Long Database Report

Forest Inventory and Analysis Database of the United States of America (FIA)

Andrew N. Gray, Thomas J. Brandeis, John D. Shaw, William H. McWilliams & Patrick D. Miles

Abstract: Extensive vegetation inventories established with a probabilistic design are an indispensable tool in describing distributions of species and community types and detecting changes in composition in response to climate or other drivers. The Forest Inventory and Analysis Program measures vegetation in permanent plots on forested lands across the United States of America (GIVD ID NA-US-001). Plot sizes and protocols for measuring tree species are standardized across the country. Additional standardized protocols have been implemented to measure the abundance of non-tree vascular plant and epiphytic lichen species. Research using this and related regional datasets have provided new insights into the key biophysical drivers of community composition and their importance at different spatial scales. Studies have also explored regional differences in species diversity patterns, documented the importance of non-native species, and described the importance of environment and management on the distribution of selected species. Although representation of locally rare community types may be low, the probabilistic sample ensures that ecological drivers are regionally significant and that results are representative of a region as a whole. Rereasurement of permanent plots provides direct evidence of vegetation change and enables detection of impacts due to climate, natural disturbance, and forest management.

Keywords: climate change; diversity; long-term dataset; permanent plot; species distribution; tree growth rate.


Abbreviations: FIA = Forest Inventory and Analysis program; NRCS = Natural Resources Conservation Service; USDA = United States Department of Agriculture.

Received: 4 October 2010 – Accepted: 19 September 2011 – Co-ordinating Editor: Robert K. Peet.

Introduction

Databases of vegetation observations collected over large areas and long time periods have been and will continue to be essential for describing changes in species composition in response to climate or other drivers. These databases are often assembled from a variety of original sources created for a variety of objectives (e.g., U.S. Long-Term Ecological Research network, Smithsonian Tropical Research Institute). The key attribute for these datasets is that each plot sample is a complete enumeration of the species present in the community, and that the sample contained in the plot is representative of the plant community. In much of Europe, this type of vegetation sampling follows the relevé approach, where plots are subjectively placed in representative locations of stands, and appropriate plot sizes are selected after creating species-area curves (Braun-Blanquet 1932).

Although forest inventories are usually initiated to provide estimates of the abundance of natural resource commodities, they have proven to be valuable for describing distributions of species and community types and quantifying changes over time. To avoid potential sources of bias, forest inventories are usually established with a probabilistic design so that any point on the landscape has a quantifiable probability of being sampled with a plot. As a result, plot sizes are usually the same for all vegetation types, and some plots may straddle multiple plant communities. Fixed sampling intensities can result in few samples of rare community types, although intensification of plot samples in well-defined zones is sometimes used to reduce that problem.

The objective of this paper is to describe a comprehensive database of the forests of the United States of America created and maintained by the Forest Inventory and Analysis (FIA) program of the United States Department of Agriculture (USDA) Forest Service. The FIA program is the most comprehensive and consistent inventory and monitoring program in the country (National Research Council Committee to Evaluate Indicators for Monitoring Aquatic and Terrestrial Environments 2000). Permanent field sample points have been installed on forest lands across all ownerships and measurements taken periodically. The focus of the inventory and of this paper is on tree species. However, data on all vascular plants (Schulz & Dobelbower 2012) and on epiphytic lichens (Jovan 2012) are also being collected on a subset of plots and will be briefly described as well.
### Inventory methods

Inventory methods have evolved since the first nationally-mandated inventories of the 1930s (Frayer and Furnival 1999). Early efforts were based on maps of forest type and condition. Beginning in the 1960s in most of the country, systematically-placed grid points (with either random starting corners for the grids or random off-sets for each plot) were installed as permanent sample points. Ten “subplot” sample points were established around each grid point and trees were selected for measurement using point sampling (Bitterlich 1948), where selection probability increased with tree diameter (LaBau et al. 2007). The selection angle varied regionally to maximize efficiency, and fixed-area plots were often used for small trees as well. Initially, subplots were systematically moved from the original layout if needed to ensure all data pertained to the same stand. However, concerns about introducing bias by undersampling stand edges led to systematic establishment of subplots by the 1990s, regardless of differences in stand characteristics or land type within the plot footprint. By the 1990s, most regions were measuring five subplots per grid point. The current inventory design was adopted nationally in 2000 to provide more uniform and consistent methods across the country. The base grid of sample points was designed to ensure one plot in each 2,430 ha hexagon of a layer that spanned the nation (Bechtold and Patterson 2004). Additional plots have been installed with cooperation from states and private land owners to meet regional objectives (e.g., disease agents, technology, and not protected from harvesting). The plot grids extend across all lands and all ownerships, and are measured on a panel system where a geographically-distributed systematic subset of the grid is measured every year. A full cycle of measurement takes from five years in eastern states to 10 years in western states. States have been added to this “annualized” inventory system as funding has become available; as of 2010, the 48.6 million ha boreal forest of interior Alaska has yet to be included. The total area of forest land in the United States is estimated at 304 million ha (Smith et al. 2009), which results in 125,100 forested grid points at the standard density.

The FIA inventoried annual plot consists of a cluster of four sample points, with three of the points located 36.6 m from the central point at 0, 120, and 240 degrees azimuth (Fig. 1). Tree sampling at each point occurs on nested fixed-area plots. Tree “seedlings” (>15 cm tall and <2.5 cm in diameter at breast height [DBH, at 1.37 m height]) are counted by species on 13.5 m² (2.1 m radius) “microplots”. Tree “saplings” (DBH >2.5 and <12.7 cm) are individually measured on the same microplot. Trees >12.7 cm DBH are measured on 168 m² (7.3 m radius) “subplots”. The national plot design also includes an optional 1,012 m² (17.9 m radius) “macroplot” which has been used in some Pacific coastal states to more precisely sample large trees (>61 or >76 cm diameter, depending on location). In addition to live trees, standing dead trees are measured with the same plots and diameter criteria. The national manual of field procedures provides detailed instructions on these and additional measurements (USDA Forest Service 2007). In addition to the required core measurements, each FIA region has additional field measurements they collect to meet regional objectives (e.g., disease agents, understory vegetation, down wood).

The population and attributes of interest have also changed over time. Early inventories were focused on measuring commercial tree species on land that was productive, able to be harvested with existing technology, and not protected from harvest. The population of interest has increased with interest in conservation and management of additional resources and vegetation types. The current sample population is all “forest

---

**Forest Inventory and Analysis Database of the United States of America (FIA)**

**Scope:** Data on tree species are collected from lands qualifying as forested (>0.4 ha in size and >10% canopy cover) with permanent plots systematically located across the country. Online data for some areas dates back to 1966; all regions are currently represented except for boreal Alaska. Plot designs have changed over time; some older measurements are exact remeasurements. All measurements since 2000 use the same core design. Additional data on stand condition and structure are also collected.

**Status:** completed and continuing  
**Period:** 1966-2012

**Database manager(s):** Patrick Miles (pmmiles@fs.fed.us)  
**Owner:** USDA Forest Service Forest Inventory and Analysis National Program

**Web address:** http://flia.fs.fed.us/tools-data/default.asp

<table>
<thead>
<tr>
<th>GIVD Database ID:</th>
<th>NA-US-001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last update:</td>
<td>2012-07-19</td>
</tr>
</tbody>
</table>

**Availability:** free online  
**Database format(s):** MS Access, Oracle  

**Online upload:** no  
**Online search:** yes

**Environmental data:** altitude: 100%; slope aspect: 100%; slope inclination: 100%; surface cover other than plants (open soil, litter, bare rock etc.): 6%; soil pH: 6%; other soil attributes: 6%; land use categories: 100%

**Geographic localisation:** point coordinates less precise than GPS, up to 1 km: 95%; political units or only on a coarser scale (>10 km): 5%

**Plot type(s):** nested plots; time series  
**Plot-size range:** 14-1013 m²  
**Non-overlapping plots:** 538,428  
**Plot observations:** 2,097,177

**Countries:** US: 100.0%

**Forest:** 100% — Non-forest: aquatic: 0%; semi-aquatic: 0%; arctic-alpine: 0%; natural: 0%; semi-natural: 0%; anthropogenic: 0%

**Guilds:** all vascular plants: 1%; only trees and shrubs: 99%; lichens (terricolous or aquatic): 1%

**Estimate of existing plots:** 538,428  
**Completeness:** 100%

**Total plot observations:** 2,097,177  
**Number of sources:** 3  
**Valid taxa:** 465

**Analysis Database:** Database description and users manual version 4.0 for Phase 2.  
**Web address:** http://www.treesearch.fs.fed.us/pubs/37446

**Information as of 2012-07-21; further details and future updates available from http://www.givd.info/ID/NA-US-001**
land*: land areas ≥0.4 ha in size that support, or recently supported, 10 percent stocking (or 10% canopy cover) of any tree species and are not primarily managed for a non-forest land use. FIA plots are installed at every base grid location with forest land within the plot footprint that is accessible (i.e., permission granted by owners and not hazardous to sample).

Because the plot design is fixed around the sample grid location, plots can sample multiple land-use conditions, vegetation types, and stand age classes, termed “condition classes” which are classified in the field using a hierarchy of criteria. All collected data are identified to the condition class on which they were sampled.

One out of every 16 FIA grid locations (or one point per 38,800 ha) is also identified as a “forest health” plot, where additional national protocols have been developed to measure various “indicators” of forest condition and relative degradation from pollution or management. Two of the indicators that are of potential interest to vegetation researchers are the lichen community and vegetation diversity indicators (Woodall et al. 2010a). Abundance of epiphytic macrolichens is estimated in 36.6 m radius plots around plot center and samples collected for identification by specialists (Jovan 2012). Vascular plant species are recorded on three 1 m² quadrats within each subplot, and on the subplot as a whole (Schulz et al. 2009, Schulz & Dobelbower 2012).

Quality assurance methods for the FIA inventory include regular field crew training and certification, evaluation and scoring of collected data by quality assurance experts, and independent remeasurement of a subset of plots by different crews. Regular reports on data quality are used to improve training, adjust procedures, and report measurement error to data users. All data are currently collected on field computers, which minimize errors by prompting field crews to fill in missing values, and which contain substantial edit checks to ensure measurements and change information are consistent and logical. For example, warnings are produced for species records entered on plots outside the currently-known range for those species.

**Database description**

FIA data are freely available on-line at http://fia.fs.fed.us/tools-data/default.asp. All national core variables, and some regional variables, are populated. Databases can be downloaded for individual states which include example queries to create various data summaries, and online tools are available to conduct custom queries. Documentation is available that describes variables, relationships among tables, and sample data queries (Woudenber et al. 2010).

Tree species data are contained in tree and seedling tables, and contain calculated variables (e.g., volume, biomass) in addition to field data (e.g., subplot number, condition class, species, diameter, height, crown class). The condition table describes the land type (e.g., forest or nonforest), stand type (e.g., dominant species and tree size class), owner type, reserve status, and disturbance or management history for each condition class found on the plot at each measurement. The plot table contains information on plot location, measurement date, and inventory cycle.

The FIA program is prohibited from publicly releasing actual plot locations to protect the privacy of private landowners and the integrity of the plot sample. Plot locations in the database have been randomly offset up to 1.6 km from the actual location. In addition the data for a proportion of the private locations have been switched between pairs of plot locations having the same forest type and owner class (but not the same owner) within a county. Private owners are identified only by general category (e.g., industrial forestry owner or not), but more detail or agency names are available for plots managed by public agencies. Several options exist for researchers to use actual plot coordinates in analyses, as long as results cannot be associated with an individual private owner (http://fia.fs.fed.us/tools-data/spatial/default.asp).

As of summer 2010, there were more than 460,000 records for forested plots in the FIA database. Many of these records are remeasurements over time of the same plot locations. The earliest records in the database date back to 1966 (Fig. 2). Although most of the FIA regions have digital data dating to the 1980s, not all have converted their data to current national database standards to enable loading on the national website. Although data are available extending to the 1960s, much of it is in hardcopy files and data for only a few states have been digitized, reformatted, and loaded in the national database. Digitizing this historical data is underway.

As discussed under Inventory methods, much of the data collected prior to 2000 used point-sampling methods where tree selection probabilities varied within and among regions. Although the compilation of summary attributes (e.g., volume and biomass) can account for these differences, it is important to note that assessments of relative diversity and community composition can be affected by the differences in plot protocols. However, since 1999 through 2009, 123,821 forested plot locations had been measured at least once with the nationally-consistent annualized plot design.
Plate: Vegetation types featured by the vegetation-plot database GIVD NA-US-001.

A: *Pinus-Juniperus* forest in northern Arizona, March (Photo: M. North).

B: Mixed northern hardwoods forest in northern Maine, October (Photo: A. Gray).

C: Old-growth *Thuja-Pseudotsuga* forest in western Washington, September (Photo: A. Gray).
Applications of FIA databases

The FIA program periodically produces mandated state-level and national reports that summarize the status and trends of forest area, tree composition, biomass, resource use, and projections of future resource availability (e.g., Conner et al. 2009, Pugh et al. 2009, Campbell et al. 2010, DeBlander et al. 2010, Smith et al. 2009). The FIA database is also used in studies on a wide range of topics, including resource management alternatives, wildlife habitat characterization, disturbance and fragmentation effects, invasive species, and mensuration techniques (Gray & Azuma 2005, Blackard et al. 2008, Long & Shaw 2009, Oswalt & Oswalt 2010).

The primary interest of the Global Index of Vegetation Databases is on species distributions, floristics, and community composition. A substantial number of publications have been developed from FIA data on these topics in recent years. The distributed probabilistic FIA sample lends itself well to analyses of tree distributions in relation to geography, and climate (e.g., Rehfeldt et al. 2006, Evans & Cushman 2009). Empirical associations between climate and species distributions lend themselves well to statistical or physiology-based predictions of potential effects of future climate change (e.g., Iveson & Prasad 1998, McKenney et al. 2007, Coops et al. 2009). Gradient models of species, climate, and geography have been applied with nearest-neighbor imputation to spatially predict vegetation composition and structure over large regions (Ohmann & Gregory 2002), which has been of great interest to land managers. Comparison of distributions of adults and juveniles from FIA plots has also been applied to evaluate current trends in species migration (Woodall et al. 2009).

The complete sample of tree species on FIA plots also lends itself well to studies of community composition. Studies have examined patterns of species richness across the nation (Waring et al. 2006) and the niche overlap of species within regions (Rehfeldt et al. 2008). The potential drivers associated with different community assemblages have been examined in tree species-rich regions (Brandeis 2006, Brandeis et al. 2009) and for areas where data on non-tree woody species was collected (Ohmann & Spies 1998). Examination of tree species’ growth rates and potential growth rates in relation to conspecifics provided insights into competitive effects on the shape of species’ realized niches (Canham et al. 2006).

The ready availability of a vegetation database from nationally-consistent, long-term forest inventories in the United States have led to a recent increase in the application of FIA data to address ecological and resource management questions. Although inter-agency efforts have tested field and statistical techniques, comparable monitoring efforts have not materialized to date for other vegetation types in the United States. Maintaining a consistent network of forest monitoring plots will undoubtedly prove invaluable to understanding changes in species distributions and community composition in the future.

Acknowledgements

Thanks to the many dedicated field crews, state agencies, and university cooperators who have improved and implemented the inventory over the years.
References


Andrew N. Gray* (agray01@fs.fed.us)
Forestry Sciences Lab 3200 SW Jefferson Way
Corvallis, OR 97331, UNITED STATES

Thomas J. Brandeis (tjbrandeis@fs.fed.us)
FIA HQ 4700 Old Kingston Pike
Knoxville, TN 37919, UNITED STATES

John D. Shaw (jdshaw@fs.fed.us)
Forest Inventory & Analysis 507 25th Street
Ogden, UT 84401-2450, UNITED STATES

William H. McWilliams (wmcmilliams@fs.fed.us)
Northern Research Station 11 Campus Blvd., Suite 200
Newtown Square, PA 19073, UNITED STATES

Patrick D. Miles (pmiles@fs.fed.us)
Northern Research Station 1992 Folwell Ave
St. Paul, MN 55108, UNITED STATES

*Corresponding author