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 (2014) 26–38

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Influence of lidar, Landsat imagery, disturbance history, plot location accuracy, and plot size on accuracy of imputation maps of forest composition and structure

Harold S.J. Zald [a,*], Janet L. Ohmann [a], Heather M. Roberts [a], Matthew J. Gregory [a], Emilie B. Henderson [b], Robert J. McGaughey [c], Justin Braaten [a]

[a] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA
[b] Institute for Natural Resources, Oregon State University, Portland, OR 97207, USA
[c] USDA Forest Service, PNW Research Station, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Received 17 July 2013
Received in revised form 18 December 2013
Accepted 19 December 2013
Available online 19 January 2014

Keywords:
Lidar
Landsat time series
Nearest-neighbor imputation
Disturbance
Forest composition and structure

ABSTRACT

This study investigated how lidar-derived vegetation indices, disturbance history from Landsat time series (LTS) imagery, plot location accuracy, and plot size influenced accuracy of statistical spatial models (nearest-neighbor imputation maps) of forest vegetation composition and structure. Nearest-neighbor (NN) imputation maps were developed for 539,000 ha in the central Oregon Cascades, USA. Mapped explanatory data included tasseled-cap indices and disturbance history metrics (year, magnitude, and duration of disturbance) from LTS imagery, lidar-derived vegetation metrics, climate, topography, and soil parent material. Vegetation data from USDA Forest Service forest inventory plots was summarized at two plot sizes (plot and subplot) and geographically located with two levels of accuracy (standard and improved). Maps of vegetation composition and structure were developed with the Gradient Nearest Neighbor (GNN) method of NN imputation using different combinations of explanatory variables, plot spatial resolution, and plot positional accuracy. Lidar vegetation indices greatly improved predictions of live tree structure, moderately improved predictions of snag density and down wood volume, but did not consistently improve species predictions. LTS disturbance metrics improved predictions of forest structure, but not to the degree of lidar indices, while also improving predictions of many species. Absence of disturbance attribution (i.e. disturbance type such as fire or timber harvest) in LTS disturbance metrics may have limited our ability to predict forest structure. Absence of corrected lidar intensity values may also have lowered accuracy of snag and species predictions. However, LTS disturbance attribution and lidar corrected intensity values may not be able to overcome fundamental limitations of remote sensing for predicting snags and down wood that are obscured by the forest canopy. Improved GPS plot locations had little influence on map accuracy, and we suggest under what conditions improved GPS plot locations may or may not improve the accuracy of predictive maps that link remote sensing with forest inventory plots. Subplot NN imputation maps had much lower accuracy compared to maps generated using response variables from larger whole plots. No single map had optimal results for every mapped variable, suggesting map users and developers need to prioritize what forest vegetation attributes are most important for any given map application.

1. Introduction

Forest management and conservation have grown increasingly complex, involving consideration of a wide array of ecological, economic, and societal values. Issues such as old growth conservation, wildlife habitat management, timber extraction, forest restoration, fuel reduction, and wildfire risk assessment often involve multiple interacting objectives, values, and threats (e.g. climate change, wildfires, and insect outbreaks) spanning broad spatial scales and long ecological gradients. In this complex policy and decision-making environment, quantitative information is required about forest vegetation conditions over large landscapes that is highly detailed with respect to multiple vegetation attributes, and spatially complete (i.e. mapped) (Spies et al., 2007).

Remotely sensed data are ideally suited to meet the need for spatially complete data about forests over large landscapes. Regional maps of forest cover are often based on multispectral satellite imagery (Cohen, Maiersperger, Spies, & Oetter, 2001; Hansen et al., 2003; Woodcock et al., 1994). However, maps from satellite imagery alone cannot provide

* Corresponding author at: 321 Richardson Hall, Oregon State University, Corvallis, OR 97331, USA. Tel.: +1 541 758 7759; fax: +1 541 737 1393.
E-mail addresses: harold.zald@oregonstate.edu (H.S.J. Zald), janet.ohmann@oregonstate.edu (J.L. Ohmann), heather.roberts@oregonstate.edu (H.M. Roberts), matt.gregory@oregonstate.edu (M.J. Gregory), emilie.henderson@oregonstate.edu (E.B. Henderson), bmcgaughey@fs.fed.us (R.J. McGaughey), justin.braaten@oregonstate.edu (J. Braaten).

0034-4257/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.12.013

the level of detail about forest composition and structure often required for many forest management and research applications. Measurements on field plot inventories often contain highly detailed ecological data, but only at sampled locations so they lack complete spatial coverage. As such, there is considerable interest in integrating field plots with remotely sensed data to generate maps with the spatial coverage of remotely sensed imagery and the ecological detail of field plots (Ohmann & Gregory, 2002; Tomppo, 1991; Tomppo, Goulding, & Katila, 1999).

One approach to integrating field plots and remotely sensed data is nearest-neighbor (NN) imputation, which has been widely used in forest inventory, monitoring, decision-support, and ecological research (Gjertsen, 2007; Moeur et al., 2011; Ohmann et al., 2012; Pierce, Ohmann, Wimberly, Gregory, & Fried, 2009; Reese et al., 2003; Spies et al., 2007; Tomppo et al., 2008; Wilson, Lister, & Riemann, 2012). Imputation is a method for filling in missing data by substituting values from donor observations (Eskelson et al., 2009). In forestry applications, imputation is used to estimate forest characteristics for large areas where a set of mapped explanatory variables are available for the entire spatial extent and these variables are related to a more detailed set of response variables only available for a limited sample of the study area. Response variables are usually measures of forest composition or structure derived from a sample of field plots, while mapped explanatory variables can include multispectral satellite imagery and other spatially complete datasets (i.e. climate, topography, etc.). In NN imputation, either a single donor observation (plot) can be chosen to fill in a given missing observation [$k = 1$], or multiple donor observations can be averaged to fill in a given missing observation [$k > 1$]. A major strength of NN imputation where $k = 1$ is the retention of the co-variance structure of multiple response variables, because each prediction links to a set of response values within a single plot.

As noted above, NN imputation mapping often relies on satellite imagery as mapped explanatory data (Ohmann & Gregory, 2002; Wilson et al., 2012). In particular, Landsat imagery (individual spectral bands and/or vegetation indices) is attractive for regional forest mapping due to its low cost, global coverage, long temporal record, and large scene-sizes, as well as spectral and spatial resolutions compatible with characterizing vegetation attributes (Cohen & Goward, 2004). However, Landsat and other passive optical sensors have limited sensitivity to vertical and below-canopy vegetation structure (Lu, 2006), and the information content in Landsat imagery is known to saturate in forests with high leaf area indices (Turner, Cohen, Kennedy, Fassnacht, & Briggs, 1999). These limitations of Landsat and other passive optical sensors pose problems for NN imputation mapping of forest attributes such as stand density, snags, and down wood (Eskelson, Temesgen, & Hagar, 2012; Pierce et al., 2009), which are important for carbon inventory and assessment, wildland fuels, and wildlife habitat.

Compared to Landsat and other passive optical sensors, Light Detection and Ranging (lidar) data can better represent the three-dimensional structure of forest canopies, and has been widely used to characterize vegetation cover and structure (see reviews by Dubayah & Drake, 2000; Lefsky, Cohen, Parker, & Harding, 2002; Reutebuch, Andersen, & McGaughey, 2005). Additionally, lidar does not suffer as much as Landsat imagery from declines in sensitivity and accuracy in forests with high leaf area index. The cost of lidar acquisition has declined dramatically over the past decade, such that lidar is increasingly available for large landscapes. Lidar has the potential to greatly improve NN imputation maps of forest structural attributes compared to maps developed using Landsat imagery or other passive optical sensors. Recent studies using lidar have had promising results at moderate spatial resolutions ($\leq$30 m pixels) and relatively small spatial extents (<60,000 ha); predicting presence/absence of snags and understory attributes (Martinuzzi et al., 2009), and imputation mapping of live tree structural attributes (Falkowski et al., 2010; Hudak, Crookston, Evans, Hall, & Falkowski, 2008). However, no published studies have determined if lidar can improve regional NN imputation mapping of forest attributes such as snag and down wood abundance, or species composition.

In addition to lidar data, advances utilizing the Landsat time series (LTS) may also improve accuracy of NN imputation maps. With the recent opening of the Landsat archive (Woodcock et al., 2008), there has been a proliferation of research in multi-temporal change detection and disturbance mapping (Huang et al., 2010; Kennedy, Yang, & Cohen, 2010; Masek et al., 2008). LTS disturbance metrics (such as time since, magnitude of, and duration of disturbance) may improve the accuracy of NN imputation indirectly, since many trends in forest composition and structure are closely related to disturbance history (Franklin et al., 2002; Oliver, 1980; Spies, 1991). This contrasts with lidar's direct characterization of forest structure. LTS disturbance metrics have been shown to have comparable predictive power to lidar for live basal area and aboveground biomass, and superior predictive power for dead basal area and aboveground biomass (Pflugmacher, Cohen, & Kennedy, 2012), suggesting many accuracy improvements that lidar can bring to imputation mapping might also be reached using LTS disturbance metrics. LTS disturbance metrics also have the advantage of complete spatial coverage and dramatically lower costs compared to lidar. LTS disturbance metrics have been used in NN imputation mapping of forests (Ohmann & Gregory, 2002), but no published studies have determined if LTS disturbance metrics can obtain predictions of comparable accuracy to lidar within the context of regional multivariate NN imputation maps of forest composition and structure. An additional advantage of LTS imagery for NN imputation mapping is it permits pixel-level normalization of multi-date images (Kennedy et al., 2010). This is an important consideration when minimization of year-to-year spectral variability and seamless multi-scene image mosaics are desired to relate to plot data collected over multiple years and across large spatial extents (Ohmann, Gregory, & Roberts, 2013).

For NN imputation and other methods linking field plots to remotely sensed data, accuracy of plot locations and plot size are important considerations. Studies relating lidar data to forest structure often do so using field plots geo-referenced using GPS receivers that manufacturers market as being capable of sub-meter accuracy when used under ideal conditions (Falkowski et al., 2010; Hudak et al., 2008; Kane et al., 2010; Pflugmacher et al., 2012). Although users and receiver manufacturers refer to GPS "accuracy", it is important to note that the "accuracy" statistic reported by GPS post-processing software is really "the precision of the solution" (i.e. a modeled estimate of geographic position), and computed GPS positions can still deviate from true geographic positions even when very high precision (aka accuracy) is reported. The accuracy of plot locations can be evaluated by comparing the GPS results with the "true" location obtained using high-order survey methods, but such comparisons are rarely made.

Unlike studies where research plots are geo-referenced using high precision GPS receivers, regional NN imputation mapping typically relies on existing plot networks such as the USDA Forest Service's Forest Inventory and Analysis Program (FIA) (Ohmann et al., 2012; Wilson et al., 2012). A variety of methods (i.e. recreational grade GPS receivers, map interpretation, and photo interpretation) have been used to determine the geographic locations of FIA plots. FIA plots located with recreational grade GPS receivers (the most common method used for locating FIA plots) have positional accuracy averaging 5–20 m, but some plots can have positional errors exceeding 20 m (Cooke, 2000; Hoppus & Lister, 2005). Additionally, FIA plots are comprised of multiple fixed-radius subplots, and NN imputation can be conducted using either larger plots (i.e. aggregates of subplots) or smaller individual subplots (McRoberts, 2009). Simulations suggest accuracy of plot locations and plot size can strongly impact the accuracy of lidar-derived estimates of forest biomass (Frazer, Magnussen, Wulder, & Niemann, 2011) and Landsat-derived estimates of forest area (McRoberts, 2010), but the impacts of plot location accuracy and plot size (in this study referring to whole plots versus individual subplots) on prediction accuracy have not been examined within the

context of NN imputation mapping over a large spatial extent using lidar data or LTS disturbance metrics.

The primary objective of this study was to determine if lidar and LTS disturbance metrics can improve the accuracy of regional NN imputation maps of forest vegetation composition and structure. An additional objective of this study was to determine if plot location accuracy and plot size influenced prediction accuracy of NN imputation maps.

2. Methods

2.1. Study area

The study was conducted on 539,300 ha in the central Oregon Cascades, USA (Fig. 1). Forest land within the study area is administered by the Deschutes National Forest. NN imputation maps only apply to the 89% of the study area with at least 10% tree cover. The study spans large climatic, physiographic, geologic, and vegetation gradients. Climate varies from maritime in the west with wet cool winters and cool dry summers; to more continental in the east with more variable temperatures, shorter frost-free seasons, and greatly reduced precipitation. Precipitation increases and temperature decreases with increasing elevation. The physiography of the study area includes glaciated plateaus, buttes, volcanic cones, mountains, and canyons. Geologically, the study area contains lava flows, ash deposits, breccia, pyroclastic deposits, glacial deposits, and lacustrine and fluvial sediments. Soils types are primarily andisols, inceptisols, and spodisols. Coniferous tree species dominate the study area, with species composition generally associated with large-scale climatic and topographic gradients. Forest zones within the study area range from high elevation timberline to low elevation juniper (*Juniperus occidentalis*) woodlands. See Franklin and Dyrness (1988), Orr, Orr, and Baldwin (1992), USDA-SCS (1993), and USDA-NRCS (1995) for more detailed descriptions of the region.

Fire is the dominant natural disturbance in the study area. Historic mean fire return intervals range from as low as 10 years in some juniper woodlands and dry ponderosa pine forests, to over 500 years in subalpine forests at high elevations (Agee, 1993; Agee, 1994). Fire severity is also highly variable within the study area, with higher severity fires generally associated with higher elevations, higher fuel loadings, and longer fire return intervals (Agee, 1993). Wildfire suppression and changing climatic conditions have altered the fire regime in the region over the past century (Everett, Schellhaas, Keenum, Spurbeck, & Ohlson, 2000; Heyerdahl, Brubaker, & Agee, 2001; Littell, McKenzie, Peterson, & Westerling, 2009; Wright & Agee, 2004). Despite wildfire suppression efforts, several large fires have occurred within the study area in the past decade. Timber harvesting has also been an important disturbance agent in the study area over the past half century, with harvest strategies shifting from clearcutting before 1970, transitioning to overstory thinning from 1970 to 1990, and an emphasis on understory thinning and fuel reduction since approximately 1990. Finally, outbreaks of insects such as the mountain pine beetle (*Dendroctonus ponderosae*) and western spruce budworm (*Choristoneura occidentalis*) have resulted in tree mortality in the region over the past three decades (Meigs, Kennedy, & Cohen, 2011).

2.2. Forest composition and structure response data from field plots

Forest composition and structure data were collected between 2004 and 2009 on 251 plots in two forest inventories: the USDA Forest Service's Forest Inventory and Analysis Program (FIA) and Region 6 Current Vegetation Survey (CVS). Although collected data, plot size, number of subplots, and layout of subplots differ slightly between FIA and CVS plots (Fig. 2, Supplemental Fig. S1), plot design and measurement protocols are comparable between the two inventory programs (Max et al., 1996; USFS, 2003). From the 251 candidate plots, 19 were excluded because they had been disturbed since last measured, leaving

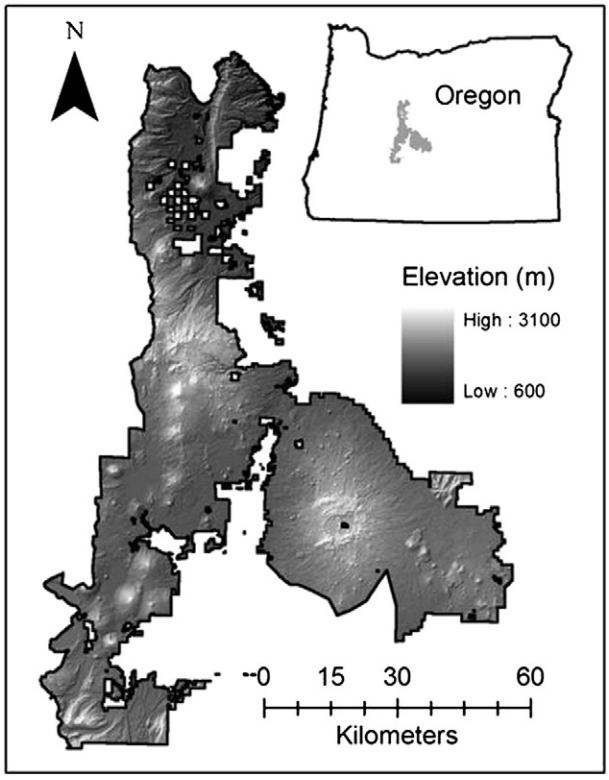

Fig. 1. Study area in the central Oregon Cascades, USA.

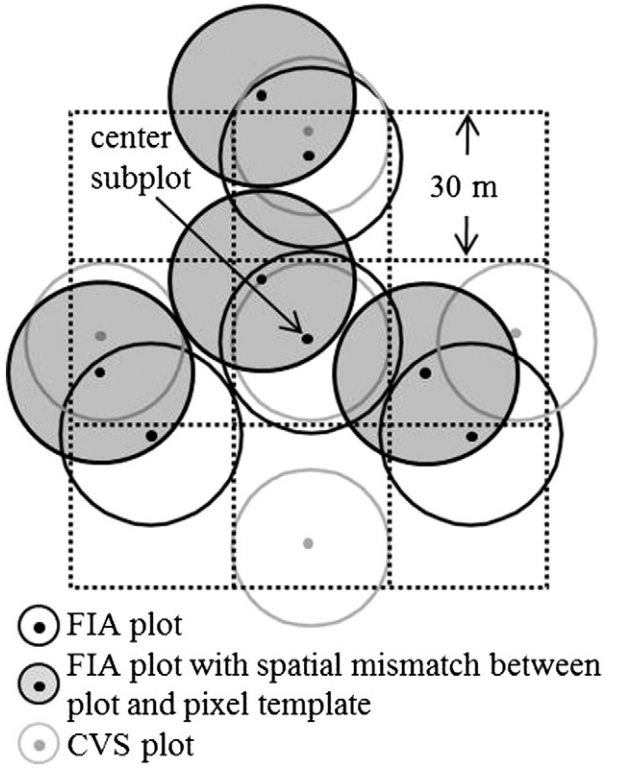

Fig. 2. Diagram showing plot design of FIA (black outline) and CVS (gray outline) plots. Dashed grid overlain on plot diagram depicts a 3 × 3 window of 30 m pixels used as a template to relate mapped explanatory variables to plots. The offset FIA plot (black outline with gray fill) demonstrates an example of a spatial mismatch between a plot and the template. For additional details on plot design and sampling, see Supplemental Fig. S1.

232 plots for modeling. These modeling plots had not been disturbed since field assessment, the majority (73%) had been assessed since 2007, and plots assessed between 2004 and 2006 were predominantly in mature forests, all factors that minimized potential impacts of temporal mismatches between plots and remotely sensed data. Twenty four response variables were calculated from sampled plots; percent canopy cover for the 12 most prevalent tree species (found on at least five plots), and 12 variables for various live tree, snag, and down wood structural attributes (Table 1).

Two sets of plot geographic locations (standard locations and improved GPS locations) were used to relate plots and subplots to mapped explanatory variables. Standard locations of plot centers were collected at time of plot assessment using a variety of methods with varying accuracy (recreational grade GPS, interpretation of digitized maps or aerial photographs). Improved GPS locations of plot centers were collected during 2010–2011 with a Trimble GeoXT GPS receiver, logging an average of 418 points (range of 146–2250 points) for each plot location. Improved GPS locations were differentially corrected and had an average horizontal precision of solution of ±0.91 m. Euclidean horizontal distance between standard and improved plot locations averaged 18.18 m (14.33–22.02 m 95% CI), but 11, six, and three plots had distances greater than 30, 60, and 90 m, respectively. See the Supplemental Figure S2 for additional information on differences between standard and improved GPS plot locations.

To examine the role of plot size in NN imputation accuracy, the 24 response variables were calculated at both the plot scale (aggregate of all subplots) and subplot scale (only the center subplot of each plot). Using only center subplots maintained the same sample number as whole plots for model comparisons, while avoiding fine-scale spatial autocorrelation in response and explanatory variables that could confound analyses if all subplots within a plot were treated as independent sample units. Species percent cover variables were arcsine square root transformed and structural variables were standardized by their ranges. A dummy variable with a value equal to the smallest value for any transformed response variable on a plot or center subplot was added to allow inclusion of plots on which no live trees, snags, or down wood were tallied.

2.3. Mapped explanatory data

Mapped explanatory variables (Table 2) included climate variables (1981–2010 normals) derived from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM, Daly et al., 2008), topographic and solar radiation variables calculated from lidar-derived digital elevation models (DEMs), regional geologic information, geographic coordinates (easting and northing), lidar-derived vegetation metrics, as well as tasseled cap indices (Crist & Cicone, 1984; Kauth & Thomas, 1976) and disturbance metrics derived from LTS mosaics developed using the LandTrendr algorithms (Kennedy et al., 2010). All mapped explanatory data were generated at or resampled to 30 m pixel resolution.

Airborne discrete return lidar data was collected by Watershed Sciences, Inc. (Corvallis, Oregon USA) during the fall and summer of 2009 and 2010. Lidar was collected from 900 to 1300 m above ground level by fixed wing aircraft equipped with Leica ALS50 Phase II and ALS60 laser scanners. Laser scanners had a 105 kHz pulse rate, scan angle of ±14° from nadir, and 50% scan swath overlap. Average pulse return density exceeded 8/m^2. A 1 m resolution DEM was developed from the lidar point cloud by Watershed Sciences Inc. using TerraScan software (Soininen, 2004) following methods previously described in the electronic supplemental material of Zald, Spies, Huso, and Gatziolis (2012). The DEM was resampled to 30 m, from which topographic and solar radiation explanatory variables were derived. Metrics of vegetation structure were derived from the lidar point cloud using the grid metrics function in FUSION/LDV (McGaughey, 2013). The grid metrics function generates over 50 gridded variables characterizing vegetation height, cover, vertical distribution, and intensity. The primary purpose of this study was to assess the predictive accuracy of NN imputation using lidar, not to assess all the possible combinations of lidar

Table 1
Codes and descriptions of species and structural response variables, along with species prevalence, structural mean and range, by plots and subplots.

Code	Species	Prevalence (percent)	
		Plot	Subplot
ABAM	*Abies amabilis*	2.2	1.7
ABGRC	*Abies grandis* and *Abies concolor**	23.7	14.2
ABPRSH	*Abies procera* and *Abies magnifica var. shastensisnn**	3.4	2.2
JUOC	*Juniperus occidentalis*	4.3	2.2
PIAL	*Pinus albicaulis*	5.2	1.7
PICO	*Pinus contorta*	61.6	48.7
PIEN	*Picea engelmannii*	2.6	2.2
PILA	*Pinus lambertiana*	5.2	2.2
PIMO3	*Pinus monticola*	10.3	4.3
PIPO	*Pinus ponderosa*	70.7	56.5
PSME	*Pseudotsuga menziesii*	12.9	8.6
TSME	*Tsuga mertensiana*	15.1	11.6
		Mean (range)	
Code	Structural variable	Plot	Subplot
BA	Basal area (m3/ha) of all live trees ≥2.5 cm dbh	21.65 (0–76.19)	21.35 (0–82.24)
QMD	Quadratic mean diameter (cm) of dominant and codominant trees	30.71 (0–87.38)	34.11 (0–124.97)
HCB	Height (m) to crown base	4.56 (0–29.47)	4.36 (0–23.77)
TPH_3_25	Density (trees/ha) of live trees 2.5–25 cm dbh (trees/ha)	926.96 (0–7224.44)	785.17 (0–6806.04)
TPH_25_50	Density (trees/ha) of live trees 25–50 cm dbh	86.61 (0–409.22)	89.26 (0–416.38)
TPH_GE_50	Density (trees/ha) of live trees ≥50 cm dbh	16.62 (0–127.97)	16.74 (0–188.1)
STPH_12_25	Density (snags/ha) of snags 2.5–25 cm dbh	37.44 (0–470.34)	31.28 (0–475.87)
STPH_25_50	Density (snags/ha) of snags 25–50 cm dbh	18.97 (0–227.59)	21.74 (0–237.93)
STPH_GE_50	Density (snags/ha) of snags ≥50 cm dbh	2.66 (0–59.43)	3.62 (0–138.72)
DVPH_12_25	Volume (m3/ha) of down wood 12.5–25 cm diameter at large end and ≥3 m long	14.47 (0–109.48)	14.07 (0–309.2)
DVPH_25_50	Volume (m3/ha) of down wood 25–50 cm diameter at large end and ≥3 m long	29.43 (0–286.54)	32.03 (0–681.15)
DVPH_GE_50	Volume (m3/ha) of down wood 25–50 cm diameter at large end and ≥3 m long	19.18 (0–489.43)	20.25 (0–962.22)

Note: Species followed by * are lumped together due to difficulties distinguishing these species in the field.

Table 2
Descriptions of mapped explanatory variables used in GNN models.

Variable group	Variable description
Climate	Mean annual precipitation (natural logarithm, mm) Mean annual temperature (°C)
	Mean maximum temperature of hottest month (August) (°C)
	Percentage of annual precipitation falling during the growing season (June - August)
	Mean minimum temperature of coldest month (December) (°C)
	Growing season moisture stress, the ratio of mean temperature (°C) to precipiation (natural logarithm, mm), May - September
Topography	Elevation (m)
	Cosine transformation of aspect (degrees)
	Weighted cumulative potential relative radiation during the growing season (Pierce, Lookingbill, & Urban, 2005)
	Slope (%)
	Topographic position index, calculated as the difference between a cell's elevation and the mean elevation of cells within a 150-m radius window
	Topographic position index, calculated as above but with a 450-m radius window
Parent material	Total ash deposition (feet), primarily from Mt. Mazama in the eastern Cascades (unpubl. data from Mike Simpson)
	Ice transported deposits (categorical)
	Pyroclastic deposits (categorical)
	Pumice desposits (categorical)
Landsat time series	Brightness axis from tasseled cap transformation
	Greenness axis from tasseled cap transformation
	Wetness axis from tasseled cap transformation
Landsat time series	Years since disturbance, from multitemporal Landat TM analysis
Disturbance metrics	Magnitude of disturbance (percent change in canopy cover), from multitemporal Landsat TM analysis
	Duration of disturbance (years), from multitemporal Landsat TM analysis
Lidar	Height (m) of 95th percentile of vegetation height returns
	Height (m) of average vegetation height returns
	Canopy cover above 2-m (percent), calculated as the proportion of first returns greater than a lower height limit of 2-m above ground in the digital terrian model
	Standard deviation of vegetation height (m)
Location	Universal Transverse Mercator easting (m)
	Universal Transverse Mercator northing (m)

vegetation metrics. Within this context, we focused on a small number of vegetation metrics characterizing vegetation height, cover, and vertical distribution (Table 2) that have consistently been found to predict many attributes of forest structure in the Pacific Northwest region (Falkowski et al., 2010; Hudak et al., 2008; Kane et al., 2010).

Tasseled cap indices and disturbance metrics were calculated from LTS mosaics (WRS-2 44/30, 45/29, and 45/30) using the LandTrendr segmentation algorithms (Kennedy et al., 2010). Using images that are cloud-free, geometrically corrected, and radiometrically normalized, the LandTrendr algorithms identify time segments for each pixel that describe sequences of disturbance and growth, and minimize annual variability from differences in sun angle, phenology, and atmospheric effects. LandTrendr algorithms also accomplish multi-year normalization at the pixel level, enabling mosaicking across large regions with less pronounced scene boundaries than associated with traditional normalization. From the LandTrendr data we generated mosaics of temporally smoothed tasseled cap indices, as well as disturbance history metrics from 1985 to 2010 characterizing years since disturbance (relative to 2010), magnitude of disturbance (as change in canopy cover), and duration of disturbance (in years).

To assign values from mapped explanatory variables to plots, each plot was represented as a template of pixels configured to approximate that plot's layout on the ground (Fig. 2). Whole plots were represented by a 3 × 3 template of 30 m pixels, while subplots were represented by a single 30 m pixel. Values of explanatory variables were extracted for templates representing whole plots and subplots, centered on both standard and improved GPS plot locations. Values of tasseled cap indices were extracted from images corresponding to the year each plot was measured.

2.4. Nearest-neighbor imputation mapping

Forest composition and structure were mapped across the study area using Gradient Nearest-Neighbor (GNN) imputation as described by Ohmann and Gregory (2002) and Ohmann, Gregory, Henderson,

Table 3
Model names and combinations of mapped explanatory variables, response data, and plot coordinates by model.

	1	2	3	4	5	6	7
	Plt-Std	Plt-lidar	Plt-disturb	Plt-lidar-disturb	Plt-lidar-disturb-GPS	Subplt-lidar-disturb	Subplt-lidar-disturb-GPS
Mapped explanatory variable groups							
Climate	†	†	†	†	†	†	†
Topography	†	†	†	†	†	†	†
Parent material	†	†	†	†	†	†	†
LTS fitted tasseled cap	†	†	†	†	†	†	†
LTS disturbance metrics			†	†	†	†	†
Lidar vegetation metrics		†		†	†	†	†
Field plot type							
Plot	†	†	†	†	†		
Subplot						†	†
Plot coordinates							
Standard	†	†	†	†		†	
Improved GPS					†		†

† denotes mapped explanatory variable group, field plot type (plot versus subplot), and type of plot coordinates (standard versus improved GPS) used in each model.

and Roberts (2011). Seven GNN maps were developed to quantify the influence of lidar and LTS disturbance metrics, plot locations (standard versus improved), and plot size (plot versus subplot) on prediction accuracy (Table 3). Map 1 (Plt-Std) is also referred to as the standard GNN map because its combination of explanatory variables, use of whole plots, and standard plot locations has been widely applied throughout the Pacific Northwest region (Moeur et al., 2011; Ohmann & Gregory, 2002; Ohmann et al., 2012). Compared to the standard GNN map, maps 2–7 include lidar vegetation and/or LTS disturbance metrics, response data calculated at different plot sizes (plot versus subplot), and explanatory data extracted for standard versus improved GPS plot locations.

Each map of forest composition and structure was created by first quantifying the relationship of multivariate response variables to mapped explanatory variables using canonical correspondence analysis (CCA, ter Braak, 1986). CCA models were created using the R package vegan (Oksanen et al., 2012; R Core Team, 2012). For species composition maps, all species from Table 1 were included as response variables in CCA. For structure models, we first included all structural response variables in CCA models, but found that using only live structural response variables in CCA (and predicting snags and down wood as ancillary variables) resulted in the highest prediction accuracy for all structure response variables. Neighbor selection in GNN was based on weighted Euclidean distance within CCA multivariate gradient space using the R package yaImpute (Crookston & Finley, 2008). GNN was implemented by assigning a single nearest-neighbor plot to each map pixel using the R package SDMap (Henderson unpublished, available upon request). All GNN maps were generated at 30 m resolution.

2.5. Accuracy assessment

Predictions of forest composition and structure were generated using a leave-one-out cross validation approach. From each GNN map, predicted response variables were extracted from map pixels associated with plot locations, using a 3 × 3 template of 30 m pixels for maps developed using plot-level response data, and a single 30 m pixel for maps developed using subplot-level response data. Leave-one-out cross validation was implemented using the R package SDMap. A variety of metrics to assess prediction accuracy were calculated using plot observations and cross-validated predictions.

To evaluate predictions of each structure variable in each map, we calculated r^2 and the normalized root mean squared difference (nRMSD, RMSE divided by the range of observed values). Normalization of RMSE allows for better comparison of models across structure variables with widely different data ranges and r^2 values. We also followed many of the protocols suggested by Riemann, Wilson, Lister, and Parks (2010) to assess differences in data distributions and spatial patterns of GNN predictions compared to plot-based observations. The bias and random error of predicted structure variables were quantified by calculating the systematic and unsystematic agreement coefficients (AC_{sys} and AC_{uns}) following Ji and Gallo (2006). AC_{sys} represents the difference between observed and predicted values that can be predicted by a simple linear model (bias from the 1:1 line), and AC_{uns} represents differences which appear to be random. Differences in observed versus predicted data distributions were quantified by calculating empirical cumulative distribution functions (ECDFs) for observed and predicted structure variables, and the Kolmogorov-Smirnov statistic (KS, Lopes, Reid, & Hobson, 2007) was calculated to assess the maximum distance between observed and predicted ECDF values. We used hexagons with a 12.5 km spacing (13,532 ha per hexagon) to assess the spatial patterns of prediction accuracy for structure variables. This hexagon size resulted in 57 hexagons averaging 4 plots per hexagon, the minimum number of plots used by FIA to define strata for estimation (Scott et al., 2005). Observed and predicted values for each structure variable were averaged for plots contained in each hexagon. Maps were then generated to visualize mean observed values and percent relative difference between observed and predicted values by hexagon. The relatively small number of plots and study area size precluded the use of larger hexagons or calculation of prediction confidence intervals by hexagon as suggested by Riemann et al. (2010).

Species predictions were evaluated as binary values (presence/absence) using the True Skill Statistic (TSS, Allouche, Tsoar, & Kadmon, 2006), as well as sensitivity and specificity (Fielding & Bell, 1997) for each species in each map. Sensitivity (also known as the true positive rate) was calculated as the number of true positives divided by the sum of true positives and false negatives; while specificity (also known as the true negative rate) was calculated as the number of true negatives divided by the sum of true negatives and false positives. Finally, the areal representation of each species in each map was calculated by comparing the proportion of mapped area each species was predicted to occupy in relation to the proportion of plots each species was observed in from the systematic plot sample.

3. Results

3.1. Vegetation structure

Prediction accuracy varied by structure variable type (i.e. live trees, snag, and down wood), inclusion or exclusion of lidar and LTS disturbance explanatory variables, and use of plot versus subplot response data (Fig. 3). R^2 ranged from 0.0 to 0.77, with the highest values for structure variables that included live trees (BA, QMD, and HCB), intermediate values for live tree density variables (TPH_3_25, TPH_25_50, and TPH_GE_50), lower values for snag density variables (STPH_12_25, STPH_25_50, and STPH_GE_50), and the lowest values ($r^2 < 0.28$) for down wood variables (DVPH_12_25, DVPH_25_50, and DVPH_GE_50). Inclusion of lidar explanatory variables consistently improved predictions of all structure variables. Inclusion of LTS disturbance metrics also improved predictions of all structure variables, although not as much as inclusion of lidar. The combination of lidar and LTS disturbance metrics (GNN maps Plt-lidar-disturb and Plt-lidar-disturb-GPS) improved r^2 values for 7 of the 12 structure variables compared to the map developed with only lidar. However, r^2 values for snags and down wood were on average, highest for maps including both lidar and LTS disturbance metrics. Subplot-based maps exhibited consistently lower r^2 values. The plot-based map using improved GPS plot locations (Plt-lidar-disturb-GPS) improved r^2 values for only 5 of the 12 structure variables compared to the map using the same explanatory variables but with standard plot locations (Plt-lidar-disturb), and the subplot-based map using improved GPS plot locations (Subplt-lidar-disturb-GPS) also had higher r^2 values for only 5 of the 12 structure variables compared to the map using the same explanatory variables and standard plot locations (Subplt-lidar-disturb).

Patterns in model-strength illustrated by nRMSD were similar to those shown by r^2; except values are reversed (lower nRMSD is better). Values for nRMSD ranged from 0.08 to 0.30. For all structure variables, maps using lidar explanatory variables had the lowest nRMSD, and the combination Lidar and LTS disturbance explanatory variables lowered nRSMD for 6 of the 12 structure variables. Values for nRSMD were higher for subplot-based maps, except for down wood response variables. Improved plot locations did not improve nRMSD values compared to maps using standard plot locations.

AC_{sys} ranged from 0.26 to 1.0, were consistently high (>0.92) among maps for all live tree variables, but varied for snag densities and down wood volumes. For snag and down wood predictions, there were few trends for AC_{sys} values in relation to map type. AC_{uns} values ranged from −14.08 to 0.76, and had higher values for live tree variables, intermediate values for snags, and lower values for down wood. AC_{uns} values were consistently higher for maps that included lidar explanatory variables, while the addition of LTS disturbance variables did not improve AC_{uns} values over maps using only lidar variables. Subplot-based maps had the lowest AC_{uns} values for 8 of 12 structure variables. Improved GPS plot locations did not influence AC_{uns} values. KS values had a

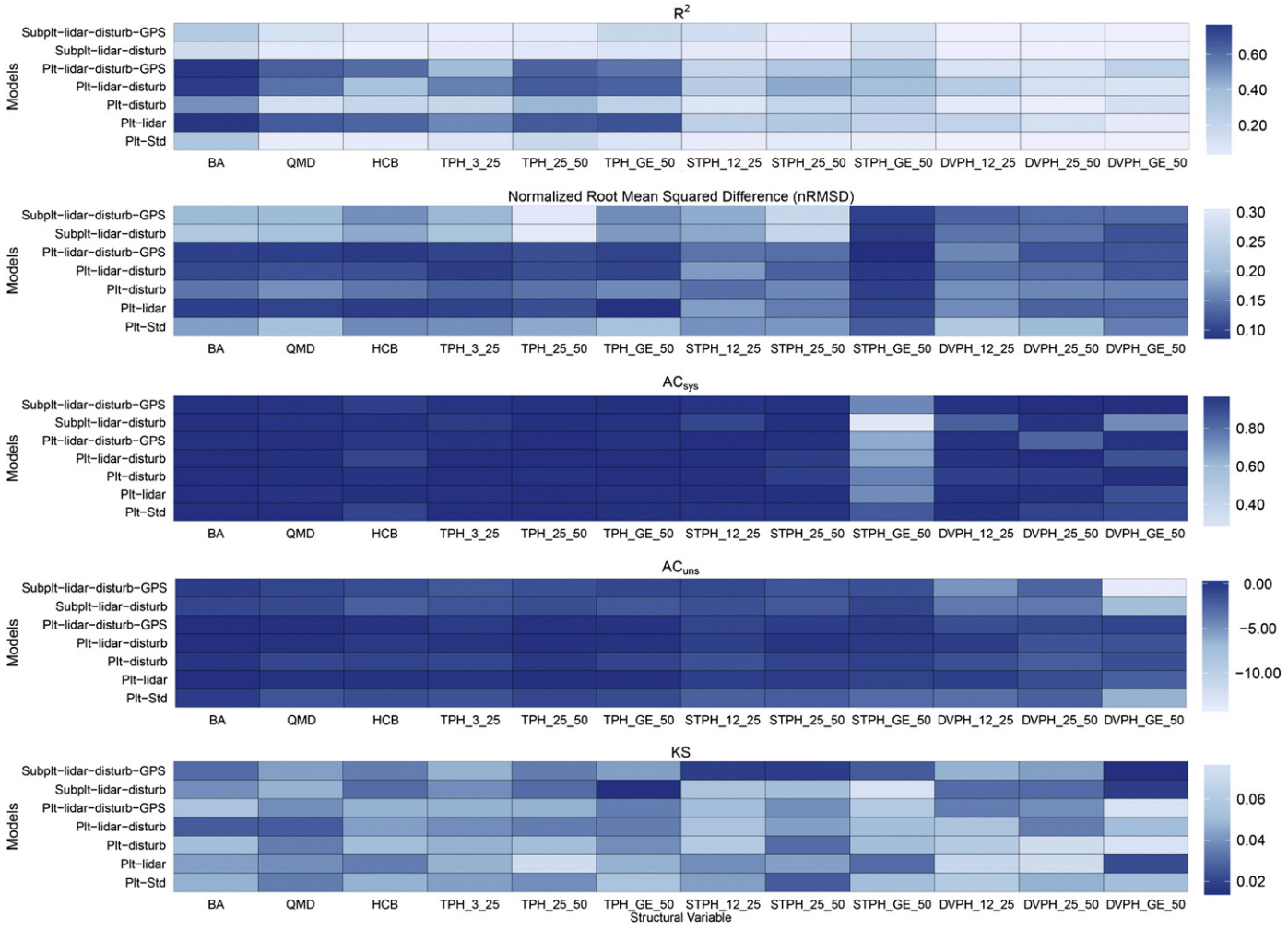

Fig. 3. Accuracy metrics (r^2, nRSMD, AC_{sys}, AC_{uns}, and KS) for each structural variable and GNN map type. See Table 1 for structural variable descriptions and Table 3 for descriptions of GNN map types.

narrow range of 0.01 to 0.11, with no consistent trends between variables or map types.

Hexagon maps of relative differences between predicted and observed structure variables displayed consistent trends in relation to specific structure variables, map types, and geographic location (Fig. 4, Supplemental Fig. S3–S13). For most live tree structure variables, maps developed using lidar and/or LTS disturbance metrics had more hexagons with low (−25 to 25%) relative differences between map predictions and plot-based observations. However, whether the "best" map included lidar, LTS disturbance metrics, or both varied by the specific live tree structure variable. For live structure variables, subplot-based maps had more hexagons with high relative differences, and the use of improved GPS plot locations did not influence relative differences. In contrast to most live structure variables, snag and large down wood variables had the lowest relative differences for subplot-based maps. However, higher accuracy of subplot-based maps for snag and down wood variables in this hexagon-based metric may result from strong geographic clustering of very low snag density and low down wood volume in the eastern portion of the study area. As with live tree variables, use of improved GPS plot locations did not influence relative differences for snag density and down wood volume.

3.2. Species composition

Accuracy of species presence/absence predictions varied by map type and species, but there were consistent patterns primarily related to the inclusion or exclusion of lidar and LTS disturbance variables, as well as the use of plot versus subplot response data (Fig. 5). TSS values ranged from −0.03 to 0.75, and were generally higher for more prevalent species. Maps using LTS disturbance metrics as explanatory variables had the highest TSS values for 8 of the 12 species. TSS values were consistently lower for subplot-based maps, and this pattern was most pronounced for the least prevalent species. The plot-based map using improved GPS plot locations improved TSS values for only 5 of 12 species compared to the map using the same explanatory variables and standard locations. Improved GPS locations had a similar result for subplot-based maps, with improved TSS values for only 4 of 12 species compared to the subplot-based maps using standard locations.

Sensitivity displayed trends similar to TSS, varying by model and species, ranging from 0.0 to 0.93, with higher sensitivity for more prevalent species. Maps using LTS disturbance metrics had higher sensitivity for eight of the twelve species (Fig. 5). The plot-based map using improved GPS locations improved sensitivity for only 5 of the 12 species compared to the plot-based map using the same explanatory variables and standard plot locations. The subplot-based map using improved GPS plot locations only improved sensitivity for 2 of the 12 species compared to the subplot-based map using the same explanatory variables and standard plot locations. Specificity was generally higher than TSS or sensitivity, ranging from 0.61 to 1.0. Specificity was associated more with species prevalence (higher specificity with lower species prevalence) than with map type. Unlike TSS and sensitivity, specificity did not display consistent trends between plot-based versus subplot-

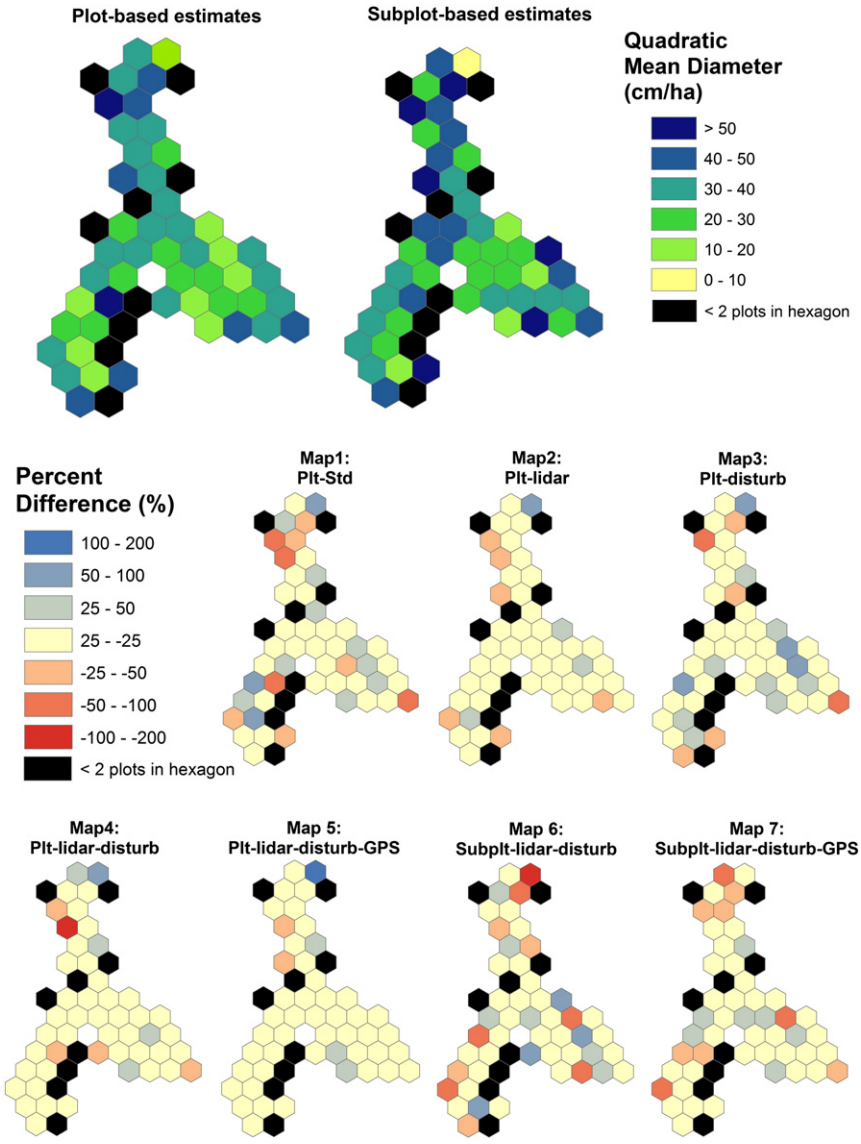

Fig. 4. Geographic distribution of relative percent difference between observed and predicted quadratic mean diameter (QMD) at the 12.5 km hexagon scale. See Table 3 for descriptions of GNN map types. See the Supplemental Fig. S3–S13 for maps of relative percent difference for other structural variables listed in Table 1.

based maps or maps using standard versus improved GPS plot locations. While specificity did vary among maps (especially for species with moderate to high prevalence), there were no consistently high- or low-performing maps across a plurality of species.

Difference of areal representation between map predictions and sampled plots varied largely by species and species prevalence rather than map type (Fig. 5). The most prevalent species (*Pinus ponderosa*, *Pinus contorta*, and *Abies grandis/concolor*) had the narrowest range of difference of areal representation (−9.83% to 4.22%), suggesting these species are least sensitive to map type and most accurately predicted. Species with moderate prevalence (*Tsuga mertensiana*, *Pseudotsuga menziesii*, and *Pinus monticola*) had an increased range of difference of areal representation (−27.27% to 22.93%), while the remaining species with low prevalence had the widest range of difference of areal representation (−50.56% to 40.98%). Most species tended to be either over-predicted or under-predicted across all map types. Subplot-based maps consistently over- or under-predicted mapped area compared to plot-based maps. Otherwise, trends in areal representation did not appear to be associated with the inclusion or exclusion of lidar, LTS disturbance metrics, or improved GPS plot locations. Aerial representation is a regional accuracy assessment, whereas all other metrics (except for KS statistic) examining species composition focused on plot scale accuracy. Results of regional and plot scale accuracy assessments suggest lidar and LTS disturbance metrics improved plot scale accuracy, but not regional accuracy.

4. Discussion

4.1. Influence of lidar-derived vegetation metrics on prediction accuracy

Inclusion of lidar-derived vegetation metrics greatly improved NN predictions of live tree structural attributes, moderately improved predictions of snag densities and down wood volume, but had inconsistent effects on species predictions. Large improvements in live structure predictions using lidar were not surprising, as other studies have found lidar can successfully predict forest structure variables such as basal area and stem density with r^2 values of 0.86–0.95 (Holmgren, 2004; Hudak et al., 2006; Means et al., 2000; Næsset, 2002). These studies outperformed NN imputation using lidar in this study (r^2 = 0.77 for basal area, 0.53–0.67 for stem densities in different size classes). However, there are many confounding factors that complicate direct comparisons with our work. These other studies were situated in relatively small spatial extents in different forest types, used fixed-radius plots, often used many lidar-

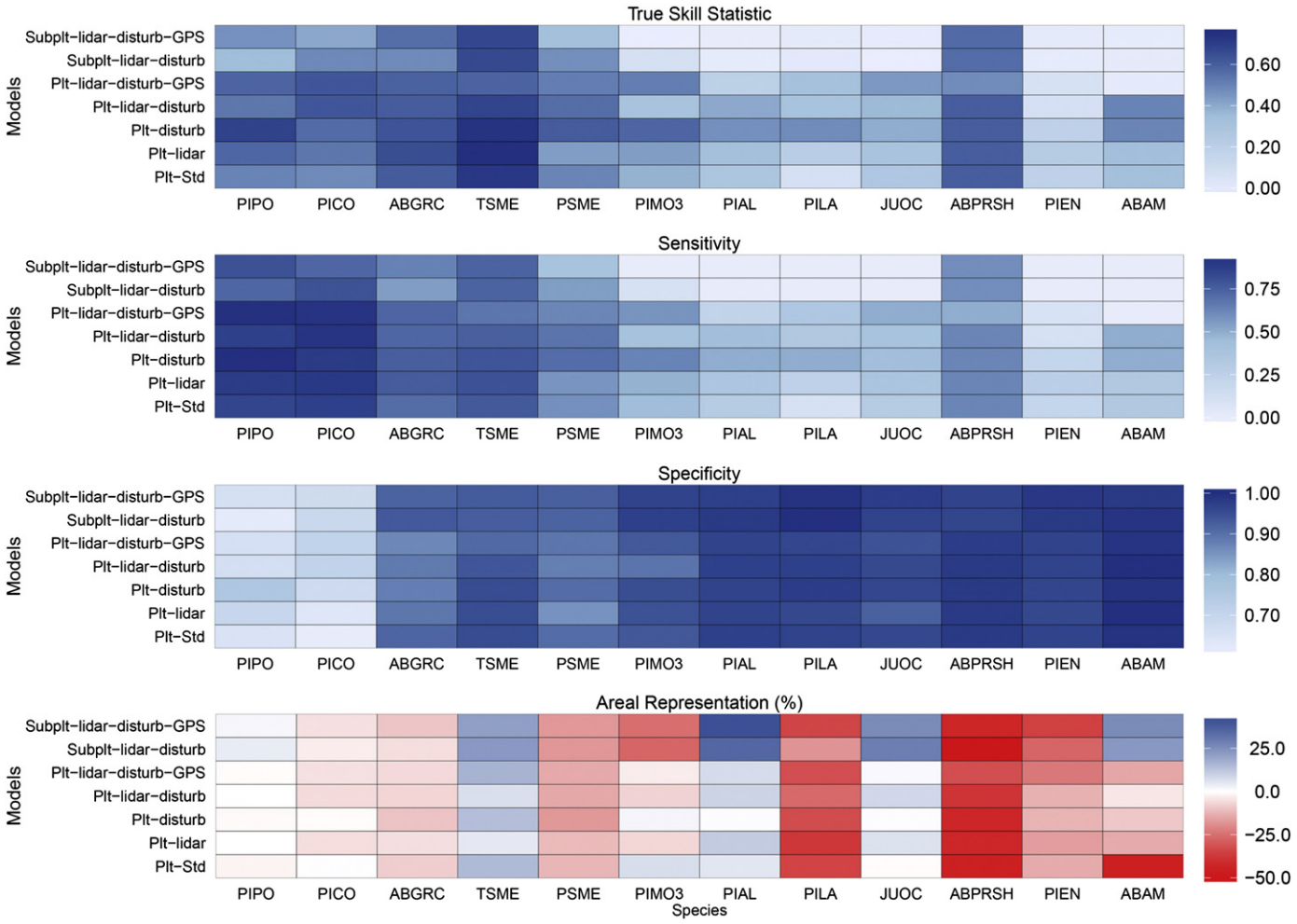

Fig. 5. Accuracy metrics (true skill statistic, sensitivity, specificity, and areal representation) for each species (ordered in declining prevalence from left to right) and GNN map type. See Table 1 for composition variable descriptions and prevalence values, and Table 3 for descriptions of GNN map types.

derived explanatory variables, and predicted individual univariate response variables. In contrast, our study was conducted over a larger regional spatial extent, used inventory plots of different radii, used relatively few lidar-derived explanatory variables, and generated multivariate predictions of response variables. Furthermore, differences in how lidar data is spatially related to plots may also confound direct comparisons between out study and prior research. Despite limitations for direct comparison, our study demonstrates that inclusion of relatively few lidar-derived vegetation metrics can greatly improve simultaneous, multivariate predictions of live forest structural attributes within the context of NN imputation mapping. The primary benefit of NN imputation when using a single nearest neighbor [$k = 1$] is the conservation of plot-level covariance among multiple response variables, resulting in ecologically realistic co-occurrence of multiple response variables. However, conservation of plot-level covariance may come at the cost of lower prediction accuracy for individual response variables compared to univariate statistical approaches (Henderson, Ohmann, Gregory, Roberts, & Zald, in press). In spite of this potential tradeoff, this study clearly demonstrates that inclusion of lidar data can greatly improve regional NN imputation mapping of live tree forest structure.

Inclusion of lidar-derived vegetation metrics moderately improved predictions of snag densities. Martinuzzi et al. (2009) reported high accuracy using lidar to make categorical predictions of snag presence/absence, and Kim et al. (2009) had success using lidar to distinguish between live tree and snag biomass, but it is difficult to directly compare these results to continuous predictions of snag densities in our study. Lidar's application to snag prediction is fundamentally limited by the small reflective surface area of snags compared to the live forest canopy, which is further complicated by the high canopy cover across parts of this study. Additionally, the absence of lidar intensity values may limit our ability to predict snags. Lidar intensity values have proven useful in predicting snags (Kim et al., 2009; Pflugmacher et al., 2012), but use of intensity values often requires corrections for differences in receiver gain and range that occur during data acquisition. Unfortunately, corrected intensity values were not available for this study, and visual inspection of gridded intensity values revealed significant terrain effects and flight line patterns which precluded the use of uncorrected intensity values.

Predictions of down wood volume also improved with inclusion of lidar, but were still low compared to live tree and snag attributes. Low accuracy for down wood predictions is consistent with both high sampling error associated with plot measurements, as well as difficulties in detecting and measuring down wood underneath forest canopies using remote sensing. FIA and CVS plots use relatively short transects nested within subplots for measuring down wood (Max et al., 1996; USFS, 2003), and short transects can increase sampling error of down wood (Harmon & Sexton, 1996; Woldendorp, Keenan, Barry, & Spencer, 2004). Other studies have used lidar to achieve higher prediction accuracies for down wood by focusing on intensity values, near-ground point filtering, and object-based image analysis (Blanchard, Jakubowski, & Kelly, 2011; Pesonen, Maltamo, Eerikäinen, & Packalèn, 2008). As previously mentioned, corrected intensity values were not available for this study. Additionally, the object-based image analysis approach used by Blanchard et al.

(2011) was conducted in disturbed forests with low canopy cover, and it is doubtful this approach would be effective in the closed-canopy forests within our study area.

Previous studies using lidar to detect and/or measure down wood have found that using live trees as explanatory variables did not improve predictions. Pesonen et al. (2008) found that the inclusion of live trees as explanatory variables did not improve lidar-based down wood predictions, while Blanchard et al. (2011) cite low correlations between live trees and down wood as evidence live trees are not important predictors of down wood. However, these studies examined down wood as univariate response variables, rather than as part of a covarying suite of forest structural attributes. In our plot data, down wood volume was correlated with many live tree and snag variables, albeit at relatively low levels (Supplemental Table S1). Within the context of NN imputation mapping in our study where $k = 1$, it appears that by improving predictions of live tree structure, including lidar explanatory data also improved down wood (and snag) predictions due to the covariance between live and dead structure variables.

In contrast to structural predictions, lidar did not improve predictions of species presence/absence. On one hand, this is unsurprising because species distributions are primarily driven by large-scale gradients of climate, topography, and soil parent materials (Ohmann & Gregory, 2002; Ohmann et al., 2011). On the other hand, finer-scale patterns of species composition do occur in relation to succession over time after disturbance in western Oregon (Schoonmaker & McKee, 1988; Spies, 1991). We expected lidar to contain information pertinent to these finer-scale patterns because it has been shown to characterize structural complexity associated with succession and stand development (Kane et al., 2010). However, in this study the linkages between stand structure and species composition may not have been strong enough to improve species-level predictions. One possible explanation for this shortcoming is the wide range of forest types, structural conditions, and disturbance histories found in the study area, which may not be adequately characterized by the lidar variables used or by our plot sample. As with snag predictions, lidar intensity values may improve species predictions (Donoghue, Watt, Cox, & Wilson, 2007).

4.2. Influence of LTS disturbance metrics on prediction accuracy

Inclusion of LTS disturbance metrics consistently improved predictions of structural attributes, although less than the inclusion of lidar variables. This is in contrast to the findings of Pflugmacher et al. (2012), who found that LTS metrics of disturbance and recovery predicted live biomass almost as well as lidar (r^2 of 0.80 versus 0.88), and outperformed lidar in predicting dead biomass (r^2 of 0.73 versus 0.23). Complicating direct comparisons, Pflugmacher et al. (2012) used a different field plot design, predicted different structure variables, and predicted them as univariate response variables. However, a likely explanation for the modest improvements in prediction accuracy with LTS disturbance metrics in our study is the absence of disturbance attribution (i.e. identification of disturbance type). LTS disturbance metrics used in this study (magnitude, duration, and years since disturbance) can be very similar for different types of disturbance agents, despite having strikingly different impacts on forest structure. For example, a clearcut and a high severity wildfire may both be high magnitude, short duration disturbance events. However, a clearcut removes most trees, resulting in little snag and down wood recruitment; while wildfire converts live trees to snags, which become down wood over time (Dunn & Bailey, 2012; Franklin et al., 2002).

Although disturbances such as clearcuts and wildfire may initially have distinct spectral signatures, additional information such as disturbance size, shape, and distribution are often required to distinguish between them (Cohen et al., 2002). Further complicating the relationship of spectral reflectance to disturbance types in this study, plots are often sampled many years after a disturbance, and are matched to Landsat imagery by their sampling date. Plots surveyed several years after a disturbance may contain recovering forests, so Landsat imagery will reflect the young trees rather than the initial disturbance. Put another way, a 10-year old fire may closely resemble a 10-year old clearcut in Landsat imagery even though they may differ greatly in structural dimensions like downed wood and snags. Within the context of this study, the absence of disturbance attribution means plots that have been logged or burned may be positioned close to each other in CCA gradient space, resulting in similar likelihoods of being chosen as nearest neighbors for either type of disturbance, thereby lowering prediction accuracy of structural attributes. LTS disturbance metrics were only recently developed for our study area, and disturbance attribution has yet to be completed and validated. However, the work of Pflugmacher et al. (2012) suggests incorporating LTS disturbance attribution into future NN imputation may significantly improve predictions of forest structure. It should be noted that, even in the absence of disturbance attribution, LTS disturbance metrics moderately improved all structural predictions including snag density and down wood volume. Improvements in snag and down wood predictions may simply result from covariance with live tree structure, but it may also suggest LTS disturbance history can be important in predicting some structure variables even without knowledge of the specific disturbance agent.

Inclusion of LTS disturbance metrics improved prediction accuracy for many species. These included a fire-adapted species (*Pinus contorta*) associated with more recent and higher magnitude disturbances, shade-intolerant and shade-intermediate species (*J. occidentalis*, *Pinus albicaulis*, *Pinus lambertiana*, *P. monticola*, *P. ponderosa*, and *P. menziesii*) that often establish after disturbances, and a shade-tolerant species (*Abies amabilis*) that was negatively associated with disturbance (Supplemental Fig. S14). As previously mentioned, species distributions are driven by broad-scale bioclimatic gradients, as well as local-scale successional gradients over time. The inclusion of LTS disturbance metrics appeared to moderately improve predictions of many species by providing information related to successional conditions, even though the disturbance type is not known. Typically, species distribution models (SDMs) focus more on the bioclimatic and/or physiographic niche (Elith & Leathwick, 2009; Guisan & Thuiller, 2005; Iverson & Prasad, 1998) than on the importance of disturbance in determining species distributions. As LTS disturbance metrics become widely available, their incorporation into SDMs has the potential to improve SDM predictions at regional to global spatial extents. Furthermore, predicting distributions of multiple species simultaneously using NN imputation has the benefit of generating more ecologically realistic maps of species diversity and community composition (Henderson et al., in press; Ohmann et al., 2011).

4.3. Influence of improved plot locations on prediction accuracy

Improved GPS plot locations had little influence on predictions of forest attributes, suggesting that standard locations of regional inventory plots may be sufficient for large-landscape NN imputation mapping of forest composition and structure with lidar and Landsat imagery. However, our findings are in contrast to simulations which found that accuracy of plot locations did influence predictions of forest attributes (Frazer et al., 2011; McRoberts, 2010). In the context of NN imputation and other methods linking field plots to remotely sensed data, we suggest that many factors need to be considered to determine if plot positional accuracy affects prediction accuracy, including: 1) accuracy of plot locations in relation to the spatial resolution of explanatory data, 2) uncertainty in accuracy of plot locations, 3) spatial mismatches between plots and explanatory data, 4) co-registration error between mapped explanatory variables, and 5) the scale of spatial autocorrelation of both forest response variables and mapped explanatory variables.

The average difference between standard and improved GPS locations was (18.18 m), over half the width of a pixel of Landsat imagery used in this study. Within the context of the plot templates used to relate explanatory data to plots, the average difference between standard

and improved plot locations changed a majority of pixel assignments for subplots, but only a minority of pixel assignments for whole plots (Supplemental Fig. S15), suggesting that standard plot locations are compatible with the spatial resolution of whole plots and Landsat imagery. However, it is important to note that improved GPS "accuracy" (±0.91 m in this study) is really the precision of the solution, not deviation from true geographic location. GPS receivers like the one used in this study are classified as sub-meter GPS receivers by manufacturers. However, this class of GPS receivers produce sub-meter accuracy only in open-canopy, clear-sky conditions, and have much poorer accuracy (average ± 2 m, outliers approaching 8 m) under forest canopy (Wing, Eklund, John, & Richard, 2008). This indicates that while the improved GPS plot locations in this study are an improvement over standard plot locations, they might have too much positional error to influence NN imputation in forests with high canopy cover. Accuracy of plot locations also needs to be considered in the context of how plots are spatially related to explanatory variables. Some studies precisely match spatial footprints of plots to lidar data (Andersen, McGaughey, & Reutebuch, 2005; Erdody & Moskal, 2010), but our study used the common approach in NN imputation to represent plots as templates of pixels at the same resolution as the imagery used. This approach can result in spatial mismatches in size and location between plots and explanatory data (Fig. 2), which could reduce the influence of improved plot locations. In addition to errors associated with accuracy of plot locations and spatial mismatches between plot templates and explanatory data, co-registration errors can exist between different explanatory data sources. For example, reported positional accuracy of orthorectified L1G Landsat scenes is ±50 m (Tucker, Grant, & Dykstra, 2004), and other geospatial datasets will have different levels of positional accuracy. Furthermore, explanatory data from multiple sources often have different native spatial resolutions and geographic projections, and conversion to a common resolution and projection can generate additional spatial errors. Finally, spatial autocorrelation of forest response variables and mapped explanatory variables may reduce the importance of accurate plot locations. Small errors in plot locations may result in spatial mismatches between forest response and mapped explanatory variables if either vary at a fine spatial resolution. This may be true in ecotones with short biophysical gradients (Zald et al., 2012), or in landscapes with abundant edge features and split land cover conditions on plots (Cooke, 2000). However in this study, response and explanatory variables may have spatial autocorrelation at distances greater than the error of standard plot locations, negating potential gains from more accurate plot locations. It is important to note that each of these factors may not individually negate the influence of plot location accuracy, but additive effects may equal or exceed the influence of plot positional accuracy, and any given mapping application may have a different combination of factors in which accurate plot locations may or may not improve prediction accuracy.

4.4. Influence of plot size on prediction accuracy

Plot size was strongly associated with prediction accuracy, with subplot-based maps having greatly reduced prediction accuracy. There are several methodological explanations for poor prediction accuracy using subplots. Species had lower prevalence on subplots than on plots. More prevalent species occupy larger areas and have wider niches, resulting in map predictions with higher sensitivity and lower specificity. Our study found that declines in accuracy between plot- and subplot-based maps were highest for the least prevalent species, suggesting that smaller subplots imposed a sampling effect on species-area curves (sensu Hill, Curran, & Foody, 1994). Smaller plots had a lower probability of detecting rarer species, and subsequently reduced prediction sensitivity. This sampling effect likely applies to rare forest structures, such as large dead wood, as well. Reduced plot size may also alter the relationship between response variables and remotely sensed explanatory variables. For example, lidar-derived vegetation metrics were extracted for the canopy above a given plot template, but these metrics can be influenced by overhanging branches from trees whose stems are outside the plot, and the opposite can be true as well. This can result in lidar metrics that do not represent the trees measured on plots, and this mismatch becomes more problematic as plot size decreases. However, a potential confounding factor in our analysis of plot size is that FIA and CVS plots differ somewhat in the total area and transect length sampled for live trees, snags, and down wood (Supplemental Fig. S1). CVS plots have a larger sampling area that may better detect rare species, trees, and snags while FIA plots have longer transect lengths that may better detect rare down wood. We used the same template of plot pixels to represent FIA and CVS plots, so spatial mismatches we previously discussed in relation to accuracy of plot positions may also be important in relation to plot size. Unfortunately, we did not have adequate sample sizes to quantify the influence of plot design (FIA versus CVS) on prediction accuracy, but this may be an important consideration for regional mapping applications when plot data comes from multiple inventory programs with different plot sizes and measurement protocols.

Compared to methods that average observations, NN imputation using a single nearest neighbor [$k = 1$] may be especially sensitive to sampling effects, because more rare attributes should be included in maps when imputing multiple nearest neighbors. One of the biggest strengths of NN imputation where $k = 1$ is the conservation of plot-level co-variance for response variables, but like any empirical approach NN imputation using a single nearest neighbor will propagate sample-based errors such as measurement errors or inadequacies of the sample design.

5. Conclusions

Integrating lidar and the LTS disturbance metrics into NN imputation has the potential to greatly improve mapping of forest composition and structure over large landscapes. This study found that standard plot locations of regional forest inventory plots are sufficient for large-extent NN imputation mapping using lidar. While the best overall forest maps included both lidar and LTS disturbance metrics, no single map type provided the best predictions of all species composition and structure attributes. This suggests that map developers and users need to be aware of what compositional and structural attributes are most important for a given map application, and develop or select maps to emphasize those attributes accordingly. For example, lidar clearly improved structural predictions, LTS disturbance metrics improved species predictions, and the combination of lidar and LTS disturbance metrics resulted in the most accurate snag predictions.

It is also important to recognize that even the inclusion of lidar and LTS disturbance metrics resulted in only moderately accurate predictions of snags and down wood. Given the ecological importance of snags and down wood in many forest ecosystems (Burrascano, Keeton, Sabatini, & Blasi, 2013; Harmon et al., 1986), it may be desirable to use other methods for mapping these structural attributes (Pesonen et al., 2008; Kim et al., 2009; Martinuzzi et al., 2009; Blanchard et al., 2011). However, mapping these variables independently would likely come with the cost of disrupting covariance relationships with other variables that is a key strength of NN imputation mapping when $k = 1$. Alternatively, incorporation of lidar intensity values and LTS disturbance types as explanatory variables may provide many of the same benefits. Finally, this study suggests that LTS disturbance metrics may improve models of species distributions, and species distribution models should incorporate disturbance history as regional and global disturbance maps are increasingly available.

Acknowledgments

We thank Mike Simpson and personnel from the Deschutes National Forest for providing improved GPS coordinates for FIA and CVS

inventory plots. Funding for this study was provided by the USDA Forest Service Pacific Northwest Research Station, and a grant from the Western Wildland Environmental Threat Assessment Center. LandTrendr-processed imagery used in this study was developed under a project supported by a grant from the USDA National Institute for Forests and Agriculture.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2013.12.013.

References

Agee, J. K. (1993). *Fire ecology of Pacific Northwest forests*. Washington, D.C: Island Press.

Agee, J. K. (1994). *Fire and weather disturbances in terrestrial ecosystems of the eastern Cascades*. Pacific Northwest Research Station Portland, OR: USDA Forest Service.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*, 1223–1232.

Andersen, H. -E., McGaughey, R. J., & Reutebuch, S. E. (2005). Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment*, *94*, 441–449.

Blanchard, S. D., Jakubowski, M. K., & Kelly, M. (2011). Object-based image analysis of downed logs in disturbed forested landscapes using lidar. *Remote Sensing*, *3*, 2420–2439.

Burrascano, S., Keeton, W. S., Sabatini, F. M., & Blasi, C. (2013). Commonality and variability in the structural attributes of moist temperate old-growth forests: A global review. *Forest Ecology and Management*, *291*, 458–479.

Cohen, W. B., & Goward, S. N. (2004). Landsat's role in ecological applications of remote sensing. *Bioscience*, *54*, 535–545.

Cohen, W. B., Maiersperger, T. K., Spies, T. A., & Oetter, D. R. (2001). Modelling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. *InternationalJournal of Remote Sensing*, *22*, 2279–2310.

Cohen, W. B., Spies, T. A., Alig, R. J., Oetter, D. R., Maiersperger, T. K., & Fiorella, M. (2002). Characterizing 23 years (1972–95) of stand replacement disturbance in western Oregon forests with Landsat imagery. *Ecosystems*, *5*, 122–137.

Cooke, W. H. (2000). Forest/non-forest stratification in Georgia with Landsat Thematic Mapper data. In Ronald E. McRoberts, Gregory A. Reams, & Paul C. Van Deusen (Eds.), *Proceedings of the First Annual Forest Inventory and Analysis Symposium; Gen. Tech. Rep. NC-213* (pp. 28–30). : St. Paul, MN: US Department of Agriculture, Forest Service, North Central Research Station.

Crist, E. P., & Cicone, R. C. (1984). Application of the Tasseled Cap concept to simulated thematic mapper data (transformation for MSS crop and soil imagery). *Photogrammetric Engineering and Remote Sensing*, *50*, 343–352.

Crookston, N. L., & Finley, A. O. (2008). yalmpute: An R package for kNN imputation. *Journal of Statistical Software*, *23*, 1–16.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*, 2031–2064.

Donoghue, D. N., Watt, P. J., Cox, N. J., & Wilson, J. (2007). Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment*, *110*, 509–522.

Dubayah, R. O., & Drake, J. B. (2000). Lidar remote sensing for forestry. *Journal of Forestry*, *98*, 44–46.

Dunn, C. J., & Bailey, J.D. (2012). Temporal dynamics and decay of coarse wood in early seral habitats of dry-mixed conifer forests in Oregon's Eastern Cascades. *Forest Ecology and Management*, *276*, 71–81.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 677–697.

Erdody, T. L., & Moskal, L. M. (2010). Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment*, *114*, 725–737.

Eskelson, B., Temesgen, H., & Hagar, J. (2012). A comparison of selected parametric and imputation methods for estimating snag density and snag quality attributes. *Forest Ecology and Management*, *272*, 26–34.

Eskelson, B. N. I., Temesgen, H., Lemay, V., Barrett, T. M., Crookston, N. L., & Hudak, A. T. (2009). The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scandinavian Journal of Forest Research*, *24*, 235–246.

Everett, R. L., Schellhaas, R., Keenum, D., Spurbeck, D., & Ohlson, P. (2000). Fire history in the ponderosa pine/Douglas-fir forests on the east slope of the Washington Cascades. *Forest Ecology and Management*, *129*, 207–225.

Falkowski, M. J., Hudak, A. T., Crookston, N. L., Gessler, P. E., Uebler, E. H., & Smith, A.M. S. (2010). Landscape-scale parameterization of a tree-level forest growth model: A k-nearest neighbor imputation approach incorporating LiDAR data. *Canadian Journal of Forest Research*, *40*, 184–199.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*, 38–49.

Franklin, J. F., & Dyrness, C. T. (1988). *Natural vegetation of Oregon and Washington.*

Franklin, J. F., Spies, T. A., Pelt, R. V., Carey, A.B.., Thornburgh, D. A., Berg, D. R., et al. (2002). Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *Forest Ecology and Management*, *155*, 399–423.

Frazer, G., Magnussen, S., Wulder, M., & Niemann, K. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment*, *115*, 636–649.

Gjertsen, A. K. (2007). Accuracy of forest mapping based on Landsat TM data and a kNN-based method. *Remote Sensing of Environment*, *110*, 420–430.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*, 993–1009.

Hansen, M., DeFries, R., Townshend, J., Carroll, M., Dimiceli, C., & Sohlberg, R. (2003). Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm. *Earth Interactions*, *7*, 1–15.

Harmon, M. E., Franklin, J. F., Swanson, F. J., Sollins, P., Gregory, S., Lattin, J., et al. (1986). Ecology of coarse woody debris in temperate ecosystems. *Advances in Ecological Research*, *15*, 302.

Harmon, M. E., & Sexton, J. (1996). *Guidelines for measurements of woody detritus in forest ecosystems*. Office Seattle (WA): US LTER Network.

Henderson, E. B., Ohmann, J. L., Gregory, M. J., Roberts, H. M., & Zald, H. (2014). All for one or one for all: Should plant communities be mapped one species at a time or all species at once? *Applied Vegetation Science* (in press).

Heyerdahl, E. K., Brubaker, L. B., & Agee, J. K. (2001). Spatial controls of historical fire regimes: A multiscale example from the interior west, USA. *Ecology*, *82*, 660–678.

Hill, J. L., Curran, P. J., & Foody, G. M. (1994). The effect of sampling on the species-area curve. *Global Ecology and Biogeography Letters*, *97–106*.

Holmgren, J. (2004). Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scandinavian Journal of Forest Research*, *19*, 543–553.

Hoppus, M., & Lister, A. (2005). The status of accurately locating forest inventory and analysis plots using the global positioning system. In Ronald E. McRoberts, Gregory A. Reams, & Paul C. Van Deusen (Eds.), *Proceedings of the seventh annual forest inventory and analysis symposium; Gen. Tech. Rep. WO-77* (pp. 179–184). Portland: ME: US Department of Agriculture, Forest Service.

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, *114*, 183–198.

Hudak, A. T., Crookston, N. L., Evans, J. S., Falkowski, M. J., Smith, A.M., Gessler, P. E., et al. (2006). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing*, *32*, 126–138.

Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., & Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment*, *112*, 2232–2245.

Iverson, L. R., & Prasad, A.M. (1998). Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs*, *68*, 465–485.

Ji, L., & Gallo, K. (2006). An agreement coefficient for image comparison. *Photogrammetric Engineering and Remote Sensing*, *72*, 823–833.

Kane, V. R., Bakker, J.D., McGaughey, R. J., Lutz, J. A., Gersonde, R. F., & Franklin, J. F. (2010). Examining conifer canopy structural complexity across forest ages and elevations with LiDAR data. *Canadian Journal of Forest Research*, *40*, 774–787.

Kauth, R. J., & Thomas, G. (1976). The tasselled cap—A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *LARS Symposia* (pp. 159).

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*, 2897–2910.

Kim, Y., Yang, Z., Cohen, W. B., Pflugmacher, D., Lauver, C. L., & Vankat, J. L. (2009). Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sensing of Environment*, *113*, 2499–2510.

Lefsky, M.A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, *52*, 19–30.

Littell, J. S., McKenzie, D., Peterson, D. L., & Westerling, A. L. (2009). Climate and wildfire area burned in western US ecoprovinces, 1916–2003. *Ecological Applications*, *19*, 1003–1021.

Lopes, R. H., Reid, I., & Hobson, P. R. (April 23–27, 2007). The two-dimensional Kolmogorov–Smirnov test. *XI international workshop on advanced computing and analysis techniques in physics research, Nikhef, Amsterdam, the Netherlands.*

Lu, D. (2006). The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing*, *27*, 1297–1328.

Martinuzzi, S., Vierling, L. A., Gould, W. A., Falkowski, M. J., Evans, J. S., Hudak, A. T., et al. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, *113*, 2533–2546.

Masek, J. G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., et al. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, *112*, 2914–2926.

Max, T. A., Schreuder, H. T., Hazard, J., Oswald, D.D., Teply, J., & Alegria, J. (1996). The Pacific Northwest region vegetation and inventory monitoring system. *USDA Forest Service Research Paper PNW-RP-493*.

McGaughey, R. (2013). *FUSION/LDV: Software for LIDAR data analysis and visualization*. : US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

McRoberts, R. E. (2009). A two-step nearest neighbors algorithm using satellite imagery for predicting forest structure within species composition classes. *Remote Sensing of Environment*, *113*, 532–545.

McRoberts, R. E. (2010). The effects of rectification and Global Positioning System errors on satellite image-based estimates of forest area. *Remote Sensing of Environment*, *114*, 1710–1717.

Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L., & Hendrix, C. J. (2000). Predicting forest stand characteristics with airborne scanning lidar. *PE & RS-Photogrammetric Engineering & Remote Sensing*, *66*, 1367–1371.

Meigs, G. W., Kennedy, R. E., & Cohen, W. B. (2011). A Landsat time series approach to characterize bark beetle and defoliator impacts on tree mortality and surface fuels in conifer forests. *Remote Sensing of Environment*, *115*, 3707–3718.

Moeur, M., Ohmann, J. L., Kennedy, R. E., Cohen, W. B., Gregory, M. J., Yang, Z., et al. (2011). Northwest forest plan — The first ten years (1994–2003): Status and trends of late-successional and old-growth forests. In F.S. U.S. Department of Agriculture, & Pacific Northwest Research Station (Eds.), Portland, OR: U.S. Department of Agriculture, Forest Service.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, *80*, 88–99.

Ohmann, J. L., & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research*, *32*, 725–741.

Ohmann, J. L., Gregory, M. J., Henderson, E. B., & Roberts, H. M. (2011). Mapping gradients of community composition with nearest-neighbour imputation: Extending plot data for landscape analysis. *Journal of Vegetation Science*, *22*, 660–676.

Ohmann, J. L., Gregory, M. J., & Roberts, H. M. (2013). Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. *Remote Sensing and Environment*., http://dx.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.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., et al. (2012). *Package 'vegan' — Community Ecology Package, version 2.0–4.*

Oliver, C. D. (1980). Forest development in North America following major disturbances. *Forest Ecology and Management*, *3*, 153–168.

Orr, E. L., Orr, W. N., & Baldwin, E. M. (1992). *Geology of Oregon.* IA: Kendall/Hunt Publishing Company Dubuque.

Pesonen, A., Maltamo, M., Eerikäinen, K., & Packalèn, P. (2008). Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *Forest Ecology and Management*, *255*, 3288–3296.

Pflugmacher, D., Cohen, W. B., & Kennedy, R. E. (2012). Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. *Remote Sensing of Environment*, *122*, 146–165.

Pierce, K. B., Lookingbill, T., & Urban, D. (2005). A simple method for estimating potential relative radiation (PRR) for landscape-scale vegetation analysis. *Landscape Ecology*, *20*, 137–147.

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*, 1901–1916.

R Core Team (2012). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Reese, H., Nilsson, M., Pahlén, T. G., Hagner, O., Joyce, S., Tingelöf, U., et al. (2003). Countrywide estimates of forest variables using satellite data and field data from the national forest inventory. *AMBIO: A Journal of the Human Environment*, *32*, 542–548.

Reutebuch, S. E., Andersen, H. -E., & McGaughey, R. J. (2005). Light detection and ranging (LIDAR): An emerging tool for multiple resource inventory. *Journal of Forestry*, *103*, 286–292.

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*, 2337–2352.

Schoonmaker, P., & McKee, A. (1988). Species composition and diversity during secondary succession of coniferous forests in the western Cascade Mountains of Oregon. *Forest Science*, *34*, 960–979.

Scott, C. T., Bechtold, W. A., Reams, G. A., Smith, W. D., Westfall, J. A., Hansen, M. H., et al. (2005). *Sample-based estimators used by the forest inventory and analysis national information management system.* The enhanced forest inventory and analysis program—National sampling design and estimation procedures, 43–67.

Soininen, A. (2004). *Terrascan user's guide.* Helsinki, Finland: Terrasolid.

Spies, T. A. (1991). *Plant species diversity and occurrence in young, mature, and old-growth Douglas-fir stands in western Oregon and Washington.* USDA Forest Service general technical report PNW-GTR-285. : Pacific Northwest Research Station.

Spies, T. A., Johnson, K. N., Burnett, K. M., Ohmann, J. L., McComb, B. C., Reeves, G. H., et al. (2007). Cumulative ecological and socioeconomic effects of forest policies in coastal Oregon. *Ecological Applications*, *17*, 5–17.

ter Braak, C. J. (1986). Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology*, *67*, 1167–1179.

Tomppo, E. (1991). Satellite image-based national forest inventory of Finland. *International Archives of Photogrammetry and Remote Sensing*, *28*, 419–424.

Tomppo, E., Goulding, C., & Katila, M. (1999). Adapting Finnish multi-source forest inventory techniques to the New Zealand Preharvest inventory. *Scandinavian Journal of Forest Research*, *14*, 182–192.

Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., & Katila, M. (2008). Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment*, *112*, 1982–1999.

Tucker, C. J., Grant, D.M., & Dykstra, J.D. (2004). NASA's global orthorectified Landsat data set. *Photogrammetric Engineering and Remote Sensing*, *70*, 313–322.

Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M. (1999). Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, *70*, 52–68.

U.S. Department of Agriculture Forest Service [USFS]. (2003). Forest inventory and analysis national core field guide, vol. 1. *Field data collection procedures for phase 2 plots, version 2.0. Washington, DC.*

U.S. Department of Agriculture Natural Resources Conservation Service [USDA-NRCS] (1995). *Soil survey geographic (SSURGO).* Data base: Data use information. Fort Worth, Texas: National Cartography and GIS Center.

U.S. Department of Agriculture Soil Conservation Service [USDA-SCS] (1993). State soil geographic data base (STATSGO). *Soil conservation service, U.S. Department of Agriculture. Miscellaneous publication no. 1492.* Washington, D.C: U.S. Government Printing Office.

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.

Wing, M. G., Eklund, A., John, S., & Richard, K. (2008). Horizontal measurement performance of five mapping-grade Global Positioning System receiver configurations in several forested settings. *Western Journal of Applied Forestry*, *23*, 166–171.

Woldendorp, G., Keenan, R., Barry, S., & Spencer, R. (2004). Analysis of sampling methods for coarse woody debris. *Forest Ecology and Management*, *198*, 133–148.

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., et al. (2008). Free access to Landsat imagery. *Science*, *320*, 1011.

Woodcock, C. E., Collins, J. B., Gopal, S., Jakabhazy, V. D., Li, X., Macomber, S., et al. (1994). Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. *Remote Sensing of Environment*, *50*, 240–254.

Wright, C. S., & Agee, J. K. (2004). Fire and vegetation history in the eastern Cascade Mountains, Washington. *Ecological Applications*, *14*, 443–459.

Zald, H. S., Spies, T. A., Huso, M., & Gatziolis, D. (2012). Climatic, landform, microtopographic, and overstory canopy controls of tree invasion in a subalpine meadow landscape, Oregon Cascades, USA. *Landscape Ecology*, *27*, 1197–1212.