



Tree species distribution in the United States Part 1

Rachel Riemann ^a, Barry T. Wilson ^b, Andrew J. Lister ^c, Oren Cook^{a*} and Sierra Crane-Murdoch^{a†}

^aUSDA Forest Service, Northern Research Station, Troy, NY, USA; ^bUSDA Forest Service, Northern Research Station, St. Paul, MN, USA; ^cUSDA Forest Service, Northern Research Station, Newtown Square, PA, USA

ABSTRACT

The distribution and local abundance of tree species constitute basic information about our forest ecosystems that is relevant to understanding their ecology, diversity, and relationship to people. The US Forest Service conducts a forest inventory across all forest lands in the United States. We developed geospatial models of forest attributes using this sample-based inventory which make this information available for an even wider variety of applications. From these modeled datasets, we created a series of maps for 24 US states in an effort to connect more people to trees, the datasets, and the scientific research behind them. Presenting these maps in an attractive way invites engagement. The sidebar text is presented in accessible scientific language that clearly defines terms, guides readers in interpreting the maps and histograms, and provides source details and links. The resulting maps are inviting, informative, and accessible to a broad range of people of different ages and backgrounds.

ARTICLE HISTORY

Received 30 March 2018
Revised 8 August 2018
Accepted 15 August 2018

KEYWORDS

Tree species; basal area; education; engagement; forest inventory; map design

1. Introduction

Trees occur in a wide variety of ecosystems, from swamplands to mountain tops. Their geographic distribution and growth are primarily limited by a combination of abiotic factors like climate, soil, and moisture gradients, as well as by competition with other plants and species interactions with animals and fungi. Individual tree species have evolved distinct forms and capabilities to survive and thrive in different conditions. Thus, the distribution of trees and individual tree species can tell us a lot about the types of ecosystems that may be present, including associated plant species, wildlife, and hydrologic and soil processes. Individual tree species may also have high economic value, or cultural use, or have variable susceptibility to insect pests or human activities because of their characteristics or location. The US Forest Service (USFS) Forest Inventory and Analysis program (FIA) collects forest tree data using a spatially-balanced network of over 350,000 forest inventory field plots distributed across the 48 contiguous United States (CONUS), Hawaii, and parts of Alaska. Data collected at each periodically-revisited plot include information on the status of and change in tree size, structure, and health (Bechtold & Patterson, 2005), as well as information on the temporal dynamics of other site factors such as land use, cover, site quality and forest type. The inventory data contained within the FIA database are a representative, nationally consistent,

high-quality statistical sample of US forest lands, and have long been used for reporting on forest status and trends and for calibration and validation of forest mapping efforts.

Building on previous work by Ohmann and Gregory (2002), Blackard et al. (2008), Riemann Hershey (2000), and others, Wilson, Lister, and Riemann (2012) developed an efficient approach for modeling and mapping forest inventory attributes, such as tree species distributions, over large spatial domains using FIA plot data. These modeled raster datasets make the FIA data more accessible and versatile for visual and geospatial analysis. However, geospatial datasets alone are not sufficient. Although some in the research community know and use the data, many other potential users may not be aware that such information exists or where to find it. Many may not even be aware that the USFS has such an extensive plot network for monitoring forests. Increasing awareness and use requires more effective links between the information available and these broader user communities. An initial tree species distribution map was created for New York, USA in 2014 by Riemann, Wilson, Lister, Cook, and Crane-Murdoch (2014) and this endeavor revealed considerable desire for this type of resource throughout the education and outreach communities. The New York map was requested for a variety of educational efforts, including public workshops on invasive species or forest management; teacher training in

CONTACT Rachel Riemann rriemann@fs.fed.us USDA Forest Service, Northern Research Station, 425 Jordan Road, Troy, NY 12180, USA

*Currently founder/owner of Red Team Solar, Greenwich, New York.

†Currently freelance writer, Hood River, WA.

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the national Project Learning Tree curriculum; traveling exhibits to inner-city schools; upper elementary school through college-level science, biology, living environment, earth science and forestry classes; and for raising awareness through display, distribution, and use as a discussion backdrop by environmental education and forest landowner organizations. This experience highlighted a strong demand for maps of tree and forest information for education and outreach. To fill this need, we have created maps for 23 additional states, completing an area covering the entire north-eastern quarter of the United States.

As shown in the Main Maps, the primary goal of our project was to produce maps that facilitate access to this wealth of forest and tree species information by teachers, students, conservation educators, natural resource extension educators, and the public. Specific goals included:

- Bring attention to the forests and tree species near which students and the public live and work.
- Create an attractive, engaging visual resource that educators can use to enhance lesson materials and bring the topic of the local natural world into classroom curricula, activities, and discussions – supporting lessons on topics such as scientific data summary and writing, botany, natural resource monitoring, ecology, statistics, math, geography, cartography, or landscape planning.
- Introduce two forms of scientific data summary and presentation – the histogram and the map, and the relationship between the two.
- Introduce some of the basic elements of scientific writing and organization (e.g. methods, sources, references, citation information).
- Publicize the freely available forest information currently collected by the USFS FIA.
- Support natural resource career opportunity exploration activities by highlighting the variety of cooperating natural resource scientists and professionals in the public and private sectors that were behind a project like this.

2. Methods

2.1. Data

Each of the embedded map graphics was produced using individual species datasets available from [Wilson et al. \(2012\)](#). These datasets were modeled using a combination of ecological ordination and nearest neighbor data mining techniques in a predictive machine-learning approach. The method integrated FIA plot data with continuous raster data describing relevant environmental parameters (topography and climate), vegetation phenology derived from dense time series of MODIS satellite imagery (MODerate-resolution Imaging

Spectroradiometer), and finer spatial resolution tree canopy cover data from the National Land Cover Dataset (NLCD). The resulting 250-m pixel size raster datasets provided nationally consistent information on the location, relative abundance, and distribution for 324 individual tree species covering CONUS. Much of the covariance among tree species distribution patterns found on the forest inventory plots is retained in the modeled raster datasets, maintaining general ecological consistency among the tree species datasets. A suite of assessment metrics was applied to each of the modeled datasets at multiple scales ([Riemann, Wilson, Lister, & Parks, 2010](#)). The results of these accuracy assessments are discussed in [Wilson et al. \(2012\)](#) and [Riemann and Wilson \(2014\)](#) and are available along with the modeled data in an online raster data warehouse ([Wilson, Lister, & Riemann, 2013](#)).

Basal area is the area of the cross-section of a tree trunk at 1.37 m (4.5 feet) above the ground. Basal area per hectare is a measure that integrates both tree size and number of trees. The data were first translated from basal area (m²) per hectare of forest land into proportion of total basal area, by dividing each species basal area per hectare by the total live tree basal area per hectare. Basal area proportion is commonly used as a measure of the dominance of that species relative to all trees found in that area.

The modeled raster datasets estimate tree species occurrence across all lands, even in areas where tree cover may be too sparse to be considered forest, whether due to environmental conditions or urban or agricultural land uses. However, as a result such areas likely did not have sufficient FIA plot data to support modeling to the same level of accuracy ([Wilson et al., 2012](#)). To visually mask out these nonforest areas in the maps, we used a forest proportion dataset that was used in the development of the tree species datasets to translate modeled tree basal area into tree basal area per hectare of forest land (Wilson, unpublished dataset). A minimum threshold of at least 40% forest land use was used to define forest areas for these maps, effectively masking out both nonforest areas and areas of sparse tree cover that could not be modeled sufficiently accurately with the available data.

The 2006 version of the National Land Cover Dataset (NLCD) ([Fry et al., 2011](#)) was used for the water overlays in most states. The only exceptions were the Kansas, Minnesota, North Dakota and South Dakota maps in which the National Hydrography Dataset (NHD) was used ([U.S. Geological Survey, 2016](#)). In these four states, the broad regions of very small lakes visible in the NLCD dataset were too visually distracting, particularly with respect to the highly scattered distribution of forest land in those regions. In these four states, the vector NHD dataset was preferred because of the ability to specify a minimum size limit of 5 km² on the lake polygons displayed.

The histograms accompanying each map represent the count of individual pixel values in each basal area proportion class. The histograms were generated from the tree species data after the water and nonforest overlays had been applied.

A 30-m shaded relief map developed from USGS National Elevation Data (NED) (ESRI, 2014) was used to add hillshade to the central map. This was done to add visual depth to the map and take advantage of the frequently strong relationship between topography and landscape scale spatial patterns of trees and species to provide additional information in the central map.

2.2. Study area

Although the tree species geospatial datasets are available for the entire CONUS, this paper focuses on maps developed for the 24 northeastern states, which range from heavily forested in the east to sparse woodlands in the Great Plains states to the west. This area was chosen to correspond with the administrative footprint of the Northern Research Station FIA Unit where the species modeling work originated, and includes approximately 25% of the US land area.

2.3 Map design

The primary feature of each map is a circle of individual species maps around a central land use map (Figure 1). The central map displays the distribution of forest versus non-forest land use and topography within the state (Figure 2). The individual species maps display each species' relative dominance in terms of its proportion of the total live tree basal area for each pixel (Figure 3). Thresholds defining the color classes reflect those used by the widely used Society of American Foresters' guide to forest cover types of the United States and Canada (Eyre, 1980) to identify those areas in which an individual tree species 'occurs only occasionally' (less than 5% of the total live tree basal area), is a minor component (5–20% of the total live tree basal area), a major component (20–50% of the total live tree basal area), or a dominant species (>50% of the total live tree basal area) in the stand. Associated with each tree species map is a corresponding histogram which provides additional information about the proportion of the total forest land area occurring in each relative dominance class (Figure 4). The histogram colors and the range of values they span also serve as the legend for interpreting the map. The circular design, primarily with tree leaves, tree names, and individual state maps, is a vehicle used to reinforce the sense of connectedness between tree species, and encourages younger students to look for patterns where the individual tree species maps are similar and different. The individual tree species map at the top of the circle

is the species with the greatest amount of basal area in the state. All other individual species maps are arranged as much as possible to group species that commonly occur together in the landscape. The number of species included varies by state based on the size of each state and the distribution of forest within it. The tree leaves (Ellis, L. 2014, Tree leaf paintings), add familiarity to the maps (almost every child has seen a tree leaf), and introduce another aspect of tree species similarity and variety. Though simple, the tree leaf artwork contributes substantially to increasing each map's effectiveness with engaging new users as well as providing additional information.

In order to increase the utility of the maps for multiple age groups and experience levels, a column of explanatory text is included along the left side of the maps (Figure 1). The goal was to provide enough information so that even readers unfamiliar with forestry, natural resources, or histogram data summaries could appreciate and learn from the information presented in the maps, while at the same time providing gentle exposure to some elements of the scientific method. To that end, the text starts with only basic information and gradually increases in complexity using principles of scientific writing and organization. An introduction to reading and interpreting the map elements is provided first, along with some engaging statistics. This text begins very simply so as to be readily understood by even upper elementary students and their teachers, and indeed such students may stop after the first three paragraphs and the example histogram. Following this is a brief description of the data and the methods behind the datasets used in the maps, so older students are prompted to ask questions about where the underlying data come from. Finally, links are provided for those high school and college students interested in using the geospatial datasets in their own projects and analyses. In addition, the map provides full disclosure of all the organizations, roles and individuals who contributed to the creation of this information so students get a sense of the variety of skills and careers related to natural resources that exist and are associated with a project like this.

3. Conclusions

The modeled raster datasets developed by Wilson et al. (2013) provide detailed information on the relative abundance and distribution of individual tree species across the CONUS. These datasets differ from other mapping efforts in the number of tree species modeled, methodological consistency across the country and across tree species, retained covariance structure among tree species, and in the comparative accuracy assessment results available for each tree species at multiple scales. The raster format of these data serves to translate the forest inventory information into a format



Figure 3. Example of individual species inset map depicting each species' relative dominance in terms of its proportion of the total live tree (e.g. chestnut oak in Pennsylvania).

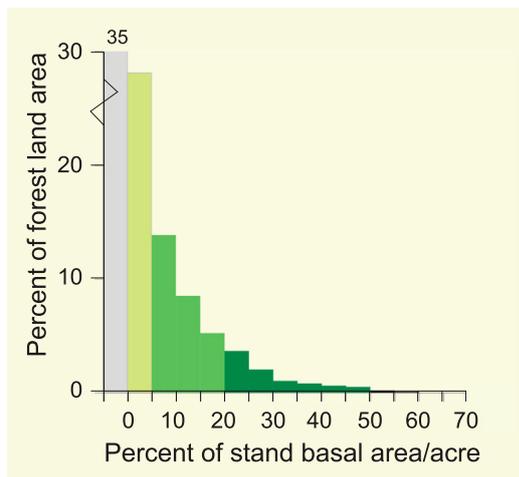


Figure 4. Example of histogram/legend associated with each species inset map (e.g. chestnut oak in Pennsylvania).

- (1) The maps are designed to make information collected by the USFS FIA program accessible to people with a broad spectrum of environmental science or forestry experience.
- (2) The visually appealing map design, made attractive by a circular layout and the addition of leaf graphics, draws people in even if they did not think they were interested in trees.
- (3) The maps are designed to feel initially very simple and easy to understand, keeping people engaged, while providing an increasing amount and depth of information as users study them.
- (4) Real-world research can feel remote from classroom activities and public interests. These maps are designed to clearly introduce the people and science behind the end product, as well as provide information and links for classes or individuals to ask their own follow-on research questions.

- (5) These maps represent another tool to connect students, educators and the public with their local ecosystems via learning about trees, which can be thought of as 'charismatic mega-flora.'

Numerous efforts and initiatives exist to try to engage youth and the general public in outdoor experiences, familiarize them with their surrounding natural world, engage youth with science and connect them with natural resource science professionals. These maps and the geospatial datasets behind them provide these forward-thinking efforts with information and resources from their public agencies in the United States and provide the public agencies with a much-needed link to the public they serve.

Plans to further automate the creation of these maps and extend this effort to the remaining states in the United States are in progress.

Software

The individual maps were produced in ESRI ArcMap 10.3.1. The histogram bars were generated in Microsoft Excel 2013. All graphic elements (maps, histograms, leaves, scale bar) were additionally processed in GIMP 2.8.6 to create final versions with transparent backgrounds. The final product was assembled in Adobe Illustrator CC 2018.

Acknowledgements

Thanks are due to the FIA field crews for collecting the necessary field plot data and/or doing quality assurance plots, to the FIA data processing and compilation staff who process the raw data into archive quality databases, to the information management staff who write the programs to

support all of the above, and to Linda Ellis for creating the leaf graphics for all the maps.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Rachel Riemann  <http://orcid.org/0000-0002-2973-525X>
 Barry T. Wilson  <http://orcid.org/0000-0002-1940-7682>
 Andrew J. Lister  <http://orcid.org/0000-0002-0412-7039>

References

- Bechtold, W. A., & Patterson, P. L. (Eds.). (2005). *The enhanced forest inventory and analysis program – national sampling design and estimation procedures (report No. GTR-SRS-80)*. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. <https://doi.org/10.2737/srs-gtr-80>
- Blackard, J. A., Finco, M. V., Helmer, E. H., Holden, G. R., Hoppus, M. L., Jacobs, D. M., ... Tymcio, R. P. (2008). Mapping US forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*, 112, 1658–1677. Retrieved from <https://www.treearch.fs.fed.us/pubs/29860>
- ESRI. (2014). World Shaded Relief [Data set]. Retrieved from http://goto.arcgisonline.com/maps/World_Shaded_Relief
- Eyre, F. H. (1980). *Forest cover types of the United States and Canada*. Washington, DC: Society of American Foresters.
- Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., ... Wickham, J. (2011). Completion of the 2006 national land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 77, 858–864. Retrieved from <https://www.mrlc.gov/nlcd2006.php>
- Ohmann, J. L., & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon USA. *Canadian Journal of Forest Research*, 32, 725–741. Retrieved from <https://www.treearch.fs.fed.us/pubs/4747>
- Riemann, R., & Wilson, B. T. (2014). Characterization and visualization of the accuracy of FIA's CONUS-wide tree species datasets. In A. M. Shortridge, J. P. Messina, S. Kravchenko, & A. Finley (Eds.), *Accuracy 2014: 11th international symposium on spatial accuracy assessment in natural resources and environmental sciences* (pp. 100–106). East Lansing, MI: International Spatial Accuracy Research Association. Retrieved from: <https://www.treearch.fs.fed.us/pubs/46519>
- Riemann, R., Wilson, B. T., Lister, A. J., Cook, O., & Crane-Murdoch, S. (2014). Modeled distributions of 12 tree species in New York (Report No. RMap NRS-5). Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. [Map]. <https://doi.org/10.2737/NRS-RMAP-5>
- Riemann, R., Wilson, B. T., Lister, A. J., & Parks, S. (2010). An effective assessment protocol for continuous geospatial datasets of forest characteristics using USFS Forest Inventory and Analysis (FIA) data. *Remote Sensing of Environment*, 114, 2337–2352. <https://doi.org/10.1016/j.rse.2010.05.010>
- Riemann Hershey, R. (2000). Modeling the spatial distribution of ten tree species in Pennsylvania. In H. T. Mowrer, & R. G. Congalton (Eds.), *Quantifying spatial uncertainty in natural resources* (pp. 119–135). Chelsea, MI: Ann Arbor Press.
- U.S. Geological Survey. (2016). National Hydrography Dataset [Data set]. Retrieved October 2016, from <https://nhd.usgs.gov/data.html>
- Wilson, B. T., Lister, A. J., & Riemann, R. I. (2012). A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *Forest Ecology and Management*, 271, 182–198. Available at: <http://doi.org/10.1016/j.foreco.2012.02.002>
- Wilson, B. T., Lister, A. J., & Riemann, R. I. (2013). *Modeled tree species distributions across the contiguous United States*. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. [Data set]. Available at: <http://doi.org/10.2737/RDS-2013-0013>