USING FIESTA, AN R-BASED TOOL FOR ANALYSTS, TO LOOK AT TEMPORAL TRENDS IN FOREST ESTIMATES

Tracey S. Frescino, Paul L. Patterson, Elizabeth A. Freeman, and Gretchen G. Moisen[1]

Abstract.—FIESTA (Forest Inventory Estimation for Analysis) is a user-friendly R package that supports the production of estimates for forest resources based on procedures from Bechtold and Patterson (2005). The package produces output consistent with current tools available for the Forest Inventory and Analysis National Program, such as FIDO (Forest Inventory Data Online) and EVALIDator. FIESTA was developed as an alternative data retrieval and reporting tool that is functional within the R environment, allowing customized applications and compatibility with other R-based analyses. FIESTA generates estimates and percent sample errors of the estimates for area, population totals, and ratios, while allowing user-defined boundaries, stratification schemes, and data filters. The features of the tool are demonstrated using temporally sensitive data over diverse areas.

INTRODUCTION

Traditional analytical tools of the Forest Inventory and Analysis (FIA) Program have been developed for reporting state-level estimates of forest resources. The growing popularity of the FIA Program, the diverse range of questions and data requests from our customers, and the ongoing advancements of remote sensing technology have expanded our analytical needs. We now require capabilities to monitor annual data trends, report on different areas of interest, and incorporate diverse ancillary data layers for reducing estimation variances. In essence, we need a flexible estimation engine that we can use to answer many forest estimation questions using many different sources of ancillary data.

_____

[1] Forester (TSF), U.S. Forest Service, Rocky Mountain Research Station, Forest Inventory and Analysis Program, 507 25th St., Ogden, UT 84401; Statistician (PLP), Ecologist (EAF), and Research Forester (GGM), U.S. Forest Service, Rocky Mountain Research Station. TSF is corresponding author: to contact, call 801-625-5402 or email at tfrescino@fs.fed.us.

FIESTA

FIESTA (Forest Inventory Estimation for Analysis) is a package developed in the R environment (R Development Core Team 2011). R is a powerful statistical computing and graphical program language that is becoming more popular within the natural resource analysis community. It is free software that provides a flexible, cross-platform environment for statistical tool development and application through R package-building. The comprehensive library of contributed package applications along with R's base functions and extensive user-community provides powerful resources for analysts including statistical, spatial, and graphical tools. The main objective of FIESTA is to provide a data-retrieval and reporting tool for FIA analysts and scientists that is more flexible than existing FIA tools and is adaptable to analysts' changing needs.

The major components of the FIESTA package are:

1. Extracting data from the FIA database for a user-defined area of interest.
2. Acquiring user-defined stratification layers.

3. Generating area estimates and standard errors within user-defined populations.

4. Generating tree attribute and ratio estimates: for example, the number of aspen trees per acre in the Lodgepole Pine forest type.

Specific features of FIESTA include: extracting data from the national or regional FIA National Information Management System database (NIMS) (Woudenberg et al. 2011); allowing for user-defined populations, strata, and data filters; ability to generate tabled estimates with sample error of area, population totals, and estimation of ratios; and providing additional tools for data exploration and spatial data manipulation, such as frequency tables and bar plots as well as raster clipping and pixel or polygon data extraction. FIESTA can be run through a graphical user interface environment or by supplying specific input parameters defined by the user. Because it is developed in R programming language, it also can easily interact with other R functions and packages, opening up many opportunities for further customization and development of tools within FIESTA.

EVALIDator AND FIESTA COMPARISON

EVALIDator is FIA's Web-based or personal computer tool for producing tables of population estimates including sampling errors (this tool is available at www.fia.fs.fed.us/tools-data/other/default.asp). EVALIDator was used to validate the accuracy of output from FIESTA. For example, Table 1 shows a comparison of the output from EVALIDator and the output from FIESTA for estimating cubic feet per acre by forest type in Montana. The percent difference is less than 1 percent for the estimates, most likely from rounding errors. The percent difference in the sampling error is also less than 1 percent in most cases. The larger differences are in types where the sampling error is already high.

FIESTA EXAMPLES

FIESTA originated from a request from the San Carlos Apache Indian Reservation in southeastern Arizona to support a cost-effective woodland inventory to

Table 1.—Comparison of EVALIDator and FIESTA. The difference in per-acre estimates and sampling error of net volume of live trees (at least 5 inches diameter) by forest type on forest land, Montana, 2003-2009. Subset of actual table.

Forest type	EVALIDator Estimates (ft³/acre)	FIESTA Estimates (ft³/acre)	Difference Estimates (%)	EVALIDator Error (%)	FIESTA Error (%)	Difference Error (%)
Rocky Mtn. juniper	328.797	327.603	0.363	8.425	8.427	-0.0237
Juniper woodland	358.052	357.705	0.097	23.955	23.281	2.8136
Douglas-fir	1835.605	1820.685	0.813	2.694	2.687	0.2598
Ponderosa pine	963.91	957.834	0.630	5.27	5.264	0.1139
Engelmann spruce	3038.053	3011.842	0.863	6.32	6.301	0.3006
Grand fir	2486.062	2464.538	0.866	18.416	18.325	0.4941
Subalpine fir	1669.218	1654.858	0.860	4.949	4.922	0.5456
Mountain hemlock	3428.542	3402.067	0.772	21.038	21.074	-0.1711
Lodgepole pine	2210.947	2192.483	0.835	3.312	3.304	0.2415
Western hemlock	8401.806	8323.888	0.927	15.173	15.101	0.4745
Western redcedar	4225.783	4196.477	0.694	9.299	9.251	0.5162
Western larch	2722.67	2698.901	0.873	7.252	7.198	0.7446
Limber pine	517.243	513.428	0.738	23.943	24.184	-1.0066
Whitebark pine	1906.364	1888.815	0.921	7.466	7.477	-0.1473

quantify woodland resources on its 1.8 million acres. It is currently difficult to generate estimates using FIA data for populations other than a state using FIA's estimation tools. With FIESTA, it is possible to extract FIA plots within a specific area, such as the San Carlos Reservation boundary, and produce estimates based on a given stratification layer and the total area of the population.

We generated estimates of area by forest type groupings (Table 2) and area, basal area, and trees per acre of woodland types using FIA plot data and a spatial layer of biological plant communities that was provided for stratification. The resulting information

was helpful for an initial look at quantifying the forest resources within the boundary and identifying areas or forest types having high sampling errors that would need further sampling to minimize variance.

Another topic of interest in the Western region is looking at temporal changes of species distributions. This information is useful for monitoring trends and potential changes in individual species distributions through time. It is currently difficult to generate estimates by year from FIA's estimation tools, but with the programming and customization features of FIESTA, it is a simple routine that can be applied to any species.

We used FIESTA to look at estimates of live net cubic-foot volume annually for lodgepole pine (Pinus contorta) and aspen (Populus tremuloides) species in Colorado, Montana, and Utah (Figs. 1 and 2). The scale of the figures reflects the amount of each species in the respective state. A simple function was developed to generate bar plots of estimates by inventory year using the total plots sampled each year and a simple statewide stratification scheme. The figures are useful as initial data exploratory tools but are not conclusive evidence for real changes on the ground.

Table 2.—Area estimates for San Carlos Apache Indian Reservation forest type groups

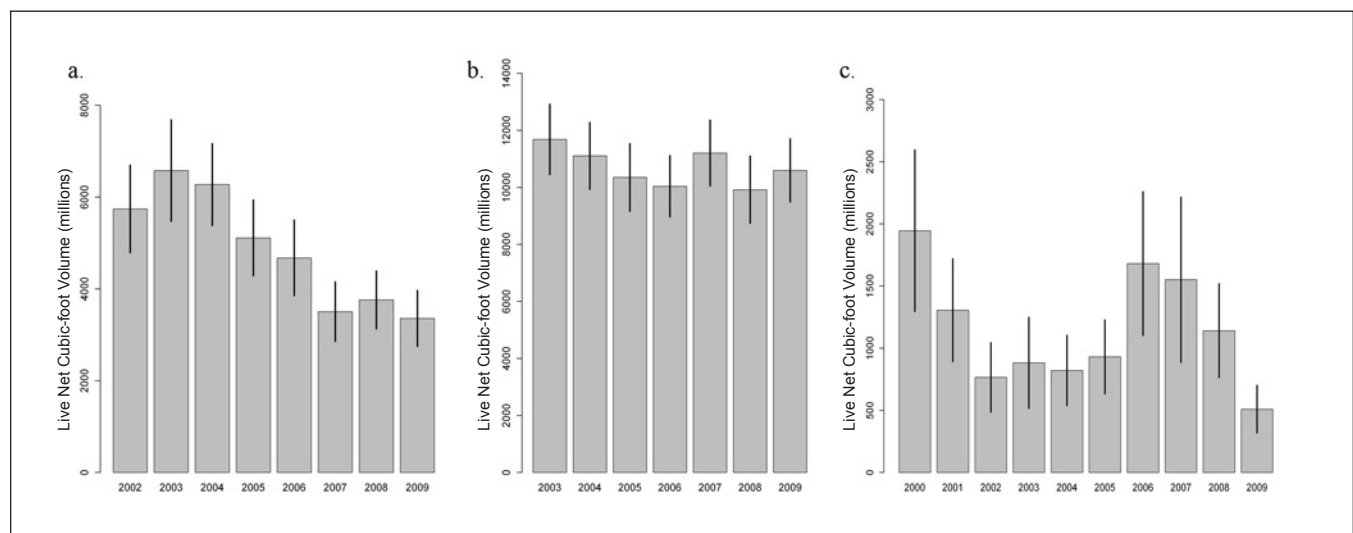| Class	Number of Plots	Area (acres)	Sampling error (%)
Nonforest	173	1,073,858	4
Ponderosa	14	78,786	24
PJandJ	61	378,460	10
Oak	30	161,079	16
Non-stocked	21	123,989	21
Mesquite	5	30,870	45
Cotton/Misc.	2	12,445	71

Figure 1.—Net cubic-foot live volume for lodgepole pine by inventory year for three states in the Interior West: a. Colorado; b. Montana; c. Utah.
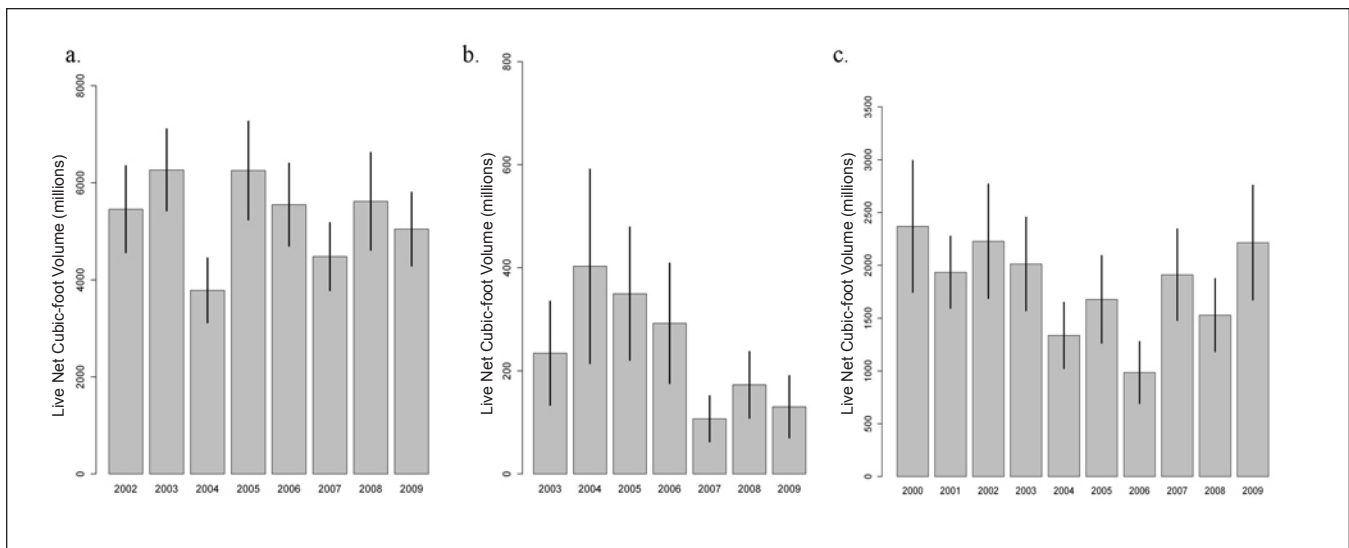
Figure 2.—Net cubic-foot live volume for aspen by inventory year for three states in the Interior West: a. Colorado; b. Montana; c. Utah.

STATUS AND FUTURE OF FIESTA

The FIESTA package was designed as a base program for accessing the FIA NIMS database and generating estimates of forest attributes using user-defined boundaries, strata, and filters.

The functionality and efficiency of the base program will be periodically analyzed and updated for speed and consistency. Updates will also occur to reflect changes in the NIMS database schema, including the addition of new reference tables or regional tables, changes in variable names, or general structural changes. The base program will be continuously augmented with additional data exploration tools and spatial tools to assist with estimation.

Plans for current and future work with FIESTA, incorporating additional customized capabilities and tools for different needs, are as follows:

- Adding small area estimation capabilities to generate estimates for areas within user-defined boundaries that contain only a few FIA field-sampled plots.

- Adding compatibility with R packages, such as ModelMap (Freeman 2009) and PresenceAbsence (Freeman and Moisen 2008), to coordinate estimates of FIA attributes with pixel-based mapping of the same attributes.
- Adding a mechanism to compensate for nonresponse within the FIA sample.
- Adding functionality to integrate remotely sensed observations, such as photo interpretation from large-scale aerial photographs or measurements from Landsat or Light Detection and Ranging.
- Exploring alternatives to the moving average.
- Communication with Forest Vegetation Simulator to generate estimates of future growth simulations.

ACKNOWLEDGMENTS

LITERATURE CITED

Bechtold, W.A.; Patterson, P.L., eds. 2005. **The enhanced Forest Inventory and Analysis program - national sampling design and estimation procedures.** Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Freeman, E.A. 2009. **ModelMap: an R package for modeling and map production using Random Forest and Stochastic Gradient Boosting.** Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. Available at http://CRAN.R-project.org/. [Date accessed unknown].

Freeman, E.A.; Moisen, G.G. 2008. **PresenceAbsence: an R package for presence-absence analysis.** Journal of Statistical Software. 23(11): 1-31.

R Development Core Team. 2011. **R: a language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing. Available at www.R-project.org. [Date accessed unknown].

Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M.; LaPoint, E.B.; Turner, J.A.; Waddell, K.L.; Boyer, D.; Christensen, G.; Ridley, T. 2011. **The Forest Inventory and Analysis database: database description and users manual version 5.1 for Phase 2.** U.S. Department of Agriculture. Available at http://fia.fs.fed.us/library/database-documentation/current/ver5/FIADB_user%20manual_5-1_p2_10_2011.pdf. [Date accessed unknown].