USDA United States
Department of
Agriculture

Forest
Service

North Central
Research
Station

General Technical
Report **NC -213**

Proceedings of the First Annual Forest Inventory and Analysis Symposium

San Antonio, Texas
November 2-3, 1999

North Central Research Station
U.S. Department of Agriculture—Forest Service
1992 Folwell Avenue
St. Paul, Minnesota 55108
Manuscript approved for publication June 12, 2000
2000

Proceedings of the 1st Annual Forest Inventory and Analysis Symposium

November 2-3, 1999

San Antonio Airport Hilton

San Antonio, Texas

Edited By:

Ronald E. McRoberts
Gregory A. Reams
Paul C. Van Deusen

PREFACE

The national Forest Inventory and Analysis (FIA) program of the USDA Forest Service officially initiated annual forest inventories following the mandate of the Agricultural Research, Extension, and Education Reform Act of 1998 (Farm Bill). This Act represents a milestone in a decade of dramatic change for the FIA program. The early part of the decade witnessed two annual inventory pilot studies with similar objectives but somewhat different approaches. In the early 1990s, the Annual Forest Inventory System (AFIS) was initiated in the North Central region to establish the capability of producing standard statewide inventory estimates on an annual basis. In the mid-1990s, the Southern Annual Forest Inventory System (SAFIS) was initiated in the Southern region with extensive state support. With passage of the Farm Bill, a national annual inventory system with a common plot layout, a common sampling design, and common core estimates has emerged. Following the 2000 field season, all FIA regions are expected to have implemented the new system in at least one state.

Throughout the 1990s, research on sampling, remote sensing, estimation, and database methods to support and enhance annual forest inventories was conducted by a variety of Forest Service, university, and forestry-related scientists. The results of these research efforts were reported at conferences and were published in a variety of conference proceedings and journals. However, no single conference or publication documented the approaches to annual forest inventories or the corresponding new methods that had been proposed and developed. Thus, the primary objective of the 1st annual FIA symposium and these proceedings was to document progress by the research stations of the USDA Forest Service and their partners in developing and implementing a national annual forest inventory system. The presentations at the symposium and the papers submitted for publication in these proceedings have achieved that objective. We thank the presenters for documenting their contributions and convey special thanks to those who documented them in writing and submitted them for publication.

Ronald E. McRoberts
Gregory A. Reams
Paul C. Van Deusen
St. Paul, Minnesota

TABLE OF CONTENTS

Page

Background for AFIS, the Annual Forest Inventory System
Ronald E. McRoberts 1

Moving to an annual inventory in the Pacific Northwest
David Azuma 5

The hexagon/panel system for selecting FIA plots under an annual inventory
Gary J. Brand, Mark D. Nelson, Daniel G. Wendt, and Kevin K. Nimerfro 8

Pros and cons of the interpenetrating panel design
Paul C. Van Deusen 14

Using classified Landsat Thematic Mapper data for stratification in a statewide forest inventory
Mark H. Hansen and Daniel G. Wendt 20

Forest/non-forest stratification in Georgia with Landsat Thematic Mapper data
William H. Cooke 28

Comparison of three annual inventory designs, a periodic design, and a midcycle update design
Stanford L. Arner 31

SAFIS area estimation techniques
Gregory A. Reams 32

Diameter growth models using FIA data from the Northeastern, Southern, and North Central Research Stations
Veronica C. Lessard, Ronald E. McRoberts, and Margaret R. Holdaway 37

Evaluating imputation and modeling in the North Central region
Ronald E. McRoberts 43

Annual forest inventory: an industry perspective
Roger Lord 49

Maine's annual inventory: state perspectives
Kenneth M. Laustsen 55

BACKGROUND FOR AFIS, THE ANNUAL FOREST INVENTORY SYSTEM

Ronald E. McRoberts

ABSTRACT.—The Annual Forest Inventory System, AFIS, was jointly proposed and developed in the early 1990s by the Forest Inventory and Analysis programs of the North Central and Rocky Mountain Research Stations of the USDA Forest Service and the Forestry Division of the Minnesota Department of Natural Resources. The objective of AFIS was to establish the capability of producing standard statewide inventory estimates on an annual basis. The context in which AFIS was proposed was defined by two crucial constraints: the average annual cost could be no greater than that of periodic inventories and the precision of estimates could be no less than that obtained from periodic inventories. The system designed to achieve the objective while simultaneously satisfying the constraints included an annual sample of measured field plots, satellite-based remote sensing for area estimation and vegetation change detection, growth and mortality models for updating plots measured in previous years, and computerized databases of field plot information.

INTRODUCTION

The Renewable Forest and Rangeland Resources Planning Act of 1978 requires that the USDA Forest Service conduct inventories of forest land in the United States to determine its extent and condition and the volume of timber, timber growth, and timber removals. Under the auspices of the agency's national Forest Inventory and Analysis (FIA) program, five regional research stations conduct periodic, statewide forest inventories and publish summary reports for individual states. The data and reports resulting from these inventories have been the primary source of information on the status, trends, and use of public and privately owned forest lands in the Eastern United States. This information has been crucial to estimating current forest resource information and monitoring forest ecosystems.

The timeliness, precision, and spatial attributes of products resulting from the Forest Service's periodic approach to forest inventories have been issues of concern to users of inventory data. FIA clients have noted a variety of deficiencies associated with these periodic inventories: (1) the precision of estimates decreases over time due to factors such as changes in land use and tree growth, mortality, and removals; (2) the point-in-time nature of estimates is compromised by the necessity of conducting inventories in heavily forested states over multiple years; (3) the estimates are difficult to integrate across state boundaries because they may be conducted at dates differing by as many as 10 years; and (4) immediate estimates of the effects of catastrophic events such as ice storms, hurricanes, fire, and insect infestations are usually impossible.

In addition, the consequences of management alternatives are often difficult to determine. Changes in the management of public lands have resulted in substantial reductions in timber harvest on National Forest lands in some regions (USDA-FS 1998) with the anticipated response being a sharp increase in timber harvest in other regions (USDA-FS 1995). As a result, many segments of the populace are expressing concerns regarding the ecological and economic sustainability of U.S. forests: some segments are concerned with maintaining biological diversity; some are concerned with satisfying future economic and societal needs; while some are simply concerned that forest management practices conform to their personal value systems. These concerns demand integrated assessments based on current and accurate data to identify trends, relate trends to causes, and evaluate the consequences of alternative management strategies.

FIA clients have recognized these deficiencies and concerns and have proposed solutions such as increasing the sampling intensity, reducing the period between inventories, and conducting midcycle updates. While these solutions might resolve some of the deficiencies, they are expensive to implement and constitute a piecemeal approach to dealing with the problems inherent in periodic inventories.

BACKGROUND

In 1991, FIA scientists at the North Central Research Station (NCRS), USDA Forest Service, proposed research to develop procedures for conducting annual statewide forest inventories. At about the same time, the Resource Assessment Unit, Division of Forestry, Minnesota Department of Natural Resources (MN DNR) initiated investigations into the use of remote sensing technology to support continuously updated inventory on a production basis. Recognizing their mutual interest, the two units contacted the FIA unit at the Rocky Mountain Research Station (RMRS), which had national responsibilities for FIA techniques research. The three-way collaboration resulted in a vision for a comprehensive forest inventory system incorporating an annual sample of measured field plots, satellite-based remote sensing, a computerized database of plot measurement data, and a set of growth and mortality models to update information for plots measured in previous years. This system, the Annual Forest Inventory System (AFIS), had as its overarching objective the development of a set of procedures for forest inventory that would be capable of producing standard statewide forest inventory estimates on an annual basis. Development of AFIS was guided by the following set of constraints:

1. Cost: the cost of implementing and maintaining AFIS in a state would not exceed the average annual cost of conducting periodic inventories.
2. Precision: the precision of the annual AFIS inventory estimates would be as great or greater than those obtained from periodic inventories.
3. Plots: AFIS would maintain the existing network of FIA plots as much as possible.
4. Design: AFIS would modify but not radically redesign FIA procedures.

THE AFIS APPROACH

To a large degree, the first two constraints dictated much of the form that AFIS would take. The 1990 periodic inventory of Minnesota (Miles *et al.* 1995) reported approximately 16.7 million acres of forest land and featured estimates based on data for 13,618 plots. Of these plots, 10,212 plots had been measured between 1986 and 1991, and 3,406 plots had been updated to 1990 using the STEMS growth and mortality models (Belcher *et al.* 1982). Funding for this inventory came from the Forest Service and the State of Minnesota with the latter contributing funds to measure twice as many additional plots as were funded for measurement by the Forest Service. The 3,404 plots whose measurement was funded by the Forest Service were described as a single intensity inventory, while the total of 10,212 plots (1/3 of 10,212 plots) was described as a triple intensity inventory. Thus, the cost constraint for a single intensity inventory was quantified as 262 plots per year, the quotient of 3,404 plots and the 13 years since the previous inventory.

The precision of the triple intensity inventory resulted from data for 13,618 plots distributed over approximately 16.7 million acres. Using the previous 1:2 Federal to State funding ratio, 4,539 of these plots were allocated to the Forest Service's single intensity inventory. Thus, the precision constraint for a single intensity inventory was quantified as no more than 3,675 acres per plot, the quotient of 16.7 million acres and 4,539 plots. Note the apparent incompatibility between the precision constraint of 3,675 acres per plot and the approximately 63,600 acres per plots resulting from the cost constraint of 262 plots per year.

The AFIS solution for simultaneously satisfying the annual cost and precision constraints was to use models to update to the current year information for a large proportion of plots measured in previous years. In particular, of the 4,539 plots for which data would be required annually, 262 plots would be measured while current information for the remaining 4,277 plots would be obtained by using models to update previous measurements. Based on more than a decade of experience, FIA scientists at NCRS were confident that the models would produce adequate results, on average over large areas, for plots that had not had substantial vegetation loss due to mortality, harvest, or other factors. However, the adequacy of results for disturbed plots would be questionable. Thus, the AFIS approach to annual inventories required identification of disturbed plots so that data for them could be obtained from field measurements rather than from model updating. The proportion of forest area in Minnesota losing substantial vegetation each year was estimated at 2 percent or less. Therefore, of the 262 plots to be measured annually, approximately 91 plots (2 percent of 4,539 plots) would be selected because they had experienced substantial disturbance, while the remaining 171 would be selected from among undisturbed plots.

At the rate of 262 plots per year, slightly more than 17 years would be required to complete a single measurement of all the 4,539 plots necessary to satisfy the single intensity precision constraint. However, some plots measured early in a 17-year cycle would be expected to experience vegetation disturbance due to harvest or other removals later in the cycle and therefore would require a second measurement before the end of the cycle. Thus, the time to complete at least one measurement of all plots was extended to 20 years.

MN DNR proposed using remote sensing techniques with satellite data obtained from the Landsat-5 Thematic Mapper as a means of predicting the disturbance status of plots. The basis of the predictions would be differences observed between two sets of TM data for the same area, separated by one or more years, but for the same time of year. Initially, a new set of imagery and a set from 4 years in the past were obtained, indices of vegetation change were calculated for each pixel, and a disturbance status was predicted for each plot. All plots predicted to have experienced substantial vegetation loss would be selected for measurement within the succeeding 4 years, the planned interval between acquisitions of new imagery and predictions of vegetation change.

In summary, current information for 4,539 plots would be required annually to satisfy the precision constraint. Of these 4,539 plots, information for approximately 4,255 plots would be based on measurements in previous years that had been updated to the current year using models, while field measurements would be obtained for the remaining approximately 262 plots. The 262 plots measured in a year would include a random 25 percent sample of those predicted to have experienced substantial vegetation loss based on the most recent remote sensing analyses. The number of these disturbed plots selected each year was expected to be approximately 91 plots. The remaining plots to be measured in a year, expected to be approximately 171 plots, would be randomly selected from among the undisturbed plots with probability of annual selection inversely related to the time since last measurement; i.e., plots that had been recently measured would have a low probability of selection, while plots whose remeasurement interval was approaching 20 years would be selected for remeasurement with a probability approaching one.

FUNDING AND SUPPORT

Funding to support the AFIS research effort was contributed primarily by the USDA Forest Service and the State of Minnesota and secondarily by states and forest industry groups in the Lakes States. NCRS and RMRS contributed the salaries and operating expenses for the equivalent of four full-time scientists whose efforts included construction of a computer database, data editing, development of the sampling design, initial work on construction of diameter growth models and analysis of uncertainty, and preliminary data analyses. MN DNR contributed the salaries and operating expenses for scientists conducting remote sensing research on techniques for using satellite data to distinguish between forest and nonforest lands and to detect forest areas that had lost substantial vegetation. In addition, MN DNR funded the measurement of approximately 750 plots per year from 1992 to 1999. This latter funding was crucial to initiating AFIS while simultaneously proceeding with periodic inventories in the other states for which NCRS has inventory responsibilities. Without this commitment by MN DNR, AFIS would not have been financially feasible. Additional funding to support the modeling research was obtained from the States of Wisconsin and Michigan and forest companies and corporations including ABT, Blandin Paper, Champion International, Colonial Craft, Louisiana Pacific, Mead, Potlatch, Tenneco, and Weyerhaeuser. The latter additional funding was held and disbursed by the Great Lakes Forest Alliance.

EPILOGUE

In the mid-1990s, the Southern Research Station (SRS) implemented an annual inventory system that was both similar and dissimilar in key aspects to the NCRS system. Although the NCRS effort was initiated before the SRS effort, the political and industrial support generated by SRS was primarily responsible for placing annual forest inventories on the national FIA agenda. With passage of the 1998 Farm Bill, formally known as the Agricultural Research, Extension, and Education Reform Act of 1998, Congress required the USDA Forest Service to conduct annual forest inventories in all states. Two features of national annual forest inventories emerged as a result of the Farm Bill: the federally funded base sample would feature one FIA field plot per approximately 6,000 acres and 20 percent of these plots would be measured each year. Based on 16.7 million acres of forest land in Minnesota, the Federal base sample would require approximately 2,800 plots of which approximately 560 would be measured each year. These two requirements of the Farm Bill, one plot per 6,000 acres and measurement of 560 plots per year, greatly relaxed the original AFIS constraints of one plot per 3,675 acres and measurement of 262 plots per year. Relaxation of these constraints had an important impact on annual inventory requirements in the North Central region. Because trees in this region generally grow relatively little in the 5 years between measurements, not updating information for plots measured in previous years is expected to have little detrimental impact on annual estimates. Thus, the roles of models for updating purposes and remote sensing for disturbance detection are non-essential for annual inventories under the Farm Bill, whereas they were absolutely essential under the AFIS constraints. As a result, the AFIS concept of disturbance-based sampling has been abandoned in favor of a systematic distribution of plots across all ownerships and forest cover types, and inventory estimates based on 5-year moving averages have been accepted as defaults, but with provision for optional enhancements using updating techniques based on growth models, imputation, or other methods.

LITERATURE CITED

Belcher, D.W.; Holdaway, M.R.; Brand, G.J. 1982. **A description of STEMS - the stand and tree evaluation and modeling system**. Gen. Tech. Rep. NC-79. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Miles, P.D.; Chen, C.M.; Leatherberry, E.C. 1995. **Minnesota forest statistics, 1990, revised**. Resour. Bull. NC-158. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

USDA Forest Service (USDA-FS). 1995. **The 1995 RPA timber assessment update.** Gen. Tech. Rep. RM-259. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

USDA Forest Service (USDA-FS). 1998. **Report of the Forest Service: fiscal year 1997**. Washington, DC.

ABOUT THE AUTHOR

Ron McRoberts is a Mathematical Statistician and Group Leader with the FIA program at the North Central Research Station, USDA Forest Service, St. Paul, MN.

MOVING TO AN ANNUAL INVENTORY IN THE PACIFIC NORTHWEST

David Azuma

ABSTRACT.—The process of moving toward an annual inventory in the Pacific Coast states began with educating the individual states as to what might be involved in the annual system. The states and some industry groups voiced concerns about inventorying unproductive or reserved lands on an annual basis. The states in particular were concerned about the ability to estimate periodic change with an annual system. The discussion presents these concerns and other possible problems that the Pacific Northwest may face when moving to an annual inventory system.

INTRODUCTION

The process of implementing the 1998 Farm Bill in the Pacific Northwest started with exposing our states (Alaska, California, Oregon, and Washington) and cooperators to the possible changes in the inventory. We explained the five panel design that was put forth in the Farm Bill and the reasons that the South and North were moving to that design. In general, our states are more interested in the change between our periodic inventories than the actual point estimate. At that time all three states were happy with the present 10-year periodic cycle. In fact they would rather we used the additional funds for the annual inventories on other projects within their states. It was apparent from our original meetings that the states would not be interested in funding an annual inventory on a 5-year cycle.

With the exception of Washington in 1988, our states do not generally give us funds to measure plots. The cooperation with our states is usually for extra analysis or additional variables. The State of Washington was interested in an intensification of the number of plots in 1988, but has no interest in spending dollars to increase the number of plots in our upcoming inventory. The State of Oregon cites data from our 1994 inventory (McKay 1998) concerning the small amount of change being measured and doesn't see the reason to go to annual inventories. The states also do not see a need to be able to capture catastrophic events with an annual inventory, since there have only been two such events in the last 40 years in the Pacific Northwest (the Mt. St. Helens eruption, and the Columbus Day storm in the 1960s).

Concern that we need a more frequent evaluation of change is offset by our inventories showing little change. Bolsinger *et al.* (1997) estimated a less than 1 percent loss in softwood growing-stock volume between 1980 and 1991 in Washington. McKay *et al.* (1998) estimated a 3-percent increase in softwood growing-stock volume and less than a 1-percent increase in timberland area between 1986 and 1994 in western Oregon. Waddell and Bassett (1994 and 1997 a,b,c,d, five California reports) found a decrease of approximately 3 percent in primary forest area but an increase of about 7 percent in growing-stock volume between 1984 and 1994 in California (table 1).

There are also concerns from industry and state groups as to why we would inventory the National Parks on an annual cycle. The National Parks also have some reservations about us visiting their land on a yearly basis, and a similar concern was expressed by the Native American community. There are approximately 4.1 million acres of reserved forest lands in California, Oregon, and Washington and a combined Native American acreage of 1.6 million acres. There is also a general concern about an annual inventory of 6.2 million acres of juniper/pinyon lands in Oregon and California where little change occurs and growth is minimal. A similar concern exists in the interior of Alaska, where there are 62 million acres of black spruce. Coastal Alaska also has 4.6

Table 1.—*Changes in area and volume in California, western Oregon, and Washington in the latest inventories on non-Federal timberland*

State	Year[a]	Growing-stock volume	Timberland area
		Million cubic feet	*Thousand acres*
California	1984	24,390	8,247
	1994	26,135	7,971
Western Oregon	1986	19,456	6,729
	1994	19,824	6,758
Washington	1980	39,331	11,939
	1991	39,122	11,452

[a] The year value is the ending year of the inventory.

million acres of reserved land in parks and wilderness. The questionable lands in California, Oregon, and Washington (parks, Native American, and juniper) together represent approximately 12 million acres of possible problems within the annual panels.

In August 1999, we held a client meeting to discuss how we would move to annual inventories. We presented several possible ways of implementing an annual system, including panel, and annualized periodic. We did not get an overwhelming positive response: the State of Washington would like more data on its own lands; Oregon didn't see an advantage to moving to an annual system; and California didn't see a problem with going annual. The key was that no one was going to come forward with money to buy the cycle down from 10 to a smaller number of years. Most of the data users present either were not sure how it would affect their use of the data or thought it would not affect them.

We plan to complete the State of Washington on a periodic design in the next 3 years. This will give us an updated starting point to move into an annual inventory. We will start a 10-percent panel in Oregon in the summer of 2000 and move onto the 6,000 acre per plot hexagon grid in Oregon. If funding continues to progress, we will start an annual system in southeast Alaska in 2002, and in Washington and California in 2003.

LITERATURE CITED

Bolsinger, C.L.; McKay, N.; Gedney, D.R.; Alerich, C. 1997. **Washington's public and private forests**. Resour. Bull. PNW-218. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 144 p.

McKay, N.; Bolsinger, C.L.; Lettman, G.L.; Mei, M.A.; Azuma, D.L. 1998. **Timber resource trends on non-federal timberland in western Oregon between 1984-86 and 1994**. Salem OR: Oregon Department of Forestry. 126 p.

Waddell, K.L.; Bassett, P.M. 1994. **Timber resource statistics for the North Coast Resource Area of California**. Resour. Bull. PNW-214. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 50 p.

Waddell, K.L.; Bassett, P.M. 1997a. **Timber resource statistics for the Sacramento Resource Area of California**. Resour. Bull. PNW-220. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 50 p.

Waddell, K.L.; Bassett, P.M. 1997b. **Timber resource statistics for the Central Coast**

Resource Area of California. Resour. Bull. PNW-221. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 45 p.

Waddell, K.L.; Bassett, P.M. 1997c. **Timber resource statistics for the North Interior Resource Area of California**. Resour. Bull. PNW-222. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 49 p.

Waddell, K.L.; Bassett, P.M. 1997d. **Timber resource statistics for the San Joaquin and Southern Resource Areas of California**. Resour. Bull. PNW-224. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 51 p.

ABOUT THE AUTHOR

David L. Azuma is a Research Forester with the Foresty Inventory and Analysis program at the Pacific Northwest Station, Forestry Sciences Laboratory, USDA Forest Service, Portland, OR.

THE HEXAGON/PANEL SYSTEM FOR SELECTING FIA PLOTS UNDER AN ANNUAL INVENTORY

Gary J. Brand, Mark D. Nelson, Daniel G. Wendt, and Kevin K. Nimerfro

ABSRACT.—Forest Inventory and Analysis (FIA) is changing to an annual nationwide forest inventory. This paper describes the sampling grid used to distribute FIA plots across the landscape and to allocate them to a particular measurement year. We also describe the integration of the FIA and Forest Health Monitoring (FHM) plot networks.

INTRODUCTION

In 1998, Federal legislation (Agricultural Research, Extension, and Education Reform Act of 1998 - PL 105-185) was passed that requires major changes in the way Forest Inventory and Analysis (FIA) conducts inventories of the nation's forest resources. This legislation resulted from concerns expressed by FIA clients that changes were needed in existing FIA methods (Van Deusen *et al.* 1999, Gillespie 1999).

A fundamental change that the legislation requires is an annual inventory of each state, with 20 percent of the plots within a state measured each year. In contrast, FIA inventories have historically been conducted within a single state over 1 to 3 years; each state has been re-inventoried every 6 to 8 years in the South and every 11 to 18 years in the rest of the country (Gillespie 1999).

In addition to FIA, the Forest Health Monitoring (FHM) program also collects data on our nation's forests. FHM data are collected annually on a 4-year cycle. Given the overlap in the FIA and FHM programs, we have an opportunity to increase the efficiency of data collection by merging the two programs (Gillespie 1999). The remainder of this paper describes the sample design for implementing the annual inventory and how it has been modified to accommodate the integration of the FIA and FHM programs.

CONSTRUCTING THE HEXAGON SAMPLING FRAMEWORK

One advantage of an annual inventory is the increased ability to quickly measure the effects of events that occur over large areas, such as hurricanes, ice storms, and windstorms. To do so requires a spatially regular distribution of plots across the landscape measured each year. The FHM program has addressed this same need for regularly distributed plots by using a lattice of hexagonal cells as a sampling framework (Scott *et al.* 1993). A base hexagon positioned over the conterminous United States was subdivided into approximately 28,000 hexagons whose centers are about 16.9 mi (27 km) apart (White *et al.* 1992). One field plot was selected for each hexagon, usually the existing FIA plot closest to the center of the hexagon. Each of the hexagons was assigned to one of four panels; a panel corresponds to a given measurement year of the cycle. After the fourth panel is measured, the cycle is repeated. One of the advantages of this framework is that it is unlikely to be aligned with regularly spaced landscape features.

Because of these desirable features, we explored the possibility of using the FHM framework as the basis for the FIA annual inventory sampling framework. To meet its mandated maximum sampling errors, the FIA program requires a sampling intensity of one plot per approximately 6,000 acres (M. H. Hansen 1998, pers. comm.). By creating a new lattice of hexagonal cells where each hexagon is 1/27 the size of an FHM hexagon, the desired sampling intensity is achieved (A. R. Olsen 1998, pers. comm.). The size of each FIA hexagon is 5,937.2 acres.

Staff of the Western Ecology Division of the U.S. Environmental Protection Agency National Health and Environmental Effects Research Laboratory in Corvallis, OR, performed a 27-factor enhancement of the FHM hexagons, resulting in more than 360,000 FIA hexagons within and adjacent to the border of the conterminous United States. To minimize distortion of the area associated with each hexagon, the Lambert azimuthal equal-area projection was used when creating the FHM and FIA hexagons. Figure 1 shows the spatial arrangement of an FHM hexagon and the FIA hexagons. Attributes included for each hexagon were a unique 8-digit hexagon ID, a hierarchical ID that can be used to decipher how the hexagon was generated, and another ID that can be used to determine the U.S. Geological Survey 7.5 minute quadrangle containing the center of the hexagon. We also determined the state and county where each FHM and FIA hexagon center was located based on 1:100,000 U.S. Bureau of the Census TIGER/Line files.

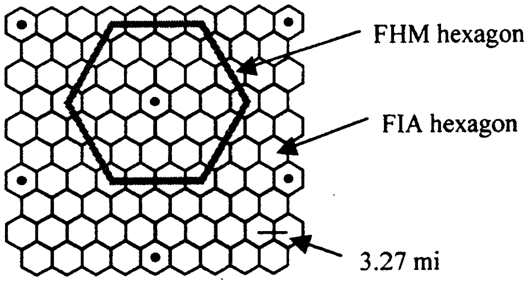


Figure 1.—*The FIA hexagon lattice. Each black dot is at the center of an FHM hexagon.*

By assigning one plot to each FIA hexagon, we create a regular spatial distribution of plots across the landscape. The 1998 legislation requires that 20 percent of the plots be measured each year. To distribute the hexagons temporally, each is assigned to one of five panels. The arrangement shown in the next column (fig. 2) distributes the hexagons among the five panels in such a way that no adjacent hexagons belong to the same panel. The plots in hexagons from panel one were measured from the fall of 1998 through the summer of 1999 in Indiana, Iowa, Minnesota, and Missouri. Each of the remaining panels is assigned to succeeding years; panel one will be measured again from the fall of 2003 through the summer of 2004. As annual inventories begin in new states, we start with the same panel that is being measured that year in states already under an annual inventory. Although the intent in the Eastern U.S. is to operate on a 5-year cycle, funding or ecological conditions may require other cycles. In particular, 7-year (as an eastern option) and 10-year (as a western standard) cycles have been proposed. These two panel arrangements are shown in figure 3.

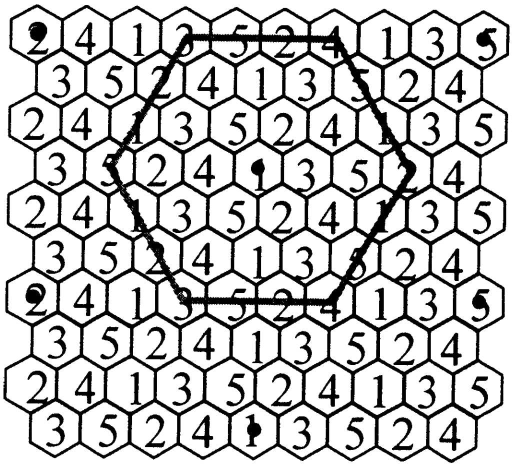


Figure 2.—*Assignment of hexagons to one of five panels (shown by number).*

SELECTING PLOTS

There are many ways to select a ground plot for each FIA hexagon. We followed two guiding principles for determining plot selection procedures. The plot selected for an FIA hexagon:

1. must be located in that hexagon and
2. should be an FIA ground plot, if one exists, thereby retaining as much historical information as possible.

However, to satisfy the first principle, the geographic location of existing plots must be known. Therefore, the first step in plot selection was to establish the latitude and longitude for all existing plots. This was most often done by transferring marked plot locations on aerial photos to geo-referenced satellite imagery. The next step was to spatially overlay the plot locations and the FIA hexagons in a GIS application. The distance from the plot to the center of its hexagon was also computed and recorded. A database management procedure then assigned one plot to each hexagon based on the following criteria:

1. if the hexagon contains an FHM plot, select it;
2. if not, then select the FIA plot within the hexagon that is closest to the center of the hexagon;

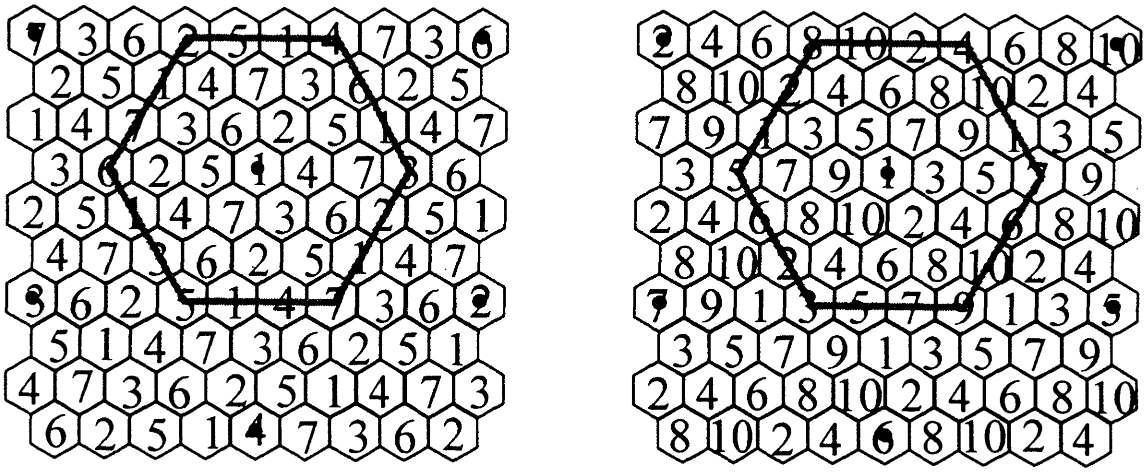


Figure 3.,—*Seven-panel (left) and ten-panel (right) arrangements.*

3. if there are no FHM or FIA plots in the hexagon, select the center of the hexagon as the location for a new plot (some regions may choose a location near the center).

In some of our states we had to adjust the probability of selection because of unequal sampling intensities in the previous inventory. For example, in Wisconsin reserved areas were sampled more intensively than other areas. Figure 4 illustrates the selection of plots for various situations encountered.

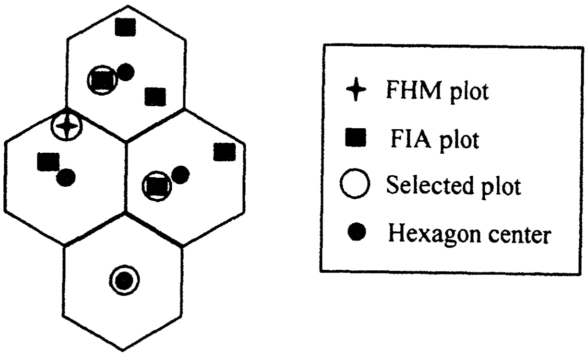


Figure 4.—*Example results of plot selection criteria.*

INTEGRATING FIA AND FHM SAMPLING FRAMEWORK

Although FHM is a national program, FHM plots have not yet been established in all states. For example, in the North Central region, FHM plots have not been established in five (Iowa, Kansas, Nebraska, North Dakota, and South Dakota) of our 11 states. However, in states where there are existing FHM plots, it is important to retain these plots, not only to keep them as plots selected for their respective FIA hexagons, but also to measure them in their same temporal order.

As noted earlier, the FHM plots were measured on a 4-year cycle, whereas the legislation mandating annual inventories specified a 5-year cycle. In addition to a 4-year cycle, one-third of the FHM plots were also measured in two consecutive years (overlap plots). To maintain the existing temporal intensity of FHM plot measurements over a 5-year cycle, the FHM program elected to increase the number of plots established in most states by 67 percent. Half of the additional plots are needed to make up for the overlap plots (years one to four). The other half are needed for the fifth year. The only exceptions were in Maryland and Minnesota where state funding permitted an original sampling intensity three times greater than in other states.

We obtained a grid of both new and original locations and panel assignments for FHM plots from the Forest Health Monitoring Program, USDA Forest Service, Research Triangle Park, NC (William D. Smith 1999, pers. comm.). The grid point designates the desired approximate location for an FHM plot, but not necessarily the ultimate location of the plot. In states where FHM plots had not been selected, this grid was regularly spaced across the state and equally distributed among the five panels. In states where FHM plots had been selected, additional grid points were systematically interspersed among the original grid points. Old locations kept their original FHM panel assignment. Approximately 50 perecent of the new grid points were assigned to a new fifth panel, and the rest were spread evenly among the other four panels. In Maryland and Minnesota, no new grid points were needed and the panel assignment was based on the FIA panel assigned to that location. That is, the existing FHM plots were simply distributed among five panels.

We chose an FIA hexagon for every FHM grid point within the conterminous U.S. based on the following rules:

1. if there is an existing FHM plot associated with an FHM grid point, then choose the FIA hexagon containing the FHM plot as the FIA/FHM hexagon; change its panel to match the panel of the FHM grid point (fig. 5),
2. if an existing FHM plot is not associated with the FHM grid point, then choose the nearest FIA hexagon of the same panel and in the same state as the FHM grid point to be the FIA/FHM hexagon (fig. 6),
3. if none of the nearby FIA hexagons (nearer than 8,500 m) are of the same panel and state as the grid point, then choose the one containing the FHM grid point as the FIA/FHM hexagon; change its panel to match the panel of the FHM grid point (fig. 7). This condition occurs along state borders and coastlines.

Rules 1 and 3 change the panels assigned to FIA hexagons and therefore disrupt the original FIA pattern. However, because there are only about 1/16 as many FHM grid points as FIA hexagons, and not all of the FHM grid points will cause the panel of the FIA hexagon to change, this disruption was considered acceptable.

CONCLUSIONS

The hexagon/panel system is one way to distribute FIA plots systematically across the conterminous United States and through time. One plot is selected for each FIA hexagon. Existing FHM and FIA plots are selected whenever possible. To maintain the existing temporal order of FHM plots, some perturbation of the FIA panels was accepted and incorporated into the system. The expected results will be a

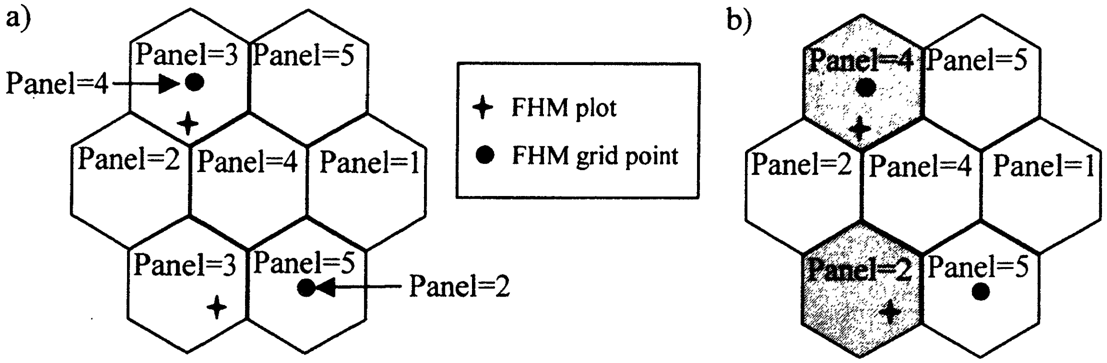


Figure 5.—*Examples where rule 1 changes the original panel of the FIA hexagon a) to a different panel b). The shaded hexagons have different panels.*

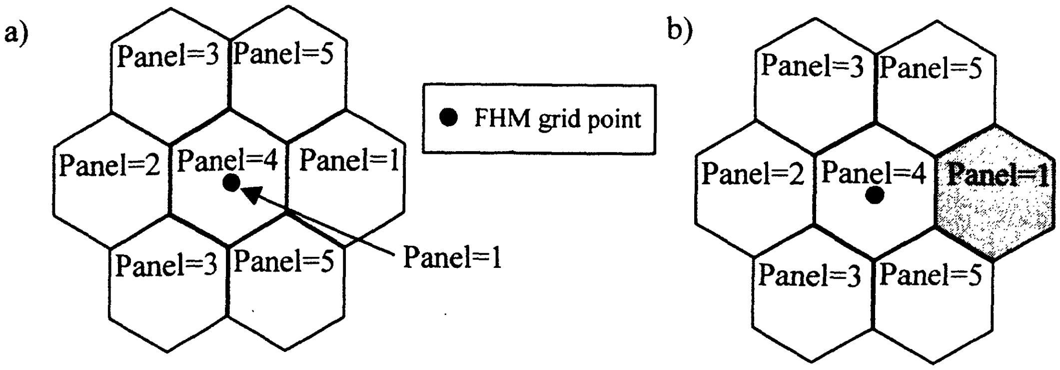


Figure 6.—*Implementation of rule 2. The panel of the FHM grid point in a) results in the shaded hexagon b) becoming the FHM/FIA hexagon.*

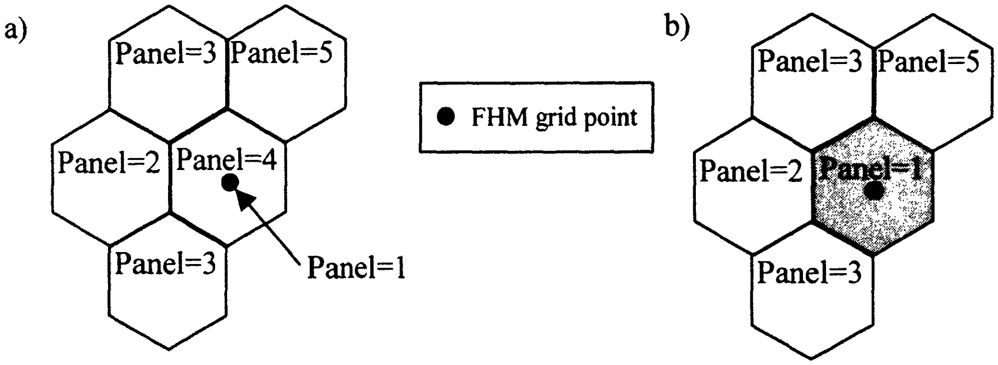


Figure 7.—*Before a) and after b) implementing rule 3. The panel of the shaded hexagon has been changed.*

consistent inventory of all forested lands that preserves historic data. This system will incorporate the FIA and FHM forest inventory efforts, comply with legislative mandate, and provide a framework for future forest inventories.

ACKNOWLEGMENTS

We thank Anthony R. Olsen, Suzanne Pierson, and Barb Rosenbaum of the Environmental Protection Agency in Corvallis, OR, for providing the FIA hexagons. We also thank the following people from the USDA Forest Service: Ronald McRoberts and Mark Hansen of the North Central Research Station in St. Paul, MN, and Andrew Gillespie of the Resource Valuation and Use Staff in Washington, DC, for their suggestions and input in designing the hexagon/panel system; and William D. Smith of the FHM program in Research Triangle Park, NC, for the original and new FHM grid locations.

LITERATURE CITED

Gillespie, A.J.R. 1999. **Rationale for a national annual forest inventory program.** Journal of Forestry. 97: 16-20.

Scott, C.T.; Cassell, D.L.; Hazard, J.W. 1993. **Sampling design of the U.S. National Forest Health Monitoring Program**. In: Proceedings of the Ilvessalo symposium on national forest inventories; 1992 August 17-21; Helsinki, Finland. Res. Pap. 444, Finnish Forest Research Institute: 150-157.

Van Deusen, P.C.; Prisley, S.P.; Lucier, A.A. 1999. **Adopting an annual inventory system: user perspectives**. Journal of Forestry. 97: 11-14.

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

ABOUT THE AUTHORS

The authors are employees of the North Central Research Station, U.S. Department of Agriculture, Forest Service, St. Paul, MN.

PROS AND CONS OF THE INTERPENETRATING PANEL DESIGN

Paul C. Van Deusen

ABSTRACT.—The interpenetrating sample design has been selected for the USDA Forest Service's Annual Forest Inventory System. The advantages and disadvantages of this design are discussed by considering alternatives such as the formerly used periodic design, a concentrated grid design, and disturbance based sampling. Factors considered for each design include fulfilling 1998 Farm Bill requirements, relative cost, ease of implementation, and analysis options. Each design alternative has positive and negative attributes, but the interpenetrating design most clearly facilitates implementation of the new annual inventory system.

INTRODUCTION

A survey designer has an array of choices to confront when deciding how samples will be allocated in the field. Sample allocation also affects the options for analyzing the data. Survey design involves both of these choices, but the emphasis here will be on the sample allocation aspect.

Sample allocation decisions for USDA Forest Service Forest Inventory and Analysis (FIA) surveys must take into account the public-use nature of FIA data. First and foremost, the data should be amenable to standard analyses. Sample allocation should lead to robust data in the sense that the data are not optimized only for a limited purpose. For example, the data should not be collected in such a way that they are optimal for estimating forest growth but inadequate for estimating current volume by species. Unfortunately, it is inevitable that a design to optimize for variable A leads to suboptimality for variable B.

With the above factors in mind, FIA selected the interpenetrating design for the annual forest inventory system. This design allocates field plots to five panels that each provide systematic coverage for a state or any other region of interest. The systematic coverage implies that no variable is favored at the expense of another, and a number of analysis methods are valid.

SAMPLE DESIGN ALTERNATIVES

Each of four design alternatives will be discussed, followed by a section that suggests minor modifications to improve the interpenetrating (INT) design. The old FIA periodic (PER) design, which is being phased out, is discussed first. The INT and PER designs weight all plots equally, and both lend themselves to simple analysis options. The concentrated grid (CON) design can be viewed as a hybrid of INT and PER. A CON design would divide each state into five regions, and one region would be measured each year. The fourth design being considered is disturbance based sampling (DIS), which results in an annual system where disturbed plots are sampled with higher probability than undisturbed plots. As such, the DIS design is the only one that attempts to optimize for certain characteristics.

Periodic Design

FIA has employed the periodic design since the program began in the 1930s. Ideally, under the PER design, all plots in a state are measured in the same year. In practice, it may take 3 or 4 years to complete large, heavily forested states like Georgia or Maine. This design puts all of FIA's attention on the current few states where field work is underway. It provides a snapshot of the state that has maximum accuracy immediately following the field work and then deteriorates over the years until the next measurements are available. Estimates are derived under the PER design by averaging over all the data, which implicitly assumes that all plots were measured at the same time. The periodic system worked well for the first 50 years of FIA's existence, but began drawing criticism around 1990 primarily because the survey cycle was too long.

The first Blue Ribbon Panel report (American Forest Council 1992) (BRP I) did not call for dismantling the PER design, but requested reduction of the survey cycle from 10 years to 5 years. Some progress was made on this, but flat budgets and increasing demands on FIA ultimately caused the cycle to become longer than ever.

Some of the positive aspects of this design are:

- The design can be funded on a 5-year cycle.
- This design maintains the status quo so no other changes are required.
- Simple analysis options are available.

Some negative aspects of this design are:

- Attempts to improve timeliness under this design following BRP I failed.
- Budgets and activities within a state fluctuate wildly over time.
- Cross-state analysis is difficult because adjacent states are measured in different years.

Interpenetrating Design

The INT design was originally developed for the Southern Annual Forest Inventory System (SAFIS) pilot study, which began in 1995. The INT design is similar to the National Forest Health Monitoring design and calls for annual measurement of panels that consist of plots that systematically cover the region of interest. This design appealed to many southern state foresters who consequently supported SAFIS. The INT design made SAFIS somewhat compatible with the original Annual Forest Inventory System (AFIS) pilot study that began in the Lake States in 1992. The use of different designs for AFIS and SAFIS gave FIA the opportunity to study two alternative ways of "going annual." AFIS used the DIS design discussed below.

The same plots used for the PER design, laid out on the national FIA grid, are used for the INT design with minor exceptions. The annual panels for the INT design consist of roughly equal numbers of plots that are systematically distributed over the FIA grid for each panel. Therefore, data from an annual panel could be analyzed using methods used for the PER design. However, the precision of the estimates obtained from a single INT panel would be less than that obtained from the full PER sample. Alternative estimation procedures (Reams and Van Deusen 1999) that use data from previously measured panels can significantly improve INT design precision. Multiple imputation (Rubin 1987, Van Deusen 1997), which involves updating unmeasured plots with models or database matching, is one viable option. A moving average estimator can also incorporate measurements from all panels without complications due to updating.

The second Blue Ribbon Panel report (American Forest and Paper Association 1998) (BRP II) concluded that FIA should move to an annual INT design that would measure 20 percent of the plots in a state annually. Subsequently, the 1998 Farm Bill mandated that FIA adopt the INT design and produce a strategic plan (USDA Forest Service 1999) for implementation.

Some of the positive aspects of this design relative to the PER design are:

- It meets the 1998 Farm Bill requirements.
- Cross-state analyses are temporally consistent.
- Budgets don't fluctuate annually by state.
- The data can be analyzed by a number of approaches.
- The States are move involved.
- New computer programs will be developed for data management and analysis.

Some negative aspects of this design relative to the PER design are:

- Longer travel time between plots is required.
- The precision is lower in any given year.
- New software for data management and analysis is required
- Requirement of more state involvement could be problematic.

Disturbance Design

The disturbance sampling design (DIS) was developed for the AFIS pilot study in the Lake States. This design allocates sampling effort to plots with probability proportional to disturbance. The design called for measuring all disturbed plots each year and then taking a random or systematic sample of undisturbed plots. Disturbance would be detected via remote sensing. This design would be very good for determining the amount and impact of disturbance, but it leads to more complicated analysis options than either the INT or PER designs. Any analysis would have to differentiate between plots that were measured because they were disturbed versus the randomly chosen undisturbed plots. Proponents feel that the DIS design could be more economical to implement than a rigid INT design where 20 percent of the plots are measured each year.

The DIS design depends strongly on remote sensing to detect disturbance. This capability is available for Minnesota courtesy of the state Department of Natural Resources, but not necessarily for other states. The DIS design also depends on models to make predictions for unmeasured, undisturbed plots. However, it is statistically problematic to incorporate modeled plots that are selected with a different probability than the measured disturbed plots. One must account for the fact that models predict expected plot means rather than individual plot values. Therefore, treating modeled predictions like actual measurements leads to understating the true variance. Multiple imputation (Rubin 1987) is one way to incorporate variance into the process and to use models in a valid manor. This approach requires making several predictions for each plot and incorporating variability into the predictions. However, the systematically different handling of disturbed and undisturbed plots under the DIS design complicates the use of multiple imputation. In effect, the DIS design creates two strata: a disturbance stratum and a non-disturbance stratum. The complications arise, in part, because these strata change each year (Van Deusen 1993). Resulting change estimates will involve plots that were measured with probabilities and stratum that change over time. Incorporating modeled estimates under the INT design is much easier, specifically because of the equal probability (systematic) plot selection process.

Some of the positive aspects of this design relative to the INT design are:

- It can be very economical.
- Sampling is optimized for disturbed areas.
- It uses remote sensing to improve sampling efficiency.

Some negative aspects of this design relative to the INT design are:

- Statistical analysis is difficult.
- It is optimal for disturbance, but sub-optimal for growth.
- It depends on remote sensing to detect disturbance.
- It depends on models.

Concentrated Grid Design

The concentrated grid (CON) design has been proposed as a compromise between the INT and PER designs. A CON design calls for measuring an equal portion of the plots each year by dividing each state into five concentrated zones. In this way, annual measurements would be taking place in each state, but each within-state zone would be under a periodic survey. The CON design is very similar to the PER design, which divided states into survey units that were usually measured one at a time. Some would argue that it also meets the Farm Bill requirements, even though it circumvents the spirit of the Farm Bill. The CON design might also allow for reduced travel costs relative to the INT design. The CON design would make it difficult to produce state-level reports because plots in different parts of the state are measured in different years.

Some of the positive aspects of this design relative to the INT design are:

- It may meet the Farm Bill requirements.
- Travel costs could be lower.
- Precision for sub-regions is higher for a given year.
- It is similar to the PER design and therefore involves less change.

Some negative aspects of this design relative to the INT design are:

- It may not meet the spirit of the Farm Bill.
- It makes cross-region analyses difficult.
- It is more periodic than annual in nature.

MODIFICATIONS TO THE INTERPENETRATING PANEL DESIGN

The INT panel design has been chosen for the new annual forest inventory system being implemented by FIA. Alternative designs are of academic interest, but FIA has already made substantial commitment to the INT design. While the INT design has many desirable attributes, it has some aspects that can be legitimately criticized. The purpose of this section is to suggest minor modifications to the standard INT design to rectify limitations that it can impose on analysis and logistics options.

Strict adherence to the INT design would eventually lead to having only 5-year growth intervals in the database. However, estimates will be made annually, which implies that the INT design is not design unbiased. In other words, estimates from a rigid INT design for intervals other than 5 years would depend on models/assumptions. Fluctuating budgets and special surveys may create additional problems with a rigid INT design. For example, it would be difficult to measure 20 percent of the plots annually in a year when the budget is reduced. It would be equally problematic to measure more than 20 percent of the plots in a budget increase year without deviating from the basic design. However, a simple alteration to the basic INT design can alleviate these problems. Consider the possibility of creating small clusters of adjacent plots. Each cluster would contain one plot for each panel being maintained under the INT design. For example, the basic design that meets Farm Bill requirements has five panels, so each cluster would have five plots (fig. 1). Panel assignments could be rotated within clusters at periodic intervals as a simple way to obtain design unbiasedness. Thus, the plots change panel membership on a periodic basis. This ensures that a mix of growth intervals is always being measured.

The second problem with the basic INT design can be alleviated by creating "extra" panels, preferably in increments of five (fig. 2). If there were 10 panels, for example, it would still be possible to measure 20 percent of the plots each year by measuring two panels. Measuring 3 of 15 or 4 of 20 panels would also work. Extra panels become advantageous when the need arises to deviate from annually measuring 20 percent of the plots. If the budget decreases under a 15-panel system, one can drop back to measuring either one or two panels rather than the usual three per year. Alternatively, the number of panels can be increased in a good budget year.

The extra panel approach adds flexibility to the basic INT design, so that fluctuating budgets or special surveys can be seamlessly accommodated. Rotation of within cluster plot-to-panel

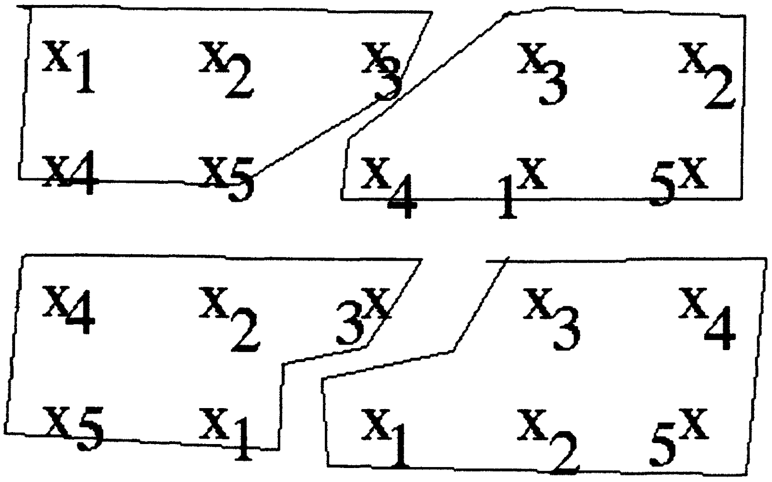


Figure 1.—*A five-panel design showing four clusters of five plots each. The plot location is represented by an x. The panel assignment is given by the number next to the plot.*

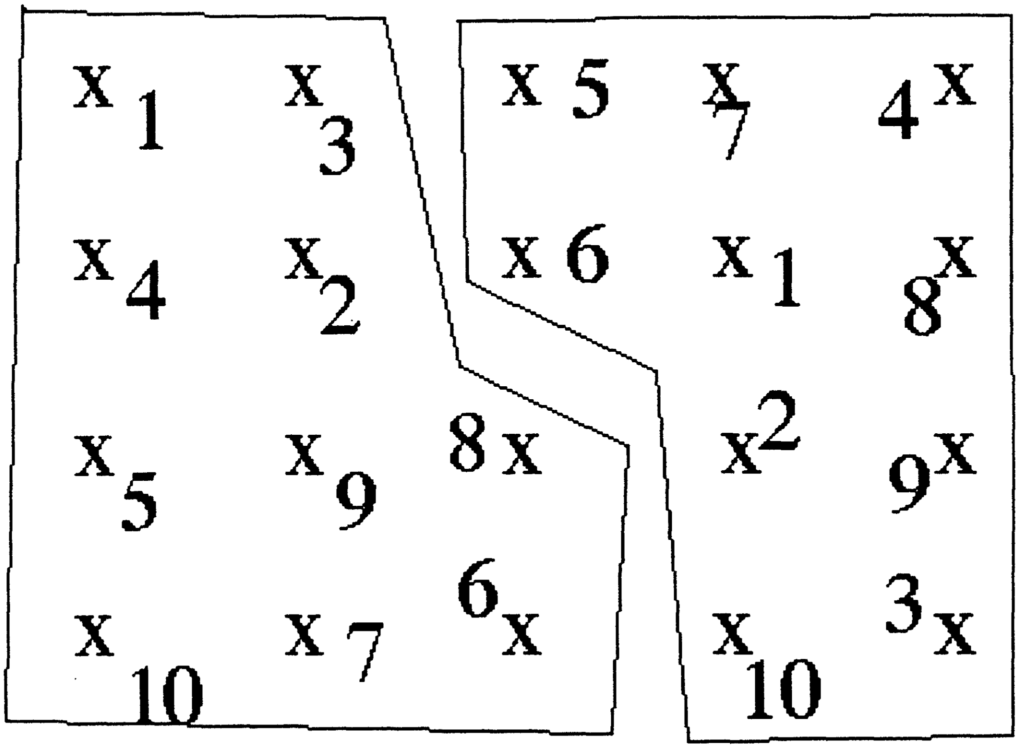


Figure 2.—*A 10-panel design showing four clusters of 10 plots each.*

assignments allows for a range of measurement intervals to be present in the database. This adds the desirable feature of design unbiasedness to the resulting estimates and can work within the context of extra panels. These ideas are discussed in somewhat more detail in Van Deusen (2000).

SUMMARY

Four sampling designs that could be used by FIA have been briefly discussed, with emphasis on the INT design that has already been selected for the new annual forest inventory system. All designs could operate with plots laid out on the national FIA grid using traditional field procedures. Recent changes made to FIA plot configuration and measurements were not required for the annual inventory system. For example, the decision to change from variable radius plots to fixed area plots with mapping (Scott and Bechtold 1995) was made prior to the 1998 Farm Bill. Current plans also call for fitting FIA field plot locations to a triangular grid (Roesch and Reams 1999). This will result in equal plot intensity nationally and will facilitate formation of five panels for the INT design.

Analysis options for either the PER, INT, or CON designs have much overlap because all use systematic, equal-probability sampling. The DIS design selects disturbed and undisturbed plots with different probabilities, and depends on remote sensing and models to be effective. The remote sensing and modeling capabilities required for the DIS design are not available at this time in each state, which precludes the use of this design at the national level. However, the advantages that can accrue from modeling and remote sensing can also be realized under the INT design. The INT design does not require models or remote sensing to be effective, but they can be used if available. The INT design can use models within the context of a procedure like multiple imputation to improve the precision of estimates and obtain valid confidence intervals. Therefore, the choice of the INT design by FIA is a prudent one.

LITERATURE CITED

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

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

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: Wiley.

Scott, C.T.; Bechtold, W.A. 1995. **Techniques and computations for mapping plot clusters that straddle stand boundaries**. Forest Science Monograph. 31: 46-61.

USDA Forest Service. 1999. **A strategic plan for forest inventory and monitoring**. Washington, DC.

Van Deusen, P.C. 1993. **Sampling moving strata on two occasions**. Canadian Journal of Forest Research. 23: 96-100.

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

Van Deusen, P.C. 2000. **Alternative sampling designs and estimators for annual surveys**. In: Hansen, M.H.; Burk, T.E., eds. Integrated tools for natural resource 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 Research Station.

ABOUT THE AUTHOR

Paul C. Van Deusen is a Principal Research Scientist for the National Council of Air and Stream Improvement (NCASI) at Tufts University in Medford, MA.

USING CLASSIFIED LANDSAT THEMATIC MAPPER DATA FOR STRATIFICATION IN A STATEWIDE FOREST INVENTORY

Mark H. Hansen and Daniel G. Wendt

ABSTRACT.—The 1998 Indiana/Illinois forest inventory (USDA Forest Service, Forest Inventory and Analysis (FIA)) used Landsat Thematic Mapper (TM) data for stratification. Classified images made by the National Gap Analysis Program (GAP) stratified FIA plots into four classes (nonforest, nonforest/forest, forest/nonforest, and forest) based on a two pixel forest edge buffer zone. Estimates based on two-phase sampling for stratification were made at the county level. This procedure differed from methods used in previous inventories where stratification was based on the stereoscopic examination of aerial photo plots. Changes in plot design, sampling intensity, and population parameters between 1986 and 1999 make it impossible to attribute differences in sampling errors entirely to this change in methods. The stratified sample estimates based on TM data provided good estimates and greatly reduced costs by eliminating the need for thousands of aerial photos and manual interpretation of several hundred thousand photo plots.

INTRODUCTION

FIA statewide inventories provide estimates of forest resource parameters such as forest area and timber volume, growth, removals, and mortality estimates at the state, unit, and county level. FIA has used two-phase sampling for stratification (also called double-sampling) for a number of years. Cochran (1977) presents a good general description of double-sampling for stratification, and Loetsch and Haller (1964) present it in a forest inventory context.

This paper describes how the FIA program at the North Central Research Station (NCFIA) used Landsat TM imagery to replace aerial photos for phase one estimates in two statewide inventories. Using digital data to replace manual interpretation of aerial photos greatly reduced the work required to complete these inventories. First, we describe the methods used in the past and then the methods used in 1998. Finally, results from these two inventories are presented to show differences related to changes in phase one procedures.

BACKGROUND

Two-Phase Sampling for Stratification

NCFIA uses two-phase sampling for stratification. Phase one is a large sample used to estimate the size (area) of each stratum and phase two is a sub-sample of the phase one plots that estimates the mean within each stratum. Population level estimates (means and totals) are weighted sums of within stratum estimates. The variance of the estimate is reduced by selecting strata that have low within stratum variances and by increasing sampling. Increasing phase one intensity decreases the variance of the estimated stratum size estimates, and increasing the phase two intensity decreases the variance of the estimated within stratum estimates.

In two-phase sampling for stratification, it is very important that phase one determines strata for all plots (both phase one and phase two plots). It is also important that this assignment to a stratum does not change in phase two when the plot is measured and that the procedures used to determine this classification be identical for all phase one plots regardless if they are also phase two plots. Often in forest inventory applications, the strata have names similar or identical to ground classifications that are of interest in the final estimate. For example, the class of sawtimber may be one stratum. In the phase one sample, a plot is assigned to the sawtimber class based on aerial photo classification. This plot could be a phase two plot and thus sent to the field where it may or may not be ground classified as sawtimber.

Regardless of its ground classification, in the estimation procedure it must be included in the sawtimber stratum.

1986 Inventory Procedures

In 1986, NHAP photos (1:40,000) were assembled into township mosaics, and a systematic grid (one plot per 190.4 acres) was overlaid on each township mosaic. These phase one photo plots were classified by land use, forest type, and stand-size density. A total of about 250,000 photo plots formed the basis for the 1986 stratification. The photo classifications were collapsed into the 11 strata in table 1 for the final estimation.

Table 1.—*Aerial photo (size-density) classes used for stratification in 1986*

1	Sawtimber high	6	Seedling/sapling high
2	Sawtimber low	7	Seedling/sapling low
3	Poletimber high	8	Questionable
4	Poletimber med	9	Nonforest with trees
5	Poletimber low	10	Nonforest without trees
		11	Water

A systematic sample (every 17th) of photo plots was selected as a phase two sample and further examined to measure the parameters of interest. The plot design used in 1986 was a cluster of ten 37.5 basal area factor (BAF) sample points distributed over a 1-acre area. The plot was arranged with all 10 sample points in a single land use (forest or nonforest) as determined by plot center. Under this plot design, each plot represents a binary observation for all area estimates.

Problems With 1986 Inventory Procedures

Two procedural problems were identified that could produce bias in the estimates. One problem relates to methods used on phase two plots that were obviously nonforest on the photos. The other problem was an apparent inconsistency in the classification of phase two plots.

In 1986 phase two plot locations were defined by a pin prick on the photo; however, plots classified in stratum 10 (nonforest without trees) and 11 (water) were treated somewhat differently. All strata 1-9 phase two plots were field visited and permanently marked by a stake in the ground. In strata 10 and 11 there were very few field visits and the aerial photo classification was considered the correct observation. This eliminated visiting thousands of nonforest plots that were obviously nonforest from the photo. It assumes no error in the location and classification of these plots and no conversion to forest between the dates of photography and inventory. These assumptions were fairly safe due to a stable agricultural economy, but the lack of a field check of these two strata probably resulted in a small underestimation of forest land in 1986. Plots in these two strata were field visited only in a few counties where the photos were fairly old or where it was thought there could be significant changes from nonforest to forest.

If nonforest plots are not visited, the only permanent record of the location in two strata (10 & 11) is a pin prick. A stake in the ground is the permanent record of the location in strata 1-9. With remeasurement, errors in the transfer of a pin prick from an old photo to a new photo can result in an observation of nonforest to forest change but not a forest to nonforest change because old forest plots have stakes marking their location. Transferring locations from one photo to another is never perfect, especially with 1:40,000 scale photos from different years that are not ortho-corrected. With remeasurement these errors would bias the sample towards forest land.

A second problem observed with the 1986 procedures is an inconsistency in the photo classification of phase two plots. Double sampling for stratification assumes phase two is a random sub-sample of phase one. Ideally, the photo interpreters would first classify all the phase one plots (without knowing which are phase two plots) and then select the phase two plots. A selection system transparent to the interpreters was not implemented because of the difficulties involved. The systematic nature of the sample and the need to pin prick and collect additional information from the photo on phase two plots made it very inefficient to keep the identity of the phase two plots secret. The lack of independence became apparent in the questionable stratum. A plot should be classed questionable if the interpreter cannot accurately make a forest or nonforest determination. If classifications were done without prior knowledge, the expected ratio of phase two plots to phase one plots is 1:17 in every stratum. The observed ratio in Illinois was 1:8.11, about half the expected value in the questionable stratum. A chi-squared test for lack of independence was significant at the 10^{-10} level. It appears that the interpreters used this stratum more frequently on phase two photo plots than on non-phase two photo plots. In training and supervising the interpreters, we stressed the importance of consistency. The problem appears to be a tendency to do a "better" job at interpreting plots that will be sent to the field over those that will not. Interpreters know someone will be visiting the site they are classifying and must look at it just a little bit more closely.

Ensuring complete independence in the selection of ground plots in a remeasurement sample would greatly increase the phase one effort. It would require two people to work on every township mosaic. Much of the work would be a duplication of effort. One person would transfer the phase two plot locations (pin pricks) from the old photo to the new photo and then locate and label the other phase one plots (those that are not phase two plots) on the new photo. The second person would then classify each of these pin pricks. This would involve placing 17 times more pin pricks and labels on each photo. NCFIA uses photos that are borrowed from other agencies; however, if we were to pin prick and mark the photos every 191 acres rather than every 3,250 acres, we would most likely be forced to purchase photography, greatly increasing our costs.

Classified Digital Imagery for Stratification

Computer aided classification of digital imagery can efficiently map large areas independent of the phase two plot selection. Numerous platforms, sensors and spectral bands, pixel sizes, and classification algorithms are available and much research is ongoing. It is beyond the scope of this paper to review the work that has been done in this area as it relates to forest inventory applications. NCFIA has cooperated in and/or supported research efforts in remote sensing with various groups including MN DNR, Univ. of MN, IL Natural History Survey, IN Univ., WI DNR, Rand Corp., EPA, other FIA projects, and other USDA programs. NCFIA has also acquired the equipment and personnel necessary to process digital image data and has been using these capabilities on special projects in Nebraska and South Dakota. TM data have been used in most of these projects. Currently it appears to be the most promising available data source for use by FIA. The combinations of low cost, multi-spectral capabilities, long-term availability, and appropriate pixel size are among the reasons it has been the primary data source for most large-scale classification efforts investigated by FIA. The minimum size of a forest area defined by FIA is 1 acre, thus sensors with a pixel size larger than an acre can easily miss small areas of forest land, especially where forest is not the primary land cover.

The classification of TM data for an entire state is a very large project and should lead to a product that can be used in many applications. For NCFIA to obtain imagery and classify it for the sole purpose of stratification of the phase two plots would probably exceed the total cost of our current phase one procedures. As methods are improved and the cost of remote sensing coverage decreases this will probably change. NCFIA has typically borrowed photos, or in some states new photos were taken by a cooperating state agency with FIA and other uses in mind. Following this approach of using the best available photos for stratification we began looking for available classified imagery for stratification as we planned the 1998 inventories. Aerial photo plot sampling would be the fall back should the data not provide the level of accuracy needed.

METHODS AND DATA

1998 Stratification Using GAP Classified Landsat TM Data

NCFIA obtained two digital maps derived from TM data. These were made by the National GAP Analysis Program (GAP) (Scott *et al.* 1993). More information about GAP can be found at www.gap.uidaho.edu/gap/ and www.epa.gov/mrlc/.

The base data sets in Illinois were obtained over approximately 5 years (1991 to 1995) and classified into 20 land cover categories by the Illinois Natural History Survey. Four of these categories describe woodland and forest land; the rest are nonforested ranging from water, marsh, and grasses to agricultural and urban. The classification loosely followed an Anderson level 2 (Anderson *et al.* 1976) scheme, and the minimum mapping unit for classification was a single 28.5 x 28.5 m pixel. The data sets in Indiana were taken between 1988 and 1994 and classified into 18 land cover categories by Indiana University. Five of these categories describe woodland and forestland; the rest are similar to those for Illinois. The classification scheme followed the UNESCO system (UNESCO 1973) with a minimum mapping unit of 1 ha (2.47 acres). A sample portion of this image is shown in figure 1 with the 18 classes represented by different gray tones.

The classifications for both states were grouped into binary forest/nonforest images (fig. 2). Furthermore, since FIA defines forestland as being at least 1 acre in size, the Illinois forest classes were clumped and sieved to a one-acre minimum unit. The Indiana data were left at a one-ha minimum unit. This difference in minimum mapping units was of concern to us; however, the raw classified Indiana data were not available to us and in the final analysis we did not detect any problems as a result of the 1-ha mapping unit in Indiana. To improve the identification of plots that were likely to be misclassified or straddle a forest/nonforest edge, two new classes were created. Pixels near the boundary of a forest-nonforest edge were identified. Any forest pixel within two pixels of a nonforest pixel was placed in the forest/nonforest class, and any nonforest pixel within two pixels of a forest pixel was placed in the nonforest/forest class. This provided the four classes used for stratification in the estimation process. Figure 3 shows the same area with the final four classes delineated.

1998 Phase Two Field Procedures

Between 1986 and 1998, field procedures for NCFIA plots changed. The 1998 plot design consists of four 1/24 acre fixed area subplots distributed over an acre. Another major change is that plots can now sample more than one land use. A plot that straddles two land uses is considered a partial observation of each. Under this system, area estimation is not a binary procedure; instead observations can fall anywhere in the range 0-1. For example, a plot that lands 60 percent on forest land and 40 percent on nonforest land becomes an observation of .6 in the estimation of forest area.

Aerial photos were used in this inventory, but only to assist in plot location and observation. It was necessary only to obtain photos for the actual phase two plots rather than complete photo coverage. This greatly reduced the number of photos needed and enabled us to purchase these photos rather than borrow them. The first step in field plot work was to transfer every phase two plot location from the old photos to new photos. Plots were examined to determine if a field visit was necessary. Again, plots that were obviously nonforested were not sent out for field check, but any plot that appeared to possibly have trees on it was remeasured by a field crew. This nonforest interpretation from the photo should not be confused with the nonforest classification in the TM data. Because of the change in plot

Figure 1.—*A part of the classified Landsat TM image produced by Indiana State University for GAP. This classification contains the original 18 classes identified by GAP.*

Figure 2.—*The same part of the image with classes 8-15 combined into one class (forest - gray) and the other classes combined into a second class (nonforest - white).*

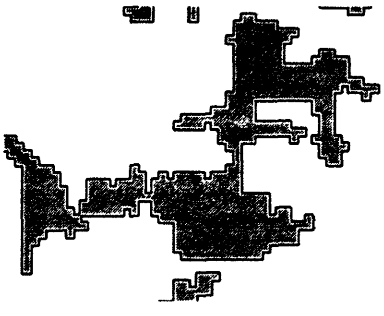


Figure 3.—*The same part of the image with the final 4 strata. Nonforest pixels within 2 pixels of forest (nonforest/forest - black) and forest pixels within 2 pixels of nonforest (forest/nonforest - light gray).*

design, pin pricked plot locations that were obviously nonforest without trees (not visited under the old design) but that fell within 200 feet of a forest area required a field visit.

Linking the phase two plots to the classified Landsat TM data requires accurate location information for every phase two plot. A GPS unit was used to obtain this information on every plot that was field visited. Obtaining location information on nonforest plots without trees that were not field visited involved a digitizing procedure where the plot locations from the aerial photos were transferred to a georeferenced image file. The UTM coordinate obtained through digitizing or the GPS unit was used to link each plot to a specific pixel making it possible to assign each plot to a stratum. Errors in this UTM coordinate as well as errors associated with the georeferenced image file contribute to the overall estimates of sampling error. In the inventory we used the best methods we could to reduce this source of error. In this report we do not attempt to examine the contribution of this single source of error to the overall sampling error of the final estimates.

RESULTS

The described procedures produced estimates with sampling errors at or below the national accuracy standard for FIA inventories. Estimation was done on a county basis; however, in some predominantly nonforest parts of each state, several counties were grouped to create populations containing at least 30,000 acres of forest land. In counties within a National Forest, lands owned by the Forest Service were treated as a population separate from those that were not. This differed from the procedure used in 1986, where estimation was done by treating each forest inventory unit excluding Forest Service lands as populations (there are three forest inventory units in Illinois and four in Indiana) with National Forest lands making a final population in each state. In both inventories, sampling errors were computed for individual populations using two-phase sampling for stratification estimators described in Cochran (1977). In the 1998 estimates, the numbers of phase one plots (pixels) are so large that stratum areas can be considered known without error and the estimates are equal to stratified sampling estimates. State total sampling errors for estimates of forest area, growing-stock volume and growth and sawtimber volume and growth from both the 1986 and 1998 inventories are summarized in table 2, along with the number of phase one and phase two plots.

The differences in sampling errors cannot be entirely attributed to the change in stratification procedures. Changes in plot design, sampling intensity, and population parameters between 1986 and 1999 also have major effects. The new plot design samples trees 5 inches dbh and larger on a fixed area plot (equal probability) rather than a variable radius plot (probability proportional to basal area) so larger trees were sampled with a lower probability in 1998. This can have a major effect on

Table 2.—*Selected sampling errors and number of plots (phase one and two) for the 1986 and 1998 Illinois and Indiana inventories. Sampling errors were computed using double sampling for stratification equations (Cochran 1977).*

	Illinois		Indiana	
Sampling errors	**1986**	**1998**	**1986**	**1998**
Forest area	0.94%	1.49%	1.00%	1.52%
Growing-stock volume	1.99%	2.28%	1.57%	2.18%
Sawtimber volume	2.50%	2.57%	1.86%	2.47%
Growing-stock growth	3.36%	2.09%	3.42%	2.04%
Sawtimber growth	5.27%	2.47%	5.47%	2.39%
Plots	**1986**	**1998**	**1986**	**1998**
Number of phase one plots	194,815	179,674,504	126,629	104,057,965
Number of phase two plots (total)	10,847	11,521	11,440	6,326
Number of phase two (ground visits)	1,342	2,114	2,430	1,847
Percent ground visits	12.37%	18.35%	21.24%	29.20%

the variance of the observed volume per acre measurements in stands with large diameter trees. The change to a plot design with one or more conditions had an impact on the observed variance in area estimates. Also, in 1986, the State of Indiana provided additional resources to NCFIA to increase the intensity of phase two. This additional funding was not available in 1998 and the phase two sampling intensity was substantially reduced.

Figures 4-7 show the average observed mean and standard deviation of forest area and growing-stock volume estimates by stratum for both the 1986 and 1998 inventories. These figures show that the classified TM data provided reasonable stratification. The TM classification was not prefect in the identification of nonforest resulting in nonzero observations of forest land in stratum 4. This stratum (nonforest) is by far the largest stratum in both states, and the impact of a good, but less than perfect, stratification of nonforest lands in phase one has a large impact on the variance of any item that is zero on nonforest land. Under the 1986 system, where most nonforest phase two plots were not field checked and assumed to be a nonforest observation, the true mean and standard deviation within strata 10 and 11 shown in figures 4 and 6 are probably underestimates. The true mean and standard deviation are probably not quite as high as stratum 4 in figures 5 and 7. If, in 1986, as few as 1 in 1,000 nonforest without trees phase one plots were misclassified and were actually forest on the ground, the sampling errors on forest area estimates in 1986 would exceed those of the 1998 shown in table 2. Under stratified sampling, a small error in a very large stratum can have as much impact as a large error within a very small stratum.

The 1998 strata 2 and 3 were designed to contain the plots near forest edges and performed well. Over 60 percent of the ground plots with both a forest and nonforest condition were in one of these strata.

The change from considering each unit a population to considering counties or county groups as populations had significant effects on our county-level estimates. In 1986, all estimates were developed at the unit level and prorated back to the county on the basis of phase one data. This procedure used all phase two ground plots from the entire unit to estimate within stratum means for every county and masked real differences between counties and underestimated the true sampling errors of the county level estimates. The direct development of county-level data for 1998 estimates provides estimates of true differences between

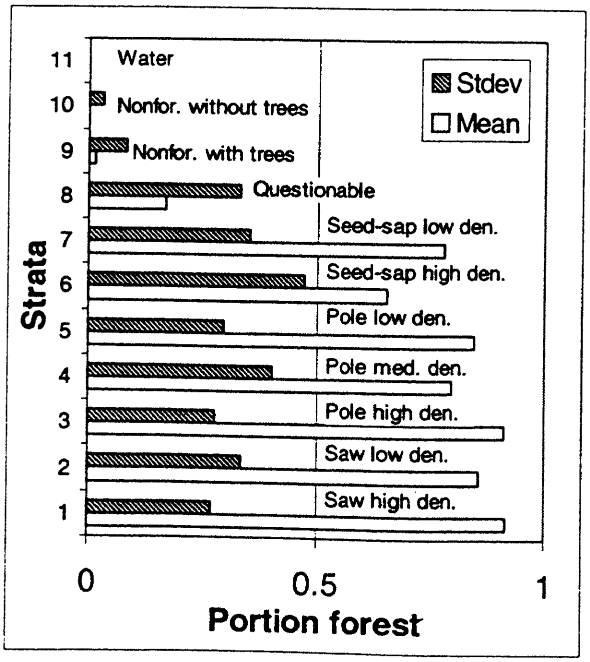


Figure 4.—*1986 forest area data by the 11 aerial photo strata used in 1986.*

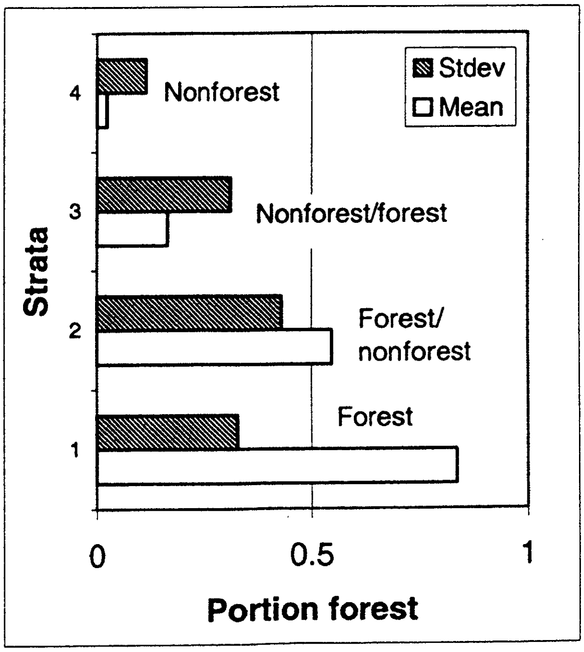


Figure 5.—*1998 forest area data by the 4 Landsat TM strata used in 1998.*

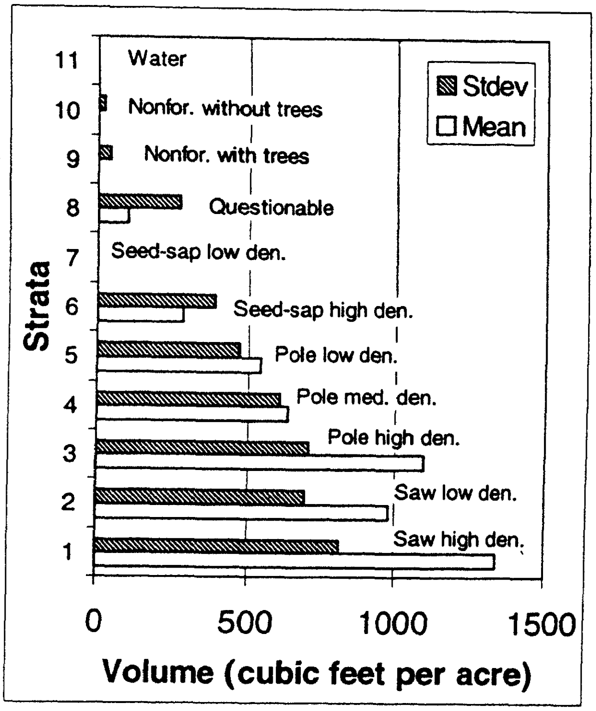


Figure 6.—*1986 growing-stock volume data by the 11 aerial photo strata used in 1986.*

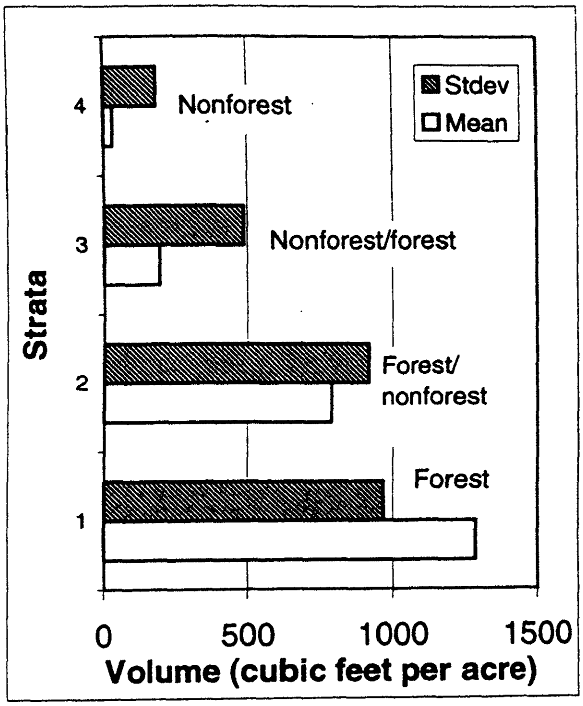


Figure 7.—*1998 growing-stock volume data by the 4 Landsat TM strata used in 1998.*

counties and the sampling errors reported are not based on assumptions that may not be valid. This level of estimation was possible because of the relatively high resolution (number of phase one sample plots per county) of the TM data compared to the photo plot sampling used in 1986.

CONCLUSIONS AND RECOMMENDATIONS

Consistency in classification across large areas is most important in an application such as this. FIA provides estimates at many levels (county, unit, state, regional, and national). A system that does a good job of classifying forest land in one part of the state but not another or classifies one forest type but not another may significantly bias estimates for some users. The approach we used here, to identify only a few strata (four) and produce estimates for relatively small areas (counties or groups of counties), was selected to reduce any inconsistencies in classification that could exist in the data due to scene differences across the state.

One major advantage with this approach to stratification is that it enabled us to conduct a stratification completely blind of the ground plot locations, removing a source of bias that could not be quantified. This can only be done if the image data are referenced and accurate plot location data are available for phase two plots. Errors in referencing and/or plot locations will be an additional source of error and contribute to misclassification. Further, if a classified image is going to be used for stratification, then the ground plot data cannot be used as an aid in the classification, such as training sets in supervised classification.

Although the TM data classified by GAP was a good source of phase one data, we need to keep looking at other sources. Since FIA is an ongoing project, new data for stratification are needed on a periodic basis, especially in areas where the forest landscape is changing. GAP does not currently have the long-term commitment that FIA needs. The classified data being produced by EPA's Landscape Characterization in the Environmental Monitoring and Assessment Program (EMAP) are one possible source of classified data that should be considered. We need to continue to search for cooperators who are interested in land cover classification at or above the state level on a continuous basis. FIA needs to continue to develop its own capabilities to classify remote sensing data. Because it is such a large effort to gather the data and do this kind of classification cooperators will always be needed in these efforts.

As inventories are repeated, the identification of changes in classification will provide additional strata and improve our estimates of change. If strata for forest to nonforest and nonforest to forest change over time can be created, then estimates of land use change over time will benefit. Other strata that help in the estimation of growth, removals, and mortality may also be possible.

LITERATURE CITED

Anderson, J.R.; Hardy, E.; Roach, J.; Witmer, R. 1976. **A land use and land cover classification system for use with remote sensor data**. Prof. Pap. 964. Washington, DC: U.S. Geological Survey. 28 p.

Cochran, W.G. 1977. **Sampling techniques**. New York: John Wiley. 413 p.

Loetsch, F.; Haller, K.E. 1964. **Forest inventory, volume I: statistics of forest inventory and information from aerial photographs**. Munchen: BLV Verlagsgesellschaft, 436 p.

Scott, J.M.; Davis, F.; Csuti, B.; Noss, R.; Buterfield, B.; Groves, C.; Anderson, H.; Caicco, S.; D'Erchia, F.; Edwards, T.C., Jr.; Ulliman, J.; Wright, R.G. 1993. **Gap analysis: a geographic approach to protection of biological diversity**. Wildlife Monograph. 123: 1-41.

UNESCO. 1973. **International classification and mapping of vegetation.** Paris, France: UNESCO. 35 p.

ABOUT THE AUTHORS

Mark H. Hansen is a Research Forester and Daniel G. Wendt is a Supervisory Computer Specialist at North Central Research Station, USDA Forest Service, St. Paul, MN.

FOREST/NON-FOREST STRATIFICATION IN GEORGIA WITH LANDSAT THEMATIC MAPPER DATA

William H. Cooke

ABSTRACT.—Geographically accurate Forest Inventory and Analysis (FIA) data may be useful for training, classification, and accuracy assessment of Landsat Thematic Mapper (TM) data. Minimum expectation for maps derived from Landsat data is accurate discrimination of several land cover classes. Landsat TM costs have decreased dramatically, but acquiring cloud-free scenes at optimum seasons for vegetation discrimination is still problematic. FIA plot locations determined from hand-held GPS units can vary ± 5-20 m. Landsat pixels can also vary ± 25 m. These spatial inaccuracies restrict the use of pixels on feature edges and decrease the usefulness of plots that have split conditions. Current research at the USDA Forest Service's, Southern Research Station involves aggregating forest types in the lab based on field plot measurements of dominant, co-dominant, and intermediate trees. We believe this methodology is most appropriate for tying FIA field plot data to the satellite imagery. We are testing methodological approaches for image processing that can satisfy the dual goals of repeatability and timeliness.

INTRODUCTION

Typically, remote sensing efforts at the Southern Research Station (SRS) of the USDA Forest Service have focused on large area estimates of forested and non-forested lands. Proportions of forested and non-forested lands within pixels of Advanced Very High Resolution Radiometer Data (AVHRR) have been predicted using high resolution Landsat Thematic Mapper (TM) data in a multiple regression scenario (Zhu and Evans 1994). To date, Forest Inventory and Analysis (FIA) managers have asked remote sensing analysts What can you do for me? with respect to rapid large area analysis for simplistic land cover conditions. I believe that our partners and cooperators in the Southern Annual Forest Inventory System (SAFIS) want more than delineation of forested from non-forested lands and attendant acreage calculations. At a bare minimum, we should be able to discriminate among broad land cover classes including pine, hardwood, scrub, grass, cultivated, and inert. As a remote sensing analyst, my question to FIA is What can you do for me? Or, how can plots taken under an annual inventory system be used to train and validate remotely sensed data to produce useful maps? Remote sensing efforts that benefit FIA should extend well beyond Phase I estimates for stratification. Remote sensing also plays an important role as a tool for providing timely information on natural disasters and for getting information about forest conditions in inaccessible areas. These applications of remote sensing should not be overlooked for funding.

BACKGROUND

Achieving the goal of providing the land cover classes of interest requires classification of large amounts of TM data with FIA field plots serving as the basis for classification and verification of those classes. Acquisition of large amounts of TM data is much less costly now that Landsat 7 has been successfully deployed. Full-scene (185 km x 185 km) costs have decreased from $4,000 to $600 per scene. However, acquiring cloud-free scenes at optimum seasons for vegetation discrimination is still problematic. Also problematic are the radiometric differences between adjacent scenes. Figure 1 illustrates the radiometric differences that exist among four full TM scenes in the Piedmont of Georgia.

Investigation by Zhiliang Zhu (U.S. Geological Survey, EROS Data Center, pers. comm.) indicates that normalizing the radiometric components of adjacent TM scenes before

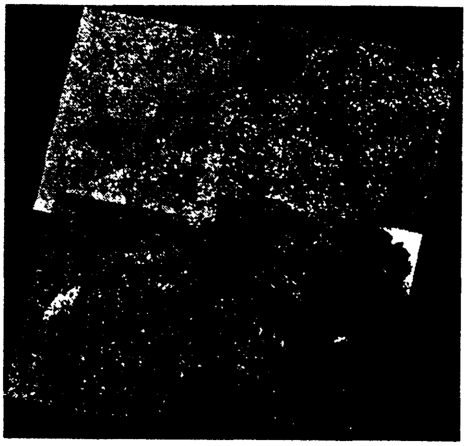


Figure 1.—*Radiometric differences for four winter scenes in Georgia.*

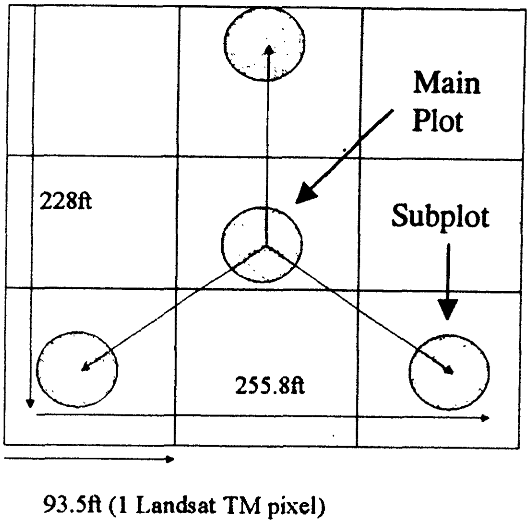


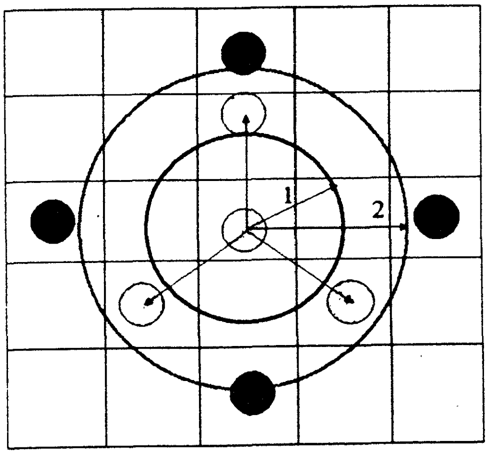
Figure 2.—*Comparison of the FIA plot design with Landsat TM data.*

classification can result in up to a 50 percent loss in reflectance characteristics important for automated classification efforts. At SRS, classification is done before scenes are mosaicked together. Although this approach maximizes classification accuracy on a scene-by-scene basis, it can result in a discontinuity of classification results between adjacent images. Higher per scene accuracy results from this methodology, but a more visually pleasing map product results from pre-classification scene normalization and mosaicking. Ultimately, there is a trade off between higher map accuracy versus more aesthetically pleasing map products. Managers and data users should be educated about this aesthetic problem if maximum information content is the desired outcome.

METHODS

Scientists at SRS are currently examining the usefulness of FIA plot variables for training and for verifying Landsat TM imagery. SAFIS inventories employ Global Positioning System (GPS) receivers to acquire geographic coordinates for the center plot of the four-plot cluster design (Rockwell Avionics 1996). Figure 2 illustrates how this four-plot cluster compares to a nine-pixel window of Landsat TM imagery.

On the surface, the correspondence between the FIA plot design and Landsat TM data appears conveniently located within this nine-pixel window. But misregistration of the imagery and sources of error in the GPS measurement require a model more like the one shown in figure 3. Each subplot falls close to a pixel center, but GPS coordinates are accurate to ± 5 m to ± 20 m depending on averaging techniques, time of acquisition, number of satellites acquired, and overhead "line-of-sight" (Rockwell Avionics 1996). Each TM pixel's registration to its "real-world" location is assumed to be about ± 1 pixel (28.5 m) in relatively flat terrain, possibly greater in steeply dissected terrain. Misregistration errors can be additive in a worst-case scenario (20 m + 28.5 m = 48.5 m). The reality is that the main plot falls somewhere within a 25-pixel window. The full



1 = pixel misregistration
2 = maximum GPS misregistration

Figure 3.—*Pixel and GPS misregistration problems.*

cluster of four plots falls within a 7 x 7 pixel window or greater. This translates into roughly an 11-ac (4.4 ha) ground area. This inherent "slop" in location makes training of automated classifiers and accuracy assessment procedures more difficult. These spatial registration problems will likely restrict the use of pixels on feature edges and limit the potential usefulness of plots that have split land cover conditions. The possibility of deriving an edge class is being investigated by SRS scientists. Classification techniques being used for wall-to-wall TM efforts in Georgia are variations on methods used by Coppin and Bauer (1994) and Cooke (1991).

Classification Techniques

1. Stratify TM scenes by physiographic/ecological condition.
2. Use National Wetlands Inventory data to mask wetlands.
3. Use Census data to mask high-density urban areas.
4. Allow low-density urban areas to be classified.
5. Use differential highway masks.
6. Use edge detection spatial filtering algorithms to locate and eliminate some edge pixels before classification.
7. Classify the data using these TM channels:
    a. Raw data channels 3, 4, and 5
    b. 1st Principal Component
    c. Brightness and Greenness components of the Kauth-Thomas transformation
    d. Ratio of channels 3 and 4 (NDVI).
8. Classify 75 classes using unsupervised techniques to reduce class variance.
9. Aggregate in classes (Pine, HW, Brush, Inert, at a minimum), then iteratively reclassify if necessary.
10. Aggregate classes by following methods developed by Linda Garnett for her Master's Thesis.
11. Use a 5 x 5 majority scan to filter out "salt and pepper" pixels.
12. Assess accuracy/refine classifications for areas > 25 pixels using FIA plots.
13. Assess supervised classifications for accuracy with FIA plots for cross validation.

RECOMMENDATIONS

Work is needed to determine which plot variables are appropriate for training and verifying TM classifications. Current efforts involve aggregating from individual tree data for dominant and co-dominant species. We believe that the satellite "view-from-above" makes these crown classes most likely to provide useful information for modeling land cover. Plot-level variables like forest type are subject to field-level interpretation. A forest type designation calculated in the lab from the dominant and co-dominant members of a stand is more likely to be representative of crown reflectance in TAM data and is easily reproducible in the lab.

LITERATURE CITED

Cooke, W.H. 1991. **Use of temporal Landsat data for discrimination of selected forest species composition group**. Mississippi State University. M.S. thesis.

Coppin, P.R.; Bauer, M.E. 1994. **Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features**. IEEE Transactions on Geoscience and Remote Sensing. 32(4): 918-927.

Rockwell Avionics. 1996. **Precision Lightweight GPS Receiver PLGR+96 and PLGR FEDERAL Precise Positioning Service (PPS)**. Operations and Maintenance Manual.

Zhu, Z.; Evans, D.L. 1994. **U.S. forest types and predicted percent forest cover from AVHRR data**. Photogrammetric Engineering and Remote Sensing. 5: 525-531.

ABOUT THE AUTHOR

William Cooke is a Research Forester, Southern Research Station, USDA Forest Service, Forest Inventory and Analysis program, Starkville, MS.

COMPARISON OF THREE ANNUAL INVENTORY DESIGNS, A PERIODIC DESIGN, AND A MIDCYCLE UPDATE DESIGN

Stanford L. Arner

ABSTRACT.—Three annual inventory designs, a periodic design, and a periodic measurement with midcycle update design are compared using a population created from 14,754 remeasured Forest Inventory and Analysis plots. Two of the annual designs and the midcycle update design allow updating of plots using sampling with partial replacement procedures. Individual year and moving average estimates are determined. The moving average estimates are compared to both the population means of the most recent year used in the average, and to population averages covering the same period as the estimate. Comparisons for net cubic-foot volume per acre and annual change in volume are based on root mean square error (RMSE) and estimator bias. Among annual designs, the rotating panel design produced the smallest RMSE for volume. For multiple year comparisons, the rotating panel and periodic designs resulted in the smallest RMSE's, while for single year comparisons, the periodic design resulted in the smallest RMSE. For annual change, the smallest RMSE's were produced by the periodic design, while among annual designs, the rotating panel design resulted in the smallest RMSE's for multiple year comparisons, and the rotating panel and balanced annual partial remeasurement designs resulted in the smallest RMSE's for single year comparisons.

ABOUT THE AUTHOR

Stanford L. Arner is a Supervisory Forester at the Northeastern Research Station, USDA Forest Service, Newtown Square, PA.

SAFIS AREA ESTIMATION TECHNIQUES

Gregory A. Reams

ABSTRACT.—The Southern Annual Forest Inventory System (SAFIS) is in various stages of implementation in 8 of the 13 southern states served by the Southern Research Station of the USDA Forest Service. Compared to periodic inventories, SAFIS requires more rapid generation of land use and land cover maps. The current photo system for phase one area estimation has changed little over the last four decades and provides area estimates within the precision requirements of the FIA program. A stated goal of the national FIA program is to eventually replace photo interpretation with digital satellite classification because the photo system cannot produce maps of forest and nonforest area, and it takes an enormous amount of time to photo interpret the phase one photo plots. Using automated classification procedures for TM satellite data, we anticipate that the time to complete phase one will decline and wall-to-wall maps will be available. In the interim period of switching to satellite data, the photo system must be modified to provide current estimates of inventory. A method being used by Southern FIA is documented.

INTRODUCTION

Historically, the Southern FIA program has produced forest area estimates using a variation of double sampling. The process consists of interpreting a large number of sample plots on aerial photographs and subsampling a proportion of the plots on the ground. The aerial photo sampling is referred to as phase one and is used to estimate the percent of the total area occurring in forest and nonforest subpopulations. The phase two samples are the FIA ground plots that provide the basic mensurational data used to further stratify forest area and estimate timber volume, growth, mortality, and removals.

The Southern Annual Forest Inventory System (SAFIS) requires rapid generation of land use and land cover maps for the southern United States. The current photo system for phase one area estimation has changed little over the last four decades and provides area estimates within the precision requirements of the FIA program. The photo method does have two shortcomings for the annual inventory program. First, although the photo method can provide estimates of forestland down to the county level, the method cannot produce maps of forest and nonforest area and distribution. Second, it takes a considerable amount of time (up to a year per state) to photo interpret an entire state. A stated goal of SAFIS and the national FIA program is to eventually replace photo interpretation with digital satellite classification to address the two shortcomings of the photo system. In the interim, the following photo-based procedures are being used to estimate forest area under the annual inventory system in the Southern FIA program.

USING DOUBLE SAMPLING FOR AREA ESTIMATION

Currently the Southern FIA is using a double sample to estimate forest area. Frayer and Furnival (1999) provide a chronology of how double sampling for stratification has been implemented nationwide by FIA. Cochran (1977) and De Vries (1986) provide statistical references for estimation when employing a double sample. Estimates of timberland area are based on forest and nonforest interpretation of a large number of plots on aerial photos and a smaller sample of ground plots. There is approximately 1 photo plot per 230 acres across the South. The current definition of forest in FIA is land 1 acre in size, 120 feet wide, and at least 10 percent stocked by forest trees of any size, or formerly having had such tree cover and not currently developed for nonforest use.

The photo interpretation points are arranged in a 5 x 5 grid where one of the photo plots is spatially coincident with a phase two ground plot. The current phase two ground plots are conceptually distributed on a 3 x 3 mile grid. In addition to the double sample that occurs with each FIA ground plot, there are intensification plots arranged on a 3 x 6 mile grid. Intensification plots are used to increase sample size for forest and nonforest ground truth type calls. An intensification plot is simply a spatially coincident photo plot and field truth location. This results in 173.6 ground plots and 86.8 intensification plots (used to correct area estimates) per million acres. Combined, there are 260.4 (173.6 + 86.8) ground reference samples per million acres and 4,347.8 photo interpretation points per million acres. Thus, an approximate 6 percent field sample of the photo plots is used to form the double sample estimate of forest area.

A quick review of how forest area is determined using a periodic survey will naturally lead into modifications necessary to operate under the new annual panel system. For detailed reference information on the new annual panel system, see Reams and Van Deusen (1999) and Roesch and Reams (1999). To estimate forest area, several types of information are needed. First, the total area estimate and census water estimate by county are obtained from the U.S. Census Bureau. Census land is computed as census total area minus census water. Forest area ($\hat{A}_f$) is then computed as:

$$\hat{A}_f = \hat{P}'_f \times A_t \qquad (1)$$

where

$$\hat{P}'_f = (\hat{P}_f \ x \ C_f) + (\hat{P}_n \ x \ C_n)$$

and, A_t is census land area,

$\hat{P}_f$ is the proportion of phase one photo plots in forest, P_n is the proportion of phase one photo plots in nonforest, and

$$C_f = \frac{\textit{number of plots correctly photo interpreted forest}}{\textit{total number of plots photo interpreted forest}}$$

and,

$$C_n = \frac{\textit{no. of plots photo interpreted as nonforest but are forest}}{\textit{total number of plots photo interpreted as nonforest}}$$

Assuming the following confusion matrix (table 1),

Table 1.—*Two-way contingency table where the diagonal elements represent correct classifications as compared to ground truth, and the off-diagonal elements represent plots with incorrect photo classification*

Grounds plots	Photo interpreted forest	Photo interpreted nonforest	Total
Forest	108	2	110
Nonforest	3	81	84
Total	111	83	194

the proportion of forest area ($\hat{P}'_f$) is then estimated as,

$$\hat{P}'_f = \frac{(\textit{no. points forest} \times C_f) + (\textit{no. points nonforest} \times C_n)}{\textit{total no. points photo interpreted}}$$

Assuming that 1,962 photo points were interpreted as forest and 1,288 photo points interpreted as nonforest for a total of 3,250 interpreted points, the proportion of forest area would be estimated as,

$$\hat{P}'_f = \frac{(1962 \times .973) + (1288 \times .024)}{3250} = .5969$$

Proportion of nonforest is simply $1 - \hat{P}'_f$.

Forest area (equation 1) is thus equal to 0.5969 x A_t. The variance of forest area is determined as follows,

$$\sigma^2 (\hat{P}'_f) = \frac{(\hat{P}_f)(\hat{P}_n)}{n} (C_f - C_n)^2 + \left[\frac{(\hat{P}_f)^2 (C_f)(1 - C_f)}{m_1} \right] + \left[\frac{(\hat{P}_n)^2 (C_n)(1 - C_n)}{m_2} \right]$$

where, $\hat{P}_f$ = proportion forest, $\hat{P}_n$ = proportion nonforest, C_f = proportion correctly classified forest, C_n = proportion incorrectly classified nonforest, m_1 = number of forest field locations, m_2 = number of nonforest field locations, and n = number of photo points interpreted. Inserting values for each variable results in,

$$\sigma^2(\hat{P}_f') = \frac{(.6037)(.3963)}{3250}(.973 - .024)^2 + \left[\frac{(.6037)^2(.973)(1-.973)}{110}\right] + \left[\frac{(.3963)^2(.0241)(1-.0241)}{84}\right]$$

Completing the above mathematical operations, $\sigma^2(\hat{P}_f') = 0.00019729$ and the s.e. of the estimate is 0.014.

MODIFICATION OF CORRECTION FACTORS FOR ANNUAL SYSTEM

Because the Southern FIA program is dependent on the National Aerial Photography Program (NAPP) for acquisition of 1:40,000 scale photography that is flown on an approximate 5-year cycle, new photography will not be available on an annual basis. For the Southern annual survey, only one-fifth of the present number of phase two ground plots are remeasured in any one year. If relying solely on the FIA ground plots for the double sampling estimate of forest area, this results in approximately 35 ground truths per million acres. The Southern FIA program has additional ground truths arranged on a 3 x 6 mile grid, and if we assume one-fifth of these plots are visited every year, this equates to approximately 17 ground truths per million acres. Thus, in any given year, 52 ground truths per million acres are available for forest area estimation. Because new photography is available on about a 5-year cycle, all photo plots will be interpreted as available, and will be completed in one calendar year. There will be 52 ground truths per million acres available per year, and this will be too small a sample size for reliable estimation at the county level because the average size of counties in most states in the South ranges from 200,000 to 350,000 acres. The reason for stating that the sample size for a panel (year) is too small for reliable estimation at the county level, is because the off-diagonal elements of the misclassification matrix will often be zero.

One remedy for this situation is to estimate forest area at the multi-county level based on 100 percent measurement of phase one photo interpretations but only 20 percent of the phase two ground plots (truths). Counties must be aggregated such that the misclassification matrix for the multi-county area has at least 200 ground truths per year. This number of ground truths usually ensures that the off-diagonals are not zero.

If one were to develop a correction factor on one panel (20 percent of the ground plots) of data, a reasonable recommendation to group counties either by immediate adjacency, percent forest, or percent change in forest area. Because misclassification errors are more often related to percent forest area, this option is being developed and is favored by Southern Station analysts. Using western (survey unit 1) Tennessee as an example, aggregation of counties totaling at least 2.5 million acres is recommended (fig. 1). The resulting misclassification matrix and correction factors are then applied

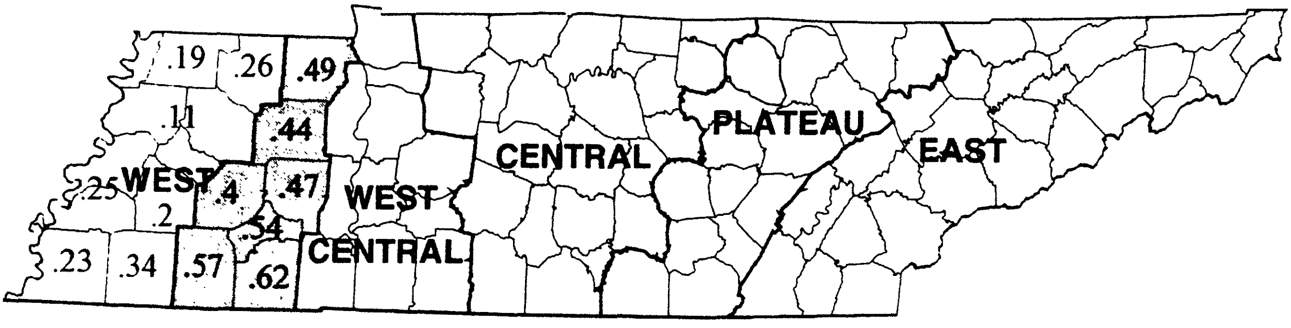


Figure 1.—*Example of how to group counties in survey unit 1, Tennessee for development of correction factors based on one panel of data. Shaded counties are aggregated for a common correction factor.*

individually to each county. Thus, a set of counties will have the same set of correction factors applied to each county's unique estimate of percent forest area from the phase one sample. In a recent application, counties were aggregated by survey unit and percent forest area class with class 1 (0-35 percent), class 2 (36-55 percent), class 3 (56-65 percent), class 4 (66-75 percent), and class 5 (76-100 percent). Forest area by county is determined then as (percent forest land x census land) as given previously in equation (1).

If estimating area based on multiple panels (years) of ground truths, there is the need to either assume that previous panel ground truths are correct, or revisit those plots or a subset of the previous panel of plots, since the forest/nonforest classification may have changed. Because of cost factors, the assumption made is that previous year panels or a subset of the previous year panel plots will not be visited. Assuming the same amount of field effort every year, the number of counties or the land area that must be aggregated each year for calculation of correction factors is cut in half in year two, a third in year three, and so on. By the time all five panels have been measured, we are back at the intensity of ground truths analysts are accustomed to under the periodic system. As many analysts well know, even with 100 percent (all five panels) measurement of ground plots, small counties are often combined because of small sample sizes. The need to combine counties for correction factors is thus a one-time adjustment with implementation of the five panel system.

USING SATELLITE DATA FOR AREA ESTIMATION

As stated earlier, the goal of SAFIS and the national FIA program is to eventually replace photo interpretation with digital satellite classification to address the two shortcomings of the photo system. An area estimation method that replaces photo interpretation with digital satellite data follows.

The three most commonly used satellite sensors are the Landsat Thematic Mapper (TM), the French Systeme Probatoire pour l'Observation de la Terre (SPOT), and Advanced Very High Resolution Radiometer (AVHRR). We have concentrated on the use of TM classifications because TM has greater spectral resolution relative to SPOT and better spectral and spatial resolution compared to AVHRR (Wynne *et al.* 2000).

To estimate map class area totals and variances, we use two-phase or double sampling, where the less accurate data are the map whose accuracy is in question and the more accurate but costly data are the FIA ground plot. The less accurate data are complete in that each map pixel has been classified.

A sampling scheme designed to evaluate and correct for map area misclassification is as follows: A sample of *n* points/pixels is located on the map, and the true and map categories are determined for each point. The *n* points are allocated as a simple random sample. This results in a two-way contingency table where n_{ij} is the number of points in the sample whose true category is *i* and whose map category is *j*.

There is an important difference between using satellite-derived maps and aerial photos for this process. The satellite-derived thematic map allows us to know the actual map marginal probabilities, which can be used as additional constraints in a maximum likelihood estimation process (Card 1982, Van Deusen 1996, Reams and Van Deusen 1999). This reduces the variance of estimates of true map category proportions. Formulas for estimating the true probabilities of interest are given in Card (1982), along with variance estimates. Methods for estimating change in category proportions and variances between two times are given in Van Deusen (1994). The estimators for the true map proportions are the same for simple random sampling or stratified sampling of map pixels. However, variance estimates are different under the two sampling strategies.

In a pilot study in central Georgia, a winter cloud-free TM scene was classified and statistical estimates of forest and nonforest area were derived using the methods of Card (1982). Estimates of percent forest and nonforest from the classified TM scene compare quite favorably to FIA survey cycle six estimates of forest area based on the photo method. The FIA cycle six survey for central Georgia indicates that 68.7 percent of survey unit 3 is in forest (Thompson 1989), and the estimate based on TM data indicates that 69.4 percent of the scene is forest (Reams and McCollum, in review). Variance estimates are comparable to those derived from the photo-based method that uses standard double sampling for stratification statistical estimates. With the wall-to-wall TM classification, it is important to use the known map marginals to reduce variance of the final estimates (Card 1982, Reams and Van Deusen 1999).

LITERATURE CITED

Card, D.H. 1982. **Using known map category marginal frequencies to improve estimates of thematic map accuracy**. Photogrammetric Engineering and Remote Sensing. 48(3): 431-439.

Cochran, W.G. 1977. **Sampling techniques**. New York: Wiley. 413 p.

De Vries, P.G. 1986. **Sampling theory for forest inventory**, New York: Springer-Verlag. 399 p.

Frayer, W.E.; Furnival, G.M. 1999. **Forest survey sampling designs: a history**. Journal of Forestry. 97(12): 4-10.

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.

Reams, G.A.; McCollum, J.M. (In review). **Estimates of forest area using thematic map data**. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

Roesch, F.A.; Reams, G.A. 1999. **Analytical alternatives for an annual inventory system**. Journal of Forestry. 97(12): 33-37.

Thompson, M.T. 1989. **Forest statistics for Central Georgia, 1989**. Resour. Bull. SE-105. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 51 p.

Van Deusen, P.C. 1994. **Correcting bias in change estimates from thematic maps**. Remote Sensing of the Environment. 50: 67-73.

Van Deusen, P.C. 1996. **Unbiased estimates of class proportions from thematic maps**. Photogrammetric Engineering and Remote Sensing. 62(4): 409-412.

Wynne, R.H.; Oderwald, R.G.; Reams, G.A.; Scrivani, J.A. 2000. **Optical remote sensing for forest area estimation**. Journal of Forestry. 98(5): 31-36.

ABOUT THE AUTHOR

Gregory A. Reams is a Supervisory Mathematical Statistician, Southern Research Station, Forest Inventory and Analysis Program, USDA Forest Service.

The author thanks Bill Bechtold, Ray Sheffield, and Mike Thompson for review and suggestions that improved the manuscript.

DIAMETER GROWTH MODELS USING FIA DATA FROM THE NORTHEASTERN, SOUTHERN, AND NORTH CENTRAL RESEARCH STATIONS

Veronica C. Lessard, Ronald E. McRoberts, and Margaret R. Holdaway

ABSTRACT.—Nonlinear, individual-tree, distance-independent annual diameter growth models are presented for species in two ecoregions defined by R.G. Bailey in the northern Lake States and in parts of the central and southern regions of the U.S. The models were calibrated using Forest Inventory and Analysis (FIA) data from undisturbed plots on land classified as timberland across all ownership categories. The data were generally from stands of mixed age and mixed species. The dependent variable is average annual diameter growth and the independent variables include crown ratio, crown class, stand basal area, stand basal area larger than the subject tree, physiographic class, and latitude and longitude of plot locations. The models have minimal bias and may be recalibrated easily to include new data sets.

INTRODUCTION

In response to the 1998 Farm Bill, formally known as the Agricultural Research, Extension, and Education Reform Act, the North Central Research Station (NCRS) of the USDA Forest Service, has developed an annual inventory system featuring a hexagonal grid system of FIA plots to be measured in 5-year inventory cycles, with 20 percent of the plots to be measured each year (Brand *et al.* 2000). Diameter growth models for individual trees provide a method to update the information on FIA plots not measured in the current year.

A study was conducted to develop individual-tree, distance-independent, diameter growth models using FIA data collected in previous inventories. The criteria for the models, calibrated for major species groups within two ecoregions of the central part of the United States, were that they would produce minimal bias in their estimates and be of a form that can be recalibrated easily with inventory data collected under the annual system.

ECOREGION PROVINCES

The diameter growth models were calibrated for two ecoregions: the 212-Laurentian Mixed Forest Province and the 222-Eastern Broadleaf Forest (Continental) Province defined by Bailey (1995) (fig. 1). The western contiguous portion of the Laurentian Mixed Forest Province on the U.S. side of the U.S./Canadian border falls within the northern half of the Lake States. Province 212 is a subdivision of the Warm Continental Division and is characterized by snowy, cold winters and warm summers. Most precipitation occurs in the summer but is plentiful throughout the year. Province 212 is a transition zone between the boreal forest and the broadleaf deciduous forest zones, and its habitant species include members of both zones.

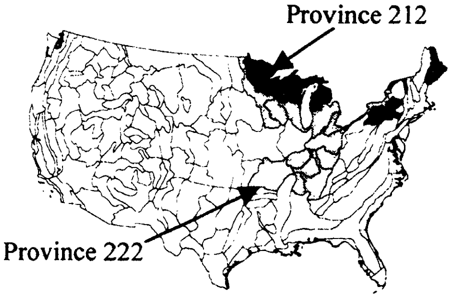


Figure 1.—*Ecoregions of North America (Bailey* et al. *1994).*

The Eastern Broadleaf Forest Province is a subdivision of the Hot Continental Division. Most precipitation in Province 222 occurs during the growing season and generally decreases in quantity and adequacy as distance from the Atlantic Ocean increases. This province favors drought-resistant oak-hickory associations. Province 222 lies to the east of the prairie regions, south and west of Province 212 in the northern areas, and west of the Appalachian Mountains in the southern regions. It extends from the Minnesota/Canadian border in the north to Missouri and Tennessee in the south.

DATA

The diameter growth models were calibrated using FIA data from undisturbed plots on land classified as timberland over all ownership categories. Timberland is defined as non-reserved forestland that is producing or is capable of producing 20 cubic feet of industrial wood per year. The FIA survey design and data collection are described by Hansen *et al.* (1992). ArcView GIS was used to overlay Bailey's ecoregion map on the FIA plot locations to select plots within each of the provinces. Growth models for Province 212 were calibrated using FIA data from the following states (the numbers in parentheses refer to the year of the inventory): Michigan (1966, 1980, 1993), Wisconsin (1968, 1983, 1996), and Minnesota (1977, 1990, 1993). Models for Province 222 were calibrated using data from parts of Michigan (1966, 1980, 1993), Wisconsin (1968, 1983, 1996), Minnesota (1977, 1990, 1993), Illinois (1962, 1985, 1998), Indiana (1967, 1986, 1998), Iowa (1974, 1990), Ohio (1978, 1990), Missouri (1972, 1989), Kentucky (1974, 1987), and Tennessee (1989, 1996).

Trees included in the calculation of total plot basal area per acre (BA) and total plot basal area per acre greater than the subject tree (BAL) for time 1 were restricted to all living trees recorded on the plot in the first inventory. Trees included in the calculation of BA and BAL at time 2 were restricted to the remeasured trees that were alive and included in the plot at time 1 and still alive at time 2. Plot variables associated with individual trees were maintained with each tree record. The data were sorted by species and the diameter growth models were fit to the data. Only remeasured trees that were alive during both inventories were used to calibrate the individual-tree diameter growth models. For states with data from three inventories, data were extracted separately for growth periods from the first to the second inventory and from the second to the third inventory, and the two growth periods were treated as separate observations.

The data were split into two databases used to calibrate and validate the models. Every fourth plot was systematically assigned to the validation database. The remaining 75 percent of the plots were used for calibration of the models. Associated plot information was retained with each tree record, and tree records were sorted into species groups.

Diameter at breast height (DBH) is used as the major predictor variable in the growth model for predicting change in DBH. The average observed change in DBH, calculated as the ratio of the difference in DBH measurements at the two inventories and the number of years between inventories, is the dependent variable. Both individual tree and plot variables are considered for use as growth predictors. Individual tree variables include initital crown ratio (CR) and initial crown class (CC). Crown ratio is the percentage of total tree height that is crown and is assigned in FIA data to one of nine categories, where each of the category values 1-8 represents 10 percent interval widths and the final category represents 81-100 percent. Crown class is recorded in FIA data in five categories ranging from nonsuppressed to suppressed.

Variables related to the plot information include BA, BAL, physiographic class (PC), latitude (LAT), and longitude (LNG). Information about competition within the stand is given by BAL and BA, while physiographic class gives information related to site soil and water conditions that affect site productivity. Latitude and longitude are surrogates for climatic conditions in the regional models.

MODELING METHODOLOGY

Mathematical Form of the Diameter Growth Model

The form of the diameter growth models is the product of two components, an average DBH growth model and a modifier. The average model is based on a two-parameter gamma probability density function using DBH as the independent variable to predict diameter growth rates. The modifier is a product of exponentials, each of which incorporates a single additional independent variable. The modifier gives greater accuracy and precision to the growth model by allowing the predicted growth values to increase or decrease from those given by the average model. The form of the diameter growth model is

$$\Delta DBH = AVE(DBH) * MOD(X_1, \ldots, X_7) \quad , \quad (1a)$$

where

$$AVG(DBH) = \beta_1 \exp(-\beta_2 DBH)\, DBH^{\beta_3}, \quad (1b)$$

and

$$MOD(X_1, \ldots, X_7) = \prod_{i=1}^{7} e^{f_i(X_i)} \quad . \quad (1c)$$

The X_is and the f_is are defined in table 1. The functions incorporate the difference of the observed and average ecosystem values of the variables. The greater the deviation of the observed variable's value from the average ecosystem value, the greater the impact that variable has on the model prediction. The values, -90 and 46, in the functions 6 and 7 are the average values of longitude and latitude in Province 212 data. The average longitude and latitude values are -88 and 40, respectively, for Province 222.

The parameterization of the gamma probability density function used in the average component (1b) of the model is a simplified form of that given by Johnson and Kotz (1970). In the formulation, β_2 serves as the scale parameter that defines the spread of the distribution, and the shape parameter, β_3, establishes the peakedness of the curve. The parameter, β_1, is a multiplier to better adjust the model fit to the data either upward or downward in conjunction with the rates of average annual growth. The parameter estimation routine failed to converge for some species. For those species, either β_2 or β_3 was set to 0 in separate calibrations of the model, thus changing the gamma function to a power function or an exponential function, respectively. Of the two functions, the power function generally provided the better fit, based on the mean square error (MSE) of the resulting fits of the models to the data.

In addition to the independent variables used in the diameter growth models (DBH, CR, CC, BA, BAL, PC, LAT, and LNG), a number of other independent variables were explored for inclusion in the diameter growth models by Holdaway (2000). These included the number of trees per acre, average stand diameter, and the ratio of DBH to average stand diameter. Our goal was to build diameter growth models that would explain the most variability while using the smallest number of variables (and associated parameters) to achieve a parsimonious model. No variable was included in the model if the asymptotic 95 percent confidence interval for the parameter estimate associated with the variable included zero. Additionally, a comparison criterion, given by Linhart and Zucchini (1986), accounting for both an estimate of the MSE and the number of estimated parameters was used to determine the variables for inclusion in the models.

Weighted Regression

Heterogeneity in the variation of the residuals was adequately addressed using linear regression to find a functional form relating standard deviation of the residuals to predicted growth as (McRoberts *et al.* 2000):

$$E[\ln(\hat{\sigma})] = \alpha_1 + \alpha_2 \ln(\Delta \widehat{DBH}) \quad , \quad (2)$$

where E(.) represents the statistical expectation, $\Delta \widehat{DBH}$ is the average predicted annual diameter growth for predicted diameter growth classes, $\hat{\sigma}$ is the standard deviation of the

Table 1.—*The functional forms for variables used in the modifier (the exponential of the sum of these functions) are presented*

Variables	i	Functional form $f_i(X_i)$
Crown ratio	1	$\beta_4 * (CR - 4)$
Plot basal area larger than the subject tree	2	$\exp(\beta_5 * (BAL - 50)) - 1$
Plot basal area	3	$\beta_6 * (BA - 100)$
Crown class	4	$\beta_7 * (\exp(CC/3) - 2.718)$
Physiographic class	5	$\beta_8 * (PC - 5) + \beta_9 * (PC - 5)^2$
Longitude	6	$\beta_{10} * (LNG + 90)$
Latitude	7	$\beta_{11} * (LAT - 46)$

residuals for predicted diameter growth classes, and the αs are parameters to be estimated.

Bias Assessments

The fit of the models was verified using the calibration data set, while the data from the 25 percent of all plots initially set aside were used to validate the models. Verification and validation were carried out by applying the models to the data. Predicted average annnual growth was compared with actual average annual growth over the interval between measurements. The resulting residuals (observed minus predicted average annual growth for the interval between measurements) from the calibration and validation data were analyzed separately.

Annualized residual medians, median ratios, standard deviations, and r^2 values were calculated to examine the prediction bias of the models. Relative bias presents the median bias as a percent of the median observed growth and gives perspective to the importance of bias from a biological viewpoint. Relative bias was calculated as 100 percent times the ratio of the median annualized growth residual to the median annualized observed growth. The r^2 values are calculated as the square of the coefficient of correlation between the annualized predicted and observed values of diameter growth.

RESULTS

Parameter Estimates

Model parameter estimates for 10 of the most frequently occurring tree species in the calibration data sets for Province 212 and Province 222 are given in tables 2 and 3, respectively. Variables in the modifier component of the model not found to be important for a particular species were assigned parameter values of 0, which is analogous to multiplying the model by 1.

Model Verification and Validation

The results of the analysis of the residuals, calculated as the difference of the observed and predicted annual diameter growth values for the observations in the calibration and validation data sets, are given in tables 4 and 5, respectively. Negative values of the median residuals and relative bias indicate overestimation of the models.

CONCLUSIONS

For most species, observed average annual change in DBH ranges from 0.05 to 0.15 inches per year, and DBH is measured to the nearest 0.1 inch. The values of the median residuals for the models applied to the calibration and validation data for Province 212 and Province 222 show the models are relatively unbiased, with median overestimation generally near or less than 0.01 inch per year. The median bias

Table 2.—*Parameter estimates for the annual diameter growth models (1a-c) are given for the 10 most frequently occurring species groups in the 212-Laurentian Mixed Forest Province. Variables to which the parameters, $\hat{\beta}_1$-$\hat{\beta}_{11}$, correspond are given in the column headings.*

Species	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$	$\hat{\beta}_{11}$
	DBH	*DBH*	*DBH*	*CR*	*BAL*	*BA*	*CC*	*PC*	*PC²*	*LNG*	*LAT*
Softwoods											
Black spruce	0.0432	0.0321	0.2194	0.1952	0	0	-0.0447	-0.1514	0	0	0
Balsam fir	0.0497	0.0829	0.6517	0.1402	0	-0.0012	-0.0589	-0.1125	0	-0.0354	-0.1074
Tamarack	0.0333	0	0.4205	0.1570	0	0	-0.1069	-0.1262	0	0	0
N. white-cedar	0.0326	0.0202	0.4353	0.1575	0	-0.0005	-0.0667	-0.0895	0	-0.0113	-0.0586
Hardwoods											
N. red oak	0.0724	0	0.2237	0.0349	-0.0029	0	-0.1447	0	0	-0.0246	-0.1090
Hard maple	0.0293	0.0566	0.8377	0.0959	0	-0.0010	-0.1955	0	0	-0.0215	-0.1595
Soft maple	0.0509	0.0167	0.4651	0.1281	0	-0.0014	-0.1274	0	-0.0307	-0.0544	-0.1363
Black ash	0.0535	0	0.2278	0.1323	0	-0.0006	-0.1042	-0.0727	0	-0.0295	-0.1598
Quaking aspen	0.0902	0.0333	0.3930	0.1144	-0.0023	0	-0.1084	0	-0.0280	-0.0261	-0.0940
Paper birch	0.0639	0	0.1284	0.1330	-0.0021	0	0	0	-0.0616	-0.0565	-0.2244

Table 3.—*Parameter estimates for the annual diameter growth models (1a-c) are given for the 10 most frequently occurring species groups in the 222-Eastern Broadleaf Forest (Continental) Province. Variables to which the parameters, $\hat{\beta}_1$-$\hat{\beta}_{11}$, correspond are given in the column headings.*

Species	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_{10}$	$\hat{\beta}_{11}$
	DBH	*DBH*	*DBH*	*CR*	*BAL*	*CC*	*PC*	*LNG*	*LAT*
Hardwoods									
Select white oak	0.0658	0.0105	0.3150	0.0372	-0.0031	-0.1959	0.1441	0.0264	-0.0200
Other white oak	0.0813	0	0.1103	0.0575	-0.0025	-0.2107	0.1509	0.0823	0
Select red oak	0.0910	0.0163	0.3374	0.0422	-0.0022	-0.1773	0.1458	0.0242	-0.0087
Other red oak	0.0871	0.0290	0.3832	0.0685	-0.0036	-0.1378	0.1145	0.0485	-0.0082
Select hickory	0.0630	0	0.1656	0.0523	-0.0018	-0.2359	0	0.0197	0
Other hickory	0.0522	0.0253	0.4692	0.0809	-0.0016	-0.2380	0.3244	0.0246	0.0683
Hard maple	0.0643	0.0547	0.5735	0.1021	-0.0035	-0.1963	0.1204	0.0157	0
Soft maple	0.0841	0.0313	0.5007	0.0793	-0.0026	-0.1842	0.1006	0	-0.0365
White and green ash	0.0621	0.0340	0.5834	0.0972	-0.0013	-0.1673	0	0.0282	0
Elm	0.0428	0.0535	0.8250	0.1166	-0.0015	-0.0541	0.1033	0.0522	0.0829

Table 4.—*Residual analysis of the diameter growth models fit to FIA calibration (Cal) and validation (Val) data from the 212-Laurentian Mixed Forest Province*

Species	**Observations**		**Median residual**		**Standard deviation**		**Relative bias**		**r^2**	
	Cal	**Val**	**Cal**	**Val**	**Cal**	**Val**	**Cal**	**Val**	**Cal**	**Val**
	Number		*In/yr*		*In/yr*		*Percent*			
Softwoods										
Black spruce	5,921	2,184	-0.006	-0.005	0.036	0.034	-13.2	-13.3	0.297	0.266
Balsam fir	5,916	1,974	-0.008	-0.005	0.051	0.051	-11.0	-6.8	0.383	0.413
Tamarack	2,572	986	-0.011	-0.009	0.050	0.045	-20.0	-15.5	0.294	0.287
Northern white-cedar	7,496	2,497	-0.006	-0.005	0.040	0.041	-10.0	-8.9	0.280	0.288
Hardwoods										
Northern red oak	2,900	991	-0.009	-0.010	0.060	0.055	-7.0	-7.8	0.299	0.388
Hard maple	7,214	2,551	-0.006	-0.005	0.054	0.056	-7.2	-6.0	0.354	0.339
Soft maple	6,487	2,075	-0.008	-0.006	0.061	0.065	-7.3	-6.3	0.328	0.303
Black ash	3,591	1,311	-0.006	-0.003	0.041	0.040	-10.5	-5.7	0.207	0.202
Quaking aspen	12,264	4,033	-0.008	-0.007	0.066	0.066	-5.7	-5.0	0.239	0.228
Paper birch	5,981	1,763	-0.007	-0.006	0.046	0.043	-9.8	-9.3	0.173	0.185

is generally within 15 percent of the median observed growth rate, as given by the relative bias (tables 4 and 5). Some larger percentages are due to the small median observed growth rates in the denominators of the relative bias ratios for slower growing species.

The low values of the squared coefficients of correlation, r^2, indicate that much unexplained variation remains. This is not surprising when one considers that the data, used in both the calibration and validation of the models, were generally collected on uneven-aged, mixed-species stands that range over an entire ecosystem. Future work on these models will investigate incorporation of climatic data such as temperature and precipitation. This information should contribute to better fitting models by explaining more of the uncertainty.

ACKNOWLEDGMENTS

We want to thank Doug Griffith of the Northeastern Research Station and Joseph Glover of the Southern Research Station for kindly providing FIA data from Kentucky and Ohio and from Tennessee, respectively. We would also like to thank Mark Nelson and Dan Wendt for their suggestions and assistance in applying Arc-Info to this research.

Table 5.—*Residual analysis of the diameter growth models fit to FIA calibration (Cal) and validation (Val) data from the 222-Eastern Broadleaf Forest (Continental) Province*

Species	Observations		Median residual		Standard deviation		Relative bias		r^2	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
	Number		*In/yr*		*In/yr*		*Percent*			
Hardwoods										
Select white oak	5,726	1,784	-0.007	-0.010	0.061	0.061	-5.7	-8.4	0.326	0.349
Other white oak	1,708	475	-0.005	-0.003	0.052	0.051	-6.3	-3.6	0.291	0.315
Select red oak	2,185	717	-0.013	-0.010	0.082	0.086	-8.7	-6.0	0.215	0.223
Other red oak	1,474	1,628	-0.009	-0.015	0.073	0.074	-6.4	-10.7	0.239	0.258
Select hickory	1,400	396	-0.008	-0.004	0.049	0.058	-10.0	-4.9	0.238	0.266
Other hickory	1,616	535	-0.006	-0.008	0.053	0.054	-7.8	-9.4	0.358	0.362
Hard maple	2,317	764	-0.007	-0.007	0.069	0.067	-8.8	-7.6	0.365	0.333
Soft maple	1,602	466	-0.013	-0.018	0.103	0.110	-8.8	-13.8	0.304	0.260
White and green ash	1,685	451	-0.010	-0.008	0.087	0.082	-6.8	-5.6	0.270	0.289
Elm	1,256	401	-0.011	-0.010	0.088	0.098	-12.2	-10.0	0.372	0.360

LITERATURE CITED

Bailey, R.G. 1995. **Description of the ecoregions of the United States**. 2d ed. Revised and expanded. Misc. Publ. 1391. Washington, DC: U.S. Department of Agriculture, Forest Service. With separate map at 1:7,500,000.

Bailey, R.G.; Avers, P.E.; King, T.; McNabb, W.H., eds. 1994. **Ecoregions and subregions of the United States (map)**. Washington, DC: U.S. Geological Survey. Scale 1:7,500,000, colored. Prepared for the USDA Forest Service.

Brand, G.J.; Nelson, M.D.; Wendt, D.G.; Nimerfro, K.K. 2000. **Technical aspects of the hexagon/panel system**. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C., eds. Proceedings of the 1st annual FIA symposium; 1999 November 2-3; San Antonio, TX. Gen. Tech. Rep. NC-213. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 8-13.

Hansen, M.H.; Frieswuk, T.; Glover, J.F.; Kelly, J.F. 1992. **The Eastwide forest inventory data base: users manual**. Gen. Tech. Rep. NC-151. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Holdaway, M.R. 2000. **The AFIS tree growth model for updating annual forest inventories in Minnesota**. In: Hansen, M.H.; Burk, T.E., eds. Integrated tools for natural resources 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 Research Station: 507-514.

Johnson, N.L.; Kotz, S. 1970. **Continuous univariate distributions**. New York: Houghton Mifflin.

Linhart, H.; Zucchini, W. 1986. **Model selection**. New York: John Wiley.

McRoberts, R.E.; Holdaway, M.R.; Lessard, V.C. 2000. **Comparing the STEMS and AFIS growth models with respect to the uncertainty of predictions**. In: Hansen, M.H.; Burk, T.E., eds. Integrated tools for natural resources 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 Research Station: 539-548.

ABOUT THE AUTHORS

Veronica C. Lessard is a Research Forester and Ronald E. McRoberts and Margaret R. Holdaway are Mathematical Statisticians at the North Central Research Station, USDA Forest Service, St. Paul, MN.

EVALUATING IMPUTATION AND MODELING IN THE NORTH CENTRAL REGION

Ronald E. McRoberts

ABSTRACT.—The objectives of the North Central Research Station, USDA Forest Service, in developing procedures for annual forest inventories include establishing the capability of producing annual estimates of timber volume and related variables. The inventory system developed to accomplish these objectives features an annual sample of measured field plots and techniques for updating data for plots measured in previous years. This paper describes and evaluates the feasibility of updating techniques and compares the bias and precision of the annual estimates they produce. The analyses indicated that simple, plot-level imputation and modeling techniques produced adequately unbiased and precise estimates of basal area per acre for large area estimates.

INTRODUCTION

The Renewable Forest and Rangeland Resources Planning Act of 1978 requires that the USDA Forest Service conduct inventories of forest land in the United States to determine its extent and condition and the volume of standing timber, timber growth, and timber removals. Passage of the Agricultural Research, Extension, and Education Reform Act of 1998 further requires that the Forest Service conduct annual forest inventories in all states with 20 percent of plots to be measured in each state each year.

Forest Inventory and Analysis (FIA) precision standards (USDA-FS 1970) require a sampling intensity of one plot for approximately every 6,000 acres in the North Central region. To satisfy this requirement, the geographical sampling hexagons established for the Forest Health Monitoring Program (White *et al.* 1992) were divided into 27 smaller FIA hexagons, each containing approximately 5,937 acres. An equal probability grid of field plots, designated the Federal base sample, was constructed by establishing a plot in each FIA hexagon. The Federal base sample was systematically divided into five interpenetrating, non-overlapping panels. Each year the plots in a single panel are selected for measurement with panels selected on a 5-year, rotating basis.

At least three approaches to calculating annual FIA estimates from the Federal base sample have been considered. The simplest approach is to use the data from the 20-percent panel of plots measured in the current year. Although these estimates reflect current conditions, their precision may be unacceptable for some variables due to the small annual sample size. An alternative is to use the data for all plots obtained from the five most recent panels of measurements and employ a moving average estimator. This alternative increases precision because data for all plots are used for estimation; the disadvantage is that the estimates do not reflect current conditions but rather a moving average of conditions over the past 5 years. A third approach is to update to the current year data for plots measured in previous years and then base estimates on the data for all plots. If the updating procedures are unbiased and sufficiently precise, this alternative provides nearly the same precision as the average of all plots but without the adverse effects of using out-of-date information. Two categories of updating techniques, imputation (Rubin 1987) and modeling, are of general interest and were evaluated using a specially created annual database of tree information.

ANNUAL DATABASE

Observations of the same 101,398 trees on 5,086 FIA plots for both the 1977 (Spencer 1982) and 1990 (Miles *et al.* 1995) Minnesota inventories were used to evaluate the updating techniques. These plots represent approximately 14.7 million acres of timberland. (In an

FIA context, timberland is defined as forest land that is capable of producing in excess of 20 ft^3 per acre per year of industrial wood crops under natural conditions and that is not associated with urban or rural development (Miles *et al.* 1995).) Plots included in the 1977 inventory were measured between 1974 and 1978; plots included in the 1990 inventory were measured between 1986 and 1991. These plots are termed variable radius plots due to the use of point sampling techniques that select trees with probability proportional to cross-sectional area rather than proportional to the frequency of occurrence in the population (Myers and Beers 1971). Thus, the number of trees in the population represented by a sample tree, termed the tree factor, varies by tree and is calculated as a scaling constant divided by the square of the tree diameter. Tree factors are used to expand the measurements of sample trees to per unit area estimates.

Based on observations of the individual trees, an 11-year database of annual diameters at breast height (DBH) (4.5 ft) and annual status with respect to survival, mortality, and harvest for each tree was created. Construction of the database required distributing total growth between inventories over varying numbers of years for individual trees in each of three categories: (1) trees alive at both inventories; (2) trees that died between inventories due to causes other than harvest; and (3) trees that were harvested between inventories. For trees alive in both inventories, average annual DBH growth was calculated by dividing the total growth in DBH over the measurement interval by the number of years between measurements. Measured DBH for the 1977 inventory was assigned to year 0, and DBHs for the 10 subsequent years were calculated by adding the average annual growth to the previous year's DBH. For trees that died due to causes other than harvest, a year of mortality between 1 and 10 was randomly selected and assigned to the tree independently of years of mortality assigned for other trees on the same plot. For harvested trees, a year of harvest between 1 and 10 was randomly selected and assigned to the tree but with the provision that all trees harvested on the same plot were harvested in the same year. For both mortality and harvested trees, measured DBH for the 1977 inventory was assigned to year 0, and DBHs for subsequent years up to the year of mortality or harvest were calculated by adding previous year's DBH and predictions of annual diameter growth obtained from individual tree diameter growth models (McRoberts and Lessard 2000). Although these procedures create greater uniformity in annual DBH growth than would be observed, the effects of differences between actual and calculated growth are expected to have minimal impact on evaluations of the updating techniques. Alternatives would require either annual measurement or destructive sampling of all trees. The former alternative would be prohibitively expensive and would risk the masking of actual DBH growth by DBH instrument measurement error; the latter alternative would be prohibited by landowners if not also ecologically disastrous.

Evaluations of the updating techniques were based on plot basal area per acre (BA), a variable representing the sum, scaled to a per acre basis using tree factors, of the cross-sectional areas of live tree boles at breast height.[1] Calculation of unbiased estimates of change in basal area per acre (ΔBA) is difficult using data from variable radius plots (Van Deusen *et al.* 1986). One technique fixes tree factors at the time of the first measurement and bases estimates of ΔBA on the increase in the cross-sectional areas of surviving trees and losses in BA due to mortality. This technique excludes contributions to ΔBA of new trees entering the sample. A second technique recalculates tree factors at every measurement, thus allowing new trees entering the sample to contribute to the ΔBA estimates. However, recalculation of tree factors excludes contributions to ΔBA of the growth of surviving trees, because the product of their cross-sectional areas and tree factors remains constant. A consequence of both techniques is that ΔBA is underestimated.

Although complex approaches to unbiased estimation of ΔBA using variable radius plots have been proposed, they were not considered for this study because evaluation of the updating techniques did not require absolutely precise ΔBA values. For this study, the constant tree factor technique was selected because it incorporates the growth of surviving trees, a primary interest in the construction of these updating techniques. Therefore, using

---

[1] *Unless otherwise noted, all future references to basal area (BA) and annual change in basal area (ΔBA) are understood to be on a per acre basis.*

the database of annual tree diameter values and tree factors corresponding to the year 0 DBHs, BA was calculated each year for each plot, and ΔBA was calculated each year for each plot as the difference between BA for the current and previous years.

UPDATING TECHNIQUES

Both imputation and plot-level models were investigated as a means of updating data for plots measured in previous years. Imputation for this application was a three-step process: (1) plots measured in the current year were placed into similarity groups; (2) plots measured in previous years were matched to a group of similar plots measured in the current year; and (3) values from the group of similar plots measured in the current year were selected to replace missing values for plots measured in previous years. For this application, plots were grouped on the basis of similarity in previous year's BA. The groups were created by first ordering all plots measured in the current year with respect to previous year's BA and then creating groups of 20 consecutive plots beginning with the plot with lowest previous year's BA. Plots measured in previous years were then matched to a group of plots measured in the current year on the basis of previous year's BA, whether it was obtained as a measurement or as an updated estimate. For each plot measured in a previous year, a plot was randomly selected with replacement from the group of 20 similar plots measured in the current year, and the latter plot's average annual ΔBA since last measurement was imputed to the former plot; this technique is hereafter referred to as IMPUTE.

Two model-based updating techniques were also investigated. For both modeling techniques, ΔBA for a plot was assumed to be related to both previous year's BA and to the current survival, mortality, or harvest status of trees on the plot. Thus, based on the annual status of trees, all plots were placed into one of three categories: (1) survival (no mortality or harvest); (2) mortality (at least one mortality tree); and (3) harvest (at least one harvested tree). There were no plots in the 1990 inventory data that had experienced both mortality and harvest since the 1977 inventory. For each category of plots, a simple model of the relationship between ΔBA and previous year's BA was selected, and its parameters were estimated using weighted regression techniques (table 1). In practice, the annual survival, mortality, and harvest status of plots will not be known. Thus, models for predicting the status of plots were also developed. First, all plots in the annual database were ordered with respect to previous year's BA and then placed into groups of 250 consecutive plots beginning with the plot with the lowest previous year's BA. For each group, the proportions of plots in the survival, mortality, and harvest categories were calculated. Simple models of the probabilities of survival, mortality, and harvest were then selected, and their parameters were estimated using maximum likelihood procedures (table 1). With this technique, hereafter referred to as PREDICT, the survival, mortality, and harvest status of each plot measured in a previous year was predicted using random numbers and the status models. Then, given the predicted status, ΔBA for the plot was predicted using the ΔBA models.

Although model predictions of survival, mortality, and harvest status were expected to be unbiased, the combined effects of their uncertainties and those of the ΔBA predictions risked increasing the variability of the annual mean estimates around the means of the annual database values, the standard errors of these

Table 1.—*The models*

Prediction	**Category**	**Model form**
Change in annual basal area, ΔBA	Survival	$E(\Delta BA)=\beta_1[1-\exp(\beta_2 BA)]$
Change in annual basal area, ΔBA	Mortality	$E(\Delta BA)=\beta_1+\beta_2 BA$
Change in annual basal area, ΔBA	Harvest	$E(\Delta BA)=\beta_1+\beta_2 BA$
Change in annual basal area, ΔBA	Disturbed	$E(\Delta BA)=\beta_1+\beta_2 BA$
Annual probability, P_{surv}	Survival	$E(P_{surv})=\exp(\beta_1 BA^{\beta_2})$
Annual probability, P_{mort}	Mortality	$E(P_{mort})=1-\exp(\beta_1 BA^{\beta_2})$
Annual probability, P_{harv}	Harvest	$E(P_{harv})=1-E(P_{sur})-E(P_{mor})$

means, or both. Thus, a second model updating technique, based on the assumption that satellite-based remote sensing techniques can be used to accurately detect plots that have experienced substantial disturbance, was developed (Befort 2000). Disturbance for this technique may be due to either mortality or harvest; no distinction is made. Disturbance using remote sensing techniques can be confidently detected for plots satisfying two criteria: previous year's $BA \geq 30$ ft^2/acre, and $(\Delta BA/BA) \leq -0.3$ (Befort, pers. comm.[2]). Using the annual database values, a simple model of the relationship between ΔBA and previous year's BA for plots satisfying these criteria was selected, and its parameters were estimated using weighted regression techniques (table 1). With this technique, hereafter referred to as REMOTE, updating again involves prediction of both status and ΔBA. First, plots measured in previous years that satisfied the remote sensing disturbance detection criteria were identified, and their ΔBA was predicted using the model constructed for this technique. For the remaining plots measured in previous years, survival, mortality, and harvest status and ΔBA were predicted in the same manner as for the PREDICT technique. However, considerably fewer plots required status prediction with the REMOTE technique.

For both modeling techniques, the uncertainty due to the residual variation around the estimated ΔBA curves was incorporated into the ΔBA predictions. For each estimated curve, distributions of the residuals for narrow categories of predicted ΔBA were estimated. In application, whenever a value of ΔBA was predicted, a corresponding residual from the appropriate distribution was randomly generated and added to the prediction. Thus, the estimates of standard errors of mean BA estimates obtained using the model updating techniques include the uncertainty of the model predictions due to residual variation.

SIMULATING THE INVENTORY

The feasibility of the updating techniques and the bias and precision of their annual BA estimates were evaluated by using the annual database as the basis for simulating the process of annually inventorying the 14.7 million timberland acres in Minnesota. For each of 250 simulations, 2,476 plots from among the 5,086 timberland plots were randomly selected to mimic the annual inventory intensity of 5,937 acres per plot. Each simulation was initiated with a simulated complete inventory of the 2,476 timberland plots by beginning with the annual database year 0 values. On a rotating basis, 20 percent of these plots were selected for measurement each year. Simulated measurement of a plot consisted of replacing its estimated BA value with the value for the appropriate year in the annual database. For the remaining 80 percent of plots for which measurement was not simulated in the current year, data were updated using each of the three inventory techniques. Each year, the mean BA across all plots and the standard error of the mean were calculated for each updating technique and for annual database of values; the latter estimates were designated TRUE. Following the simulations, the median values of the distributions of the annual means and the standard errors of the means were determined for each technique.

RESULTS

Evaluations of the updating techniques entailed comparing the median values for the 250 simulations of the estimated means of annual BA across all plots and the standard errors of the means obtained using the three updating techniques to the corresponding annual means and standard errors obtained from the annual database values. A comparison of the TRUE and IMPUTE annual means revealed that the imputation technique produced estimates that exhibited negligible bias with respect to the TRUE values (Table 2). In addition, the similarity between the TRUE and IMPUTE standard errors indicated that the IMPUTE technique quite accurately estimated the uncertainty in the TRUE means. A comparison of the median values of the TRUE, PREDICT, and REMOTE annual means revealed that neither modeling technique exhibited conspicuous bias (table 2). A comparison of the median standard errors of the means indicated that both modeling techniques adequately estimated the TRUE standard errors. As expected, the variability of the PREDICT annual means around the TRUE means was greater than for the REMOTE means.

A further comparison of the updating techniques was made by calculating the root mean

---

[2] *William Befort, Division of Forestry, Minnesota DNR, November 9, 1999.*

Table 2.—*Median values of annual means and standard errors of means for 250 simulations*

Year	TRUE		IMPUTE		PREDICT		REMOTE	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	56.24	0.74	56.12	0.76	56.38	0.73	56.09	0.74
2	55.53	0.73	55.23	0.75	55.39	0.73	55.22	0.73
3	54.07	0.73	54.28	0.75	54.34	0.72	53.88	0.73
4	53.14	0.72	53.41	0.74	53.30	0.72	52.93	0.72
5	51.64	0.72	52.14	0.73	52.16	0.72	51.64	0.72
6	50.80	0.72	51.31	0.72	51.31	0.71	50.75	0.72
7	49.72	0.72	49.66	0.72	50.05	0.71	49.55	0.72
8	48.48	0.72	48.73	0.72	49.12	0.71	48.47	0.71
9	45.97	0.72	47.28	0.72	47.75	0.71	46.59	0.71
10	45.24	0.72	45.91	0.73	46.54	0.71	45.52	0.71

square error of the squared deviations of the updated annual means from the corresponding TRUE annual means for years 5-10 (table 3). The first 4 years were excluded in this comparison, because annual means for these years retained a component of the year 0 complete inventory. The resulting 5th percentile, median, and 95th percentile values for distributions of root mean square errors indicated that the differences between the IMPUTE and REMOTE means were small with respect to root mean square deviation, although the REMOTE results were somewhat better than the IMPUTE results. The similarity of results for these updating techniques may be partially attributed to the large area represented by the aggregation of data over this large number of plots; it is yet to be determined if these results hold for smaller areas.

CONCLUSIONS

Several conclusions emerged from these analyses. First, the simple, plot-level updating techniques were not only feasible, but they produced acceptable estimates of both annual BA means and standard errors of the means for large area estimates. Second, although the REMOTE technique produced somewhat better results, the quality of the means obtained with the modeling and imputing techniques relative to the TRUE means was similar. Third, because 5-year ΔBA is usually small compared to BA 5 years in the past and because the uncertainty in ΔBA predictions is small compared to the natural variation in BA among plots, ΔBA appears to be an appropriate quantity to use as the basis for updating. Fourth, but less conclusively, a combination of disturbance detection using remote sensing procedures and model predictions of survival, mortality, and harvest status appeared to be a better alternative than using only model predictions of status. Finally, additional testing is appropriate to determine if these large area results hold for smaller numbers of plots representing smaller timberland areas.

LITERATURE CITED

Befort, W. 2000. **Change detection, stratification, and mapping for continuous forest inventory**. 1999 proceedings of the Statistics and the Environment Section; annual meeting of the American Statistical Association; 1999 August 8-12; Baltimore, MD.

Table 3.—*Root mean square deviation: updated versus TRUE annual means*

Statistic	TRUE	IMPUTE	PREDICT	REMOTE
5th percentile	0.00	0.58	0.66	0.29
Median	0.00	0.68	1.01	0.33
95th percentile	0.00	0.77	1.34	0.54

Alexandria, VA: American Statistical Association.

McRoberts, R.E.; Lessard, V.C. 2000. **Estimating uncertainty in annual forest inventory estimates.** 1999 proceedings of the Statistics and the Environment Section; annual meeting of the American Statistical Association; 1999 August 8-12; Baltimore, MD. Alexandria, VA: American Statistical Association.

Miles, P.D.; Chen, C.M.; Leatherberry, E.C. 1995. **Minnesota forest statistics, 1990, revised**. Resour. Bull. NC-158. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Myers, C.C.; Beers, T.W. 1971. **Point sampling and plot sampling compared for forest inventory**. Res. Bull. 877. August, 1971. West Lafayette, IN: Purdue University, Agriculture Experiment Station.

Rubin, D.B. 1987. **Multiple imputation for nonresponse in surveys**. New York: Wiley.

Spencer, J.S., Jr. 1982. **The fourth Minnesota forest inventory: timber volumes and projections of timber supply**. Resour. Bull. NC-57. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

USDA Forest Service (USDA-FS). 1970. **Operational procedures**. For. Serv. Handb. 4809.11, Chapter 10:11.1-1—11.1-3. Washington, DC: U.S. Department of Agriculture, Forest Service.

Van Deusen, P.C.; Dell, T.R.; Thomas, C.E. 1986. **Volume estimation from permanent horizontal points**. Forest Science. 32(2): 415-422.

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

ABOUT THE AUTHOR

Ron McRoberts is a Mathematical Statistician and Group Leader with the Forest Inventory and Analysis program of the North Central Research Station, USDA Forest Service, St. Paul, MN.

ANNUAL FOREST INVENTORY: AN INDUSTRY PERSPECTIVE

Roger Lord

ABSTRACT.—The Forest Inventory and Analysis Program serves important public interests by providing credible data for informed public forest policy debates as well as feedback to the forest-based economic market. This feedback, which affects timber price expectations, helps ensure resource sustainability by promoting better investment decisionmaking within the forest products sector. Industry's use of FIA data is illustrated by the types of analysis performed by Boise Cascade Corporation. Key needs include more timely and consistent data across all forestlands, improved spatial resolution, and integration of socioeconomic variables that affect timber availability. Concerns about the implementation of the annual inventory program required by the 1998 Farm Bill, particularly in the western states, are discussed.

IMPORTANCE OF FOREST INVENTORY INFORMATION

After two Blue Ribbon Panel reports and countless other discussions, it should by now be unnecessary to point out the value of the Forest Inventory and Analysis program in monitoring the status and health of the Nation's forests. The program represents the only continuous inventory system that quantifies the condition of forest ecosystems across the United States. The critical information provided by the program can promote informed discussion of public forest policy issues and can serve as the basis for sound business decisions within forest industry.

Public Forestry Issues

The FIA program provides the information base upon which we can intelligently address important public issues such as forest sustainability, ecosystem health, land use, and timber policy.

Intense public debate rages in virtually every region of the country over forestland use and management. In the coastal Pacific Northwest, the spotted owl controversy, old growth protection, and threatened and endangered species continue to focus attention on the management of both public and private timberlands. In the Inland Northwest, forest health and sustainability, roadless areas, and endangered species issues dominate. In the Rocky Mountain States, forest health issues such as the decline of aspen stands in Colorado are debated. In the South, it is the impact of chip mills and clearcutting of hardwood stands as well as wetland protection that has recently drawn the most attention. Wetlands, endangered species, and forest sustainability are important issues in the Lake States, while forest fragmentation, acid rain, and loss of wildlife habitat are important in the Northeast. Finally, on a national scale, the U.S. must address the forest sustainability issue through the internationally accepted forest sustainability criteria and indicators.

Objective forest resource data from the FIA program are essential to developing a fact-based, intelligent discussion of these issues. Yet, FIA data are useful only when they are current, consistent, and reliable. When they are not, public policy debate quickly degenerates into emotions, perceptions, and opinions.

Sound Business Decisions

Business decisions within the forest products sector are often strongly influenced by expectations of future timber prices. The decision to add new manufacturing capacity, or reconfigure or close existing facilities, for example, is essentially an analysis of the expected cost competitiveness of the facility and return on investment. Fiber costs, which can be up to 80 percent of the variable costs of production of some wood products, are key to this analysis. As another example, determining the appropriate level of investment in productivity-enhancing forest management practices also hinges on expectations of future timber prices. Whether a particular silvicultural practice earns an adequate financial return is highly dependent on the expected increase in value at harvest. Price expectations provide the industry with the market feedback signals it needs to make adjustments to demand (e.g., via changes in mill capacity or development of technology to decrease fiber use) and supply (e.g., through silvicultural investments, genetic research, or development of new supply sources) to maintain competitive raw material costs.

Price expectations depend on changes in the available supply of fiber relative to demand. Accurate projections of future changes in timber prices and availability, in turn, are dependent on accurate, current resource data. To the extent that data are out-of-date and misleading, price expectations will be inaccurate and inappropriate business decisions may be made as a result.

There are obvious public benefits to good business decisions within the industry. Questionable business investments, such as adding new manufacturing capacity in an already supply-constrained woodbasket, may have a negative impact not only on the individual company involved, but also on resource sustainability. Thus, it is in the public interest as well as the industry's interest to maintain current resource inventories.

HOW BOISE CASCADE USES FIA DATA

Current Uses

To understand how FIA data are used by industry, it may be helpful to review the use of that data within one company. Boise Cascade Corporation has timberlands and manufacturing operations in Washington, Oregon, Idaho, Minnesota, Louisiana, and Alabama. Thus, it has direct business interests in FIA data from four of the five program regions. Current use of FIA data within the Timberland Resources department at Boise Cascade can be broken down into four general categories:

- Broad regional resource monitoring
- Detailed woodbasket modeling
- Support of investment analysis
- Regulatory impact assessment

To better understand our general operating environment, Timberland Resources undertakes regional resource studies for the Pacific Northwest, Lake States, and Southern U.S. on a regular basis. In the Northwest and Lake States, we have developed a projection system called *Dynamic Forest Simulator™* to project FIA plot data for up to 20 years using a combination of stand table projection and individual tree diameter growth and mortality modeling. The current model covers California, Oregon, Washington, Idaho, and Montana as well as North Dakota, Minnesota, Wisconsin, and Michigan. Users can select sample plots based on geographic location and plot characteristics, update the plot data to the current year, and project into the future by simulating historical and projected harvest levels. Harvest can be specified by owner, species, and diameter class, or it can default to historic harvest patterns. Using a different approach in the U.S. South, we have looked at the impact of increased silvicultural intensity, urbanization and forest fragmentation, and wetland and coastal zone regulations on future available supply versus biological inventory.

In addition to regional studies, we also undertake more detailed modeling in the woodbaskets where we operate. For example, we have developed linear program-based resource allocation models for our southern woodbaskets in which FIA data are used as the basis for the supply side of the model. The demand side is based on an individual mill database of the primary wood processors within the woodbasket and surrounding area. FIA data are used to predict supply to each mill from within and outside of the designated woodbasket. Marginal wood cost curves (supply curves) are also developed to predict delivered wood cost to each mill.

FIA data also support investment analysis efforts, including financial analysis of both manufacturing investments, and timberland and silvicultural investments. For example, we use FIA data to analyze potential fiber supply and develop delivered fiber cost estimates for proposed manufacturing facilities. FIA data also help set the timber price scenarios for our timberland planning models by which we develop our silvicultural and harvesting plans.

Another use of FIA data is in the assessment of the impact of regulatory changes on timber supply. In the West, much attention has been focused on the scenario analysis in which we looked at the impacts of alternative degrees of riparian and endangered species restrictions on available timber supplies. In the South, we have examined the implications of wetland and coastal zone management regulations using FIA data.

Future Uses

There are at least two areas of analysis in which we would like to be able to use FIA data in the future. The first is more spatially explicit timber supply analysis. The FIA sample plot framework is only pseudo-spatial at best. Currently, we know only that a plot is representative of forested acres somewhere in the county, but we don't know *where* in the county the represented acreage is. Higher spatial resolution to forest resource data is needed to better address issues of availability such as the impacts of urbanization and land use regulation. Better spatial resolution of the data would also be more helpful in regions with large counties and more scattered forest cover such as the Inland Northwest.

A second area of future need is data to support landscape analysis. Boise Cascade has completed ecosystem management projects in Idaho, Washington, and Minnesota and has successfully classified its fee-owned land base into an ecological classification scheme known as an ecosystem diversity matrix (Haufler *et al.* 1996). This system classifies land into Habitat Type Class (e.g., Warm-Dry Douglas-fir) to describe the potential vegetation type and Vegetative Growth Stage (e.g., Medium Tree, Multi-Story) that describes stand structure. The resulting acreage matrix describes the habitat contribution of company lands and can be projected over time using Boise Cascade's forest planning models. Company fee land, however, is only one component of the ecological landscape. For example, in Idaho, Boise Cascade owns 200,000 acres of the 5.9 million acre Southern Idaho Batholith ecoregion. To fully describe the landscape and depict the contribution of company lands within the context of the entire landscape, we need similar data across other ownerships within the ecological region. Aside from access issues, data in the detail acquired on company lands would be prohibitively expensive to collect across all acres. As an alternative, FIA data, particularly augmented with more spatially explicit remote sensing data, could provide the information base for this type of large landscape analysis.

KEY DATA NEEDS

Current, Consistent, and Comprehensive Data

Current, consistent, comprehensive FIA data are required to meet the challenges of addressing public forest policy and business issues. Data older than 5 years are simply not credible with decisionmakers or the public. Table 1 shows the age of FIA data Boise Cascade regularly uses in analysis. On an acreage-

Table 1.—*Age of key FIA data used by Boise Cascade*

State	Area of timberland (Thousand acres)	Latest survey	Previous survey	Approximate midpoint
Alabama	21,932	1990	1982	1986
Louisiana	13,783	1991	1984	1987.5
Mississippi	18,587	1994	1987	1990.5
Tennessee	13,265	1989	1980	1984.5
Texas	11,774	1992	1986	1989
Minnesota	14,723	1990	1977	1983.5
Idaho	21,427	1991	1981	1986
Oregon-Eastside	2,978	1992	1987	1989.5
Oregon-Westside	6,777	1986	1976	1981
Washington-Eastside	4,008	1992	1980	1986
Washington-Westside	9,581	1991	1979	1985
Total	138,836			
Weighted Avg Date		1990.8	1981.9	1986.4
Weighted Avg Age as of 1/1/2000		9.2		13.6

weighted basis, the average plot is over 9 years old and the average to the last period midpoint (the relevant age for growth and removals estimates) is nearly 14 years. Within the 13 southern states, average age to plot midpoint is about 11 years. Since that time, we estimate that the southern timber harvest has increased by 20 to 25 percent. Harvests in the Pacific Northwest meanwhile have fallen by about 40 percent across all ownerships, and harvesting patterns have changed. Yet these changes are not yet being fully reflected in FIA statistics because of the long cycle times.

However, long cycle time is only part of the problem. We can also no longer afford the delays of sometimes up to 2 to 3 years in the release of FIA data and analysis once field collection has been completed. The value of the data decreases rapidly, perhaps exponentially, with time. As has been demonstrated, particularly by the Southern Station, technology allows significant improvements in data collection, editing, management, and release. These technologies need to be fully exploited. Further, allowing raw field data to sit on the shelf for 6 months or more because of a lack of analysts is difficult to understand given the cost of data collection and value of timely release. Stations should be adequately staffed with analysts to efficiently process and release the data.

Consistency within and between FIA administrative and survey units is also essential, as the Blue Ribbon Panel reports have pointed out. Rarely in my industrial experience is FIA data from only one survey unit or state used. We regularly merge data from two or more states in our analyses. Much time and frustration has been spent uncovering the differences in data collection methods, definitions, and coding between the various FIA data sets.

Finally, it is very important that data provide a wall-to-wall coverage across all ownerships, including National Forests, and all classes of forested land. To understand the ecological condition of our forest resource, we must necessarily have consistent data for all forests regardless of commercial availability for harvest.

Increased Emphasis on New Techniques

Several other papers in this workshop have already covered this topic in some detail. Techniques such as remote sensing and GIS offer potential for both reducing program costs and enhancing the spatial component of the data. More work is needed on growth and change modeling so that FIA statistics can be updated through modeling and perhaps measured less frequently. The Farm Bill also requires that the FIA program develop 20-year projections of forest conditions, yet it appears that little work has been directed to this effort.

Moving Beyond Biological Inventory

From an industrial perspective, I believe we must get beyond the practice of using biological inventory as a measure of timber supply. We can encourage this change by capturing socioeconomic aspects that relate to the availability of inventory for commercial use. The role of FIA in this would be to integrate socioeconomic data into the plot database. For example, linking plot data with measures of urbanization, such as the Rural-Urban Continuum Codes developed by the USDA Economic Research Service or population density statistics from the Bureau of Census, would allow users to take into account population and land use pressures when analyzing resource data. It may also be useful to develop a sub-classification of the timberland definition (e.g., suburban timberland) to take into account factors affecting availability and to draw further attention to the availability issue.

IMPLEMENTATION OF THE FARM BILL

Inequity Between FIA Programs in the Eastern and Western U.S.

The 1998 Farm Bill addressed many of the issues raised previously in this paper and sets forth a framework for an improved FIA program. Boise Cascade, however, is concerned about progress to date in implementing aspects of the annual forest inventory program. To be sure, a great deal of progress is being made as we have seen in the papers presented at this workshop and I do not want to minimize the tremendous amount of work that has been accomplished and change that has already taken place. However, Boise Cascade is particularly concerned about implementation in the West and continued inadequate funding and support from the Chief's Office.

The FIA Program is currently administered by five experiment stations including the Southern, Northeastern, North Central, Rocky Mountain, and Pacific Northwest Research Stations. The lion's share of funding for the FIA program historically has gone to the stations in the East and particularly to the Southern and North Central Stations. Meanwhile, the FIA programs in the West (including the Pacific Northwest and Rocky Mountains) and the Northeast have languished for lack of funding and staff.

The Forest Service's current plan to implement the Farm Bill program only perpetuates this inequity by adopting a base program that calls for measuring 15 percent of the sample plots annually in the East but only 10 percent in the West. This arbitrary decision is more a product of agency politics and culture than science. It makes little scientific or statistical sense from the standpoint of providing the consistent and timely base of forest resource information.

To support this decision, it has been argued that trees grow slower in the West and therefore change is less dynamic. This is fallacious reasoning on at least two counts. First, FIA data refute this argument. Data on net annual growth and timberland acreage by region in Powell *et al.* (1993) indicate that net annual growth per acre of timberland is actually higher in the West than the East (table 2). Per acre growth rates on a cubic foot basis are higher in the Pacific Northwest and Pacific Southwest than they are in the Southern U.S.

Table 2.—*Growth per acre on timberland, by region*

Region	Net annual growth	Timberland	Net annual growth per acre
	MM Ft^3	*MM Ac*	Ft^3*/Ac/Yr*
Northeast	3,093	79.4	38.92
North Central	2,269	78.4	28.96
Southeast	4,323	84.8	50.98
South Central	5,509	114.5	48.10
Great Plains	98	3.5	27.81
East	**15,292**	**360.6**	**42.40**
Intermountain	2,074	59.1	35.09
Alaska	270	15.1	17.90
PNW	2,904	37.9	76.73
PSW	1,087	16.9	64.30
West	**6,334**	**128.9**	**49.13**

Second and more importantly, tree growth is only one component of change in forest ecosystems and should not dictate inventory intensity. Other forest and landscape-scale change agents make the western forests potentially as dynamic if not more dynamic than many eastern forests. These change agents include changes in harvesting patterns, forest management systems, insect and disease outbreaks, fire, and urbanization and development.

The base FIA program should be defined as that level of effort that ensures that the desired national standard of statistical accuracy is achieved in each region and state. The starting point should be the annual sample size required for a scientifically acceptable estimate at state level using only that year's data. Science should set the base program effort in each state, not politics.

I want to be very clear that I am not arguing for a reallocation of static funding from the Southern and North Central Stations to the other Stations. The point of the Blue Ribbon Panel reports as well as the Farm Bill legislation was that the FIA program was inadequate nationwide and needed to be raised to a higher standard of consistency, comprehensiveness and timeliness. This is true in every region. All FIA units need to be brought up to a new, higher, common level of performance. This is essentially a funding issue and there is little evidence that the FIA program has received the called-for level of priority from the U.S. Forest Service Chief's Office. Despite its national importance and broad support, the FIA program represents only about 1 percent of the agency's budget.

Partner funding through the state forestry agencies is another area of concern to Boise Cascade. Differences in agency focus and landownership patterns between the states in the West and East bring into question the feasibility of reliance on state funding. For example, given that 70 percent of the forest land in Idaho is federally owned, is it reasonable to expect the state Department of Lands to partner with the FIA program at the same level as another state where 95 percent or more of the timberland is privately owned? Finally, we believe that securing and maintaining consistent, continuous funding from 40 or more state legislatures will greatly complicate the task of providing a stable and capable FIA program.

The annual FIA system holds promise to at last bring the program up to the level of quality envisioned by the two Blue Ribbon Panels, the National Association of State Foresters, and other stakeholders. Progress is being made but the annual program has not yet been embraced nationwide and there appear to be many technical questions and funding inequities that have yet to be addressed. Industry will continue to monitor closely how the Forest Service resolves these important issues and truly begins implementing a consistent nationwide program.

LITERATURE CITED

Haufler, J.B.; Mehl, C.A.; Roloff, G.J. 1996. **Using a coarse-filter approach with species assessment for ecosystem management**. Wildlife Society Bulletin. 24(2): 200-208.

Powell, D.S.; Faulkner, J.L.; Darr, D.R.; Zhu, Z.; MacCleary, D.W. 1993. **Forest resources of the United States, 1992**. Gen. Tech. Rep. RM-234. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

ABOUT THE AUTHOR

Roger Lord is Senior Forest Economist, Timberland Resources Planning and Development, Boise Cascade Corporation, Boise, ID.

MAINE'S ANNUAL INVENTORY: STATE PERSPECTIVES

Kenneth M. Laustsen

ABSTRACT.—In 1999, Maine became the first northeastern state to begin implementing the USDA Forest Service's annual inventory system as directed by PL 105-185, the Agricultural Research, Extension, and Education Reform Act of 1998. The Maine Forest Service, in collaboration with Forest Inventory and Analysis program of the Northeastern Research Station of the USDA Forest Service is currently measuring Panel #1, a 20-percent component of the annual inventory design. This paper covers five major topics: the implementation plan, training and measurement progress, data analysis, reporting goals, and conclusions.

IMPLEMENTATION PLAN

The current implementation plan for Maine's annual inventory system, dated April 30, 1999, grew from a meeting of the Advisory Committee convened by the Maine Forest Service (MFS) at the University of Maine on November 24, 1998. This meeting, attended by a wide array of stakeholders, was held to outline the critical needs for implementing the new inventory system. Additional meetings and teleconferences of USDA Forest Service's Northeastern Research Station (NERS) and MFS specialists focused on the core variable listing and the development of a field guide. Then, in February 1999, the Advisory Committee reconvened to finalize the list of measurement variables and other procedural details.

This plan was fluid over the 6-month planning horizon, changing as the new national measurement protocols were proposed, other information needs evaluated, and variable tradeoffs discussed. With the initiation of training in April 1999, its content became formalized into a study plan.

The Advisory Committee met again in late May 1999 for an update on training, early measurement progress, potential changes, and a field visit to a simulated measurement plot. This committee is expected to become actively involved again early in the year 2000, for the discussion of analyses and reporting of the Panel #1 measurements.

As part of the implementation plan, MFS agreed to supply the personnel requirements and get the program jump-started in 1999. The future expectation is for NERS to fund their 75-percent proportion of the required 20-percent annual measurement expenses, based on the U.S. Congress providing sufficient appropriations for a 7-year annual inventory system (15 percent) in the East. However, MFS does not intend to continue to fund the staff allocation into the future.

TRAINING AND MEASUREMENT PROGRESS

In April 1999, two NERS sponsored and five MFS sponsored crews were trained and certified over a 2-week period by David Alerich, Jason Morrison, and other NERS personnel.

Following crew certification, field measurement of Panel #1 plots began immediately in late April. The graph on the next page displays both weekly and overall-to-date crew plot measurement production in Maine's Panel #1 (fig. 1).

Currently, 88.5 percent of plot measurements have been completed for the year, and the measurement season is expected to conclude around the end of November. More importantly to date, crew production has averaged 3.3 completed plots per week.

Panel #1 had its fair share of start-up problems to overcome. In terms of crew efficiency, the greatest impact was the unavailability of the full complement of Panel #1 plot tallysheets at the start of the measurement season. Plots trickled in over a 2-month period, delaying preparatory work like landowner contacts and increasing crew allocation inefficiencies. Also,

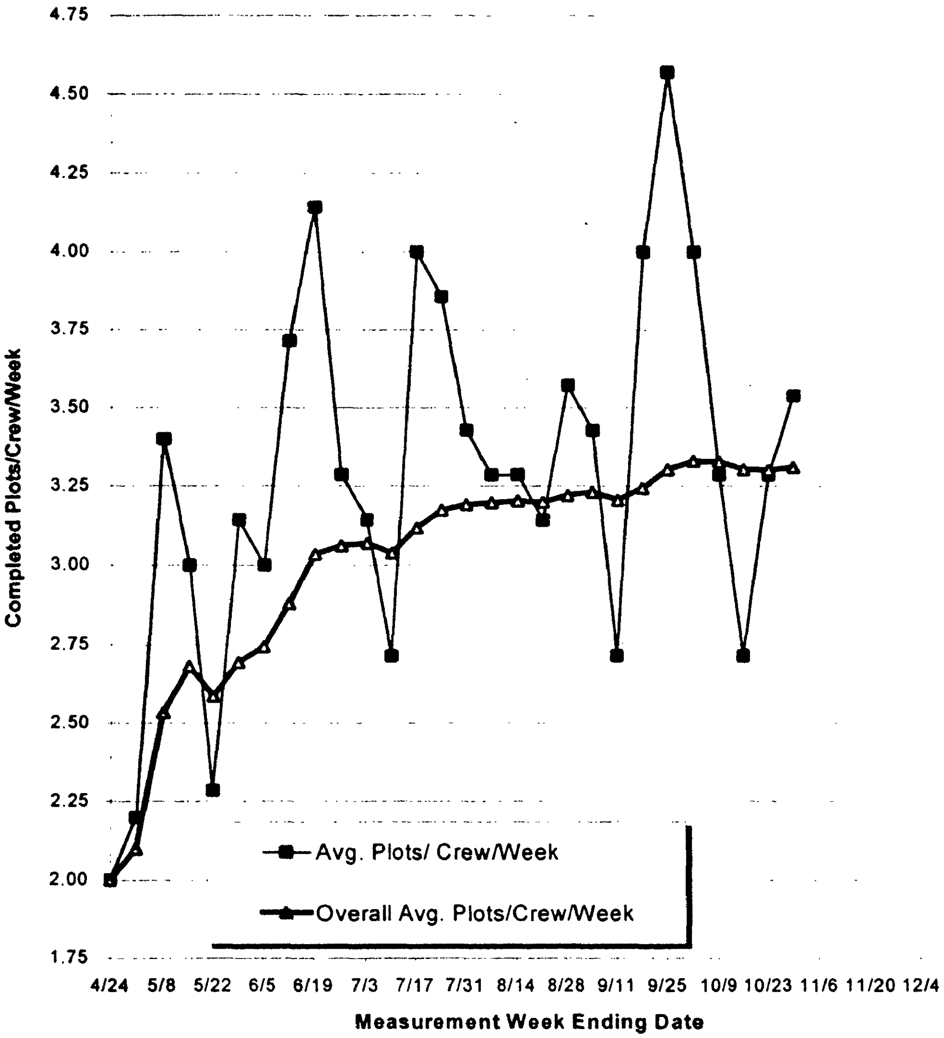


Figure 1.—*Plot measurement productivity.*

data entry and analysis was handcuffed by the unavailability of portable data recorders. When the recorders finally became available in August, 50 percent of the plots were already measured, and MFS decided not to introduce them at that time to maintain measurement progress and data reliability.

One additional problem cropped up late in the measurement season that could be described as a minor difficulty with major implications. On October 1, the NERS Quality Assurance/ Quality Control Coordinator and the MFS Field Supervisor were verifying completed plots when they collectively realized that the real plot count for Panel #1 was 21 more than originally scheduled. This number constituted an average week's production for the seven field crews, and its discovery at that time of the year in Maine caused a slight panic.

Two of the more controversial issues involving the annual inventory system are the core plot design and the core variables. MFS supports the goal of creating data collection procedures for a single, consistent, and uniform framework across all FIA units, in accordance with the Blue Ribbon Panel recommendation (American Forest and Paper Association 1998). Beyond that level, regions and states can each add other variables or data collection procedures to incorporate special information needs. Maine, for example, has 1 additional mil-acre seedling/sapling plot and is doing the full suite of Forest Health Monitoring crown damage coding on all measured trees. Finally, MFS is adamant that on the average, a two-person crew must complete a plot measurement in a 1-day visit. Based on experiences in the first year, we are almost there.

DATA ANALYSIS

In September 1999, I requested a subset of validated plots, preferably a single FIA unit, to begin data familiarization, graphic template construction, and statistical testing/comparison to the 1995 periodic inventory. That request mirrors a separate mandate for MFS, in that the State Legislature is requiring the publication of an annual inventory report. It was disappointing to learn that only very limited data entry for completed plots had begun by a point in the season when total panel measurements were 80 percent completed. Maine plot data, recorded on paper tallysheets, were being used to beta test the portable data recorder programming for data entry and error checking.

One of the major issues that the Advisory Committee discussed in detail was the continuing lack of remeasurement data for component of change analysis and trends. Only 50 percent of the plots measured in the 1995 periodic inventory are part of the new annual 5-year Panel sample. Furthermore, the change to the core plot design meant that only the 1/24-acre concentric overlay of subplot #1 in the new cluster would contain remeasured information from the 1995 1/6-acre plot area. As a result, only 13 percent of the data are of remeasurement quality. The Advisory Committee was well aware of this predicament and the inherent analytical weakness, but the alternative of aggressively collecting additional remeasurement data on the 1995 1/6-acre plots would have increased measurement time beyond the one plot/crew/day productivity threshold. All parties agreed to recognize the lack and to move onward with just the core plot design.

REPORTING GOALS

At this point in the inventory process of Panel #1, the remaining goals are pretty distinct, at least from my viewpoint:

- Complete data collection
- USFS provides keypunched data, common tables, analyses, and descriptive statistics
- Begin creation of templates on a state-level analysis of inventory
- Reconcile mutual reporting responsibilities.

I have a mandate to make a March 2000 presentation to the Maine Legislature on Panel #1's progress and gross comparisons of population estimates to the 1995 periodic inventory. No excuses will be accepted for not meeting that mandate.

CONCLUSIONS

In conclusion, MFS and I are confident that the spirit of collaboration and cooperation expressed and demonstrated to date will continue, and that both MFS and NERS will meet their own unique, independent, and collective needs in supplying new inventory information to the public.

LITERATURE CITED

American Forest and Paper Association. 1998. **The report of the second Blue Ribbon Panel**. Washington, DC. (Available at www.srsfia.usfs.msstate.edu/wo/brp2.htm)

USDA Forest Service, Forest Inventory and Analysis. 1998. **Strategic plan for forest inventory and monitoring**. Washington, DC. (Available at www.srsfia.usfs.msstate.edu/wo/strategy_setup.htm)

USDA Forest Service, Forest Inventory and Analysis and Maine Forest Service. April 30, 1999. **An implementation plan for Maine's annual inventory system**. Radnor, PA. Unpublished document.

ABOUT THE AUTHOR

Kenneth M. Laustsen is a Biometrician with the Maine Forest Service in Augusta, ME.

McRoberts, Ronald E.; Reams, Gregory A.; Van Deusen, Paul C.
2000. **Proceedings of the first annual forest inventory and analysis symposium**; 1999 November 2-3; San Antonio, TX. Gen. Tech. Rep. NC-213. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 57 p.

Documents progress in developing the techniques required for full implementation of the national Forest Inventory and Analysis program's new annual forest inventory system.

KEY WORDS: Annual forest inventory, interpenetrating sample, remote sensing, growth models, imputation, 1998 Farm Bill.