpopulations determined by individual tree characteristics (e.g., tree size, presence of disease or conks, species). The term *condition class* will refer to area-based subpopulations and are classes that can be mapped on the ground (e.g., stand size, productivity class, forest type, ownership).

Scott and Bechtold (1995) and Williams and Schreuder (1995) present a ratio of means (RM) and a Horvitz-Thompson (HT) estimator, respectively, but these estimators and their relative merits have not been adequately compared. This article outlines some protocols for double sampling for stratification, highlights some of the assumptions that are used, introduces a mean of ratios (MR) estimator, and compares all three estimators. Properties addressed in the comparison are:

- relevant theoretical properties,
- bias and variance of the estimators, and
- effects of choice of estimator on sampling protocol and inventory compilation.

Results of a simulation study on a small mapped population are used to illustrate the bias and efficiency of each estimator.

Michael S. Williams, Rocky Mountain Research Station, USDA Forest Service, 2150 A Center Drive, Fort Collins, Colorado 80526-2098—Phone: (970) 295-5974; Fax: (970) 295-5927; E-mail: usdafs@lamar.colostate.edu.

Acknowledgments: The author thanks the Editor, Associate Editor, anonymous reviewers, Stan Zarnoch, and Bill Bechtold for their patience, insightful comments, and suggestions.

Manuscript received March 4, 1999. Accepted November 29, 2000.

Description of the Sampling Design

Two common approaches exist for areal sampling designs, which will be referred to as the tessellated and continuous population approaches. The former approach assumes that an area frame (Sarndal et al. 1992, p. 12) is constructed by tessellating the land area into N nonoverlapping plots of equal area, from which a probability sample of plots is selected. Given this area frame, estimates can be derived using finite population cluster sampling theory (cf. Cochran 1977, Chap. 9 and 10, Sarndal et al. 1992, Chap. 3). One problem with this approach is that circular plots, which are used in virtually all forest inventories, cannot be used to divide a forest into nonoverlapping plots.[1] The consequence of this fact is that a positive probability of inclusion cannot be assigned to every element of the population. Despite its lack of statistical rigor, this approach is still quite popular.

That continuous population sampling differs from finite population cluster sampling is a concept that dates back to at least the 1950s. Pielou (1977, p. 115) describes three different relationships that can exist between the area sampled and the organisms within that area. For the first relationship, the land area is divided into sampling units by some mechanism, such as roads that define city blocks or individual forest stands, and each organism belongs to one of the sampling units. This relationship describes the tessellated population approach for which finite population cluster sampling is appropriate. The second situation describes most forest populations where organisms (trees) inhabit a continuous area where there is no natural division into sampling units. Thus, to study individuals on a per unit area basis, the unit must be arbitrarily defined by a randomly located plot. The third situation differs from the second by allowing for the movement of organisms within the area. The term *continuous population sampling*, as suggested by Bartlett (1986), is used to refer to the second situation where randomly located plots are used.

The notion that the theory behind continuous population forest sampling designs differs from that of traditional finite population cluster sampling, covered in texts such as Cochran (1977, Chap. 9 and 10) and Thompson (1992, Chap. 12 and 13), is not a new concept. Pielou (1977), Ripley (1981, Chap. 3 and 6), Bartlett (1986), Campbell (1993), Eriksson (1995), Stevens (1997), and Gregoire (1998) discuss numerous topics relating to the sampling of a continuous space. For our comparison, a continuous population sampling approach is used. Relevant references to both finite and continuous population sampling results are given. Differences in sampling protocol and area estimation for the two approaches will be discussed as necessary.

First Phase

The first phase of a continuous population sampling design is used to stratify the land area into forest and nonforest strata. The total area of the population is assumed known and is denoted by A_T. Given A_T, and assuming no classification errors of the sample points, the total forested area is determined by calculating the proportion of the sample points, $\hat{P}_F$, that fall on land that is classified as forest. The sample points for this proportion may be generated from a grid sample with a random start. However, the estimator of forested area, $\hat{A}_F$, and its variance are generated assuming points are uniformly distributed over the population. Estimates derived from systematic sampling can be problematic when there is periodicity in the spatial pattern of the population. Milne (1959) studied such concerns in forest populations and concluded that "the danger to centric systematic sampling from unsuspecting periodicity is so small as to be scarcely worth a thought."

Second Phase

The second phase plots are a subsample of points that fell in the forest stratum during the first phase. Most large-scale inventories use a cluster of plots or prism points to reduce costs and estimator variance. The cluster plot used in this study resembles the FIA ground plot and consists of four 1/60th ha (1/24 ac) fixed-area subplots. On each subplot, all trees greater than 12.7 cm (5 in.) diameter at breast height (dbh) are tallied. Each subplot also contains an additional 1/750 ha (1/300 ac) microplot where trees with dbh less than 12.7 cm are tallied. In practice, the microplot is offset from the center to avoid trampling. For simplicity, the microplot was not offset in our simulation study. To distinguish between the two plot types, the terms subplot and microplot are used to denote the 1/60th and 1/750 ha plots, respectively. The HT estimator requires that plots are included in the sample only when the center point of the cluster plot falls in the forest stratum. For this study, the same sampling protocol is used for the RM and MR estimators, where inference is model-based. All subplots that cover more than one condition class are mapped for forest/nonforest and condition class boundaries. The boundaries are used to determine area of forest and areas by condition class for each plot.

Estimation Methods

The focus of the article is estimating area by condition class and various attributes derived from the trees within the forest stratum such as total basal area and number of trees. First, the estimation of forested area (first phase) is discussed, and then three estimators are proposed for deriving total values for the various attributes (second phase). To further illustrate the differences between methods, the estimators are given for an arbitrary condition class c and domain l. Two of the three estimation methods have been discussed by Scott and Bechtold (1995) and Williams and Schreuder (1995). The third estimator is presented, and standardized notation for all methods is given. The performance of the estimators is compared using the continuous population approach.

---

[1] It is often thought that triangles, rectangles, or hexagons must be used to perform the regular division of a forest. However, a system of transformations was developed by M.C. Escher (cf. Bool and Locher 1982) to perform such divisions with other shapes.

Notation

The following variables are associated with area estimation:

C = number of condition classes to be estimated

c = index for each condition class $c = 1, \ldots, C$

h = index for each cluster plot $h = 1, \ldots, m$

i = index for each subplot

A_T = total land area

$\hat{A}_F$ = estimated area of forested land

$\hat{A}(c)$ = estimated area of forested land in condition class c

$\hat{P}_F$ = estimated proportion of A_T which is forested

$\hat{P}(c)$ = estimated proportion of A_F which is in condition class c

a_{hi}^{mic} = area of the microplot covering forestland on subplot i of cluster plot h

a_{hi}^{sub} = area of the subplot covering forestland on subplot i of cluster plot h

$a_{hi}^{mic}(c)$ = the portion of a_{hi}^{mic} in condition class c

$a_{hi}^{sub}(c)$ = the portion of a_{hi}^{sub} in condition class c

The notation used to describe tree attributes is:

L = number of domains to be estimated

j = index for each tree

k = number of subplots in the cluster with $k = 4$ for this study

l = index for each domain $l = 1, \ldots, L$

m = number of second phase cluster plots in the forest stratum

$K_h(F)$ = the set of subplot indices of cluster plot h whose center point falls in the forest

$K_h(\mathrm{c})$ = the set of subplot indices of cluster plot h whose center point falls in condition class c

$k_h(F)$ = number of subplots of cluster plot h whose center point falls in forest

$k_h(\mathrm{c})$ = number of subplots of cluster plot h whose center point falls in condition class c

$y_{hij}(c, l)$ = attribute of interest for tree j on subplot i of cluster plot h in condition class c and domain l

n_{hi} = number of trees tallied on both micro or subplot i of cluster plot h

n_{hi}^{mic} = number of trees tallied on microplot i of cluster plot h

n_{hi}^{sub} = number of trees tallied on subplot i of cluster plot h

Area Estimation by Condition Class

The forested area in condition class c is estimated by

$$\hat{A}(c) = \hat{A}_F \hat{P}(c) = \hat{A}_F \frac{\sum_{h=1}^{m} \sum_{i=1}^{k} a_{hi}^{sub}(c)}{\sum_{h=1}^{m} \sum_{i=1}^{k} a_{hi}^{sub}}, \tag{1}$$

with $a_{hi}^{sub}(c)$ being the area of subplot i of clusterplot h covering condition class c derived by mapping the boundary of the condition class. The forested area by condition class can also be derived from the microplot. This alternative estimator is

$$\hat{A}'(c) = \hat{A}_F \hat{P}'(c) = \hat{A}_F \frac{\sum_{h=1}^{m} \sum_{i=1}^{k} a_{hi}^{mic}(c)}{\sum_{h=1}^{m} \sum_{i=1}^{k} a_{hi}^{mic}}$$

Horvitz-Thompson Estimator

This estimator, used in Williams and Schreuder (1995), can be formally described in a number of different ways (Eriksson 1995). In this study, each tree is assigned an inclusion probability, π (the term expansion factor is often used in forestry and is the inverse of the inclusion probability), which is used to generate estimates of the population total using a Horvitz-Thompson estimator (cf. Cochran 1977, Sec. 9A.7, Thompson 1992, Sec. 6.2). The estimator for the total of any attribute in condition class c and domain l is given by:

$$\hat{Y}_{HT}(c,l) = \frac{1}{m} \sum_{h=1}^{m} \sum_{i=1}^{k} \sum_{j=1}^{n_{hi}} y_{hij}(c,l) / \pi_{hij}, \tag{2}$$

where

$$y_{hij}(c,l) \begin{cases} y_{hij} & : \ \textit{if} \ \ y_{hij} \text{ is in condition class c and domain } l \\ 0 & : \ \text{otherwise} \end{cases},$$

π_{hij} is the inclusion probability for the tree, and n_{hi} is the number of trees that were selected on both micro- and subplot i of cluster plot h.

For cluster plot designs, the true π values are too time-consuming to calculate for each tree and standard boundary correction technique, such as the mirage technique (Gregoire 1982), and are also difficult to implement and not likely to be used. Thus, when the goal is to assess performance in a real world setting, it would be unfair to use the true π values for comparing this estimator to the alternatives. Therefore, two approximate Horvitz-Thompson estimators are proposed that could be implemented with less trouble. For both of these estimators, each subplot in the cluster is used to generate an estimate of the total as if it were a single plot and the individual estimates are averaged together to generate an estimate of the total for the cluster plot. The most significant consequence of this action is that trees are not tallied on subplots whose center-point falls outside the boundary of the forest stratum.

Two approximations to the true values are reasonable. The first approximation is

$$\pi'_{hij} = \begin{cases} k_h(F)a_{hi}^{mic} \,/\, \hat{A}_F & : \quad d < 12.7 \ \text{cm} \\ k_h(F)a_{hi}^{sub} \,/\, \hat{A}_F & : \quad d \geq 12.7 \ \text{cm} \end{cases}$$

depending on whether or not the tree is tallied on the microplot. For this approximation, the inclusion probabilities are exact when $\hat{A}_F = A_F$ and for trees that are more than one cluster radius away from the forest stratum boundary. The second approximation uses the tree-centered approach described in Husch et al. (1982. Sec. 14-3). For each tree in the population of interest, a circular plot of area a^{mic} or a^{sub}, depending on the dbh, is centered around the tree. The inclusion probability is then approximated by

$$\pi'' = \frac{k_h(F)a''}{\hat{A}_F},$$

where a'' is the area of the circle that is within the population boundary. When $\hat{A}_F = A_F$, the π'' are exact inclusion probabilities for sampling designs that use only a single plot ($k = k_h(F) = 1$), rather than a cluster of plots. These π''s are also equal to the true inclusion probabilities for trees that are more than one cluster radius away from the forest stratum boundary.

The motivation for these estimators is as follows. The π'' approximation is used because if the cluster plots were broken up and the individual subplots were spread uniformly across the forested area, an unbiased HT estimator based on the $\Sigma_{h=1}^{m} k_h(F)$ subplots is possible. Thus, this approximation assumes the location of plots within the forest area do not need to be completely random. The π' approximation makes the same assumption but uses a^{mic} and a^{sub} to approximate a''. This additional approximation is used because it would be very time-consuming for field crews to measure a'' for every tree.

The two approximate Horvitz-Thompson estimators based on these π approximations are

$$\hat{Y}'_{HT}(c,l) = \frac{1}{m}\sum_{h=i}^{m}\sum_{i \in K_h(F)}\sum_{j=1}^{n_{hi}} y_{hij}(c,l) \,/\, \pi'_{hij}$$

and

$$\hat{Y}''_{HT}(c,l) = \frac{1}{m}\sum_{h=i}^{m}\sum_{i \in K_h(F)}\sum_{j=1}^{n_{hi}} y_{hij}(c,l) \,/\, \pi''_{hij}$$

Note that changes in condition classes at the cluster-plot level influence neither π nor its approximations. These quantities are only affected when a plot straddles the forest/nonforest stratum boundary. In this situation, the value of the π approximations represents the reduced likelihood of trees along the boundary of the forest being selected. The approximations suggested can cause a bias because π's are no longer fixed. The effects of these approximations are discussed later and are compared to the other estimators in the simulation study.

The Horvitz-Thompson estimator also differs from the other two in how the condition class attributes are estimated. The Horvitz-Thompson estimator expands the attributes of trees on each cluster plot to give an estimate of the total for the forested area, while the other two methods generate average per-unit-area estimates directly. As a result, per-unit-area estimates by condition class can only be generated for the population using a ratio estimator as described in Williams and Schreuder (1995). For example, if an estimate of basal area per hectare for the hardwood condition class was desired, the estimate of total basal area for the hardwood condition class [Equation (2) or one of the approximations] is divided by the estimate of the area in the hardwood condition class [Equation (1)].

Ratio of Means Estimator

The ratio of means estimator (RM) was first suggested by Scott and Bechtold (1995) and further developed by Bechtold and Zarnoch (1999) for mapped plots. Its continuous population analog is similar with a few key differences. While I am unaware of specific references in the sample survey literature, this type of sampling has been discussed in the context of spatial statistics and quantitative ecology, where it is often referred to as *quadrat sampling*. Ripley (1981, Chap. 6) summarizes quadrat sampling methods for determining the population total. He defines the intensity of the population, denoted by λ, as the total number of units divided by the total area. He states: “Random positioning of the quadrats allows us to take averages over the randomization. Then the count per unit area from the samples is an unbiased estimator of the population count divided by the total area of the region within which the samples are taken.” This statement may be misleading because while it may be correct for populations that can be described by a stationary point process (Stoyan et al. 1995, Sec. 4.6) or ones with no edge (e.g., those spread over a sphere), it may not be applicable to sampling forest populations because of the spatial properties of the trees along the forest boundary.

For this method, the sampling unit is the plot, and the area of the plot for a given condition class is tallied with trees located on that portion of the plot. Thus, both the tree tally and plot area within the condition can be viewed as random variables. For all trees tallied on the subplot, the tree attributes y and subplot areas $a^{sub}(c)$ are summed across all plots in condition class c and domain l to form the ratio $\hat{R}(c,l)$, which is the per-unit-area estimator of y for all trees with $dbh \geq 12.7$ cm. Information for trees with dbh < 12.7 cm on the microplots is summed similarly to obtain the ratio $\hat{R}'(c,l)$. The additive RM estimator of the total in condition class c and domain l is given by

$$\hat{Y}_{RM}(c,l) = (\hat{R}'(c,l)\hat{P}'(c) + \hat{R}(c,l)\hat{P}(c))\hat{A}_F.$$

Note that this estimator uses the estimated proportion of the population in condition class c from both the micro- and subplot. While it seems illogical to use $\hat{P}'(c)$, because its variance will certainly be larger than that of $\hat{P}(c)$, additivity across condition classes cannot be maintained using only $\hat{P}(c)$. While additivity is not necessary, it is a useful tool in maintaining confidence in public databases because it ensures that the estimates for all individual condition classes and domains can be summed to match the estimate of the overall total.

Expanding this estimator yields:

$$\hat{Y}_{RM}(c,l) = \left(\frac{\sum_{h=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{n_{hi}^{mic}} y_{hij}^{mic}(c,l)}{\sum_{h=1}^{m}\sum_{i=1}^{k} a_{hi}^{mic}(c)} \hat{P}'(c) + \frac{\sum_{h=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{n_{hi}^{sub}} y_{hij}^{sub}(c,l)}{\sum_{h=1}^{m}\sum_{i=1}^{k} a_{hi}^{sub}(c)} \hat{P}(c) \right) \hat{A}_F.$$

where

$$y_{hij}^{mic}(c,l) = \begin{cases} y_{hij} & : \text{if } y_{hij} \text{ is in condition class } c \text{ and domain } l \text{ and tallied on the microplot} \\ 0 & : \text{otherwise} \end{cases}$$

and

$$y_{hij}^{sub}(c,l) = \begin{cases} y_{hij} & : \text{if } y_{hij} \text{ is in condition class } c \text{ and domain } l \text{ and tallied on the subplot} \\ 0 & : \text{otherwise} \end{cases}$$

Substituting the expressions for $\hat{P}'(c)$ and $\hat{P}(c)$ from the previous section and canceling like terms yields

$$\hat{Y}_{RM}(c,l) = \left(\frac{\bar{y}^{mic}(c,l)}{\bar{a}^{mic}} + \frac{\bar{y}^{sub}(c,l)}{\bar{a}^{sub}} \right) \hat{A}_F, \tag{3}$$

where

$$\bar{y}^{sub}(c,l) = 1/m\Sigma_{h=1}^{m}\Sigma_{i=1}^{k}\Sigma_{j=1}^{n_{hi}^{sub}} y_{hij}^{sub}(c,l)$$

and

$$\bar{a}^{sub} = 1/m\Sigma_{h=1}^{m}\Sigma_{i=1}^{k} a_{hi}^{sub},$$

with $\bar{y}^{mic}(c,l)$ and $\bar{a}^{mic}$ similarly defined. This shows that the additive RM estimator of the forest total is composed of the average of the *y* attributes in condition class *c* and domain *l* for the micro- and subplots, respectively, divided by their respective average area of the plot covering the forest stratum and is an application of the results in Ripley (1981, Chap. 6). Thus, the apparent need to map both the micro- and subplot by condition class is unnecessary for estimating population totals. Instead, each tree must be assigned to its appropriate condition class and domain and only the forest and nonforest boundaries require mapping because this is all that is required to compute a^{mic} and a^{sub}. Of course, ignoring the condition class boundaries may compromise other analyses such as per area estimates by condition class.

The ratio estimator has been extensively studied, though not in the context of double sampling or as the sum of two ratios, as given in (3). If only trees on the subplot are considered, and all condition class and domains are considered, the ratio estimator

$$\hat{R} = \frac{\bar{y}^{sub}}{\bar{a}^{sub}}$$

would be optimum under the tessellated population approach (Cochran 1977, Sec. 6.7), in the sense of being the best linear unbiased estimator, if:

1. The relationship between the sum of the cluster plot

$$y_{h..}^{sub} = \sum_{i=1}^{k} \sum_{j=1}^{n_{hi}^{sub}} \sum y_{hij}^{sub}$$

and the forested plot area

$$a_{h.}^{sub} = \sum_{i=1}^{k} a_{hi}^{sub}$$

is a straight line through the origin.

2. The variance of $y_{h..}^{sub}$ conditional on $a_{h.}^{sub}$ is proportional to $a_{h.}^{sub}$.

Assuming the results hold in the continuous population setting, results from spatial statistics and stochastic processes can be used to determine when these two conditions will be met by studying the RM estimator for the total number of trees. For this estimator, $y_{hij} = = 1$, with $y_{h..}$ being the number of trees counted on a cluster plot with forested area $a_{h.}^{sub}$. If the forest is viewed as a realization of a point process, the first condition holds when the point process describing the forest has a constant intensity because

$$y_{h..}^{sub} = \lambda a_{h.}^{sub} + \varepsilon \tag{4}$$

where $E[\varepsilon] = 0$ and $Var[\varepsilon] = \sigma^2 v(a_{h.})$ and λ is the expected number of points per unit area (Ripley 1981, p. 13). The second condition is satisfied provided the point process describing the forest is consistent with a stationary Poisson process because $Var[y_{h..}] = Var[\varepsilon] = \lambda a_{h.}^{sub}$ (Ripley 1981, p. 102). The Appendix shows that the RM estimator is unbiased with respect to this model under the continuous population approach. A reasonable model for the variance of other point process models is $Var[y_{h..}] = \lambda(a_{h.}^{sub})^r$. If the trees are aggregated, the count per unit area is more variable, which implies $r > 1$. Regularity in the spatial pattern implies $r < 1$. When other attributes are estimated, such as basal area, the two conditions will hold provided the location of trees is consistent with a Poisson process, and the size of the attribute is spatially uncorrelated. Penttinen et al. (1995) discusses analytical methods for describing these populations.

There are some important implications associated with the model-based properties of the RM estimator. For example, the rules for selecting the *m* plot locations are not as well defined as those of the HT estimator because the Poisson process assumption places few restrictions on the size, shape or location of $a_{h.}^{sub}$ (Taylor and Karlin 1994, Sec. 5.2 and 5.5). Therefore, the *m* plot locations could be purposively selected from the interior of the population, or at the other extreme, they could be any *m* random plot locations such that any portion of the plot samples part of the forest stratum. The estimators for both of these examples are unbiased with respect to Equation (4). However, the variance of the two ratios will be very different. To illustrate, rearrange Equation (4) into a ratio. This yields

$$\frac{y^{sub}_{h..}}{a^{sub}_{h.}} = \lambda + \varepsilon',$$

where $E[\varepsilon'] = 0$ and $Var[\varepsilon'] = \lambda / a^{sub}_{h.}$. Thus, the sampling designs that minimize the variance of the ratio are those that maximize the area sampled, which are designs where plots are selected that do not intersect the boundary. This result is consistent with the finite population ratio estimator results given by Royall and Cumberland (1981). For our study, the same m plot locations are used for all three estimators. This design is not optimal for the RM estimator, but it does provide a component of randomization that is necessary to avoid concerns of personal bias in the sample selection. For most natural populations the model given by (4) will not be satisfied because of edge-effect, which is a consistent difference in forest structure between the interior and boundary of the forest. The model failure occurs because λ is a function of the distance to the forest boundary when edge-effect is present. Wales (1972) and Chen et al. (1992) studied edge-effects for forest stands and give the following list of significant differences between the interior and the edges

- different stocking density, number of stems, and basal area
- different growth rates
- elevated rates of mortality
- different numbers of seedlings and saplings
- different responses of tolerant and nontolerant species

The effects of departure from the model $y^{sub}_{h..} = \lambda a^{sub}_{h.} + \varepsilon$ on the tessellated population ratio estimator has been studied extensively by Royall and coauthors (Royall and Herson 1973a, 1973b, Royall and Cumberland 1981, Royall 1992). They found the ratio estimator extremely vulnerable to bias for even subtle model failures. A review of the sample survey and spatial statistics literature has been unable to locate similar results for continuous population sampling. The simulation study is used to illustrate how edge-effect influences the continuous population RM estimator in a similar manner.

Astute readers may have noticed that Scott and Bechtold [1995, Equation (19)] use the total area of the micro- and subplot in their derivations of forest area, rather than the portion of the plot covered by forest (a^{mic} and a^{sub}, respectively). This occurs because the equal area tessellated population approach divides the area into strata along the boundary of the sampling units, unlike the continuous population approach, which uses the forest/nonforest boundary. Some consequences of this difference are:

- Strata under the tessellated sampling approach are more heterogeneous than those of a continuous population design because some forest and non-forest area are included in every stratum.
- The first phase sample of a tessellated population is not used to estimate area of forest and nonforest. Instead, it simply divides the total area into strata that are predominantly forested and nonforested.
- Because a tessellation of A is almost sure to have some portion of forestland in every stratum, estimates of forest attributes must be calculated for every stratum and then added across all strata.
- The second phase sample size is larger for the tessellated population approach because every plot that covers a portion of forestland must be visited. For a continuous population design, a cluster plot can be installed only when the center point of the cluster falls in the forest stratum.
- Under the tessellated population approach, all tree inclusion probabilities are equal provided the population can truly be tessellated into plots of equal area.

Mean of Ratios Estimator

The mean of ratios (MR) estimator is similar to the HT estimator except that the expansion factors are random variables and are dependent on the area of the plot falling in a given condition class. Thus, rather than expanding each tree by the total plot sizes (a^{mic} or a^{sub}), the proportion of the plot falling in condition class c is used. Thus, the estimator for any attribute of interest in condition class c and domain l is given by;

$$\hat{Y}_{MR}(c,l) = \hat{A}_f \frac{1}{mk_h(c)} \sum_{h=1}^{m} \sum_{i \in k_h(c)} \sum_{j=1}^{n_{hi}} y_{hij}(c,l) / a_{hi(c)},$$

where

$$a_{hi}(c) = \begin{cases} a^{mic}_{hi}(c) & : \ \text{dbh} < 12.7 \ \text{cm} \\ a^{sub}_{hi}(c) & : \ \text{dbh} \geq 12.7 \ \text{cm} \end{cases}$$

and $\hat{Y}_{MR}(c,l) = 0$ if $k_h(c) = 0$. This is one of many possible MR estimators. Others can be derived by averaging the data in different ways. This one was chosen because it is equivalent to the $\hat{Y}'_{HT}$ estimator when all condition classes are combined to estimate the population total. These two estimators are substantially different when estimating specific condition classes.

Optimum properties in the tessellated population setting are discussed by Schreuder et al. (1993) and are similar to those of the RM estimator with the MR estimator being optimal when:

1. A graph of the relationship between the attribute $y_{h..}$ and the plot area $a_{h.}$ forms a straight line through the origin.
2. The variance of $y_{h..}$ conditional on $a_{h.}$ is proportional to $a^2_{h.}$.

Assuming the results hold in the continuous population setting, the first property is satisfied for a point process with constant intensity λ as in Equation (4). The second condition requires an extreme degree of clustering, which is unlikely to occur in natural populations. If the population has an edge-effect, neither of these properties will be satisfied. The bias of the MR estimator can also be a concern because

$$\text{Bias}[\hat{Y}_{MR}] \approx m\text{Bias}[\hat{Y}_{RM}],$$

implying that the bias does not decrease with sample size as is the case for the RM estimator (Tin 1965).

The performance of this MR estimator is known to be inferior to that of the other two estimators when estimating a specific condition class *c*. This occurs because it simultaneously estimates both the size of the condition class (e.g., $\hat{A}(c) = \hat{A}_F / a_{hi}^{sub}(c)$) and the average number of points in the condition class, which makes the variance of the estimator very large. This estimator is included only to illustrate the potential problems with its use. It is not developed to the same extent as the HT and RM estimators.

Variance Estimators

The Variance Estimator for HT Estimators

Williams and Schreuder (1995) list classical variance estimators for forest means and totals and for various tree, area, and boundary attributes. These estimators were based on the assumption of a known forested area, A_F. Accounting for the estimation of A_F is relatively easy. Writing Equation (2) as

$$\hat{Y}_{HT} = \hat{A}_F \bar{Y}_{HT} = \hat{A}_F \frac{1}{A_F m} \sum_{h=1}^{m} \sum_{j=1}^{k} \sum_{j \in In_{hi}} y_{hij} / \pi_{hij},$$

where $\hat{Y}_{HT} = \hat{A}_F \bar{Y}_{HT}$ is a product estimator with $\bar{Y}_{HT}$ the total per hectare estimator. Because $\bar{Y}_{HT}$ and $\hat{A}_F$ should be essentially uncorrelated, the variance is

$$\text{var}[\hat{A}_F \bar{Y}_{HT}] = \hat{A}_F^2 \,\text{var}[\bar{Y}_{HT}] + \bar{Y}_{HT}^2 \,\text{var}[\hat{A}_F].$$

The estimated variance for a tree-based attribute using the approximate HT estimators is given by

$$\text{var}[\hat{Y}_{HT}] \approx \sum_{h=1}^{m} (\hat{y}_h - \hat{Y}_{HT})^2 / m(m-1),$$

where $\hat{y}_h$ is the total per hectare estimate of the *h*th cluster. This estimator will only be an approximate variance estimator for $\hat{Y}'_{HT}$ and $\hat{Y}''_{HT}$ because of the approximated inclusion probabilities and the variability in *m*, which is usually not accounted for in forest surveys that use a systematic grid. The classical variance estimator for $\hat{A}_F$ is defined analogously.

Variance Estimates for RM Estimators

Classical variance estimators given by Bechtold and Zarnoch (1999) and Zarnoch and Bechtold (2000) are briefly summarized. Applying a Taylor's series approximation to Equation (3) yields

$$\text{var}[\hat{Y}_{RM}] \approx \hat{A}_F^2 \,\text{var}\left[\frac{\bar{y}^{mic}}{\bar{a}^{mic}}\right] + \left(\frac{\bar{y}^{mic}}{\bar{a}^{mic}}\right)^2 \text{var}[\hat{A}_F]$$
$$+ \hat{A}_F^2 \,\text{var}\left[\frac{\bar{y}^{sub}}{\bar{a}^{sub}}\right] + \left(\frac{\bar{y}^{sub}}{\bar{a}^{sub}}\right)^2 \text{var}[\hat{A}_F].$$

This approximation ignores both the possibility of a correlation existing between $\hat{A}_F$ and the per acre estimators and data collected on the micro- and subplot, which is reasonable for the simulation study where trees with dbh < 12.7 cm were generated independently from the those with dbh ≥ 12.7 cm. Finite population sampling theory suggests this estimator is a reasonable approximation provided that $m > 30$ and that the coefficient of variation for $\bar{a}_{mic}$ and $\bar{a}_{sub}$ is small. Underestimates in excess of 30% can occur for small sample sizes (cf. Cochran 1977, Sec. 6.9 and 6.17) and when the $var[y_{h..}]$ is not proportional to $a_{h.}$ (Royall and Eberhardt 1975). Alternative classical variance estimators, such as those described by Royall and Cumberland (1981), require information that is not available in large-scale forest inventories.

Simulation Study

Because few analytical results are available to suggest which of the estimators would be best in terms of bias and variance, a simulation study was performed to evaluate their properties. The data collected from large-scale inventories are not useful in a simulation study because the true population values are unknown. A small mapped data set was used to compare the performance of the three estimators.

The data were collected in central New Jersey in 1985 (Gregoire and Scott 1990) on an area of 5.37 ha and comprised two adjacent hardwood stands containing 4,744 trees. Square 15.2 m × 15.2 m (50 × 50 ft) plots were laid out and all trees of dbh ≥ 12.7 cm (5 in.) were measured for diameter and their x and y coordinates were recorded. The data set relates to two stands of different structure, one comprised of mature hardwoods of sawlog size and the other a stand of smaller hardwoods that were regenerating after an earlier disturbance. The terms *hardwood* and *regeneration* will be used to refer to the two stands. These stands were separated by a dirt road, which divided the population roughly in half and had grown over at the time of measurement. Conditions within each stand were fairly heterogeneous, but there was no visual indication of an edge-effect at the forest boundary. The original data were augmented with additional trees because no information on trees of dbh < 12.7 cm was collected in 1985. Thus information for 5,000 seedlings and saplings was added to the data set, their (x,y) coordinates and diameters being generated from a uniform distribution. Another addition to the data set was a strip of nonforested area surrounding the orignal data. This strip added an additional 13.28 ha to the data set. This strip was added so the process of stratifying the population into forest and nonforest areas could be mimicked. Figure 1 depicts the data set used. The seedling data were excluded, and the diameter of each circle represents the relative size of each tree.

Forest area, $\hat{A}_F$, was estimated by simulating an aerial photography survey which consisted of randomly selecting 100 points from a uniform distribution over the 18.65 ha block and classifying each point as forest or nonforest. The proportion of area under forest was then estimated by

$$\hat{P}_F = \frac{\text{The number of points falling in forest}}{100},$$

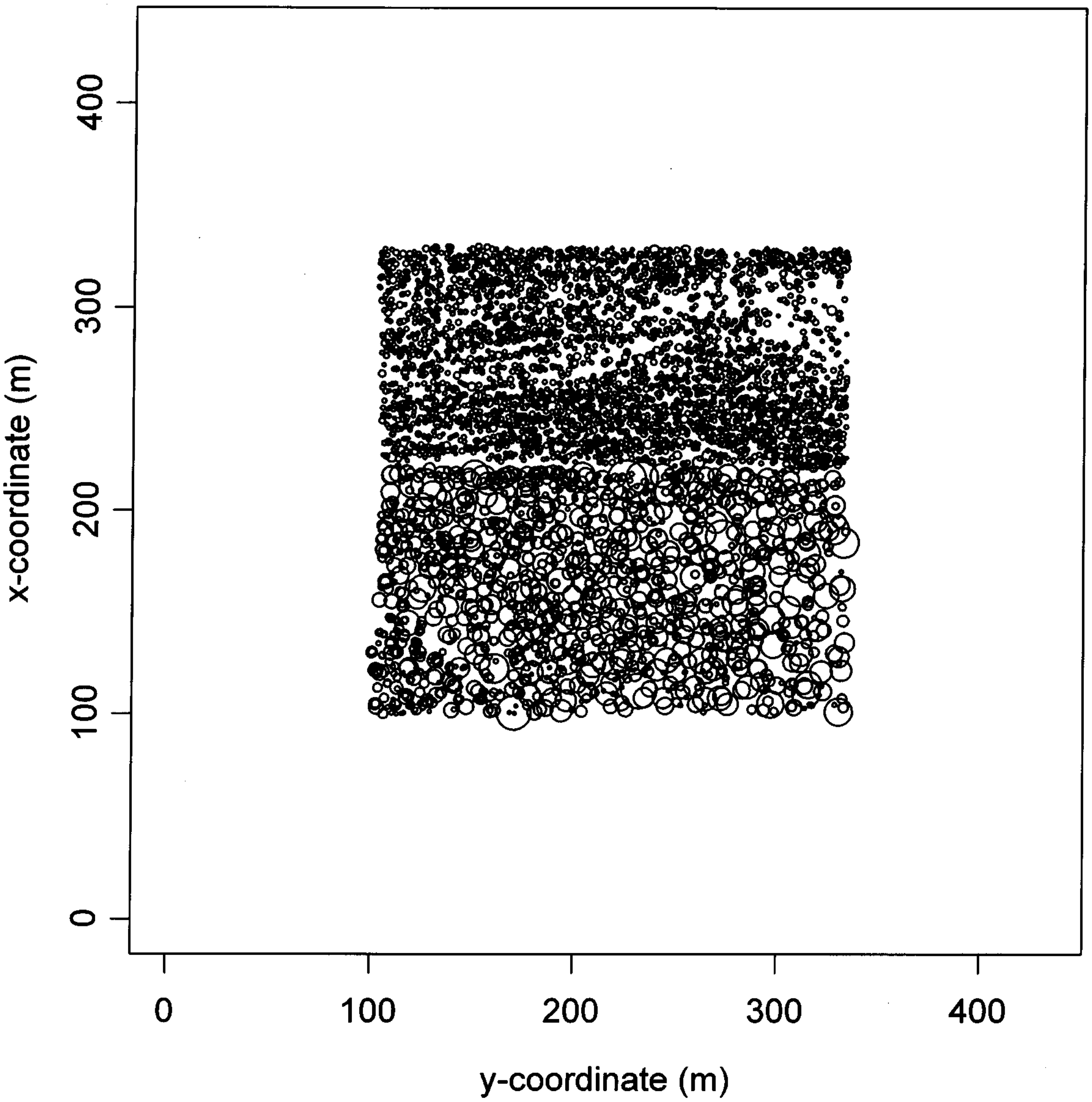

Figure 1. Plan of the data set. The size of each circle shows relative size of each tree. The 5,000 artificially generated trees with dbh < 12.7 cm are not included.

with $\hat{A}_F = 18.65\hat{P}_F$. A sample of the block of forest was drawn by selecting the location of every fifth point within the forested portion of the data set. This approach provided a sample of *m* uniformly distributed cluster plot locations over the forested area and a random second phase sample size. One advantage of this approach is that the variability of the second phase sample size would be more similar to that achieved when sampling from a systematic grid than from a purely random sample. The portion of each micro- and subplot falling in the sawlog and regeneration stands was calculated using the methods described in Scott and Bechtold (1995), and all trees on these plots were tallied. Estimates of the total basal area and its variance were derived from the sample information. Estimates were made for the whole population as well as for the domains defined by tree size (dbh < 12.7, 12.7 ≤ dbh < 43.2, and dbh ≤ 43.2) and the two condition classes (sawlog and regeneration) and every combination thereof. The sampling process was repeated M = 4000 times. The average number of clusters over the M samples was approximately m = 5.3. The mean and variance of the M samples were calculated and used to determine the magnitude of the biases in both the estimates of the total basal area ($\hat{Y}$) and its estimate of variance (var[$\hat{Y}$]) for each of the three estimators.

Additional simulations were performed to illustrate the bias of the estimators when the population of interest has an edge-effect. In each of these, the diameters of all trees within 5 m of the forest/nonforest boundary were multiplied by a fixed amount q = 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0. The diameters of 736 trees were affected, which represented about 7.5% of the population. To illustrate that the bias does not significantly decrease with increases in sample size, another set of simulation studies was performed where the number of ground plots was nearly tripled. For these simulations the m value averaged over the 4,000 sampled was 16.7. The larger m value was generated by increasing the first phase sample size from 100 to 300 points. While these sample sizes seem small, they are reasonable for some strata in many large-scale

Table 1. Simulation results for the approximate Horvitz-Thompson (HT) estimator $\hat{Y}'_{HT}$ by size class and condition class. The number of simulations performed was 4,000. Estimates of total basal area are expressed as a percentage of the true value with 100.0 indicating an unbiased estimate. Simulation standard errors across the 4,000 simulations are given in the curved brackets and are expressed as a percentage of the true basal area. The averages of the classical variance estimates are given in the square brackets and are also expressed as a percentage of the true basal area. Classical variance estimators are unbiased when the numbers in the square and round brackets are identical.

	Size class			
Condition class	All	0 < dbh < 12.7	12.7 < dbh < 43.2	dbh > 43.2
All	98.8	99.2	98.3	98.9
	(24.6)	(33.0)	(39.3)	(32.3)
	[24.4]	[32.9]	[39.5]	[32.1]
Sawlog	98.9	98.1	95.6	98.9
	(41.5)	(55.1)	(56.3)	(41.6)
	[41.2]	[55.9]	[56.7]	[41.4]
Regeneration	98.6	100.5	98.5	98.7
	(44.2)	(60.7)	(44.9)	(45.8)
	[44.3]	[62.1]	[45.2]	[45.6]

surveys. For example, they approximate the range of sample sizes for each county for each panel in the new annualized FIA inventory (Reams and Van Deusen 1999).

Comparison of the Estimators

The discussion is divided into two parts, the first addressing the performance of the each of the estimators in the simulation study and then the practical implications of using the estimators for compiling inventory information.

Bias and Variance

The bias, variance, and average of the classical variance estimates for the data set with no edge-effect are presented in Tables 1–4. The approximate HT estimators tend to underestimate the totals by 1 to 2%. The RM estimator exhibits little if any bias for most condition classes and domains, the exception being the estimator of trees in the sawlog stand whose dbh ranged from 12.7 to 43.2 cm. The MR estimator produced both under- and overestimates as large as 5%. Tables 1–3 illustrate the smaller variance of the RM estimator when compared to the approximate HT estimator for fixed-area sampling. The MR estimator exhibits a much larger variance when used to estimate totals by condition class. In most condition classes, the percentage increase is small. However, the standard error of the MR estimator, for trees in the sawlog stand with 12.7 cm ≤ dbh < 43.2 cm, was about twice as large as for the approximate HT and RM estimators. This occurs because plots that straddle the condition class boundary can have trees tallied when the area of the plot in the condition class, $a_{hi}(c)$, is very small. The resulting estimate is much larger than the true total for the condition class, producing a large outlier and inflating the variance. In practice, the variance of the MR estimator may be larger than that indicated in this study because the dirt road separating the regeneration and sawlog stands acted as a buffer zone that limited the number of trees tallied on one of these small sections.

The averages of the classical variance estimates for both approximate HT estimators (Tables 1 and 2) are nearly identical to the simulation variance across the 4,000 replications. This indicates that the approximated inclusion probabilities and the random sample size have little effect on the performance of the variance estimator. The classical variance estimator for the RM estimator underestimates the simulated variance (Table 3). An additional simulation study was preformed to determine the cause of the bias. In the study, the true A_F was substituted for the estimated $\hat{A}_F$ and the second phase sample size was fixed. This did not reduce the bias. Thus, the bias can be attributed to the ratio estimator. This result is consistent with the theory for ratio estimation, and is

Table 2. Simulation results for the approximate Horvitz-Thompson (HT) estimator $\hat{Y}''_{HT}$ by size class and condition class. The number of simulations performed was 4,000. Estimates of total basal area are expressed as a percentage of the true value with 100.0 indicating an unbiased estimate. Simulation standard errors across the 4,000 simulations are given in the curved brackets and are expressed as a percentage of the true basal area. The averages of the classical variance estimates are given in the square brackets and are also expressed as a percentage of the true basal area. Classical variance estimators are unbiased when the numbers in the square and round brackets are identical.

	Size class			
Condition class	All	0 < dbh < 12.7	12.7 < dbh < 43.2	dbh > 43.2
All	99.5	98.0	98.0	99.9
	(23.3)	(42.6)	(42.7)	(33.4)
	[23.4]	[42.0]	[42.8]	[33.3]
Sawlog	100.1	98.9	98.8	100.1
	(44.5)	(71.0)	(69.8)	(44.8)
	[44.6]	[71.0]	[69.8]	[44.9]
Regeneration	98.1	97.1	98.0	98.3
	(48.2)	(75.3)	(49.4)	(52.3)
	[48.3]	[75.4]	[49.7]	[52.2]

Table 3. Simulation results for the ratio of means (RM) estimator by size class and condition class. The number of simulations performed was 4,000. Estimates of total basal area are expressed as a percentage of the true value with 100.0 indicating an unbiased estimate. Simulation standard errors across the 4,000 simulations are given in the curved brackets and are expressed as a percentage of the true basal area. The averages of the classical variance estimates are given in the square brackets and are also expressed as a percentage of the true basal area. Classical variance estimators are unbiased when the numbers in the square and round brackets are identical.

	Size class			
Condition class	All	0 < dbh < 12.7	12.7 < dbh < 43.2	dbh > 43.2
All	100.2	99.4	98.5	100.6
	(21.7)	(31.6)	(37.7)	(32.6)
	[18.2]	[27.9]	[36.2]	[27.5]
Sawlog	100.6	98.9	94.3	100.7
	(40.0)	(54.7)	(49.4)	(41.1)
	[37.3]	[52.1]	[47.3]	[37.8]
Regeneration	99.2	99.8	98.9	99.6
	(41.8)	(58.9)	(43.3)	(43.6)
	[40.3]	[58.6]	[42.0]	[42.3]

not surprising because, on average, only 5.3 points were sampled and the variance approximation requires at least 30 points (Cochran 1977, p. 162). Other possibilities include:

- The existence of two completely different stand types in this small population creates the potential for the numerator of the RM estimator to follow a bimodal distribution when, in estimating the bias of the RM estimator, it is assumed that the numerator and denominator are distributed as bivariate normal random variables (see Cochran 1977, Eq. 6.43).
- The distribution of plot area (denominator) could also play a role in the underestimation problem. Plot areas (a^{mic} and a^{sub}) are constant provided the plots do not intersect the boundary of the forest stratum, where plot area is decreased to represent the reduced area of the forest sampled. The result is that the distribution of plot area is highly skewed to the left. To illustrate, the areas of 2,000 sample plots were generated from the simulation study. The first four moments of the distribution were calculated and adjusted so that they could be compared to those of a standard Normal distribution. The sample moments $\hat{\mu} = 0, \hat{\sigma}^2 = 1$, $\hat{\mu}_3 = -4.5$ and $\hat{\mu}_4 = 23.7$ were compared to those for a standard normal which are $\mu = 0$, $\sigma^2 = 1$, and $\mu_3 = 0$. Now, combining the facts that $\hat{\mu}_3$ and $\hat{\mu}_4$ are substantially larger than those for a standard Normal and the Taylor's series approximation for the variance is only of order $1/\sqrt{m}$, suggests that the contribution of higher order terms in the Taylor's series approximation may no longer be trivial. Thus, very large sample sizes may be required to satisfactorily reduce the bias. Additional testing has shown $m = 100$ to be adequate for this population and a number of others.

These observations suggest that forest surveys using RM estimators may represent a difficult situation for the Taylor's series classical variance estimator. The jackknife estimator, suggested by Royall and Cumberland (1981) and other authors, would likely be a better choice.

When the diameters along the boundary of the population were changed, both the approximate HT and RM estimators produced biases in estimates of all condition and size classes (Figures 2 and 3). The RM estimator produced biases of the population total ranging from 3.4% when $q = 0.25$, to –7.6% when $q = 2.0$. Estimates for the domain with dbh ≥ 43.2 cm in the sawlog stand showed biases as large as –17.6 %, whereas those in other domains showed less. The first approximate HT estimator $(\hat{Y}'_{HT})$ exhibited a smaller bias in almost every situation, with values ranging for 0.1 to –4.3% for the whole population. For the domain consisting of trees with dbh ≥ 43.2 cm in the sawlog stand, the maximum bias was –11.4 %. The second estimator $(\hat{Y}''_{HT})$ exhibited a maximum bias of less than 2% when estimating basal area for the whole population. When the domain consisting of trees with dbh ≥ 43.2 cm in the sawlog stand was considered, the maximum bias

Table 4. Simulation results for the mean of ratios (MR) estimator by size class and condition class. The number of simulations performed was 4,000. Estimates of total basal area are expressed as a percentage of the true value with 100.0 indicating an unbiased estimate. Iterated standard errors are given in the curved brackets and are expressed as a percentage of the true basal area.

	Size class			
Condition class	All	0 < dbh < 12.7	12.7 < dbh < 43.2	dbh > 43.2
All	98.8	99.2	98.3	98.9
	(24.6)	(33.0)	(39.3)	(32.3)
Sawlog	102.1	99.8	105.2	102.0
	(42.4)	(57.9)	(96.5)	(42.5)
Regeneration	103.5	103.1	101.8	105.5
	(46.2)	(65.9)	(45.8)	(53.2)

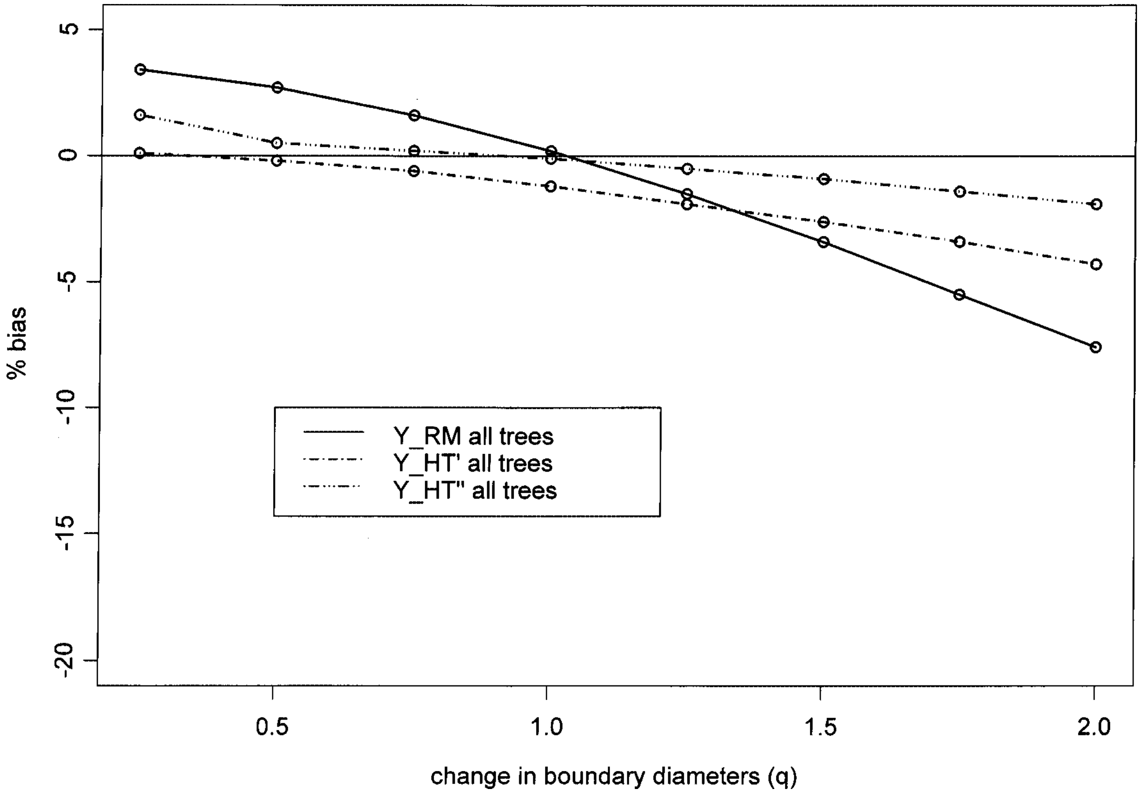

Figure 2. Bias of the HT and RM estimators for all trees. Average sample size is 5.3.

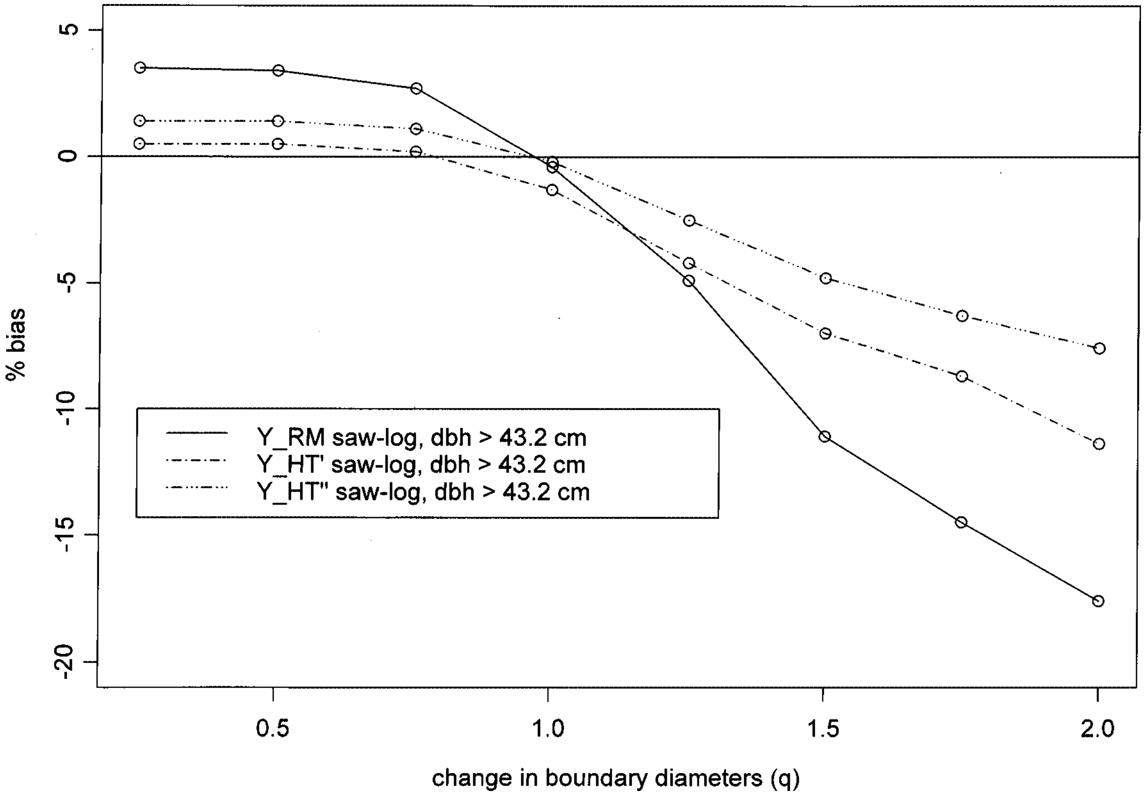

Figure 3. Bias of the HT and RM estimators for the trees in the sawlog stand with dbh > 43.2 cm. Average sample size is 5.3.

was –7.6%. The approximate HT estimators exhibit this bias because of the approximation of the probabilities of selection along the boundary.

To illustrate the effect of sample size on the RM estimator, the simulations were run again using a sample three times larger. Figure 4 is a comparison of sample sizes. Interestingly, the bias of the RM estimators actually increased for small q values when the sample size was increased. Because the bias is nearly identical for both sample sizes, it is concluded that it is almost solely the result of the edge effect. The small size of the data set and the extreme edge-effect used in these populations contribute to the magnitude of the bias in the estimators. For real populations, the bias of most attributes is expected to be smaller. However, the estimator for any attribute that is almost exclusively associated with the forest boundary [e.g., number of trees affected by Kudzu (*Pueraria lobata* Willd.) infestation] is likely to be have a large bias.

Additivity

An additive estimator is one whose individual domain and condition class estimates, regardless of the number of classifiers, sum to the estimate for any or combination of domains and condition classes, and to the estimate of the population as a whole. For public databases, such as FIA, additivity is important for maintaining public confidence and for highlighting possible errors in compiling the data.

Additivity for the MR estimator requires a common denominator, which can not be achieved for condition classes. Any HT estimator and the RM estimator given by Equation (3) are additive.

Sampling Protocols

The approximate HT estimators use a slightly different sampling protocol along the boundary of the population. For all methods, the cluster plot is installed only when the center point of the cluster falls in the forest stratum. For the RM estimator, all subplots that cover a portion of the forest are installed, and all trees are measured. For the approximate HT estimators, each subplot is installed only when the center point of the subplot falls in the forest stratum. This constitutes a small saving in time and effort at each cluster plot. However, this savings could be offset by the information needed to approximate inclusion probabilities along the forest/nonforest boundary for the $\hat{Y}''_{HT}$ estimator. This would require mapping the boundary past the edge of the subplot.

Data Storage and Compilation

The traditional method of storing forest survey data has been to use ASCI files where the attributes of each tree are stored in a single record, with one of the attributes being the expansion factor. For the MR method, the expansion factor changes for every condition class. Thus, the area of each condition class must be available for every tree record and collapsed whenever estimates for a combination of condition classes are desired.

The choice of finite or continuous population approach, regardless of the estimator, also has an effect on the structure of the database. Only plots in the forest stratum need to appear in the database when the continuous population approach is used. For the tessellated approach, every cluster plot, regardless of stratum, must be represented in the database. This is necessary because a portion

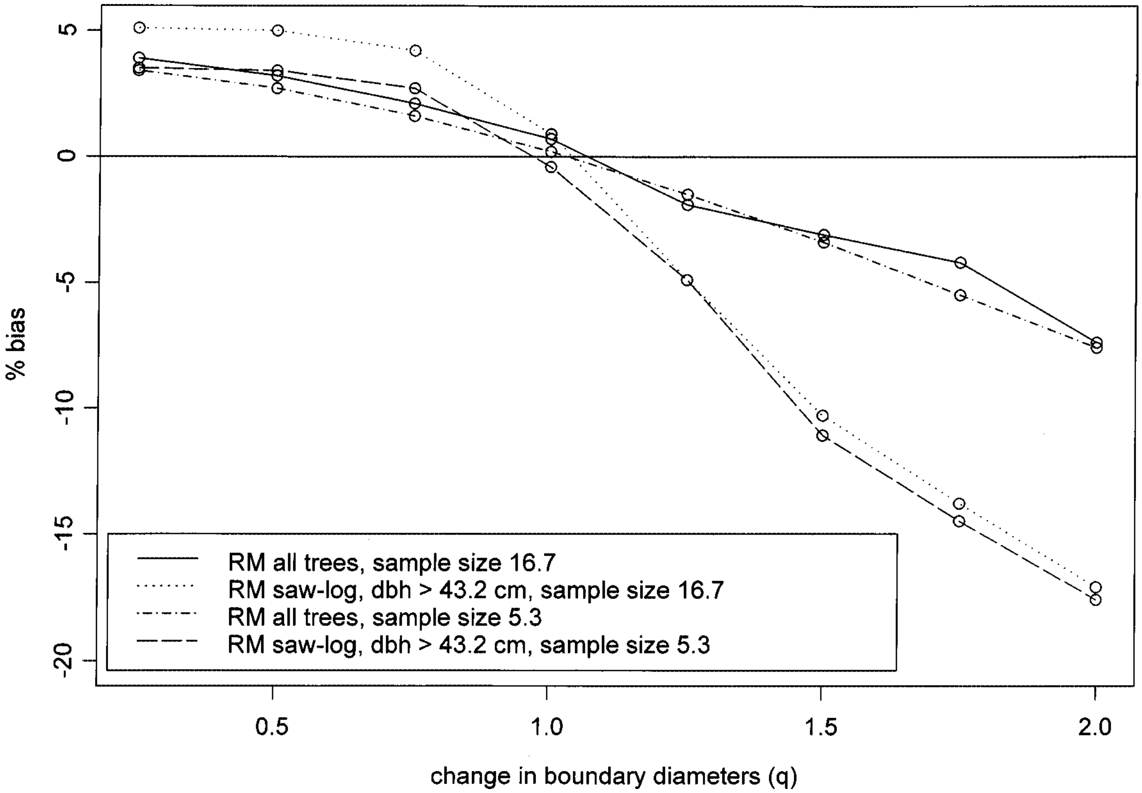

Figure 4. Comparison of the bias of the RM estimators using two different sample sizes. The bias of the estimators for all trees and trees in the sawlog stand with dbh > 43.2 cm is given.

of forestland will likely occur in every stratum and estimates of forest attributes must be calculated for every stratum and added across all strata.

The HT estimator has been used to derive the majority of results described in the literature (e.g., Gregoire 1993). A key example is the estimators for components of change between two periods in time. Results for the RM estimator are limited to topics covered in the papers of Scott and Bechtold (1995) and Bechtold and Zarnoch (1999) and Zarnoch and Bechtold (2000). The author is unaware of any such results for the MR estimator. The results of the HT estimator could be adapted to both estimators with minimal effort.

Summary and Conclusions

The analytical properties for all three estimators and the simulation study show that the MR estimator is clearly the least attractive. The simulation study also shows that the efficiency of the RM estimator is greater than that of the approximate HT estimators for the populations studied. The efficiency of these two estimators may differ the most in extremely fragmented populations because a larger proportion of the subplots could fall outside the boundary of the forest stratum, which would reduce the amount of information collected for the HT estimator. In such cases, the RM estimator would likely produce estimates with a smaller variance. However, an extremely fragmented population is also more likely to have a large degree of edge-effect (Chen et al. 1992). The primary drawback for the RM estimator is its potential to incur a bias due to edge-effect and the poor performance of the classical variance estimator. The jackknife variance estimator has been shown to perform significantly better in many situations and should be considered as an alternative. Of the three estimators, only a true HT estimator [Equation (2)] can produce reliable estimates of both the population total and its variance in all situations. However, it is probably unrealistic to assume that boundary correction techniques can be practically implemented for a cluster plot design. The approximate HT estimators tested exhibited a smaller bias in the presence of edge-effect and the classical variance estimator performed well. However, they exhibited a larger variance than the RM estimators.

A final conclusion is that it is imperative that methods appropriate for either the tessellated or continuous population sampling not be combined.

Literature Cited

Bartlett, R.F. 1986. Estimating the total of a continuous population. J. Stat. Plann. Infer. 13:51–66.

Bechtold, W.A., and S.J. Zarnoch. 1999. Field methods and inventory techniques for mapped inventory plot. *In* Toward a unified framework for inventorying and monitoring forest ecosystem resources, Aguirre-Bravo, C. and C.R. Franco (eds.). USDA For. Serv. RMRS-P-12.

Bool, F.H., and J.L. Locher. 1982. M. C. Escher, his life and complete graphic work: With a fully illustrated catalogue. H.N. Abrams, New York. 349 p.

Campbell, K. 1993. Sampling the continuum: The spatial support of environmental data. P. 11–19 *in* 1993 Proc. Stat. Comput. Sect., Am. Stat. Assoc.

Chen, J., J.F. Franklin, and T.A. Spies. 1992. Vegetation responses to edge environment in old-growth Douglas-fir forests. Ecol. Applic. 2:387–396.

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

Eriksson, M. 1995. Design based approaches to horizontal-point-sampling. For. Sci. 41(4):890–907.

Gregoire, T.G. 1982. The unbiasedness of the mirage correction procedure for boundary overlap. For. Sci. 28:504–508.

Gregoire, T.G. 1993. Estimation of forest growth from successive surveys. For. Ecol. Manage. 56:267–268.

Gregoire, T.G. 1998. Design-based and model-based inference in survey sampling: Appreciating the difference. Can. J. For. Res. 28:1429–1447.

Gregoire, T.G., and C.T. Scott. 1990. Sampling at the stand boundary: A comparison of the statistical performance among eight methods. *In* Research in forest inventory, monitoring, growth and yield. Burkhart, H.E., et al. (eds.). Publ. No. FWS-3-90, Sch. of For. and Natur. Resour., Virg. Polytech. Inst. and State Univ., Blacksbuurg, VA.

Husch, B., C.I. Miller, and T.W. Beers. 1982. Forest mensuration. Ed. 3. Wiley, New York. 402 p.

Milne, A. 1959 The centric systematic area-sample treated as a random sample. Biometrics 15:270–297.

Penttinen, A., D. Stoyan, and H.M. Henttonen. 1992. Marked point processes in forest statistics. For. Sci. 38:806–824.

Pielou, E.C. 1977. Mathematical ecology. Wiley, New York. 385 p.

Reams, G.A., and P.C. Van Deusen. 1999. The southern annual forest inventory system. J. Agric. Biol. Env. Stat. 4:345–359.

Ripley, B.D. 1981. Spatial statistics. Wiley, New York. 252 p.

Royall, R.M. 1992. Robustness and optimal design under prediction models for finite populations. Surv. Methodol. 18:179–185.

Royall, R.M., and W.G. Cumberland. 1981. An empirical study of the ratio estimator and estimators of its variance. J. Am. Stat. Assoc. 76:66–77.

Royall, R.M., and K.R. Eberhardt. 1975. Variance estimates for the ratio estimator. Sankhya. Ser. C. 37:43–52.

Royall, R.M., and J.H. Herson. 1973a. Robust estimation in finite populations I. J. Am. Stat. Assoc. 68:880–889.

Royall, R.M., and J.H. Herson. 1973a. Robust estimation in finite populations II: Stratification on a size variable. J. Am. Stat. Assoc. 68:890–893.

Sarndal, C.E., B. Swensson, and J. Wretman. 1992. Model assisted survey sampling. Springer-Verlag, New York. 694 p.

Schreuder, H.T., T.G. Gregoire, and G.B. Wood. 1993. Sampling methods for multiresource forest inventory. Wiley, New York. 446 p.

Scott, C.T., and W.A. Bechtold. 1995. Techniques and computations for mapping plot clusters that straddle stand boundaries. For. Sci. Monogr. 31:46–61.

Stevens, D.L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics 8:167–195.

Stoyan, D., W.S. Kendall, and J. Mecke. 1995. Stochastic geometry and its applications, Ed. 2. Wiley, New York. 436 p.

Taylor, H.T., and S. Karlin. 1994 An introduction to stochastic modeling. Academic Press, New York. 566 p.

Thompson, S.K. 1992. Sampling. Wiley, New York. 343 p.

Tin, M. 1965. Comparison of some ratio estimators. J. Am. Stat. Assoc. 60:294–307.

Wales, B.A, 1972. Vegetation analysis of north and south edges in a mature Oak-Hickory forest. Ecol. Monogr. 42:451–471.

Williams, M.S., and H.T. Schreuder. 1995. Documentation and evaluation of growth and other estimators for the fully mapped design used by FIA: A simulation study. For. Sci. Monogr. 31:26–45.

Zarnoch, S. J., and W.A. Bechtold. 2000. Estimating mapped-pot forest attributes with ratios of means. Can. J. For. Res. 30:688–697.

APPENDIX

Show the ratio of means estimator is model unbiased for a homogeneous Poisson processes. For simplicity, consider only the subplot portion of (3) and let $\hat{Y}_{RM}$ estimate the number of trees.

Assume a homogeneous Poisson process of intensity $\lambda = Y/A_F$, where Y is the total number of trees in the forested area. Define $y^{sub}_{h..}$ be the number of trees within a randomly located plot of area $a^{sub}_{h.}$. Then $y^{sub}_{h..}$ has a Poisson distribution with mean and variance equal to $\lambda a^{sub}_{h.}$. Thus, the model describing the relationship between the number of trees and plot area is

$$y^{sub}_{h..} = \lambda s^{sub}_{h.} + \varepsilon_h$$

where $E[\varepsilon_h] = 0$, $Var[\varepsilon_h] = \lambda a^{sub}_{h.}$ and $Cov[\varepsilon_h, \varepsilon_{h'}] = 0$. For simplicity let the number of cluster plots be m=1 and assume that the area of forest is known ($\hat{A}_F = A_F$).

Taking the expected value of the Taylor's series expansion yields

$$E[\hat{Y}_{RM}] \approx A_F \left[\frac{E[y^{sub}_{h..}]}{E[a^{sub}_{h.}]} - \frac{1}{E[a^{sub}_{h.}]^2} Cov[y^{sub}_{h..}, a^{sub}_{h.}] + \frac{E[y^{sub}_{h..}]}{E[a^{sub}_{h.}]^3} Var[a^{sub}_{h.}] \right]$$

Substituting $y^{sub}_{h..} = \lambda a^{sub}_{h.} + \varepsilon$ yields

$$AF \left[\frac{\lambda E[a^{sub}_{h.}]}{E[a^{sub}_{h.}]} - \frac{\lambda}{E[a^{sub}_{h.}]^2} Var[a^{sub}_{h.}] + \frac{\lambda E[a^{sub}_{h.}]}{E[a^{sub}_{h.}]^3} Var[a^{sub}_{h.}] \right]$$
$$= A_F \lambda = Y$$

Thus, regardless of sample size, the ratio of means estimator is unbiased with respect to the model describing a homogeneous Poisson process. This results also holds for any point process with constant intensity λ (Stoyan et al. 1995, p. 102–103).