4 Sample-Based Estimators Used by the Forest Inventory and Analysis National Information Management System

**Charles T. Scott, William A. Bechtold,
Gregory A. Reams, William D. Smith,
James A. Westfall, Mark H. Hansen, and
Gretchen G. Moisen[1]**

4.1 Panels and Estimation

This chapter outlines **prescribed core**[2] procedures for deriving **population** estimates from **attributes** measured in conjunction with the **Phase 1** and **Phase 2** samples. These estimation procedures also apply to those **Phase 3** attributes in common with Phase 2. Given the sampling frame and **plot** design described in the previous two chapters, many estimation approaches can be applied. In fact, one goal of the overall design is to maximize flexibility, so Forest Inventory and Analysis (FIA) data can be used to address a variety of analytical needs.

Much of the flexibility inherent in the Enhanced FIA design is derived from the way **panels** are combined for analysis. This topic is addressed in chapter 5. For estimations involving a single panel or a **periodic survey**, the approach to estimation would proceed directly as outlined in chapter 4. When the estimation is for some combination of annual panels, then the estimation procedures discussed herein may require modification depending on the method used to combine the panels. Related modifications are discussed in conjunction with the two specific methods for combining panels presented in chapter 5 (secs. 5.2.1 and 5.2.2).

---

[1] Charles T. Scott, FIA Program Manager, USDA Forest Service, Northeastern Research Station, Newtown Square, PA 19073; William A. Bechtold, Research Forester, USDA Forest Service, Southern Research Station, Asheville, NC 28802; Gregory A. Reams, FIA National Program Manager, USDA Forest Service, 1601 North Kent Street, Arlington, VA 22209; William D. Smith, Research Quantitative Ecologist, USDA Forest Service, Southern Research Station, Research Triangle Park, NC 27709; James A. Westfall, Research Forester, USDA Forest Service, Northeastern Research Station, Newtown Square, PA 19073; Mark H. Hansen, Research Forester, USDA Forest Service, North Central Research Station, St. Paul, MN 55108; and Gretchen G. Moisen, Mathematical Statistician, USDA Forest Service, Rocky Mountain Research Station, Ogden, UT 84401.

[2] First use of a glossary term in each chapter is in bold face.

4.2 Phase 1

Stratification is a statistical tool used to reduce the variance of attributes of interest by partitioning the population into homogenous **strata**, such as **forest** and **nonforest**. It may also involve partitioning out a highly variable but small portion of the population.

To use stratified sampling methods, the strata sizes (weights) must be either known or estimated. As discussed in chapter 2, strata weights commonly are estimated by classifying points on aerial photographs or **pixels** on satellite imagery. Full enumeration of an entire landscape (population) is possible with wall-to-wall satellite imagery, resulting in strata weights that can be treated as known values in the variance formulations. Whether the strata weights are known or estimated, strata classifications (e.g., forest and nonforest) from remote sensing are never perfectly accurate. Although this increases the variance of the resulting population estimates, it does not introduce any bias due to the interaction between Phases 1 and 2 in the estimation process.

4.2.1 Satellite Classification and Known Strata Weights

Using wall-to-wall satellite imagery and computer-aided classification, the population can be divided into strata of known size—typically forest, **noncensus water**, nonforest, or combinations of these. In this case the classified imagery divides the total area of the population (A_T) into pixels of equal and known size (e.g., 30 m square with Landsat TM), and the classification assigns one of H strata to each of these pixels. Here, the stratum weight, W_h $(h = 1,\ldots,H)$, typically is determined as the number of pixels classified as stratum h divided by the total number pixels in the population of interest. The weights are known quantities that are fixed by the classification system and selected strata. As used here, "pixel" indicates either an individual pixel, or a non-overlapping block of multiple pixels. Each stratum, h, then contains n_h ground plots, each selected with known probability, where the Phase 2 attributes of interest are observed. Note that the strata sample sizes, n_h, are random because ground-plot locations are chosen prior to stratification.

Satellite classification systems separate the reflectance values from each pixel into a set of H distinct values (i.e., ranges). Such systems can range from very complicated functions dependant on a number of different reflectance band values and ancillary data layers, to very simple step functions of a single reflectance band. A variety of automated classification schemes are available, but these rarely match FIA stratification requirements. These must usually be modified or additional classifications must be performed to produce strata that are relevant to FIA estimation needs.

When using **classified satellite imagery** for stratification, one must know the location of each plot and each pixel, so that plots and pixels can be linked in the estimation process. FIA assigns each plot to one and only one stratum using the pixel corresponding to plot center. Typically this is done using a global positioning system (GPS) instrument for ground plots and georeference information associated with the imagery for pixels. If GPS field data are not available, plot locations can be digitized manually from aerial photos or maps.

Most of the estimation presented here is based on the assumption of equal probability sampling, where all elements of the population have the same probability of being sampled by a ground plot. This assumption can be violated if ground-plot information is used to help classify satellite imagery. Inevitably, ground-plot data will be used in the development of some classification algorithms. Breidt and Opsomer (2004) show that for algorithms based on general linear models, this approach provides valid results and variances. Other types of classification algorithms (e.g., nonparametric) have not been similarly evaluated.

It is also important to remember the importance of the pixel-plot link when changing or considering a change in the stratification algorithm. Two similar algorithms may divide a population into very similar strata that have similar or identical names. The application of a different classification system will not only change the W_h (even when they have identical names), but also the strata assignments of ground plots.

4.2.2 Aerial Photography and Estimated Strata Weights

In some cases, especially when using aerial photos, it is not practical or cost effective to divide imagery into strata of known sizes. Strata sizes then may be estimated using photo-plot sampling with manual interpretation, and applying **double sampling for stratification** as opposed to the **stratified estimation** that occurs when strata weights are known. When double sampling for stratification is used in conjunction with satellite imagery or other spatial layers, only a sample of pixels is classified, as opposed to the wall-to-wall estimation described in the previous section. This relieves some of the computational burden of more complex computer-aided classification methods. The loss of precision when such methods are applied to a subset of pixels is minimal when large Phase 1 sample sizes are used (Moisen and Edwards 1999).

Under double sampling for stratification, a set of sample points (i.e., Phase 1 plots) is placed on aerial photo or satellite coverage of the population. Phase 1 plots are then assigned to strata by photo interpretation or computer-aided classification algorithm. Phase 1 plots are usually placed on a grid, typically every 200 to 300 acres. The Phase 2 and Phase 3 plots usually are a subset of the Phase 1 photo plots. Any Phase 2 and Phase 3 plots that are not must also be classified, but their values are not used in developing the estimates of strata sizes. Here:

n' = total number of Phase 1 plots (pixels) sampled in the population

n'_h = number of Phase 1 plots (pixels) classified as belonging to stratum h ($h = 1,\ldots,H$, H = total number of strata)

4.2.3 Combining Small Strata and Populations

Because sample sizes are random with respect to strata, some may contain too few plots to compute a reliable mean and variance. Experience indicates that a minimum of four Phase 2 plots per stratum are required. If less than four, then similar strata must be collapsed (combined) until the minimum is attained. Because stratification schemes may vary regionally to accommodate resource differences, there are no prescribed rules by which strata are combined.

The minimum-sample-size rule also affects some small populations or populations with little forest (e.g., some counties in the Eastern or Great Plains regions of the United States). Populations with fewer than 12 forested Phase 2 plots should thus be combined with adjoining counties (i.e., populations) forming what is termed as a "**super-county**." The use of super-counties will be especially important when using only a subset of panels.

4.3 Phase 2

This chapter builds on chapter 3 by specifying the population estimators used for calculation of area and other attributes of interest associated with **mapped plots**. Generally, these attributes are summed for each plot (after adjustment to correct for any plots that overlap the population boundary), and then averaged across plots in the stratum. The strata means are then combined using the Phase 1 strata weights to form a population mean. This mean is then expanded to a total using the total land area in the population. This approach to estimation is first described for area attributes, and then other attributes, because the formulas are somewhat different. Examples of the estimation procedures discussed below are provided in the supplementary documents "Examples of FIA Point-in-Time Estimation Procedures for Several Common Cases" and "Examples of FIA Change-Component Estimation Procedures for Several Common Cases" at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.3.1 Estimation of Area by Stratum by Domain (Row and Column) Attributes

For a given attribute of interest (e.g., forest area), estimates of population totals and **domain** totals (e.g., area by forest type and stand size class) are produced similarly. The only difference is that different restrictions (filters) are placed on each cell of an output table depending on which domain (row or column) the cell is in. Mapped **condition classes** (e.g., forest type) are often used to specify domains, but other attributes also can be used (e.g., species).

Row and column attributes must be discrete categorical variables. Continuous attributes can be converted into discrete categories by dividing them into classes (e.g., diameter classes). The attribute of interest is "summed" into a cell only when it satisfies the row and column requirements (i.e., when it is in the domain of interest). Each plot and tree has an associated **indicator function**, δ_{hijkd} (or δ_{hijtd}), which is 1 if the attribute is in the domain d of interest, or 0 if not. For example, when estimating the area in northern hardwood (row) sapling stands (column), the indicator function is 1 when the specified forest type and stand-size requirement are satisfied; otherwise it is 0. This method is described in Cochran (1977) for estimating domain (cell) means.

4.3.1.1 For Each Table Cell, Compute the Attribute of Interest for Each Plot

Each plot is assigned to only one stratum based on the Phase 1 stratification of the plot center. For area estimation, the attribute of interest is the proportion of the plot in the **domain** of interest:

$$P_{hid} = \frac{\sum_{j}^{4} \sum_{k}^{K_{hij}} a_{mhijk} \delta_{hijkd}}{a_m \overline{p}_{mh}} \tag{4.1}$$

where

P_{hid} = proportion of plot i in the domain of interest d, for plots assigned to stratum h, adjusted for stratum h plots that overlap the population **boundary**

a_{mhijk} = mapped area (acres) of **subplot** (**macroplot**) j covering condition k on plot i assigned to stratum h (The area is computed using the largest area mapped, which is the subplot except in the Pacific Northwest (PNW) where the macroplot is used.)

δ_{hijkd} = zero-one domain indicator function, which is 1 if condition k on subplot (macroplot) j of plot i assigned to stratum h belongs to the domain of interest d

K_{hij} = the number of conditions that exist on subplot (macroplot) j of plot i assigned to stratum h

a_m = total area of the largest-sized plot on which area attributes are mapped (i.e., four times the subplot or macroplot area)

$\bar{p}_{mh}$ = mean proportion of stratum h mapped plot areas falling within the population (see equation 4.2)

Table margins (row and column totals) are treated like any other cell. Additivity is a property of the table construction, not the estimators. Most tables will be additive (i.e., the cells in the table body will add to totals in the margins such as area by forest type and stand size). However, a table of the area containing combinations of species and diameter classes will not be additive, because each cell represents the number of acres on which the particular species and diameter class combination occurs.

4.3.1.2 Adjustment for Partial Plots Outside the Population

Equation 4.1 essentially treats all areas sampled on stratum h plots equally—every square foot is expanded equally whether it is part of a partial plot or not. For a given population, dividing by $\bar{p}_{mh}$ in equation 4.1 adjusts the plot observations to account for any portions of stratum h plots falling outside the population. Note that $\delta_{hijkd} = 0$ for condition classes outside the population in equation 4.1, because conditions outside the population are never conditions of interest. Reasons for condition classes being outside the population include:

- Partial plots that straddle an international boundary (i.e., Canada or Mexico)
- Partial plots that straddle a mapped ownership population boundary (e.g., national forest)
- Whole or partial plots within the population that are nonsampled (e.g., denied access or hazardous conditions)
- Whole or partial plots in **census water** (unless census water is estimated, i.e., included in A_T) (Note that FIA currently subtracts census water from A_T, but anticipates that census water will be estimated when precise digitized census water boundaries become available from the U.S. Census Bureau. If and when census water is estimated from the FIA sample, a Phase 1 stratum will likely be created for it.)

If all plots are entirely within the population boundaries, then $\overline{p}_{mh}$ would be 1. Otherwise, the average mapped-plot area actually sampled within the population is divided by the standard plot area, a_m. This approach was taken as a way of handling the potential bias introduced by ignoring portions of plots straddling population boundaries. Essentially, this creates a buffer around the population to ensure that plots which are only partially inside the population of interest are included in the estimation. More area is sampled, but more plots are also taken into the sample. Multiplying the totals by the larger area is on average equivalent to the adjustment made by using $\overline{p}_{mh}$:

$$\overline{p}_{mh} = \sum_{i}^{n_h} \sum_{j}^{4} \sum_{k}^{K_{hij}} \frac{a_{mhijk} \delta_{hijk}}{a_m n_h} \tag{4.2}$$

where

n_h = number of ground plots with Phase 1 assignments to stratum h (For initial area tables, this includes all plots sampled with any portion of a subplot (macroplot) in the population. For subsequent tables, any plots that are entirely nonsampled are excluded.)

δ_{hijk} = zero-one in-sample indicator function, which is 1 if condition k on subplot (macroplot) j of plot i assigned to stratum h is within the boundaries of the population (Nonsampled areas are included in initial area tables in order to estimate their areas, but are zero otherwise. Missing values are also treated as zero values.)

Clearly, there are cases where $\overline{p}_{mh}$ will not be 1, meaning that $\overline{p}_{mh}$ is a constant for a given stratum, but varies between strata and populations. The variation due to nonsampled plots and plots extending beyond population boundaries is expected to be small enough to be ignored. One particular area of concern is the checkerboard ownership pattern in the West where National Forest System (NFS) boundaries may be treated as population boundaries. FIA is evaluating the frequency of plots that straddle NFS boundaries and may switch to the **ratio-of-means** estimators described by Zarnoch and Bechtold (2000), if necessary.

4.3.1.3 Compute Strata Means and Variances

Plot values are averaged within each stratum. In the case of **simple random sampling**, this is the final estimate because simple random sampling is just stratified sampling with a single stratum.

The stratum mean is the sum of the plot observations, P_{hid}, divided by the number of plots in the stratum, n_h:

$$\overline{P_{hd}} = \frac{\sum_{i}^{n_h} P_{hid}}{n_h} \tag{4.3}$$

with variance estimator:

$$v\left(\overline{P_{hd}}\right) = \frac{\sum_{i}^{n_h} P_{hid}^2 - n_h \overline{P_{hd}}^2}{n_h\left(n_h - 1\right)} \tag{4.4}$$

Note that $\overline{P_{hd}}$ is the mean of field-based observations in domain d within each Phase 1 stratum. This means that some plots classified as forest on the ground may have been assigned to a nonforest stratum and vice versa. These are not viewed as "misclassifications", but as strata with less than ideal homogeneity. The estimators remain unbiased.

4.3.2 Estimation of Population Totals and Their Variances

Generally, individual counties are the populations of interest (i.e., the basic building blocks for estimation). Counties may be divided into **subpopulations** that are processed independently. This is the case when a portion of a county has an intensified Phase 2 sampling grid, has enumerated acreages, or has a measurement **cycle** that differs from the rest of the county. These scenarios are not uncommon when sampling land owned by public agencies [i.e., NFS, Bureau of Land Management (BLM), and National Park Service (NPS)]. Because populations and subpopulations are mutually exclusive, estimated totals are additive. Likewise, because different populations and subpopulations are independent, the associated variance estimates are also additive. Thus, totals from groups of counties can be combined to formulate State and regional totals; or segments of NFS land, by county, can be combined to yield totals for a specific national forest. County areas provided by the U.S. Census Bureau, which are used in the estimation of population totals, are updated at least every 10 years. NFS and BLM provide similar area totals for their lands, and totals by county if they have intensified or otherwise altered the sampling effort on their lands.

As first noted in equation 4.2, nonsampled plots present an estimation problem that requires more attention than plots that are otherwise out of the population. Because field crews cannot reliably ascertain the actual land use of such plots, initial area tables in FIA reports will report the proportion of total area designated as nonsampled. However, in all subsequent tables, these plots (or the nonsampled portions of them) will be processed as if they were out of the population. This approach either reduces the sample size (n_h), or the mean proportion of stratum h observed plot areas falling within the population, or both. The result is to adjust upward the number of acres represented by the accessed plots, or portions of them, in each stratum. Thus the area that could not be accessed is redistributed based on the accessible plots within each stratum. This essentially replaces nonsampled plots with the strata means and increases the strata variances due to the reduction in sample size. This approach has the advantage of simplicity, but has the potential to incur bias if the nonsampled plots are not representative of the rest of the population. Other methods of accounting for nonsampled areas are under investigation, including remote sensing both for direct measurement of a subset of attributes and for use in identifying similar plots for imputation purposes.

In the simple random sampling case, there is only one stratum, so totals are estimated by multiplying the population mean by the total area in the population. FIA rarely uses simple random sampling, but it has been used for Forest Health Monitoring (FHM) and is useful in responding to special requests where stratification data may be lacking. Processing protocols for the FIA plot design under a simple random sampling approach using ratios of means are discussed in detail by Zarnoch and Bechtold (2000).

To estimate the total area in each domain when the population has been stratified, the strata means are averaged using the strata weights and then multiplied by the total land area in the population. The estimated total is given for the stratified estimation and double sampling for stratification cases, respectively, as:

$$\hat{A}_d = A_T \sum_h^H W_h \overline{P_{hd}} = A_T \overline{P_d} \tag{4.5}$$

and

$$\hat{A}_d = A_T \sum_h^H \frac{n'_h}{n'} \overline{P_{hd}} = A_T \overline{P_d}$$

where

A_T = total area in the population in acres

$\overline{P_d}$ = estimated proportion of the population in the domain of interest d

W_h = weight for stratum h (i.e., the proportion of the population area, A_T, that is in stratum h)

An approximation of the variance of the total area in the stratified estimation (and the simple random sampling) case, where strata sizes are known, is adapted from equation 5A.42 in Cochran's (1977) finite sampling framework. An estimate of this population variance (finite population correction factors ignored) is given by:

$$v\left(\hat{A}_d\right) = \frac{A_T^{\,2}}{n}\left[\sum_h^H W_h n_h v\left(\overline{P_{hd}}\right) + \sum_h^H \left(1 - W_h\right)\frac{n_h}{n} v\left(\overline{P_{hd}}\right)\right] \qquad (4.6)$$

The first term is for stratification, assuming proportional allocation, and the second term reflects the fact that the sample sizes are not fixed by strata in advance.

Double sampling for stratification applies when the strata weights are estimated. The variance of total area in this case is adapted from equation 12.32 in Cochran (1977), again ignoring finite population correction factors. The second term accounts for the fact that strata sizes are estimated:

$$v\left(\hat{A}_d\right) = A_T^2\left\{\sum_h^H \left(\frac{n'_h - 1}{n' - 1}\right)\frac{n'_h}{n'} v\left(\overline{P_{hd}}\right) + \frac{1}{n' - 1}\sum_h^H \frac{n'_h}{n'}\left(\overline{P_{hd}} - \overline{P_d}\right)^2\right\} \qquad (4.7)$$

As noted above, totals and their variances from different populations are additive. Thus the variance of a total across multiple populations or subpopulations is the sum of their variances. Examples of how to apply equations 4.1 through 4.7 are given in spreadsheet form in the supplementary document "Examples of FIA Point-in-Time Estimation Procedures for Several Common Cases" at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm (see sections "plot summary", "problem number 1", and "problem number 2").

4.3.3 Estimation of Other Attributes

Population totals for attributes other than area usually are calculated by summing attributes to the plot level and then averaging at the stratum level. Indicator functions are used to identify the attribute of interest (e.g., total volume of white oak) in the domain of interest (e.g., oak-pine stands). The attribute of interest is summed for each plot and then divided by the observed plot area and the mean proportion of stratum *h* observed plot areas falling within the population, yielding an estimate of the attribute of interest on a per-unit-area basis:

$$y_{hid} = \frac{\sum_{j}^{4}\sum_{t} y_{hijt}\delta_{hijtd}}{a_o\overline{p}_{oh}} \tag{4.8}$$

where

y_{hijt} = attribute of interest for tree t on macroplot, subplot, or **microplot** j of plot i assigned to stratum h

δ_{hijtd} = zero-one domain indicator function, which is 1 if tree t on subplot j of plot i assigned to stratum h belongs to the domain of interest d

a_o = total area normally used to observe the attribute of interest on a plot (i.e., four times the microplot, subplot, or macroplot area)

$\overline{p}_{oh}$ = mean proportion of stratum h observed-plot areas falling within the population (see equation 4.9)

In equation 4.8, dividing by $\overline{p}_{oh}$ adjusts the plot observations to account for any portions of stratum h plots falling outside the population:

$$\overline{p}_{oh} = \sum_{i}^{n_h}\sum_{j}^{4}\sum_{k}^{K_{hij}} \frac{a_{ohijk}\delta_{hijk}}{a_o n_h} = \frac{1}{a_o n_h}\sum_{i}^{n_h}\sum_{j}^{4}\sum_{k}^{K_{hij}} a_{ohijk}\delta_{hijk} \tag{4.9}$$

where

a_{ohijk} = area normally used to observe the attribute of interest (microplot, subplot, or macroplot j) covering condition k on plot i assigned to stratum h

In equation 4.8, dividing by the observed plot area and by the proportion of plots outside the population allows attributes such as number of trees across all diameter classes to be summed across plot types while accounting for any differences in the proportion of the various plot types that are outside the population. For example, if y_{hid} is the number of trees 1.0-inch diameter at breast height (**d.b.h.**) and larger sampled on microplots and subplots, then equation 4.8 should include a term for each plot type. This adjustment is the main difference between the estimators described herein and the ratio-of-means estimators outlined by Zarnoch and Bechtold (2000), where population totals are computed on the basis of each plot size separately, then summed for all plot sizes. Because totals for the latter alternative are not independent, the variance is complicated by the need to include a covariance term among plot sizes. We opted for a simpler approach, where all estimates are combined at the plot level and then treated as a single attribute of interest when calculating population totals. Using this method, the variance is much easier to compute, and the resulting estimate (equation 4.10) is equivalent to the separate approach. When combining subplot and microplot values, the attribute of interest is computed as:

$$y_{hid} = \frac{\sum_{j}^{4}\sum_{t} y_{hijt}\,\delta_{hijtd}}{a_o \overline{p}_{oh}} + \frac{\sum_{j}^{4}\sum_{t} y'_{hijt}\,\delta'_{hijtd}}{a'_o\,\overline{p}'_{oh}} \tag{4.10}$$

where

y'_{hijt} = attribute of interest for tree t on microplot j of plot i in stratum h

δ'_{hijtd} = zero-one domain indicator function, which is 1 if tree t on microplot j of plot i assigned to stratum h belongs to the domain of interest d

a'_o = total microplot area

$\overline{p}'_{oh}$ = mean proportion of stratum h microplot areas falling within the population (equation 4.9)

The plot attributes from either equation 4.8 or equation 4.10 are then summed across all plots within each stratum and divided by the total number of plots in the stratum to yield the stratum mean of the y_{hid}:

$$\overline{Y_{hd}} = \frac{\sum_{i}^{n_h} y_{hid}}{n_h} \tag{4.11}$$

with estimated variance:

$$v\left(\overline{Y_{hd}}\right) = \frac{\sum_{i}^{n_h} y_{hid}^2 - n_h \overline{Y_{hd}}^2}{n_h\left(n_h - 1\right)} \tag{4.12}$$

As was the case for area, the strata means are averaged using the strata weights, then multiplied by the total land area. The estimated total is given for the stratified estimation and double sampling for stratification cases, respectively, as:

$$\hat{Y}_d = A_T \sum_{h}^{H} W_h \overline{Y_{hd}} = A_T \overline{Y_d} \tag{4.13}$$

or

$$\hat{Y}_d = A_T \sum_{h}^{H} \frac{n'_h}{n'} \overline{Y_{hd}} = A_T \overline{Y_d}$$

where

$\overline{Y_d}$ = population mean of the attribute of interest in the domain of interest *d*

As with the area estimate when strata weights are known, an approximation to the variance of the attribute total in the stratified estimation (and the simple random sampling) case was developed in the same way as for equation 4.6:

$$v\left(\hat{Y}_d\right) = \frac{A_T^2}{n}\left[\sum_{h}^{H} W_h n_h v\left(\overline{Y_{hd}}\right) + \sum_{h}^{H} \left(1 - W_h\right)\frac{n_h}{n} v\left(\overline{Y_{hd}}\right)\right] \qquad (4.14)$$

The estimated variance of the attribute total in the double sampling for stratification case is again adapted from equation 12.32 in Cochran (1977):

$$v\left(\hat{Y}_d\right) = A_T^2\left\{\sum_{h}^{H}\left(\frac{n'_h - 1}{n' - 1}\right)\frac{n'_h}{n'} v\left(\overline{Y_{hd}}\right) + \frac{1}{n' - 1}\sum_{h}^{H}\frac{n'_h}{n'}\left(\overline{Y_{hd}} - \overline{Y_d}\right)^2\right\} \qquad (4.15)$$

As noted in section 4.3.2, totals and their variances from different populations are additive for a given domain. Thus the variance of a total across multiple populations or subpopulations is the sum of their variances. Although the additivity property is also true for means, sums of means generally are not useful. To obtain a mean over populations, the totals should be added and divided by the total area of the populations. The result is a mean weighted by the population areas. An example of how to apply equations 4.8 through 4.15 is given in problem number 3 of the supplementary document "Examples of FIA Point-in-Time Estimation Procedures for Several Common Cases" at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.3.4 Estimation of Ratios

Often, interest is not in the totals but in the attribute of interest expressed on a per-acre, per-condition (stand), or per-tree basis. An approach that is also compatible with the aforementioned estimates of population totals is the ratio-of-means estimator, wherein the numerator is the estimated attribute total and the denominator depends on the ratio to be estimated. The three cases can all be estimated using one of the following general formulas. The first is for stratified estimation where strata weights are known and the second for double sampling where strata weights are estimated:

$$\hat{R}_{dd'} = \frac{\hat{Y}_d}{\hat{X}_{d'}} = \frac{\sum_h^H W_h \overline{Y_{hd}}}{\sum_h^H W_h \overline{X_{hd'}}}$$

or (4.16)

$$\hat{R}_{dd'} = \frac{\hat{Y}_d}{\hat{X}_{d'}} = \frac{\sum_h^H \frac{n'_h}{n'} \overline{Y_{hd}}}{\sum_h^H \frac{n'_h}{n'} \overline{X_{hd'}}}$$

The strata means for the denominator, $\overline{X_{hd'}}$, are computed in the same manner as the numerator. Note that the numerator and denominator have different domains of interest, with *d* being a subset of *d′* (e.g., 12-inch oaks in oak-hickory stands). If the denominator is an area attribute, then $x_{hid'}$ is estimated using equation 4.1 and the area attribute replaces P_{hid}. If the denominator is a tree or other attribute, then $x_{hid'}$ is estimated using equation 4.8 and the attribute replaces y_{hid}. The choice of the individual plot observations, $x_{hid'}$, is described in sections 4.3.4.1 through 4.3.4.3 for three common situations.

The variance estimator from equation 5.6.10 in Särndal and others (1992) is:

$$v(\hat{R}_{dd'}) = \frac{1}{\hat{X}^2_{d'}} \left[v(\hat{Y}_d) + \hat{R}^2_{dd'} v(\hat{X}_{d'}) - 2\hat{R}_{dd'} \operatorname{cov}(\hat{Y}_d, \hat{X}_{d'}) \right] \quad (4.17)$$

In the stratified estimation case, the right-hand-side variances are computed using equation 4.6 for area attributes and equation 4.14 for other attributes. The covariance is estimated as:

$$\operatorname{cov}(\hat{Y}_d, \hat{X}_{d'}) = \frac{A_T^2}{n} \left[\sum_h^H W_h n_h \operatorname{cov}(\overline{Y_{hd}}, \overline{X_{hd'}}) + \sum_h^H (1 - W_h) \frac{n_h}{n} \operatorname{cov}(\overline{Y_{hd}}, \overline{X_{hd'}}) \right] \quad (4.18)$$

where

$$\operatorname{cov}(\overline{Y_{hd}}, \overline{X_{hd'}}) = \frac{\sum_i^{n_h} y_{hid} x_{hid'} - n_h \overline{Y_{hd}}\,\overline{X_{hd'}}}{n_h (n_h - 1)} \quad (4.19)$$

In the double sampling for stratification case, the variance is approximated by assuming the covariance can be computed similarly to the area attributes (equation 4.7) and other attributes (equation 4.15)

$$\mathrm{cov}\left(\hat{Y}_d, \hat{X}_{d'}\right) = A_T^2 \left\{ \sum_h^H \left(\frac{n'_h - 1}{n' - 1}\right) \frac{n'_h}{n'} \mathrm{cov}\left(\overline{Y_{hd}}, \overline{X_{hd'}}\right) + \frac{1}{n' - 1} \sum_h^H \frac{n'_h}{n'} \left(\overline{Y_{hd}} - \overline{Y_d}\right)\left(\overline{X_{hd'}} - \overline{X_{d'}}\right) \right\} \qquad (4.20)$$

Equations 4.16 and 4.17 can be used for all three estimation of ratio cases—to express values on a per-acre, per-tree, or per-condition basis.

4.3.4.1 Estimation on a Per-Acre Basis

Often interest is in expressing the attribute of interest on a per-acre basis. This can be estimated by dividing $\hat{Y}_d$ in equation 4.13 by the total surface area in the population, A_T. Because the area is known, the variance of the ratio is simply the variance of the total (equation 4.14) divided by the square of the total area, A_T.

However, interest is more commonly in the attribute total expressed on a per-forested-acre basis—a ratio estimate. The denominator of equation 4.16 can be the estimate of total forest area $(\hat{A}_F)$ computed using equation 4.5, or the area of another domain such as the area in a specific forest type. When the denominator is derived from equation 4.5, the $x_{hid'}$ is equal to P_{hid} in equation 4.1 which is then used to compute the X values in equations 4.16 through 4.20.

Indicator functions are used to specify the domains and attributes of interest that define y_{hi} and x_{hi}. These may or may not change for various cells in the tables, depending on the ratio of interest. For example, the value of $x_{hid'}$ can be the same for all values in a table, such as those used in computing $\hat{A}_F$. This is useful for estimating additive tables, such as stand tables with numbers of trees per acre by species and diameter class. Alternatively, the value of $x_{hid'}$ might change for each cell, such as the area by forest type and stand size, so that the attribute average can be estimated for each combination (domain). An example of how to apply equations 4.16 through 4.20 for estimation on a per-acre basis is given in problem number 4 of the supplementary document "Examples of FIA Point-in-Time Estimation Procedures for Several Common Cases" at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.3.4.2 Estimation on a Per-Condition (Per-Stand) Basis

A few attributes are only observed at the condition (or stand) level, such as stand age. To compute average stand age requires a slight modification of the approach to the y_{hid} attribute. The approach is to weight the condition attribute, y_{hik}, by the area in the condition of interest on the plot:

$$y_{hid} = \frac{\sum_{j}^{4} \sum_{k}^{K_{hij}} y_{hik} a_{mhijk} \delta_{hijkd}}{a_m \overline{p}_{mh}} \tag{4.21}$$

The denominator reflects the proportion of the area sampled in the condition of interest:

$$x_{hid'} = \frac{\sum_{j}^{4} \sum_{k}^{K_{hij}} a_{mhijk} \delta_{hijkd'}}{a_m \overline{p}_{mh}} \tag{4.22}$$

The result is the average of the attribute of interest weighted by the area in which it occurs. These values of y_{hid} and $x_{hid'}$ are then used in equations 4.16 through 4.20. An example of how to apply equations 4.16 through 4.22 for estimation on a per-stand basis is given in problem number 5 of the supplementary document “Examples of FIA Point-in-Time Estimation Procedures for Several Common Cases” at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.3.4.3 Estimation on a Per-Tree Basis

For some applications, the attribute of interest is an individual tree attribute, such as average tree height or conks per tree, where the ratio of interest is expressed on a per-tree basis. The ratio estimator then becomes the estimate of the population total for the attribute of interest, $\hat{Y}_d$, divided by the total number of trees, $\hat{X}_{d'}$, in the population. Thus, y_{hid} is the sum of tree attributes observed on plot *i* in stratum *h* in the domain of interest; and $x_{hid'}$ is the number of trees observed on plot *i* in stratum *h* in the domain of interest. Those values are then used in equations 4.16 through 4.20. An example of how to apply equations 4.16 through 4.20 for estimation on a per-tree basis is given in problem number 6 of the supplementary document “Examples of FIA Point-in-Time Estimation Procedures for Several Common Cases” at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.3.5 Computation of Sampling Errors

Sampling errors are used by FIA to reflect the accuracy of the estimates. Expressed on a percentage basis in order to enable comparisons between the precision of different estimates, sampling errors are computed by dividing the estimate into the square root of its variance:

$$S.E.\% = 100\frac{\sqrt{v\left(\hat{Y}_d\right)}}{\hat{Y}_d} \tag{4.23}$$

Assuming normality of the distribution of estimates, the percent sampling error can be used to compute an approximate 67 percent confidence interval. If the sampling error is doubled, then an approximate 95 percent confidence interval can be formed. A better approximation is achieved by using the appropriate Student's *t*-values.

4.3.6 Components of Change

FIA inventories are designed to measure net change over time, as well as the individual **components of change** that constitute net change (e.g., growth, removals, mortality). Change estimates are computed for two sequential measurements of each inventory panel. Upon remeasurement, a new initial inventory is established for remeasurement at the next scheduled inventory. As such, computation of change components is not intended to span more than one inventory cycle. Rather, the change estimation process is repeated cycle by cycle. This simplifies field protocols and ensures that change estimation is based on short and relatively constant time intervals (e.g., 5 years). Change estimates for individual panels are combined across multiple panels in the same manner as panels are combined to obtain current inventory parameters such as total standing volume.

FIA recognizes the following components of change as prescribed core variables; they usually are expressed in terms of growing-stock or all-live volume, where t is the initial inventory of a measurement cycle, and $t+1$ is the terminal inventory:

G_S = survivor growth—the growth on trees tallied at time t that survive until time $t+1$.

I = ingrowth—the volume of trees at the time that they grow across the minimum d.b.h. threshold between time t and time $t+1$. The estimate is based on the size of trees at the d.b.h. threshold which is 1.0 inch for all-live trees and 5.0 inches for growing-stock trees. This term also includes trees that subsequently die (i.e., ingrowth mortality), are cut (i.e., ingrowth cut), or diverted to nonforest (i.e., ingrowth diversion); as well as trees that achieve the minimum threshold after an area reverts to a forest-land use (i.e., reversion ingrowth).

G_I = growth on ingrowth—the growth on trees between the time they grow across the minimum d.b.h. threshold and time $t+1$.

R = reversion—the volume of trees on land that reverts from a **nonforest land** use to a **forest land** use (or, for some analyses, land that reverts from any source to **timberland**) between time t and time $t+1$. The estimate is based on tree size at the midpoint of the measurement interval. Tree size at the midpoint is modeled from tree size at time $t+1$.

G_R = reversion growth—the growth of reversion trees from the midpoint of the measurement interval to time $t+1$. Tree size at the midpoint is modeled from tree size at time $t+1$. This term also includes the subsequent growth on ingrowth trees that achieve the minimum diameter threshold after reversion.

M = mortality—the volume of trees that die from natural causes between time t and time $t+1$. The estimate is based on tree size at the midpoint of the measurement interval (includes mortality growth). Tree size at the midpoint is modeled from tree size at time t.

G_M = mortality growth—the growth of trees that died from natural causes between time t and the midpoint of the measurement interval. Tree size at the midpoint is modeled from tree size at time t. This term also includes the subsequent growth on ingrowth trees that achieve the minimum diameter threshold prior to mortality.

C = cut—the volume of trees cut between time t and time $t+1$. The estimate is based on tree size at the midpoint of the measurement interval (includes cut growth). Tree size at the midpoint is modeled from tree size at time t. Trees felled or killed in conjunction with a harvest or silvicultural operation (whether they are utilized or not) are included, but trees on land diverted from forest to nonforest (diversions) are excluded.

G_C = cut growth—the growth of cut trees between time t and the midpoint of the measurement interval. Tree size at the midpoint is modeled from tree size at time t. This term also includes the subsequent growth on ingrowth trees that achieve the minimum diameter threshold prior to being cut.

D = diversion—the volume of trees on land diverted from forest to nonforest (or, for some analyses, this may also include land diverted to **reserved forest land** and **other forest land**), whether utilized or not, between time t and time $t+1$. The estimate is based on tree size at the midpoint of the measurement interval (includes diversion growth). Tree size at the midpoint is modeled from tree size at time t.

G_D = diversion growth—the growth of diversion trees from time *t* to the midpoint of the measurement interval. Tree size at the midpoint is modeled from tree size at time *t*. This term also includes the subsequent growth on ingrowth trees that achieve the minimum diameter threshold prior to diversion.

FIA recognizes the following components of change as **prescribed optional variables**:

CI = cull increment—the net reduction in growing-stock volume due to reclassification of growing stock trees to cull trees between two surveys. Cull increment is the volume of trees that were growing stock at time *t*, but cull at time *t*+1. The estimate is based on tree size at the midpoint of the measurement interval (includes cull increment growth). Tree size at the midpoint can be modeled from tree size at time *t*, time *t*+1, or both.

G_{CI} = cull increment growth—the growth to the midpoint of the measurement interval between time *t* and *t*+1 of trees that were growing stock at time *t*, but cull trees at time *t*+1. Tree size at the midpoint can be modeled from tree size at time *t*, time *t*+1, or both.

CD = cull decrement—the net gain in growing-stock volume due to reclassification of cull trees to growing stock trees between two surveys. Cull decrement is the volume of trees that were cull at time *t*, but growing stock at time *t*+1. The estimate is based on tree size at the midpoint of the measurement interval. Tree size at the midpoint can be modeled from tree at time *t*, time *t*+1, or both.

G_{CD} = cull decrement growth—the growth from the midpoint of the measurement interval to time *t*+1 on trees that were cull at time *t*, but growing stock at time *t*+1. Tree size at the midpoint can be modeled from tree size at time *t*, time *t*+1, or both.

Except for *R*, *D*, G_D, and G_R, all components listed above are computed from plot areas where land use is defined as forest at both time *t* and time *t*+1. Note that it is not possible to measure the terminal d.b.h. of all trees that were cut, died, or diverted to a nonforest land use. To minimize potential bias associated with the growth of these trees, estimates of G_M, G_C, G_D, *D*, *R,* and G_R are modeled on the basis of the measurement-interval midpoint. The midpoint is calculated as $\Delta_t / 2$, where Δ_t is the time in years (rounded to the nearest 10th) between measurements for an individual plot. Models to predict midpoint tree sizes are developed regionally and may include a variety of factors, including terms to account for slowed growth on mortality trees.

The use of midpoint tree sizes creates special situations where careful attention is needed to account for all change components. Particularly notable are ingrowth trees that assume a status other than survivor. For instance, between time t and time t+1, a tree may cross the 5.0-inches d.b.h. threshold (ingrowth), grow to 5.4-inches d.b.h., and then die. Under those circumstances, there would be three components of change: (1) Ingrowth (I), the volume at 5.0-inches d.b.h.; (2) Mortality growth (G_M), the volume growth from 5.0-inches d.b.h. to 5.4-inches d.b.h.; and (3) Mortality (M), the volume at 5.4-inches d.b.h. Similar circumstances occur where ingrowth trees are associated with cutting, reversions, and diversions. Such situations implicitly require that midpoint tree sizes be modeled for all trees 1.0-inch d.b.h. and larger in order to check for trees that may have crossed the tree diameter threshold before removal or death.

For reporting growth and change, the individual components are usually combined as follows, and expressed either in terms of growing-stock or all-live volume:

Gross ingrowth = $I + R$

Accretion = $G_S + G_I + G_R + G_M + G_C + G_D$

Gross growth = gross ingrowth + accretion

Mortality = M

Removals = $C + D$

Net growth = gross growth – mortality

Net change = net growth – removals

The above terms for accretion and net growth are modified as follows for FIA regions that elect to produce additional output containing optional expressions of cull increment and decrement. Note that these optional terms are relevant only when components of change are expressed in terms of growing-stock volume:

Accretion = $G_S + G_I + G_R + G_M + G_C + G_D + G_{CI} + G_{CD}$

Net growth = gross growth – mortality + CD – CI

In addition to volume, all change components may also be expressed in terms of basal area or weight. More commonly, some (*I, R, M, C, D, CI,* and *CD*) are occasionally expressed as numbers of trees.

A variety of estimators have been proposed for the various components of change (Beers and Miller 1964, Gregoire 1993, Roesch and others 1989, Van Deusen and others 1986). When only one fixed-size plot is involved, estimation of change components is very straightforward and most estimators are equivalent. This is the case for most FIA reporting purposes because volumes generally are reported in terms of growing stock, which is based on trees 5.0-inches d.b.h. and larger that are only recorded on the subplot (except PNW which also uses macroplots).

For regions using the macroplots, or for expressions of growth involving microplot trees (e.g., growth of trees 1.0-inch d.b.h. and larger), change estimation is complicated by trees that grow from one plot size to the next. This requires techniques designed for variable-radius plots. Historically, FIA units have used one of two methods for calculating components of change for variable-radius plots—Beers and Miller (1964) or Van Deusen and others (1986). Because the Van Deusen estimator is more appropriate for prism sampling and FIA has moved away from prism sampling, we have decided to use the simpler Beers-Miller approach.

The Beers-Miller estimator weights all survivor growth (G_S) on the basis of plot size at time *t*:

$$G_S = s_2 - s_1 \tag{4.24}$$

where

s_2 = tree size at time *t*+1 weighted on the basis of plot size at time *t*

s_1 = tree size at time *t* weighted on the basis of plot size at time *t*

Ingrowth, *I*, is defined as those trees on the microplot that grew across the 1.0-inch threshold:

$$I = s_{dbh=1.0} \tag{4.25}$$

where

$s_{dbh=1.0}$ = the size of an ingrowth tree at the 1.0-inch d.b.h. threshold and growth on ingrowth as

$$G_I = s_2 - s_{dbh=1.0} \tag{4.26}$$

Both s_2 and the size of the tree at 1.0-inch d.b.h. ($s_{dbh=1.0}$) are weighted on the microplot basis for trees on the microplot that were < 1.0-inch d.b.h. at time *t* but greater than 1.0 inch at time *t*+1. Note that the Beers-Miller estimator ignores trees that grow onto the subplot from outside the microplot.

When estimating change, individual trees are placed into the appropriate change-component category(s). Tree attributes associated with the change-component of interest are then summarized to the plot level using equations

4.8 and 4.10, annualized by dividing plot-level periodic values by the number of years between the initial and terminal measurements of each plot, and summarized to the population level as specified in equations 4.11 and 4.13. The components of change are converted to average annual values as follows, where Δ_t is the time in years (rounded to the nearest 10th) between measurements for an individual plot:

Annual gross ingrowth = $(I + R) / \Delta_t$

Annual accretion = $(G_S + G_I + G_R + G_M + G_C + G_D) / \Delta_t$

Annual gross growth = annual gross ingrowth + annual accretion

Annual mortality = M / Δ_t

Annual removals = $(C + D) / \Delta_t$

Annual net growth = annual gross growth – annual mortality

Annual net change = annual net growth – annual removals

Annualized values for accretion and net growth for regions that include the optional expressions of cull increment and decrement would be:

Annual accretion = $(G_S + G_I + G_R + G_M + G_C + G_D + G_{CI} + G_{CD}) / \Delta_t$

Annual net growth = annual gross growth – annual mortality + $(\mathrm{CD} - \mathrm{CI}) / \Delta_t$

Observe that some of the change components pertain to trees on conditions that remained in forest for an entire inventory cycle; some are based on trees and areas that become forest between the initial and terminal inventory of a cycle (reversions); and some pertain to trees on conditions removed from the forest land base (diversions). Improved estimates of change can be obtained by stratification on the basis of both initial and terminal land use: (1) Forest to forest, (2) Forest to nonforest, (3) Nonforest to forest, and (4) Nonforest to nonforest. This might also be expanded to include the finer subsets of forest recognized by FIA such as timberland, reserved forest land, and other forest land.

This requires classification of the same Phase 1 points at both time *t* and time *t*+1—either photo interpretation of the same photo plots if photography was used or classification of the same pixels if satellite imagery is used. Note that the same collapsed strata should be used for both current and change estimates if the results are to be compatible.

Once the strata weights are assigned, estimation by domains within strata follows the same procedure outlined for current inventory parameters. For standard reporting purposes, domains identified or partitioned by forest attributes (e.g., owner group) usually are defined on the basis of initial classifications for forest-nonforest and forest-forest parameters; and domains for nonforest-forest inventory attributes are defined on the basis of terminal classifications. Examples of computational procedures for change estimation are provided in the supplementary document "Examples of FIA Change-Component Estimation Procedures for Several Common Cases" at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.4 Expansion Factors

When periodic inventories and flat files were the FIA standards, it was convenient to calculate a small set of expansion factors by which individual plot-level or tree-level observations could be converted to their population-level equivalents. This allowed population totals to be obtained via summation, which greatly simplified the estimation process. Expansion factors were popular with external FIA clients, many of whom used this concept to build their own processing systems. The tradeoff for such simplicity is that the use of expansion factors precludes the ability to calculate variances. At best, the variances of estimators derived from expanded values can only be approximated, and these approximations are known to be poor (Alegria and Scott 1991).

Expansion factors are less practical with panelized inventory systems, which are designed to increase analytical flexibility by allowing panels to be combined in a variety of ways. Each different panel combination produces a unique set of expansion factors, rendering expansion factors associated with panel systems less stable than those produced by periodic systems.

The use of expansion factors is discouraged because they prohibit accurate variance estimation and they no longer have the advantage of simplicity. However, there is still a demand for them, and it may take a while to convert processing systems to the estimation procedures specified in this chapter. Therefore, FIA will continue to offer expansion factors until a demand is no longer apparent. The derivation of expansion factors is described in the supplementary document "Computation of FIA Plot Expansion Factors" at http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

4.5 Phase 3

4.5.1 Estimation Procedures Used for FHM (Phase 3) Data

The FHM Program did not utilize any Phase 1 stratification prior to merging with FIA. In order to avoid reporting regional and State totals that conflicted with those reported by FIA, FHM statistics usually were presented as population-level means based on either simple random sampling or Generalized Leased Squares (GLS). Zarnoch and Bechtold (2000) developed one such approach based on simple random sampling and ratio-of-means estimators. Smith and Conkling (2004) used GLS estimation procedures, where population-level current mean values and annual change estimates are obtained from linear models for repeated measurements (Gregoire and others 1995, Urquhart and others 1993, Van Deusen 1996). The estimate of current value is:

$$y_t = b_{t-1} + b_1\left(Year_t - Year_{t-1}\right) \tag{4.29}$$

where

y_t = predicted value of the attribute at year *t*

b_{t-1} = the mean value of attribute at year *t*-1

b_1 = the annual change in *y* from Year *t*-1 to Year *t*

Both b_{t-1} and b_1 are computed by estimated generalized least squares using Proc MIXED® (SAS Institute 1999).

4.5.2 Combining with Phase 1

Because the FIA and FHM inventory systems have merged, it is now possible to combine Phase 1 and Phase 3 data using stratified random estimation and double sampling for stratification. We are studying use of Phase 1 stratification to enhance the estimation of attributes unique to Phase 3 and plan to document estimation procedures for attributes specific to Phase 3 in subsequent manuscripts.

4.5.3 Linking with Phase 2

For inventory parameters common to both Phase 2 and Phase 3, usually there is no advantage in generating population estimates from the smaller Phase 3 subset, because the estimates will not match and the Phase 3 variances will be larger due to a reduced *n*. Phase 3 plots, therefore, should be combined with Phase 2 plots when estimating inventory parameters common to both phases.

For those inventory parameters that are unique to Phase 3, it is possible to model relationships between unique Phase 3 parameters and parameters common to both Phase 2 and Phase 3. Such models can then be used to extrapolate estimates of Phase 3 parameters onto the more intensive Phase 2 sampling network.

4.6 Literature Cited

Alegria, J., Scott, C.T. 1991. Generalized variance function applications in forestry. Res. Note NE–345. Radnor, PA: U.S. Department of Agriculture Forest Service, Northeastern Forest Experiment Station. 4 p.

Beers, T.W.; Miller, C.I. 1964. Point sampling: research results, theory, and application. Res. Bull. 786. Lafayette, IN: Purdue University. Agricultural Experiment Station. 56 p.

Breidt, F.J.; Opsomer, J.D. 2004. Endogenous and post-stratification in surveys: classifying with a sample-fitted model. Preprint Series #02-23. Ames, IA: Department of Statistics, Iowa State University. 25 p.

Cochran, W.G. 1977. Sampling techniques. New York: John Wiley. 428 p.

Gregoire, T.G. 1993. Estimation of forest growth from successive surveys. Forest Ecology and Management. 56: 267-278.

Gregoire, T.G.; Schabenberger, O.; Barret, J.P. 1995. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Canadian Journal of Forest Research. 25: 137-156.

Moisen, G.G.; Edwards T.C., Jr. 1999. Use of generalized linear models and digital data in a forest inventory of northern Utah. Journal of Agricultural, Biological, and Environmental Statistics. 4: 372-390.

Roesch, F.A.; Green, E.J.; Scott, C.T. 1989. New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Science. 35: 281-293.

Särndal, C.E.; Swensson, B.; Wretman, J. 1992. Model Assisted Survey Sampling. New York: Springer-Verlag. 694 p.

SAS Institute. 1999. SAS/STAT® user's guide. Version 8. Cary NC: SAS Institute Inc. 2552 p. Vol. 2.

Smith, W.D.; Conkling, B.L. 2004. Analyzing forest health data. Gen. Tech. Rep. SRS-077. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 33 p.

Urquhart, N.S.; Overton, W.S.; Birkes, D.S. 1993. Comparing sampling designs for monitoring ecological status and trends: impact of temporal patterns. In: Statistics for the Environment. London: John Wiley: 71-85.

Van Deusen, P.C. 1996. Incorporating predictions into an annual forest inventory. Canadian Journal of Forest Research. 26(10): 1709-1713.

Van Deusen, P.C; Dell, T.R.; Thomas, C.E. 1986. Volume growth estimation from permanent horizontal points. Forest Science. 32: 415-422.

Zarnoch, S.J.; Bechtold, W.A. 2000. Estimating mapped-plot attributes with ratio of means. Canadian Journal of Forest Research. 30: 688-697.