ECOGRAPHY

Software note

'FIESTA': a forest inventory estimation and analysis R package

Tracey S. Frescino[1], Gretchen G. Moisen[1], Paul L. Patterson[1], Chris Toney[1] and Grayson W. White[2,3]

[1]US Dept of Agriculture, Forest Service, Rocky Mountain Research Station, Forest Inventory and Analysis (FIA), Riverdale, UT, USA
[2]RedCastle Resources Inc., Contractor to the US Dept of Agriculture, Forest Service Geospatial Technology and Applications Center, Salt Lake City, UT, USA
[3]Dept of Forestry, Michigan State Univ., East Lansing, MI, USA

Correspondence: Tracey S. Frescino (tracey.frescino@usda.gov)

Ecography
2023: e06428
doi: 10.1111/ecog.06428

Subject Editor: Brody Sandel
Editor-in-Chief: Miguel Araújo
Accepted 16 February 2023

Ecologists are increasingly relying on national forest inventories to address a wide variety of issues. The 'FIESTA' R package (Forest Inventory ESTimation and Analysis) is a tool that enables customized investigations using the extensive sample-based inventory data collected across all lands in the US by the US Dept of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) Program. To date, the complex nature of the FIA inventory constrains many users to conduct only limited analyses through existing tools with pre-specified geographic boundaries, timeframes, and auxiliary data under a single statistical estimation process. Yet, the rapid evolution of available remotely sensed data and statistical methods present the opportunity to conduct spatial and temporal analyses of forest attributes that are much more relevant to many pressing ecological, environmental, economic, and social issues in the US, The 'FIESTA' package was developed to augment the current set of available tools by providing a flexible platform that accommodates evolving technologies and leading-edge estimation techniques. The package contains a collection of functions that can query FIA databases, summarize sample-based inventory data, extract and aggregate auxiliary spatial data, and generate estimates with associated variances. The 'FIESTA' R package is available on CRAN (https://cran.r-project.org/package=FIESTA).

Keywords: estimation, FIA, FIESTA, forest inventory, model-assisted, small area estimation

Background

Ecologists are increasingly using long-term monitoring data provided by national forest inventories to help answer questions at different spatial and temporal scales across ecosystems heavily influenced by humans (Seeing 2022). Within the United States, the national forest inventory is conducted by the Forest Inventory and Analysis (FIA) Program of the US Dept of Agriculture, Forest Service (www.fia.fs.usda.gov). This program was initiated in 1928 with mission to collect, analyze, and report on the status and trends of America's forests (LaBau et al. 2007). Although the FIA program was originally intended to monitor forestry-centric population parameters like forest land area, tree volume, growth, and mortality; the use of FIA data in more diverse ecological analyses began many decades ago. Rudis (2003) provides an annotated bibliography 1976–2001 of studies using FIA data in applications related to biomass, dead wood, aesthetics, geographic context, nonforest influences, owner attitudes, range lands, recreation, tropical issues, water quality, vegetation habitat typing, and wildlife. Tinkham et al. (2018) extended the Rudis (2003) review to more recent applications of forest inventory data to study not only forest products and growth, but forest-related aspects of the carbon cycle, climate, fuels and fire, insects and pathogens, invasive species, and habitat suitability. In recent years, the use of FIA data in ecological analyses continues to grow to applications for shifting species distributions (Stanke et al. 2021, Hill et al. 2021), effects of drought (Anderegg et al. 2020), effects of mountain pine beetle outbreaks (Thompson 2017), climatic effects on the relationship between biodiversity and productivity (Fei et al. 2018), downed and dead wood (Woodall et al. 2019), shifts in forest density (Woodall and Weiskittel 2021), protecting biodiversity with strategic forest reserves (Law et al. 2021), the role of fungal networks in biodiversity (Averill et al. 2022, Carteron et al. 2022), and more. Stable, consistent, and long-established, FIA data offer a wealth of information to ecologists but must be used with caution and in collaboration with forest inventory specialists to ensure the data are used appropriately (Seeing 2022).

The 1998 Farm Bill (Public Law 105-185) brought significant changes to the focus and methods of FIA including: a switch from periodic to annual surveys, inclusion of remotely sensed data, and an expansion of data collection beyond timber products. These changes resulted in a prescribed process for data collection, compilation, and statistical estimation (what we hereafter call 'FIA standard estimators') documented in Bechtold and Patterson (2005). Regional FIA units conduct state-level inventories to annually update an extensive, publicly available national FIA database and produce estimates of area, volume, growth, removals, and conditions of forest resources. Publicly available tools are provided to generate state-level estimates with FIA standard estimators and serve an important role in delivering statistically-sound estimates from regional to national investigations (https://apps.fs.usda.gov/fia/datamart/datamart.html). However, due to the complex nature of FIA data, it has been increasingly difficult to use this valuable information to its full potential without a deep knowledge of the FIA database, data collection techniques, and many nuances of regional differences and changes in definitions through time. Adding to the complexity, new estimation techniques are evolving quickly in the FIA program (Westfall et al. 2022). This evolution is motivated by new information needs about land use and land cover change, remote or inaccessible regions, and smaller geographic areas and shorter time intervals (USDA 2014, Farm Bill, Public Law 113-79). Tools are needed to assist FIA data users in navigating the complexity of the data and allow them to take advantage of enhanced statistical techniques, increased computing capacity, as well as improved remotely sensed data.

The 'FIESTA' R package (www.r-project.org) was developed to provide a flexible platform for accommodating new information needs and FIA's evolving estimation strategies – closing the gap between complex ecological questions and statistically defensible estimates across spatial and temporal domains. 'FIESTA' was designed in R's open-source environment (www.r-project.org) to provide transparency of FIA estimation routines, while making use of R's statistical programming capabilities, cross-platform interoperability, and extensive library of available packages (Comprehensive R Archival Network (CRAN); https://cran.r-project.org). The 'FIESTA' package contains a collection of functions that can query FIA databases, summarize sample-based inventory data, extract and aggregate auxiliary spatial data, and generate estimates with associated variances. While other R packages exist with goals of increasing accessibility to the FIA databases and estimation using FIA data, most notably 'rFIA' (Stanke et al. 2020), 'FIESTA' functions widen the options for querying FIA data and include efficient tools for compiling spatial data for different estimation strategies. Although additional functionality results in complex 'FIESTA' functions, it provides the necessary flexibility for research and production environments.

'FIESTA' is increasingly being used in applications which require non-standard estimates from FIA data. For example, it serves as the estimation engine for generating land use and land cover change estimates from FIA's Image-based Estimation Program (Frescino and Patterson 2017). It has provided a flexible platform for comparing different estimation strategies to improve estimates of forest disturbance (Schroeder et al. 2014). It has facilitated statistical investigations comparing both model-assisted and small area (SA) estimation methods improving the reliability of forest inventory information (e.g. Rintoul et al. 2020, White et al. 2021, Wojcik et al. 2022). It has been used to create the first dashboards of numerous SA estimates for counties, ecosubsections, and watersheds across the conterminous US (Frescino et al. 2022a).

Recognizing the growing need for non-standard spatial and temporal estimates of forest attributes from forest land managers, ecologists, policy makers and scientists, this manuscript documents the public release of the 'FIESTA' package (Frescino et al. 2022b) for use by a broader community. Here, we describe the general functionality of 'FIESTA' and associated package, 'FIESTAutils' (Frescino et al. 2022c), including a description of the functions and structure of 'FIESTA' for generating standard and non-standard estimates using sample-based FIA data. Further details and examples can be found in the 'FIESTA' package vignettes. A brief description of the FIA annual sample design and FIA terminology is provided in the Supporting information while details of the FIA annual sampling design, plot configuration, and standard estimators can be found in Bechtold and Patterson (2005). Knowledge of the FIA sample design, sampling protocol, terminology, and database design is essential for generating accurate and meaningful output.

Methods and features

The 'FIESTA' package relies on a separate package, 'FIESTAutils', that provides functions for data wrangling, spatial data analysis, and statistical modeling. 'FIESTA' functions are organized by theme and are named with a corresponding prefix. There are three main themes to characterize the functions: 1) core functions; 2) estimation modules; and 3) analysis functions. The core functions include tools for extracting, exploring, and compiling database and spatial information (DB*, dat*, sp*). These functions may be used independently but are also necessary components of the estimation modules. The estimation modules combine multiple functions to enable different estimation strategies. The estimation modules include statistically based estimators and their variance estimators or provide integration with other R estimation packages (www.r-project.org) with additional steps to check and format data for package inputs and outputs. The analysis functions are wrapper functions that combine multiple functions for using the estimation modules in custom or routine analyses. The analysis functions reside in the 'FIESTAnalysis' package which depends on 'FIESTA' and 'FIESTAutils' packages, and is available on GitHub (https://github.com/USDAForestService/FIESTAnalysis). The analysis functions provide tools for compiling inputs to estimation modules, automating standard reporting, performing temporal analyses, and comparing estimates. These specific functions will not be described in the paper because they will be evolving quite rapidly in response to this public release of 'FIESTA'. However, they are mentioned here as a placeholder for the evolution of the package. A schematic for the overarching structure of 'FIESTA' is given in Fig. 1, and more information of each theme will be addressed in later sections.

A user wanting to conduct a customized analysis can use 'FIESTA' to think through and carry out the acquisition of their information needs, including: the sampling population or area of interest; the time frame of the sampled plot data used; the estimation response, or plot attribute; the type of estimator that fits best for the analysis; and auxiliary information that may improve the estimate by reducing variance. The core functions and modules may be used interchangeably through available 'FIESTAnalysis' functions or custom-designed analysis functions to accomplish desired results.

The 'FIESTA' package relies heavily on the 'data.table' (Dowle and Srinivasan 2019) R package for overall speed and efficiency of data processing and the 'sf' (Pebesma 2018) R package for spatial manipulation of vector data. Functions from 'methods', 'grDevices', 'graphics', and 'utils', are imported for data exploration, summarizing, and graphical display. Other imported and suggested packages are discussed in the applicable sections following.

Core functions

The 'FIESTA' DB functions query and extract data from FIA's online publicly available DataMart (https://apps.fs.usda.gov/fia/datamart/datamart.html). 'FIESTA' functions download, compile, and query these files for a user-defined set of requested states using functions from imported packages, 'DBI' (Wickham and Müller 2018), 'httr' (Wickham 2018), and 'sqldf' (Grothendieck 2017). Because of FIA's confidentiality mandate to protect the privacy of landowners (Section 1770 of the Food Security Act of 1985 (7 U.S.C. 2276(d)) in the Fiscal Year 2000 Consolidated Appropriations Bill (PL 106-113), as well as protecting the scientific integrity of FIA's sample design, the exact coordinates of the sample plot locations are not included in the public database. While public coordinates are often enough for a meaningful analysis, one may need access to exact coordinates for FIA plots for their particular analysis. 'FIESTA' does not provide access to these protected coordinates; however, access can be requested through FIA's Spatial Data Services (www.fia.fs.usda.gov/tools-data/spatial/index.php). If granted, actual coordinates can easily be aligned with plot identifiers in 'FIESTA', and a user can proceed seamlessly with analyses using exact coordinates. A full understanding of the FIA database is recommended when extracting plots for a non-standard population boundary, as there may be differences within the area, such as: sampling intensities; response characteristics; or measurement time frames.

The 'FIESTA' dat functions assist with customizing data variables as well as summarizing and visualizing FIA data, including aggregating tree attributes (e.g. volume, biomass, trees per acre) and adding classified variables for estimation or exploratory analyses. These functions are called from the estimation modules but may also be run independently.

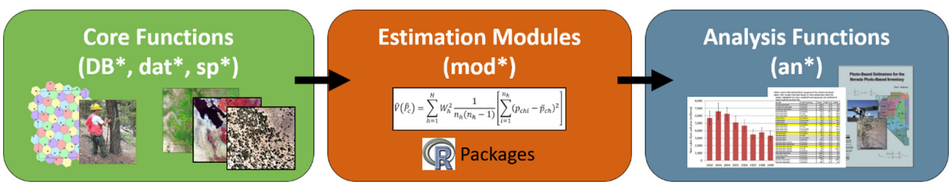

Figure 1. Schematic of the overarching structure of 'FIESTA' and 'FIESTAnalysis'. The core functions include functions for querying, exploring, and compiling data from the publicly available FIA field plot data, from aerial photo-plot data, as well as from satellite-based datasets. The estimation modules combine field- and/or photo-based sample data with the auxiliary remotely sensed data to produce estimates of forest attributes and their estimated variances though internal functions as well as other R estimation packages (www.r-project.org). These modules contain steps to check and format data for package inputs and outputs. The analysis functions are wrapper functions for using the estimation modules in customized or routine analyses, including in some cases automated tables, graphics, and reports.

The 'FIESTA' sp functions are for extracting and summarizing spatial vector or raster data. These tools provide capability for including auxiliary information to improve precision in estimates. Functions for spatial processing use the geospatial data abstraction library (GDAL) for low-level access to raster and vector geospatial data formats (GDAL/OGR contributors 2019). We use the packages 'rgdal' (Bivand et al. 2013) and 'sf' (Pebesma 2018) for their GDAL bindings in R, along with their associated libraries of utility functions. Certain spatial functions are supported by C++ code via the *Rcpp* interface (Eddelbuettel and Francois 2011, Eddelbuettel 2013) to increase performance when processing larger datasets. These include internal functions for rasterizing vector data, counting unique combinations across thematic raster layers, and calculating mean and variance from raster data in one pass. All internal spatial functions reside in the 'FIESTAutils' dependent package. Conversion of *rgdal* functionality is in development in recognition of the anticipated retirement of *rgdal* in late 2023 (https://cran.r-project.org/web/packages/rgdal/index.html).

Estimation modules

The 'FIESTA' estimation modules (mod*) include: 1) the green book (GB) module which uses FIA standard estimators with either standard or customized inventory time spans, estimation domain units, or stratification schemes; 2) the photo-based (PB) module which allows users to supplement FIA's traditional inventory with land use and land cover change data collected from aerial photographs; 3) the model-assisted (MA) module which uses model-assisted estimators for alternative methods for estimation, allowing flexibility for using more auxiliary information to improve precision; and 4) the SA module which uses SA estimators for constructing reliable estimates in spatial or temporal domains that contain only a few FIA plots.

Each module includes an associated function (mod*pop) which compiles a set of population data for an area, or population of interest. The functions check data inputs and calculate adjustment factors to account for plots with partial nonresponse (e.g. inaccessible condition). A set of population data can be used for multiple estimation responses and/or filters.

Green-book module

The 'FIESTA' GB module calculates population estimates and their sampling errors based on Bechtold and Patterson (2005), referred to as the 'green book' in forest inventory circles. For FIA's nationally consistent, semi-systematic annual sample design, see chapter four within the GB (Scott et al. 2005). FIA's sample design is based on two phases: the first phase uses remotely sensed data to post-stratify the land area to increase precision of estimates; while the second phase obtains photo and ground observations and measurements for a suite of information for plots, each located within hexagons, approximately 6000 acres in size, and tiled across the US (see Supporting information for more FIA terminology and details on the FIA plot design). The associated estimators and variance estimators are used for area and tree attribute totals with the assumption of a simple random design with post-stratification. Adjustment factors are calculated by estimation unit and strata to account for plots with partial nonresponse.

The GB module produces area (e.g. acres) and tree (e.g. biomass) estimates by domain (e.g. forest type) and per-acre or per-tree ratio-of-means estimates (e.g. trees per acre) within domain. In addition, the module includes adjustments for partial nonresponse, supports post-stratification for reducing variance, and reports by estimation unit or a summed combination of estimation units. The GB module can be used for generating FIA's standard state-level estimates or estimates using customized populations or stratification strategies.

Photo-based module

Sampling high-resolution imagery, such as large-scale photography, has made it possible to enhance FIA's data collection–extending it to nonforest lands and providing more timely information on disturbance and change estimates.

The 'FIESTA' photo-based (PB) module calculates population estimates and associated sampling errors based on Patterson (2012). In contrast to FIA's traditional GB estimators which were constructed based on the finite sampling paradigm using sample plots with distinct area, the photo-based estimators were constructed based on the infinite sampling paradigm, along with the concept of a support region. The sample is the set of plot centers and the information from the support region (the photo plot (Cordy 1993, Stevens and Urquhart 2000), which may include a collection of point-wise observations) is assigned to the plot centers. The photo interpreted points are used as a sample of the support region and the measurements are used to estimate the information from the support region. The PB module includes non-ratio estimators for area and percent cover estimates by domain and ratio-of-means estimators for area and percent cover estimates within domain, and supports post-stratification for reducing variance.

The PB module was developed to generate estimates from supplementary remotely-sensed data collected across the FIA grid (Frescino et al. 2009), but has been extended to other probabilistic sample datasets that includes information from a surrounding support region. Similar to the GB module, customized strata, population boundaries, and domain categories can also be used for estimation.

Model-assisted estimation module

FIA's standard estimation strategies assume simple random sampling and only accommodate a limited set of post-stratification options. With current resource questions and monitoring concerns, together with advancements of auxiliary data products, modeling techniques, and programming capabilities, new strategies are in development for estimation, including the use of a broader set of model-assisted estimators.

The 'FIESTA' MA estimation module was set up to integrate with current model-assisted estimators available on CRAN, specifically from the 'mase' (model-assisted survey estimators) R package (McConville et al. 2018). The module includes functional steps for checking, compiling, and formatting FIA plot data and auxiliary spatial information, ensuring FIA adjusted data are summarized to the domain and plot levels, for input to *mase.* The 'mase' package allows a user to choose from not only estimators frequently used in forest inventory applications such as the Horvitz–Thompson, ratio, and post-stratified estimators, but also more flexible estimators falling under the broad class of generalized regression estimators with extension to lasso, ridge, and elastic net estimators to accommodate variable selection and collinearity issues (McConville et al. 2020). The MA module also allows user to opt for bootstrap variance estimates for the model-assisted estimators that *mase* provides.

Small area estimation module

The FIA inventory sample design was developed for broad-scale monitoring, with an intensity of approximately one sample plot per 6000 acres (24 km^2). It is therefore difficult to estimate forest attributes over small areas or populations, such as Natl Forest districts, small counties, or areas affected by disturbances. Small 'areas' may also refer to other domains such as time intervals or forest classifications for which there are too few sample plots. Numerous strategies for small area estimation (Rao and Molina 2015) have been developed, documented, and packaged on CRAN to use auxiliary information and modeling to enhance estimation techniques for small areas, some of which we have integrated into 'FIESTA'.

There are many different approaches to small area estimation, but currently 'FIESTA' uses the Empirical Best Linear Unbiased Prediction (EBLUP) estimation strategy through both the 'JoSAE' (Breidenbach 2018) and 'sae' (Molina and Marhuenda 2015) packages and a hierarchical Bayesian (HB) strategy through the 'hbsae' (Boonstra 2012) package. Borrowing strength (e.g. plots and auxiliary data) from similar areas within a larger but ecologically similar area, a linear mixed model is fit to describe the relationship between the forest attribute of interest and auxiliary remotely sensed data (fixed effects) with adjustments for the small areas of interest (random effects). The models can be fit at the unit level which relates plot observations to plot-intersected covariates, or at the area level which relates mean area observations to mean area covariates. Simple area- and unit-level EBLUPs and HB predictions are output from 'FIESTA', with additional small area estimators planned for the future.

Example

The examples below follow Fig. 2, which describes the process for generating estimates through 'FIESTA' for an area of interest (AOI). We demonstrate three different 'FIESTA' modules: GB, MA, and SA. Here we demonstrate how we can use 'FIESTA' with FIA data and auxiliary information to characterize the AOI using different FIA attributes and filters and to answer a variety of ecological questions. More examples are provided in FIESTA's extensive vignettes

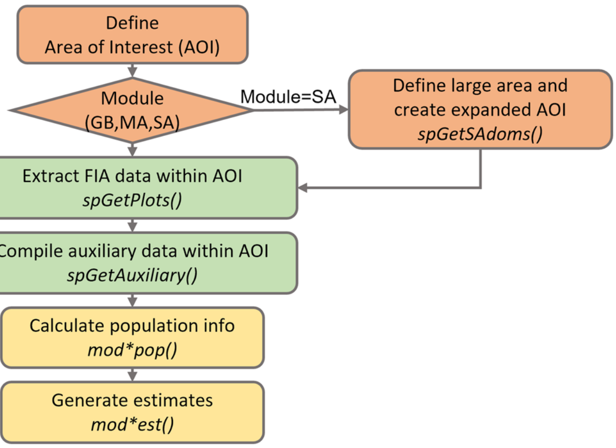

Figure 2. Flowchart describing the process for generating estimates using 'FIESTA'.

(*vignette(package="FIESTA")*). Examples comparing estimates from 'FIESTA' at national levels for watersheds, ecosubsections, and counties, along with links to associated dashboards, can be found in Frescino et al. (2022a).

Define area of interest

We use the Middle Bear-Logan Watershed, hydrologic unit code (HUC) level eight (16010203), as the defined AOI. The watershed is located in northern Utah crossing over to Idaho and encompasses multiple ownerships. It is approximately 1500 km^2 in size, ranging from 1340 to 3030 m a.s.l. and is the primary source of water across the Bear River Basin. The watershed is essential for Logan's growing population, providing water and recreation, and is habitat to many different species of wildlife (https://bearriverinfo.org/watershed-description/middle-bear-logan/index).

For the SA module, we use an additional function (*spGetSAdoms*) that includes an automated process for selecting a larger area of representative domains from which to borrow strength for model development. In this example, the Overthrust Mountains (M331D) Ecomap Section boundary (Cleland et al. 2007) is used to extract ecologically representative data to help improve estimates within the Middle Bear-Logan Watershed.

Extract FIA data within AOI (*spGetPlots*)

We use the most current data from FIA's publicly available DataMart ver. 2.0.1, last updated on 14 November 2022 (https://apps.fs.usda.gov/fia/datamart/datamart.html). Data were extracted on 8 December 2022 and subset to the M331D Section boundary. For GB and MA modules, plot data are extracted within the AOI boundary. For the SA module, we extract data within the expanded AOI, M331D Ecomap Section boundary. All plot data are stored in SQLite database as Supporting information with this paper.

Table 1. Auxiliary data file names and descriptions used in example.

Attribute	Description
M331D_dem_90m.img	LANDFIRE 2010 30 m resolution DEM–Resampled to 90 m (www.landfire.gov/elevation.php) [Downloaded:2013]
M331D_tcc16_90m.img	2016 Natl land cover dataset (NLCD), analytical tree canopy cover – Resampled to 90 m (https://data.fs.usda.gov/geodata/rastergateway/treecanopycover) [Downloaded:04/06/2019]
M331D_ppt_90m.img	PRISM mean annual precipitation – 30 year normals (1991–2020) (mm*100) (https://prism.oregonstate.edu) [Downloaded:12/05/2021]
M331D_def_90m.img	TOPOFIRE mean annual climatic water deficit – 30 year normals (1981–2010) (Holden et al. 2019)
M331D_tnt_90m.img	A tree/non-tree lifeform mask, recoded from LANDFIRE 2014 existing vegetation type (EVT), and resampled 90 m based on majority value (www.landfire.gov/evt.php)

Compile auxiliary data within AOI (*spGetAuxiliary*)

The auxiliary information was extracted and summarized from several different data sources, listed in Table 1. We use the perturbed coordinates included in the FIA Datamart for assigning auxiliary data values to the plot locations. All data are subset to the M331D Section boundary and stored as Supporting information with this paper.

Calculate population info (*mod*pop*)

The module population data is necessary for generating corresponding module estimates. For GB and MA modules, we use the same plot data and auxiliary information extracted within the AOI boundary using the *spGetPlots* and *spGetAuxiliary* 'FIESTA' functions. We define the categorical tree/non-tree lifeform auxiliary layer as the strata information used for GB post-stratification estimates. For the SA module, we use the plot data and auxiliary information from the same functions using the expanded AOI boundary.

Generate estimates (*mod**)

The *mod*tree* 'FIESTA' functions generate estimates including summarized tree data. In this example, we demonstrate use of the respective population data for generating total tree estimates using the three different modules, setting a random seed to assure reproducibility of the estimators. We then look more closely at the estimates from each module for live tree basal area. In Table 2, we see that the additional auxiliary information resulted in slightly smaller percent standard errors and the SA area-level method resulted in even smaller percent standard errors for three of the four HUC 10 boundaries.

Finally, we use the same set of population data to produce estimates for several different variables using the SA module and generate maps of the estimates to help characterize the Middle Bear-Logan Watershed (Fig. 3).

Discussion

This example demonstrates how we can easily move through the steps described in Fig. 2 to characterize and understand an AOI using 'FIESTA'. By comparing these variables as we do in Fig. 3, we can get a more extensive understanding of the forest health across the AOI. Within the Middle Bear-Logan Watershed, we can see patterns, such as the distribution of live tree basal area and aboveground carbon being highest in the south-eastern area of the watershed, in HUC 1601020302, and the largest square footage of standing dead basal area found in the northeast HUC, 1601020303.

It is notable that multiple estimates can be generated using different methodology, with some estimates more precise than others. It is important to keep in mind that this example is based on the perturbed public coordinates and may misrepresent the true value of adding extracted auxiliary data sources. In addition, one surprising artifact from this example is the large percent standard error from the SA module (with respect to the MA module) for HUC 10, 1601020304. This may be a consequence of the MA estimator's variance estimator being negatively biased in situations with limited degrees of freedom, leading to the potential of an overestimate of precision for the MA estimator (McConville et al. 2017).

Conclusion

The 'FIESTA' R package was designed as a research tool for generating estimates of sample-based inventory data. The main objective of 'FIESTA' is to enhance and add to the suite of current FIA estimation tools by providing a flexible platform to accommodate customized inputs, different estimators, and auxiliary remotely sensed information, thus opening up opportunities for using FIA's extensive database for new spatial and temporal analyses related to forests in the US The tool was developed in the R environment to take advantage

Table 2. Estimates (*est), percent standard errors (*pse), and number of plots greater than 0 (NBRPLT.gt0) for estimates of live tree basal area (sqft) within the Middle Bear-Logan Watershed Area of Interest by hydrologic unit code (HUC) 10 boundary.

huc10	NBRPLT.gt0	AREAUSED	GBest	GBpse	MAest	MApse	SAest	SApse
1601020301	6	139 043	2 355 956	41.77	2 708 913	35.09	3 915 136	21.12
1601020302	20	183 677	9 621 684	22.16	8 900 244	20.46	9 153 829	14.87
1601020303	20	159 768	8 994 584	17.83	8 124 436	16.26	8 727 938	12.89
1601020304	5	83 325	1 437 779	57.27	1 749 450	27.27	1 451 326	41.10

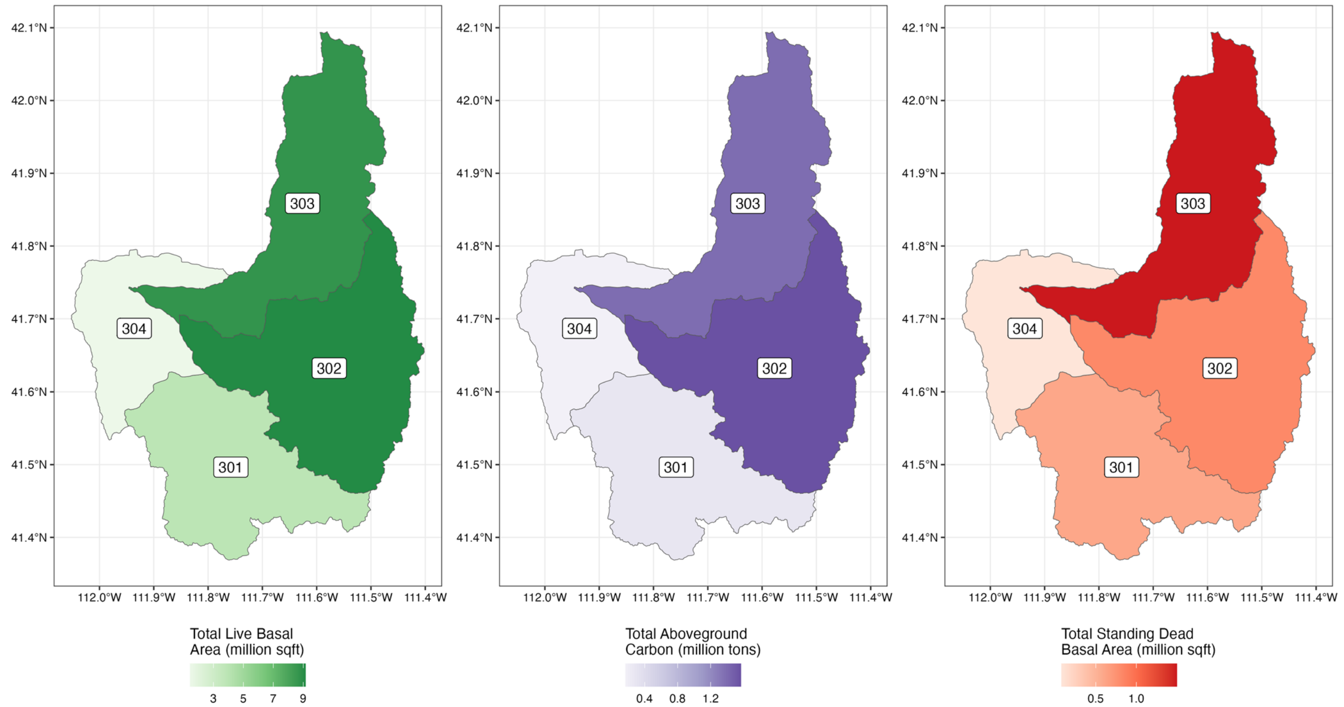

Figure 3. Maps of total live basal area (left); total aboveground carbon (center); total standing dead basal area (right) estimates generated from SA module. Labels represent the last three digits of the associated hydrologic unit code (HUC).

of the powerful and efficient programming capabilities of R and for easy integration with other available R packages or other R-compatible software. The 'FIESTA' package will continue to evolve with alternative strategies for Natl FIA estimation and extension to international collaborations.

The 'FIESTA' ver. 3.5.0 code was written and prepared by a US Government employee on official time, and therefore it is in the public domain and not subject to copyright. The 'FIESTA' R source code is available on CRAN (https://cran.r-project.org/package=FIESTA) and will be evolving through time. Citation information can be found using the following code: citation('FIESTA').

To cite 'FIESTA' or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for 'ver. 1.0':

Frescino, T. S., Moisen, G. G., Patterson, P. L., Toney, C. and White, G. W. 2023. 'FIESTA': a forest inventory estimation and analysis R package. – Ecography 2023: e06428 (ver. 1.0).

Funding – This funding was supported by the USDA Forest Service, Forest Inventory and Analysis Program and the USDA Geospatial Technology and Applications Center.

Author contributions

Tracey S. Frescino: Conceptualization (lead); Methodology (lead); Project administration (lead); Software (lead); Writing – original draft (lead); Writing – review and editing (lead). **Gretchen G. Moisen**: Conceptualization (equal); Methodology (equal); Software-Supporting, Writing – original draft-Supporting, Writing – review and editing. **Paul L. Patterson**: Conceptualization (equal); Methodology (equal). **Chris Toney**: Conceptualization-Supporting, Methodology-Supporting, Software-Supporting, Writing – original draft-Supporting, Writing – review and editing. **Grayson W. White**: Conceptualization-Supporting, Methodology-Supporting, Software-Supporting, Writing – original draft-Supporting, Writing – review and editing.

Transparent peer review

The peer review history for this article is available at https://publons.com/publon/10.1111/ecog.06428.

Data availability statement

Data are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.4tmpg4ffw (Frescino 2023).

Supporting information

The Supporting information associated with this article is available with the online version.

References

Anderegg, W. R. L., Trugman, A. T., Badgley, G., Konings, A. G. and Shaw, J. 2020. Divergent forest sensitivity to repeated extreme droughts. – Nat. Clim. Change 10: 1091–1095.

Averill, C., Fortunel, C., Maynard, D. S., van den Hoogen, J., Dietze, M. C., Bhatnager, J. M. and Crowther, T. W. 2022. Alternative stable states of the forest mycobiome are maintained through positive feedbacks. – Nat. Ecol. Evol. 6: 375–382.

Bechtold, W. A. and Patterson, P. L. 2005. The enhanced forest inventory and analysis program—national sampling design and estimation procedures. – Gen. Tech. Rep. SRS-80. U.S. Dept of Agriculture, Forest Service, Southern Research Station, 85p.

Bivand, R. S., Pebesma, E., Gomez-Rubio, V. 2013. Applied spatial data analysis with R, 2nd edn. – Springer.

Boonstra, H. 2012. hbsae: hierarchical Bayesian small area estimation. – https://CRAN.R-project.org/package=hbsae.

Breidenbach, J. 2018. JoSAE: unit-level and area-level small area estimation. – http://CRAN.R-project.org/package=JoSAE.

Carteron, A., Vellend, M. and Laliberté, E. 2022. Mycorrhizal dominance reduces local tree species diversity across US forests. – Nat. Ecol. Evol. 6: 370–374.

Cleland, D. T., Freeouf, J. A., Keys, J. E., Jr., Nowacki, G. J., Carpenter, C. and McNab, W. H. 2007. Ecological subregions: sections and subsections of the conterminous United States [1:3,500,000] [CD-ROM]. Sloan, A. M., cartog. Gen. Tech. Report WO-76. – U.S. Dept of Agriculture, Forest Service.

Cordy, C. B. 1993. An extension of the Horvitz–Thompson theorem to point sampling from a continuous universe. – Stat. Probab. Lett. 18: 353–362.

Eddelbuettel, D. and Francois, R. 2011. Rcpp: seamless R and C++ integration. – J. Stat. Softw. 40: 1–18.

Eddelbuettel, D. 2013. Seamless R and C++ integration with Rcpp. – Springer.

Fei, S., Jo, I., Guo, Q., Wardle, D. A., Fang, J., Chen, A., Oswalt, C. M. and Brockerhoff, E. G. 2018. Impacts of climate on the biodiversity-productivity relationship in natural forests. – Nat. Commun. 9: 5436.

Frescino, T. 2023. Example data from: FIESTA: a forest inventory estimation and analysis R package. – Dryad Digital Repository, https://doi.org/10.5061/dryad.4tmpg4ffw.

Frescino, T. S., Moisen, G. G., Megown, K. A., Nelson, V. J., Freeman, E. A., Patterson, P. L., Finco, M., Brewer, K. and Menlove, J. 2009. Nevada photo-based inventory pilot (NPIP) photo sampling procedures. Gen. Tech. Rep. RMRS-GTR-222. – U.S. Dept of Agriculture, Forest Service, Rocky Mountain Research Station, 30p.

Frescino, T. S. and Patterson, P. L. 2017. Number of points needed for image-based change estimation. In: Healey, S. P. and Berrett, V. M. (eds), Doing more with the core: proceedings of the 2017 Forest Inventory and Analysis (FIA) Science Stakeholder Meeting; 2017 October 24–26; Park City, UT. Proc. RMRS-P-75. – U.S. Dept of Agriculture, Forest Service, Rocky Mountain Research Station, pp. 16–21.

Frescino, T. S., McConville, K. S., White, G. W., Toney, J. C. and Moisen, G. G. 2022a. Small area estimates for national applications: a database to dashboard strategy for FIA using FIESTA. – Front. For. Global Change 5.

Frescino, T. S., Moisen, G. G., Toney, J. C. and White, G. W. 2022b. FIESTA: forest inventory estimation and analysis. R package ver. 3.5.0. – https://CRAN.R-project.org/package=FIESTA.

Frescino, T. S., Toney, J. C. and White, G. W. 2022c. FIESTAutils: utility functions for forest inventory estimation and analysis. R package ver. 1.1.4. – https://github.com/USDAForestService/FIESTAutils.

GDAL/OGR contributors. 2019. GDAL/OGR geospatial data abstraction software library. – Open Source Geospatial Foundation.

Grothendieck, G. 2017. sqldf: manipulate R data frames using SQL. R package ver. 0.4-11. – https://CRAN.R-project.org/package=sqldf.

Dowle, M. and Srinivasan, A. 2019. Data.table: extension of 'data.frame'. R package ver. 1.12.2. – https://CRAN.R-project.org/package=data.table.

Hill, A. P. and Field, C. B. 2021. Forest fires and climate-induced tree range shifts in the western US. – Nat. Commun. 12: 6583.

Holden, Z. A., Jolly, W. M., Swanson, A., Warren, D. A., Jencso, K., Maneta, M. and Landguth, E. L. 2019. TOPOFIRE: a topographically resolved wildfire danger and drought monitoring system for the conterminous United States. – Bull. Am. Meteorol. Soc. 100: 1607–1613.

LaBau, V. J. 2007. A history of the forest survey in the United States: 1830–2004. – U.S. Dept of Agriculture, Forest Service.

Law, B. E., Berner, L. T., Buotte, P. C., Mildrexler, D. J. and Ripple, W. J. 2021. Strategic forest reserves can protect biodiversity in the western United States and mitigate climate change. – Commun. Earth Environ. 2: 254.

McConville, K. S., Breidt, J. F., Lee, T. C. M. and Moisen, G. G. 2017. Model-assisted survey regression estimation with the Lasso. – J. Surv. Stat. Methodol. 5: 131–158.

McConville, K. S., Tang, B., Zhu, G., Cheung, S. and Li, S. 2018. mase: model-Assisted Survey Estimation. R package ver. 0.1.2. – https://cran.r-project.org/package=mase.

McConville, K. S., Moisen, G. G. and Frescino, T. S. 2020. A tutorial in model-assisted estimation with application to forest inventory. – Forests 11: 244.

Molina, I. and Marhuenda, Y. 2015. sae: an R package for small area estimation. – R J. 7: 81–98.

Patterson, P. L. 2012. Photo-based estimators for the Nevada photo-based inventory. Res. Pap. RMRS-RP-92. – U.S. Dept of Agriculture, Forest Service, Rocky Mountain Research Station, 14p.

Pebesma, E. 2018. Simple features for R: standardized support for spatial vector data. – R J., 10: 439–446.

Rao, J. N. and Molina, I. 2015. Small area estimation. – John Wiley and Sons.

Rintoul, M. A., Maebius, S., Alvarado, E., Lloyd-Damnjanovic, A., Toyohara, M., McConville, K. S., Moisen, G. G. and Frescino, T. S. 2020. An alternative post-stratification scheme to decrease variance of forest attribute estimates in the Interior West. – In: Brandeis, T. J. (eds). Celebrating progress, possibilities, and partnerships: proceedings of the 2019 Forest Inventory and Analysis (FIA) Science Stakeholder Meeting; November 19–21, 2019; Knoxville, TN. e-Gen. Tech. Rep. SRS-256. – U.S. Dept of Agriculture Forest Service, Southern Research Station, pp. 268–276.

Rudis, V. A. 2003. Comprehensive regional resource assessments and multipurpose uses of forest inventory and analysis data, 1976 to 2001: a review. Gen. Tech. Rep. SRS-70. – U.S. Dept of Agriculture, Forest Service, Southern Research Station, 129p.

Schroeder, T. A., Healey, S. P., Moisen, G. G., Frescino, T. S., Cohen, W. B., Huang, C., Kennedy, R. E. and Yang, Z. 2014. Improving estimates of forest disturbance by combining observations from Landsat time series with U.S. forest service forest inventory and analysis data. – Remote Sens. Environ. 154: 61–73.

Seeing the forest for the trees. 2022. – Nat. Ecol. Evol. 6: 341.

Stanke, H., Finley, A. O., Weed, A. S., Walters, B. F., and Domke, G. M. 2020. rFIA: an R package for estimation of forest attributes with the US forest inventory and analysis database. – Environ. Model Softw. 127: 104664.

Stanke, H., Finley, A. O., Domke, G. M., Weed, A. S., and MacFarlane, D. W. 2021. Over half of western United States' most abundant tree species in decline. – Nat. Commun. 12: 451.

Stevens, D. L. and Urquhart, S. 2000. Response designs and support regions in sampling continuous domains. – Environmetrics 11: 13–41.

Thompson, M. T. 2017. Assessing the impact of a mountain pine beetle infestation on stand structure of lodgepole pine forests in Colorado using the forest inventory and analysis annual forest inventory. – J. For. 15: 270–275.

Tinkham, W. T., Mahoney, P. R., Hudak, A. T., Domke, G. M., Falkowski, M. J., Woodall, C. W. and Smith, A. M. 2018. Applications of the United States forest inventory and analysis dataset: a review and future directions. – Can. J. For. Res. 48: 1251–1268.

USDA. 2014. Farm bill. – https://www.congress.gov/113/plaws/publ79/PLAW-113publ79.pdf.

Westfall, J. A., Coulston, J. W., Moisen, G. G. and Andersen, H. E. (eds). 2022. Sampling and estimation documentation for the enhanced forest inventory and analysis program: 2022. Gen. Tech. Rep. NRS-GTR-207. – U.S. Dept of Agriculture, Forest Service, Northern Research Station, 129p.

White, G. W., McConville, K. S., Moisen, G. G. and Frescino, T. S. 2021. Hierarchical Bayesian small area estimation using weakly informative priors in ecologically homogeneous areas of the interior Western forests. – Front. For. Glob. Change 4: 178.

Wickham, H. 2018. httr: tools for working with URLs and HTTP. R package ver. 1.4.0. – https://CRAN.R-project.org/package=httr.

Wickham, H. and Müller, K. 2018. DBI: R database interface. R package ver. 1.0.0. – https://CRAN.R-project.org/package=DBI.

Wojcik, O., Olson, S. D., Nguyen, P. V., McConville, K. S., Moisen, G. G. and Frescino, T. S. 2022. GREGORY: a modified generalized regression estimator approach to estimating forest attributes in the interior Western US. – Front. For. Global Change 230.

Woodall, C. W., Monleon, V. J., Fraver, S., Russell, M. B., Hatfield, M. H., Campbell, J. L. and Domke, G. M. 2019. The downed and dead wood inventory of forests in the United States. – Sci. Data 6: 180303.

Woodall, C. W. and Weiskittel, A. R. 2021. Relative density of United States forests has shifted to higher levels over last two decades with important implications for future dynamics. – Sci. Rep. 11: 18848.