United States
Department of
Agriculture

Forest Service

Southern
Research Station

Research Note
SRS–16

July 2008

Forest Inventory and Stratified Estimation: A Cautionary Note

John Coulston

Abstract

The Forest Inventory and Analysis (FIA) Program uses stratified estimation techniques to produce estimates of forest attributes. Stratification must be unbiased and stratification procedures should be examined to identify any potential bias. This note explains simple techniques for identifying potential bias, discriminating between sample bias and stratification bias, and determining the magnitude of the effect of stratification bias on forest area estimates. The key recommendation is that checks and balances should be incorporated into the FIA processing system to reduce the likelihood of bias caused by stratification.

Keywords: Bias, FIA, national land cover dataset, photointerpretation, proportional allocation.

Introduction

The Forest Service, U.S. Department of Agriculture national Forest Inventory and Analysis (FIA) Program provides estimates of forest and tree attributes at moderate to broad spatial scales (e.g., at scales of multiple county regions and States). In 2000, the FIA Program implemented a nationally consistent annualized survey design. The design is assumed to produce random equal probability samples (McRoberts and Hansen 1999). The FIA design incorporates three types of sampling—Phase 1 (P1), Phase 2 (P2), and Phase 3 (P3) (Bechtold and Patterson 2005). The goal of the P1 sample is to independently stratify the total area and assign each P2 and P3 plot to a stratum. The P2 sample refers to FIA's network of permanent forest mensuration field plots. The intensity of P2 is about one plot per 6,000 acres of total area (forest and nonforest). Additional forest health information is collected on one-sixteenth of the P2 plots, and this information is referred to as the P3 sample.

Because of funding constraints and natural variability among forest variables, the P1 sample is used along with stratified estimation techniques to improve the precision of estimates (Bechtold and Patterson 2005). Ancillary information used for stratification includes the National Land Cover Dataset (NLCD), Gap Analysis Program (GAP) classification, digital aerial photography, and ecoregion sections (McRoberts and others 2006). In fact, any independent map-based product can be used for stratification (e.g., soil type maps and elevation class maps). When ancillary map-based data is used to successfully group similar forest inventory plots together in the same stratum, estimates with smaller variance are produced. However, the ancillary information must be independent of the FIA field plot data; and the stratification must be performed without introducing bias. The purpose of this note is to describe how to identify potential bias and examine its potential magnitude.

Background

The FIA survey grid is a 27x intensification of the Environmental Monitoring and Assessment Program (EMAP) survey grid (White and others 1992). The EMAP survey grid and the FIA survey grids (as well as intensifications of these grids) are triangular, isotropic, and systematically cover the conterminous United States. The FIA Program has moved to nationally standardized estimation techniques (stratified estimation) and a standardized National Information Management System (NIMS). Stratified estimation is a statistical technique that can reduce the variance of estimates without increasing sample size (Cochran 1977). To implement stratified estimation, each P2 plot must be assigned to a stratum and

Supervisory Research Forester, U.S. Forest Service, Southern Research Station, Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37919.

the proportion of the total area (e.g., FIA estimation unit) in each stratum must be identified. When ancillary information is used to stratify FIA P2 plots, the probability that a P2 plot is assigned to a stratum is equal to the proportion of that stratum in the total area of interest (i.e., the probability is proportional to area). This results in proportional allocation of P2 plots to each stratum.

Ancillary data derived from remote sensing are generally used for P1. For example, the total area of an FIA estimation unit may be proportioned into forest and nonforest based on manual interpretation of digital aerial photography. Each P2 plot would also be classified as forest or nonforest based on the same photography. This information is then used to estimate the weight of each stratum (forest and nonforest) and identify which P2 plots are included in each stratum. Misclassifications are inherent in the classification of any remote sensing medium. In this case, misclassifications are not viewed as errors but as resulting in less homogenous strata. The result of unbiased (i.e., random) misclassifications is decreased correlation between the ancillary data used for stratification and the attribute being estimated, leading to an increase in the variance of the estimate.

Of the several P1 media that can be used, the NLCD and digital aerial photography are the most common. When digital aerial photography is employed, the P1 photopoints are based on a 729X intensification of the EMAP sampling grid or a 5 by 5 cluster of photopoints surrounding each P2 plot. When either photointerpretation grid is employed, each photopoint represents about 220 acres. When using double sampling for stratification, the digital aerial photography is sampled by manual interpretation at each photopoint and P2 ground plot. Because of the potential for interpreter bias, the same person, using the same computer and the same imagery, should classify each P1 photopoint and P2 ground plot. When the NLCD is used for P1, each pixel represents about 0.222 acres and a spatial overlay is used to assign P2 ground plots to strata.

Bechtold and Patterson (2005) document the NIMS estimation procedures. When stratified estimation is employed, the estimate of total forest area ($\hat{A}_f$) is:

$$\hat{A}_f = A_t \sum_{h=1}^{H} W_h \overline{P}_h \qquad (1)$$

where

A_t = the total area of interest (acres)
W_h = the weight of stratum h (the proportion of A_t occupied by stratum h)
$\overline{P}_h$ = the mean plot-level proportion forest in stratum h

The variance of the estimate of total forest area [$v(\hat{A}_f)$] based on double sampling for stratification used with photo interpretation (where W_h is sampled as opposed to known), is estimated by:

$$v(\hat{A}_f) = A_t^2 \left[\sum_{h=1}^{H} \left(\frac{n'_h - 1}{n' - 1} \right) \frac{n'_h}{n'} v(\overline{P}_h) + \frac{1}{n' - 1} \sum_{h=1}^{H} \frac{n'_h}{n'} \left(\overline{P}_h - \overline{P} \right)^2 \right] \qquad (2)$$

where

n' = the total number of P1 photopoints
n'_h = the total number of P1 photopoints in stratum h

$$\overline{P} = \sum_{h=1}^{H} W_h \overline{P}_h$$

$v(\overline{P}_h)$ = within stratum variance of the proportion of forest (equation 3):

$$v(\overline{P}_h) = \frac{\sum_{i=1}^{n_h} P_{hi}^2 - n_h \overline{p}_h^2}{n_h (n_h - 1)} \qquad (3)$$

where

P_{hi} = the proportion of forest on the i^{th} plot within stratum h
n_h = the number of P2 plots in stratum h

When a P1 medium such as the NLCD is used, W_h is assumed to be known (as opposed to sampled); and the variance estimator is:

$$v(\hat{A}_f) = \frac{A_t^2}{n} \left[\sum_{h=1}^{H} W_h n_h v\left(\overline{P}_h\right) + \sum_{h=1}^{H} (1 - W_h) \frac{n_h}{n} v\left(\overline{P}_h\right) \right] \qquad (4)$$

where

n = the total number of P2 plots

The FIA Program has traditionally used the percent sampling error (equation 5) to reflect the accuracy of estimates:

$$S.E.\% = 100 \frac{\sqrt{v(\hat{A}_f)}}{\hat{A}_f} \qquad (5)$$

where

S.E.% = the percent sampling error

Ninety-five percent confidence intervals are then constructed by $\hat{A}_f \pm t(\hat{A}_f \text{S.E.\%})$ where t is the two-sided student t-value with $\left(\sum_{h=1}^{H} n_h - 1\right)$ degrees of freedom.

Methods

Our goals are to (1) identify potential bias, (2) discriminate between sample bias (preferential location of P2 plots in a particular land use) and stratification bias (systematic error in P1), and (3) examine the magnitude of the effects that stratification bias may have on forest area estimates. To determine whether bias (sample bias or stratification bias) exists, the P1 stratification information is used. Because of the systematic grid used by FIA, the probability that a P1 photopoint (or pixel) is assigned to stratum h is A_h/A_t where A_h is the area of stratum h in A_t. Therefore, the proportion of P1 photopoints (or pixels) assigned to stratum h (p_{p1h}) is also approximately A_h/A_t. The probability that a P2 field plot is assigned to stratum h based on its photopoint classification is also A_h/A_t, and the proportion of P2 plots assigned to stratum $h(p_{p2h})$ is also approximately A_h/A_t. Therefore, $p_{p1h} \approx p_{p2h}$ and the assumption of proportional allocation holds. When the difference between p_{p1h} and p_{p2h} is consistently different in direction, either the stratification is biased or the plot locations are biased (e.g., plots were preferentially located in forest). Bias can be quantified by:

$$b_h = p_{p1h} - p_{p2h} \tag{6}$$

where

b_h = bias in stratum h
p_{p1h} = the proportion of stratum h based on P1
p_{p2h} = the proportion of stratum h based on P2

When $b_h = 0$, the stratified estimate of A_f (equation 1) will equal the simple random sample estimate of A_f (equation 1, where $H = 1$). However, if $b_h > 0$, a disproportionately small number of P2 plots carry a disproportionately large weight (W_h). Conversely, when $b_h < 0$, a disproportionately large number of P2 plots carry a disproportionately small weight (W_h). When b_h is statistically significant from zero and therefore $p_{p1h} \neq p_{p2h}$, the proportional allocation assumption has been violated.

In practice, p_{p1h} will generally not be exactly equal to p_{p2h} because P1 and P2 have different sampling intensities and different standard errors. However, when $H = 2$, we can statistically test the hypothesis $p_{p1h} = p_{p2h}$ using equation 7 (Steel and others 1997).

$$\chi^2 = \frac{(n_{11}n_{22} - n_{12}n_{21})^2 v}{n_{.1}n_{.2}n_{1.}n_{2.}} \tag{7}$$

where

χ^2 = chi-square with 1 degree of freedom
n_{11} = number of P1 points in stratum 1
n_{22} = number of P2 points in stratum 2
n_{12} = number of P1 points in stratum 2
n_{21} = number of P2 points in stratum 1
v = number of P1 and P2 points
$n_{.1} = n_{11} + n_{21}$
$n_{.2} = n_{12} + n_{22}$
$n_{1.} = n_{11} + n_{12}$
$n_{2.} = n_{21} + n_{22}$

When $H > 2$ the hypothesis $p_{p1h} = p_{p2h}$ can be tested using the normal approximation of equation 7:

$$Z_h = \frac{p_{p1h} - p_{p2h}}{\sqrt{\frac{p_{p1h}(1-p_{p1h})}{n_{p1}} + \frac{p_{p2h}(1-p_{p2h})}{n_{p2}}}} \tag{8}$$

where

Z_h = Z score for stratum h
n_{p1} = the number of P1 points classified
n_{p2} = the number of P2 points classified

If we fail to reject the null hypothesis that $p_{p1h} = p_{p2h}$, then we can assume that proportional allocation holds. If we reject the null hypothesis that $p_{p1h} = p_{p2h}$, then we have identified bias and should move to the second step, discriminating between sample bias and stratification bias.

Several options are available to identify whether the bias was caused by sample bias or stratification bias. The most cost-effective option is to examine the distribution of plots based on an alternative land cover or land use map such as GAP or NLCD. To accomplish this, the map should be collapsed to the same thematic resolution as the P1 stratification being examined (e.g., the map might be reclassified to forest and nonforest land cover). Each P2 plot is then classified based on the alternative map, and b_h from equation 6 is then evaluated and the hypothesis $p_{p1h} = p_{p2h}$ is tested using equation 7. If we reject the null hypothesis, $p_{p1h} = p_{p2h}$, then there is an indication that the sample may be biased. However, if we fail to reject the null hypothesis, $p_{p1h} = p_{p2h}$, then the locations of the P2 plots were not preferentially placed in a certain land use; and the P2 sample may be considered unbiased at the spatial and thematic scale of the alternative land use map. Rejecting the null hypothesis at this step implies that the stratification was biased.

When the P1 stratification is biased, then analysts will likely want to know the impact of the biased stratification on the estimate of forest area. The simplest option is to forego stratification and set $H = 1$, which produces estimates of the total and its variance based on a simple random sample. This removes the influence of the biased stratification on the estimate of the total but in most cases substantially increases the variance estimate, thus decreasing the certainty associated with the estimate of the total. The second option is to use the alternative land use map from the previous step. This also removes the influence of the biased stratification on the estimate of the total but in most cases will not increase the variance estimate as dramatically as setting $H = 1$. The potential magnitude of the bias can then be quantified by subtracting the biased forest area estimate from both the upper and lower confidence bounds of an unbiased estimate (either the simple random sample estimate or the stratified estimate based on the NLCD in this case).

Case Study

The forest area estimates for Tennessee (2000–2004) were calculated using equation 1 and photointerpretation to estimate stratum weights and plot stratum assignments. FIA has traditionally divided the State into five geographic regions for reporting purposes (fig. 1). These survey units serve as estimation units for the State. Based on previous State reports (Birdsey 1983, May 1991, Murphy 1972, Schweitzer 2000), the forest area estimates for 1999 and 2005 deviated from historical estimates (fig. 2). Forest area estimates for 1999 and 2005 were based on different statistical techniques; however, both inventories were based on the current FIA plot design (Bechtold and Patterson 2005). The 1999 forest area estimate was based on double sampling for area (Reams 2000), and the 2005 estimate was based on stratified estimation using forest and nonforest strata. However, the same P1 photo interpretation was used for both estimates. Focusing on the 2005 estimate, bias was identified and statistically tested using equations 6 and 7 respectively (table 1). b_{forest} ranged from 1.6 percent in estimation unit 2 to 4.6 percent in estimation unit 5. The hypothesis $p_{p1h} = p_{p2h}$ was rejected for estimation units 1 and 5 and the entire State at or below $a = 0.012$ (table 1).

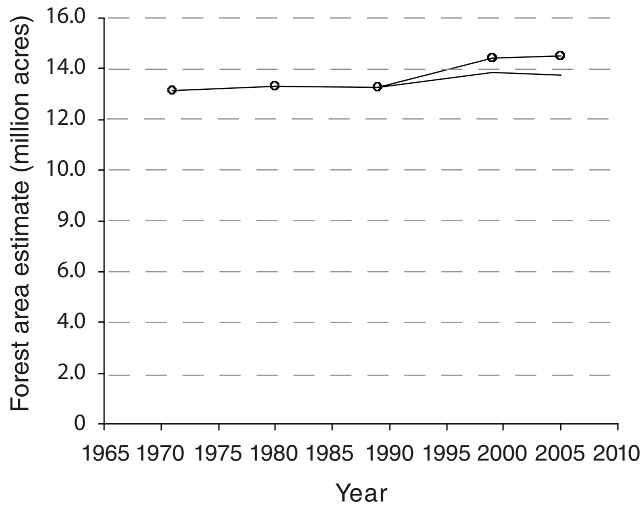

Figure 2—Forest area estimates for Tennessee. The circles represent the original estimates. Note the increase in 1999. The solid line denotes the trajectory of forest area estimates when the 1999 and 2005 estimates are corrected for stratification bias.

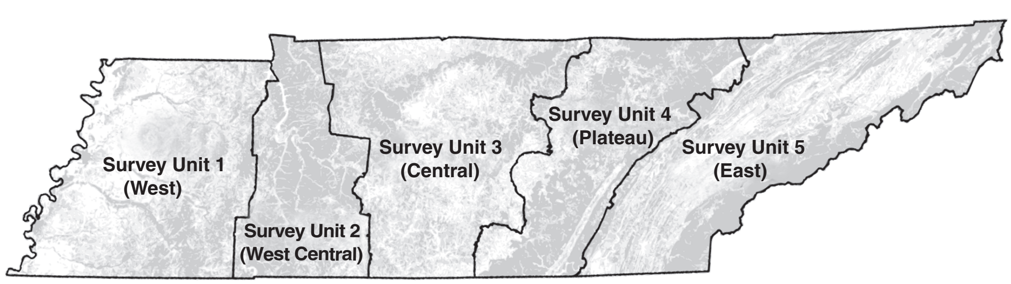

Figure 1—FIA survey units for Tennessee, with forest (gray) cover based on the National Land Cover Dataset.

Table 1—Bias in the forest stratum based on equation 6 and associated statistical test based on equation 7 for each survey unit in Tennessee

Estimation unit	b_{forest}	χ^2	p-value
Photo Interpreted Phase 1			
1	0.038	6.37	0.012
2	0.016	0.69	0.407
3	0.025	2.59	0.107
4	0.018	1.10	0.294
5	0.046	9.66	0.002
State	0.032	18.70	0.000
NLCD			
1	0.021	1.98	0.160
2	0.001	0.00	0.966
3	-0.027	3.14	0.077
4	0.022	1.66	0.197
5	-0.018	1.45	0.229
State	-0.002	0.04	0.833

The first step was to determine if the P2 plot locations were preferentially placed in a certain land use. We used the 2001 NLCD to examine this issue. The NLCD has 21 land use/land cover classes. The thematic resolution of the NLCD map was reduced to forest and nonforest (fig. 1), and each P2 plot was assigned a forest or nonforest classification based on spatial overlay. The difference between p_{p1h} and p_{p2h} was then examined and the hypothesis $p_{p1h} = p_{p2h}$ was tested. b_{forest} ranged from -2.7 percent in estimation unit 3 to 2.2 percent in estimation unit 4 (table 1). The null hypothesis, $p_{p1h} = p_{p2h}$, was provisionally accepted at a = 0.05 for each of the estimation units and for the entire State. The conclusion at this step was that the P2 plot locations did not significantly oversample one stratum (forest or nonforest) at the spatial resolution of the NLCD and the thematic resolution of a forest-nonforest classification. Therefore, the original P1 stratification was assumed to be biased.

To examine the potential impact of the biased P1 stratification on the estimates of forest area, we first set H = 1 to compute estimates of forest area based on simple random sampling (fig. 3). The difference between the total forest area estimates at the State level was 683,113 acres. Differences between forest area estimates were greatest for estimation units 1, 3, and 5, and these units are identified as having the largest amount of stratification bias (table 1). However, on average the variance estimates tripled. Because the NLCD stratification was unbiased, forest area estimates were made using the NLCD for stratification. Based on the 95 percent confidence intervals of the forest area estimates when using the NLCD for stratification, the potential bias in unit 5 was between approximately 272,000 and 384,000 acres (fig. 3). The potential bias in units 1 and 3 was also high.

Discussion and Conclusions

The goal of this research note was to describe methods to identify bias in P2 plot locations, identify bias in P1 stratification, and examine the potential magnitude of stratification bias. Bias identified in the P2 plot locations is a serious issue that is not addressed here. Stratification bias, however, is easily identifiable and can be corrected either by treating the data as a simple random sample or by using an alternative map for stratification (e.g., NLCD).

Stratification is an important statistical tool for reducing variance, but the stratification must be implemented without introducing bias. Here we describe stratification bias, which differs from classification error. Classification error occurs with all classified remotely sensed data. Unbiased classification error will lead to increased variance but unbiased estimates. However, when there is systematic classification error, then classification bias can occur.

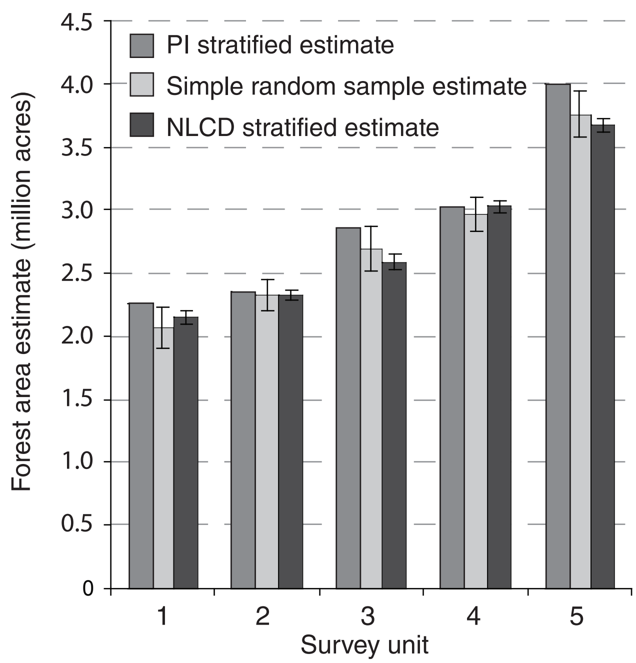

Figure 3—Forest area estimates for Tennessee (2005) by survey unit. The error bars denote 95 percent confidence intervals.

Classification bias can lead to stratification bias, which causes biased estimates. This point was demonstrated in the case study where stratification based on photointerpretation was biased. This research note also demonstrates the influence of stratification on variance. The influence of the stratification can be quantified by dividing the variance of an estimate from the unbiased stratified estimate by the variance of an estimate based on simple random sampling. This computation yields a value called the design effect (Kish 1995). In the case study presented in this research note, stratifying using the NLCD had a design effect of 0.43 for forest area estimates at the State level. This means that the stratified estimate had 43 percent of the variance that the simple random sample had.

The FIA Program strives to provide reliable estimates of forest attributes. Forest area estimates receive particular scrutiny from State cooperators. Bias introduced during the stratification process influences the reliability of estimates and causes variance estimates to be unreliable so care should be taken when developing the stratification. The bias identified here in the case study was introduced when digital aerial photography was interpreted manually; but similar problems can occur when other stratification media, such as the NLCD, are employed. As a practical matter, it is unlikely that $b_h = 0$ across all levels of aggregation used by FIA because the spatial scale of the P2 plot sample differs from the spatial scale of the stratification media. However, the allowable amount that b_h deviates from zero should be tested using the suggested chi-square statistical test. Furthermore, testing of the stratification should be done as early as possible in the processing stage.

Literature Cited

Bechtold, W.A.; Patterson, P.L., eds. 2005. The enhanced Forest Inventory and Analysis Program—national sampling design and estimation procedures. Gen. Tech. Rep. SRS–80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 85 p.

Birdsey, R.A. 1983. Tennessee forest resources. Resour. Bull. SO–90. New Orleans: U.S. Department of Agriculture Forest Service, Southern Forest Experiment Station. 35 p.

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

Kish, L. 1995. Survey sampling. New York: John Wiley. 643 p.

May, D.M. 1991. Forest resources of Tennessee. Resour. Bull. SO–160. New Orleans: U.S. Department of Agriculture Forest Service, Southern Forest Experiment Station. 65 p.

McRoberts, R.E.; Hansen, M.H. 1999. Annual forest inventories for the North Central region of the United States. Journal of Agricultural, Biological, and Environmental Statistics. 4(4): 361-371.

McRoberts, R.E.; Holden, G.R.; Nelson, M.D. [and others]. 2006. Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis Program of the USDA Forest Service. Canadian Journal of Forest Research. 36: 2968-2980.

Murphy, P.A. 1972. Forest resources of Tennessee. Resour. Bull. SO–35. New Orleans: U.S. Department of Agriculture Forest Service, Southern Forest Experiment Station. 33 p.

Reams, G.A. 2000. SAFIS area estimation techniques. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C., eds. Proceedings of the first annual Forest Inventory and Analysis symposium. Gen. Tech. Rep. NC–213. St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Research Station: 32-36.

Schweitzer, C.J. 2000. Forest statistics for Tennessee 1999. Resour. Bull. SRS–52. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 78 p.

Steel, R.G.D.; Torrie, J.H.; Dickey, D.A. 1997. Principles and procedures of statistics: A biometrical approach. 3rd ed. New York: McGraw-Hill. 666 p.

White, D.; Kimerling, A.J.; Overton, W.S. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. Cartographic and Geographic Information Systems. 19: 5-22.