
Access and Use of FIA Data Through FIA Spatial Data Services

Elizabeth LaPoint¹

Abstract.— Forest Inventory and Analysis (FIA) Spatial Data Services (SDS) was established in May 2002 to facilitate outside access to FIA data and allow use of georeferenced plot data while protecting the confidentiality of plot locations. Modification of the Food Security Act of 1985 legislated the protection of information on plot location and ownership. Penalties were put in place for violations. Because of this change in the law, many customers have been served by FIA SDS, and demand for spatial analyses continues to grow. More than 130 requests for spatial data or other information have been received from academia, State and local governments, Federal agencies, and forest industry. This article describes how projects progress from inception to completion, including security concerns.

Introduction

Forest Inventory and Analysis (FIA) has been collecting data and reporting on the status of the Nation's forests for more than 70 years. Sample plots are located across the landscape and are revisited periodically. In recent years, plot locations have been recorded using Global Positioning System (GPS) technology, which has resulted in new uses for FIA data.

The Privacy Issue

With passage of the fiscal year (FY) 2000 Consolidated Appropriations Bill (Public Law 106-113), the Food Security Act of 1985 (7 U.S.C. 2276(d)) was modified, making it illegal to reveal information on FIA plot locations or ownership. Penalties for violating the law can include fines up to \$10,000 and/or a year in jail.

Even before the change in the law, FIA treated ownership information and plot location in a confidential manner for the following reasons: (1) to maintain goodwill because FIA relies on private landowners for access to their property, and they must understand that FIA is not concerned with any regulatory or taxation issues; (2) to eliminate unnecessary site visits that might damage or alter the plot; and (3) to ensure management decisions are not influenced by knowledge of a plot's location.

Background

FIA Spatial Data Services (SDS) was created in May 2002 to assist customers in accessing and using FIA data spatially within the bounds of existing legislation. FIA SDS also provides assistance with Geographic Information System (GIS) technology and in linking to non-FIA data, in addition to attempting to answer users' questions about the data. Although FIA SDS was established partly in response to the change in the privacy law, FIA's national management realized that users needed a single contact point for questions related to the FIA data confidentiality.

Almost one-third, 32 percent, of the requests for data have been from academia. Other consumers include other Federal agencies such as the Environmental Protection Agency and Bureau of Land Management as well as non-FIA U.S. Department of Agriculture Forest Service employees and State agencies. Academia and Federal agencies account for 63 percent of the data requests received in FY 2003.

Data Request Process

After customers initially contact FIA SDS, the data request is typically revised or refined for some period of time. Often, requesters are unaware of the kinds of data collected by FIA, how these data are collected, or the differences in data collection between States and FIA units and/or privacy issues related to the use of FIA data. Protecting the location of the plots and landowner privacy are the primary concerns of FIA SDS with

¹ Forester, Northeastern Research Station, U.S. Department of Agriculture, Forest Service, Newtown Square, PA 19073. Phone: 610-557-4049; fax: 610-557-4250; e-mail: elapoint@fs.fed.us.

every data request—for example, ensuring that the data provided cannot be back-engineered to reveal plot location or ownership information.

As part of refining a request, FIA SDS describes the relevant data to the requester with respect to the specific area of interest. Revising a data request, whether for security purposes or other reasons, can take weeks to months depending on the request's complexity, the requester's motivation level, the extent of communication between FIA SDS and the requester, and the work backlog at FIA SDS.

FIA data limitations are described while discussing the data request with the requester; for example, the variations in the quality of plot location data are pointed out. Originally, plots were georeferenced by digitizing their positions on aerial photographs or maps. These digitized positions are rarely as accurate as the GPS coordinates currently being collected for sample plots. Unfortunately, plot locations with no GPS coordinates still exist. GPS-derived plot locations also are subject to inaccuracies due to user error, signal degradation, or satellite configuration. Some of these problems should be resolved with the national standardization of GPS data collection methods and increased recognition of the importance of accurate locations.

Another limitation is that some areas of the Nation have few, if any, FIA plots. In the past, plots were not always established in wilderness areas or areas with no timberland, for example, western Texas and parts of California. Some areas, such as interior Alaska, had no plots established due to their remoteness.

Another issue that may affect a data request is that inventories in neighboring States may have occurred as many as 10 years apart, and, therefore, the data requested may not be current. However, FIA's recent transition to an annual inventory will result in more timely information.

After the data request has been refined and meets security requirements, the request is forwarded to the appropriate FIA program manager(s) for approval. Often, the security review by FIA SDS entails processing a portion of the data request and examining the results. Each FIA unit can refuse a data request or recommend additional refinement, use its own staff to process the request, or approve the request and redirect it to SDS for action.

When a request is approved, SDS completes its processing and forwards the results to the program manager(s) for final review. If the release of the data is approved, the results are

forwarded to the customer. If a request is denied, SDS will work with the customer to accept an alternative that best meets the customer's needs.

Revisions to the data request and communications among the requester, FIA SDS, and FIA management make up most of the time spent on a request. As data requests become more complex, the amount of time required to supply the data also will increase. Data requests that cover multiple States or FIA units can become more complicated due to differences in the data between States or units.

During FY 2003, 95 data requests were received, and an average of 5.5 requests were filled each month. Almost from its inception, FIA SDS has experienced a work backlog due to needed refinements in requests and/or communication time lags between SDS and the customer. In FY 2003, 67 percent of the requests received were filled within 4 weeks; the overall average fulfillment time was 7 weeks.

A customer can work with FIA SDS at its office at Newtown Square, PA, or submit a request at any regional FIA location.

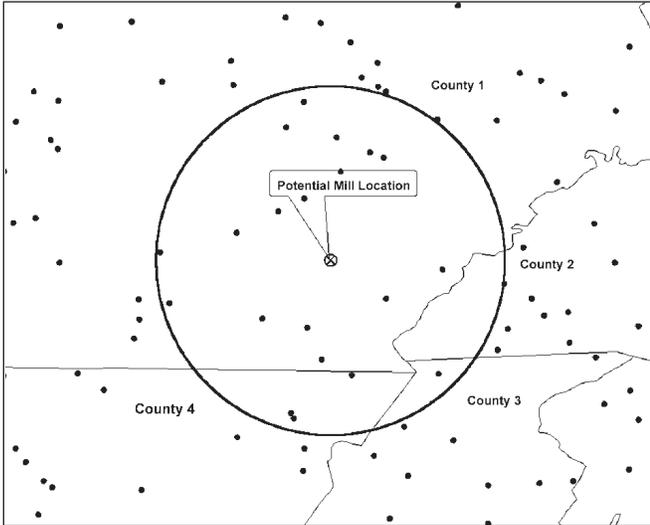
Examples of Requests

Simple Example

A common request is for a data summary for plots that fall within the requester's area of interest, for example, a summary of growth and removals within a given distance of a mill location. Because the data provided to the requester is summarized, no concerns exist related to disclosing information on plot location or ownership. The only security concerns are ensuring that the area of interest is of sufficient size to prevent plot location disclosure, and that the area of interest covers at least three private landowners. Current restrictions on FIA Mapmaker (http://ncrs2.fs.fed.us/4801/fiadb/fim_tab/wc_fim_tab.asp), an Internet-based data query and mapping application, prohibit circular retrieval with a radius less than 25 miles (1,256,637 acres). Generally, tables are created only if at least 12 forested plots fall within the area of interest; i.e., a circle with a minimum of 72,000 forested acres.

Creating tables of growth and removals at the county level can create security problems. Figure 1 depicts circular retrieval on a map containing county lines. If summarized data are provided

Figure 1.—Fifty-mile-radius circle retrieval around potential mill location. The single plot in lower right quadrant could reveal location or ownership information if county level summaries are provided. (Plots shown are fictitious.)



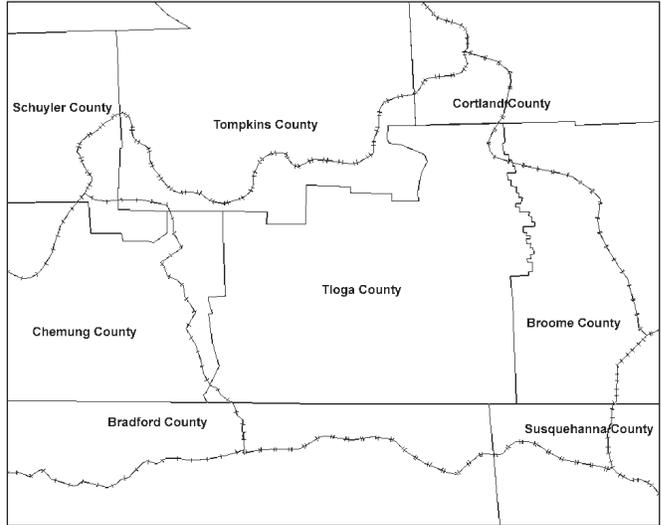
for the entire circle, the security issue is moot. If the customer wants county-level summaries, however, a security issue arises because overlaying the circular area onto the county layer creates a sliver polygon. The sliver may be small enough to locate the plot on the ground, or the owner of that area may be known. FIA SDS assumes that the customer has access to GIS technology and is able to overlay the circle onto the county layer.

Complex Example

A more complex request is one that entails overlaying plot locations onto a polygon layer to associate a polygon label or attribute with each plot—for example, the customer wants to assign a soil polygon to each FIA plot. The ideal solution is to provide summarized data for polygons and restrict the data to polygons with a minimum of three plots, which avoids creating sliver polygons. Figure 2 shows how overlaying a watershed layer onto a county layer can create polygons that would reveal plot location or ownership.

A possible solution to the security problem is to ensure that at least three privately owned plots exist in each polygon, although the three plots may be owned by the same person. Another solution is to make the polygon larger than common land holdings in the area of interest. This minimum size varies greatly between regions. For example, a polygon that encompasses

Figure 2.—Overlaying a watershed layer onto a county layer can create sliver polygons that can reveal plot location or ownership information; i.e., the southeast corner of Schuyler County and the northwest corner of Susquehanna County.



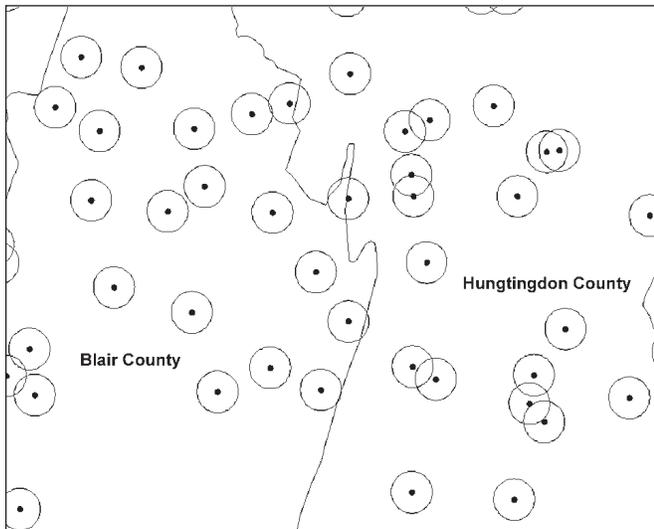
100 acres on the coast of Rhode Island likely will contain the required three private owners, but that same parcel size may be woefully inadequate in north-central Maine. Ideally, FIA SDS would provide data summarized for the polygons rather than provide the polygon ID for each FIA plot.

FIA SDS also handles requests involving remotely sensed imagery or other raster data. These are often data requesters sending in satellite imagery they have previously classified, SDS overlays the FIA plots and then returns information on the classes of the imagery. These requests are similar to the previous example with soil polygons. Providing summarized information for each class, or ensuring both a minimum of three privately owned plots in each class/county combination and that each class/county combination covers enough acreage to prevent disclosure of plot locations, avoids security issues.

Fuzzing and Swapping

To make using spatial data more accessible to customers, the FIA database has coordinates available for downloading. The coordinates associated with the plots are altered to protect ownership information and prevent locating plots on the ground.

Figure 3.—Plots with a 1-mile-radius circle depict the area from which a new “fuzzed location” would be randomly selected. (Plots shown are fictitious.)



The process of altering the plots entails “fuzzing” the plot locations and then swapping coordinates for a certain percentage of plots. Fuzzing involves creating a buffer area of 0.5 to 1.0 miles around each plot and randomly selecting a point within that circle as the “new” coordinate for that plot (fig. 3). This procedure is performed for all plots to prevent users from locating the true plot locations.

To protect landowner privacy, the location coordinates of up to 20 percent of privately owned forested plots are swapped with similar plots in the same county or supercounty (aggregation of two or more adjacent counties in the same State). For example, if plots A and B are selected as a swapping pair, plot A’s data will be assigned to plot B’s location, and plot B’s data will be assigned to plot A’s location. The plots are determined to be similar based on criteria established by each FIA unit. Plots usually will not be swapped outside their county.

Customers can download the fuzzed and swapped data and perform their own spatial analyses. In some cases, data requests are not significantly affected by using fuzzed and swapped coordinates, for example, when evaluating mill locations. In others, customers may want to use the fuzzed and swapped coordinates to examine and refine their data needs before submitting a data request.

Summary

FIA SDS has made great strides in improving access to FIA’s spatial data. FIA SDS was begun as a pilot project; since then, however, the need for this service has been demonstrated. By establishing FIA SDS, FIA has shown a commitment to its customers. As FIA’s commitment to spatial products continues to grow, so will the variety of spatial tools and spatial data available to FIA customers.