

---

## Sole: Online Analysis of Southern FIA Data

Michael P. Spinney, Paul C. Van Deusen, and Francis A. Roesch<sup>1</sup>

**Abstract.**—The Southern On Line Estimator (SOLE) is a flexible modular software program for analyzing U.S. Department of Agriculture Forest Service Forest Inventory and Analysis data. SOLE produces statistical tables, figures, maps, and portable document format reports based on user selected area and variables. SOLE's Java-based graphical user interface is easy to use, and its R-based analysis engine is fast and stable. Each of the program's components (data retrieval, statistical analysis, and output) can be modified individually. This adaptability encourages outside development of analysis algorithms.

### Introduction

The Forest Inventory and Analysis (FIA) program of the U.S. Department of Agriculture (USDA) Forest Service has been collecting forest inventory data since 1930. Database structure and inventory design have evolved to accommodate available technology and satisfy user demand. The most recent version of the FIA Database (FIADB), available at [http://www.ncrs2.fs.fed.us/4801/FIADB/fiadb\\_documentation/FIADB\\_DOCUMENTATION.htm](http://www.ncrs2.fs.fed.us/4801/FIADB/fiadb_documentation/FIADB_DOCUMENTATION.htm), includes new ecological variables and older variables with modified definitions. The size of the FIADB can be intimidating to some users, and the database structure can confuse others. Readily accessible, easy-to-use tools are necessary to ensure proper analysis of FIA data.

FIA data are analyzed periodically by the USDA Forest Service and published as a standard suite of tables for a State or region. FIA has recently converted from a periodic to an annual inventory system and now publishes reports every 5 years. This schedule

may not be satisfactory for users who require more up-to-date reports or who are interested in information for areas other than States or regions. Production of customized reports requires users to download data, replicate the relational database, and then correctly query and summarize the data. Not all users with special data needs are willing to invest this amount of effort.

Web-based FIA analysis tools simplify the production of customized reports by eliminating the need to understand the underlying database structure. FIA's MapMaker (<http://www.ncrs2.fs.fed.us/4801/fiadb/fim17/wcfim17.asp>) is the USDA Forest Service's official Web-based analysis tool that uses standardized FIA analysis algorithms. Another online tool, the Southern On Line Estimator (SOLE; <http://ncasi.uml.edu/SOLE>) has been developed cooperatively by the USDA Forest Service and the National Council for Air and Stream Improvement. SOLE, a Web-based tool for analyzing annual FIA data, has a Java-based graphical user interface and an R-based analysis engine that employs both standardized FIA and alternative algorithms.

### Sole Structure Overview

Flexibility has been the top priority during the development of SOLE's structure. Each of the three main components of the program (data retrieval, statistical analysis, and output) was developed to be independently modifiable, which facilitates outside development of analysis algorithms.

The user completes a query by sequentially selecting tabs at the top of the user interface. After choosing States in the **State Selection** window, the user progresses through the **SOLE Map, Variable Selection, Filters** (optional), **Analysis**, and/or **Mapping** tabs. The application progresses automatically as the user completes requirements associated with each tab. Detailed instructions can be found in the help files linked to each tab.

---

<sup>1</sup> Michael P. Spinney is a Quantitative Forestry Analyst, and Paul C. Van Deusen is a Principal Research Scientist, National Council for Air and Stream Improvement, 600 Suffolk Street, Fifth Floor, Lowell, MA 01854. E-mail: [mspinney@ncasi.org](mailto:mspinney@ncasi.org); Web site: <http://ncasi.uml.edu>. Francis A. Roesch is a Mathematical Statistician, U.S. Department of Agriculture, Forest Service, Southern Research Station, P.O. Box 2680, Asheville, NC 28802-2680.

## Using the Program

### State Selection

The application begins with the **State Selection** window. States for which annual data are available have black backgrounds, and States for which only periodic data are available are “grayed out” (i.e., unavailable for analysis). When a State is selected, its name is listed, and the State is highlighted with green. When done selecting States, the user clicks the orange bar at the bottom of the window to proceed to the **SOLE Map** tab, where individual counties may be selected (fig. 1).

### SOLE Map

Individual counties can be chosen with the **Select by County** button, or all counties in the State can be chosen with the **Select All Counties** button. Clicking on the **Retrieve Data** button loads the necessary data files into the statistical program, R (R Development Core Team 2003). After the data files have been loaded, the next group of tabs, **Variable Selection** and **Filters**, become available (fig. 2).

### Variable Selection

Currently, only FIADB volume variables are available for analysis. These are grouped in 2 ways: Quantitative (continuous data, i.e., any volume or biomass estimate), and Qualitative (categorical data, such as productivity class). In addition, Quantitative Variables are grouped based on diameter greater or less than 5 inches. The diameter breakdown preserves the precision of volume estimates by ensuring that no data for trees less than 5 inches diameter at breast height are included in merchantable volume estimates (fig. 2). All analyses require selection of one Quantitative and at least one Qualitative Variable.

### Filters

All data for the area of interest are analyzed unless analysis is restricted by an optional filter. The filters currently offered include stand size, stand origin, all live stocking, growing-stock stocking, specific forest type, site productivity class, physiographic class, ownership group, owner, reserved class, and measurement year. Users can customize their analyses by specifying particular levels of filters. Filter status is displayed at the top of each analysis.

Figure 1.—SOLE map selection tab, for county selection and data retrieval.

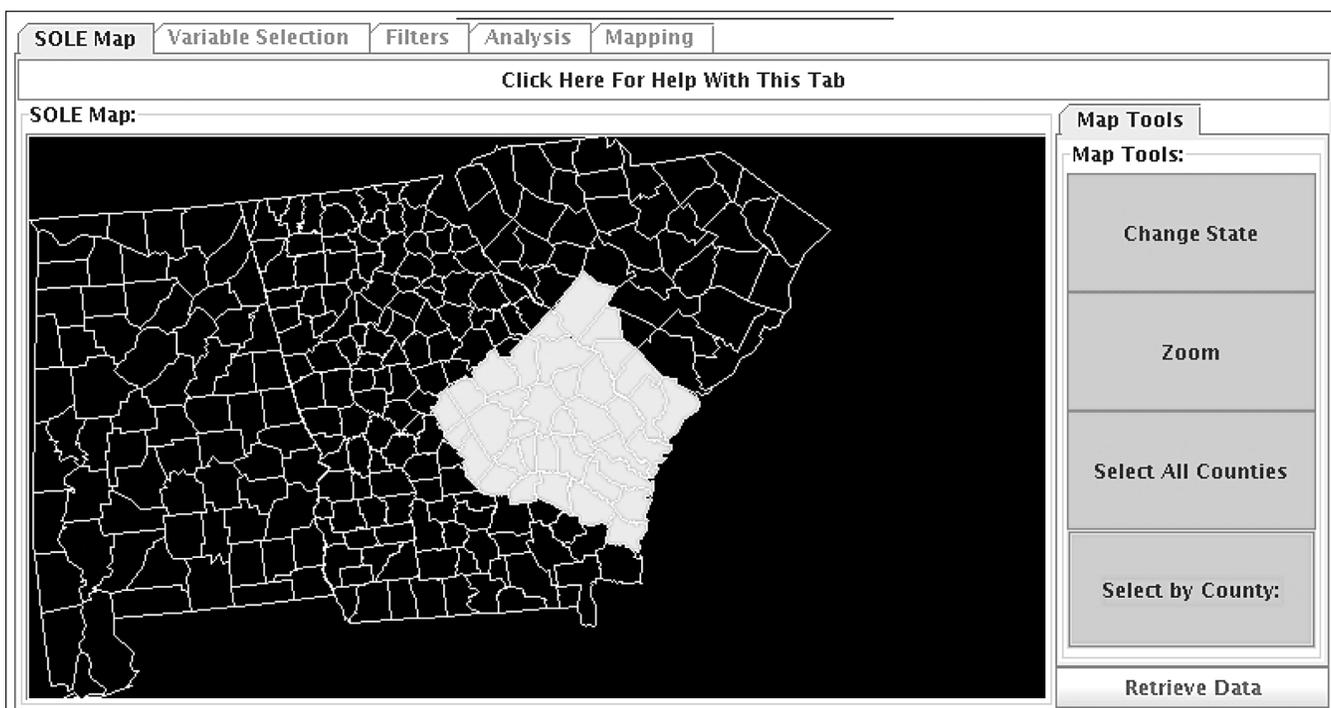
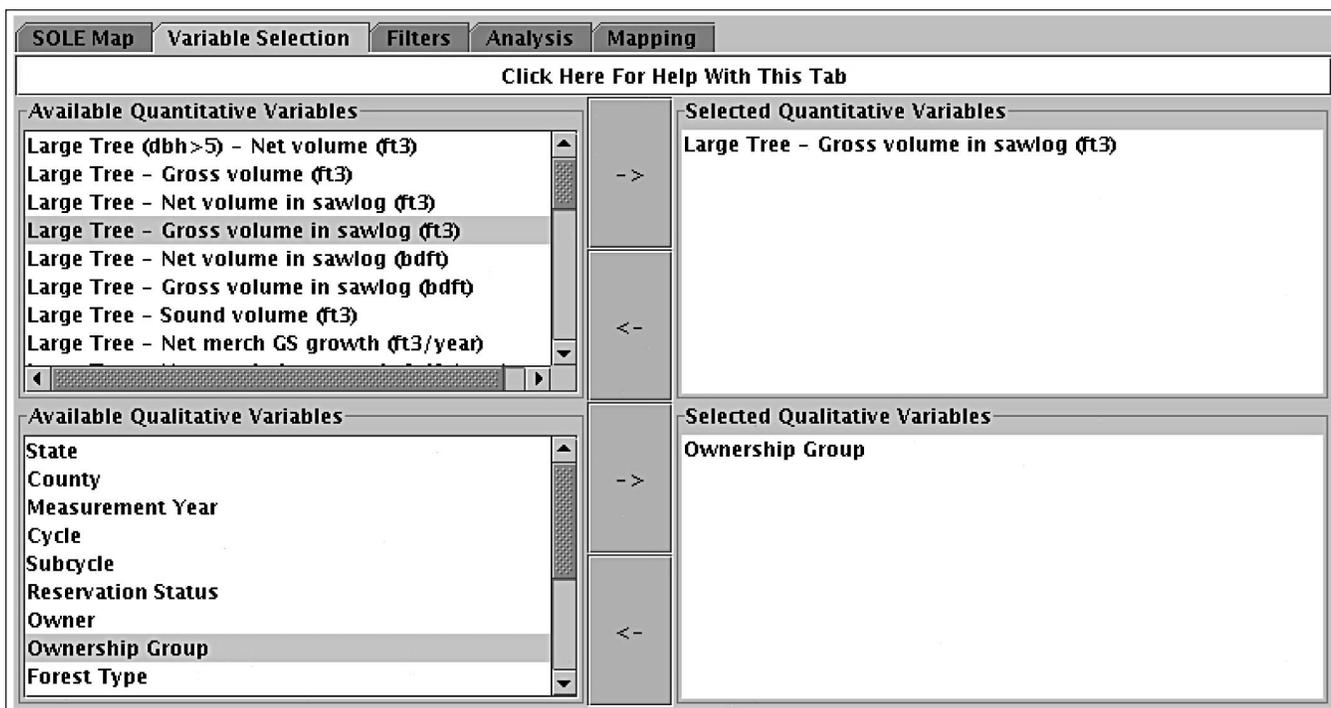


Figure 2.—SOLE variable selection tab.



### Analysis

Charts and tabular analyses are launched from the **Analysis** tab, as illustrated in fig. 3. The header of each analysis indicates which filters, analysis, and variables have been selected.

**Tabular Analysis.**—SOLE can now calculate the means on which the population summary statistics are based by either a moving average (MA) or a mixed estimator (ME) method (Van Deusen 2001). The MA estimator, the default estimator for FIA analysis, provides a complete estimate for a specific 5-year period. MA analysis results in the following seven tables:

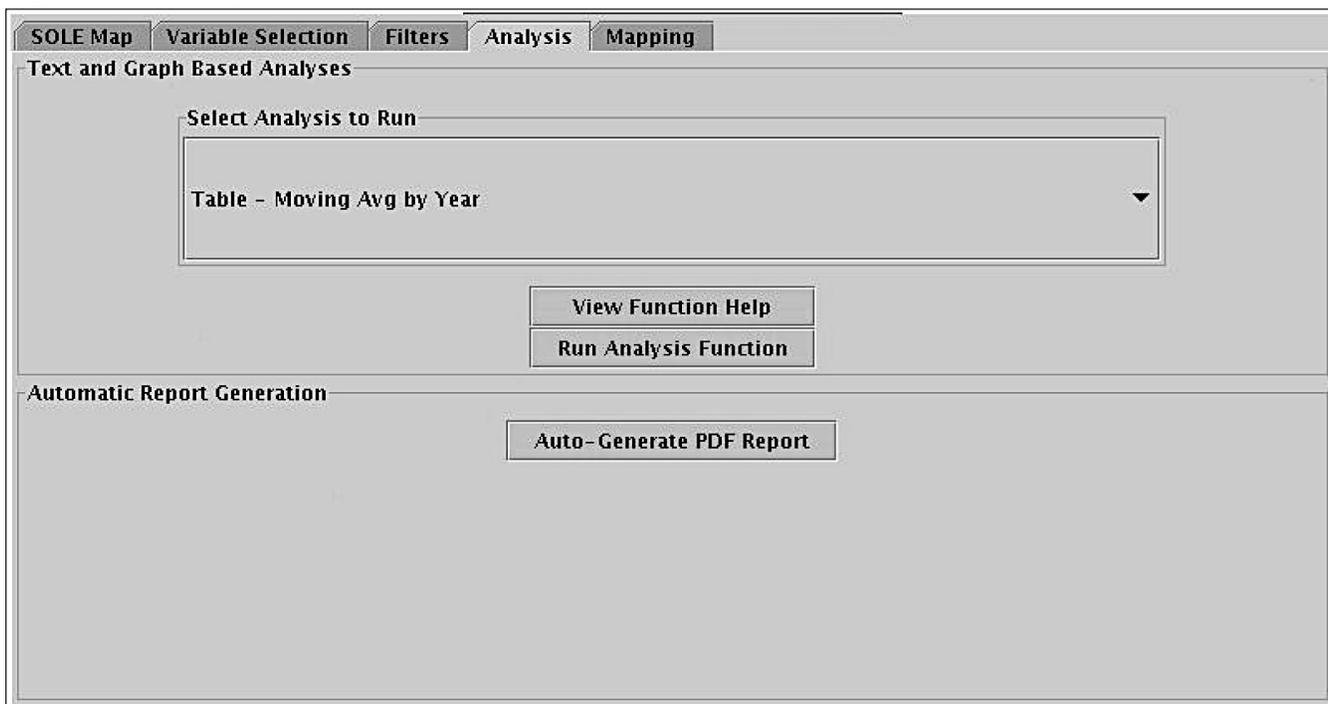
1. MA (Quantitative Variable of interest per acre).
2. Standard error of each element of table 1.
3. Sampled area (acres) represented by the plots used in the calculation of table 1 (summarizes the FIADB variable EXPCURR).
4. Estimated population total for the Quantitative Variable of interest (equals MA multiplied by area).
5. Standard error of the total estimate.
6. Total number of plots used in calculation, by each level of the Qualitative Variable.
7. Number of plots by measurement year.

ME analysis results in the following six tables that have measurement years as columns:

1. ME mean Quantitative Variable of interest by year (per acre).
2. Standard error per year.
3. Sampled area (acres) represented by the plots used in the calculation of table 1.
4. Estimated population total for the Quantitative Variable of interest.
5. Standard error for the total estimate.
6. Adjusted sample size by measurement year.

The MA provides a complete estimate for each specific 5-year period. When the MA is used, change can be determined after a second MA can be calculated in year 6. In this case, the panel 1 data that are measured in year 1 are dropped from the estimator, and the panel 1 data measured in year 6 are added. Note that 80 percent of the data are common to the two consecutive MA estimates. Both the MA and ME can provide full-data change estimates, but to which specific years the MA difference estimator should be applied is not obvious. Another reason the ME estimator might be preferred is its ability to model explicit

Figure 3.—SOLE analysis selection tab.



trends, which permits the use of all available plots to estimate rates of change for any point in time covered by the panel series.

**Graphical Analysis.**—Graphs are often the most comprehensible means of presenting data characteristics. SOLE offers bar charts, box plots, pie charts, and XY plots. Bar and pie charts show a basic distribution of the Quantitative Variable by each level of the Qualitative Variable. XY plots offer some insight into the data distribution and frequency based on the number and distribution of points with respect to the axes. Box plots convey the most statistical information about the data because they show the mean, interquartile range, and outliers for each level of the Qualitative Variable.

**Map Analysis.**—County-level maps can be created from the **Mapping** tab. Basic maps of the mean and median are supplemented by more complex ratio maps, which display the ratio of filtered to unfiltered data. For example, the user can filter for a specific ownership group and then view a map of the proportion of volume on land owned by that group to the volume on all land owned by all groups.

**Portable Document Format Report.**—At the simplest level, the portable document format (more commonly known as PDF) report could be completely contained such that the user selects an area of interest and then generates a report containing a pre-determined combination of text and tabular, graphical, or map analyses describing forest attributes for that area. Potentially, this feature could be developed to produce an array of standard FIA report tables at the click of a button, giving the user the ability to produce current reports at each FIADB update.

## Summary

Web-based FIA analysis tools are essential for proper analysis of FIA data. SOLE provides a simple interface that allows users to obtain customized analytical results. SOLE is constructed in a modular fashion that makes it easy to add new capabilities. SOLE provides a wide range of options for users who want to analyze FIA annual inventory data and obtain graphical and tabular results. Flexibility in each component of SOLE ensures that SOLE remains highly adaptable to changes in database structure and user needs.

---

## Acknowledgments

This project is supported by a USDA Forest Service grant through the Southern Research Station.

## Literature Cited

R Development Core Team. 2003. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org>.

Van Deusen, P.C. 2001. Alternatives to the moving average. In: Reams, G.A.; McRoberts, R.E.; Van Deusen, P.C., eds. Proceedings of the second annual Forest Inventory and Analysis symposium; 2000 October 17–18; Salt Lake City, UT. Gen. Tech. Rep. SRS-47. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 90–93.