Forest Service
U.S. DEPARTMENT OF AGRICULTURE

Pacific Northwest Research Station | General Technical Report PNW-GTR-1018 | November 2023

Guidance for Forest Management and Landscape Ecology Applications of Recent Gradient Nearest Neighbor Imputation Maps in California, Oregon, and Washington

David M. Bell, Matthew J. Gregory, Marin Palmer, and Raymond Davis

Authors

David M. Bell is a research forester, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331; **Matthew J. Gregory** is a senior faculty research assistant, Oregon State University, Department of Forest Ecosystems and Society, 321 Richardson Hall, Corvallis, OR 97331; **Marin Palmer** is a biometrician, U.S. Department of Agriculture, Forest Service, Pacific Northwest Region, 1220 SW 3rd Avenue, Portland, OR 97204; **Raymond Davis** is the old forests and northern spotted owls monitoring lead for the Northwest Forest Plan Interagency Regional Monitoring Program, U.S. Department of Agriculture, Forest Service, Pacific Northwest Region, 3200 SW Jefferson Way, Corvallis, OR 97331.

Cover photo: Mount Shuksan, Washington, as seen from the High Divide Trail. Photo by Edmund Lowe.

Guidance for Forest Management and Landscape Ecology Applications of Recent Gradient Nearest Neighbor Imputation Maps in California, Oregon, and Washington

David M. Bell, Matthew J. Gregory, Marin Palmer, and Raymond Davis

U.S. Department of Agriculture, Forest Service
Pacific Northwest Research Station
Portland, Oregon
General Technical Report PNW-GTR-1018
November 2023

This publication was developed by the Landscape Ecology Modeling, Mapping, and Analysis team, a collaboration between the U.S. Department of Agriculture (USDA) Forest Service Pacific Northwest Research Station and Oregon State University, with the assistance of the USDA Forest Service Pacific Northwest Region's Biometrics Program and the Northwest Forest Plan Federal Interagency Monitoring Program.

Abstract

Bell, David M.; Gregory, Matthew J.; Palmer, Marin; Davis, Raymond. 2023. Guidance for forest management and landscape ecology applications of recent gradient nearest neighbor imputation maps in California, Oregon, and Washington. Gen. Tech. Rep. PNW-GTR-1018. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 41 p. https://doi.org/10.2737/pnw-gtr-1018.

For many forest landscape ecology and ecological monitoring projects, forest structure and composition data availability at the correct scale are often limiting factors, motivating the development of map products based on remote sensing. Gradient nearest neighbor (GNN) imputation is a flexible framework for generating multivariate, annual, wall-to-wall maps of forest structure and composition for landscape, regional, and national applications. This report provides guidance on the appropriate use of forest structure and composition maps generated from satellite imagery, physical environment, and forest inventory data using the GNN modeling and mapping framework. We describe the GNN modeling and mapping framework associated with the generation and delivery of updated maps in 2020 (GNN-2020) that provide forest attribute status and trend data from 1986 to 2017. In relation to the GNN-2020 map data, we describe (1) the accuracy assessment reporting that accompanies all maps, (2) basic concepts regarding the strength of relationships between forest attributes and the geospatial predictor variables used in mapping, (3) the role of the number of nearest neighbors in map accuracy, (4) factors affecting the appropriate spatial and temporal scales for using the maps, (5) the types of changes in forest attributes that can be reasonably assessed with the maps, and (6) the challenges in creating categorical or classified maps based on the GNN-2020 data. Our goal is to provide enough background and guidance to use GNN products appropriately and effectively.

Keywords: Forest structure and composition, gradient nearest neighbor imputation, GNN, monitoring, vegetation mapping.

Contents

Introduction

In many forest landscape ecology and ecological monitoring projects, the availability of forest structure and composition data is often a limiting factor, motivating the development of map products based on remote sensing. There is a need to quantify both status and trends in forest vegetation attributes at a variety of scales—from forest stands (tens to hundreds of hectares) to landscapes (thousands to hundreds of thousands of hectares) to regions (more than a million hectares)—across all forest lands to support many applications, including forest planning, monitoring, and research. Along the Pacific coast of the United States, the need became more pressing in the 1990s with the listing of the northern spotted owl as threatened under the Endangered Species Act and subsequent implementation of the Northwest Forest Plan. In response, researchers developed methods to leverage satellite remote-sensing and forest inventory data to generate a diverse suite of forest vegetation structure and composition maps useful for monitoring late-successional and old-growth forests and habitats of several species of conservation interest. This capacity to map forest inventory data at relatively fine scales now supports forest monitoring and planning across federal agencies, state governments, municipalities, and research contexts. With the need for the forest attribute maps comes a responsibility for appropriate guidance in the use of those data.

This report provides guidance on the appropriate use of forest structure and composition maps generated from satellite imagery, physical environment, and forest inventory data using the gradient nearest neighbor (GNN) imputation framework. GNN is a flexible framework for generating multivariate, annual, wall-to-wall maps of forest structure and species composition for landscape, regional, and national applications. Although GNN is a single modeling and mapping framework, it generates a complex assortment of forest structure and composition data that vary in accuracy and precision at multiple spatial and temporal scales. This report is motivated by, and most directly addresses, the most recent iteration of GNN maps (hereafter referred to as GNN-2020), but many of the concepts described apply more broadly to other vegetation maps that were generated using similar data or methods. GNN-2020 maps are based on long-standing forest inventory (U.S. Department of Agriculture [USDA] Forest Service Forest Inventory and Analysis [FIA] program) and satellite remote-sensing (Landsat) programs. Determining the most appropriate usage of GNN-2020 data remains challenging, even for experienced users. Here, we describe the framework used to generate the GNN-2020 maps and provide guidance regarding the appropriate application of those data for forest management and landscape ecology. Because the eventual decision on what constitutes sufficient accuracy in vegetation maps for a given application is the responsibility of the data user, we focus on informing users on how to interpret previous research and current accuracy assessments to help them identify the best forest structure and composition variables to use and the spatial and temporal scales at which questions can be addressed using GNN-2020 maps.

Background

Mapped vegetation data form the basis of many aspects of forest management, monitoring, conservation, species recovery efforts, and landscape ecology. Resource managers can use vegetation maps to answer such questions as the following:

- What are the distributions of forest types on the landscape?
- How many trees of a certain size are present?
- Where are the late-successional forests or the aspen groves?
- How are these things changing?

These questions in turn help answer questions about ecological conditions for a variety of fish and wildlife species as well as botanical communities. Although plot-based inventory data can also answer many of these same questions, the answers are often limited to coarse spatial scales and temporal resolutions. Maps are required to inform questions where detailed spatial pattern matters, such as for mapping habitat connectivity. They must have the appropriate detail to examine spatial patterns and relationships to key drivers, such as topography, disturbance history, and climate. Maps are also a compelling way to communicate; for example, managers and stakeholders can use maps to visualize where forest density is greatest and how it has changed over time. As the USDA Forest Service's 2012 Planning Rule emphasizes adaptive management, and thus monitoring, as a core component of planning (Nie 2018, Ryan et al. 2018), vegetation mapping will remain a central basis for decision making on National Forest System lands and beyond. Furthermore, broader scale monitoring can often be the first analytical step for assessing changes, trends, and future desired conditions at a spatial scale that has larger implications.

Apart from the spatial patterns that vegetation maps can reveal, the long temporal record of multispectral satellite imagery facilitates change detection and trend analysis. Annual Landsat time series imagery composites (i.e., best pixel within a given year) are now widely used (e.g., Zhu 2017), and the long-term perspective (Landsat 5, 7, and 8 cover 1984 to the present) they provide stretches back before the current annual national forest inventory design used by the FIA program was widely applied (circa 2001) (Bechtold and Patterson 2005). Over broad areas, annual maps allow for trend analysis to identify long-term changes in resources. For example, Landsat-based change detection quantifies changes in disturbance regimes in recent decades (Schleeweis et al. 2020), sometimes identifying major changes in the key drivers of forest change (Cohen et al. 2016, Schroeder et al. 2017).

As with any existing vegetation map or inventory, users must first identify their information needs in relation to their questions before they can determine the appropriate data source and attributes/indicators and whether it meets their required level of certainty. This is true of GNN mapping for California, Oregon, and Washington and related maps that were released in 2020 (Bell et al. 2021) (GNN-2020), which form the basis for the 25-year late-successional and old-growth forest monitoring report for the Northwest Forest Plan (NWFP) (Davis et al. 2022).

GNN Imputation Mapping

As the methodological basis for a regional mapping and monitoring framework, statistical modeling techniques that leverage satellite remote sensing, such as Landsat imagery for vegetation mapping, reflect developments in remote sensing and scientific or management needs (Banskota et al. 2014). Nearest neighbor imputation is well-suited for mapping complex vegetation characteristics, as there are no distributional assumptions about the data being modeled, and by imputing plots to pixels, the method can preserve relationships between many forest structure and composition characteristics (McRoberts 2001, Moeur and Stage 1995). In particular, the maintenance of covariation among multiple response variables needed to map late-successional and old-growth forest became a major justification for the use of nearest neighbor imputation as a component of NWFP effectiveness monitoring for late-successional and old-growth forests (Davis et al. 2015, 2022; Ohmann and Gregory 2002; Ohmann et al. 2012). The synergistic drivers (technological development and management needs) were particularly strong in the Pacific Northwest region of the United States. In this section, we provide a general overview of the characteristics of the GNN workflow, which leverages methodological developments dating back over several decades. This basic understanding of the mapping approach provides insights into the appropriate use of maps generated by GNN for planning, monitoring, and estimation.

Overview of GNN Methodology

At the center of regional vegetation monitoring and mapping in California, Oregon, and Washington lies an integrated remote-sensing and forest inventory framework that was developed over several decades. This framework links forest inventory data with remote sensing through an empirical modeling approach to produce 30-m (pixel resolution) maps of forest vegetation characteristics across all forests, thus supporting landscape assessment and analysis (fig. 1). In this section, we give an overview of the components of this system used to generate GNN maps: reference data, geospatial predictor data, GNN modeling and imputation, and nonforest masking. Our work builds on a 20-year legacy of GNN development in the Pacific Northwest (Bell et al. 2015a, 2021; Davis et al. 2015, 2022; Kennedy et al. 2018a; Ohmann and Gregory 2002; Ohmann et al. 2012) as well as nearest neighbor imputation research in forest sciences across the globe (Chirici et al. 2016).

The reference data: USDA Forest Service Forest Inventory and Analysis data— The availability of field-based vegetation observations is foundational to reliable vegetation mapping (Lister et al. 2020). Though there are many potential sources of forest data, we used forest inventory plot data (hereafter, plot data) for California, Oregon, and Washington forests—from the USDA Forest Service's FIA and Pacific Northwest Region monitoring plots that had at least 50 percent forested conditions as the source of forest attribute information. The 50-percent-forest-condition

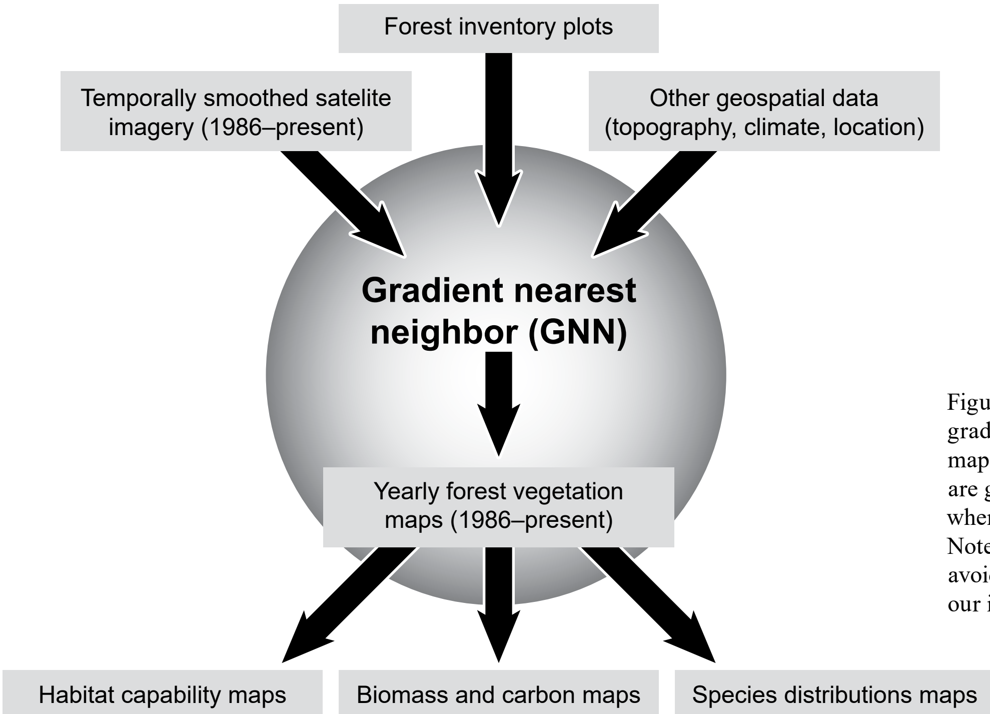

Figure 1—Conceptual framework for gradient nearest neighbor vegetation map production and applications. Maps are generated for 1986 onward as that is when Landsat imagery became available. Note that we exclude 1984–1985 to avoid some imagery artifacts related to our image processing.

threshold differs from the FIA definition of forest and was selected to ensure that most of the remote-sensing signal for a given plot is associated with forest and not, for example, grassland, shrubland, or developed land.

FIA plots provide a broadly distributed, nationally consistent statistical sample of forests supporting broad-scale research and monitoring (fig. 2), such as species distributional responses to climate change (e.g., Bell et al. 2014, Canham and Thomas 2010, Martin and Canham 2020, Monleon and Lintz 2015), carbon storage (e.g., Smith et al. 2019), and vegetation cover types as part of wildlife habitat monitoring (Mellen-McLean et al. 2017). Although previous iterations of GNN mapping used data generated through various sampling designs (e.g., current annual FIA design, past periodic FIA design, regional forest monitoring plots), GNN-2020 uses only measurements derived from the annual FIA plot design, which was adopted in California, Oregon, and Washington in 2001 (Bechtold and Patterson 2005). FIA plots use a clustered, nested, four-subplot design with the largest components (macroplots) covering a total of 0.4 ha and a 9-Landsat-pixel (3- by 3-pixel) footprint (0.81 ha) centered on the central subplot roughly, but not completely, overlapping the entire measured area (fig. 3). At FIA plots, live and dead tree characteristics are measured, including individual tree species, diameter, and mortality status, among other attributes (see Burrill et al. 2021 for details on all available data). Those attributes can be summarized for each plot, providing a multifaceted dataset that characterizes forest structure and composition at the plot level. Of course, measurements from plots or portions of plots may also be aggregate across larger areas (e.g., counties) or conditions (e.g., forest types) to generate unbiased estimates of forest attributes or areas (Bechtold and Patterson 2005).

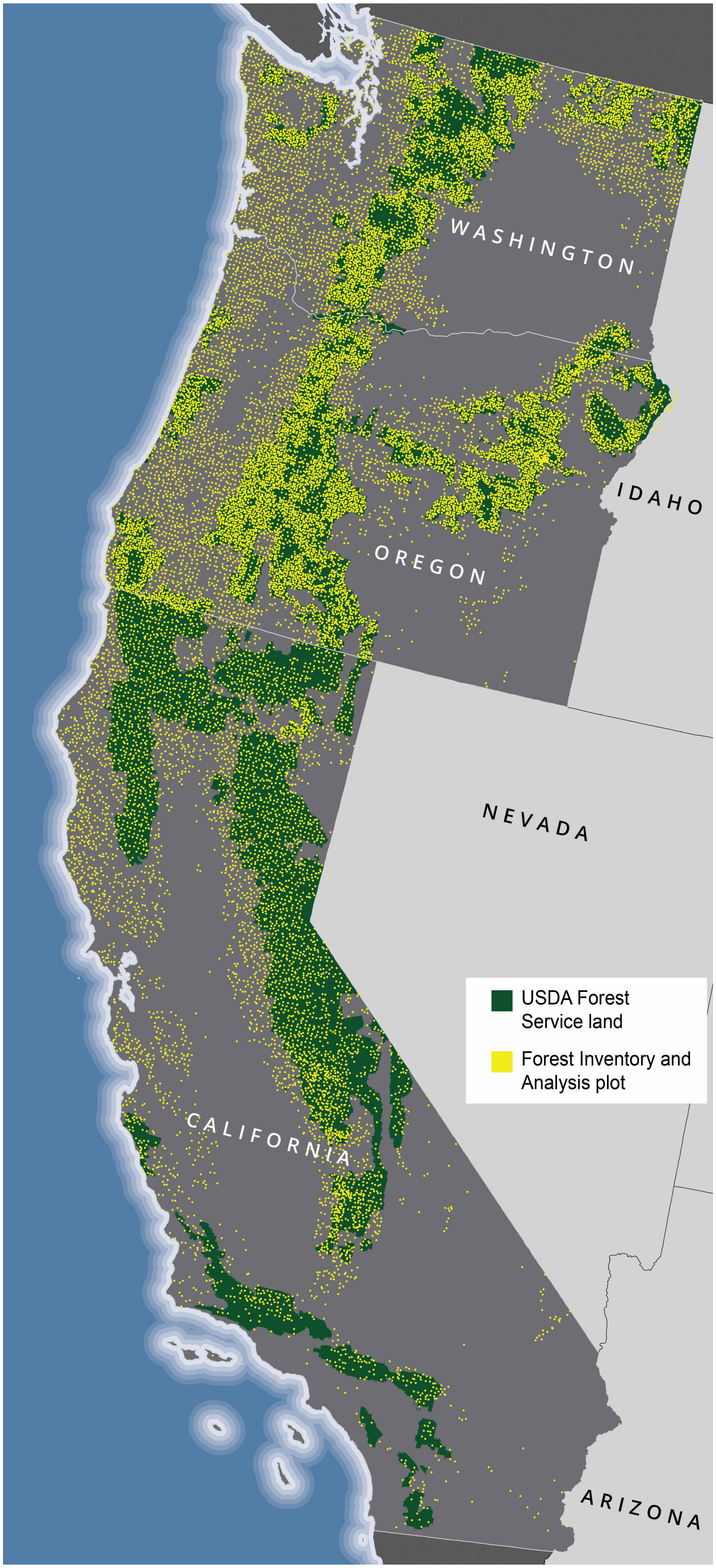

Figure 2— USDA Forest Service Forest Inventory and Analysis plot distribution across California, Oregon, and Washington. USDA Forest Service lands have higher plot density in Oregon and Washington outside of wilderness areas.

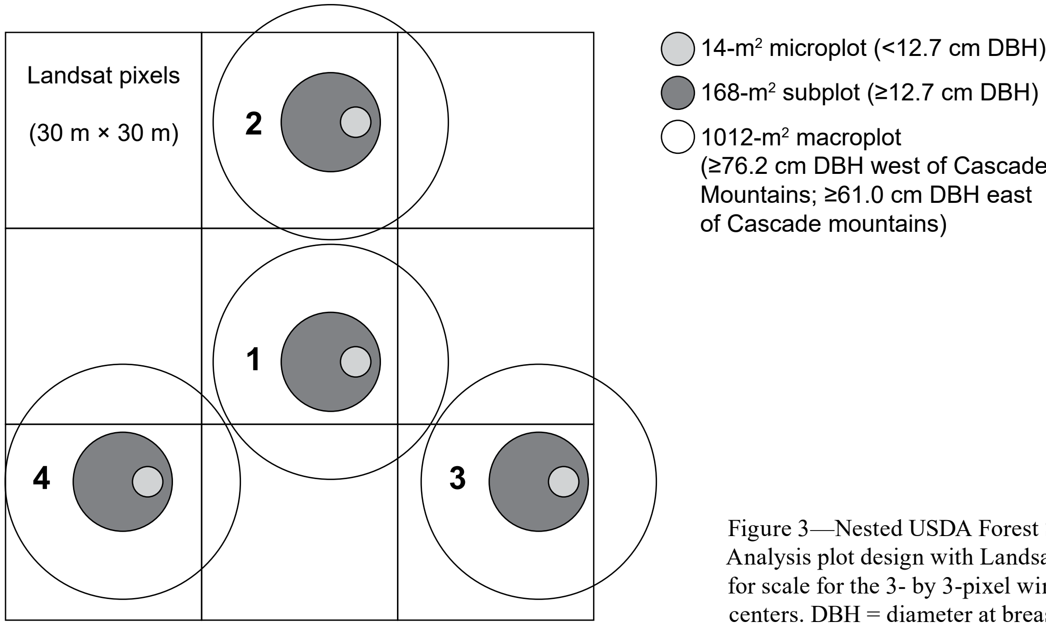

Figure 3—Nested USDA Forest Service Forest Inventory and Analysis plot design with Landsat pixel sizes (30 m by 30 m) for scale for the 3- by 3-pixel window that contains the subplot centers. DBH = diameter at breast height.

The FIA nested plot design (four subplots) (fig. 3) captures local heterogeneity in forest conditions by distributing measurements across an area that, at times, can contain various forest conditions (e.g., plot straddles the boundary between an older and younger forest). For the purposes of modeling and mapping, using a single plot design (e.g., the clustered subplots of the annual FIA design) minimizes modeling artifacts associated with incorporating other plot designs into an analysis, as would be the case if including previous periodic FIA plot sampling designs. Furthermore, using only annual FIA plot design (fig. 3) provides consistent measurement on hundreds of forest structural and compositional attributes.

Despite the clear value of FIA plot data for the study of forest conditions across broad geographies, there are several limitations as they relate to landscape and regional monitoring and planning in California, Oregon, and Washington. Spatial plot densities range from one plot for every 700–2400 ha, supporting design-based estimates of forest attributes at coarser scales (e.g., >25 000 ha) (Bell et al. 2022b). This spatial intensity of plot data is suitable for broad-scale estimation, but not necessarily for landscape assessments, and certainly not for spatial patterns for which mapping may be the only reasonable solution, such as identifying patches of a specific forest condition capable of supporting a northern spotted owl (Davis et al. 2016). Furthermore, rare forest conditions may be poorly represented in forest inventory data, though they are often locally abundant. Time intervals between plot measurements in California, Oregon, and Washington are 10 years, which may result in delayed assessment of forest change associated with episodic disturbances. Many local planning and monitoring questions may be better answered with a complementary approach that incorporates inventory data, physical environment information, remotely sensed data, and modeling.

Geospatial predictor data: Landsat satellite imagery and biophysical parameters—Because annual maps are necessary to characterize forest attribute patterns across space and time, forest attribute mapping also relies on spatially explicit data, or maps, of variables relevant to the distribution of forest attributes across pixels, stands, landscapes, and regions. Multiscale controls on vegetation distribution and local community structure depend on interacting factors, including climate, geology, disturbance, and topography (Ohmann and Spies 1998). Therefore, mapping variables used to predict forest structure and function should account for these multiscale controls on vegetation.

Climate and topography represent broad- and fine-scale drivers of moisture and temperature environments, which control tree species distributions, forest productivity, and forest structure (Latta et al. 2010, Pan et al. 2013, Rehfeldt et al. 2006). For GNN-2020, we used 800-m Parameter-Elevation Regressions on Independent Slopes Model (PRISM) normal rasters for precipitation and temperature[1] and 10-m U.S. Department of the Interior Geological Survey (USGS) digital elevation models to derive a diverse suite of climatic and topographic indices that have been found useful in characterizing forest vegetation patterns (e.g., Davis et al. 2022, Ohmann and Gregory 2002). All derived rasters are projected and resampled to 30 m into the USGS North American Datum 1983 Contiguous USA Albers projection using bilinear interpolation. In addition to these climate and topographic variables, latitude and longitude, or maritime influence on vegetation patterns (Daly et al. 2008), may be included as spatial predictors to help account for additional geographic patterns. The incorporation of latitude and longitude tends to constrain imputation and improve accuracy (Ohmann and Gregory 2002).

In contrast to climatic and topographic variables, the Landsat satellite image archive captures elements of forest condition at the time that imagery is acquired and can therefore reflect forest succession, disturbance, land use, and other past and present constraints on forest structure and composition (Cohen and Spies 1992; Hansen et al. 2001, 2013; Nguyen et al. 2020; Wulder et al. 2022; Zhu 2017; Zhu et al. 2020). Since 1984, spectrally harmonized data from three Landsat satellites with different sensors (Landsat 5 Thematic Mapper, Landsat 7 Enhanced Thematic Mapper Plus, and Landsat 8 Operational Land Imager) provide wall-to-wall observations of all forest lands globally at 30-m grain sizes (e.g., Roy et al. 2016, Vogeler et al. 2018). The multispectral data, which include visible light (red, green, and blue), near infrared, shortwave infrared, and thermal wavelengths, form the basis of a variety of vegetation mapping applications, including vegetation classification (e.g., Wickham et al. 2021), biomass mapping (e.g., Kennedy et al. 2018a, Matasci et al. 2018, Wilson et al. 2013), ecosystem productivity (e.g., Robinson et al. 2018), and disturbance mapping (e.g., Kennedy et al. 2010, Masek et al. 2008, Schroeder et al. 2017, Zhu et al. 2020). Ideally, satellite imagery used in forest monitoring is temporally consistent and represents only the trends in vegetation status through time and across

[1] Data were downloaded from the Oregon State University PRISM Climate Group's PRISM Climate Data website, http://prism.oregonstate.edu, on September 28, 2016.

space, but we know that this is not the case. To improve consistency and reduce noise in annual Landsat time series, we leveraged fitted imagery based on change detection modeling with the ensemble LandTrendr (box 1). For GNN modeling, we excluded the first 2 years of imagery (1984–1985) to ensure high-quality performance of LandTrendr disturbance mapping, and thus fitted imagery (Davis et al. 2022).

Early examination of multispectral data and subsequent GNN-2020 vegetation maps highlighted low magnitude, multidecadal shifts in several multispectral indices. Such changes were often associated with late-successional or old-growth forests that appeared stable based on aerial photography, but GNN-2020 indicated potential declines in forest attributes associated with old forests, such as live tree basal area. Similar patterns have been previously observed in GNN mapping, motivating the stabilization (i.e., use of mean spectral indices rather than the trends from LandTrendr) when forests were deemed to be stable (i.e., not disturbed

Box 1

Using Ensemble LandTrendr for Change Detection and Temporal Smoothing of Satellite Imagery

Using the LandTrendr algorithm—a method of temporally segmenting annual Landsat time series to identify trends and change points, or vertices (Kennedy et al. 2010)—on Google Earth Engine™ (Kennedy et al. 2018b), we leveraged an ensemble disturbance mapping approach—a random forest, machine learning modeling approach (Healey et al. 2018)—to combine multiple versions of the LandTrendr algorithm (Cohen et al. 2018). This LandTrendr multispectral ensemble provides improved disturbance mapping accuracy from which Landsat fitted imagery can be derived (fig. 4). Thus, ensemble LandTrendr generates yearly 30-m imagery where each band's pixel values represent interpolated values between segment endpoints (i.e., vertices). We refer to this as "fitted" imagery, as each yearly vertex is fit to the straight-line interpolation defined by ensemble LandTrendr segments. Vertices are years in which the temporal trajectory in multispectral imagery changes (i.e., slope differs before and after vertex year). Fitted imagery can be constructed for any band or index that is present or can be created from original Landsat imagery, creating a temporally smoothed set of Landsat bands and indices, such as tasseled-cap transformation indices (Crist and Cicone 1984) and the normalized burn ratio (Key and Benson 2006), for all years even when data are missing for certain pixel-year combinations. Because the fitted imagery arises from a disturbance mapping methodology, forest attribute mapping is inherently consistent with those same disturbance maps (Kennedy et al. 2018a). See Davis et al. (2022) for more information about the ensemble LandTrendr process and its role in Northwest Forest Plan effectiveness monitoring. In figure 4, we present a diagrammatic representation of the LandTrendr multispectral ensemble process as a reference to readers, but we do not explore the process deeply in this document.

There are four steps for the ensemble LandTrendr process: (1) run LandTrendr using single index or band, (2) repeat with multiple bands/indexes, (3) create datasets for training/validation, and (4) use the ensemble model to identify final vertices. Details for each step are explained in the figure 4 caption.

continued on next page

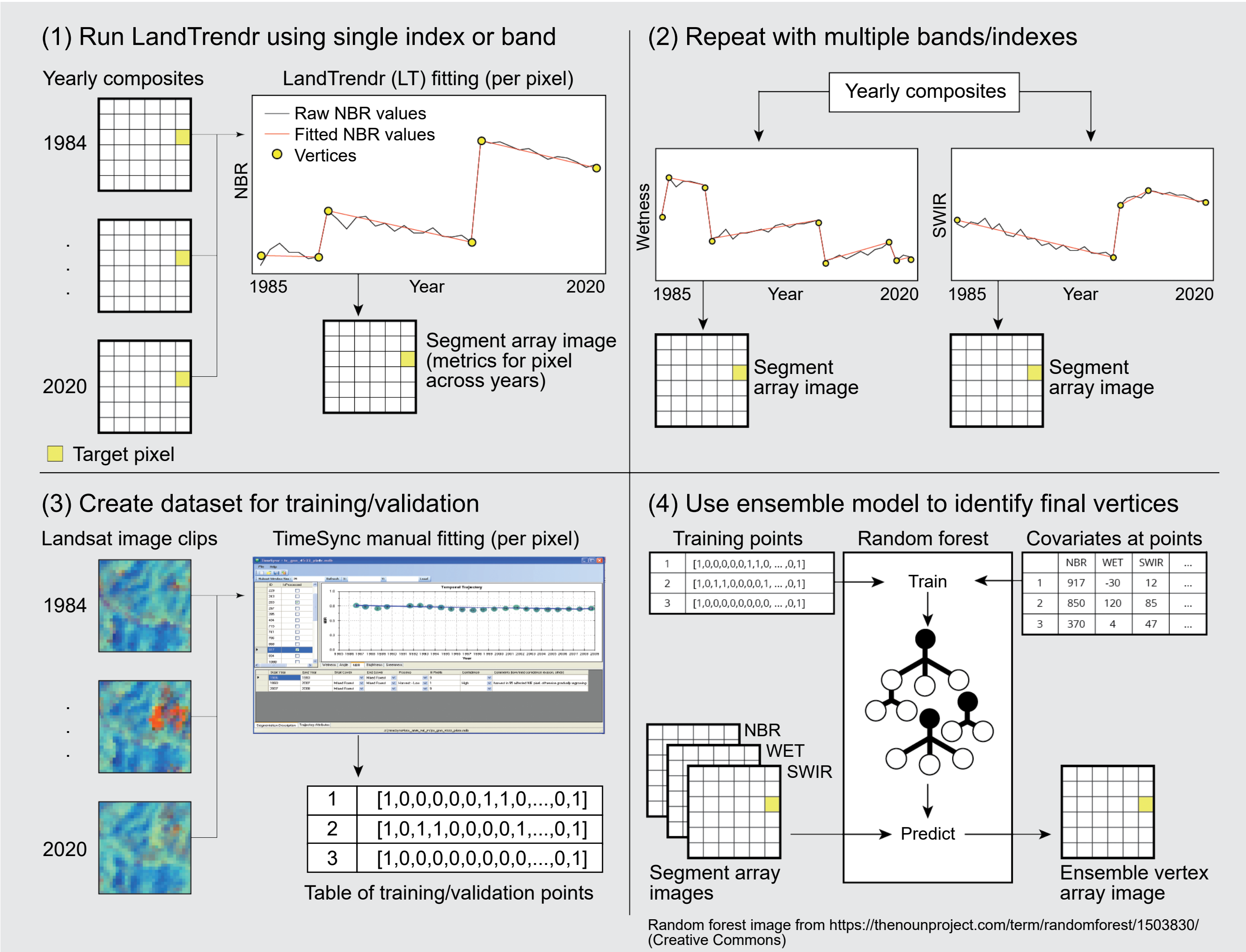

Figure 4—Conceptual figure for the ensemble LandTrendr process. Yearly composites (i.e., best available growing season pixels) of 13 spectral bands and indices are generated from the Landsat time series. The LandTrendr algorithm (Kennedy et al. 2010) is applied to each set of composites to identify years in which spectral trends change (i.e., vertices), likely because of a disturbance. An ensemble LandTrendr prediction of vertices is generated using the 13 LandTrendr outputs, visually interpreted spectral changes associated with forest succession using TimeSync (Cohen et al. 2010), and a random forest model. The resulting ensemble vertices are then used to generate yearly fitted imagery by fitting linear regressions between each vertex (sensu piecewise regression). NBR = normalized burn ratio; WET = tasseled-cap wetness index (axis 3); and SWIR = shortwave-infrared.

or recovering from disturbance) (Davis et al. 2015, Kennedy et al. 2018a). Such issues may arise because of spectral saturation, where the sensitivity of spectral reflectance decreases as forest biomass, and thus age, increase (Foody et al. 2003, Lu et al. 2012), meaning that small changes resulting from noise can drive shifts in forest attribute predictions from the modeling. Other potential drivers of multidecadal shifts in spectral indices independent of actual vegetation change include orbital drift for Landsat 5 decreasing the normalized difference vegetation index (Roy et al. 2016) or increasing exposure of non-leaf material, such as lichen and bark (Cohen and Spies 1992). We adopted an image stabilization method intended to minimize errors associated with slow spectral changes in older forests that lack any observed disturbance or recovery signal (box 2).

Box 2

Temporal Stabilization to Minimize Unrealistic Change in Old Forests

To address unrealistic spectral trends in older stable forests, we applied a temporal stabilization method to the temporally smoothed imagery. We identified stable pixels as locations where the ensemble LandTrendr indicated no disturbance from 1986 to 2017 and classified as older forest (structure similar to forests ≥80 years in age) in 1993 based on a preliminary gradient nearest neighbor (GNN) model using unstabilized imagery (fig. 5). This classification of older, stable forest was based on the old-growth structure index (OGSI), which incorporates thresholds for density of large trees, density of large snags, tree diameter diversity index, and downwood volume for different forest types in the study region (Davis et al. 2015, 2022). The OGSI was calculated for each pixel based on the forest attributes predicted for that pixel based on a preliminary GNN model using the unstabilized imagery.

We then stabilized imagery for pixels with OGSI values equivalent to those of a 200-year-old forest (i.e., old-growth forest) (Davis et al. 2015, 2022) by replacing the temporally smoothed indices from LandTrendr with the mean spectral index value across years. Previous stabilization methods held spectral indices constant across years for all pixels that the LandTrendr ensemble identified as not experiencing disturbance or postdisturbance regrowth (e.g., Kennedy et al. 2018a). Thus, mid-seral forests were generally stabilized such that they could not grow into late-successional categories, meaning that the primary mode of old-growth forest change was disturbance (e.g., loss of older forest). This may have biased results by underrepresenting gains in older forests. This could also underrepresent subtle losses (e.g., root disease); the ensemble LandTrendr used in this study should still represent some of these disturbances, though omission and commission rates exceed 20 percent when examining disturbances with the most subtle spectral signals (Kennedy et al. 2018a). In contrast, our current imagery stabilization method minimizes

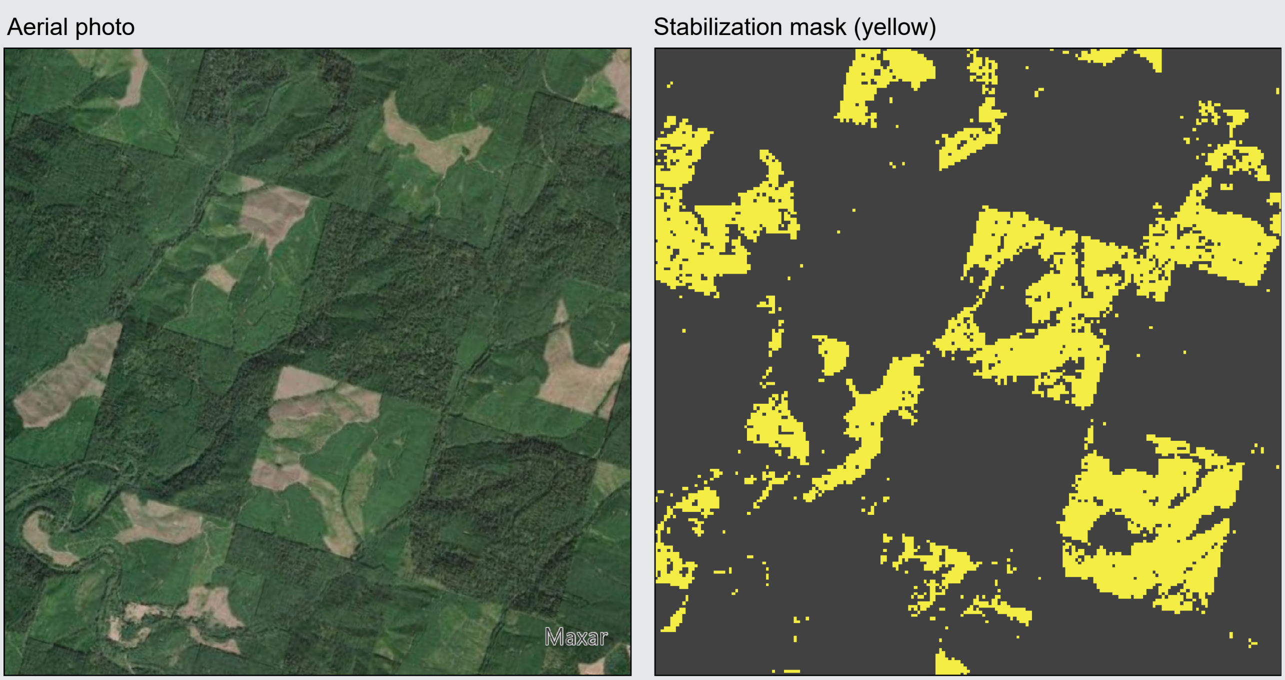

Figure 5—Example comparison of an aerial photograph taken August 18, 2020 (left) versus stable pixels (right).

continued on next page

unreasonable modeled changes in older forests, while still allowing for the transition of young stands into mature forest conditions over the course of decades. Whereas the forest area stabilized in the previous Northwest Forest Plan 20-year report was 15.7 million ha (43.1 percent of total forest area), stabilization in these models only accounted for 9.9 million ha (27.1 percent). Therefore, although our stabilization may discard some real changes associated with forest dynamics, we performed 37 percent less stabilization than in previous iterations and tended to stabilize older forest areas where spectral saturation (Foody et al. 2003, Lu et al. 2012) is more prevalent. Trends in Landsat imagery may have relatively low signal-to-noise ratios with respect to, for example, forest biomass change in undisturbed forests.

GNN modeling and imputation—
GNN provides a multivariate, nonparametric methodology to vegetation mapping (Kennedy et al. 2018a; Myroniuk et al. 2022; Ohmann and Gregory 2002; Wilson et al. 2012, 2013) that can form the basis for stand-, landscape-, and regional-scale analysis and monitoring across a broad range of conditions (e.g., Davis et al. 2015, 2016; Falxa and Raphael 2016). Although GNN can theoretically be used to model any ecosystem's vegetation, it is used here to combine geospatial data with forest inventory data to model and impute forest attributes (structure and composition) across all forest lands in the study area (fig. 6).

The framework uses canonical correspondence analysis, a method of constrained ordination (direct gradient analysis) (ter Braak 1986), to define a multivariate gradient space. Weighted Euclidean distances between plots and pixels in that multivariate gradient space are calculated such that the *k*-nearest neighbors for every pixel can be identified as those plots that minimize the distance to the pixel in question. For mapping, forest attributes are imputed to pixels based on some function of those nearest neighbors, such as the mean. The user can select differing values of *k* based on their objectives. For example, using only the nearest neighbor (i.e., $k = 1$) may help to maintain more realistic combinations of variables (e.g., species lists) by imputing only those combinations that were observed in the forest inventory (Ohmann et al. 2014). Conversely, using multiple neighbors ($k > 1$) may limit some types of mapping errors and allows for the estimation of prediction uncertainties (McRoberts 2001, 2012; McRoberts et al. 2007). Therefore, the selection of *k* depends on the questions under consideration and the variable being examined.

Spatial predictors were extracted for each plot using a 3- by 3-pixel (90- by 90-m, or 0.81 ha) footprint that encompassed the outer extent of the forested area of the field plot (fig. 3). We intersected the footprint with each spatial predictor and generated means across the 3- by 3-pixel footprint for each predictor. For satellite imagery, we extracted data for the same year as plot measurement to avoid temporal mismatches between field vegetation measurements and imagery-based measurements. Note that while the models were fit to these 3- by 3-pixel footprint

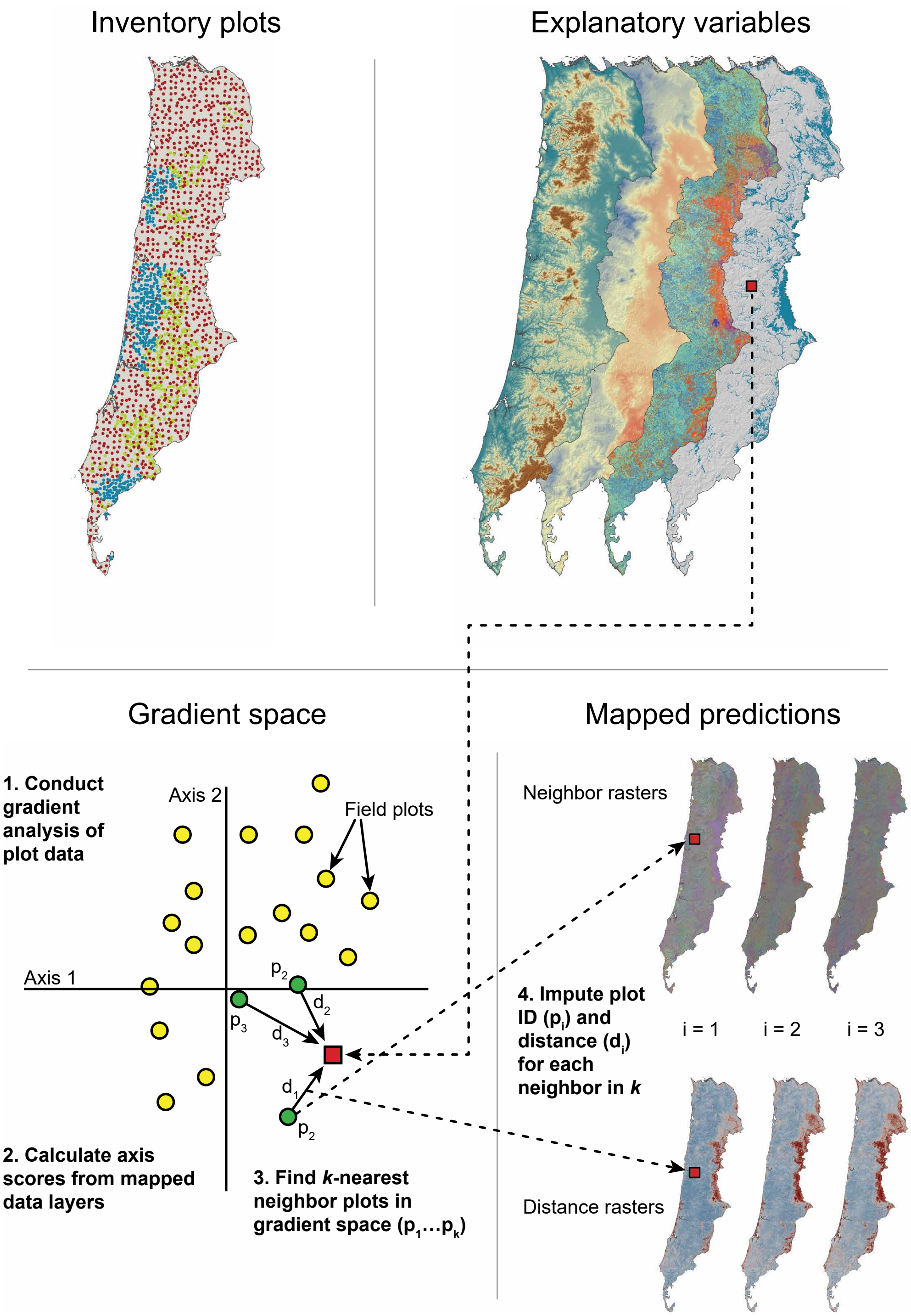

Figure 6—Overview of the gradient nearest neighbor approach for a single modeling region (Oregon Coast Range), including inputs (inventory plots and explanatory geospatial variables), the development of gradient space, and the imputation of plot data to individual pixels.

composites, mapping occurred at the 30-m-pixel scale, which results in similar or better map accuracy compared to other strategies, such as using only the central FIA subplot and overlapping pixel or a kernel density approach (i.e., weighting the central pixel more heavily) to modeling (Ohmann et al. 2012, Zald et al. 2014). Furthermore, we did not include heterogeneous plots that straddle multiple disparate vegetation conditions (e.g., harvest boundaries or forest to shrubland transitions) to ensure that we retained representative spectral signatures for the forests being imputed (i.e., minimized heterogeneity within plots). In this context, heterogeneous plots were identified as those where aerial photography, Landsat imagery, or the plot data themselves indicated more than one forest condition with dramatically different forest structural conditions, primarily determined by two researchers, Janet Ohmann and David Bell.

Nonforest mask—

Because GNN models were fit solely based on plots with at least 50-percent forest condition class, and nearly all imputed attributes were forest attributes (e.g., basal area), our application of the GNN imputation process produces outputs that apply only to forested lands (≥10-percent tree cover), though imputed forest attributes were generated, but not disseminated, for all lands regardless of forest tree cover. For most applications, we developed a 30-m resolution raster for map analyses that represents areas within the study area that are capable of developing into forests, or “forest capable lands” (Davis et al. 2015). This forest capable layer excludes urbanized areas, major roads, agricultural areas, water, lands above the tree line, snow, rock, and other persistently nonforested features.

Outside of the NWFP boundary, we used raster land cover data from both the Northwest and Southwest Regional Gap Analysis Projects (Grossman et al. 2008, Kagan et al. 2006) and mosaicked these rasters together. These datasets used NatureServe’s ecological systems (Comer et al. 2003) as the mapping categories. Forest and nonforest classes were created by grouping the NatureServe ecological systems, with forest and woodland types classed as forest and all others classed as nonforest based on the Ecological System Life Form (ESLF) (Kagan et al. 2006) codes. Members of our group (Janet Ohmann and Matthew Gregory) reviewed these ESLF classes to sort them into forest capable and nonforest capable classes to match the methods within the NWFP boundary as much as possible. Uncertain classes, such as woodland, were generally classified as forest such that GNN predictions would take precedence.

Consideration for Appropriate Use of GNN Maps

As discussed in previous monitoring reports (e.g., Davis et al. 2015), using vegetation maps, such as GNN-2020, as the basis of monitoring efforts requires a clear understanding of their uncertainties. The forest attribute maps from GNN-2020 are produced using statistical models based on remote-sensing and forest inventory data, both of which include their own errors. Some of those errors are documented elsewhere, such as the spectral saturation associated with relating Landsat multispectral imagery to forest biomass measurements (Foody et al. 2003), spatial patterns and biases in biomass mapping errors associated with nonforest (Huang et al. 2015) or the estimation of forest biomass based on tree diameters (Clough et al. 2016). In this section, we focus on a variety of concepts affecting the appropriate use of GNN-2020 (and other products generated by the GNN approach) for planning, monitoring, and estimation in forest landscapes at various spatial and temporal scales.

Reading the Accuracy Assessment Reports

When new GNN models and maps are generated, we provide GNN accuracy reports for each modeling region for the past year in the mapping cycle. These reports are available as part of user downloads for the model regions that intersect the requested geographic regions. In each modeling region, we take several approaches to quantifying accuracy at several relevant spatial scales (box 3 provides an example for live tree basal area). This accuracy assessment protocol applies a set of standardized metrics to facilitate comparison across variables and spatial scales. These metrics focus on data distributions, agreement between plot and modeled estimates, and the spatial pattern of local differences. Users should examine accuracy assessments for each attribute they use in analyses to understand if it meets their needs at the scale of analysis, as accuracy varies among forest attributes and between model regions.

Box 3

GNN-2020 Accuracy Assessment Report Example

As with previous iterations of gradient nearest neighbor (GNN) map production, GNN maps released by Bell et al. (2021) for 2020 (GNN-2020) are distributed with an accuracy assessment report that describes model parameterization and performance for commonly used predictor variables. Here, we provide an example of that accuracy assessment for the basal area of live trees ≥2.54 cm diameter at breast height. The accuracy assessment includes predicted versus observed plots, classification accuracy for binned data, area estimates of classes compared to USDA Forest Service Forest Inventory and Analysis (FIA)-based estimates, and comparisons of aggregate GNN-2020 predictions to aggregate FIA estimates at successively coarser landscape scales (8660 to 216 500 ha). We summarize each assessment below but refer the reader to the reports for a more complete explanation.

Local (plot)-scale accuracy—

For local-scale accuracy assessment, we present a scatterplot of paired observed and predicted values for all plots that were included in model development (fig. 7). Observed values come directly from the FIA plot data and are calculated across all forested conditions on a plot. Predicted values are calculated from a modified leave-one-out technique (i.e., imputed predictions, excluding the focal plot to avoid self-imputation during accuracy assessment) (Ohmann and Gregory 2002). Using a 90- by 90-m area centered on the plot location, we calculate predictions for each of the nine 30-m pixels. Following the approach for approximating a bootstrapped sample of neighbors, we extract the first seven independent candidate neighbors for each pixel, weight each neighbor based on values that approximate a bootstrapped sample, and calculate the weighted average (Bell et al. 2015a). The plot predicted value is calculated as the unweighted average across all nine pixels. We display the 1:1 line between observed and predicted values and color the dots based on a kernel density estimator to highlight frequent values in the distribution.

In addition, we present a confusion matrix based on binning the observed-predicted pairs using Jenks' natural breaks classification (Jenks 1967) (fig. 8). Both strict and fuzzy class accuracy are displayed, with fuzzy classes defined as +/− one class from the

continued on next page

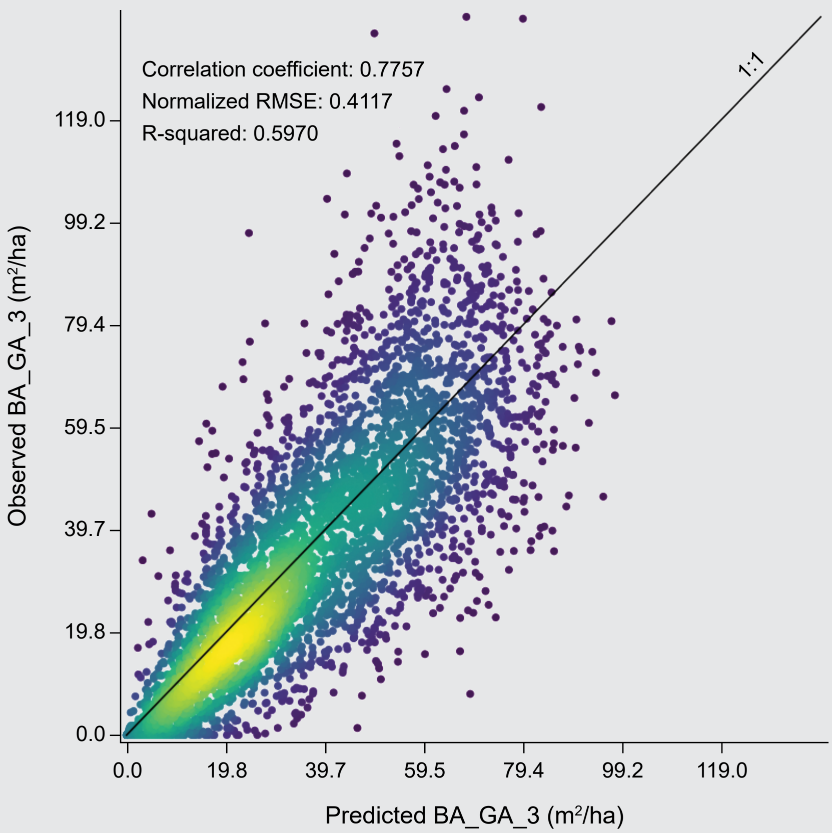

Figure 7—Scatterplot of observed versus predicted values for basal area of trees ≥2.54 cm diameter at breast height (BA_GE_3). Colors in the scatterplot are associated with a kernel density estimate of the underlying data distribution, where blue represents low density of data points and yellow represents high density of data points. Because points in dense space are often overlapping and hard to differentiate, these colors provide insight into the distribution of points at their highest density. RMSE is the root mean square error.

Observed class \ Predicted class	0.0–15.1	15.2–29.3	29.4–43.6	43.7–59.0	59.1–77.9	78.0–138.9	Total	Percent correct	Percent fuzzy correct
0.0–15.1	454	259	56	7	1	0	777	58.4	91.8
15.2–29.3	93	481	248	90	16	0	928	51.8	88.6
29.4–43.6	10	198	304	211	63	9	795	38.2	89.7
43.7–59.0	1	47	196	290	158	17	709	40.9	90.8
59.1–77.9	0	10	46	192	199	43	490	40.6	88.6
78.0–138.9	0	2	9	79	137	20	247	8.1	63.6
Total	558	997	859	869	574	89	3946		
Percent correct	81.4	48.2	35.4	33.4	34.7	22.5		44.3	
Percent fuzzy correct	98.0	94.1	87.1	79.7	86.1	70.8			88.3

Figure 8—Confusion matrix of observed versus predicted classes for basal area of trees ≥2.54 cm diameter at breast height. Classes are based on natural breaks classification of all plots in a modeling region.

continued on next page

actual class. Strict classes are shaded as darker gray, and fuzzy classes are shaded as lighter gray.

Regional scale accuracy—
For regional scale accuracy assessment, we compare forest area using three different methods (fig. 9). First, a design-based sample from the FIA plots is used to estimate nonforest area and forested area of the attribute split into six classes. Area expansion factors are associated with each plot based on the relative proportion of each condition on a plot. The plots that comprise the FIA estimate are those nearest in time to the model year, and because the FIA data are collected on a 10-year annual cycle, plot assessment dates may differ by up to 10 years from the current GNN model year. Furthermore, this simple assessment does not fully account for unsampled plots (e.g., those too dangerous to measure or where the landowner did not permit access to crews), meaning that the FIA acreage will not be the same as the official FIA estimates. For example, the production of FIA estimates of forest attributes use stratified sampling to deal with missing data. The second series shows GNN model predictions based on grouping pixel areas into the classes defined by natural breaks classification (Jenks 1967) of the plot data. The third series is based on an area estimation technique introduced by Olofsson et al. (2013). In this approach, an error matrix is constructed for all plots that were used in the design-based sample using the same bin endpoints. The error matrix is used to adjust the mapped, GNN-based estimates to approximate the relative distribution of this design-based sample. In addition, confidence intervals are constructed, which vary as a function of the agreement in the error matrix. Therefore, the GNN and error-adjusted area estimation for a given class may not exactly match the values based on the mapped forest attributes or the FIA plot data alone, but similarities among FIA, GNN, and error-adjusted area estimates should provide user confidence of broad-scale performance.

Accuracy across scales—
To assess intermediate scales of GNN model performance, we rely on the assessment protocol introduced by Riemann et al. (2010). We construct three spatial scales of hexagon tessellations (repeating patterns of a shape) across our model region—10-km (8660 ha), 30-km (78 100 ha), and

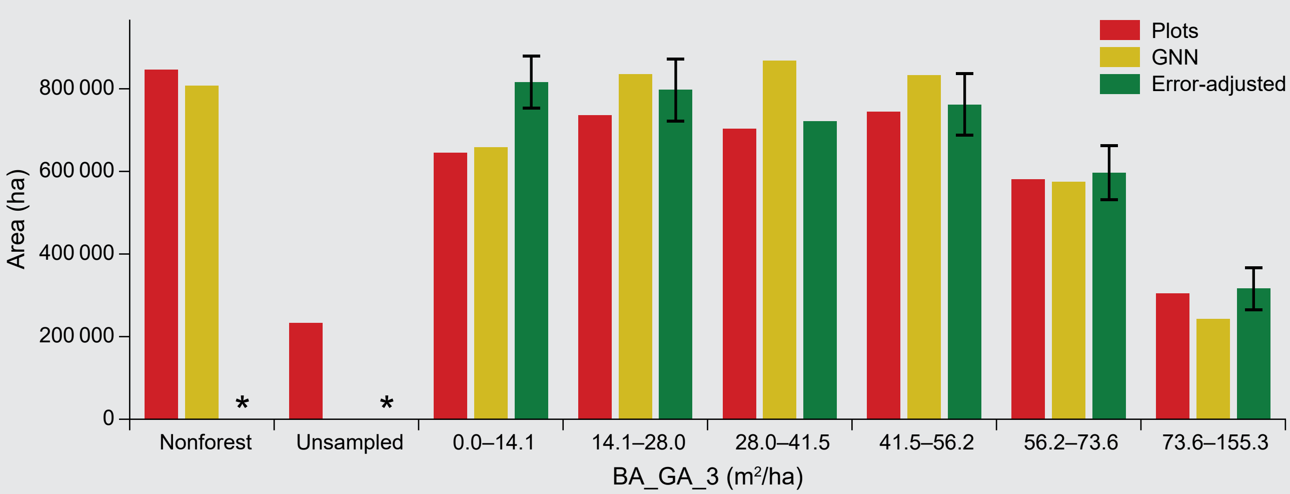

Figure 9—Area histogram of basal area of trees ≥2.54 cm diameter at breast height (BA_GE_3) based on design-based USDA Forest Service Forest Inventory and Analysis sample, map-based gradient nearest neighbor (GNN) predictions, and error-adjusted estimates. Error adjustment of map-based areas follows Olofsson et al. (2013). Asterisks indicate that the error-adjusted estimates were not performed in those cases.

continued on next page

50-km (216 500 ha) spacing between hexagon centers. For every hexagon, both the observed and predicted means are calculated across all plots within that hexagon. As the spatial scale of the hexagons increase, more plots are included in the calculation of mean observed and predicted values. The paired means for each hexagon are then displayed as scatterplots (fig. 10). Generally, as spatial scale increases, there is a tighter correspondence between hexagon observed and predicted mean values. When this behavior is observed in the scatterplot, it suggests that the correspondence between design-based estimates (plots) and model-based estimates (GNN) converge reliably as a function of spatial scale.

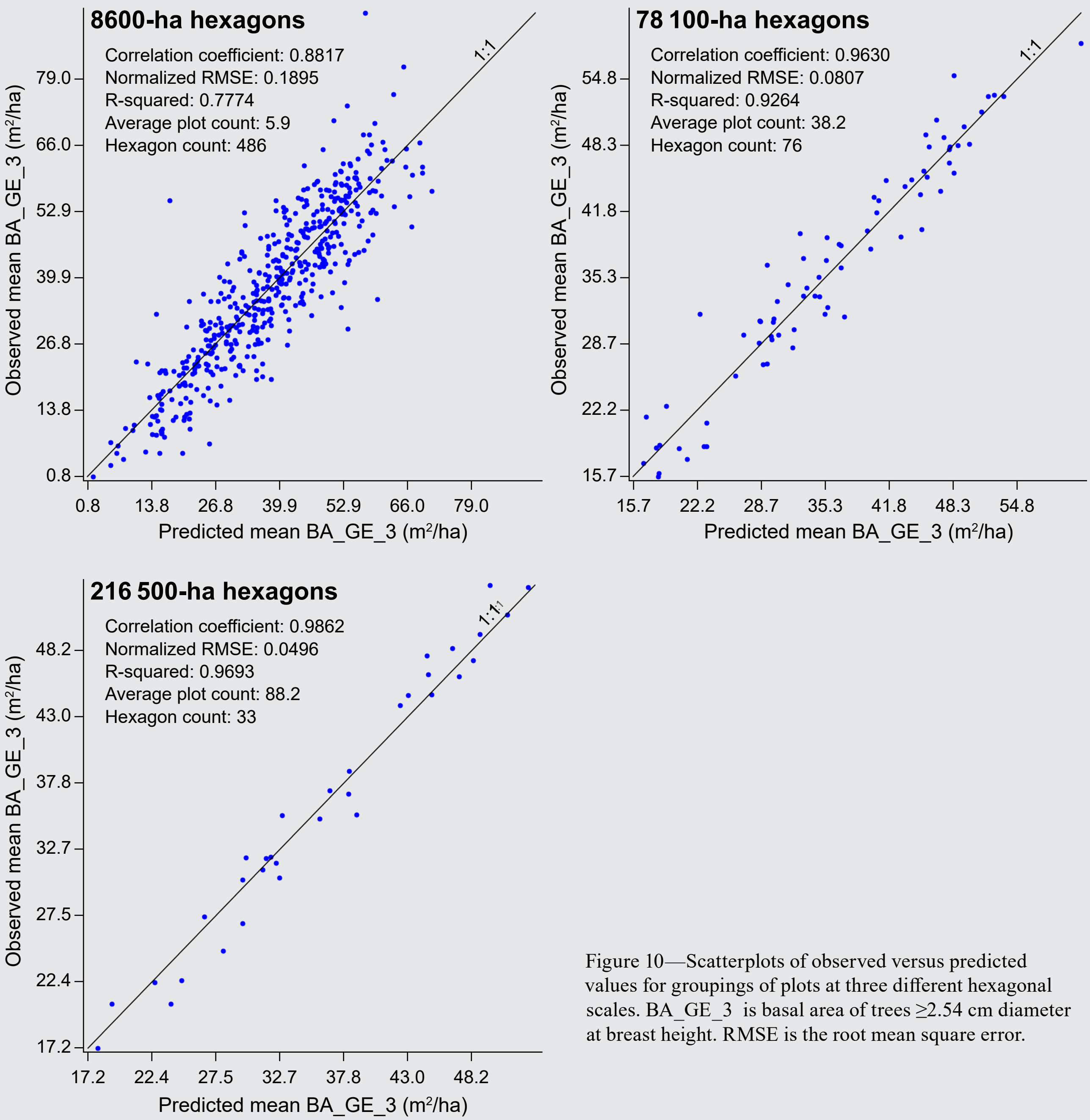

Figure 10—Scatterplots of observed versus predicted values for groupings of plots at three different hexagonal scales. BA_GE_3 is basal area of trees ≥2.54 cm diameter at breast height. RMSE is the root mean square error.

Forest Attributes and Geospatial Predictors

In addition to accuracy assessments available with the data, we encourage users to consider whether the forest attribute that they wish to examine is a characteristic related to the geospatial variables being used to generate the GNN-2020 maps: satellite imagery, climate, or topography. Our past work (Bell et al. 2015a) has indicated that forest attributes most closely related to live tree structure (canopy cover, tree size, and biomass) are better predicted than (1) dead tree components (snags and logs) that are relatively rare or not directly observed and (2) species composition (hardwood proportion), which is better represented by subannual variation in imagery (e.g., Wilson et al. 2012) as opposed to annual composites (fig. 11). Similarly, a long history of niche modeling indicates that environmental gradients determine species distributions at broader scales (e.g., Rehfeldt et al. 2006). Therefore, including climatic and topographic variables in GNN-2020 represents those distributions, though similar performance of maps generated with and without Landsat imagery indicate that fine-scale variation in species occurrence is more poorly represented (Henderson et al. 2014). However, local abundances of tree species may depend more strongly on local processes as opposed to climate (Canham and Thomas 2010), similar to disturbance dynamics; and remote sensing may contribute strongly to the mapping of forest community structure (Bell et al. 2022a, He et al. 2015), including the occurrence of rare plant species (Zimmerman et al. 2007).

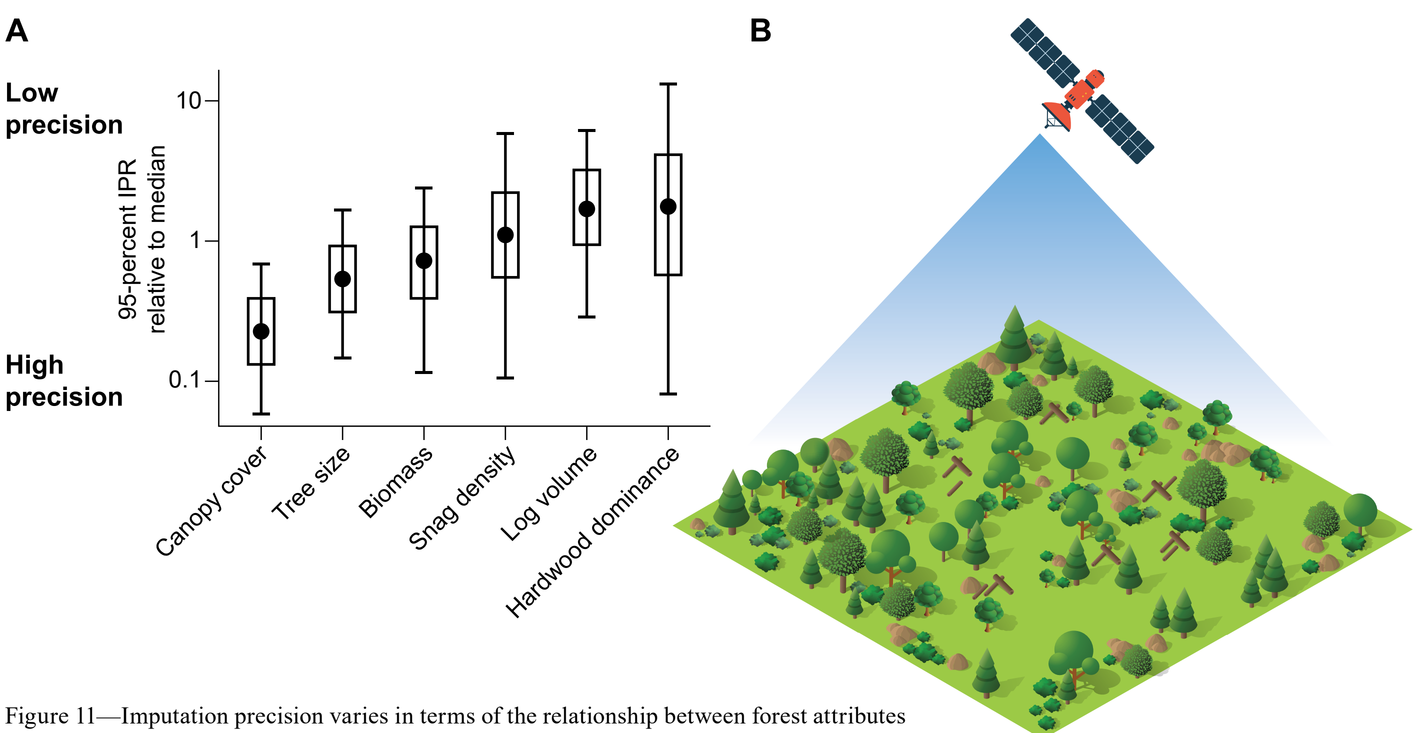

Figure 11—Imputation precision varies in terms of the relationship between forest attributes and the ability of satellite imagery to capture forest structure related to those attributes. (A) In the Oregon Cascade Mountains, six forest attributes are shown in terms of prediction precision at the pixel scale (modified from Bell et al. 2015a). The interpercentile range (IPR) is the difference between 97.5th and 2.5th percentile values across all pixels, with the box and whisker diagrams representing variation in IPR across pixels. (B) How forest structure related to those attributes may be captured by satellite imagery.

Even without direct observation of a forest attribute to the geospatial predictors, maps of those attributes may be valuable, especially when aggregated across broader areas. When a given forest attribute is not directly associated with, for example, the satellite imagery, the user should be aware that the correlation between the target attribute and the live tree forest structure captured by satellite imagery constrains the predictions. Thus, the accuracy of predictions for some forest attributes depends on their correlation with characteristics of the exposed forest canopy that are being observed by the satellites. For example, down log volume is generally low in a young, Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) plantation but can be quite variable in certain types of late-successional and old-growth forests (i.e., western hemlock [*Tsuga heterophylla* (Raf.) Sarg.]), leading to relatively high certainty in predictions in some places (young plantations) and not others (older forests). Despite the uncertainties at the scales of pixels in many forest attributes, aggregating GNN data across broader areas or within vegetation types or seral stages tends to improve prediction accuracy, as found for aboveground live forest carbon (Bell et al. 2018, Ohmann et al. 2014) and fuel type classification (Pierce et al. 2009). Furthermore, spatially coherent mapped patterns (e.g., continuous blocks of late-successional and old-growth [LSOG] forests) may provide additional confidence or may form the basis of spatially explicit analyses, such as the study of habitat connectivity (box 4). For more discussion of this concept, see the "Spatial Scale" section below.

Number of Nearest Neighbors

A central tradeoff in the development of forest attribute mapping and estimation with nearest neighbor techniques is the selection of the number of neighbors (k), which often represents a balance between individual forest attribute accuracy and multivariate attribute realism (Chirici et al. 2016, McRoberts 2012, Ohmann and Gregory 2002). Previous NWFP effectiveness monitoring reports have used k = 1 GNN (i.e., imputing the nearest neighbor only) as the basis of mapping to avoid mapping unrealistic combinations of forest conditions, such as cooccurrences of species that do not coexist in actual forests (Davis et al. 2015). However, pixel-level uncertainties can be large (Bell et al. 2015b, Loehle et al. 2015), especially when examining rare forest conditions, such as high-elevation forests recovering from wildfire (Bell et al. 2015a). Such uncertainties differ substantially depending on the forest attribute being considered, indicating that maps of predicted means and uncertainties (e.g., standard deviations) are both needed in the use of forest attribute maps; traditional accuracy assessments may not be sufficient for assessing map utility. To minimize the potential for unrealistic combinations of predictions, allow for the mapping of model precision across the study area, and dampen uncertainties related to rare forest conditions, we adopted an approximation of a bootstrap sampling approach for k = 1 GNN (Bell et al. 2015a). The bootstrap approximation uses multiple neighbors (k = 7) and applies a weighted mean, with

Box 4

Habitat Connectivity Mapping

A central motivation for our mapping work has been to improve understanding of the status and trends in late-successional and old-growth forest (LSOG) areas, with connectivity among patches guiding the application and interpretation of the GNN-2020 data. Thus, quantification of spatial patterns in LSOG (e.g., core, edge, finger, and scatter) forms the basis for applying LSOG mapping for Northwest Forest Plan monitoring (fig. 12). The quality of LSOG as a part of habitat for obligate species (e.g., northern spotted owl) depends on the spatial context of individual pixels, as species will respond differently to conditions inside a large patch versus small areas in a sea of young forest. Additionally, lone pixels not associated with a larger patch may be associated with omission (false negatives) and commission (false positives), as indicated by comparisons of LSOG spatial patterns from GNN-2020 and light detection and ranging (LiDAR) forest canopy height information (fig. 12). Where species require large vegetation patches, habitat quality modeling, including analysis of connectivity, with GNN-2020 data should focus on core and edge components because of their greater biological importance and added confidence in predictions (Davis et al. 2022).

A. LiDAR canopy height model

B. Morphological spatial patterns (OGSI 80)

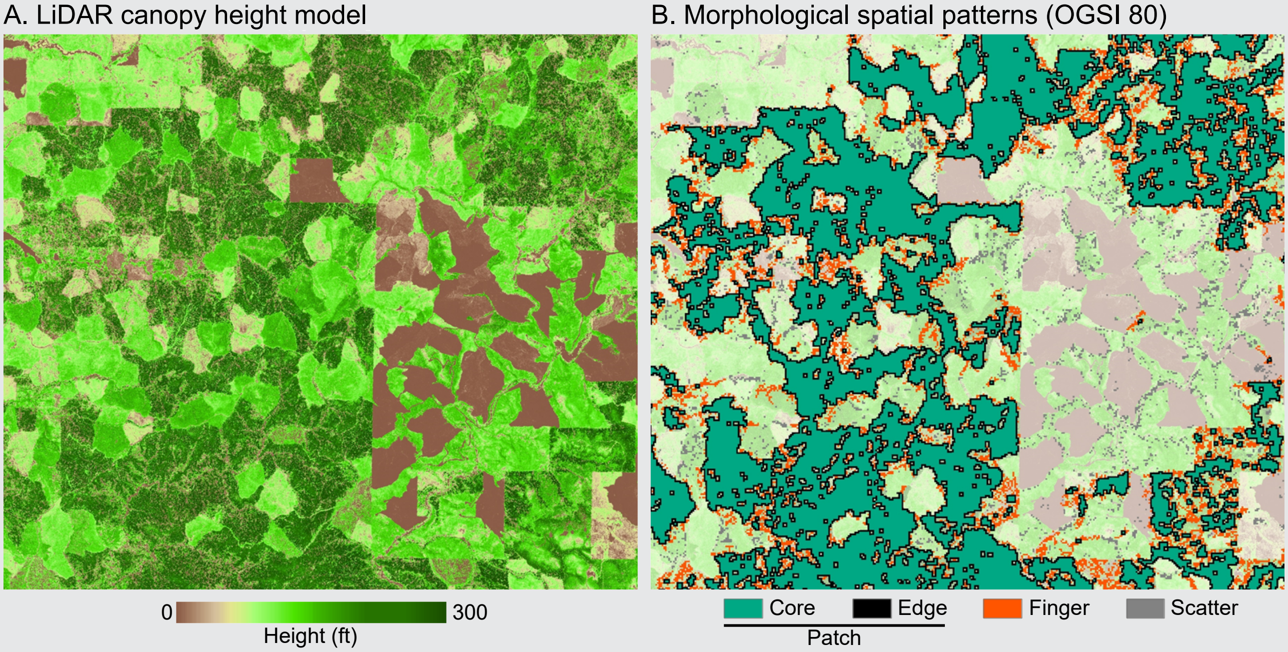

Figure 12—Example comparison of (A) forest height and (B) predicted late-successional and old-growth forest patches from GNN-2020 classified as patch core, patch edge, finder, or scatter. OGSI 80 is a classification of a continuous old-growth structure index using a ≥80-year age threshold—when forests commonly attain late-successional stand structure. Reproduced from Davis et al. (2022).

weights proportional to the probability that a bootstrap sample would result in that plot being the nearest neighbor for a pixel. Resulting predictions from the bootstrap approximation for aboveground forest biomass in California and western Oregon were highly correlated to $k = 1$ predictions but tended to exhibit lesser root mean square deviations (Battles et al. 2018). Because of the improved accuracy, we advise users to leverage GNN with $k > 1$ (e.g., bootstrap approximation) in most cases, but that they use $k = 1$ if unreasonable combinations of variables that could be created by averaging over neighbors (e.g., tree community composition) would be problematic for their analyses.

Spatial Scale

Regional- to national-scale vegetation mapping products based on multispectral remote sensing, such as Landscape Fire and Resource Management Planning Tools (LANDFIRE) (Rollins 2009), may be generated to support landscape-level assessments but still provide information as fine as the grain size of the imagery (30 m for Landsat). However, various factors influence the most appropriate spatial scale for map-based analyses. Even though the grain size of GNN maps is that of a Landsat pixel (30 m), accuracy assessments are not appropriate at scales finer than the 0.81-ha or 9-pixel footprint (Bell et al. 2015b). Analyses at these fine scales require great care. Aggregated GNN map predictions at landscape- to regional-scales (tens to thousands of square kilometers) are similar to plot-based estimates (Ohmann et al. 2014, Pierce et al. 2009). For example, although aboveground live biomass predictions in California at 30- to 90-m scales often underpredict at the high end and overpredict at the low end relative to light detection and ranging (LiDAR)-based predictions, county-level GNN predictions were equivalent to FIA design-based estimates ($R^2 = 0.97$) (Battles et al. 2018). Additionally, comparisons of model-based versus design-based estimation of forest live carbon in California for small area estimation indicated that mean predictions were not systematically biased (10 000- to 68 400-ha areas) and that below 25 000 ha, GNN-based variance estimates (McRoberts et al. 2007) were superior to design-based methods (Bell et al. 2022b). These findings support the application of spatially aggregated GNN data to monitor forest resources and support forest planning (e.g., boxes 5 and 6).

Additionally, relative gradients in forest structural attributes (e.g., low to high biomass or young to old forest) produced by GNN are similar to LiDAR-based maps at stand- to landscape-scales (hundreds to thousands of hectares), though often biased relative to the LiDAR-based maps (Bell et al. 2018). Even though GNN predictions may overpredict at low biomass and underpredict at high biomass relative to LiDAR, the similarity in forest structural gradients means that maps of modeled vegetation attributes important to wildlife based on GNN- and LiDAR-based forest attribute maps may result in similar maps (Ackers et al. 2015). We do not think that there is a set spatial scale for appropriate use of GNN but rather a variety of scales depending on the forest attribute under consideration and the

Box 5

Providing Watershed-Scale Context Around Planning Areas Using DecAID

Gradient nearest neighbor (GNN) maps are valuable in providing the landscape context around a planning area. The Decayed Wood Advisor (DecAID) is a planning tool used by wildlife biologists and ecologists as a single source for best available science to guide management of snags, partially dead trees, and downwood (Mellen-McLean et al. 2017). It is mostly used on national forests within the Pacific Northwest, but the U.S. Department of the Interior Bureau of Land Management, county extensions, and small private woodlands in Oregon have recently used DecAID to inform management decisions. Snags and down dead wood are important habitat attributes for many wildlife species. DecAID is a publicly available Web platform that incorporates three main components: (1) an analysis of vegetation inventory data or statistical summaries in the form of histograms for snag diameter, snag density, downwood diameter, and downwood percentage of cover for both reference conditions from FIA data and current conditions from GNN maps for all landscapes across Oregon and Washington grouped by wildlife habitat types (Johnson and O'Neil 2001) in 5th hydrologic unit code watersheds or larger (generally >60 000 ha); (2) a review of published literature and other available data on wildlife use of decayed wood attributes, primarily in Oregon and Washington, where studies with quantitative data through 2016 are synthesized and presented at three statistical levels of use (called tolerance levels); and (3) summaries for practitioners intended to help inform meaningful treatments resulting in the best ecological outcome for a variety of wildlife species. For example, data such as the distribution of snags by size class in late-successional ponderosa pine (*Pinus ponderosa* Lawson & C. Lawson)/Douglas-fir forest plots are easily summarized by DecAID and can serve as a benchmark for managers that need to assess specific habitat attributes.

way the GNN data are being used. We suggest that users examine the accuracy of the variables in which they are most interested in using (e.g., figs. 7 and 8) and the behavior as plots are aggregated (figs. 9 and 10) to determine the scale at which the data can be applied. Thus, users can rely on our accuracy assessments, the above descriptions of application-specific considerations, and our previous literature, in making their decision of appropriate scale for use of GNN maps. When in doubt, we propose that users err on the side of caution and use the GNN map products at coarser scales if appropriate for the question at hand.

Tracking Change Through Time

As a basis for monitoring, data need to be consistent through time (i.e., related to patterns and drivers in the same way, regardless of year) to minimize the role of methodological artifacts on map changes. Our use of temporally smoothed spectral indices generated by the ensemble LandTrendr algorithm addresses temporal consistency because we used the trends in spectral indices, rather than the raw data, to build our models, making year-to-year variation in vegetation mapping a more reasonable representation of reality (Kennedy et al. 2018a). Furthermore, the use of only FIA annual plots in our imputation process minimized methodological

Box 6

Identifying Broad-Scale Abundance of Unique Habitat Types

USDA Forest Service Forest Inventory and Analysis estimates provide accurate information about the abundance of common forest types on the landscape, but as with other systematic random, poststratified sampling efforts, they capture uncommon forest types with lower precision. When managers need to know where a specific forest type occurs within a forest, gradient nearest neighbor (GNN) maps can add key information. As part of a regional effort to understand the abundance of several unique habitats, GNN species maps were used to identify where basal area of whitebark pine (*Pinus albicaulis* Engelm.), quaking aspen (*Populus tremuloides* Michx.), Oregon white oak (*Quercus garryana* Douglas ex Hook.), or western juniper (*Juniperus occidentalis* Hook.) was identified using species-level basal area attributes (fig. 13). Zonal statistics were then used to assess the number of acres in each national forest in Oregon and Washington state where these species were predicted to occur (table 1). The 2020 GNN maps give managers an indication of where certain species are likely to occur and can make field verification efforts more efficient. Forest-level summary, even of these relatively rare species, is appropriate for

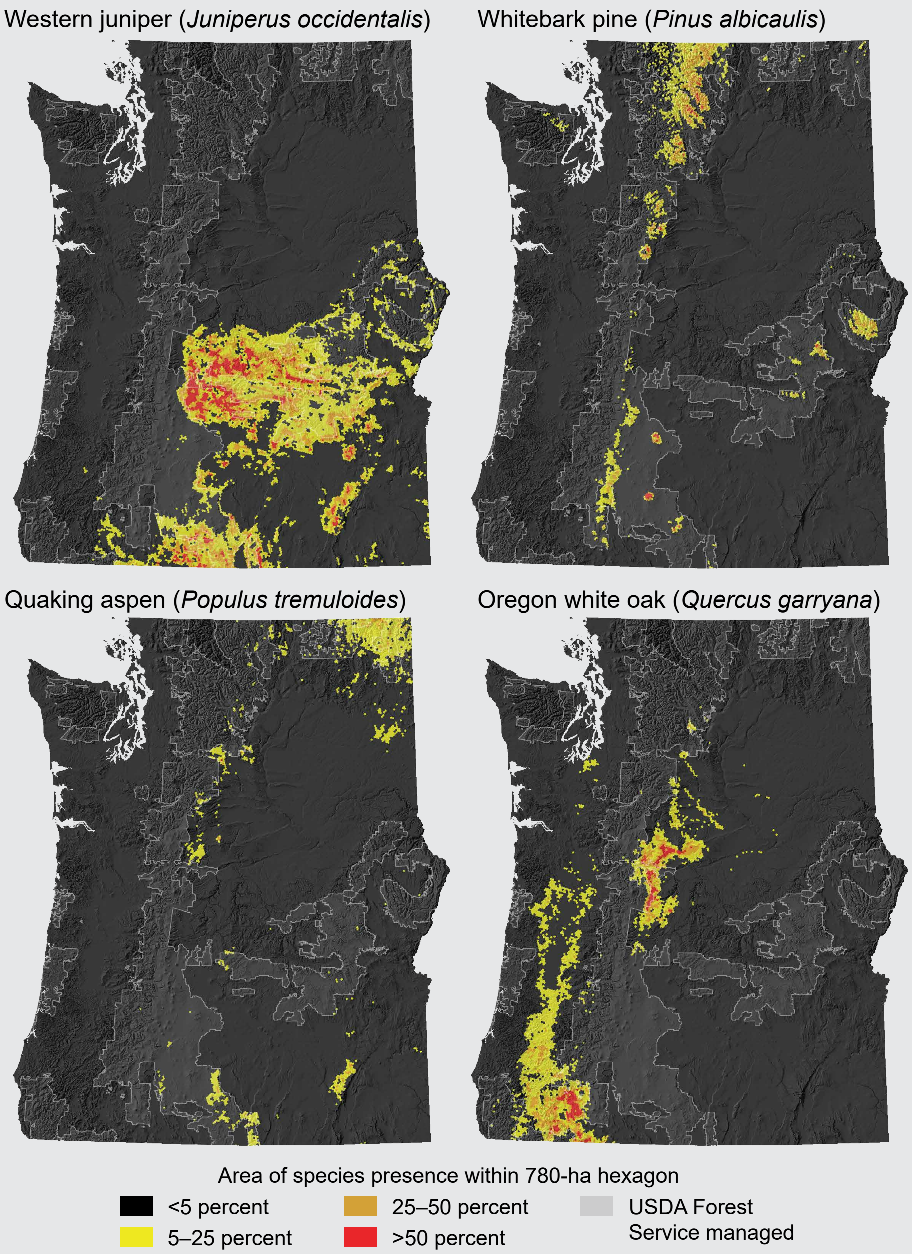

Figure 13—Distributions of four species for Oregon and Washington based on 2020 gradient nearest neighbor map data that was summarized for 780-ha hexagons to improve visual clarity of the maps.

continued on next page

Table 1—Predicted area with western juniper (*Juniperus occidentalis*) on each national forest in Oregon and Washington based on 2017 western juniper basal area maps from 2020 gradient nearest neighbor map data

Forest name	Western juniper present
	Acres
Columbia River Gorge National Scenic Area	33
Colville National Forest	0
Deschutes National Forest	39,741
Fremont-Winema National Forest	276,513
Gifford Pinchot National Forest	0
Malheur National Forest	283,599
Mount Baker-Snoqualmie National Forest	0
Mount Hood National Forest	321
Ochoco National Forest	278,094
Okanogan-Wenatchee National Forest	0
Olympic National Forest	0
Rogue River-Siskiyou National Forest	2,546
Siuslaw National Forest	0
Umatilla National Forest	69,925
Umpqua National Forest	777
Wallowa-Whitman National Forest	100,864
Willamette National Forest	21
Regional total	1,043,433

Note: The six national forests with the most western juniper acreage are in eastern Oregon; this corresponds with the western juniper map in fig. 13.

strategic planning. Here, it is important to emphasize that individual pixels in the raster are not assumed to be accurate, even for the most common species, and especially not for these unique habitat types. In this example, the quick GNN summary informed the decision to pursue mapping optimized for these unique habitats, but the specific pixel maps were not used because the accuracy for these species is not sufficient when assessing acreages over areas smaller than national forests.

artifacts in the maps resulting from changes in forest inventory plot design (e.g., shift from periodic to annual plot designs for FIA). However, by limiting our modeling and imputation to plots with the annual plot design, we also limited the time frame used for model fitting (2001–2016). As a result, to map conditions prior to 2001, we must assume that the multivariate relationships of remote-sensing and environmental predictor variables with forest inventory response variables were stationary through time but not necessarily in space (e.g., developing models for different regions). For example, a clearcut in 1985 is assumed to appear the same to the Landsat satellites as a clearcut in 2005, thus if all plots across years are used in the same model, a 2005 plot (from some other location) can be imputed to the 1985 pixel. Similarly, others have assumed that models based on reference data from one time step can be applied to other time steps for a spectrally harmonized approach (e.g., Moisen et al. 2016, Ohmann et al. 2012, Vogeler et al. 2018).

However, we do not account for a potential source of error caused by subtle shifts in spectral data, such as those observed for Landsat 5, where gradual changes in orbital path caused erroneous detections of vegetation browning and underestimations of greening (Roy et al. 2016). However, in forested ecosystems, imagery stabilization procedures, as described above, may minimize the impact of shifting spectral indices (Schleeweis et al. 2016). Thus, while uncertainties associated with the construction of a time series of vegetation mapping data remain, we have taken several steps to maximize the temporal consistency of those maps. These efforts to maximize temporal consistency provide added confidence in mapping forest vegetation trends and assessing rates of change (box 7). Note also that our imagery stabilization process may help avoid unrealistic losses of late-successional and old-growth forest in otherwise stable forests but could also

Box 7

Trends in Late-Successional Old-Growth Forests in the Northwest Forest Plan Area

Late-successional and old-growth (LSOG) forest monitoring for the Northwest Forest Plan (NWFP) Effectiveness Monitoring Program (Davis et al. 2015, 2022; Moeur et al. 2011) has been a driving force behind gradient nearest neighbor (GNN) method development in general and the 2020 GNN map (GNN-2020) data in particular. Annual mapping of older forests from GNN-2020 highlights interannual variability in predicted rates of change as well as lagged change associated with large disturbances (fig. 14). We observed general increases in LSOG forest area since the 1990 timber injunction, which led to NWFP adoption in 1994, though annual rates of increase (area in focal year divided by area in previous year) vary from 0 to 0.4 percent. Increasing trends were reversed during major fire years (e.g., 2002) when large amounts of LSOG forest were affected by wildfires; and wildfire losses, depending on the timing of the event, sometimes are not observed until the next year's Landsat imagery. Since the end of this study's period of measurement (2017), additional fires, especially the 2020 Labor Day fires (Reilly et al. 2022), have dramatically decreased the extent of LSOG. Further work is needed to determine defensible ways to explicitly assess uncertainty in change analyses.

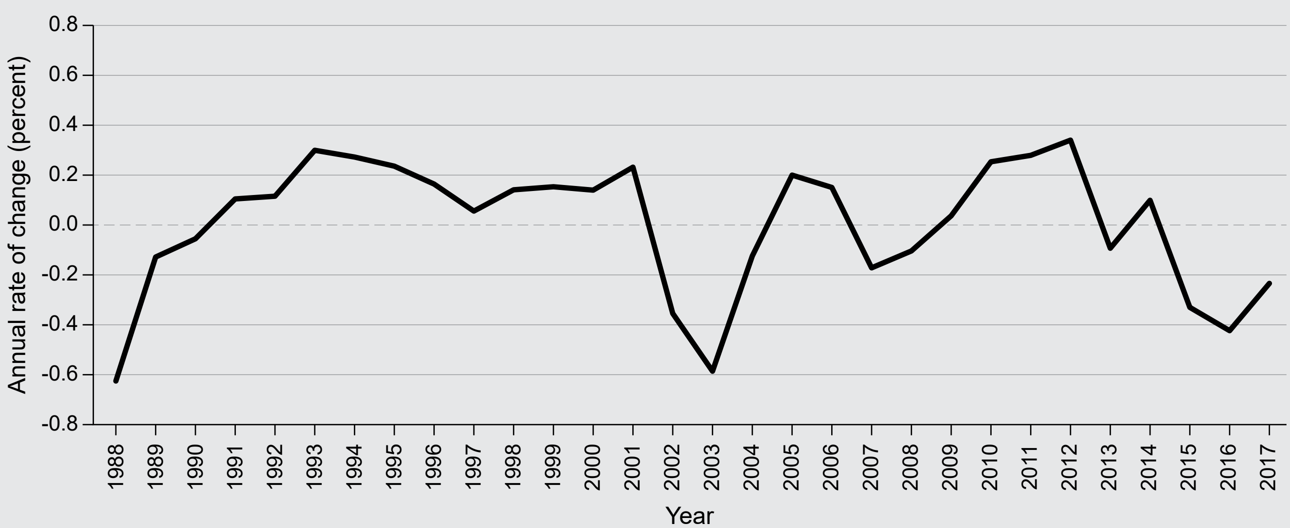

Figure 14—Predicted annual rates of change in forests with structure similar to forests >80 years of age based on 2020 gradient nearest neighbor map differences. Modified from Davis et al. (2022).

contribute biases to other forest attributes, such as underprediction of live tree biomass accumulation in those older forests.

It is also important to note other sources of uncertainty arising from regeneration harvest edges, which often have substantial shadowing, sometimes resulting in erroneous gains or losses along disturbance margins (Davis et al. 2015) and potentially indicating that LSOG forest cores are most reliably mapped. We have observed a tendency for GNN to predict increasing older forest characteristics following forest thinning disturbances, which is not consistent with basal area reductions associated with thinning; elevated frequency of canopy gaps and shadows and basal area reductions from thinnings would most likely cause little change in the old-growth structure index. Other research has noted that the tasseled-cap transformation indices can be sensitive to shadows (e.g., Baumann et al. 2014). Such errors lead to the false impression that thinning operations are rapidly reestablishing older forest, which is not silviculturally or ecologically possible for young conifer plantations over the temporal span of this monitoring cycle (a few decades). Likewise, we have observed that rapid ground vegetation reestablishment (greening) after forest stand-replacing events (e.g., high-severity fire) can result in apparent rapid gains in older forest. For some applications, such as old-growth mapping, high-severity disturbance maps can be used to mask out these rapid recovery events (Davis et al. 2022). In other words, once a stand-replacing event affected a pixel, it could not be classified as “older forest” following the disturbance event, given that it takes longer than the length of the Landsat time series for a regenerated stand to redevelop into older forest.

For remotely sensed vegetation mapping to be used as a component in monitoring activities, changes in those vegetation maps should be accurate and sensible. Our imagery stabilization, plot screening to remove heterogeneous plots, and integrated disturbance and vegetation mapping framework was explicitly designed to ensure that changes in forest attributes, especially those associated with LSOG forests, were consistent with disturbance and recovery processes occurring on the landscape (Kennedy et al. 2018a, Ohmann et al. 2014). The fitted satellite imagery takes into account disturbance history for a given pixel, providing trends in spectral indices upon which vegetation mapping is generated, but does not preclude errors in the mapping of change. Aboveground live biomass change predictions produced by GNN mapping in California and Oregon were most accurate for disturbed forests (i.e., losses of biomass), with recovering forests exhibiting moderate accuracy and stable forests exhibiting little accuracy in change predictions (Battles et al. 2018). Furthermore, local predictions (e.g., tens to hundreds of hectares) can be quite error prone (Bell et al. 2015a) even though they may represent gradients in forest attributes well (Bell et al. 2018). Skepticism is warranted for predictions that don’t make ecological sense or for areas smaller than the 3- by 3-pixel footprint (0.81 ha) (Bell et al. 2015b).

Temporal Scale

The appropriate temporal scale for examining change depends on the processes of interest (Masek et al. 2015). Acute and abrupt disturbances, such as wildfire or timber harvesting, generally occur over the course of a single year. Even chronic and protracted disturbances, such as drought stress, insect attack, or disease tend to occur over a handful of years. Thus, using Landsat-based approaches that use annual composites, such as GNN, to examine the impacts of disturbance on forest attributes can characterize changes occurring over only a few years. In contrast, it may be difficult to correctly map, and thus interpret, annual changes in slower and more subtle dynamics in forest attributes that are associated with undisturbed forests, such as the accumulation of large live or dead trees (Bell et al. 2021), biomass change (Powell et al. 2010), or the development of vegetation attributes that are important parts of northern spotted owl or marbled murrelet (*Brachyramphus marmoratus* J.F. Gmelin, 1789) habitat (Davis et al. 2016, Falxa and Raphael 2016, Lorenz et al. 2021) over decades. From year to year, pixels may blink on and off (i.e., move in and out of a classification over time) for a given forest attribute, leading to substantial noise regarding forest landscape change. If these protracted processes are of interest, users may want to examine trends over multiple decades (e.g., bookend analyses) or over large areas (e.g., watersheds) to ensure that changes are directional and sensible.

Almost all accuracy assessments of GNN-2020 have focused on mapping status, not change. In the absence of disturbance, the amount of change experienced in a year or a decade is often small compared to the initial conditions, leading to the potential for substantial uncertainty in rates of change. For example, even in a fast-growing, mature forest stand, aboveground live biomass change in Pacific Northwest forests will, on average, increase by much less than 1 percent per year (Bell and Gray 2016). However, GNN mapping often better predicts substantial changes associated with forest disturbance compared to changes within stable (undisturbed) forests (Battles et al. 2018). Additionally, mapping of small (e.g., wind) or long-duration (e.g., stress) disturbances and their consequences may be more prone to error than larger, shorter duration events (e.g., fire or harvest) (Schroeder et al. 2017). As a result, users interested in examining forest change with GNN mapping should recognize that small, relative changes in a forest attribute are innately more uncertain compared to rapid, large-magnitude changes associated with forest disturbance, especially at the pixel level, but perhaps also at broader scales. Whether a user is interested in fast or slow changes in forest resources, Web-based visualization can help users assess and interpret GNN-2020 data relevant to their particular questions (box 8).

Box 8

User-Friendly Visualization Tools for Tracking Trends Over Time

An ongoing partnership between the USDA Forest Service Pacific Northwest Research Station and Oregon State University makes continued development of user-friendly tools possible. These public-facing tools make using 2020 gradient nearest neighbor (GNN-2020) data far more accessible to managers. Google Earth Engine™ combines annual data from 30+ years and summarizes them at a user's desired scale, allowing managers to focus on the attributes and ecologically relevant processes they need to monitor. To simplify users' introduction to the GNN-2020 data and facilitate data exploration, we developed the GNN trend tool, a Web-based platform for visualizing and exporting landscape- and regional-scale data summaries (fig. 15).

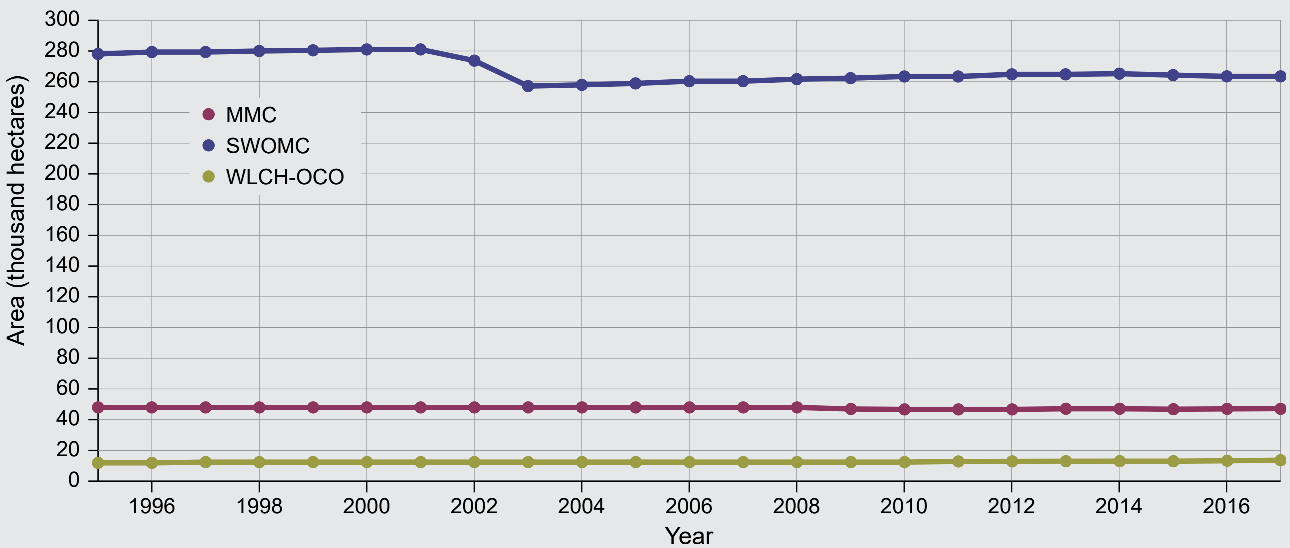

Figure 15—Example output from the gradient nearest neighbor (GNN) trend tool. In this example, the trend in area of large live trees from 1995 to 2017 is shown for the Rogue River-Siskiyou National Forest. Wildlife habitat types listed here are moist mixed conifer (MMC), southwest Oregon mixed conifer (SWOMC), and westside lowland conifer/hardwood in the Oregon coastal region (WLCH-OCO). The obvious drop in SWOMC is associated with the Biscuit Fire in 2002.

Categorical Variables and the Danger of Thresholds

Classification of continuous variables (low to high values of a forest attribute) can be used to produce categorical variables based on a set(s) of attribute threshold rule(s), for example, when binned into natural breaks for the accuracy assessment. Note that categorical variables can be sensitive to small errors and should be examined with care. For example, if a user is interested in which lands support at least five large (>50 cm DBH) snags per hectare, minor changes in the number of snags observed on plots, and thus which plot is imputed where, can determine the difference between the presence or absence of a given pixel. As a result, presences may blink on and off through time, especially at finer spatial scales. Note that similar behavior can be observed for plot data, not just map predictions. Continuous variables allow subtle gradation and empower users to customize their own rule set(s) or to use the variable as a gradient predictor variable in species distribution modeling as mentioned above.

Future Research Directions

When we highlight the strengths and limitations of the GNN-2020 and similar products, several future lines of research emerge. We are pursuing several areas of work to improve regional forest structure and composition mapping:

- **Annual GNN mapping—**The roughly 5-year cycle required for research, modeling, and mapping highlights a need to develop repeatable workflows for annual updates to ensure that GNN maps are applicable to current landscapes. For example, the most recently completed mapping covers 1986–2017; however, numerous infrequent but not unprecedented fires, including the large, wind-driven Labor Day fires (Reilly et al. 2022), dramatically altered the landscape, reducing GNN-2020 map utility in disturbed areas. Based on our experience (Bell et al. 2021, Davis et al. 2022, Myroniuk et al. 2021), we are currently working to leverage Google Earth Engine™ to update our imagery annually to form the basis for provisional GNN mapping each year.
- **Integrating stand height—**Incorporating stand height from LiDAR as a predictor in GNN mapping can dramatically improve forest structure predictions, as indicated for the Deschutes National Forest in central Oregon (Zald et al. 2014). However, LiDAR data are generally not appropriate for regional mapping and monitoring because it is not available for all lands; LiDAR acquisitions differ in terms of the sensors used, year of measurement, time of year, and several other characteristics that make compilation and consistent use of those data challenging; and remeasurement across years is still relatively rare outside of certain geographies. However, we have begun incorporating height information from statewide, biennial National Agricultural Imagery Program (NAIP) imagery (Strunk et al. 2020, 2022) as a predictor for forest structure and composition mapping at regional scales (Bell et al. in 2022a). Initial results seem to indicate substantial improvements for some forest structure predictions, such as large tree (>75 cm DBH) density, with the addition of stand height from NAIP. Research is still needed to understand the full benefits of such an approach.
- **Integrating high-resolution photography—**Another avenue for leveraging NAIP imagery is the incorporation of measures of fine-scale variation in vegetation for subpixel (<900 m^2) areas. Subpixel heterogeneity in vegetation may be particularly important for mapping postdisturbance forest structure and composition, providing support for understanding the effects of, and recovery from, wildfires.

Conclusions

The integration of repeat-measurement forest inventory data with time series remote sensing as a basis for monitoring the status and trends in forest attributes and ecosystem properties is an increasingly important strategy for supporting forest management and planning. The maturation of both nearest neighbor imputation as a key approach and the increasing access to time series of satellite imagery have facilitated the development and adoption of the GNN imputation mapping approach as an important tool for delivering complex, multivariate forest attribute data to managers. Map users will undoubtedly leverage the complex, multivariate data to address a variety of research and management questions, highlighting the need for clear guidance on the use of those data.

This report provides guidance in several areas that users should be familiar with before using GNN-2020 data. We describe the GNN method and key inputs, including FIA data, Landsat imagery, and other geospatial predictors. We describe the extensive accuracy assessment reporting performed before map deliveries, including advice regarding how to interpret various diagnostics and statistics. We discuss various concepts regarding the appropriate use of maps, including whether forest attributes might be well characterized by the geospatial predictors, the appropriate spatial and temporal scales for using the data, and the types of forest dynamics/processes that may or may not be well represented in the maps, among other things. Most importantly and broadly, we advise users of any mapping product to think carefully about their needs and consider how much uncertainty is acceptable and whether it is reasonable to expect the map in question to represent the forest attributes of interest across space and through time.

Acknowledgments

This publication was funded by the USDA Forest Service Pacific Northwest Research Station and Pacific Northwest Region. We appreciate the valuable comments of Tom DeMeo and Irene Lockman (Pacific Northwest Region), Barbara Garcia (Northern Region), and Andrew Hudak and Karen Schleeweis (Rocky Mountain Research Station), which led to significant improvements in this work.

U.S. Standard Equivalents

When you know:	Multiply by:	To find:
Centimeters (cm)	.394	Inches
Meters (m)	3.28	Feet
Kilometers (km)	.621	Miles
Hectares (ha)	2.47	Acres
Square meters (m^2)	10.76	Square feet
Trees per hectare	.405	Trees per acre

Metric Equivalents

When you know:	Multiply by:	To find:
Inches	2.54	Centimeters
Acres (ac)	.405	Hectares

Literature Cited

Ackers, S.H.; Davis, R.J.; Olsen, K.A.; Dugger, K.M. 2015. The evolution of mapping habitat for northern spotted owls (*Strix occidentalis caurina*): a comparison of photo-interpreted, Landsat-based, and LiDAR-based habitat maps. Remote Sensing of Environment. 156: 361–373. https://doi.org/10.1016/j.rse.2014.09.025.

Banskota, A.; Kayastha, N.; Falkowski, M.J.; Wulder, M.A.; Froese, R.E.; White, J.C. 2014. Forest monitoring using Landsat time series data: a review. Canadian Journal of Remote Sensing. 40(5): 362–384. https://doi.org/10.1080/07038992.2014.987376.

Battles, J.J.; Bell, D.M.; Kennedy, R.E.; Saah, D.S.; Collins, B.M.; York, R.A.; Sanders, J.E.; Lopez-Ornelas, F. 2018. Innovations in measuring and managing forest carbon stocks in California. A report for California's fourth climate change assessment. CCCA4-CNRA-2018-014. Sacramento, CA: California Natural Resources Agency. 44 p. https://www.energy.ca.gov/sites/default/files/2019-12/Forests_CCCA4-CNRA-2018-014_ada.pdf. (July 6, 2023).

Baumann, M.; Ozdogan, M.; Wolter, P.T.; Krylov, A.; Vladimirova, N.; Radeloff, V.C. 2014. Landsat remote sensing of forest windfall disturbance. Remote Sensing of Environment. 143: 171–179. https://doi.org/10.1016/j.rse.2013.12.020.

Bechtold, W.A.; Patterson, P.L.; eds. 2005. The enhanced Forest Inventory and Analysis Program—national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p. https://doi.org/10.2737/SRS-GTR-80.

Bell, D.M.; Acker, S.A.; Gregory, M.J.; Davis, R.J.; Garcia, B.A. 2021. Quantifying regional trends in large live tree and snag availability in support of forest management. Forest Ecology and Management. 479: 118554. https://doi.org/10.1016/j.foreco.2020.118554.

Bell, D.M.; Bradford, J.B.; Lauenroth, W.K. 2014. Early indicators of change: divergent climate envelopes between tree life stages imply range shifts in the Western United States. Global Ecology and Biogeography. 23(2): 168–180. https://doi.org/10.1111/geb.12109.

Bell, D.M.; Gray, A.N. 2016. Assessing intra- and inter-regional climate effects on Douglas-fir biomass dynamics in Oregon and Washington, USA. Forest Ecology and Management. 379: 281–287. https://doi.org/10.1016/j.foreco.2016.07.023.

Bell, D.M.; Gregory, M.J.; Churchill, D.J.; Smith, A.C. 2022a. Mapping with height and spectral remote sensing implies that environment and forest structure jointly constrain tree community composition in temperate coniferous forests of eastern Washington, United States. Frontiers in Forests and Global Change. 5. https://www.frontiersin.org/articles/10.3389/ffgc.2022.962816.

Bell, D.M.; Gregory, M.J.; Kane, V.; Kane, J.; Kennedy, R.E.; Roberts, H.M.; Yang, Z. 2018. Multiscale divergence between Landsat- and LiDAR-based biomass mapping is related to regional variation in canopy cover and composition. Carbon Balance and Management. 13(1): 15. https://doi.org/10.1186/s13021-018-0104-6.

Bell, D.M.; Gregory, M.J.; Ohmann, J.L. 2015a. Imputed forest structure uncertainty varies across elevational and longitudinal gradients in the western Cascade Mountains, Oregon, USA. Forest Ecology and Management. 358: 154–164. https://doi.org/10.1016/j.foreco.2015.09.007.

Bell, D.M.; Gregory, M.J.; Roberts, H.M.; Davis, R.J.; Ohmann, J.L. 2015b. How sampling and scale limit accuracy assessment of vegetation maps: a comment on Loehle et al. (2015). Forest Ecology and Management. 358: 361–364. https://doi.org/10.1016/j.foreco.2015.07.017.

Bell, D.M.; Wilson, B.T.; Werstak, C.E.; Oswalt, C.M.; Perry, C.H. 2022b. Examining *k*-nearest neighbor small area estimation across scales using national forest inventory data. Frontiers in Forests and Global Change. 5. https://www.frontiersin.org/articles/10.3389/ffgc.2022.763422.

Burrill, E.A.; DiTommaso, A.M.; Turner, J.A.; Pugh, S.A.; Christensen, G.; Perry, C.J.; Conkling, B. 2021. The Forest Inventory and Analysis database: database description and user guide for phase 2 (version 9.0.1). Washington, DC: U.S. Department of Agriculture, Forest Service. 1026 p. https://www.fia.fs.usda.gov/library/database-documentation/current/ver90/FIADB%20User%20Guide%20P2_9-0-1_final.pdf. (July 6, 2023).

Canham, C.D.; Thomas, R.Q. 2010. Frequency, not relative abundance, of temperate tree species varies along climate gradients in eastern North America. Ecology. 91(12): 3433–3440. https://doi.org/10.1890/10-0312.1.

Chirici, G.; Mura, M.; McInerney, D.; Py, N.; Tomppo, E.O.; Waser, L.T.; Travaglini, D.; McRoberts, R.E. 2016. A meta-analysis and review of the literature on the *k*-nearest neighbors technique for forestry applications that use remotely sensed data. Remote Sensing of Environment. 176: 282–294. https://doi.org/10.1016/j.rse.2016.02.001.

Clough, B.J.; Russell, M.B.; Domke, G.M.; Woodall, C.W. 2016. Quantifying allometric model uncertainty for plot-level live tree biomass stocks with a data-driven, hierarchical framework. Forest Ecology and Management. 372: 175–188. https://doi.org/10.1016/j.foreco.2016.04.001.

Cohen, W.B.; Spies, T.A. 1992. Estimating structural attributes of Douglas-fir/western hemlock forest stands from Landsat and SPOT imagery. Remote Sensing of Environment. 41(1): 1–17. https://doi.org/10.1016/0034-4257(92)90056-P.

Cohen, W.B.; Yang, Z.; Healey, S.P.; Kennedy, R.E.; Gorelick, N. 2018. A LandTrendr multispectral ensemble for forest disturbance detection. Remote Sensing of Environment. 205: 131–140. https://doi.org/10.1016/j.rse.2017.11.015.

Cohen, W.B.; Yang, Z.; Stehman, S.V.; Schroeder, T.A.; Bell, D.M.; Masek, J.G.; Huang, C.; Meigs, G.W. 2016. Forest disturbance across the conterminous United States from 1985–2012: the emerging dominance of forest decline. Forest Ecology and Management. 360: 242–252. https://doi.org/10.1016/j.foreco.2015.10.042.

Comer, P.J.; Faber-Langendoen, D.; Evans, R.; Gawler, S.C.; Josse, C.; Kittel, G.; Menard, S.; Pyne, M.; Reid, M.; Schulz, K.; Snow, K.; Teague, J. 2003. Ecological systems of the United States. Arlington, VA: NatureServe. 75 p. https://www.natureserve.org/sites/default/files/pcom_2003_ecol_systems_us.pdf. (July 6, 2023).

Crist, E.P.; Cicone, R.C. 1984. A physically-based transformation of thematic mapper data—the TM tasseled cap. IEEE Transactions on Geoscience and Remote Sensing. GE-22(3): 256–263. https://doi.org/10.1109/TGRS.1984.350619.

Daly, C.; Halbleib, M.; Smith, J.I.; Gibson, W.P.; Doggett, M.K.; Taylor, G.H.; Curtis, J.; Pasteris, P.P. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology. 28(15): 2031–2064. https://doi.org/10.1002/joc.1688.

Davis, R.J.; Bell, D.M.; Gregory, M.J.; Yang, Z.; Gray, A.N.; Healey, S.; Stratton, A.E. 2022. Northwest Forest Plan—the first 25 years (1994–2018): status and trends in late-successional and old-growth forests. Gen. Tech. Rep. PNW-GTR-1004. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 82 p. https://doi.org/10.2737/PNW-GTR-1004.

Davis, R.J.; Hollen, B.; Hobson, J.; Gower, J.E.; Keenum, D. 2016. Northwest Forest Plan—the first 20 years (1994–2013): status and trends of northern spotted owl habitats. Gen. Tech. Rep. PNW-GTR-929. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 54 p. https://doi.org/10.2737/PNW-GTR-929.

Davis, R.J.; Ohmann, J.L.; Kennedy, R.E.; Cohen, W.B.; Gregory, M.J.; Yang, Z.; Roberts, H.M.; Gray, A.N.; Spies, T.A. 2015. Northwest Forest Plan—the first 20 years (1994–2013): status and trends of late-successional and old-growth forests. Gen. Tech. Rep. PNW-GTR-911. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 112 p. https://doi.org/10.2737/PNW-GTR-911.

Falxa, G.A.; Raphael, M.G. 2016. Northwest Forest Plan—the first 20 years (1994–2013): status and trend of marbled murrelet populations and nesting habitat. Gen. Tech. Rep. PNW-GTR-933. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 132 p. https://doi.org/10.2737/PNW-GTR-933.

Foody, G.M.; Boyd, D.S.; Cutler, M.E.J. 2003. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. Remote Sensing of Environment. 85(4): 463–474. https://doi.org/10.1016/S0034-4257(03)00039-7.

Grossmann, E.B.; Kagan, J.S.; Ohmann, J.L.; May, H.K.; Gregory, M.J.; Tobalske, C. 2008. The Pacific Northwest Regional Gap Analysis Project final report on land cover mapping methods, MRLC mapzones 2 and 7. Corvallis, OR: Oregon State University, Institute for Natural Resources. https://ir.library.oregonstate.edu/concern/defaults/2801ph20s?locale=en. (July 6, 2023).

Hansen, A.J.; Neilson, R.P.; Dale, V.H.; Flather, C.H.; Iverson, L.R.; Currie, D.J.; Shafer, S.; Cook, R.; Bartlein, P.J. 2001. Global change in forests: responses of species, communities, and biomes. BioScience. 51(9): 765. https://doi.org/10.1641/0006-3568(2001)051[0765:GCIFRO]2.0.CO;2.

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.; Kommareddy, A.; Egorov, A.; Chini, L.; Justice, C.O.; Townshend, J.R.G. 2013. High-resolution global maps of 21st-century forest cover change. Science. 342(6160): 850–853. https://doi.org/10.1126/science.1244693.

He, K.S.; Bradley, B.A.; Cord, A.F.; Rocchini, D.; Tuanmu, M.N.; Schmidtlein, S.; Turner, W.; Wegmann, M.; Pettorelli, N. 2015. Will remote sensing shape the next generation of species distribution models? Remote Sensing in Ecology and Conservation. 1(1): 4–18. https://doi.org/10.1002/rse2.7.

Healey, S.P.; Cohen, W.B.; Yang, Z.; Kenneth Brewer, C.; Brooks, E.B.; Gorelick, N.; Hernandez, A.J.; Huang, C.; Joseph Hughes, M.; Kennedy, R.E.; Loveland, T.R.; Moisen, G.G.; Schroeder, T.A.; Stehman, S.V.; Vogelmann, J.E.; Woodcock, C.E.; Yang, L.; Zhu, Z. 2018. Mapping forest change using stacked generalization: an ensemble approach. Remote Sensing of Environment. 204: 717–728. https://doi.org/10.1016/j.rse.2017.09.029.

Henderson, E.B.; Ohmann, J.L.; Gregory, M.J.; Roberts, H.M.; Zald, H. 2014. Species distribution modelling for plant communities: stacked single species or multivariate modelling approaches? Applied Vegetation Science. 17(3): 516–527. https://doi.org/10.1111/avsc.12085.

Huang, K.; Yi, C.; Wu, D.; Zhou, T.; Zhao, X.; Blanford, W.J.; Wei, S.; Wu, H.; Ling, D.; Li, Z. 2015. Tipping point of a conifer forest ecosystem under severe drought. Environmental Research Letters. 10(2): 24011. https://doi.org/10.1088/1748-9326/10/2/024011.

Jenks, G.F. 1967. The data model concept in statistical mapping. International Yearbook of Cartography. 7: 186–190.

Johnson, D.H.; O'Neil, T.A. 2001. Wildlife-habitat relationships in Oregon and Washington. Corvallis, OR: Oregon State University Press. 736 p.

Kagan, J.; Ohmann, J.L.; Gregory, M.J.; Tobalske, C.; Hak, J.C.; Fried, J.S. 2006. The Pacific Northwest Regional Gap Analysis Project final report on land cover mapping methods, map zones 8 and 9. Corvallis, OR: Oregon State University, Institute for Natural Resources. https://lemma.forestry.oregonstate.edu/export/pubs/kagan_etal_2006_gap89_final_report.pdf. (July 6, 2023).

Kennedy, R.E.; Ohmann, J.; Gregory, M.; Roberts, H.; Yang, Z.; Bell, D.M.; Kane, V.; Hughes, M.J.; Cohen, W.; Powell, S.; Neeti, N.; Larrue, T.; Hooper, S.; Kane, J.T.; Miller, D.L.; Perkins, J.; Braaten, J.; Seidl, R. 2018a. An empirical, integrated forest biomass monitoring system. Environmental Research Letters. 13(2): 41001. https://doi.org/10.1088/1748-9326/aa9d9e.

Kennedy, R.E.; Yang, Z.; Cohen, W.B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—temporal segmentation algorithms. Remote Sensing of Environment. 114(12): 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008.

Kennedy, R.E.; Yang, Z.; Gorelick, N.; Braaten, J.; Cavalcante, L.; Cohen, W.B.; Healey, S. 2018b. Implementation of the LandTrendr algorithm on Google Earth Engine. Remote Sensing. 10(5): 691. https://doi.org/10.3390/rs10050691.

Key, C.H.; Benson, N.C. 2006. Landscape assessment (LA): sampling and analysis methods. In: Lutes, D.C.; Keane, R.E.; Caratti, J.F.; Key, C.H.; Benson, N.C.; Sutherland, S.; Gangi, L.J., eds. FIREMON: Fire effects monitoring and inventory system. Gen. Tech. Rep. RMRS-GTR-164-CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: 1–51. https://doi.org/10.2737/RMRS-GTR-164.

Latta, G.; Temesgen, H.; Adams, D.; Barrett, T. 2010. Analysis of potential impacts of climate change on forests of the United States Pacific Northwest. Forest Ecology and Management. 259(4): 720–729. https://doi.org/10.1016/j.foreco.2009.09.003.

Lister, A.J.; Andersen, H.; Frescino, T.; Gatziolis, D.; Healey, S.; Heath, L.S.; Liknes, G.C.; McRoberts, R.; Moisen, G.G.; Nelson, M.; Riemann, R.; Schleeweis, K.; Schroeder, T.A.; Westfall, J.; Wilson, B.T. 2020. Use of remote sensing data to improve the efficiency of national forest inventories: a case study from the United States national forest inventory. Forests. 11(12): 1364. https://doi.org/10.3390/f11121364.

Loehle, C.; Irwin, L.; Manly, B.F.J.; Merrill, A. 2015. Range-wide analysis of northern spotted owl nesting habitat relations. Forest Ecology and Management. 342: 8–20. https://doi.org/10.1016/j.foreco.2015.01.010.

Lorenz, T.J.; Raphael, M.G.; Young, R.D.; Lynch, D.; Nelson, S.K.; McIver, W.R. 2021. Status and trend of nesting habitat for the marbled murrelet under the Northwest Forest Plan, 1993 to 2017. Gen. Tech. Rep. PNW-GTR-998. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 64 p. https://www.fs.usda.gov/research/treesearch/63314. (July 6, 2023).

Lu, D.; Chen, Q.; Wang, G.; Moran, E.; Batistella, M.; Zhang, M.; Vaglio Laurin, G.; Saah, D. 2012. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. International Journal of Forestry Research. 2012: 1–16. https://doi.org/10.1155/2012/436537.

Martin, P.H.; Canham, C.D. 2020. Peaks in frequency, but not relative abundance, occur in the center of tree species distributions on climate gradients. Ecosphere. 11(6): e03149. https://doi.org/10.1002/ecs2.3149.

Masek, J.G.; Hayes, D.J.; Joseph Hughes, M.; Healey, S.P.; Turner, D.P. 2015. The role of remote sensing in process-scaling studies of managed forest ecosystems. Forest Ecology and Management. 355: 109–123. https://doi.org/10.1016/j.foreco.2015.05.032.

Masek, J.G.; Huang, C.; Wolfe, R.; Cohen, W.; Hall, F.; Kutler, J.; Nelson, P. 2008. North American forest disturbance mapped from a decadal Landsat record. Remote Sensing of Environment. 112(6): 2914–2926. https://doi.org/10.1016/j.rse.2008.02.010.

Matasci, G.; Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W.; Zald, H.S.J. 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and LiDAR plots. Remote Sensing of Environment. 209: 90–106. https://doi.org/10.1016/j.rse.2017.12.020.

McRoberts, R.E. 2001. Imputation and model-based updating techniques for annual forest inventories. Forest Science. 47(3): 322–330.

McRoberts, R.E. 2012. Estimating forest attribute parameters for small areas using nearest neighbors techniques. Forest Ecology and Management. 272: 3–12. https://doi.org/10.1016/j.foreco.2011.06.039.

McRoberts, R.E.; Tomppo, E.O.; Finley, A.O.; Heikkinen, J. 2007. Estimating areal means and variances of forest attributes using the *k*-nearest neighbors technique and satellite imagery. Remote Sensing of Environment. 111(4): 466–480. https://doi.org/10.1016/j.rse.2007.04.002.

Mellen-McLean, K.; Marcot, B.G.; Ohmann, J.L.; Waddell, K.; Willhite, E.A.; Acker, S.A.; Livingston, S.A.; Hostetler, B.B.; Webb, B.S.; Garcia, B.A. 2017. DecAID: the decayed wood advisor for managing snags, partially dead trees, and down wood for biodiversity in forests of Washington and Oregon. https://apps.fs.usda.gov/r6_decaid/views/index.html. (April 21, 2023).

Moeur, M.; Ohmann, J.L.; Kennedy, R.E.; Cohen, W.B.; Gregory, M.J.; Yang, Z.; Roberts, H.M.; Spies, T.A.; Fiorella, M. 2011. Status and trend of late-successional and old-growth forests. Gen. Tech. Rep. PNW-GTR-853. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 48 p.

Moeur, M.; Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. Forest Science. 41(2): 337–359.

Moisen, G.G.; Meyer, M.C.; Schroeder, T.A.; Liao, X.; Schleeweis, K.G.; Freeman, E.A.; Toney, C. 2016. Shape selection in Landsat time series: a tool for monitoring forest dynamics. Global Change Biology. 22(10): 3518–3528. https://doi.org/10.1111/gcb.13358.

Monleon, V.J.; Lintz, H.E. 2015. Evidence of tree species' range shifts in a complex landscape. PloS One. 10(1): e0118069. https://doi.org/10.1371/journal.pone.0118069.

Myroniuk, V.; Bell, D.M.; Gregory, M.J.; Vasylyshyn, R.; Bilous, A. 2022. Uncovering forest dynamics using historical forest inventory data and Landsat time series. Forest Ecology and Management. 513: 120184. https://doi.org/10.1016/j.foreco.2022.120184.

Nguyen, T.H.; Jones, S.; Soto-Berelov, M.; Haywood, A.; Hislop, S. 2020. Landsat time-series for estimating forest aboveground biomass and its dynamics across space and time: a review. Remote Sensing. 12(1): 98. https://doi.org/10.3390/rs12010098.

Nie, M. 2019. The Forest Service's 2012 planning rule and its implementation: federal advisory committee member perspectives. Journal of Forestry. 117(1): 65–71. https://doi.org/10.1093/jofore/fvy055.

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(4): 725–741. https://doi.org/10.1139/x02-011.

Ohmann, J.L.; Gregory, M.J.; Roberts, H.M. 2014. Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. Remote Sensing of Environment. 151: 3–15. https://doi.org/10.1016/j.rse.2013.08.048.

Ohmann, J.L.; Gregory, M.J.; Roberts, H.M.; Cohen, W.B.; Kennedy, R.E.; Yang, Z. 2012. Mapping change of older forest with nearest-neighbor imputation and Landsat time-series. Forest Ecology and Management. 272: 13–25. https://doi.org/10.1016/j.foreco.2011.09.021.

Ohmann, J.L.; Spies, T.A. 1998. Regional gradient analysis and spatial pattern of woody plant communities of Oregon forests. Ecological Monographs. 68(2): 151–182. https://doi.org/10.1890/0012-9615(1998)068[0151:RGAASP]2.0.CO;2.

Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. 2013. Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment. 129: 122–131. https://doi.org/10.1016/j.rse.2012.10.031.

Pan, Y.; Birdsey, R.A.; Phillips, O.L.; Jackson, R.B. 2013. The structure, distribution, and biomass of the world's forests. Annual Review of Ecology, Evolution, and Systematics. 44(1): 593–622. https://doi.org/10.1146/annurev-ecolsys-110512-135914.

Pierce, K.B.; Ohmann, J.L.; Wimberly, M.C.; Gregory, M.J.; Fried, J.S. 2009. Mapping wildland fuels and forest structure for land management: a comparison of nearest neighbor imputation and other methods. Canadian Journal of Forest Research. 39(10): 1901–1916. https://doi.org/10.1139/X09-102.

Powell, S.L.; Cohen, W.B.; Healey, S.P.; Kennedy, R.E.; Moisen, G.G.; Pierce, K.B.; Ohmann, J.L. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches. Remote Sensing of Environment. 114(5): 1053–1068. https://doi.org/10.1016/j.rse.2009.12.018.

Rehfeldt, G.E.; Crookston, N.L.; Warwell, M.V.; Evans, J.S. 2006. Empirical analyses of plant-climate relationships for the Western United States. International Journal of Plant Sciences. 167(6): 1123–1150. https://doi.org/10.1086/507711.

Reilly, M.J.; Zuspan, A.; Halofsky, J.S.; Raymond, C.; McEvoy, A.; Dye, A.W.; Donato, D.C.; Kim, J.B.; Potter, B.E.; Walker, N.; Davis, R.J.; Dunn, C.J.; Bell, D.M.; Gregory, M.J.; Johnston, J.D.; Harvey, B.J.; Halofsky, J.E.; Kerns, B.K. 2022. Cascadia burning: the historic, but not historically unprecedented, 2020 wildfires in the Pacific Northwest, USA. Ecosphere. 13(6): e4070. https://doi.org/10.1002/ecs2.4070.

Riemann, R.; Wilson, B.T.; Lister, A.; 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(10): 2337–2352. https://doi.org/10.1016/j.rse.2010.05.010.

Robinson, N.P.; Allred, B.W.; Smith, W.K.; Jones, M.O.; Moreno, A.; Erickson, T.A.; Naugle, D.E.; Running, S.W. 2018. Terrestrial primary production for the conterminous United States derived from Landsat 30 m and MODIS 250 m. Remote Sensing in Ecology and Conservation. 4(3): 264–280. https://doi.org/10.1002/rse2.74.

Rollins, M.G. 2009. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. International Journal of Wildland Fire. 18(3): 235–249. https://doi.org/10.1071/WF08088.

Roy, D.P.; Kovalskyy, V.; Zhang, H.K.; Vermote, E.F.; Yan, L.; Kumar, S.S.; Egorov, A. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. Remote Sensing of Environment. 185: 57–70. https://doi.org/10.1016/j.rse.2015.12.024.

Ryan, C.M.; Cerveny, L.K.; Robinson, T.L.; Blahna, D.J. 2018. Implementing the 2012 forest planning rule: best available scientific information in forest planning assessments. Forest Science. 64(2): 159–169. https://doi.org/10.1093/forsci/fxx004.

Schleeweis, K.; Goward, S.N.; Huang, C.; Dwyer, J.L.; Dungan, J.L.; Lindsey, M.A.; Michaelis, A.; Rishmawi, K.; Masek, J.G. 2016. Selection and quality assessment of Landsat data for the North American forest dynamics forest history maps of the U.S. International Journal of Digital Earth. 9(10): 963–980. https://doi.org/10.1080/17538947.2016.1158876.

Schleeweis, K.G.; Moisen, G.G.; Schroeder, T.A.; Toney, C.; Freeman, E.A.; Goward, S.N.; Huang, C.; Dungan, J.L. 2020. U.S. national maps attributing forest change: 1986–2010. Forests. 11: 653. https://doi.org/10.3390/f11060653.

Schroeder, T.A.; Schleeweis, K.G.; Moisen, G.G.; Toney, C.; Cohen, W.B.; Freeman, E.A.; Yang, Z.; Huang, C. 2017. Testing a Landsat-based approach for mapping disturbance causality in U.S. forests. Remote Sensing of Environment. 195: 230–243. https://doi.org/10.1016/j.rse.2017.03.033.

Smith, J.E.; Domke, G.M.; Nichols, M.C.; Walters, B.F. 2019. Carbon stocks and stock change on federal forest lands of the United States. Ecosphere. 10(3): e02637. https://doi.org/10.1002/ecs2.2637.

Strunk, J.L.; Bell, D.M.; Gregory, M.J. 2022. Pushbroom photogrammetric heights enhance state-level forest attribute mapping with Landsat and environmental gradients. Remote Sensing. 14(14): 3433. https://doi.org/10.3390/rs14143433.

Strunk, J.L.; Gould, P.J.; Packalen, P.; Gatziolis, D.; Greblowska, D.; Maki, C.; McGaughey, R.J. 2020. Evaluation of pushbroom DAP relative to frame camera DAP and LiDAR for forest modeling. Remote Sensing of Environment. 237: 111535. https://doi.org/10.1016/j.rse.2019.111535.

ter Braak, C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology. 67(5): 1167–1179. https://doi.org/10.2307/1938672.

Vogeler, J.C.; Braaten, J.D.; Slesak, R.A.; Falkowski, M.J. 2018. Extracting the full value of the Landsat archive: inter-sensor harmonization for the mapping of Minnesota forest canopy cover (1973–2015). Remote Sensing of Environment. 209: 363–374. https://doi.org/10.1016/j.rse.2018.02.046.

Wickham, J.; Stehman, S.V.; Sorenson, D.G.; Gass, L.; Dewitz, J.A. 2021. Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States. Remote Sensing of Environment. 257: 112357. https://doi.org/10.1016/j.rse.2021.112357.

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. https://doi.org/10.1016/j.foreco.2012.02.002.

Wilson, B.T.; Woodall, C.W.; Griffith, D.M. 2013. Imputing forest carbon stock estimates from inventory plots to a nationally continuous coverage. Carbon Balance and Management. 8: 1–15. https://doi.org/10.1186/1750-0680-8-1.

Wulder, M.A.; Roy, D.P.; Radeloff, V.C.; Loveland, T.R.; Anderson, M.C.; Johnson, D.M.; Healey, S.; Zhu, Z.; Scambos, T.A.; Pahlevan, N.; Hansen, M.; Gorelick, N.; Crawford, C.J.; Masek, J.G.; Hermosilla, T.; White, J.C.; Belward, A.S.; Schaaf, C.; Woodcock, C.E.; Huntington, J.L.; Lymburner, L.; Hostert, P.; Gao, F.; Lyapustin, A.; Pekel, J.-F.; Strobl, P.; Cook, B.D. 2022. Fifty years of Landsat science and impacts. Remote Sensing of Environment. 280: 113195. https://doi.org/10.1016/j.rse.2022.113195.

Zald, H.S.J.; Ohmann, J.L.; Roberts, H.M.; Gregory, M.J.; Henderson, E.B.; McGaughey, R.J.; Braaten, J. 2014. Influence of LiDAR, Landsat imagery, disturbance history, plot location accuracy, and plot size on accuracy of imputation maps of forest composition and structure. Remote Sensing of Environment. 143: 26–38. https://doi.org/10.1016/j.rse.2013.12.013.

Zhu, Z. 2017. Change detection using Landsat time series: a review of frequencies, preprocessing, algorithms, and applications. ISPRS Journal of Photogrammetry and Remote Sensing. 130: 370–384. https://doi.org/10.1016/j.isprsjprs.2017.06.013.

Zhu, Z.; Zhang, J.; Yang, Z.; Aljaddani, A.H.; Cohen, W.B.; Qiu, S.; Zhou, C. 2020. Continuous monitoring of land disturbance based on Landsat time series. Remote Sensing of Environment. 238: 111116. https://doi.org/10.1016/j.rse.2019.03.009.

Zimmermann, N.E.; Edwards, T.C.; Moisen, G.G.; Frescino, T.S.; Blackard, J.A. 2007. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. Journal of Applied Ecology. 44(5): 1057–1067. https://doi.org/10.1111/j.1365-2664.2007.01348.x.

Pacific Northwest Research Station

Website	https://www.fs.usda.gov/research/pnw
Telephone	(503) 808-2100
Publication requests	(503) 808-2138
FAX	(503) 808-2130
E-mail	sm.fs.pnw_pnwpubs@usda.gov
Mailing address	Publications Distribution Pacific Northwest Research Station USDA Forest Service 1220 SW 3rd Avenue Portland, OR 97204

U.S. Department of Agriculture
Forest Service
Pacific Northwest Research Station
1220 SW 3rd Avenue, Suite 1400
Portland, OR 97204

Official Business
Penalty for Private Use, $300