An Alternative View of Some FIA Sample Design and Analysis Issues

Paul C. Van Deusen[1]

Abstract.—Sample design and analysis decisions are the result of compromises and inputs from many sources. The end result would likely change if different individuals or groups were involved in the planning process. Discussed here are some alternatives to the procedures that are currently being used for the annual inventory. The purpose is to indicate that alternatives exist and that reasonable people might prefer approaches that differ from the ones selected for the annual inventory. The topics covered include panel creep, mapping, the moving average, and data security issues.

The Forest Inventory and Analysis (FIA) program of the USDA Forest Service is implementing an annual forest inventory (USDA Forest Service 1999) where a percentage of the plots are measured each year. Before the 1998 Farm Bill that mandated the annual inventory, FIA was conducting periodic surveys in each State. Each survey took 1–4 years, and then the State was revisited every 10–18 years. The 1998 Farm Bill followed two Blue Ribbon Panel reports, BRP I and BRP II (American Forest Council 1992, American Forest and Paper Association 1998). BRP I called for shortening the period between surveys from 10 to 5 years. This goal was never achieved, and cycles averaged 10 years or more when BRP II convened in 1997. The BRP II request for more timely data motivated the 1998 Farm Bill legislative mandate for an annual survey. The Farm Bill mandate implies that it is important to have annual updates and continuous data collection underway in all States. It remains to be seen whether this goal will be achieved.

Panel Creep

The interpenetrating panel design was selected because it seemed to best facilitate the intent of the Farm Bill, which was to have a regular proportion of the plots measured each year. The Farm Bill also intended that annual updates should be possible. Presumably, the purpose of annual data collection is to provide annual estimates. Hence, the original plan was to measure one panel each year. A panel consists of plots that systematically cover the State. This means that unbiased estimates of any quantity are possible from a single panel or from combinations of panels. There are also secondary benefits of the panel design, such as maintaining equal annual budgets and work loads.

The initial plan was to create 5 panels in Eastern States and 10 panels in the West. Unfortunately, resources are not available to measure one panel per year in all States. This leads to panel creep. Panel creep occurs when a panel has been only partially measured at the end of the field season. One could differentiate between planned and unplanned panel creep, but the end result is the same. A panel can't be used for a statewide update until all plots have been measured. One way to avoid panel creep is to create subpanels, which amounts to dividing a main panel into smaller panels that each systematically covers the State. The 10 panels out west could be viewed as a 5-panel system with subpaneling. A 5-panel system has 20 percent of the plots in each panel, while a 10-panel system has 10 percent of the plots per panel. Therefore, panel creep in a 10-panel system results in fewer plots being unavailable than in a 5-panel system.

A 20-subpanel design seems like a reasonable compromise. Each panel contains 5 percent of the plots, so leaving a panel uncompleted at the end of the field season is less damaging. At the beginning of the season, the crews could begin measuring as many panels as they expect to complete, say four. It might become clear that only three panels can be finished as the season progresses. A solution is to stop work on one panel so that three full panels are completed. A 20-subpanel design works well when the real goal is to measure either 5 or 10 percent of the plots each year. The current plan to complete a State in 7 years in the East doesn't fit a 20-subpanel design perfectly, but it is close. Measuring 3 subpanels out of 20 each year would be a good approximation.

[1]National Council for Air and Stream Improvement, 600 Suffolk Street, Fifth Floor, Lowell, MA 01854. E-mail: Pvandeusen@ncasi.org.

Mapping Versus Fuzzing

FIA sample plots will not always fall into a uniform forest condition and may straddle multiple conditions. At one time, FIA rotated "straddler" subplots into a uniform forest condition. This created a bias against forest conditions that were long and narrow, and procedures to replace plot rotation were developed (Birdsey *et al.* 1995). The procedure that FIA has officially adopted is referred to as "mapping." This involves mapping the boundary of each distinct condition class that occurs on a subplot, subject to certain limitations. An alternative approach, called fuzzing, is to assign each subplot to a single condition class.

An FIA plot consists of four 1/24-acre subplots in a fixed configuration, and there is one plot for every 6,000 acres. The official justification seems to be that mapping subplots results in the least amount of bias and variance for estimates of condition class volume and area, and this may be true when condition boundaries are well defined. The fuzz method involves classifying each subplot into a single condition class. There are advantages and disadvantages of fuzzing versus mapping (Hahn *et al.* 1995). A major advantage is that field procedures are simple for fuzzing, because no decisions about type boundaries are required.

With fuzzing, a subplot is classified into whatever category occupies the largest proportion of the subplot. Usually, there are no more than two classes per subplot. Suppose we are dealing with pine and hardwood classes and there are several subplots that are 50 percent in each class. Presumably, the field crews would on average call half of these subplots pine and the other half hardwood, which means the end result is unbiased. In any event, a subplot is only 1/24 of an acre, and it seems that splitting it into smaller areas via mapping is unnecessary. The tradeoff in precision may be worth the savings in field-work time, and complexity. This could be formally evaluated with existing mapped data, because the corresponding fuzzed results can be determined from the more detailed mapped data. Such an evaluation would require assuming that the mapped boundaries are distinct and correctly mapped.

The Moving Average

The data from an interpenetrating panel design can be analyzed in many ways, which is a strength of the design. FIA has chosen the 5-year moving average (MA) as the default procedure. A comparison among the MA, a one-panel mean and a mixed estimator (Van Deusen 2002) indicates that the MA works well when there is no trend in the data, but can show significant bias in the presence of a trend.

A brief review of the 5-year moving average will help clarify the problem with bias. The MA, as envisioned for use by FIA, is equivalent to averaging all plot measurements from the last 5 years in a State. For years t-4 through t, this can be written as:

$$MA_{t-4,t} = \sum_{j=t-4}^{t} w_j \bar{y}_j$$

where [$\bar{y}_j$][2] is the average of all plot values measured in year j, and w_j is a weight such that [$\sum w_j = 1$][3] . The weight, w_j, ensures that each panel is weighted according to the proportion of the total plots it contains. With an exact 20 percent sample, [w_j=0.2.] When [$\bar{y}_j$][4] represents a single panel mean for year t, it is unbiased for the true underlying value, [β_t][5] , and we can write [$\bar{y}_t = \beta_t + e_t$][6]

where e_t is a random error term. It follows that the expected value of the moving average is

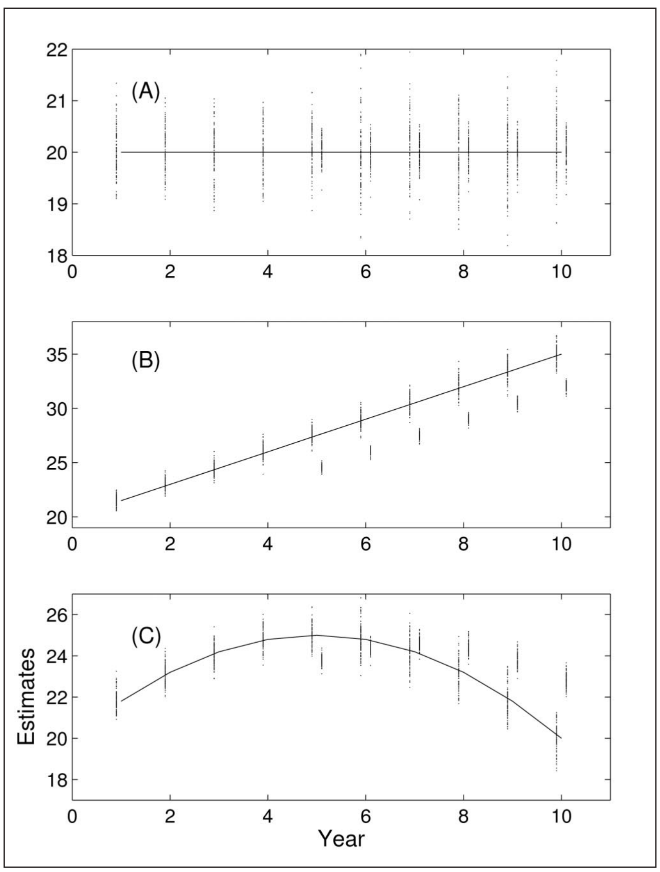
$$E(MA_{t-4,t}) = \sum_{j=t-4}^{t} w_j \beta_j$$

Therefore, [$MA_{t-4,t}$] estimates the true average over the last 5 years and is a biased estimate of the current value, [β_t][8]. If a user wants to estimate the midpoint of a 5-year period, the bias of the MA would generally be less. Most users likely want to be able to associate estimates with particular years, which is the reason an annual inventory received so much support from users.

A comparison using simulated data described elsewhere (Van Deusen 2002) shows the one-panel mean (M1) and the MA (fig. 1) for flat, linear and quadratic trends. The figure displays results from 100 simulations for years 1 through 10. The MA results can be shown only for years 5 through 10, since they require 5 years of data. The MA is clearly superior when there is no trend. However, it's not clear that the MA is superior for the linear trend since the bias is so large. The quadratic trend shows that the sign of the bias of the MA changes after the trend changes direction.

It is clear that the moving average can show significant bias. Other estimators that can follow trends and yield less bias could be implemented by FIA. The MA also has some weaknesses for estimating change. The bias in a linear trend would be subtracted out when estimating change, but the bias in a quadratic trend would be amplified. Research exists on alternative model-based methods and imputation procedures (McRoberts 2001, Reams and Van Deusen 1999, Roesch and Reams 1999, Rubin 1987), but this is an area where more research may be justified.

Figure 1.—*Comparison of the one-panel mean and the 5-year moving average with three trends. Estimates made with simulated data and 100 repetitions are shown for years 1-10 for the one-panel mean and years 5-10 for the moving average. The trends are (A) flat, (B) linear, (C) quadratic.*



Data Security Issues

There is an inherent conflict between making FIA data available to users and maintaining data confidentiality. Data security places limits on the use of FIA data that can diminish its value, but also prevents potential abuses. FIA must consider private property owners' concerns about these data. To obtain the data, FIA field crews must have permission from landowners to access private land where most FIA plots are located. These private owners are not obligated to grant access, and in most cases derive little direct benefit from FIA activities. Owners are concerned about how the information being collected on their land is being used and who has access to it.

The long-term viability of the FIA program depends on striking a balance between adequate security and providing access to users. FIA access refusal rates are somewhat less than 2 percent nationally. If owners lose confidence in FIA data security, refusal rates will go up. At worst, the FIA program could become ineffective. Likewise, FIA will lose political support if users have too little access to the data.

Data security issues were less important before data became easily accessible in digital form. This situation has changed dramatically over the past two decades with the rise of personal computers and the Internet. The first official indication of this trend may have been in section 1770 of the 1985 Food Security Act (FSA), which creates a legal mandate for USDA employees to protect the identity of individuals who provide data. The 1985 FSA had no immediate impact on FIA data confidentiality procedures, but in 1989 FIA made it a policy to release only fuzzed coordinates to the general public. The fuzzed coordinates were adjusted to be within ±1 mile of the true location. The 1985 FSA became applicable to FIA data following a 1999 amendment (H.R. 3423) that inserted a new paragraph including the Forest and Rangeland Renewable Resources Act.

The 1989 system of fuzzed coordinates was deemed to be inadequate in 2001 due to H.R. 3423, and FIA stopped releasing coordinates to the public. Before this, there were publicly available Eastwide and Westwide databases containing fuzzed coordinates. In the future, data released to the public will be aggregated, fuzzed, or rearranged so that it will not be possible to determine who owns the land that contains a particular plot or group of plots. These new restrictions may hinder some traditional industry uses of the data, such as mill supply or location surveys. They will certainly hinder research uses of FIA data that require owner and land use information. FIA needs to develop a policy that will protect private owners' rights while not alienating traditional users by unnecessarily limiting access to the data. It remains to be seen if FIA can accomplish such a difficult compromise.

Summary

FIA is still in the process of replacing the periodic inventory with an annual inventory, and many decisions must be made regarding implementation details. It is often true that an organizational decision involves considering competing options with no clearly superior choice. Regardless, the attributes of each option are weighed and a decision is made. It is clear that another group or decisionmaker might have reached a different conclusion. This paper is written in the spirit of acknowledging that reasonable people can reach different conclusions. Likewise, some of the initial decisions made when implementing the annual inventory could change as new information and viewpoints become available. The issues discussed here were panel creep, mapping versus fuzzing, the moving average, and FIA data security.

Panel creep will almost certainly occur within the annual inventory system. The discussion of panel creep should be about minimizing the impact of not being able to use plot data until a panel is completed. FIA chose to implement 5 panels in the East and 10 in the West, even though few Eastern States would be able to measure 1 panel per year. This ensures that up to 20 percent of already measured plots may be unavailable in the current year. One solution would be to create additional subpanels to decrease the number of unusable, measured plots.

Mapping is a fix for the bias caused by plot rotation. The issue here is that mapping might be unnecessarily detailed and time consuming and that a method known as fuzzing could be sufficient. Fuzzing might result in more bias than mapping, but it would simplify field work and analysis.

The moving average was chosen as the default estimator because it is easy to use. The issue here is that ease of use comes at the expense of bias. In this case, we see FIA accepting bias as a tradeoff for ease of use, whereas bias associated with fuzzing was not accepted. This seems to demonstrate the contention that different groups can reach different conclusions.

The final decisions related to the issue of data security had not been made as this paper was being written. Whatever they turn out to be, there is little doubt that reasonable people will disagree about them. Some will feel they went too far and prevented legitimate uses of the data and others will feel they were too lenient.

Literature Cited

American Forest and Paper Association. 1998. The report of the second blue ribbon panel on forest inventory and analysis. Washington, DC.

American Forest Council. 1992. The report of the first blue ribbon panel on forest inventory and analysis. Washington, DC: American Forest Council.

Birdsey, R.A.; Hahn, J.T.; MacLean, C.D.; *et al.* 1995. Techniques for forest surveys when cluster plots straddle two or more conditions. Forest Science Monograph 31. 82 p. Supplement to Forest Science 41(3).

Hahn, J.T.; MacLean, C.D.; Arner, S.L.; Bechtold, W.A. 1995. Procedures to handle inventory cluster plots that straddle two or more conditions. Forest Science Monograph 31. 82 p. Supplement to Forest Science 41(3).

McRoberts, R.E. 2001. Imputation and model-based updating techniques for annual forest inventories. Forest Science. 47: 322–330.

Reams, G.A.; Van Deusen, P.C. 1999. The southern annual forest inventory system. Journal of Agricultural, Biological, and Environmental Statistics. 4(4): 346–360.

Roesch, F.A.; Reams, G.A. 1999. Analytical alternatives for an annual inventory system. Journal of Forestry. 12: 33–37.

Rubin, D. B. 1987. Multiple imputation for nonresponse in surveys. New York, NY: John Wiley and Sons. 258 p.

U.S. Department of Agriculture. 1999. A strategic plan for forest inventory and monitoring. Washington, DC: Department of Agriculture, Forest Service. 48 p.

Van Deusen, P.C. 2002. Alternative sampling designs and estimators for annual surveys. In: Hansen, M.; Burk, T., eds. Proceedings, Integrated tools for natural researches inventories in the 21st Century, proceedings of the IUFRO conference; 1998 August 16–20; Boise, ID. Gen. Tech. Rep, NC–212. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 192–196.

Van Deusen, P.C. 1997. Annual forest inventory statistical concepts with emphasis on multiple imputation. Canadian Journal of Forest Research. 27: 379–384.