



1 **TreeMap 2020, 2022, and 2023: Improved species range mapping in a tree-level forest**
2 **dataset for the conterminous United States**

3 **Karin Riley¹, Rachel Houtman¹, Scott Zimmer¹, Lila Leatherman², Jamie Peeler^{3,4}, John**
4 **Shaw⁵, Isaac Grenfell¹, Maria Borja Arboleda², Abhinav Shrestha^{2,6}, Ian Housman², and**
5 **Mark Finney¹**

6 ¹ Missoula Fire Sciences Laboratory, Rocky Mountain Research Station, US Forest Service,
7 5775 W Highway 10, Missoula, Montana 59808, USA

8 ² RedCastle Resources, 3556 S 5600 W Suite #1-1102, Salt Lake City, UT 84120, USA

9 ³ Oak Ridge Institute for Science and Education, Oak Ridge Associated Universities, 1299 Bethel
10 Valley Road, Oak Ridge, Tennessee 37830, USA

11 ⁴ Department of Ecosystem and Conservation Sciences, University of Montana, 32 Campus
12 Drive, Missoula, Montana 59812, USA

13 ⁵ Forest Inventory and Analysis, US Forest Service, 507 25th St., Ogden, Utah 84401, USA

14 ⁶ Chair of Ecosystem Dynamics and Forest Management in Mountain Landscapes, School of
15 Life Sciences, Technical University of Munich, Freising, Germany

16 Corresponding author: Karin Riley (karin.l.riley@usda.gov)

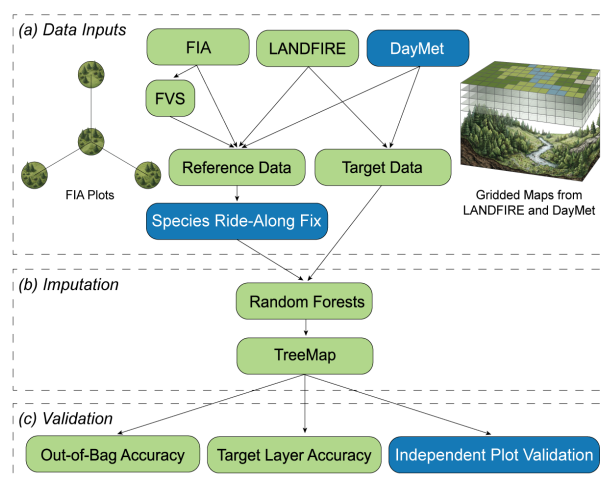
17
18
19
20
21
22
23
24
25
26
27
28



29 Abstract

30 TreeMap is a tree-level model of the forests of the conterminous United States (CONUS) that
 31 provides a spatially continuous, 30 m resolution representation of forest structure and
 32 composition. It uses a random forests imputation approach to assign the most similar Forest
 33 Inventory and Analysis (FIA) plot to each forested pixel in gridded LANDFIRE data based on a
 34 suite of predictor variables that include vegetation, topographic, biophysical, and disturbance
 35 characteristics. By linking every mapped pixel to an FIA plot identifier for all forested areas in
 36 CONUS, TreeMap bridges localized inventory measurements with national-scale analyses and
 37 provides high-resolution estimates of forest structure and composition. In this update, TreeMap
 38 2020, 2022, and 2023 incorporate climatic variables from Daymet in place of LANDFIRE
 39 biophysical variables. Further, we refined the independent validation by limiting the analysis to
 40 FIA plots where heterogeneity was administrative rather than ecological, excluding plots
 41 previously used to validate TreeMap versions that spanned heterogeneous forest conditions.
 42 TreeMap 2020, 2022, and 2023 also implement a new process constraining the pool of FIA
 43 plots available for imputation within each LANDFIRE zone to species documented within that
 44 zone and its adjacent zones, thereby reducing implausible species assignments across
 45 CONUS. TreeMap 2020, 2022, and 2023 are publicly available at [https://doi.org/10.2737/RDS-](https://doi.org/10.2737/RDS-2025-0031)
 46 [2025-0031](https://doi.org/10.2737/RDS-2025-0031) (Zimmer et al., 2025) and <https://doi.org/10.2737/RDS-2025-0032> (Houtman et al.,
 47 2025), and <https://doi.org/10.2737/RDS-2026-0038> (Zimmer et al., 2026), respectively.

48 Graphical Abstract/Key Figure



49



50 **Short Summary (limit 500 characters, non-technical)**

51 New TreeMap datasets provide imputed 30 m gridded maps with modeled tree-level detail for
52 forests across the conterminous United States in 2020, 2022, and 2023. Updated methods
53 improve climate data inputs, independent validation approaches, and species range mapping.
54 These improvements strengthen TreeMap's use in Earth system applications such as carbon
55 cycling, wildfire behavior, and habitat distribution modeling, which depend on accurate
56 representations of forest composition and structure.

57 **1. Introduction**

58 Forest inventory data provide detailed tree-level information that is essential for understanding
59 the current condition and structure of forests (Shaw, 2008; McRoberts et al., 2010). However,
60 these data are collected at discrete and dispersed sample plots rather than across spatially
61 continuous coverages, which limits their direct application to large and heterogeneous
62 landscapes (Riley et al., 2021a, 2022a). For example, the United States Department of
63 Agriculture's Forest Inventory and Analysis (FIA) program maintains a national forest inventory
64 of permanent plots distributed across the United States at a density of roughly one plot per
65 6,000 acres (Bechtold and Patterson, 2005; Burrill et al., 2024). Each FIA plot covers
66 approximately 2/3rd acres (2698 m²) and is typically re-measured every five to 10 years to track
67 changes in forest dynamics over time (Bechtold and Patterson, 2005). Although this sampling
68 design is well suited for reporting national statistics, the extensive unsampled areas between
69 plots limits their utility in spatially contiguous applications common in scientific modeling,
70 adaptive forest management, and policy development.

71 Imputed datasets overcome the spatial discontinuities inherent in forest inventory data by
72 extending detailed tree-level information across unsampled areas, thereby producing spatially
73 continuous representations of forest conditions (Riley et al., 2021a, 2022a). Broadly, imputed
74 datasets scale detailed information from limited sample datasets to wall-to-wall gridded surfaces
75 through statistical modeling. A variety of methods have been used to model forest attributes
76 across landscapes, including small area estimation (Wiener et al., 2021), regression models
77 (Anderegg et al., 2022), and machine learning algorithms (Blackard et al., 2008; Riley et al.,
78 2021a, 2022a). Among these approaches, imputation via random forests has gained particular
79 prominence due to its flexibility and predictive performance in complex forest ecosystems, as it



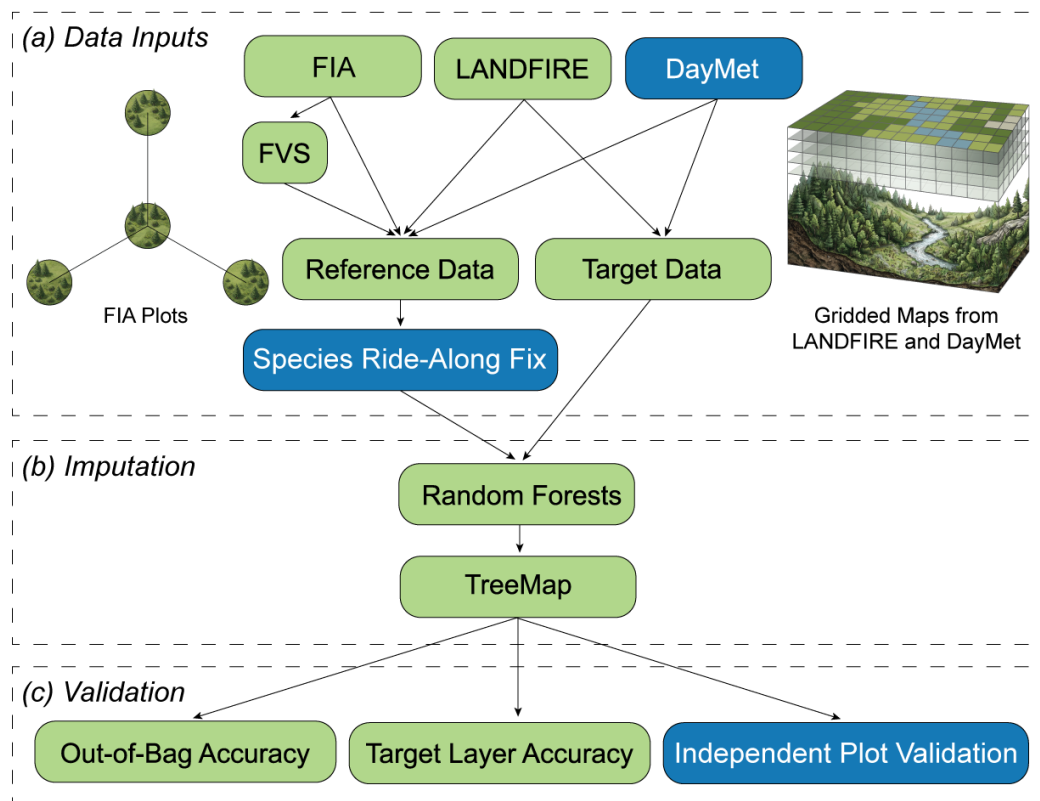
80 can model nonlinear relationships, manage missing data, and handle both categorical and
 81 continuous variables (Breiman, 2001; Cutler et al., 2007).

82 TreeMap 2014 (Riley et al., 2019) was the first national-scale forest product to apply random
 83 forests to link forest inventory data with gridded datasets to generate a spatially explicit, tree-
 84 level forest dataset for the conterminous United States (CONUS). In TreeMap 2014's imputation
 85 design, random forests was used to assign the most similar FIA plot to each pixel in a 30 m grid
 86 (Riley et al., 2021a, 2022a). Using the imputed FIA plot ID, users can link TreeMap to FIA's
 87 extensive database, which include records for individual trees, downed woody material
 88 encountered along transects, and litter and duff depth measurements (Burrill et al. 2024).
 89 Because TreeMap predicts a list of trees for each pixel, users can also derive spatial CONUS-
 90 wide estimates of forest attributes including biomass and stand density. As a result, TreeMap
 91 has since supported Earth system applications like wildfire risk assessments (Peeler et al.,
 92 2023), carbon stock analyses (Chivhenge et al., 2025; Hurteau et al., 2025), and snag hazard
 93 mapping (Riley et al., 2022b; USDA Forest Service Strategic Analytics Branch, 2026).

94 Here, we introduce TreeMap editions for 2020 (Zimmer et al., 2025), 2022 (Houtman et al.,
 95 2025), and 2023 (Zimmer et al., 2026). These updated releases reflect more recent landscape
 96 conditions across CONUS and introduce improvements to the TreeMap methodology (Figure 1).
 97 Notably, we replaced LANDFIRE biophysical variables with gridded climatic datasets from
 98 Daymet (Thornton et al., 2021, 2022) with TreeMap 2020, 2022, and 2023. The LANDFIRE
 99 program produces gridded maps at 30 m resolution of forest and fuel characteristics frequently
 100 used in fire modeling, as well as a set of topographic and disturbance grids (Finney et al., 2011;
 101 La Puma 2023). LANDFIRE's datasets are developed using field data, satellite imagery,
 102 biophysical gradients, and predictive modeling to estimate vegetation attributes (La Puma
 103 2023). We leverage many of LANDFIRE's datasets as predictors in order to create TreeMap
 104 versions that are consistent with fire modeling work (Finney et al., 2011; Riley and Loehman,
 105 2016a), including national fire risk components. In addition to updating the bioclimatic variables,
 106 we also refined the independent validation by limiting the analysis to FIA plots where
 107 heterogeneity was administrative rather than ecological, excluding plots previously used to
 108 validate TreeMap versions that spanned heterogeneous forest conditions. Finally, TreeMap
 109 2020, 2022, and 2023 implement a new process we call a "species ride-along fix", which
 110 prevents an FIA plot from being imputed in a LANDFIRE zone if the plot does not contain tree
 111 species found within that zone or adjacent zones. This adjustment prevented a small but
 112 meaningful number of ecologically implausible species assignments found in earlier TreeMap



113 versions, where plots were imputed outside their local geographic range even though their
 114 vegetation group was correctly represented.



115
 116 **Figure 1.** Conceptual model of the random forests workflow used to produce TreeMap 2020,
 117 2022, and 2023, with additions or improvements from previous TreeMap versions highlighted in
 118 blue. (a) Data inputs include observed forest attributes from FIA plots and gridded maps from
 119 LANDFIRE and Daymet, which together form the reference and target datasets. Two new
 120 additions include using gridded climate data from Daymet, as well as creating a species ride-
 121 along fix that corrected rare cases where FIA plots were imputed to locations containing tree
 122 species well outside their known ranges. (b) These data inputs were used to impute forest
 123 attributes across CONUS using a random forests model. (c) The resulting TreeMap product was
 124 validated using out-of-bag accuracy, target layer accuracy, and independent plot validation.
 125 Independent plot validation was improved from previous TreeMap versions by limiting multi-
 126 condition plots to those varying only in administrative attributes, ensuring each validation plot
 127 represented homogeneous ecological conditions at the pixel scale.



128 2. Methods

129 2.1 *Random forests*

130 Random forests (Breiman 2001; Cutler et al., 2007) is an ensemble machine learning algorithm
 131 that combines the predictions of many decision trees to produce robust predictions that
 132 outperform any single tree. Each decision tree utilizes a random sample of the training
 133 observations and partitions them by selecting the predictor variable that most effectively
 134 reduces variance in the response variable at each split in the decision tree. At each split, a
 135 different random subset of predictor variables is considered, which reduces correlation among
 136 trees and increases the diversity of the final ensemble. Importantly, this diversity makes the
 137 combined predictions more stable than those of a single tree and helps prevent overfitting
 138 (Breiman 2001; Cutler et al., 2007). Model performance is assessed using “out-of-bag”
 139 observations that are not included when decision trees are being built, which provides an
 140 estimate of prediction error.

141 TreeMap uses a random forests model to link a reference dataset consisting of FIA plot
 142 observations to spatially continuous gridded target datasets. We trained the model on the
 143 statistical relationships in the observed forest conditions using the reference dataset, which
 144 consists of one row per FIA plot listing the values of each predictor variable (a set of
 145 topographical, biophysical, vegetation and disturbance variables (Table 1)). We then applied the
 146 trained model to the target dataset, where each 30 m pixel in the target dataset was assigned
 147 the FIA plot identifier (also referred to as the control number (CN)) that most often occurred in
 148 the terminal bucket of the 300 decision trees with the pixel observations. By predicting an FIA
 149 plot identifier rather than modeling individual forest attributes (e.g. biomass or number of live
 150 trees), we ensured that each pixel inherited a complete and internally consistent set of field-
 151 measured variables. As a result, our approach produces a spatially explicit, tree-level
 152 representation of forest structure and composition across CONUS.

153 2.2 *Reference data*

154 TreeMap’s reference dataset consisted of measurements from FIA as well as a suite of
 155 bioclimatic variables (Table 1). Each FIA plot’s footprint consists of four 1/24 acre subplots
 156 (1/6th acre in total) arranged with one at the center and three positioned at 120° intervals at 120
 157 feet from the center (Bechtold and Patterson, 2005). Within each subplot, data are collected on
 158 forest attributes including but not limited to overstory trees (live trees $\geq$ 1 inch DBH and dead



159 trees ≥ 5 inches DBH), understory vegetation, downed woody debris, litter, duff, soils, and
160 disturbance (Bechtold and Patterson, 2005). A 1/300-acre microplot (1/75th acre in total) is
161 located within each subplot to inventory saplings <1 inch DBH and tally seedlings (Bechtold and
162 Patterson, 2005). FIA plots are re-measured on a rotating schedule every 5 to 7 years in the
163 East and 10 years in the West.

164 To build TreeMap's reference dataset, we applied a series of filtering steps resulting in a
165 reference dataset composed of 69,800 FIA plots. We retained only single-condition plots, which
166 represent relatively homogenous conditions across the plot; multi-condition plots, which were
167 used in the validation step, can include, for example, multiple land uses, stand size classes, or
168 forest types (Bechtold and Patterson, 2005). We also only retained plots with live tree canopy
169 cover greater than or equal to 10% to align with LANDFIRE's definition of forested cover (La
170 Puma 2023). Although true FIA plot locations are not publicly available to protect confidentiality,
171 a Memorandum of Agreement with FIA allowed us to use exact coordinates for building
172 TreeMap. Predictor variables for the reference data were derived from the FIA databases, via
173 overlay with LANDFIRE vegetation and disturbance grids or Daymet grids, using LANDFIRE's
174 AutoKey system, and utilizing the Forest Vegetation Simulator (FVS; see Table 1 for specificity;
175 Dixon 2002)

176 **Table 1.** Variables used to build the reference and target datasets for TreeMap 2020, 2022, and
177 2023.

Variable name (Units)	Source for reference dataset	Source for target dataset
Latitude (meters)	FIA program through Memorandum of Agreement	Pixel centroid of LANDFIRE 2020
Longitude (meters)	FIA program through Memorandum of Agreement	Pixel centroid of LANDFIRE 2020
Slope (degrees)	FIA DataMart, Condition table, SLOPE, section 2.4.33, or if unavailable via overlay of plot coordinates with	LANDFIRE 2020 Slope



	LANDFIRE 2020 Slope	
Northness	Cosine transformation of FIA DataMart, Condition table, ASPECT, section 2.4.34, or if unavailable via overlay of plot coordinates with LANDFIRE 2020 Slope	Cosine transformation of LANDFIRE 2020 Aspect
Eastness	Sine transformation of FIA DataMart, Condition table, ASPECT, section 2.4.34, or if unavailable via overlay of plot coordinates with LANDFIRE 2020 Slope	Sine transformation of LANDFIRE 2020 Aspect
Elevation (meters)	FIA DataMart, Plot table, ELEV, section 2.3.22	LANDFIRE 2020 Elevation
Existing Vegetation Cover (percent)	Estimated from FVS StrClass keyword, "Total_Cover" field	LANDFIRE Existing Vegetation Cover versions for 2020, 2022, and 2023
Existing Vegetation Height (meters)	Estimated from FVS StrClass keyword, "Stratum_1_Nom_Ht" field	LANDFIRE Existing Vegetation Height versions for 2020, 2022, and 2023
Existing Vegetation Type Group	Keyed from LANDFIRE AutoKey routine (see Riley et al. (2021a) for description)	LANDFIRE Existing Vegetation Type versions for 2020, 2022, and 2023
Disturbance Presence	Binary variable representing absence of disturbance or occurrence of fire, insects, or disease, derived from disturbance codes in FIA	Binary variable representing absence of disturbance or occurrence of fire, insects, or disease, derived from LANDFIRE disturbance



	DataMart, sections 2.4.37, 2.4.39, and 2.4.41 (see Riley et al. (2021a) for descriptions)	rasters (see Riley et al. (2021a) for descriptions)
Time Since Disturbance	Derived from disturbance year in FIA DataMart, sections 2.4.38, 2.4.40, and 2.4.43 (see Riley et al. (2021a) for descriptions)	Derived from LANDFIRE disturbance rasters (see Riley et al. (2021a) for descriptions)
Precipitation (mm)	Daymet annual normal from 1981-2010, extracted to plot location	Daymet annual normal from 1981-2010
Solar Radiation (W/m ²)	Daymet annual normal from 1981-2010, extracted to plot location	Daymet annual normal from 1981-2010
Snow Water Equivalent (kg/m ²)	Daymet annual normal from 1981-2010, extracted to plot location	Daymet annual normal from 1981-2010
Maximum Temperature (°C)	Daymet annual normal from 1981-2010, extracted to plot location	Daymet annual normal from 1981-2010
Minimum Temperature (°C)	Daymet annual normal from 1981-2010, extracted to plot location	Daymet annual normal from 1981-2010
Vapor Pressure (Pa)	Daymet annual normal from 1981-2010, extracted to plot location	Daymet annual normal from 1981-2010
Vapor Pressure Deficit (Pa)	Daymet annual normal from	Daymet annual normal from



	1981-2010 derived following Lobell et al. 2014, extracted to plot location	1981-2010 derived following Lobell et al. 2014
--	--	--

178 2.3 Target data

179 For TreeMap, the target data consist of CONUS-scale rasters of predictor variables that the
 180 random forests model uses to assign observations from the reference dataset to each pixel.
 181 These include topographic, vegetation, and disturbance variables from LANDFIRE (La Puma
 182 2023; Table 1), as well as bioclimatic variables from Daymet (Thornton et al., 2021, 2022). To
 183 create a forest mask, we select all pixels in CONUS from LANDFIRE's Existing Vegetation
 184 Cover that indicate 10% or higher live forest cover. Although LANDFIRE previously released
 185 updates every 2–5 years to reflect vegetation change and disturbance, annual updates have
 186 been available since 2022. Daymet provides gridded estimates of daily surface weather
 187 parameters derived from ground-based meteorological observations and interpolated using a
 188 statistical model that accounts for topographic and physiographic influences on climate,
 189 producing 1 km resolution coverage across North America (Thornton et al., 2021, 2022). We
 190 calculated 30-year climate normals using annual averages from Daymet spanning 1981–2010
 191 (Thornton et al., 2021, 2022). We selected the 1981–2010 window rather than a more recent
 192 period because it better represents the climate under which contemporary forests developed.
 193 We reprojected all climate normals to the LANDFIRE coordinate system and resampled them to
 194 30 m resolution using cubic convolution, truncating any out-of-range values to dataset minimum
 195 or maximum limits.

196 2.4 Imputation

197 We implemented the random forests model for TreeMap using the *yalImpute* package in R
 198 (Crookston and Finley, 2008). Specifically, we built separate models for each of the 67
 199 LANDFIRE zones across CONUS (Figure S1), generating 300 decision trees per zone split
 200 equally across four response variables: Existing Vegetation Cover (EVC), Existing Vegetation
 201 Height (EVH), Existing Vegetation Type Group (EVT-GP), and disturbance presence. The
 202 random forests model was training on the reference data, then observations from each pixel
 203 were run through the 300 decision trees. For each pixel, the FIA plot ID that appeared most
 204 often in the terminal node with the observations from that pixel was chosen for imputation;
 205 appearing most often in the terminal node with the pixel observations indicates greatest



206 similarity amongst the reference observations. TreeMap users can link this FIA plot identifier to
 207 the CN field in the PLOT table or the PLT_CN field in other tables of the FIA database
 208 (<https://apps.fs.usda.gov/fia/datamart/datamart.html>; Burrill et al., 2024) to access FIA's detailed
 209 field measurements.

210 *2.5 Model validation and stochasticity*

211 Because FIA plots in the reference dataset were sampled for each decision tree with
 212 replacement a number of times equal to the number of plots, some plots were randomly held out
 213 from each tree's construction and used to assess out-of-bag (OOB) accuracy. As another form
 214 of accuracy assessment, we produced confusion matrices between LANDFIRE target data and
 215 imputed EVC, EVH, EVT-GP, and disturbance presence, which allowed us to evaluate the
 216 frequency of within-class agreement for each response variable (Riley et al., 2021a, 2022a).
 217 Although the target layers served as inputs to the imputation, comparing them with the imputed
 218 serves as a form of evaluation of model accuracy. Because the imputation assigns a single
 219 reference plot based on similarity across all four response variables and the predictor variables,
 220 the random forests model may be unable to identify a plot that matches all four response
 221 variables. We compared target and imputed data for each LANDFIRE zone.

222 Although OOB accuracy and comparison of target and imputed data indicate how well the
 223 model performed, these methods do not directly evaluate the accuracy of the imputation
 224 compared to ground conditions. To assess pixel-level accuracy against ground measurements,
 225 we conducted an independent validation using multi-condition FIA plots that were not included
 226 in training the model or reference dataset (Riley et al., 2021a, 2022a). We queried multi-
 227 condition plots measured in the year of the dataset and included only those with the following
 228 characteristics: 1) single forest types (based on both the FIA FLDTYPCD and FORTYPCD
 229 variables), 2) single or no disturbance types, and 3) consistent densities (FIA variable MAPDEN
 230 == 1). This yielded 509 plots for independent validation of TreeMap 2020, 434 plots for TreeMap
 231 2022, and 362 plots for TreeMap 2023 whose conditions differed only in administrative
 232 attributes such as land ownership rather than in ecological characteristics.

233 To reconcile the spatial footprint of 30 m TreeMap pixels with 40.2 m radius FIA plots during
 234 validation, we established which TreeMap pixels would fall largely within a plot footprint by
 235 selecting pixels with centroids falling within 40.2 m of plot center. For each plot footprint, we
 236 then calculated the mean value of EVC and EVH across the selected pixels. As a validation



summary, we tallied the number of validation plots with means falling within 10% and 5 m of the EVC and EVH means, respectively. We compared EVT-GP and disturbance presence directly, summing the number of validation plots that contained at least one matching TreeMap pixel within the FIA plot footprint. Due to the potential for confusion across very similar EVT-GPs, we also compared dominant species composition by identifying the two species with the highest basal area for each validation plot and TreeMap pixel. We then counted how many validation plots had at least one TreeMap pixel within the FIA plot radius that shared one of the validation plot's top two species. Although validation with a dataset entirely independent of FIA would offer a completely independent assessment, no currently available dataset exists at the resolution and spatial extent needed to validate TreeMap.

In addition to the validation metrics described above, we evaluated how much the stochastic processes of random forests affected imputed plot assignments and target layer accuracy. For each of a subset of 16 LANDFIRE zones, we generated 20 candidate models using different random seeds and used each model to impute FIA plot identifiers to a random sample of 100,000 pixels per zone. We then compared the imputed EVC, EVH, EVT-GP, and disturbance presence values with the corresponding target layers to calculate accuracy for each target variable. We also tabulated the number of unique FIA plots assigned in each model. Finally, we quantified variability among the accuracy of candidate models within each zone using the coefficient of variation (CV) for all evaluation metrics.

2.6 Species ride-along fix

Earlier versions of TreeMap occasionally assigned FIA plots to locations containing tree species well outside their observed ranges. For example, some plots containing balsam fir (*Abies balsamea*) were imputed to pixels in Montana and Oregon, far outside the species' range. We identified two scenarios that contribute to this issue. First, the only predictor variable that addresses species directly is the EVT-GP, which is a broad categorization of the vegetation community. While the EVT-GP of a given plot may match the EVT-GP of a given pixel, the plot may contain tree species that are out of range. Second, random forests cannot always identify a plot that matches all four response variables and may select a plot that matches EVC, EVH, and disturbance presence, but not EVT-GP. In instances where there are not many plots for a given EVT-GP, the model may have to draw plots from greater distances away, potentially increasing the likelihood of including species outside their ranges.



We developed a new method we refer to as the species ride-along fix for TreeMap 2020, 2022, and 2023 to address this issue. For each LANDFIRE zone, we compiled a master list of the tree species recorded in FIA plots occurring within the zone and adjacent LANDFIRE zones. During imputation for a given LANDFIRE zone, we limited candidate FIA plots to those containing only species on that zone's species list, reducing the likelihood of implausible species assignments. Because this constraint reduced the number of candidate plots and could affect target layer accuracy, we evaluated the trade-off using reference datasets with and without the species ride-along fix. For each TreeMap vintage, we imputed plots for a random sample of 100,000 pixels in each LANDFIRE zone and compared imputed EVC, EVH, EVT-GP, and disturbance presence values with the corresponding target layers. We also developed a case study for the Willamette National Forest, an area with which the authors are familiar, to examine whether the species ride-along fix prevented implausible species assignments based on expert knowledge.

The species ride-along fix reduces the geographic pool of available FIA plots for each LANDFIRE zone, which raised the question of how far imputed FIA plots originated from their assigned pixels. Understanding plot-to-pixel distances helps TreeMap users assess the degree to which imputed forest conditions reflect local conditions and can identify LANDFIRE zones where fewer plots may be available for some combinations of EVC, EVH, EVT-GP and disturbance, resulting in less accurate assignments or assignments from more geographically distant places in some cases. To quantify distance metrics, we converted the TreeMap ID grid for each LANDFIRE zone into a table of x and y coordinates and calculated the Euclidean distance between each pixel centroid and its imputed FIA plot's origin. We then summarized distances at the zone and national levels and converted the distance estimates back to the original TreeMap grid for mapping.

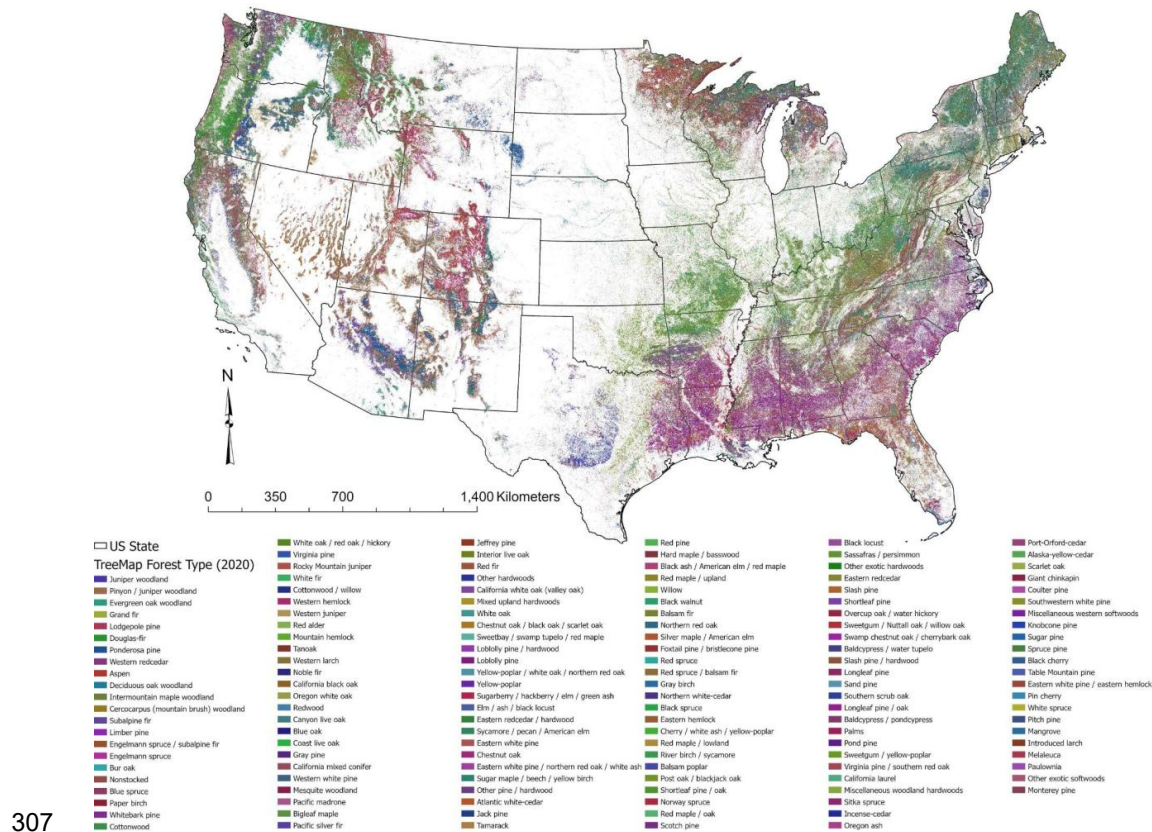
3. Results

3.1 Out-of-bag accuracy, target layer accuracy, and independent plot validation

TreeMap 2020, 2022, and 2023 generated spatially explicit tree-level datasets at the national level (Figure 2). Across all LANDFIRE zones, mean OOB and mean target layer accuracies revealed high model accuracy for all four response variables in TreeMap 2020 (Table S1) and 2022 (Tables S2). Confusion matrices also confirmed high accuracies for EVC (Tables S3-S5), EVH (Tables S6-S8), EVT-GP (Tables S10 and S11), and disturbance presence (Tables S9-S11). Across the four response variables, EVC had the lowest target layer accuracy, at 96.28%



299 in 2020, 96.13% in 2022, and 96.11% in 2023 (Table 2). Notably, we saw that target layer
 300 accuracy for EVC and EVT-GP varied more across LANDFIRE zones than EVH and
 301 disturbance presence (Figure 3). For example, in TreeMap 2020, EVC target layer accuracy
 302 ranged from 69.18% to 99.91% across zones with a standard deviation of 6.25%, while EVT-GP
 303 accuracy ranged from 77.15% to 99.93% with a standard deviation of 4.69%. For EVH and
 304 disturbance presence, we observed lower standard deviations of 2.19% and 0.38%,
 305 respectively. Accuracy measurements in TreeMap 2020, 2022, and 2023 were similar to those
 306 from TreeMap 2016 (Table 2).



308 **Figure 2.** TreeMap forest type for the 2020 landscape year across forested areas of the
 309 conterminous United States. TreeMap 2020, 2022, and 2023 use random forests to impute FIA
 310 plot identifiers to 30 m forested pixels. By linking imputed plot identifiers to FIA's extensive plot-
 311 level databases, users can generate spatially continuous maps of forest attributes, including
 312 species composition, structure, biomass, and forest type.



313 **Table 2.** Mean out-of-bag (OOB) and target layer accuracies among all LANDFIRE zones in
 314 TreeMap 2016, 2020, 2022, and 2023. OOB accuracy is equal in 2020, 2022, and 2023,
 315 because these vintages utilized the same model files.

Validation Method	Year	EVC	EVH	EVT-GP	Disturbance Presence
OOB Accuracy (%)	2016	99.30	99.96	99.65	100.00
	2020, 2022, 2023	99.47	99.98	99.68	100.00
Target Layer Accuracy (%)	2016	97.70	99.60	94.80	99.98
	2020	96.28	99.48	97.24	99.77
	2022	96.13	99.43	97.14	99.70
	2023	96.11	99.26	97.00	99.66

316

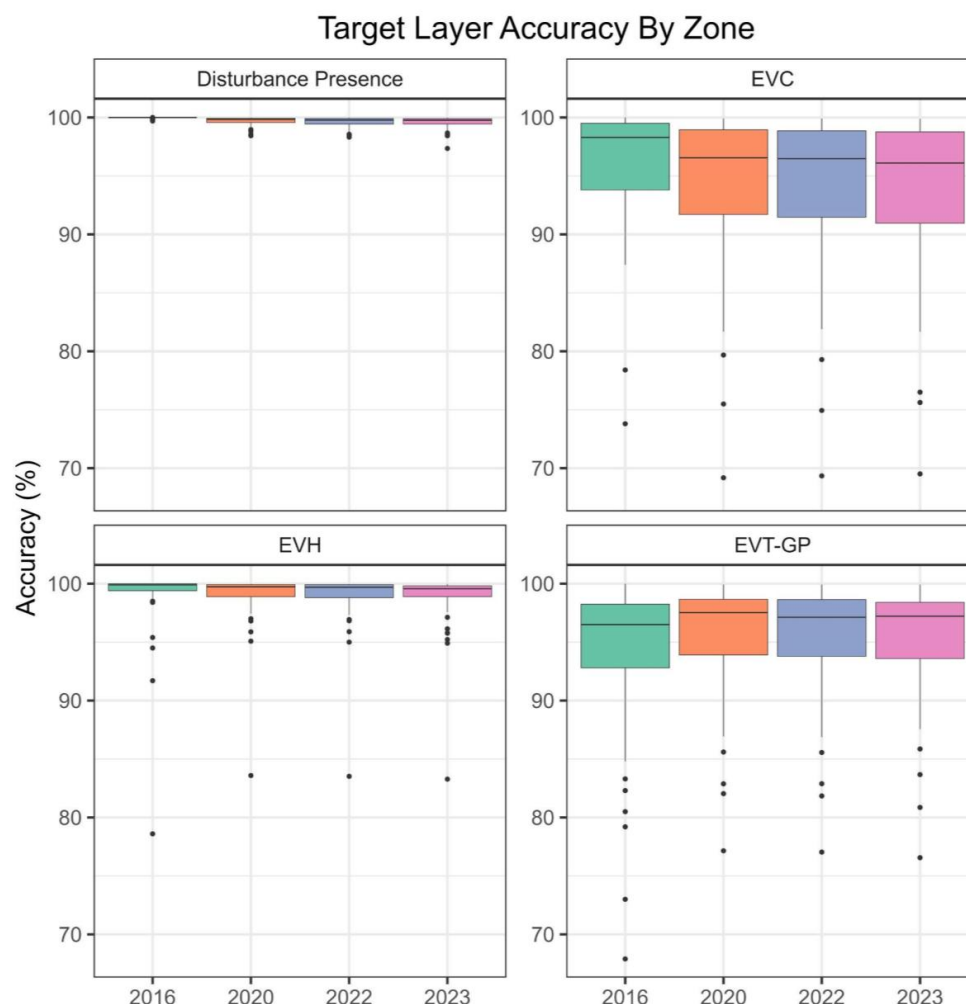


Figure 3. Boxplots of target layer accuracy across all LANDFIRE zones for TreeMap 2016, 2020, 2022, and 2023.

317 To evaluate the effect of the species ride-along fix on target layer accuracy, we compared
 318 TreeMap 2020 models built with and without the ride-along fix across all LANDFIRE zones. We
 319 observed that target layer accuracies fell slightly in models with the species ride-along fix
 320 relative to those without it. In TreeMap 2020, target layer accuracy for EVC decreased the most
 321 at 1.12%, followed by EVT-GP at 0.62%, EVH at 0.04%, and disturbance presence at 0.03%.
 322 The ride-along fix limits the number of FIA plots available for imputation, which likely contributed



323 to this slight decrease in accuracy. The unique number of FIA plots imputed per LANDFIRE
 324 zone fell by 2.93% in models built with the species ride-along fix relative to those built without it.
 325 Despite the restricted pool of plots available in the species ride-along fix models and inherent
 326 model stochasticity, 65.33% of pixels received the same FIA plot assignment in models with and
 327 without the fix.

328 To provide independent validation of model accuracy against field-measured conditions, we
 329 overlaid multi-condition FIA plots that were not used in model development with the imputed
 330 TreeMap rasters. For all pixels with centroids within the FIA plot radius, we found that 44.5-
 331 49.9% of plots had at least one pixel that matched the FIA plot value for EVC (Table 3). These
 332 values were approximately 10% lower than the TreeMap 2016 result. Because target layer
 333 accuracy only declined approximately 1% in TreeMap 2020, 2022, and 2023 compared to
 334 TreeMap 2016, this result may reflect reduced LANDFIRE accuracy for mapping vegetation
 335 cover. Similarly, EVH accuracy ranged from 71.7-77.1% in newer TreeMap versions in
 336 comparison to 80.0% in TreeMap 2016, again suggesting a possible decline in LANDFIRE
 337 vegetation height accuracy. EVT-GP was higher in TreeMap 2023 than in TreeMap 2016
 338 (61.0% versus 51.8%) but lower in TreeMap 2020 and 2022 (53.4% and 47.0%, respectively).
 339 Because only disturbed pixels, which represent a small proportion of the landscape, varied
 340 among TreeMap versions, this result may indicate that the small number of multicondition plots
 341 available for validation is too small for reliable inference. Finally, disturbance presence values
 342 were stable across all TreeMap versions at approximately 86%.

343 Additional independent validation showed improvements in species representation but lower
 344 agreement for vegetation cover and height. Accuracy for capturing at least one of the two
 345 species with highest basal area improved in newer TreeMap versions to 83.7-85.6% compared
 346 to 80.0% in the TreeMap 2016, which may correspond with the species ride-along fix. For EVC,
 347 48.6–52.7% of validation plots had mean imputed values within 10% of the FIA value, which
 348 was 10% lower than TreeMap 2016. For EVH, 50.5–56.4% of validation plots had mean
 349 imputed heights within 5 m of the FIA value, approximately 20% lower than TreeMap 2016.
 350 However, classification errors most often occurred between adjacent classes, suggesting that
 351 differences between imputed and field-measured values were frequently small. The lower EVC
 352 and EVH agreement may reflect reduced LANDFIRE accuracy for vegetation cover and height,
 353 but it could also result from the exclusion of some multi-condition FIA plots that made our
 354 validation sample less representative.



Table 3. Validation agreement between independent FIA plots and TreeMap 2020, 2022, and 2023. For each FIA validation plot, we calculated the mean TreeMap value across pixels within the FIA plot footprint and compared it with the corresponding FIA plot value. Agreement was evaluated as the proportion of plots where the mean TreeMap value matched the FIA value. For EVC and EVH, we also calculated the proportion of plots where the mean TreeMap value was within 10% or 5 m of the FIA value, respectively.

	Proportion of FIA plots where at least one TreeMap pixel in the plot footprint matched the FIA value (%)				Proportion of plots where mean TreeMap value for the plot footprint was within 10% EVC or 5 m EVH of the FIA value (%)			
Year	2016	2020	2022	2023	2016	2020	2022	2023
EVC	57.5	49.9	44.5	48.1	60.9	52.7	51.6	48.6
EVH	80.0	71.7	72.6	77.1	73.0	50.5	52.3	56.4
EVT-GP	51.8	53.4	47.0	61.0	NA	NA	NA	NA
Disturbance presence	87.4	85.1	84.8	87.0	NA	NA	NA	NA
Tree species with two highest basal areas	80.0	83.7	85.5	85.6	NA	NA	NA	NA



Number of validation plots	2,937	509	434	362	2,937	509	434	362
---------------------------------------	-------	-----	-----	-----	-------	-----	-----	-----

361

362 3.2 Model stochasticity

363 To quantify model stochasticity, we evaluated the coefficient of variation (CV) in target layer
 364 accuracies across 20 model runs in 16 sampled LANDFIRE zones. EVC had the greatest mean
 365 CV across zones at 1.29%, followed by EVT-GP at 0.44%, EVH at 0.28%, and disturbance
 366 presence at 0.07%. The CV for the unique number of FIA plots imputed between models was
 367 2.86%. Importantly, we found that target layer accuracies between model runs can vary by
 368 several percent in a given LANDFIRE zone. For instance, in zone 4 the CV for EVC was 5.24%
 369 (mean target layer accuracy = 67.19%; standard deviation = 3.52%). Overall, variability in target
 370 layer accuracy between model runs was low, and in most cases the CV of target layer accuracy
 371 was less than 1%.

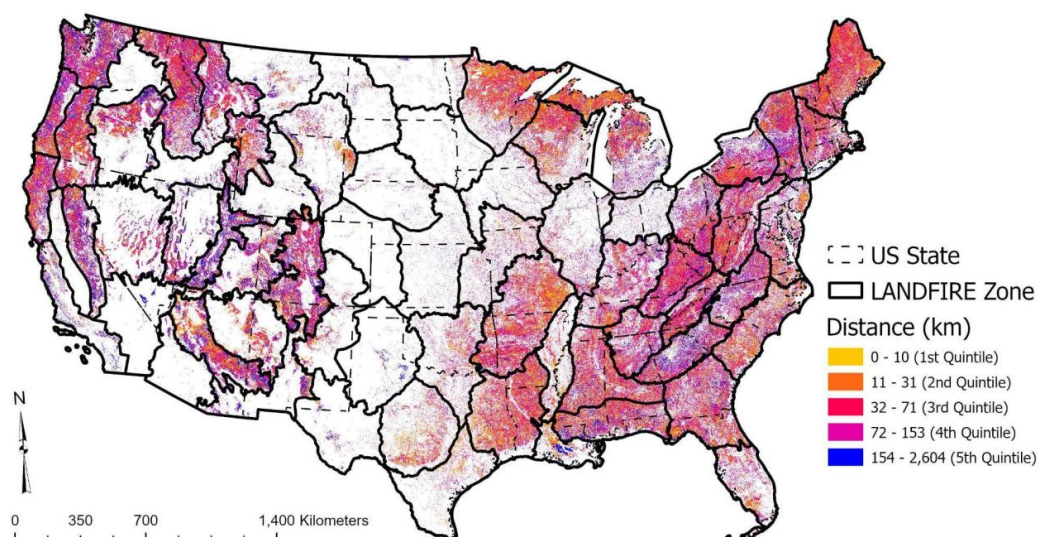
372 3.3 Distance metrics

373 The extent to which FIA plots are imputed across forested pixels captured the spatial reach of
 374 TreeMap and allowed us to evaluate imputation distances. Of the 69,800 single-condition FIA
 375 plots in the reference dataset, 64,743 (92.8%) were imputed at least once across 2,696,548,860
 376 forested pixels in TreeMap 2020 and 65,043 (93.2%) were imputed at least once across
 377 2,687,805,994 forested pixels in TreeMap 2022. In TreeMap 2023, 65,355 (93.6%) were
 378 imputed at least once across 2,720,143,517 forested pixels. These proportions were similar to
 379 proportions observed in TreeMap 2014, where 62,758 (93.5%) of 67,141 single-condition FIA
 380 plots were imputed to 2,841,601,981 forested pixels. However, they were greater than
 381 proportions observed in TreeMap 2016, where 55,609 (84.7%) of 65,652 single-condition FIA
 382 plots were imputed to 2,669,430,013 forested pixels (Riley et al. 2022a).

383 For TreeMap 2020, 2022, and 2023, 95% of pixels were assigned FIA plots sampled less than
 384 325 km, 329 km, and 333 km, respectively, from the pixel centroid (Figure 4). In TreeMap 2020,
 385 we found the average distance from each pixel to its originating plot location was 92.0 km.
 386 Similarly, in TreeMap 2022 and TreeMap 2023, the average distance was 93.3 km and 92.0 km,



387 respectively. The inclusion of plot and pixel coordinates in the reference and target datasets
 388 contributed to this spatial proximity by encouraging nearby FIA plot assignments. However, in
 389 rare instances, FIA plots representing rare or limited forest conditions were imputed to pixels as
 390 far as 2,604 km away.



391

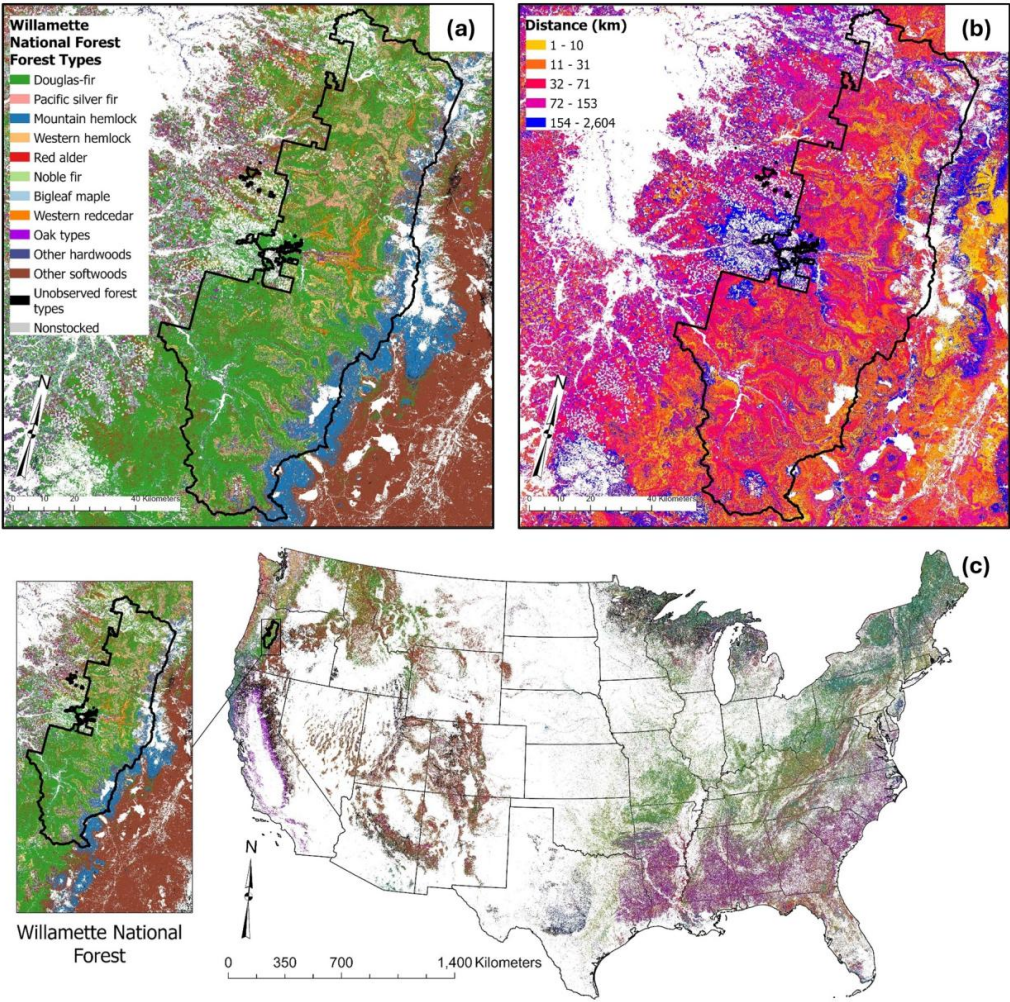
392 **Figure 4.** Distance between imputed TreeMap 2020 plots and originating FIA plot locations
 393 across CONUS.

394 3.4 Species range mapping in the Willamette National Forest

395 Analyzing forest types and distance metrics in the Willamette National Forest (Figure 5) showed
 396 that the species ride-along fix improved species mapping. The four most common forest types in
 397 the area, Douglas fir (*Pseudotsuga menziesii*), Pacific silver fir (*Abies amabilis*), mountain
 398 hemlock (*Tsuga mertensiana*), and western hemlock (*Tsuga heterophylla*), were consistent
 399 between TreeMap 2016 and 2020, comprising 84.9% and 82.3% of the total area, respectively
 400 (Figure 6). Other common forest types such as red alder (*Alnus rubra*), noble fir (*Abies procera*),
 401 bigleaf maple (*Acer macrophyllum*), and western redcedar (*Thuja plicata*) also remained stable
 402 across both versions. However, the species ride-along fix eliminated the most inappropriate
 403 forest type assignments, reducing the proportion of the area imputed to forest types not
 404 endemic to the area from 0.76% in TreeMap 2016 to 0.19% in TreeMap 2020 (Figure 7). In
 405 TreeMap 2016, balsam fir (*Abies balsamea*), California mixed conifer, and paper birch (*Betula*



406 *papyrifera*) accounted for much of this inappropriate area at 0.16%, 0.07%, and 0.07% of the
 407 total area, respectively. In TreeMap 2020, the percent area of these forest types was reduced to
 408 0%, 0.015%, and 0.001%, respectively (Table S12).

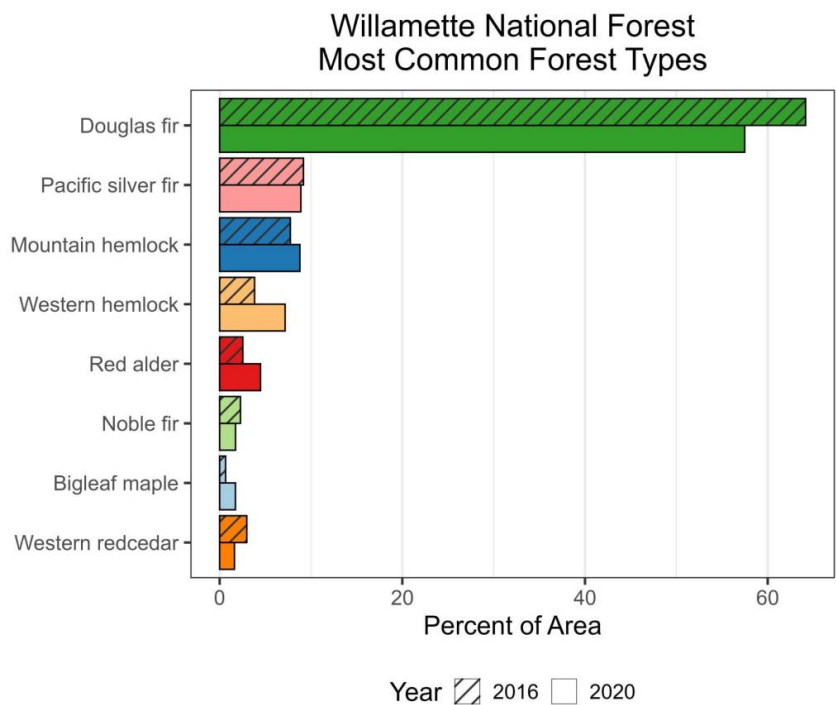


409

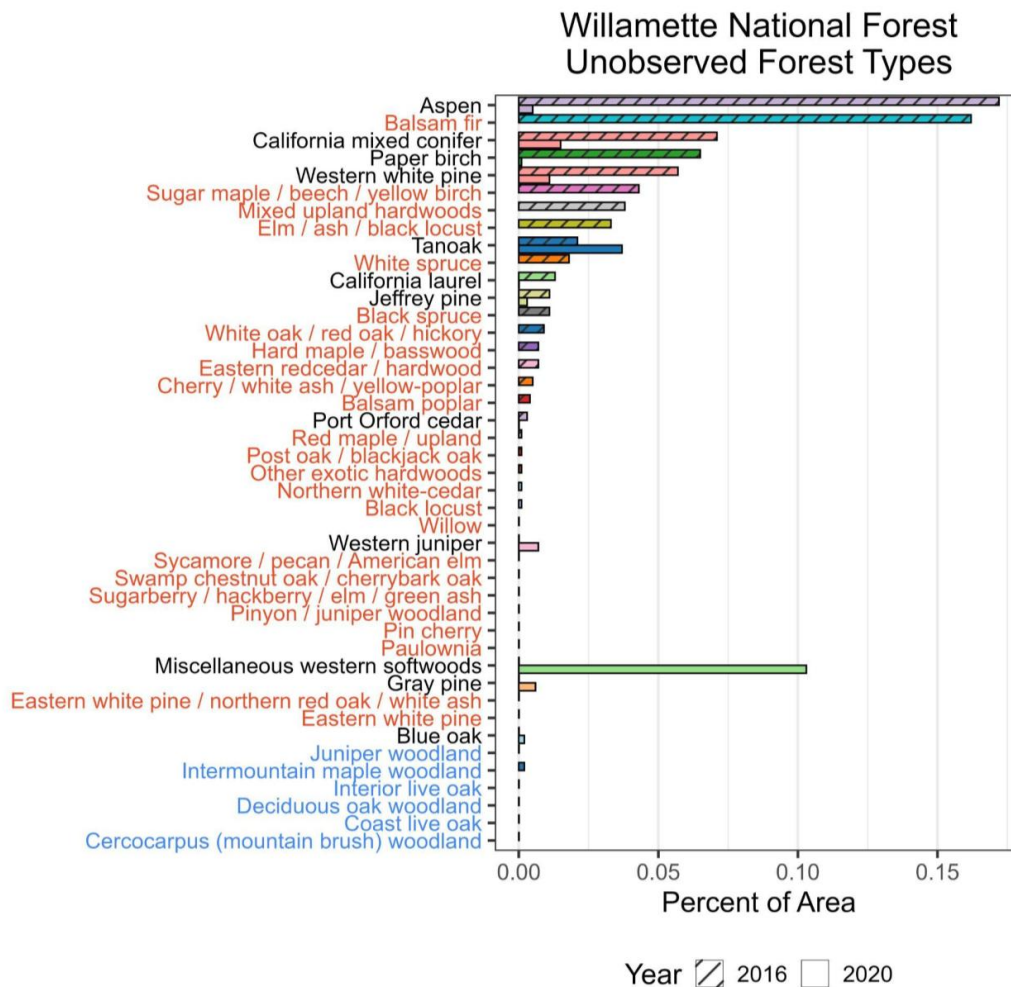
410 **Figure 5.** Case study evaluating TreeMap 2020 forest type and plot-to-pixel distance in and
 411 around the Willamette National Forest, Oregon, United States. (a) FIA-derived forest types show
 412 that the landscape is dominated by Douglas-fir, with mountain hemlock concentrated along the
 413 Cascade crest and western hemlock, Pacific silver fir, western redcedar, noble fir, bigleaf maple,
 414 and red alder also present. Forest types unobserved in FIA plots across the Cascades
 415 ecoregion account for 0.19% of the Willamette National Forest area. (b) Distance between each



416 TreeMap 2020 pixel and the originating FIA plot location assigned during imputation. (c)
417 TreeMap 2020 forest types across the conterminous United States, with the Willamette National
418 Forest shown as the case study location.
419



420
421 **Figure 6.** Percent of Willamette National Forest area imputed to the most common forest types
422 in TreeMap 2016 and TreeMap 2020. The comparison shows broadly similar dominant forest
423 types between products, with Douglas-fir representing the largest share of forested area in both
424 years.



425
426
427 **Figure 7.** Percent of Willamette National Forest area imputed in TreeMap 2016 and 2020 to
428 forest types not observed in FIA plots within the forest. Forest types labeled in black are present
429 in both TreeMap 2016 and TreeMap 2020, forest types labeled in red are present in TreeMap
430 2016 but not TreeMap 2020, and forest types labeled in blue are absent in TreeMap 2016 but
431 present in TreeMap 2020.

432



433 4. Discussion

434 4.1 Applications for an updated TreeMap and improved species range mapping

435 TreeMap 2020, 2022, and 2023 build on a growing collection of TreeMap datasets that support
 436 nationally consistent, tree-level forest data for Earth system applications (Riley et al., 2022b;
 437 Peeler et al., 2023; Dye et al., 2024; Chivhenge et al., 2025; Hurteau et al., 2025). Because
 438 error increases at finer spatial scales in the underlying LANDFIRE data and this error
 439 propagates to TreeMap, TreeMap is best applied at project, landscape or regional scales rather
 440 than for inference at the individual pixel level (Riley et al., 2021a, 2022a; Hudak et al. 2026). In
 441 addition, because methodology sometimes changes across LANDFIRE versions, LANDFIRE
 442 does not advise comparisons of the same pixel across versions (La Puma 2023). Within these
 443 recommended uses, continued improvements across TreeMap versions have expanded the
 444 range of suitable applications, including species range mapping. The species ride-along fix
 445 improved the species range mapping of TreeMap 2020, 2002, and 2023 with only a small trade-
 446 off in target layer accuracy that was comparable in magnitude to the variability observed
 447 between model runs in the stochasticity analysis, suggesting that the fix did not meaningfully
 448 degrade model performance. Despite its small accuracy cost, the species ride-along fix
 449 noticeably reduced implausible species assignments and increased the suitability of TreeMap
 450 for Earth system applications where species identity and geographic range accuracy are
 451 important. We therefore consider the overall impact of the species ride-along fix on model
 452 accuracy to be minimal relative to the ecological improvement it provides.

453 Although the species ride-along fix reduced the number of implausible species assignments,
 454 isolated cases may still occur where species are imputed outside of their geographic ranges due
 455 to zone adjacency. For example, blue oak (*Quercus douglasii*) is native to the foothill woodlands
 456 of California's central valley. Because the LANDFIRE zone containing the Willamette National
 457 Forest is adjacent to zones in California where blue oak occurs, the species remained available
 458 for imputation in that zone, resulting in approximately 0.0085% of pixels in the Willamette
 459 National Forest being assigned a blue oak forest type. For land managers working in areas
 460 where such assignments are ecologically significant, one practical solution is to reassign trees
 461 of an inappropriate species to another locally-occurring species within that genus, such as
 462 substituting a native oak species for blue oak within the Willamette National Forest.



A key strength of TreeMap is its ability to impute FIA plot IDs that link directly to the FIA DataMart (<https://apps.fs.usda.gov/fia/datamart/datamart.html>), giving users access to detailed tree-level information for each pixel. Other national-scale forest attribute products like BIGMAP (Wilson et al., 2012, 2018) summarize stand conditions through aggregate metrics such as number of trees, basal area, and total aboveground biomass, without providing individual tree records. By preserving the species, size, and status of individual trees, TreeMap supports applications that require tree-level forest structure to inform management decisions. Its linkage to the FIA DataMart also allows users to project TreeMap pixels into the future using tree-level stand projection tools such as the Forest Vegetation Simulator (Dixon, 2002).

4.2 Sources of uncertainty and error

Because random forests uses a set of decision trees, TreeMap does not produce uncertainty or error statistics in the traditional frequentist sense. Instead, we focused on three complementary validation approaches that together characterize accuracy. Specifically, OOB error rates provide an internal measure of model performance, target layer accuracy reflects model technical and structural uncertainty by comparing imputed values to the LANDFIRE input data, and independent validation at FIA plot locations additionally captures inaccuracies in the target data itself (Riley and Thompson, 2016b).

Uncertainty and error in TreeMap stem from measurement and classification errors in the data inputs (FIA, LANDFIRE and DayMet) that propagate into the final product. One key source of uncertainty is the forest mask used to define forested pixels. In one documented case, a forested area along the North Carolina coast was incorrectly classified primarily as shrub in LANDFIRE circa 2014 and consequently excluded from TreeMap 2014, but the omission was corrected in TreeMap 2016 when LANDFIRE circa 2016 reclassified the area as forest cover. Misalignment can also occur in recently disturbed areas with high tree mortality, which may fall below LANDFIRE's 10% live tree cover threshold and be excluded from TreeMap even though FIA continues to classify them as forested if they are expected to recover. Similar discrepancies can occur in sparsely forested areas, such as juniper forests in Texas and the American Southwest. For these reasons, users are advised to carefully review the forest mask for their study area to identify omissions or erroneous inclusions of forested pixels before applying TreeMap to local analyses. Beyond the forest mask, the FIA plot data itself may contain errors, including misassignment of field forest type. For instance, two plots identified in the field as Port



494 Orford-cedar (*Chamaecyparis lawsoniana*) forest type were found to lack that species entirely in
 495 the tree tally, suggesting possible misassignment of the forest type.

496 Another source of uncertainty in TreeMap stems from the inherent spatial mismatch between
 497 the footprint of FIA field plots and the pixels considered for validation assessments. One
 498 approach used in other work (Hudak et al. 2026) compares the center subplot value for a given
 499 attribute, such as aboveground biomass, to the value of the pixel in which it is located. However,
 500 this approach is limited because the center subplot area of 167 m² represents only
 501 approximately 19% of the 900 m² pixel area, meaning that the subplot characterizes less than a
 502 fifth of the pixel it is used to validate. Our approach instead uses the entire FIA plot footprint of
 503 670 m², which covers approximately 74% of a single pixel and provides a more representative
 504 sample of pixel conditions. However, this approach is also limited because of FIA's splayed plot
 505 design, which means that some subplots will fall outside the center pixel. To address this
 506 limitation, we report whether at least one pixel within the plot footprint matches the plot value for
 507 each response variable and whether the mean cover and height values of pixels within the plot
 508 radius fall within 10% and 5 m of the plot value, respectively. Ideally, if available, stem maps
 509 that include all trees comprehensively measured within 30 m Landsat pixels could be leveraged
 510 in future validation.

511 Finally, uncertainty can stem from LANDFIRE data being updated between vintages to reflect
 512 disturbance and post-disturbance dynamics. A pixel identified as having experienced a
 513 disturbance, such as wildfire, may be assigned a non-forested EVT-GP (La Puma 2023). In that
 514 case the pixel will not be included in the TreeMap vintage or it may be assigned a new EVT-GP,
 515 EVC, and EVH. LANDFIRE classifies pixels as recently disturbed for up to 10 years following
 516 disturbance, depending on the pre-disturbance lifeform and disturbance severity (La Puma
 517 2023). Following this period, a pixel is transitioned from disturbed to estimate current EVT-GP,
 518 EVC, and EVH. LANDFIRE only incorporates new satellite imagery in its datasets every 7-10
 519 years, with undisturbed pixels being identical in the interim. Due to the LANDFIRE rulesets, our
 520 use of the same random forests models, and no other TreeMap predictor variables varying
 521 annually, many pixels in TreeMap 2020, 2022, and 2023 are identical between vintages.

522 4.3 Future directions

523 One possible future direction for TreeMap is incorporating an explicit measure of biomass
 524 density as a response variable in the random forests imputation. In most forest types, canopy



cover and vegetation height provide sufficient information for the model to assign plots with reasonable size class distributions and biomass estimates. However, in forest types where biomass density varies dramatically independent of canopy cover and vegetation height, these variables alone may not adequately distinguish between structurally distinct stand types. For example, coastal redwood forests that contain old growth stands with very large trees can have substantially higher biomass density than second growth stands of similar canopy cover and the same vegetation height class in LANDFIRE. Consequently, we observed differences in biomass density imputed to coast redwood EVT-GPs across TreeMap versions, with TreeMap 2014 imputing stands with biomass densities upwards of 2000 tons/acre, while TreeMap 2016, 2020, 2022, and 2023 imputed no stands exceeding 1000 tons/acre. Moving forward, including biomass density as a response variable could help the model better distinguish these structurally distinct stand types and produce more consistent estimates.

Additionally, future TreeMap versions could build on the growing consistency of annual LANDFIRE data releases to expand TreeMap's capacity for tracking forest change. One promising direction is incorporating previously forested disturbed areas into TreeMap, which could enable applications such as monitoring post-disturbance regeneration, snag persistence, and forest type change. One important caveat for time series applications is that LANDFIRE production methods change periodically, including substantial changes between the 2014 and 2016 versions. Direct comparisons of some attributes across methods changes are suitable, but other attributes may be unreliable and limit the temporal scope of any trend analysis.

Future work could also refine the disturbance attribute to distinguish whether a pixel experienced no disturbance, fire, or insects and disease, rather than treating disturbance as binary. Earlier versions of TreeMap lacked a sufficient number of disturbed FIA plots to differentiate disturbance types. However, as FIA has continued to measure additional plots, sufficient data may now be available. Although these refinements are not expected to substantially affect TreeMap outputs, they may influence the development of imputed datasets focused on understory composition and structure. Differentiating disturbance type is important because it directly affects key ecological variables that TreeMap users are interested in calculating. For example, fuel loadings are much lower on burned plots than on plots affected by insects and disease, which could result in errors if imputed datasets are used to calculate smoke emissions.



556 Finally, because traditional uncertainty or error statistics cannot be provided by the random
557 forests methodology, a fuller analysis of the sources of uncertainty would aid TreeMap users.
558 Quantitative evaluation of sources of uncertainty is unlikely to be possible. However,
559 enumeration of sources and types of error and some estimation of their relative magnitude can
560 likely be accomplished. Potential sources include variability in FIA plot data, limitations in
561 predictor variables, and model-driven uncertainty. Clearly describing these sources would
562 improve transparency and support more informed use of TreeMap in Earth system applications.

563 **5. Data availability**

564

565 TreeMap 2020 data described in this manuscript are available at the Forest Service Research
566 Data Archive under <https://doi.org/10.2737/RDS-2025-0031> (Zimmer et al., 2025). TreeMap
567 2022 data are available at the Forest Service Research Data Archive under
568 <https://doi.org/10.2737/RDS-2025-0032> (Houtman et al., 2025). TreeMap 2023 data are
569 available at the Forest Service Research Data Archive under [https://doi.org/10.2737/RDS-2026-](https://doi.org/10.2737/RDS-2026-0038)
570 [0038](https://doi.org/10.2737/RDS-2026-0038) (Zimmer et al., 2026).

571

572 **6. Code availability**

573

574 The code used to produce TreeMap 2020, TreeMap 2022, and TreeMap 2023 is publicly
575 available at <https://github.com/firelab/gtac-treemap>.

576 **7. Conclusions**

577 TreeMap 2020, 2022, and 2023 provide 30 m resolution gridded maps of FIA plot identifiers for
578 the forested extent of CONUS, updating and improving on previous versions of TreeMap. The
579 most significant methodological improvement since prior versions is the species ride-along fix,
580 which constrains the available FIA plots for imputation to a geographic pool for each LANDFIRE
581 zone to address the issue of implausible species assignments observed in earlier TreeMap
582 versions. The Willamette National Forest case study demonstrated that this fix reduced the
583 proportion of pixels containing implausible species assignments from 0.76% in 2016 to 0.19% in
584 2020. Validation using OOB error rates, target layer accuracy, and independent validation with a
585 set of plots not used in the imputation confirmed high accuracy across EVC, EVH, EVT-GP, and
586 disturbance presence in TreeMap 2020, 2022, and 2023. Together, TreeMap 2020, 2022, and
587 2023 provide an updated, validated, and ecologically improved tree-level forest dataset suitable



for Earth system applications ranging from carbon cycling to wildfire behavior modeling across CONUS.

8. Author contributions

Karin Riley: conceptualization, methodology, software, writing - original draft, writing - review & editing, resources, supervision, project administration, funding acquisition. **Rachel Houtman:** software, validation, formal analysis, data curation, writing - original draft, writing - review & editing, visualization. **Scott Zimmer:** software, validation, formal analysis, data curation, writing - original draft, writing - review & editing, visualization. **Lila Leatherman:** software, formal analysis, data curation, writing - review & editing. **Jamie Peeler:** writing - original draft, writing - review & editing, visualization. **John Shaw:** data curation, methodology, writing - review & editing. **Isaac Grenfell:** conceptualization, methodology, software. **Maria Borja Arboleda:** data curation, writing - review & editing. **Abhinav Shrestha:** software, formal analysis, writing - review & editing. **Ian Housman:** software, writing - review & editing. **Mark Finney:** conceptualization, writing - review & editing, funding acquisition.

9. Competing interests

The authors declare the following financial interests or personal relationships which may be considered as potential competing interests: Karin Riley, Mark Finney, Scott Zimmer, John Shaw, and Rachel Houtman report financial support was provided by USDA Forest Service Rocky Mountain Research Station, Fire and Aviation Management, and IT funds. Jamie Peeler reports funding was provided by USDA Forest Service. Karin Riley, Mark Finney, Lila Leatherman, Maria Borja Arboleda, Abhinav Shrestha, and Ian Housman report financial support was provided by USDA Forest Service's Fire and Aviation Management as well as Forest Service IT funds. Other authors declare they have no known competing financial interests or personal relationships that could influence the work reported in this paper.

10. Disclaimer

The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or US Government determination or policy. This work was supported by the US Department of Agriculture, USDA Forest Service.



620 **11. Acknowledgements**

621

622 Tom Thompson at the US Forest Service's Forest Inventory and Analysis (FIA) provided true
 623 plot coordinates for this work under a Memorandum of Cooperation between FIA and the Fire
 624 Lab. Abbey Schaaf of Red Castle Resources provided project management. US Forest
 625 Service's Rocky Mountain Research Station, Fire and Aviation Management, and Forest
 626 Service IT provided funding. Lexie Peeler provided graphical design support for the conceptual
 627 figure.

628 **12. References**

- 629 Anderegg, W. R. L., Chegwidden, O. S., Badgley, G., Trugman, A. T., Cullenward, D.,
 630 Abatzoglou, J. T., Hicke, J. A., Freeman, J., and Hamman, J. J.: Future climate risks from
 631 stress, insects and fire across US forests, *Ecol. Lett.*, 25, 1510–1520,
 632 <https://doi.org/10.1111/ele.14018>, 2022.
- 633 Bechtold, W. A. and Patterson, P. L.: The Enhanced Forest Inventory and Analysis Program -
 634 National Sampling Design and Estimation Procedures, U.S. Department of Agriculture, Forest
 635 Service, Southern Research Station, Asheville, NC, <https://doi.org/10.2737/SRS-GTR-80>, 2005.
- 636 Blackard, J., Finco, M., Helmer, E., Holden, G., Hoppus, M., Jacobs, D., Lister, A., Moisen, G.,
 637 Nelson, M., and Riemann, R.: Mapping U.S. forest biomass using nationwide forest inventory
 638 data and moderate resolution information, *Remote Sens. Environ.*, 112, 1658–1677,
 639 <https://doi.org/10.1016/j.rse.2007.08.021>, 2008.
- 640 Burrill, E. A., DiTommaso, A. M., Turner, J. A., Pugh, S. A., Christensen, G., Kralicek, K. M.,
 641 Perry, C. J., Lepine, L. C., Walker, D. M., and Conkling, B. L.: The forest inventory and analysis
 642 database, FIADB user guides, volume: database description (version 9.3), nationwide forest
 643 inventory (NFI), U.S. Department of Agriculture, Forest Service, 1026 pp.,
 644 <https://research.fs.usda.gov/understory/forest-inventory-and-analysis-database-user-guide-nfi>,
 645 2024.
- 646 Chivhenge, E., Weiskittel, A. R., Woodall, C. W., D'Amato, A. W., and Daigneault, A.:
 647 Geospatial Estimation of Forest Relative Density for Carbon Stewardship Decision Support
 648 across the Continental US, *Sci. Data*, 12, 1728, <https://doi.org/10.1038/s41597-025-06012-6>,
 649 2025.



- 650 Crookston, N. L. and Finley, A. O.: yalImpute: An R Package for kNN Imputation, *J. Stat. Softw.*,
 651 23, 1–16, <https://doi.org/10.18637/jss.v023.i10>, 2008.
- 652 Dixon, G. E.: Essential FVS: A user's guide to the forest vegetation simulator, Internal Rep.,
 653 U.S. Department of Agriculture, Forest Service, Forest Management Service Center, Fort
 654 Collins, CO, 236 pp., 2002 (revised 2025).
- 655 Dye, A. W., Houtman, R. M., Gao, P., Anderegg, W. R. L., Fettig, C. J., Hicke, J. A., Kim, J. B.,
 656 Still, C. J., Young, K., and Riley, K. L.: Carbon, climate, and natural disturbance: a review of
 657 mechanisms, challenges, and tools for understanding forest carbon stability in an uncertain
 658 future, *Carbon Balance Manag.*, 19, 35, <https://doi.org/10.1186/s13021-024-00282-0>, 2024.
- 659 Houtman, R. M., Leatherman, L. S. T., Zimmer, S. N., Housman, I. W., Shrestha, A., Shaw, J.
 660 D., and Riley, K. L.: TreeMap 2022 CONUS: A tree-level model of the forests of the
 661 conterminous United States circa 2022, Forest Service Research Data Archive, Fort Collins,
 662 CO, <https://doi.org/10.2737/RDS-2025-0032>, 2025.
- 663 Finney, M. A., McHugh, C. W., Grenfell, I. C., Riley, K. L., and Short, K. C.: A simulation of
 664 probabilistic wildfire risk components for the continental United States, *Stoch. Environ. Res.*
 665 *Risk Assess.*, 25, 973–1000, <https://doi.org/10.1007/s00477-011-0462-z>, 2011.
- 666 Hurteau, M. D., Jung, C. G., Francis, E. J., Dobrowski, S. Z., Littlefield, C. E., and Parks, S. A.:
 667 Decreasing landscape carbon storage in western US forests with 2 °C of warming, *Environ.*
 668 *Res. Ecol.*, 4, 041001, <https://doi.org/10.1088/2752-664X/ae212c>, 2025.
- 669 Hudak, A. T., Bakken, J. L., Lister, A., Mauro, F., Gregory, M. J., Riley, K., Wilson, B., Santoro,
 670 M., Healey, S., Ma, L., Yu, Y., Bright, B. C., Byrne, J., Mita, R., Houtman, R. M., Shaw, J. D.,
 671 Fekety, P. A., Domke, G., Walters, B., Bell, D. M., Weiskittel, A., Atkins, J. W., Duncanson, L.,
 672 Hunka, N., Kennedy, R., Babcock, C., Saatchi, S., Raczka, B., Silva, C., Tang, H., and Hurtt, G.:
 673 A quantitative evaluation of forest aboveground biomass density map products in Oregon, USA,
 674 *Environ. Res. Lett.*, 21, 143007, <https://doi.org/10.1088/1748-9326/acf05a>, 2026.
 675
- 676 La Puma, I. P. (Ed.): LANDFIRE technical documentation, U.S. Geological Survey Open-File
 677 Report 2023–1045, 103 pp., <https://doi.org/10.3133/ofr20231045>, 2023.



- 678 McRoberts, R. E., Tomppo, E. O., and Næsset, E.: Advances and emerging issues in national
 679 forest inventories, *Scand. J. For. Res.*, 25, 368–381,
 680 <https://doi.org/10.1080/02827581.2010.496739>, 2010.
- 681 Peeler, J. L., McCauley, L., Metlen, K. L., Woolley, T., Davis, K. T., Robles, M. D., Haugo, R. D.,
 682 Riley, K. L., Higuera, P. E., Fargione, J. E., Addington, R. N., Bassett, S., Blankenship, K.,
 683 Case, M. J., Chapman, T. B., Smith, E., Swaty, R., and Welch, N.: Identifying opportunity hot
 684 spots for reducing the risk of wildfire-caused carbon loss in western US conifer forests, *Environ.*
 685 *Res. Lett.*, 18, 094040, <https://doi.org/10.1088/1748-9326/acf05a>, 2023.
- 686 Riley, K. and Thompson, M.: An Uncertainty Analysis of Wildfire Modeling, in: *Natural Hazard*
 687 *Uncertainty Assessment*, American Geophysical Union (AGU), 191–213,
 688 <https://doi.org/10.1002/9781119028116.ch13>, 2016.
- 689 Riley, K. L. and Loehman, R. A.: Mid-21st-century climate changes increase predicted fire
 690 occurrence and fire season length, Northern Rocky Mountains, United States, *Ecosphere*, 7,
 691 e01543, <https://doi.org/10.1002/ecs2.1543>, 2016.
- 692 Riley, K. L., Grenfell, I. C., Finney, M. A., and Wiener, J. M.: TreeMap, a tree-level model of
 693 conterminous US forests circa 2014 produced by imputation of FIA plot data, *Sci. Data*, 8, 11,
 694 <https://doi.org/10.1038/s41597-020-00782-x>, 2021a.
- 695 Riley, K. L., O'Connor, C. D., Dunn, C. J., Haas, J. R., Stratton, R. D., and Gannon, B.: A
 696 National Map of Snag Hazard to Reduce Risk to Wildland Fire Responders, *Forests*, 13, 1160,
 697 <https://doi.org/10.3390/f13081160>, 2022a.
- 698 Riley, K. L., Grenfell, I. C., Finney, M. A., Wiener, J. M., and Houtman, R. M.: Fire lab tree list: A
 699 tree-level model of the conterminous United States landscape circa 2014, Forest Service
 700 Research Data Archive, Fort Collins, CO, <https://doi.org/10.2737/RDS-2019-0026>, 2019.
- 701 Riley, K. L., Grenfell, I. C., Finney, M. A., and Shaw, J. D.: TreeMap 2016: A tree-level model of
 702 the forests of the conterminous United States circa 2016, Forest Service Research Data
 703 Archive, Fort Collins, CO, <https://doi.org/10.2737/RDS-2021-0074>, 2021b.
- 704 Riley, K. L., Grenfell, I. C., Shaw, J. D., and Finney, M. A.: TreeMap 2016 Dataset Generates
 705 CONUS-Wide Maps of Forest Characteristics Including Live Basal Area, Aboveground Carbon,



- 706 and Number of Trees per Acre, *J. For.*, 120, 607–632, <https://doi.org/10.1093/jofore/fvac022>,
 707 2022b.
- 708 Shaw, J.: Benefits of a strategic national forest inventory to science and society: the USDA
 709 Forest Service Forest Inventory and Analysis program, *IForest - Biogeosciences For.*, 1, 81–85,
 710 <https://doi.org/10.3832/ifor0345-0010081>, 2008.
- 711 Thornton, M. M., Shrestha, R., Wei, Y., Thornton, P. E., and Kao, S.-C.: Daymet: Annual climate
 712 summaries on a 1-km grid for North America, version 4 R1, ORNL DAAC, Oak Ridge,
 713 Tennessee, USA, <https://doi.org/10.3334/ORNLDAAC/2130>, 2022.
 714
- 715 USDA Forest Service Strategic Analytics Branch: Risk Management Assistance (RMA)
 716 Dashboard,
 717 [https://experience.arcgis.com/experience/f5ca19164c1d4ec9bd4e221611508d28/page/RMA-](https://experience.arcgis.com/experience/f5ca19164c1d4ec9bd4e221611508d28/page/RMA-Dashboard)
 718 [Dashboard](https://experience.arcgis.com/experience/f5ca19164c1d4ec9bd4e221611508d28/page/RMA-Dashboard) (last access: 24 July 2026).
 719
- 720 Wiener, S. S., Bush, R., Nathanson, A., Pelz, K., Palmer, M., Alexander, M. L., Anderson, D.,
 721 Treasure, E., Baggs, J., and Sheffield, R.: United States Forest Service Use of Forest Inventory
 722 Data: Examples and Needs for Small Area Estimation, *Front. For. Glob. Change*, 4,
 723 <https://doi.org/10.3389/ffgc.2021.763487>, 2021.
- 724 Zimmer, S. N., Houtman, R. M., Leatherman, L. S. T., Housman, I. W., Shrestha, A., Shaw, J.
 725 D., and Riley, K. L.: TreeMap 2020 CONUS: A tree-level model of the forests of the
 726 conterminous United States circa 2020, Forest Service Research Data Archive, Fort Collins,
 727 CO, <https://doi.org/10.2737/RDS-2025-0031>, 2025