Vision for the Future of FIA: Paean to Progress, Possibilities, and Partners

Richard W. Guldin,[1] Susan L. King, and Charles T. Scott[2]

Abstract. — The Forest Inventory and Analysis (FIA) program of the U.S. Department of Agriculture Forest Service has made significant progress implementing the annualized inventory in 46 States in 2004. Major increases in program performance included the availability of plot data and the plots' corresponding approximate coordinates. A mill site study and biomass models were used to compare actual versus approximate coordinates. The protocols used to protect the privacy of private forest landowners did not meaningfully alter the results. A new strategic plan for FIA will be developed for 2007–12. Through meetings with partners and customers, FIA will evaluate opportunities to broaden the information collected and analyses of this data.

Introduction

The Forest Inventory and Analysis (FIA) program of the U.S. Department of Agriculture (USDA) Forest Service has nearly completed the transition to an annualized inventory approach that incorporates forest health detection monitoring and uses state-of-the-art geospatial technologies. Web delivery of results is increasing. The principles of continuous improvement are being applied to the FIA program, focusing on those aspects identified by users as most important. The program is poised to begin a second round of strategic planning for 2007–12. Partners will play key roles in the strategic planning process to evaluate and prioritize the future possibilities and help the program achieve its goals.

Progress

In fiscal year (FY) 2003, the annualized FIA program had field operations in 46 States. Measurements were taken across the landscape, covering 71 percent of the forest land in the United States,[3] an increase of 9 percent over the area covered in FY 2002. A total of 43,034 Phase 2 (P2) plots, the traditional ground sample, and 3,740 Phase 3 (P3) plots were remeasured. P3 plots measure additional variables that indicate forest health.

Users' needs were met in a variety of ways. FIA program analysts engaged in 1,450 significant consultations with users, an increase of 41 percent from the previous year. Users made nearly 15,000 downloads of data from the FIA Web site, a 20-percent increase over the previous year. Web tracking software identified the data most frequently downloaded, and this information was used to focus continuous improvement activities.

During the past 2 years, technical specialists have revised the FIA Field Guide for Phase 2 Measurements (http://www.fia.fs.fed.us/library/field-guides-methods-proc/). A substantial number of changes in field procedures had been proposed to simplify fieldwork and make it more efficient. A major revision of the guide, version 2.0, was released in January 2004. Data recorder and compilation software were upgraded in response to the new field guide so that its protocols could be implemented during the 2004 field season.

Privacy Policy

The privacy policy adopted a year ago in response to the new legislative language in the FY 2000 Interior Appropriations Bill continues to attract the attention of external users and analysts.[4]

---

[1] Director, Science Policy, Planning, Inventory and Information, U.S. Department of Agriculture (USDA), Forest Service, P.O. Box 96090, Washington, DC 20090–6090.

[2] Operations Research Analyst and Program Manager, respectively, USDA Forest Service, Northeastern Research Station, Forest Inventory and Analysis, 11 Campus Blvd., Suite 200, Newtown Square, PA 19073.

[3] This percentage includes the acreage of forest land in interior Alaska. In prior years, the percent coverage excluded interior Alaska. With plans now in place to begin work in interior Alaska in FY 2005, reports of area covered in FY 2003 and beyond will include interior Alaska.

[4] Section 348 of H.R. 3423, Department of the Interior and Related Agencies Appropriations Act of 2000. (Nov. 17, 1999). H.R. 3423 was incorporated by cross-reference in the conference report into H.R. 3194 [Division B]. H.R. 3194, the FY 2000 Consolidated Appropriations Bill, became Public Law 106-113 on Nov. 29, 1999.

The goal of the policy is to protect the privacy of private forest landowners who allow FIA field crews to collect data on their property. The policy ensures that data for any plot cannot be linked *with certainty* to the participating private landowner.

Two-thirds of the forest land in the United States is privately owned. Permission to collect data on private lands is vital to the continued credibility of the FIA program. In recognition of the importance of private landowner participation in the FIA program, FIA was placed under the same privacy protection provisions as other critical agricultural inventory, monitoring, and census programs operated by the National Agricultural Statistical Service (NASS). A new privacy law was not created for FIA. Rather, Congress gave private forest landowners participating in the FIA program the same legal protections already enjoyed by farmers participating in the other USDA programs.

The pre-1998 FIA privacy policy was updated to comply with the law. NASS and the USDA Office of the General Counsel (OGC) were consulted to ensure that the legislative intent was faithfully implemented in the new policy. USDA's long experience with the same legislative language in other USDA programs provided a sound foundation for developing the new FIA policy.

Based on experience with the other USDA agricultural crop inventory programs operated by NASS, OGC did not believe that "fuzzing" (providing an approximate location) alone was sufficient to meet the terms of the legislation. FIA national program staff and members of the FIA Statistics Band consulted with experts from the American Statistical Association and the U.S. Census Bureau to learn what techniques they advocated for ensuring the privacy of individuals participating in surveys. The experts believed that a small amount of "swapping" (switching the locations of two similar plots) would be much more effective than coarse "fuzzing." Indeed, with a small amount of "swapping," the statistical experts believed that "fuzzing" could be radically reduced, and the combination of "swapping" with reduced "fuzzing" would improve the quality and usefulness of the publicly available data while providing the minimal amount of privacy protection required. OGC concurred.

To understand what impact the new policy might have on analyses performed with data from FIA's public database, the two techniques must first be understood. "Fuzzing" consists of randomly adjusting the latitude and longitude locations of the plot. Under the old, pre-1998 FIA privacy policy, a combination of latitude and longitude were rounded to the nearest 100 seconds. This meant that users could be certain that the actual plot location fell within the 2,010 acres surrounding the plot location contained in the FIA public database. Under the new FIA privacy policy, latitude and longitude are randomly located within one-half mile. This means that the actual plot location is masked within only a 500-acre area. Users commenting on the draft privacy policy applauded the fourfold reduction in fuzz compared to the old policy.

"Swapping" consists of exchanging the plot coordinates for a small number of *similar* plots *within close proximity and in the same county*. Swapping only occurs on private forested plots and depends on the region of the country. Between 0 and 10 percent of the forested plots are randomly selected for swapping with plots from the remaining data for a total swapping of between 0 and 20 percent. The primary criterion for swapping is based on a measure of ecological similarity. Plots with the smallest ecological difference are swapped. The variables for swapping—e.g., x and y coordinates, forest type group, and stand size—vary by region. This induces enough uncertainty as to the actual property owner to satisfy the legal requirements without introducing an unacceptable amount of error in the population estimates computed for analyses.

What are the impacts of fuzzing and swapping on analyses? In general, any analysis that requires computation of population estimates using entire counties will be completely unaffected. By definition, swapping is limited to plots in the same county. Therefore, when all FIA plots in one or multiple counties are used to compute population estimates, all the data are used. Because population estimates at subcounty scale already have a relatively high mean square error due to the small number of plots, the error contribution of swapping is likely to be small in

comparison to the error due to the small sample size.[5] No other data for the plot are swapped other than the plot coordinates. Therefore, all the other relationships within and among the variables for the plot are retained.

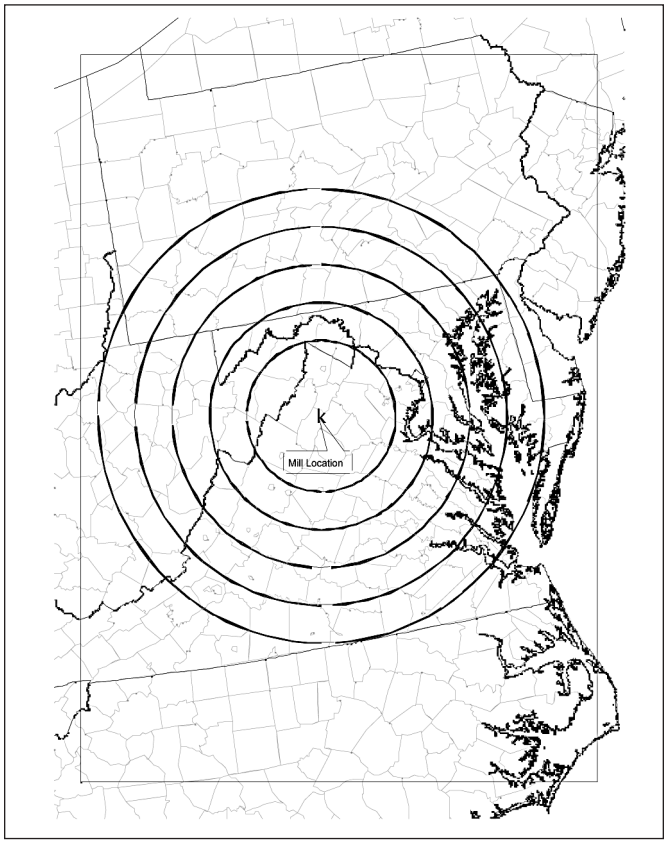
Data that have been fuzzed and swapped are not suitable for geospatial analyses in which FIA data are used to validate pixel classifications derived from satellite imagery. For this type of validation work, actual plot coordinates are required to properly register the ground data with the imagery. To serve such needs, the FIA program has created National FIA Spatial Data Services at the Northeastern Research Station and regional centers at the other four FIA units. The centers will populate data layers or prepare derived products using actual coordinates for FIA clients. Routine requests are fulfilled in several days or up to several weeks. More complex requests take longer. For especially complex and intensive data modeling, mapping, or analyses beyond the scope of Spatial Data Services, two additional options exist. Clients can visit a Spatial Data Services center to work directly with the data under the supervision of FIA staff, or FIA partners (those furthering FIA's mission) can negotiate a confidentiality agreement with the FIA program to use data for a specific purpose and time.

Some questions continue to persist about the use of the data in the FIA public database after fuzzing and swapping according to the new policy. Allegations and assertions have been made that the public database is useless, and that analyses, such as mill studies, cannot be reliably made from the public database. For this article, two special studies were conducted to test the hypothesis that the public database will produce results that are substantially different than if the actual plot locations were used.

Mill Site Case Study

A common data request to FIA from private industry is to calculate woodshed information within a specified geographic distance of a proposed mill. A prospective mill site was selected in the Shenandoah Valley of Virginia, (fig. 1). Woodsheds of five different radii—50, 75, 100, 125, and 150 miles—were evaluated to determine the acreages of forest and timberland, the numbers of live trees 1 inch and greater in diameter at breast height (d.b.h.), the number of growing-stock (GS) trees 5 inches and greater in d.b.h., and the cubic foot and board foot volumes (gross and per acre) of timber in the prospective woodsheds. Table 1 shows the results.

Figure 1.—*The five concentric circles surrounding the proposed mill location correspond to radii of 50, 75, 100, 125, and 150 miles, respectively. The circles cross State boundaries and different FIA regions, which may have used different variables for swapping. Each circle includes both complete and partial counties.*



[5] If a volume estimate is wanted for a 30,000-acre tract, for example, using data from the FIA database is likely to only yield four or five plots within that polygon. Computing population estimates for the tract based on so few plots will probably yield estimates with large mean square errors; if one of those plots happens to be swapped, the impact may be noticeable. Rather than bemoaning the potential impact of having a swapped plot in the population, however, the more important question is why one would be willing to impute total volume on the tract on the basis of such a limited sample of plots in the first place. Polygons for analysis should be large enough to yield 30 to 40 forested plots—200,000 to 250,000 forested acres—before population estimates and mean square errors are used to impute population totals.

Table 1.—*Comparison of actual versus fuzzed and swapped summaries for several radii (50, 75, 100, and 150 miles).*

Retrieval type	Area (1,000s of acres)			Trees (1,000s)		Volume (millions)		Volume/acre	
	Total	Forest	Timber-land	Live (≥ 1")	GS[a] (≥ 5")	ft³	Bd. Ft.	ft³	Bd. Ft.
				50-mi radius					
Actual	5,027	2,763	2,532	1,273,234	301,828	4,315	14,594	1,704	5,764
Fuzz/swap		2,777	2,542	1,277,453	301,729	4,322	14,632	1,700	5,756
				75-mi radius					
Actual	11,310	6,659	6,181	3,459,449	807,015	11,163	36,155	1,806	5,849
Fuzz/swap		6,639	6,161	3,424,786	802,263	11,085	35,907	1,799	5,828
				100-mi radius					
Actual	20,106	11,848	11,224	6,576,872	1,502,825	20,477	65,605	1,824	5,845
Fuzz/swap		11,850	11,223	6,570,391	1,501,532	20,496	65,773	1,826	5,861
				125-mi radius					
Actual	31,416	18,228	17,487	10,569,005	2,355,062	32,069	102,361	1,834	5,854
Fuzz/swap		18,223	17,487	10,536,468	2,350,409	32,011	102,194	1,831	5,844
				150-mi radius					
Actual	45,239	25,538	24,682	15,158,820	3,359,473	44,773	141,290	1,814	5,724
Fuzz/swap		25,568	24,722	15,176,113	3,361,850	44,780	141,288	1,811	5,715

[a] GS = growing stock.

The results show that the differences between using actual versus fuzzed and swapped plot locations are trivial for this mill site study. The differences are due to fuzzed and swapped plots in partial counties within the radii. Because the results are consistent across potential woodsheds of five different radii and for all the major variables that are normally a part of a mill site feasibility study, the trivial nature of the differences between the two sets of data do not appear to be an aberration.

Biomass Map Case Study

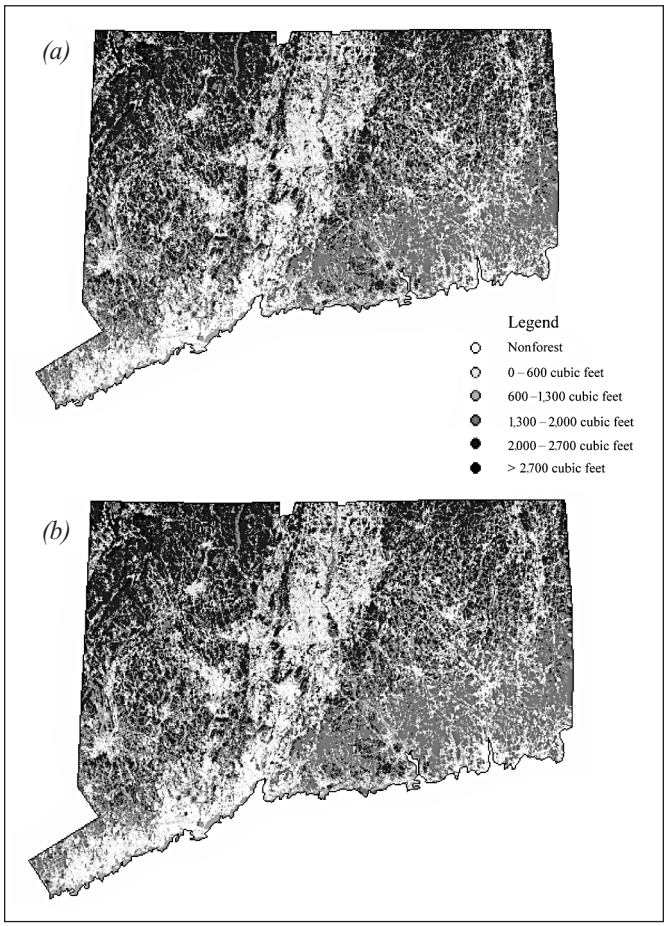
A spatial analysis was conducted to test the usefulness of the fuzzed and swapped data for producing models of mapped attributes. Using satellite imagery to model and map attributes of forests has become very popular. The spatial resolution of orbiting sensors varies significantly from less than 10 m (IKONOS, QuickBird) to 1,000 m (Advanced Very High Resolution Radiometer [AVHRR]). Several mid-resolution sensors, such as Landsat (30 m) and Moderate Resolution Imaging Spectroradiometer (MODIS) (250–500 m), offer especially useful resolution for studying forest attributes over wide areas. The twin challenges to using digital data from satellite sensors are (1) developing models to classify the images by individual pixels or groups of pixels, and (2) using ground-based data to validate the classification models. The former challenge is relatively easy. The latter challenge is more difficult. FIA plot data are among the best ground-based sets of data that can be used to validate classification models. For example, the interagency team developing the Multi-Resource Land Cover classification models relies on FIA plot data for validation of forest cover type models. FIA data also were key to developing the forest cover type map contained in the National Atlas (www.nationalatlas.gov/mld/foresti.html) and is based on a 1991 map using AVHRR imagery.

The hypothesis tested by this case study is that the publicly available plot data in the FIA database will yield model validation results no different from actual plot data. This hypothesis was tested by building a forest biomass map for Connecticut. The map was created by modeling above-ground forest biomass using road densities, satellite data, and the x and y coordinates of FIA plot data. The road density from the plot center within varying pixel radii was computed using 1:100,000 Tiger/Line road data from the U.S. Census Bureau. Two independent variables derived from satellite imagery were included in the

model. The first satellite variable was created from a reclassified forest/nonforest map acquired from a National Land Cover Data set. The second satellite variable was calculated from six of the seven Landsat Thematic Mapper bands using the tasseled cap transformation for greenness.

The R^2 for both the actual and the publicly available coordinates was 0.43. The intercepts accounted for the largest difference in the coefficients of the two models. The standard errors of the coefficients for two models were equivalent. Figures 2a and 2b reveal the similarity of the maps resulting from the model with exact coordinates and the model with the publicly available coordinates.

Figure 2.—*Biomass maps for Connecticut using (a) actual plot coordinates and (b) publicly available plot coordinates. Although the two maps appear identical, the predicted values do not always fall in the same cubic-foot volume class.*



What Can Be Inferred From the Two Case Studies?

In developing the approach to satisfy privacy concerns, FIA statisticians consulted with counterparts in several other agencies and organizations that are responsible for programs that rely on sampling. During those consultations, two points were consistently made by the other agencies and organizations: (1) swapping did not induce significant deterioration in the quality of their program's data to the detriment of their clients; and (2) nothing indicated that nontrivial differences would emerge from the algorithms adopted by FIA. The peer review that was conducted before issuing the new privacy also uncovered no problems with the approach proposed by FIA. These results bear out the wisdom of the advice FIA received and the quality of the internal testing performed before issuing the new privacy policy.

The results of the mill site case study dispel recent allegations that the approach taken by FIA to protect the privacy of partners seriously damaged the usefulness of the publicly available FIA data. The database can be used with confidence for projects such as mill site studies and woodshed analyses.

The results of the biomass map case study suggest that using the new publicly available data does not compromise the ability to model attributes, at least when using imagery of 30 m or larger resolution. The fourfold reduction in fuzz in the new public database appears to have improved the utility of the data for geospatial analyses. The "noise" in the model data set due to Global Positioning System errors and the georeferencing of Landsat pixels is probably similar to the noise induced by the fuzzing and swapping algorithms.

The FIA program is willing to conduct additional case studies, both in other regions and for different types of studies, to demonstrate the utility of the publicly available FIA data. Perhaps different results will be obtained for mill site studies in regions that are substantially different from the conditions existing in this study area (e.g., 50- to 60-percent forest cover; 90 to 95 percent of forest land being timberland; growing stock between 20 and 25 percent of total stocking; and per acre volumes of 5,500 to 6,000 board feet).[6] One reason this case study region was selected is that its conditions are very similar to

[6] Individuals wanting to partner with the FIA program in conducting additional tests of the publicly available FIA data for projects such as mill site studies should contact the author directly. The only caveat the author requests is that the results of the tests be published in a future FIA Science Symposium or equivalent outlet.

many other areas across the United States. Consequently, the results reported herein are believed to be broadly applicable. An unusual suite of conditions may lead to different results, however, and the FIA program is interested in exploring the utility of the publicly available data set in such conditions.

The biomass map case study sheds little light on the usefulness of the public FIA database for modeling with imagery of higher resolution (10 m or less).We believe, however, that locational errors likely will be of similar magnitude. It may be that actual plot coordinates are needed to work with high-resolution images over large areas, assuming that very accurate rectification of images occurs. The FIA program is very interested in conducting tests with partners using high-resolution imagery.

As more experience is gained with the publicly available database, which kinds of analyses and which spatial scales may require the use of actual coordinates will become clearer. National FIA Spatial Data Services at the Northeastern Research Station is available to assist with special needs. Over the past year, the center has consistently enabled clients to meet their deadlines, even when clients had only 3 or 4 months to complete their analytical work. Clients with short deadlines are encouraged to make those known in their initial contact with Spatial Data Services at www.fs.fed.us/ne/fia/spatial/index.html.

FIA Program Plans

Looking Ahead to 2005 and 2006

The FIA program has adopted two slogans for FY 2005 and FY 2006: "Lose No Ground" and "No State Left Behind." As additional States have been added to the annualized inventory program, FIA Program Managers have focused on postponing annualized inventory work if reasonable assurances do not exist that the annualized work can be continued in subsequent years—"Lose No Ground." Adding new States has been neither simple nor easy. Hiring, training, and retaining field crew members have been difficult at some stations and in some States. Consequently, full panels of data collection have not always been collected in a single field season. Further, during bad fire seasons when some State and Federal crew members are pulled away from FIA duties to fight wildfires, all the fieldwork anticipated has not been performed. When some fieldwork has been carried over to the next field season, known as "panel creep," Program Managers working with State forestry agency partners have worked hard to gain efficiencies and eliminate the fieldwork backlog. Recent experience shows that panel creep can be reduced and eliminated. Completing a full panel of fieldwork each season and maintaining annualized inventories in States in which operations have begun are essential components of "Lose No Ground."

In the early years of transition to annualized FIA, States were added to the program based on their willingness and ability to partner as test cases with stations. One reason that funding for the program has increased annually over the past 5 years is because Congress has seen increases in cost sharing and in-kind contributions from partners. As the program has grown and more States have been added, concerns have grown among the decreasing pool of States not yet annualized that they will be left behind with only an occasional periodic inventory. This concern has grown as the overall Federal budget has tightened to fund war and homeland-security-related needs. The long-term success of the FIA program is contingent on having all States included in the annualized program. The USDA Forest Service is committed to seeking the funds needed to implement annualized inventories in all States—"No State Left Behind."

The FIA program is on a trajectory to achieve the funding level and coverage of all 50 States outlined in the 1998 FIA Strategic Plan. The plan has been very helpful for keeping program leaders, others inside the USDA Forest Service, and all FIA partners and clients focused on the shared goals and objectives for FIA. But now that achievement of the 1998 plan's goals and objectives is imminent, the time to begin preparing a new FIA Strategic Plan is now.

Looking Ahead to 2007–12

A new FIA Strategic Plan should aim at two broad goals. The first and primary goal is to consolidate the programmatic gains achieved in 1998–2006. This means keeping the annualized inventory program working smoothly in all 50 States and territories, meeting all deadlines for releasing compilations of annual data, and producing integrated analytical reports at 5-year intervals.

The second goal is to create additional value for clients by augmenting the existing program with new features or levels of intensity. Some features may involve expanding coverage to fill critical information gaps and increasing spatial or temporal plot intensity. Others may involve creating new analyses and innovative uses of existing FIA data. In short, these twin goals emphasize better serving our core clientele by tuning up the existing program to make it better, faster, and cheaper, and adding a few carefully chosen new features to meet the most pressing emerging needs. Taken together, these actions will increase the value of the FIA program to a broader clientele.

To achieve the first goal, current business practices may need to change. For example, during the transition from periodic to annualized inventories, the task of writing data compilation software was undertaken both communally and individually. Station experts worked together to build common routines for collecting and compiling annualized field data. To compare the current situation with the most recent previous inventory (which consisted of periodic inventories of different types and ages, even within a single State and station), however, each station worked individually to build the routines to compare the past data with the present. When implementing annualized FIA in Kentucky, for example, the Southern Research Station had to work with data from the previous inventory created by the Northeastern Research Station using a different sampling scheme. Now that this transition is largely complete and a more consistent and common set of routines for compiling annualized FIA data exists, gaining additional efficiencies by further centralizing some of the data processing and compilation business processes may be possible. Additional advances from improving or reengineering other business practices may also be possible and should be explored. The improved internal discipline arising from having cross-station consistency as a core value for the new FIA program will make this change management task easier. Savings will be reinvested in improving the FIA program.

To achieve the second goal, an extended dialog is needed. Core clients remain an important source of support. A large number of potential clients need to be listened to, also. The following ideas have been presented in the past 2 years:

- Reduce the inventory cycles to 5 years everywhere.
- Move into urban areas to inventory all land.
- Broaden coverage of health issues, in particular to better characterize the impact of invasive species.
- Expand analyses to provide more information on issues for which forest vegetation structure and function are important influences, such as wildlife habitat quality.
- Intensify the grid on public and private land so that inventory data are more useful to resource managers at the forest management unit level.
- Expand the inventory to include rangeland and help characterize range condition, health, and trends.
- Broaden coverage of linear features, such as riparian zones, and link that data to the FIA grid.

Other concepts, including the following, have been expressed in less concrete terms:

- Take greater advantage of advanced satellite imagery, perhaps shifting some attributes from field plot collection as imagery provides more and more usable data.
- Develop faster, more accurate, more sophisticated change detection algorithms resulting in more timely and accurate assessments of land use/land cover dynamics.
- Develop more and better Web-based tools, such as adding more geospatial tools, so that users can customize their own analyses and make it easier for clients to build and populate their own unique data layers.
- Create better linkages between FIA data and pressing natural resource policy issues, such as fire and fuels analyses, fragmentation, sustainability reporting, certification, and other global issues.
- Build closer ties to the university community to take advantage of expertise for analyses and prototyping innovative techniques.
- Expand the use of FIA data in other parts of the USDA Forest Service for additional research and development and more State and Private Forestry programs and to improve the management of national forests and national grasslands.

All these suggestions have merit and would produce useful information for clients. All these suggestions also would

increase program costs. None can be launched solely on cost savings wrung out of improved efficiencies in the ongoing program. All would require changes—sometimes radical changes—in the existing program to be accommodated. But after the past 6 years, the FIA program has come to embrace change as invigorating and as the only route to future success.

Over the next 18 months, the FIA National Program Leader and station FIA Program Managers will convene and facilitate a series of discussions with current and potential clients and USDA Forest Service leadership over the future foci of the FIA program for 2007–12. Regional and national user group meetings this year and next will be used to launch the dialog, and additional meetings will probably be needed to ensure that all points of view are heard and to build consensus on strategies. Moving the FIA program forward will be impossible without strong support among the entire FIA community of interest.

Summary

The FIA program would not be in its current position without the strong, dedicated leadership of USDA Forest Service Research and Development employees and our partners. Together, we have accomplished great things in the past 5 years. Now is not the time to rest on our laurels or foreswear further change. Rather, FIA's continued science mission is to help clients—all Americans—see their forests in new and different ways; to wisely protect, manage, and use them; and to leave them in better condition for our children and grandchildren. We cannot accomplish this mission alone, but we can accomplish it together.