Sensitivity of FIA Volume Estimates to Changes in Stratum Weights and Number of Strata

James A. Westfall and Michael Hoppus[1]

Abstract.—In the Northeast region, the USDA Forest Service Forest Inventory and Analysis (FIA) program utilizes stratified sampling techniques to improve the precision of population estimates. Recently, interpretation of aerial photographs was replaced with classified remotely sensed imagery to determine stratum weights and plot stratum assignments. However, stratum weights based on remotely sensed data depend on many factors, such as classification algorithm and image selection. County volume estimates and associated variances were calculated over a range of stratum weight scenarios and for various number of strata. Rates of change in estimated values and variances, and their effects on percent sampling error, were examined in relation to different strata configurations.

Historically, the USDA Forest Service Forest Inventory and Analysis (FIA) program has used stratified estimation techniques to increase the precision of estimates of forest population parameters. Stratifying the population has traditionally been done by interpreting aerial photography. However, advances in technology now favor the use of satellite imagery for many of the regional FIA units. In the Northeast (NE) region, stratification is performed with data from the Landsat 7 sensor platform. Each pixel in an image is classified and the resulting classes are collapsed into groups that represent different strata. Information on sample-plot location is used to assign plots to strata, and stratum weights are determined by stratum size in relation to total area. The computed stratum weights depend on several factors. Images of the same land area taken at different times produce different stratification results. Also, many algorithms can be used to classify an image. Each algorithm produces different classification results, which affect the stratification outcome. As stratification results are utilized in the estimation process, it is of interest to investigate the degree to which different results might affect estimated values. In our study, we examined the extent to which changes in stratum weights and number of strata affect estimates of merchantable (4-inch top limit) cubic-foot volume and associated sampling errors.

Data

Data for this study are from a new FIA annual inventory system (USDA Forest Service 2002) in Maine, where data exist for about 60 percent of the sample plots. Each plot consists of four circular subplots, each with a radius of 24 feet (7.3 m). Sample plots were mapped in detail by land condition, allowing for an estimate of the area for each condition. Different conditions were delineated among forest, nonforest, water, and other variables, and boundaries were established within forested conditions for other types of changes, such as forest type, stand size, and tree density. Information collected included tree location, diameter at breast height, total and merchantable heights, percent of cull, and type and location of tree damage. Gross cubic-foot volumes to a 4-inch (10.2 cm) top limit for individual trees were computed with equations developed by Scott (1981). The volume of cull is subtracted from the predicted gross volume to obtain net volume. Net plot volumes were obtained by summing volumes of individual trees.

Methods

To estimate total cubic-foot volume for a county, weights for each stratum must be determined. This was done by reclassifying a Landsat TM-based forest/nonforest cover map (Vogelmann *et al.* 1998) using a 5 x 5 moving-window summarization. This placed each pixel into one of 26 classes (0 to 25 forested pixels in the surrounding 5 x 5 pixel box). These class-

[1] Research Forester and Supervisory Forester, U.S. Department of Agriculture, Forest Service, Northeastern Research Station, Newtown Square, PA 19073. Phone: 610–557–4043; fax: 610–557–4250; e-mail: jameswestfall@fs.fed.us.

es were then collapsed into strata and stratum weights by county were computed from map pixel counts and census-derived county boundary information (U.S. Census Bureau 2001).

Using the stratum weights derived from the satellite-image analysis, we calculated net cubic-foot volume estimates and variances by county. The estimation procedures were based on methodology described by Cochran (1977) for stratified random sampling. These estimators utilize the stratum weights to reflect the relative influence of each stratum. Having obtained the estimate and associated variance of the estimate, one can compute the percentage sampling error for the estimate of total volume:

$$SE\% = \frac{\sqrt{v(\tilde{Y})}}{\hat{Y}}(100) \qquad (1)$$

where:

$v(\hat{Y})$ = variance of the estimate of total cubic-foot volume

$\hat{Y}$ = estimate of total net cubic-foot volume

To assess how different stratum weights may affect county estimates, estimates and sampling errors for total cubic-foot volume were computed over a range of stratum weights. For each stratum, weights were systematically altered to ±0.25 of the original value in increments of 0.01. Weights of the remaining strata were increased or reduced proportional to their size to maintain the requirement that the sum of all stratum weights equals 1. For example, if the original weight of stratum 1 was 0.30, then the range of stratum 1 weighting scenarios covered 0.05 through 0.55 in increments of 0.01. The ±0.25 range was not fully realized for strata with original weight values less than 0.25 or greater than 0.75, as this would produce weights less than 0 or greater than 1.

Different numbers of strata were created by altering the four original strata. When the number of strata was to be reduced, strata were combined. For this analysis, strata 3 and 4 had similar attributes and were combined to reduce the number of strata to three. To increase the number of strata, an existing stratum was divided into two parts. Because stratum 4 was nearly twice the size of the other strata, this stratum was divided into two equal parts. Half of the plots from the original stratum 4 having the smallest plot volumes were

retained as stratum 4, and stratum 5 was created with the remaining plots. Interestingly, this division created five strata with nearly equal weights.

Analysis

The analysis was conducted for Androscoggin County in Maine, a small county (~318,000 ac) that is about 70 percent forested. The cubic-foot volumes and stratum weights used in this analysis are given in table 1. For the original configuration of four strata, the effects of changing weights of various strata on the estimate of volume are depicted in figure 1. The estimates are most sensitive to weight changes for stratum 1, where a 0.01 increase in weight decreased the estimated value roughly 1 percent. Stratum 2 is largely insensitive to changes in stratum weight. Strata 3 and 4 behave similarly, where increases in the estimated volume are nearly 0.5 percent for every 0.01 unit of change in weight. Similar trends are noted for changes in the variance of the estimate (fig. 2), although the magnitude of change is less. The exception is stratum 2, where, unlike the negligible change in the estimate, a notable trend in variance is

Table 1.—*Summary of stratum attributes for Androscoggin County, Maine*

3 Strata	Mean plot volume (ft³)	Variance of mean plot volume	Stratum weight
1	24.1	442.4	0.226
2	149.2	1,060.3	.180
3	217.2	1,096.0	.594
4 Strata			
1	24.1	442.4	.226
2	149.2	1,060.3	.180
3	233.0	2,927.1	.228
4	204.5	1,864.9	.366
5 Strata			
1	24.1	442.4	.226
2	149.2	1,060.3	.180
3	233.0	2,927.1	.228
4	87.0	719.5	.183
5	321.9	776.3	.183

Figure 1.—*Percent change in estimate of total volume when stratum weights are changed for four strata in Androscoggin County, Maine.*

Figure 3.—*Change in percent sampling error for total volume when stratum weights are changed for four strata in Androscoggin County, Maine.*
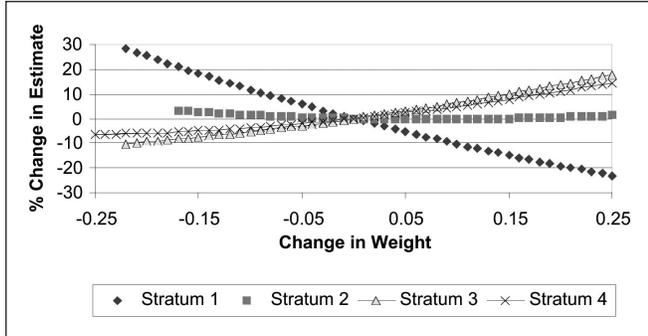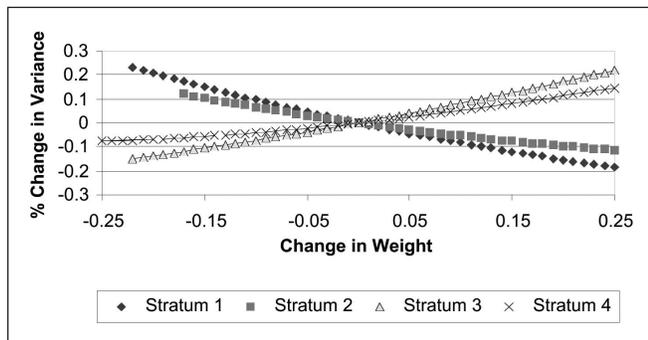

Figure 2.—*Percent change in variance of estimate of total volume when stratum weights are changed for four strata in Androscoggin County, Maine.*
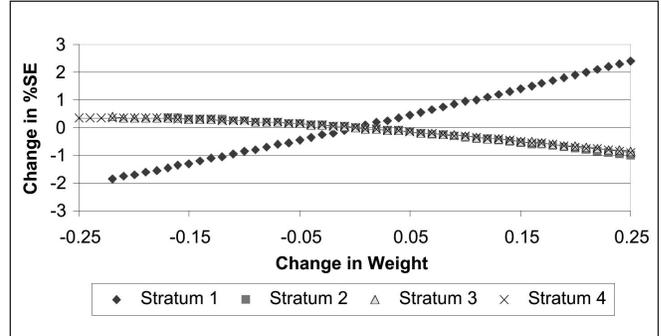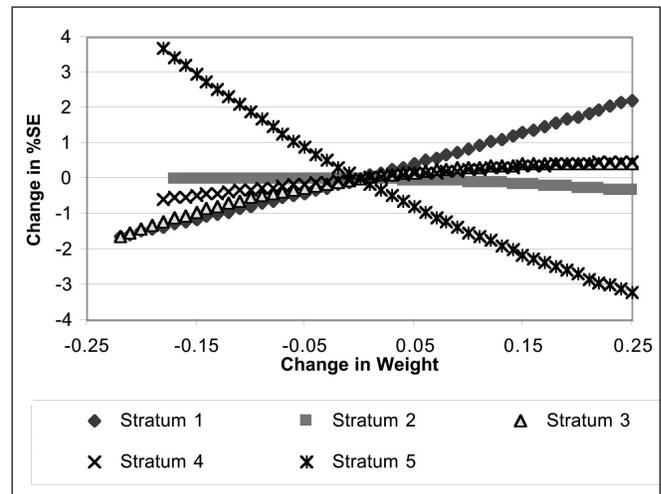

Figure 4.—*Change in percent sampling error for total volume when stratum weights are changed for five strata in Androscoggin County, Maine.*

associated with the weight changes. These changes in volume estimates and variances affect the percent sampling error (%SE). For example, figure 3 shows how the reduction in variance associated with decreasing weight for stratum 1 is exceeded by the accompanying decrease in the estimate with a net result of a smaller %SE. Strata 2 through 4 show surprisingly similar change in %SE trends. For these strata, a decrease in weight increases the %SE, as the increase in variance exceeds that of the estimate. When weight is added to these strata, the reduction in variance occurs more rapidly than the change in the estimate and %SE decreases.

Determining the number of strata also can be affected by classification algorithm, image selection, and other factors. For FIA purposes, Androscoggin County was divided into four strata (see table 1). For this study, the number of strata was altered to determine how different numbers of strata might affect estimates of volume. Where four strata were reduced to three,

changes in estimates and sampling errors were similar to the four-strata configuration. However, the sampling error was reduced from 13.8 percent to 13.2 percent (when the original set of stratum weights was used). This was due to a slight increase (0.6 percent) in the estimate and a decrease (7.6 percent) in the variance. The reduction in variance was unexpected, as reducing the number of strata generally increases variance (Cochran 1977). The result presented here occurred for several reasons. First, the two strata that were combined had similar means, which leads to the combined data providing additional observations that clustered about a similar value. Additionally, these strata have relatively few observations due to the small county size. The combination of strata increases

sample size, which has a notable effect on the variance. This phenomenon would likely not be observed when computing estimates for larger areas or where sampling intensity is higher.

When the number of strata was increased to five, the changes in estimates of volume for a given change in stratum weight were similar in magnitude to those for the three- and four-strata analyses. However, the relationship between the rates of change differed. This is apparent in the trends for change in %SE (fig. 4), where decreases are now evident when weight is decreased for strata 3 and 4. Altering of weights in the newly created stratum 5 has the greatest effect on %SE. For this stratum, decreases in weight result in increases in %SE. The %SE based on original weights for the five-strata analysis was much lower (11.0 percent) than the %SE for the original four strata (13.8 percent). The improved sampling error resulted from the decrease in variance of 37.1 percent, which resulted from creating the addition stratum.

In each of the analyses presented, differences between the original estimate and those where stratum weights were altered were not significantly different at the 95-percent confidence level for stratum weight changes less than 0.20. Where the change in stratum weight exceeded 0.20, differences between estimates were significant in some cases.

Discussion

We can determine mathematically the magnitude and direction of change in volume estimates given a particular re-weighting of strata if the mean plot volumes of the strata are known. For example, when changes in stratum weights are between only two strata, the change in the estimate is proportional to the difference between the mean plot volumes of these strata. However, when weights are altered in more than two strata, the distribution of the change in weights also affects the outcome. Rates of change are determined by how far the stratum mean deviates from the overall sample mean. The further the stratum mean is from the sample mean, the greater the rate of change. The direction of change depends on whether the stratum mean is smaller or larger than the sample mean. The estimate of volume decreases when weight is added to strata whose mean is smaller than the overall mean. Similarly, the

estimate increases as weight increases for strata with means greater than the overall mean.

It is not as clear how changing stratum weights alters variance estimates as strata usually have different mean volume estimates, while the number of plots and variances across strata may or may not be similar. The amount and direction of change in variance depends on the magnitude of differences in the variances of the individual strata. More importantly, changes in both the estimate and the variance can have a notable effect on the %SE. In instances where the mean plot volume for the strata is close to the overall mean, the change in variance drives the change in %SE. For strata with means that differ greatly from the overall mean, the change in the volume estimate is more influential.

As was illustrated for the five-strata analysis, a notable reduction in %SE is attainable. Although this reduction was essentially manufactured by the method by which the stratum was separated, it does show that certain relationships among the strata can help reduce sampling error. This does not mean that having more strata is always better as %SE also was reduced when changing from four strata to three. The key is to minimize the variances of the individual strata to the extent possible. Thus, if the merging of two strata results in a variance that is less than the sum of the original two strata, %SE will decrease. Likewise, breaking a single stratum into smaller strata can be beneficial if the sums of the variances for the smaller strata are less than the variance of the single stratum. The ability to reduce these variances will depend on both the spread of the data and number of observations. In this exercise, we used the plot data to increase the number of strata and minimize the variance within each stratum. In practice, this would not be possible as the 5 x 5 moving window summary does not give actual values for sample-plot data. Research is needed to create methodologies that provide classifications that are highly correlated with sample-plot data.

Conclusion

Stratification is an effective tool for improving the precision of FIA volume estimates. In many instances, stratified estimation procedures produce sampling errors that meet or exceed guidelines

for estimates of area and volume. However, in some situations it is difficult to obtain sufficiently low sampling errors (e.g., estimation for small areas). This may preclude drawing meaningful conclusions. Our research has shown that it is possible to improve sampling errors by refining stratification techniques. A stratification method that optimizes the number of strata and associated variances might be an efficient and effective way to obtain meaningful estimates where the number of measurement plots is small. This approach would likely be far less expensive and time consuming than increasing sampling intensity.

Literature Cited

Cochran, W.G. 1977. Sampling techniques. 3d ed. New York, NY: John Wiley and Sons. 428 p.

Scott, C.T. 1981. Northeastern forest survey revised cubic-foot volume equations. Res. Note NE-304. Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 3 p.

U.S. Census Bureau. 2001. TIGER/Line Files, Census 2000. Washington, DC: U.S. Department of Commerce, Census Bureau. Available: http://www.census.gov/geo/www/tlmetadata/tl2kmeta.txt [12 October 2001].

U.S. Department of Agriculture. 2002. Forest inventory and analysis national core field guide, volume 1: field data collection procedures for phase 2 plots, version 1.6. Washington, DC: U.S. Department of Agriculture, Forest Service. Internal report.

Vogelmann, J.E.; Sohl, T.L.; Campbell, P.V.; Shaw, D.M. 1998. Regional land cover characterization using Landsat thematic mapper data and ancillary data sources. Environmental Monitoring and Assessment. 51: 415–428.