

Mapping relative forest maturity and stand development for conservation in continental USA

Brendan Mackey*¹, Dominick A. DellaSala², Patrick Norman¹, Carly Campbell¹, Patrick Comer³ and Brendan M. Rogers⁴

¹Griffith Climate Action Beacon, Griffith University, (email) b.mackey@griffith.edu.au

²Chief Scientist, Wild Heritage, a Project of Earth Island Institute, (email) dominick@wild-heritage.org

³Chief Ecologist, NatureServe, (email) pat_comer@natureserve.org

⁴Woodwell Climate Research Center, (email) brogers@woodwellclimate.org

*co-corresponding author: (email) b.mackey@griffith.edu.au; (phone) +61 7 555 27263

Table of Contents

Data assembly and visualization	1
Expert workshops.....	1
Forest maturity cluster analysis	2
Validation analysis.....	3
Attribution analysis.....	4
Supplementary references	4

Data assembly and visualization

Spatial data sets on key forest attributes for the conterminous USA were identified through a search of published peer-reviewed literature and publicly available government databases.

Available data layers were assessed based on their relevance to the study, spatial resolution, geographic coverage, temporal relevance and accuracy.

Using the assembled spatial data, interaction web-based mapping apps (were created to help workshop elicitation. The mapping apps were created using the Google Earth Engine platform (Gorelick et al., 2017) allowing spatial datasets to be processed and viewed in real time using the program’s app viewer (Figure S11). These apps were used as an exploratory tool in the expert workshops and for post-analysis diagnostics as discussed below.

Expert workshops

A series of five online workshops focused on the major forest regions of the conterminous USA (Pacific Northwest/Pacific Southwest, Rockies, Southwest, Great Lakes, and Eastern) were conducted between September and November 2021, involving a total of over 40 forest and

ecological and conservation expert participants. The key objectives were to obtain expert advice on: (a) the most appropriate level of ecoregion and forest ecosystem classification to use for a continental analysis; (b) potential sources of error and limitation in these data; (c) the suitability of available forest data sets for modelling and analysis; and (d) to provide feedback on preliminary modelling results. We also sought advice on the range and thresholds of forest cover, height and biomass that are characteristic of mature forests in different ecosystem and ecoregion contexts. A final workshop brought all participants together to report back on adjustments from the prior workshops and to seek feedback on preliminary results.

To add into the workshop discussions, an interactive mapping app was developed that enabled participants to interactively view and interrogate key data sets that had been assembled including: Major Forest Type Groups; Ecoregions; forest cover; forest height; forest biomass; land tenure; and broad land management categories. For each Forest Type Group, a GIS layer was generated from a spatial model based on the FIA plot data (Ruefenacht et al., 2008). The forest height and cover layers were presented with integrated sliders whereby users could shift thresholds to show areas exceeding input values. Participants were able to focus on any of the forest type groups in an ecoregion and using sliding scales, map a given level of forest cover, height and biomass. Using this tool, participants were able to set thresholds to map forested areas that matched their local knowledge.

Key recommendations from the workshop included that:

(a) forest height (Potopov et al., 2019), cover (Hansen et al. 2013) and aboveground woody biomass (Harris et al., 2021) should provide useful indicators of forest maturity for a given forest type and ecoregion;

(b) the US Forest Service Major Forest Type Groups are an appropriate level of forest ecosystem representation and are mapped at sufficient spatial resolution for continental analysis, for this study. As the layer was created through the use of the forest inventory and analysis (FIA) (FIA, 2022) plots, they represent an appropriate combination of forest type aggregation and country wide consistency from extensive on-ground plot data;

(c) The identified thresholds for mature forests, which took into account different forest groups and ecoregions, were used as an initial indicative calibration of the three forest data layers; and

(d) The experts also highlighted the impacts of land use, including logging and natural disturbances especially wildfire, and insects and disease, on the level of forest maturity and stand developed.

Forest maturity cluster analysis

To generate a spatial data layer of forest maturity, the 30m pixels were clustered into ordinal classes of relative forest maturity based on the values for forest cover, height and biomass.

First, the three global layers were aligned and clipped for the conterminous USA. As each input layer was developed using the Global Land Analysis and Discovery's (GLAD) Landsat Analysis Ready Data (ARD), they shared a 30m pixel resolution. Spatial vector files of the distribution of the Major Forest Group within each EPA Level III Ecoregion were created using the GDAL

command line tool set (GDAL/OGR Contributors., 2022). To overcome computational memory constraints on Google Earth Engine, the step was processed on a local CPU. The layers - consisting of the spatial coverage of each Major Forest Group per ecoregion (N=782; hereafter Spatial Units of Analysis or SUA) – once generated were exported to Google Earth Engine. Iterating over each combination, quartile values were calculated for forest height, biomass and cover. Using these quartile values, pixels were assigned a score from 0 to 3 for each of the three forest metrics (i.e., cover, height, biomass) and summed to provide a total score of between 0 and 9 with 9 representing the most mature forest for a given SUA. The nine ordinal maturity levels were also classified into three levels of maturity: scores 1-3 Young; 4-6 Intermediate; and 7-9 Mature. Pixels with a value of 0 were classed as Indeterminate.

Due to computation constraints on Google Earth Engine, the maturity score layer was broken into 22 components before exporting for further local CPU processing. To create a seamless country wide forest maturity dataset, Orfeo toolbox's (Grizonnet et al., 2017) BandMathX was used to combine components.

An interactive mapping app was used to visually interrogate the forest maturity layer by members of the team, and experts involved in the workshops. This revealed that forest type groups bordering alpine regions were often categorized as less mature forests in forests that were known to have not been impacted by either logging or natural disturbances. Rather, they are natural have shorter and less dense canopies, and contain less woody biomass than forest of a similar type at lower elevations (Adams and Kolb, 2005; Uzoh and Oliver, 2006). To ensure that naturally marginal forests were excluded from the maturity dataset, the LANDFIRE 2016 National Vegetation Cover layer was used to mask areas of non-forest (LANDFIRE, 2016). Categories used to mask the layer included: Forest and Shrubland; Recently Disturbed or Modified Classes; and Agricultural & Developed Vegetation. Additionally, the Southeastern Coastal Plain Pocosin & Tree Bog was included due to large areas of it meeting forest definition criteria.

Validation analysis

As an approach to validating the modelled forest maturity map, data from the Forest Inventory and Analysis (FIA) were used as these comprise a large number of forest assessment plots (N= 67,141) located throughout conterminous USA (Riley et al., 2021).

FIA plot data for each SUA were extracted from the FIA database using the rFIA package (Stanke, 2021) for R (R Core Team, 2021). These data included live above ground biomass (using the biomass function), stand structure (using the standStruct function) and tree height for each plot with available data. All available FIA plot data found within each SUA were used. To account for the spatial perturbations or “fuzziness” applied to the FIA dataset (Lister, 2002), the plot points were spatially buffered by a 1.61km radius to encompass the uncertainty caused by the perturbations. Only points with a spatial buffer 95% contained within the corresponding SUA were used for validation analysis. For plot locations with multiple samples, the most recent data were used. Where plots were categorized by multiple Structural Stage Classes, the plot was assigned to the class representing the largest proportion of the plot.

The field derived metrics from the FIA plot data were compared with a sample of the pixels that comprise the modelled forest maturity map. The sample of pixels was generated as follows. First, a vector grid was built covering the SUA with cell dimensions of 0.01°x 0.01° using QGIS (QGIS.org). The PostGIS command ST_GeneratePoints was then used to generate a series of random points within each grid cell, to create a dataset of 115,400,000 randomly selected pixels covering all forested areas. At each pixel location, the values for forest height and above ground woody biomass were sampled from the layers used to create the forest maturity dataset, along with the associated SUA. Data was exported to an SQLite database which was then drawn upon for further analysis.

To visualize and compare the modelled and FIA height and biomass data for each SUA, data were grouped into maturity categories – “Young”, “Intermediate” and “Mature” for the modelled forest data and “Pole”, “Mature” and “Late” for the FIA data from Structural Stage Classifications - and summarized using min, max, median, 25% and 75% quantiles. These data were plotted for all those SUA that had a minimum of 10 FIA plots per Structural Stage Classification.

Attribution analysis

An attribution analysis was conducted to explore the impact on forest maturity, and specifically modelled Young forest (maturity scores 1-3), of three natural disturbances for which there were available consistent data for the conterminous USA. The expert workshops wildfire severity and insect and disease dieback as being important factors to consider. Additionally, tornado paths have also been identified as impacting USA forest (Harris et al., 2016)(Harris et al., 2016). A spatial dataset was assembled based on three available sources: (1) high fire severity (Finco et al. 2012); (2) very severe insect and disease outbreaks (USFS, 2022); and (3) tornado paths (NOAA, 2022). The dataset being restricted to a 20-year period between 2000-2019, which represented the longest time period for which data were available from all three sources.

We used the Monitoring Trends in Burn Severity (MTBS) program to map the burn severity and extent of large fires across all lands of the conterminous USA from 1984 to present (Eidenshink et al. 2007). For each 30m resolution pixel, we extracted the annual maximum severity level recorded for the period 2000-2019. Burn severity is recorded in four levels: unburned, low, moderate and high. Due to both the insect and disease and the tornado data being only available in vector formats, these datasets were filtered and rasterized to a 30m pixel resolution.

To calculate the total area of modelled forest severely impacted by each of the three disturbances factor, a Young forest spatial layer was culled from the modelled forest maturity map and masked against each factor. The total area of modelled Young forest impacted by each disturbance factor was then extracted for SUA using exactextractr (Baston, 2021). The area of impacted modelled Young forest was then calculated within each Ecoregion level II and III.

Supplementary references

ADAMS, H.D. & KOLB, T.E., 2005. Tree growth response to drought and temperature in a mountain landscape in northern Arizona, USA. *Journal of Biogeography*, 32(9), 1629-1640.

- BASTON, D., (2021). exactextractr: Fast Extraction from Raster Datasets using Polygons. R package version 0.7.2. <https://CRAN.R-project.org/package=exactextractr>
- FIA 2022. The Forest Inventory and Analysis (FIA) program of the USDA Forest Service. Research and Development Deputy Area of the U.S. Forest Service.
- FINCO M., B., Q., ZHANG Y., LECKER J., MEGOWN K. A. & BREWER, C. K. Monitoring trends and burn severity (MTBS): monitoring wildfire activity for the past quarter century using Landsat data. *In: MORIN R.S & LIKNES G, C., eds. Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012, 2012 December 4-6 2012 Baltimore. US Department of Agriculture, Forest Service, Northern Research Station.*
- GDAL/OGR CONTRIBUTORS., 2022. GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation. URL <https://gdal.org> , DOI: 10.5281/zenodo.5884351
- GORELICK, N., HANCHER, M., DIXON, M., ILYUSHCHENKO, S., THAU, D. & MOORE, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, 18-27.
- GRIZONNET, M., MICHEL, J., POUGHON, V., INGLADA, J., SAVINAUD, M. & CRESSON, R., 2017. Orfeo ToolBox: open source processing of remote sensing images. *Open Geospatial Data, Software and Standards*, 2(1), 1-8.
- HANSEN, M. C., POTAPOV, P. V., MOORE, R., HANCHER, M., TURUBANOVA, S. A., TYUKAVINA, A., THAU, D., STEHMAN, S. V., GOETZ, S. J. & LOVELAND, T. R. 2013. High-resolution global maps of 21st-century forest cover change. *Science*, 342, 850-853.
- HARRIS, N., HAGEN, S., SAATCHI, S., PEARSON, T., WOODALL, C. W., DOMKE, G. M., BRASWELL, B., WALTERS, B. F., BROWN, S. & SALAS, W. 2016. Attribution of net carbon change by disturbance type across forest lands of the conterminous United States. *Carbon balance and management*, 11, 1-21.
- HARRIS, N. L., GIBBS, D. A., BACCINI, A., BIRDSEY, R. A., DE BRUIN, S., FARINA, M., FATOYINBO, L., HANSEN, M. C., HEROLD, M. & HOUGHTON, R. A. 2021. Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*, 11, 234-240.
- LANDFIRE 2016. National Vegetation Classification. U.S. Department of Agriculture Forest Service and U.S. Department of the Interior.
- NOAA 2022. Storm Tornado path maps. *In: ADMINISTRATION, N. O. A. A. (ed.).*
- POTAPOV, P., LI, X., HERNANDEZ-SERNA, A., TYUKAVINA, A., HANSEN, M. C., KOMMAREDDY, A., PICKENS, A., TURUBANOVA, S., TANG, H. & SILVA, C. E. 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, 253, 112165.
- QGIS.ORG, 2022. QGIS Geographic Information System. QGIS Association. <http://www.qgis.org>

RILEY, K.L., GRENFELL, I.C., FINNEY, M.A. & WIENER, J.M., 2021. TreeMap, a tree-level model of conterminous US forests circa 2014 produced by imputation of FIA plot data. *Scientific Data*, 8(1), pp.1-14.

USFS 2022. U.S Forest Service's National Insect and Disease Detection Survey Summary Maps. U.S. Forest Service.

UZOH, F.C. & OLIVER, W.W., 2006. Individual tree height increment model for managed even-aged stands of ponderosa pine throughout the western United States using linear mixed effects models. *Forest Ecology and Management*, 221(1-3), 147-154.

HARRIS, N., HAGEN, S., SAATCHI, S., PEARSON, T., WOODALL, C. W., DOMKE, G. M., BRASWELL, B., WALTERS, B. F., BROWN, S. & SALAS, W. 2016. Attribution of net carbon change by disturbance type across forest lands of the conterminous United States. *Carbon balance and management*, 11, 1-21.