Forest Structure Patterns across Crater Lake National Park from LiDAR Data

Final Report to the National Park Service
Cooperative Agreement PNW

Task Agreement J8W07110013
Cooperative Agreement H8W07110001

Van R. Kane[1], Calvin Farris[2], Jonathan T. Kane[1],
Miles LeFevre[1], Sean Jeronimo[1],
James A. Lutz[3], Derek J. Churchill[1]

[1] University of Washington
[2]National Park Service
[3]Utah State University

Contents

1.0 INTRODUCTION

In 2010, airborne light detection and ranging (LiDAR) data was acquired for the full area of Crater Lake National Park in Oregon. The data was acquired primarily to develop highly accurate ground models to support hydrological analyses. No effort was made to collect concurrent field data to support the analysis of the forests and other vegetation that cover most of the park.

Shortly after the acquisition of the data, ecologists in the National Park Service's fire management program funded the study described in this report to use the LiDAR data to study the park's forests. Many studies of forests using LiDAR data calculate traditional forest inventory metrics such as basal area or biomass, but this was not possible for Crater Lake because of the lack of field data for the vast majority of forest types across the park. Instead, this study has built on the work of Lefsky et al. (2005) and Kane et al. (2011; 2010a; 2010b) to use only the LiDAR data measure and describe forest structure.

This study has had three sets of goals:

I Explore and develop LiDAR classification methods

1. Develop methods to identify and describe the range of forest structures across the park using the LiDAR data as the only data source
2. Determine how well LiDAR data could be used to predict field plot metrics for a small number of plots where data was coincidently collected to study the effects of prior fires within small areas of the park

II Analyze patterns of forest structure across the park

3. Use the LiDAR classification methods to map forest structure across the park and associate forest structure with different forest zones and fire history
4. Determine the influence of the biophysical environment on patterns of forest structure and fire severity

III Deliver processed LiDAR forest structure metrics to support the management of the park

5. Document a core set of LiDAR metrics that can be useful for park management applications
6. Describe the larger set of metrics produced by the FUSION software package
7. Provide a layer identifying potential tall legacy trees
8. Provide a 1 m resolution LiDAR intensity map for potential mortality mapping

In addition to the analysis of the LiDAR data, a complete fire severity atlas (section 2.3) was developed for this project (in conjunction with USGS) and previously delivered to the Park managers and the Monitoring Trends in Burn Severity (MTBS) project to become part of their publicly available database.

Over the course of the five years of the study, the analytic methods have matured considerably from those originally conceived to the set described in section 2.4 of this study. In the meantime, intermediate analysis and processed sets of LiDAR metrics have been delivered to the Park's managers beginning in 2011. This report and the associated data files represent the final products for this study. The methods to map forest openings, tree clumps, and tree clusters

are new, and the interpretation of the results is still maturing. The authors of this report plan to work with the National Park Service's ecologists and other forest ecologists to continue to analyze the forest structure patterns and publish the results in a peer reviewed science journal.

2.0 METHODS

2.1 Study area

Crater Lake National Park (CRLA) (741 km^2) lies in southern central Oregon, USA, within the Cascades mountain Range. This area has a Mediterranean climate with precipitation ranging from 769 mm to 1833 mm (PRISM 19702000 climate normal estimates) that mostly falls as snow during the winter.

The park straddles the crest of the Cascades range. West of the crest the climate is more mesic with precipitation increasing and temperatures decreasing with increasing elevations. At higher elevations, colder temperatures and deep snow packs constrain the growing season leading to energy limited forests. The eastern portions of the park lie in the rain shadow of the crest and receive less precipitation leading to water limited forests. Large portions of the park have extensive flat terrains that permit cold air pooling and also have infertile soils.

2.1.1 Forest zones

The park is forested except where local soil conditions are too dry to support trees or at the highest elevations where forests are replaced by woodlands with scattered individual trees and small tree clumps. No completed, currently accepted map of field-verified forest types exist for the park at the time of this writing, and different studies have used different criteria for identifying and mapping forest types or assigning them to field plot data. The LandFire project (http://www.landfire.gov/index.php) provides a raster map (Existing Vegetation Type) that uses a model to predict likely vegetation map units based on the NatureServe's Ecological Systems classification (Comer et al., 2003) and Landsat images. However, after review of the LandFire data set (C. Farris, personal communication), we decided that while the LandFire vegetation cover map was likely accurate in broad trends, it contained too much fine scale ("salt and pepper") granularity to be used directly and was inaccurate for several cover types.

For this study, we developed our own forest types classification using a combination of predicted forest composition maps from the LandFire project, patterns of precipitation and temperature from PRISM (Daly et al., 2008), and expert local knowledge (C. Farris, personal communications) (Fig.1). The goal was to relate patterns of forest structure to both patterns of potential biomass represented by mean precipitation and temperature as well as specific composition types relevant to park management goals. The resulting map of forest zones corresponds to local knowledge and to visual field checks at over 100 locations. However, because the map has not been formally verified through a rigorous field methodology, it should be used as an interpretive tool and not used for formal hypothesis testing.

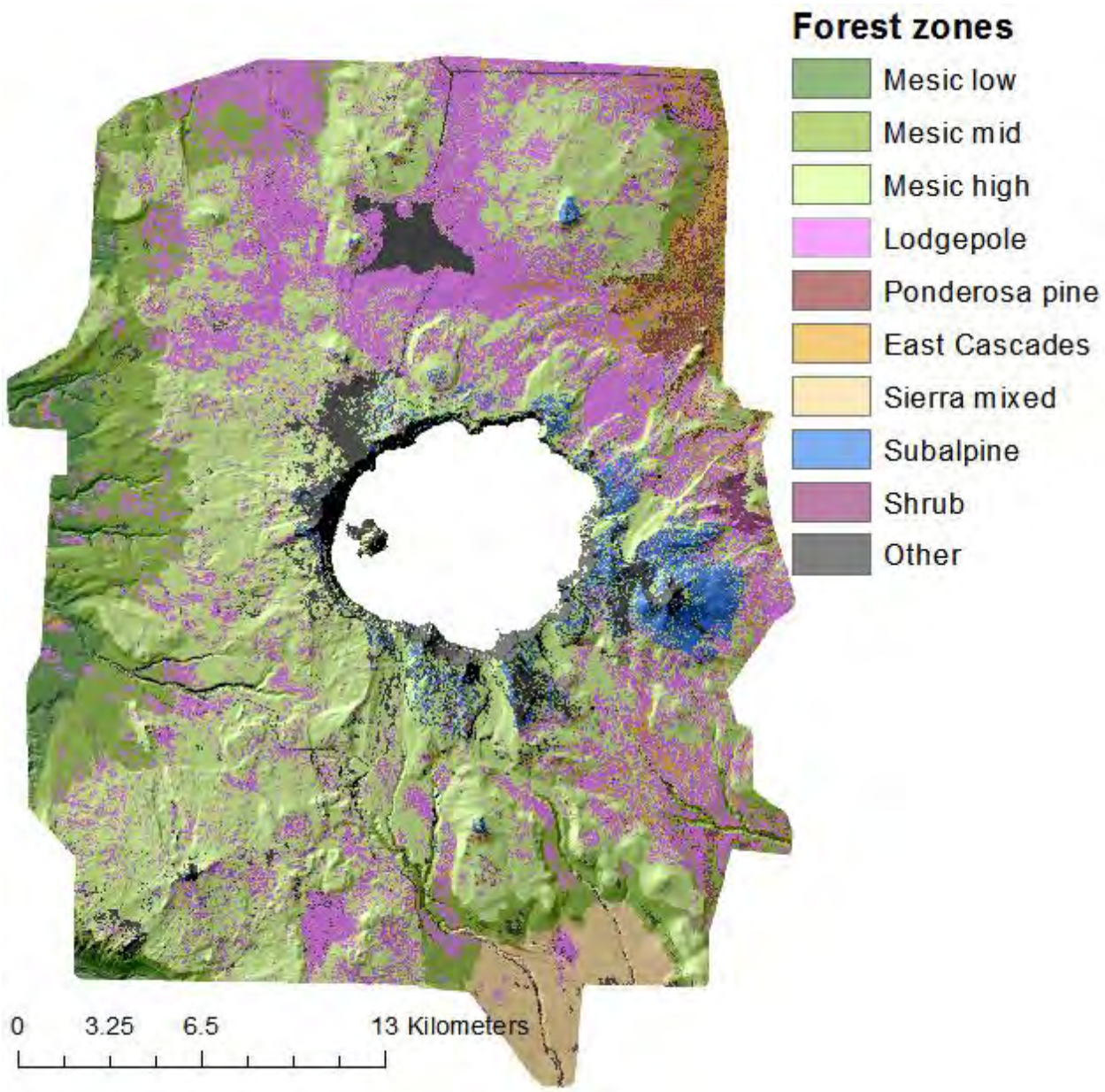


Figure 1. Forest zones identified for this study.

We chose three elevation bands that generally matched broader patterns of dominant tree species predicted by the LandFire 2010 vegetation cover map (LANDFIRE.US_130EVT): Low, < 1550 m; mid, 1550-1700 m; and high >1700 m. We further subdivided the low and high bands into mesic and xeric forests. C. Farris with the Park Service edited the resulting maps using local knowledge to identify a subalpine class from the xeric high class, refine mapping of the mono-specific lodgepole pine class largely from the xeric high class, and divide the xeric low class into an East Cascades, ponderosa pine class, and a Sierra mixed conifer class. Table 1 lists the

corresponding common tree species associated with each class from the LandFire vegetation cover map.

Table 1. Cover types used in this study with common tree species for each type.

Cover type	Common tree species
Mesic low	White fir, Engelmann spruce, Douglas fir, grand fir
Mesic mid	Mountain hemlock, Shasta red fir, lodgepole pine
Mesic high	Mountain hemlock, Douglas fir; eastern: Shasta red fir
Subalpine	Whitebark pine
Lodgepole	Lodgepole pine
East Cascades	Ponderosa pine, white fir, lodgepole pine
Ponderosa pine	Ponderosa pine, lodgepole pine
Sierra mixed conifer	Ponderosa pine, white fir

2.1.2 Fire and tree establishment patterns

Lightning ignited fires occur several times a year and are distributed throughout the park. Prior to approximately 1900, fire regimes differed between the mesic and xeric forests. In the more mesic western forests that correspond to our mesic forests, Forrestel and Stephens (2011) found a median fire return interval 37.5 years. In the mesic high forests, deep snow packs limit the fire season and fires typically burn at low severity with small pockets of high severity (Forrestel and Stephens, 2011). The xeric forests historically experienced more frequent fires with return intervals of 9-42 years (McNeil, 1975). Most historic fires appears to have burned with mixed severities and small patches of high severity fire, although episodic large high severity fires could have replaced larger areas of forests particularly in the lower mesic forests (Agee, 1993).

The results of these fire patterns when combined with other fine scale disturbance processes such as windthrow, disease, and insects was to maintain mixed-age stands across much of the park. In the mesic forest types, stands were likely structured as nearly continuous forests with clumps of trees of similar ages and enclosed gaps (Franklin and Van Pelt, 2004; Franklin et al., 2002). In the xeric forest types, most stands likely were structured as mixed-age tree clumps separated by open areas (Heyerdahl et al., 2014; Larson and Churchill, 2012; Shuffield, 2010).

Since approximately 1900, the structures of the forests have changed significantly compared to earlier patterns based on both studies within the park and similar forests within the region. First, park managers suppressed wildfires for most of the 20th century which has allowed less fire tolerant tree species such as lodgepole pine and white fir to infill within the xeric forests creating more continuous vertical and horizontal canopies than was the historic (Forrestel and Stephens, 2011; Shuffield, 2010; Taylor and Solem, 2001). Second, a large increase in the successful establishment of true fir within the lower mesic forests occurred in the decades surrounding 1900 likely leading to more dense forests than was the historic norm (Bekker and

Taylor, 2010; Forrestel and Stephens, 2011; Taylor, 1995). This second trend may be related to both fire suppression and likely favorable warmer temperatures and increased precipitation around 1900 (Forrestel and Stephens, 2011). The decades around 1900 also saw the establishment of stands of mono-specific lodgepole in areas of stronger cold air pooling and poorer soils (Forrestel and Stephens, 2011). It is unclear whether these stands were a new development for the park enabled by fire suppression or represent re-establishment of previous such stands following previous fires or widespread insect infestations. These stands today are experiencing high rates of mortality from mountain pine beetles. Only the higher elevation mesic forests appear to have seen relatively little structural change since 1990 (Forrestel and Stephens, 2011).

2.3 Fire atlas

We mapped estimated burned severities for fires from 1984 (the earliest date for data from the Landsat Thematic Mapper instrument) to 2010 using the differenced Normalized Burn Ratio (dNBR) (Key and Benson, 2006). dNBR estimates the effects of fire on the abiotic environment and vegetation, including the immediate impacts of the fire and ecosystem responses up to a year post-. Higher dNBR values signify a decrease in photosynthetic materials and surface materials holding water and an increase in ash, carbon, and exposed soil. dNBR values have been found to correlate with field-measured burn severities (NPS data, on file at Crater Lake National Park).

We obtained fire perimeters from the Park Service (C. Farris, personal communication) for fires 40 ha and larger. We selected pre-fire and one year post-fire Landsat scenes (30×30 m resolution) and calculated dNBR for each fire. Where the satellite data indicated that the fire perimeter records did not accurately correspond to the actual fire perimeter, we adjusted the fire perimeter to match the Landsat data. We provided our results to the MTBS project for inclusion in their database.

2.4 Forest structure measurements

2.4.1 LiDAR data

LiDAR data were collected by Watershed Sciences, Inc. (Corvallis, OR; later merged into Quantum Spatial) using three aircraft carrying a mixture of single Leica ALS60 sensors with a nominal flight altitude of 900 m above ground level with a scan angle of ±14° and dual mounted Leica ALS50 Phase II sensors with a nominal flight altitude of 1300 m above ground level with a scan angle of ±13º. Data were collected from August 23, 2010 to September 5, 2010 with an average pulse density of 8.39 pulses m^{-2} and 1.63 ground returns m^{-2} with up to four returns per pulse. Watershed Sciences created a 1 m resolution LiDAR-derived digital terrain model (DTM) using the TerraScan (v.10.009 & v.11.009) and TerraModeler (v.10.004 & v.11.006) software packages (Terrasolid, Helsinki, Finland). We processed the LiDAR data using the USDA Forest Service's FUSION software package (beta version derived from version 3.2, http://forsys.cfr.washington.edu/fusion.html) (McGaughey, 2014). All vegetation analyses were done with LiDAR return elevations normalized to height above ground to reflect height of vegetation canopy above ground.

2.4.2 FUSION metrics of vertical structure

We produced a set of vertical forest structure metrics from the LiDAR return data that were measured across the park using a 30×30 m (0.09 ha) grid cells (Fig. 2). Metrics for the distribution of the canopy profile were calculated as the heights at which a percentile of returns (e.g., 95th percentile height) >2 m occurred. We measured the structural heterogeneity of the forest canopy with a measure of canopy rugosity, rumple, calculated from a canopy surface model (CSM) created using the maximum return height within each 1-m grid cell smoothed with a 3 × 3 low pass smoothing filter. Canopy cover was calculated as the percentage of returns in a stratum divided by the number of returns in that stratum and all lower strata for strata >2, 2 to 8, 8 to 16, 16 to 32, and >32 m. For canopy cover >2 m, the 95th percentile of return heights (a surrogate for dominant tree height), and rumple, we also calculated values using a 3×3 smoothing algorithm in which the average value of the nine cells of a moving window was assigned to each center grid cell. This allowed us to assess trends in these metrics at both a 0.09 ha and a 0.81 ha grain.

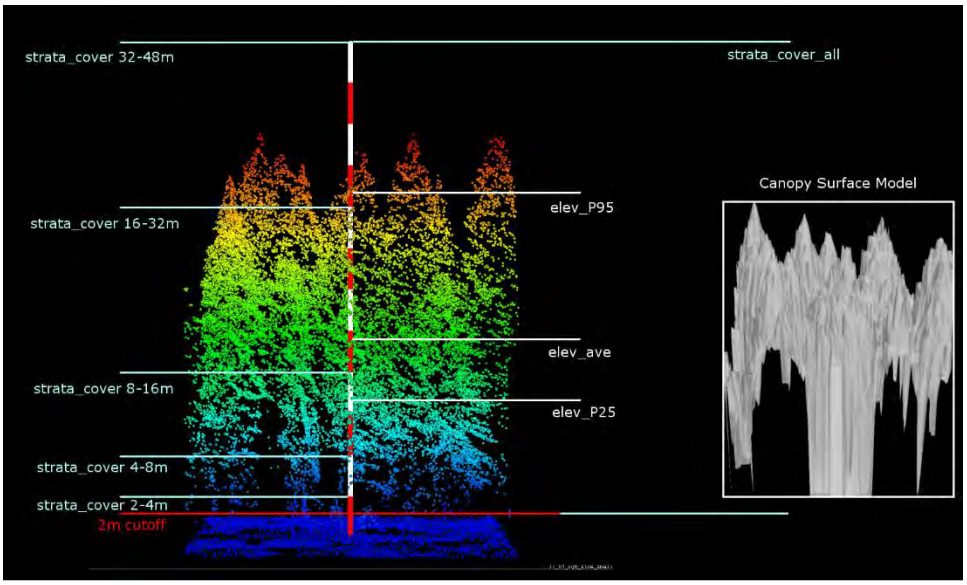


Figure 2. Illustration of key vertical structure metrics mapped across Crater Lake National Park. All of these metrics were calculated from the LiDAR return data except for rumple which was calculated from a 1 m canopy surface model. Cover and height return values were calculated for returns >2 m above the ground. Metric names shown are the same as the names of the corresponding rasters supplied with this final report.

2.4.3 Identification of tree approximate objects

In tree segmentation, an attempt is made to partition the LiDAR data into segments representing individual trees based on the geometry of the canopy surface model (Breidenbach et al., 2010). Although many studies have sought to perfect tree segmentation algorithms (e.g., Kaartinen et al., 2012; Vauhkonen et al., 2012), virtually all have come across the same major limitation: only overstory trees with direct visibility from above are reliably identified (Jeronimo,

2015). This includes the traditional classification of canopy dominants and co-dominants but also includes any tree directly visible from above with no overtopping trees regardless of the tree's height (which we refer to as trees that are immediately dominant within their horizontal space).

Acknowledging these limitations, we approach tree segmentation from the standpoint that each segment does not represent an individual tree per se, but instead represents an area of connected canopy that may encompass one to several trees. We have termed these canopy units tree-approximate objects (TAOs). Each TAO is typically composed of one immediately dominant tree and zero or more subordinate trees. The apex of the dominant tree is visible to the airborne LiDAR instrument, while the apices of the subordinate trees are hidden within or under the dominant tree's foliage. Despite not being individually delineated, the subordinate trees still do have a signature in the LiDAR point cloud, and metrics based on the vertical distribution of return heights (section 2.4.2) reflect their presence.

We carried out the segmentation using TreeSeg, a prototype tool developed for the FUSION LiDAR software package (http://forsys.cfr.washington.edu/fusion.html). This tool partitions the CSM into segments using the watershed transform (Vincent and Soille, 1991) with areas without canopy >2 m identified as openings. The watershed transform mathematically inverts the CSM and treats it as a topographical landscape: each distinct (endorheic) basin in the upturned CSM is taken to be a TAO in the unturned CSM (Fig. 3). We selected the watershed transform because it is based purely on the morphology of the canopy surface. Because it does not require assumptions or prior knowledge about crown widths, tree allometry, forest density, etc., it is well-suited to consistent performance across a wide range of forest structural conditions such as those found across the park. A study of the accuracy of tree identification across a wide range of forest structure types in the Sierra Nevada showed that TreeSeg correctly identified >80% of the largest trees that dominate stand structure (Fig. 4) (Jeronimo, 2015).

Because high resolution CSMs covering large study areas can be too large for many programs to load and process, the CSMs for the park were divided into 452 tiles. For each tile, we identified TAO's and openings using a 0.75 m canopy surface model (CSM) in which the height for each grid cell is the height of the highest LiDAR return within the grid cell area. For each tile, the TreeSeg software produced 1) a raster that identified the area of each TAO with an identifier unique to the tile and openings as a categorical class, 2) a raster that assigned the maximum height of each TAO to the grid cells within its area with openings as a categorical class, and 3) and ESRI ArcGIS shapefile showing the location of each TAO maximum height, its maximum height, and the area assigned to each TAO.

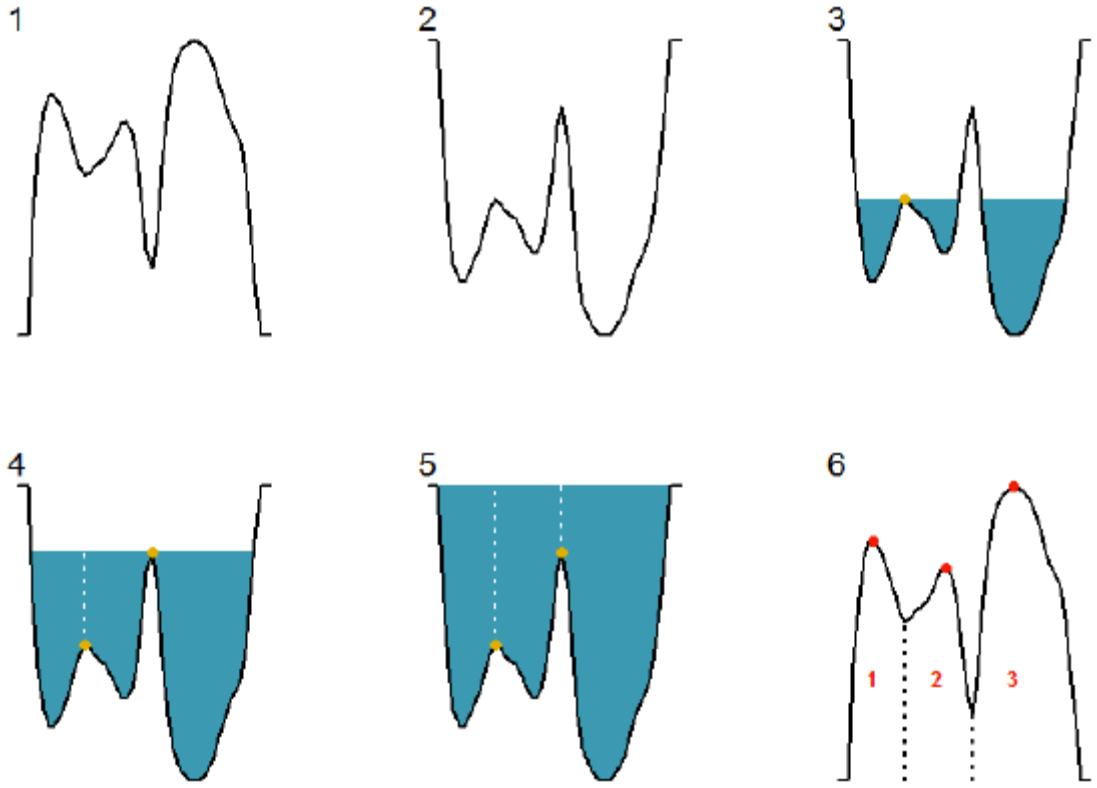


Figure 3. Overview of the watershed segmentation algorithm, presented here as a 2D concept, but actually performed in 3D. (1) The canopy surface model is draped over the LiDAR point cloud. (2) The canopy surface model is inverted. (3-5) The surface is imagined to be made of a permeable material, and is slowly lowered into water. Any time two separate pools come into contact (green points), a dam (white dashed lines) is formed. (6) The canopy surface model is righted. The dams are taken to be region boundaries, and high points within each region so defined are taken to be the treetops (red points). From Jeronimo (2015).

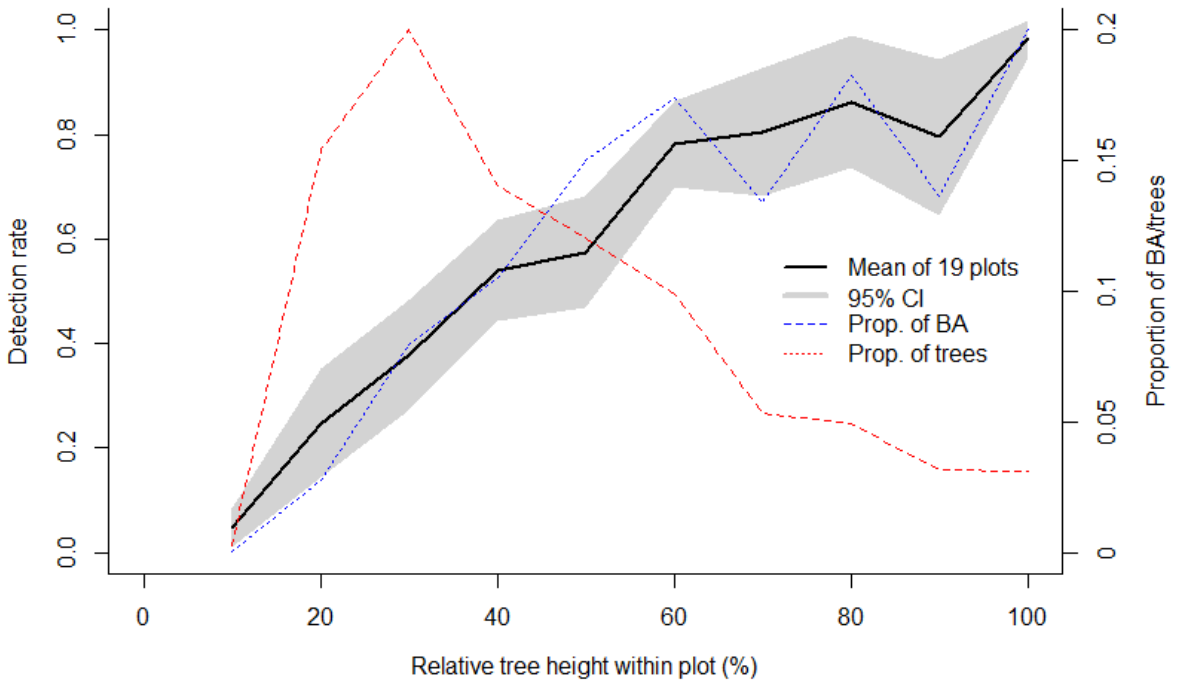


Figure 4. Detection rate for trees from 19 plots representing a wide range of forest structures in the Sierra Nevada (Jeronimo, 2015). Detection rates were generally poor for trees that were less than 60% of the maximum height in each plot. Because most trees were smaller than this, the accuracy rate as a percentage of the number of trees detected was low but was high as a percentage trees representing the majority of the basal area. Similar accuracy rates are expected for the forests of Crater Lake National Park.

2.4.4 Landscape analysis of tree clump and opening patterns

Tree clumps and openings are fundamental structures found in a wide range of forests (Attiwill, 1994; Franklin et al., 2004; Larson and Churchill, 2012; Muscolo et al., 2014). Because the stem mapping necessary to study tree clump patterns is expensive, studies of tree clump patterns have been rare and generally limited to plots ranging in size from less than a hectare to a few hectares (Larson and Churchill, 2012). The use of LiDAR data provides the opportunity to identify and examine these patterns across a wide range of forest types and spatial areas. We developed methods for this study to examine these patterns.

Because characteristic clump and opening patterns typically are expressed at scales from 0.4 ha to 1 ha (Larson and Churchill, 2012), we used 90×90 m (0.81 ha) areas to measure patterns. The analysis used a gliding window approach in which the TAOs and openings in a 0.81 ha area surrounding the center of a 30×30 m grid cell were analyzed. The center point was then moved to the next grid cell and the analysis was repeated, resulting in analyses that included openings and TAOs that overlapped with those used for adjacent grid cells. We did this so that

the analysis grain matched the 30×30 m grain of the data sets used in this study and commonly used by the Park Service. The values recorded for each 30×30 m grid cell, therefore, can be thought of as the equivalent of a 3×3 smooth of the tree clump-opening context centered on a 30 m grid. However, we spaced samples used in the statistical analyses far enough apart that the areas used for each sample did not overlap with the area for any other sample.

We identified three types of structures derived from the canopy surface model within each 90×90 m area (Fig. 5): openings, tree clusters by height strata, and tree clumps. As a preliminary step, we reclassified the 0.75 m rasters that recorded the maximum height for each TAO and openings (section 2.4.3) into height strata: <2, 2-8, 8-16, 16-32, >32 m. Openings were contiguous patches of CSM grid cells with heights <2 m, and clusters were contiguous patches of TAO's in the same stratum >2 m. Finally, we identified tree clumps as one or more adjacent tree clusters separated by an opening.

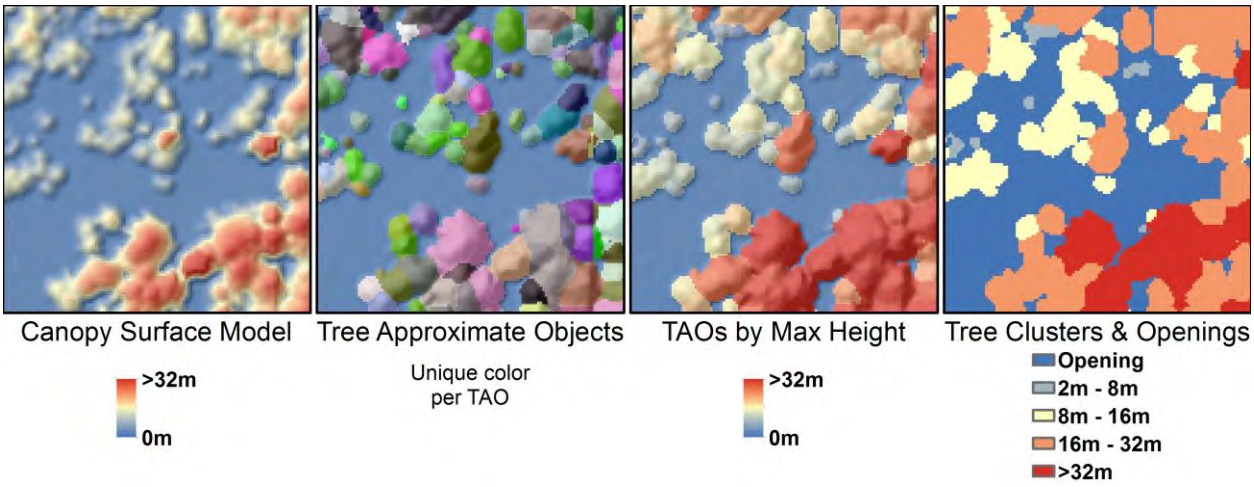


Figure 5. Identification and measurement of tree clusters, tree clumps, and openings from LiDAR-derived5canopy surface models was done in four steps: 1) create a 0.75 m resolution canopy surface model, 2) identify immediately dominant tree approximate objects (TAOs) using a watershed segmentation, 3) assign the area of each TAO the maximum height of that TAO, 4) classify each TAO by height strata (2-8, 8-16, 16-32, and >32 m; openings are <2 m). Tree clusters are one or more adjacent TAOs in the same height strata while tree clumps are one or more adjacent tree clusters separated from other tree clumps by openings.

In 90×90 m areas with essentially contiguous cover, we typically identified a single contiguous tree clump consisting of multiple tree clusters and enclosing several openings (which in this context would correspond with the usual definition of canopy gaps). In areas dominated by open areas, we typically identified one or two openings that surrounded one or more tree clumps that each consisted of one or more tree clusters. Tree clumps composed of multiple tree clusters likely result from multiple phases of tree establishment following fire suppression or disturbance processes such as fire or wind throw that resulted in gaps that subsequently filled through regeneration.

The result of this processing was to create a classified raster for each 90×90 m sample area that could be analyzed for patterns of openings, tree clusters, and tree clumps using standard landscape ecology metrics such as those included in the FRAGSTATS software package (McGarigal et al., 2012). We treated each opening, tree cluster, and tree clump as a separate patch. Grid cells in the same height strata were treated as members of the same landscape class.

We investigated numerous landscape metrics to quantify the patch, cluster, and clump patterns and selected a parsimonious set. The total area in each height stratum class (FRAGSTATS *class area* metric) measured the relative dominance of each stratum in each sample. We measured the structural complexity of the arrangement of patches using the number of patches of openings, tree clusters, and tree clumps (FRAGSTATS *patch number* metric). Because the number of potential canopy clusters would change with the area of canopy in each sample, we also computed a normalized cluster ratio as the number of clusters divided by the total area in canopy >2 m. (The fixed 0.81 ha area of each sample made the number of openings and tree clumps the equivalent of normalized values.) We also used the total length of edges (FRAGSTATS *total edge* metric) between patches (including the fixed length of the edge around each sample) as an additional measure of structural complexity. We explored using several measures of aggregation (the FRAGSTATS *aggregation, contagion, dispersion, and clumpiness* metrics), but did not use them after finding that they poorly differentiated samples possibly because these are cell-based metrics and openings, tree clusters, and tree clumps are highly aggregated at the cell level. We also explored using a number of additional metrics but dropped them because they either were highly correlated with metrics in the parsimonious set or did not usefully differentiate among patterns found in the samples.

We used a custom Python routine to process the CSMs and TAOs by tile, compute landscape metrics for 90×90 m areas centered on each 30×30 m grid cell, and assemble single 30×30 m rasters covering the entire park for each metric. The landscape metrics were computed using the same formulas as those used by the FRAGSTATS software as checked by computing the metrics using both our software and the FRAGSTATS software on a set of samples and ensuring that the results were identical.

2.4.5 Structure classes

We tested multivariate combinations of our parsimonious landscape metrics to define structure classes to identify and map the patterns of openings, tree clusters, and tree clumps across the park. We started with the metrics for class area for openings and each tree cluster height stratum and then added metrics for number of tree clumps, the normalized cluster ratio, and total edge to determine if they materially changed the classes defined. We found that using the additional metrics did not meaningfully change the classes defined using only class area, and the classes we defined were based only on the class areas.

We used the five class area (representing openings (CSM <2 m) and TAOs with height maxima in the 2-8, 8-16, 16-32, and >32 m strara) to define classes of vertical forest structure based on a random sample of 10,000 grid cells (11.7% of the study area). To account for collinearity between the metrics, we used the principle components analysis (PCA) (Legendre and Legendre, 1998) axes of variation to define the structure classes. We used hierarchical clustering to split the sample set into nine statistically distinct classes based on the PCA axes of variation values. We used Euclidean distances and Ward's linkage method within the ''hclust''

function of the R statistical package (release 2.6.1) (R Core Team, 2014) for this analysis. We used the classified random sample of 10,000 grid cells as training data to classify the vertical structure of all grid cells within the study area using the random forest algorithm (Breiman, 2001) predict function in the R statistical randomForest package (Breiman et al., 2012)

2.5 Biophysical environment

2.5.1 Topographic metrics

Topographic metrics were calculated based on terrain within moving windows centered on the 30×30 m grid cells. We could not know a priori what scales of topography would best correlate with forest structure and fire patterns, and we calculated each metric at multiple scales. For example, starting with the center point of each grid cell, we measured each topographic metric at scales of 30 m, 90 m, and 270 m for the area surrounding each grid cell (Fig. 6). The metrics were calculated using the LiDAR-derived 1 m digital terrain model (DTM) within the park and using the US Geological Survey 10 digital elevation model when the extent of the moving windows moved beyond the LiDAR data. Calculations were done using the FUSION software topometrics tool.

We calculated slope, aspect, and a solar radiation index (SRI) for scales of 30 m, 90 m, and 270 m. The topometrics tool uses a 3×3 grid of points spaced to calculate the slope, aspect, and curvature (Zevenbergen and Thorne, 1987). Slope is reported as degrees and aspect as the cosine of azimuth with south set to 1.

The SRI metric provides information comparable to heat load indices and combines slope, aspect, and latitude into a single value useful for comparing relative solar radiation loads across a study area and. SRI models solar radiation during the hour surrounding noon on the equinox (Keating et al., 2007):

$$SRI = 1 + \cos(latitude) \times \cos(slope) + \sin(latitude) \times \sin(slope) \times \cos(aspect) \quad (1)$$

where latitude and slope are in degrees and aspect is relative to south.

The topometrics tool calculates slope positions using annuli of 100 m, 250 m, 500 m, 1000 m, and 2000 m radii using an algorithm that replicates the Topographic Position Index (TPI) (Jenness, 2006; Weiss, 2001). We converted the results to standardized values. More negative TPI values indicate a position towards a valley bottom, values near zero indicate flat areas or mid-slope, and more positive values indicate a hill or ridge top.

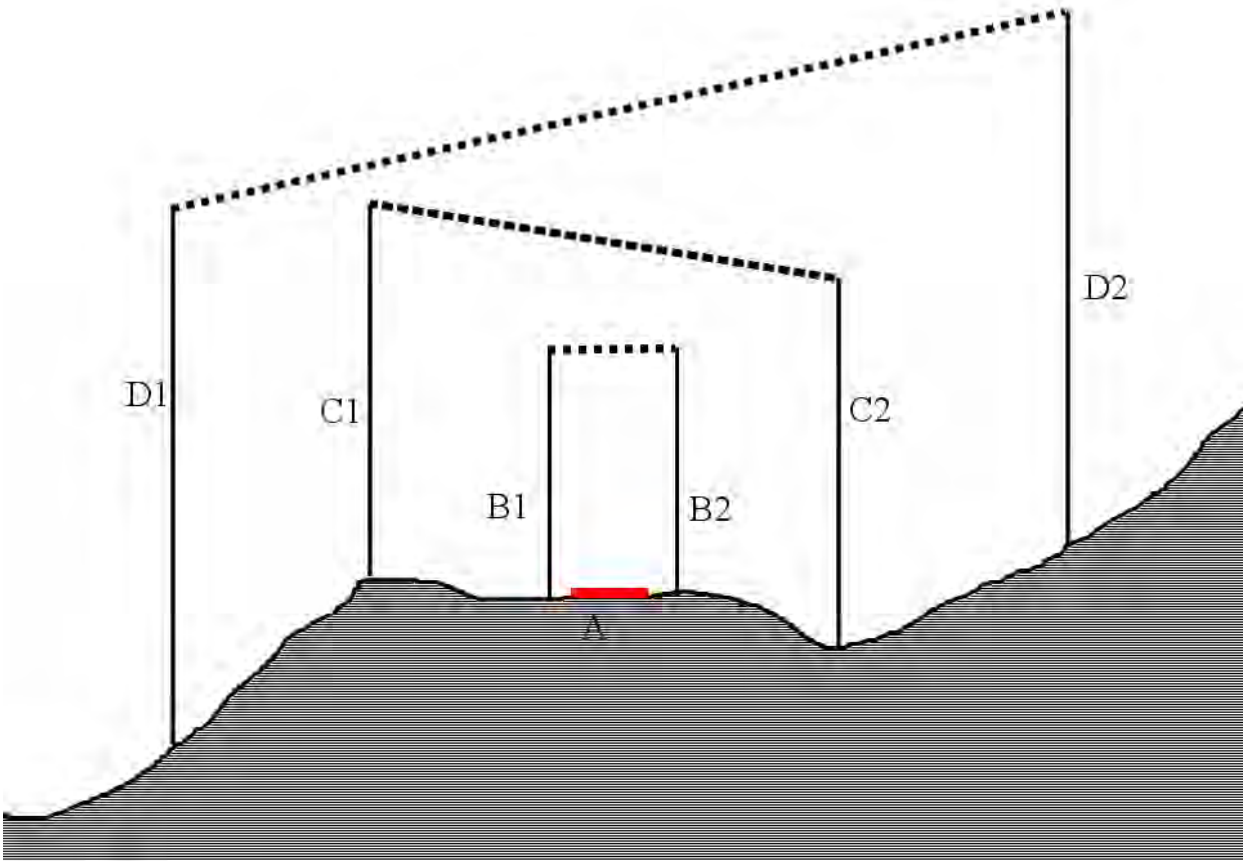


Figure 6. Illustration of the effects of window size on topography metric calculations. A profile view of an example landscape is shown above. Pixel (A) is calculated at three different scales. At the scale of the distance between lines B1 and B2, the slope for Cell A would be close to zero. At the scales of the distances between lines C1 and C2, and D1 and D2, the slopes would be roughly equal. However, the aspect for these two scales would be opposite.

2.5.2 Climate metrics

Actual evapotranspiration (AET) and climatic water deficit (Deficit) are correlated to the elevation gradient through patterns of precipitation (the western portions of the park and higher elevations generally receive more) and temperature (higher elevations are colder). AET is associated with potential biomass and potential surface fuels while Deficit is correlated with drought stress and fuel moisture (Kane et al., 2015; Miller and Urban, 1999). We used monthly climate normal data (1971–2000) for precipitation and temperature from PRISM (Daly et al., 2008) mapped at 30 arc-second (~800 m) resolution with a Thornthwaite-type (Thornthwaite and Mather, 1955) AET and Deficit calculation where potential evapotranspiration is based on temperature. We used the Dingman (2002) water balance algorithm that calculates AET and Deficit based on an exponential model of soil water depletion as implemented by Lutz et al. (2010) without their heat load modifier. To estimate soil water holding capacity within the top 200 cm, we used maps from the Natural Resources Conservation Service (SSURGO) database

(http://soils.usda.gov/survey/geography/ ssurgo/).

2.6 Random Forest Modeling

We used the random forest supervised learning algorithm (Breiman, 2001) to identify statistical relationships between fire severity, forest structure and several environmental and past management history variables. This algorithm models complex relationships between predictors and response, has the ability to work with large data sets, a wide range of data types, minimizes overfitting of data sets, and can accept spatial autocorrelation in predictor data (Breiman, 2001). This modeling is an extension of non-parametric classification and regression trees (CART) ; (Breiman et al., 1984).

Random forest modeling is used across a wide range of disciplines with increasing use in ecology (Cutler et al., 2007). Like linear regression modeling, model error-rates are reported as the proportion of total variance explained (i.e., R^2), which reports how closely model predictions are to a given response variable. Unlike most linear regression applications, the random forest algorithm calculates model error-rates by applying models to separate independent validation samples including out-of-bag samples (similar to cross-validation), or an independent test set.

Furthermore, the random forest algorithm reports the importance of each predictor, allowing us to identify the key predictor variables (e.g., harvest history) associated with fire severity patterns. We used partial dependence plots (Hastie et al., 2001) to examine the relationship of individual predictors to patterns of burn severity and forest structure.

We examined the effects of the biophysical environment on patterns of forest structure and fire severity using measurements of the biophysical environment as predictors (Table 2). To identify the key drivers of the of structure and burn severity, we selected a parsimonious subset of predictors that explained approximately 95% or more of the variance explained by the full set of predictors. We selected this subset by starting with the predictor reported as most influential from the run with all predictors and then testing the addition of all remaining predictors one at a time to see which best improved modeling results given previously selected parsimonious predictors. We iteratively repeated this process using previously selected predictors until no additional predictors improved variance explained by more than 2%.

Table 2. Response and predictor metrics used in this study to study fire and forest structure relationships to the biophysical environment using a parsimonious predictor set. All variables were mapped using 30 m (0.09 ha) grid cells although some were calculated using larger windows around each grid cell (calculation window sizes shown in parantheses) or in the case of AET and Deficit resampled from the original 270 m rasters and for AWS from a shape file.

Response metrics (scale)	Source	Units/interpretation
dNBR estimated burn severity (30 m)	MTBS[1]	relative burn severity
>95th percentile return height >2 m	LiDAR returns	meters
Class areas: 02 (open), 2-8, 8-16, 16-32, >32 m (90 m)	LiDAR TAOs[2]	percent of 0.81 ha sample area
Number of openings ≥10 m (90 m)	LiDAR TAOs[2]	count per 0.81 ha sample area
Number of tree clumps (90 m)	LiDAR TAOs[2]	count per 0.81 ha sample area
Tree cluster ratio (90 m)	LiDAR TAOs[2]	count normalized to canopy area per 0.81 ha sample area
Total edge (90 m)	LiDAR TAOs[2]	meters
Predictor metrics (scale)	Source	Units/interpretation
January mean temperature (JanT, 800 m)	PRISM[3]	ºC
Precipitation mean (PPT, 800 m)	PRISM[3]	mm water
Actual evapotranspiration (AET, 800 m)	Lutz et al. (2011)	mm water
Soil water capacity (AWS, 30 m)	SSURGO[4]	mm water
Climatic water deficit (CWD, 800 m)	Lutz et al. (2011)	mm water
Topographic position index (TPI, 4 km)	1 m LiDAR DTM & 10 m USGS[5] DEM Jenness 2006	standardized relative index
Solar radiation index (SRI, 135 m)	1 m LiDAR DTM & 10 m USGS[5] DEM	relative index

[1]Monitoring Trends in Burn Severity, [2]Tree Approximate Objects,[3]PRISM Climate Group, [4]USDA Soil Survey Geographic (SSURGO) Database,[5]United States Geologic Survey

3.0 RESULTS AND DISCUSSION

3.1 Relationship of biophysical environment to forest structure patterns

We investigated the relationship between the biophysical environment and forest structure using a parsimonious set of predictor metrics with random forest modeling. In selecting the specific best parsimonious predictor set for each forest structure metric, we found that a number of predictor metrics were common to most of the models. We selected a set of seven of these predictors as a common parsimonious predictor set and used them for our final modeling reported here. By selecting a common set of parsimonious predictors, variance explained for individual responses was typically somewhat less than if we had used a specific parsimonious set per response. However, using a common set allows easier comparison of relationships between biophysical predictors and the full set of responses.

Five of the predictors relate to the water balance, which describes the simultaneous availability of energy and water to support plant growth. Both AET and Deficit were selected as final parsimonious predictors. The three other predictors are key inputs used to model AET and Deficit: January mean temperature (which will be correlated with mean temperatures across the year), mean annual precipitation, and the soil water holding capacity. Since each forest structure metric to some degree reflects the effects of the water balance, the pattern of its values may have served to guide the random forest algorithm to effectively develop its own enhanced internal water balance model using these five predictors.

The other two parsimonious predictors described topography. The topographic position index at a 4 km scale reflects broad trends in slope position. Forest structure in this region commonly varies based on whether a stand is located in a valley bottom or other flat area, on a slope, or near a ridge, especially for more xeric locations (Bekker and Taylor, 2001; Taylor, 2000). The solar radiation index (SRI) at the 135 m scale integrates the effects of latitude, slope and aspect on the solar influx and is a similar measurement to the heat load index. The solar radiation influx effects local temperatures to impact local snow melting, soil moisture, and evapotranspiration rates.

We interpreted the relationships between predictors and responses using partial plots (Fig. 7). Each plot shows the relationship between a single response (e.g., 95th percentile of return heights) and a single predictor (e.g., AET) by varying the range of the predictor values while the values for all other predictors in the model are held to their mean values. As a result, these plots do not show interactions among predictors, and this can be reflected by an approximately flat trend line for a plot (e.g. precipitation for opening area). We report the overall variance explained by each model as a pseudo-R^2 and the normalized importance of each predictor in the model.

The seven predictors explained 61-62% of the variability in dominant tree height (95th percentile LiDAR return height >2 m), total area in openings (and also its inverse value, total area in canopy), and the total edge length of all tree cluster and opening patch edges. Dominant tree height and cover both are related to biomass. Both were positively correlated with increases in AET and the underlying factors that affect greater AET: warmer temperatures, greater precipitation, and greater capacity to store water in the soil.

The total edge length is a measure of local stand complexity and was most strongly associated with January mean temperature and AET. Peak complexity occurred at moderate values for these predictors, suggesting that both more favorable (warmer and higher AET) and unfavorable (colder and lower AET) lead to less complex stand structures. For the former case, this would be consistent with mesic conditions trending to nearly continuous canopies with relatively few but larger tree clusters. In the latter case, this would suggest that harsher conditions trending to nearly continuous openings with relatively few tree clumps and few clusters per clump.

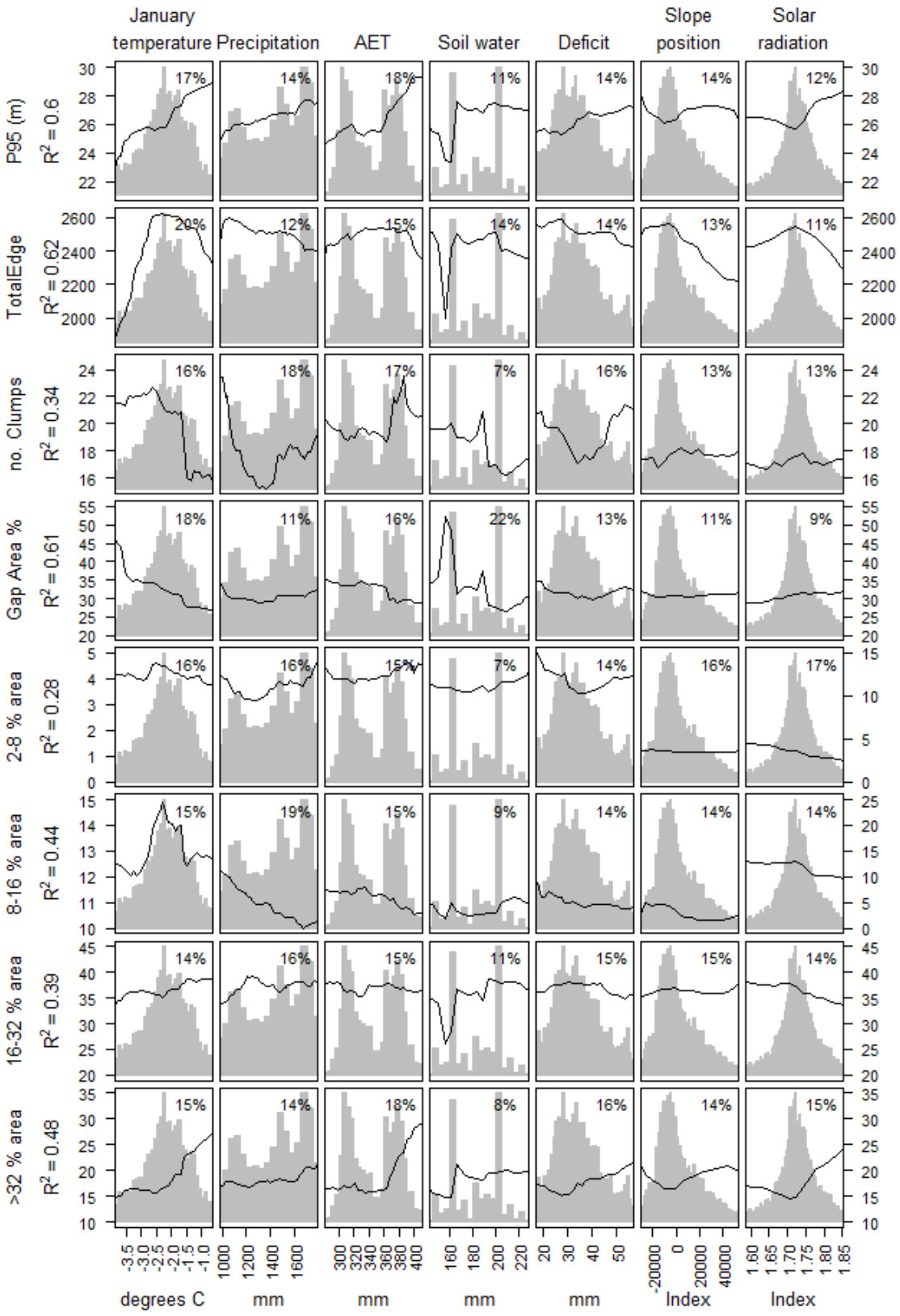


Figure 7. Partial plots for random forest models relating forest structure and estimated burn severity (dNBR) with a parsimonious set of biophysical predictors. P95 is the 95th percentile

return height, #Clumps the total number of tree clumps, the Cluster Ratio the number of tree clusters divided by the canopy area, and CA is class area for different strata. JanT is mean January temperature, PPT is mean precipitation, AET is actual evapotranspiration, AWS is the estimated soil water capacity in the top 1.5 m of soil, CWD is climatic water deficit, TPI 4km is topographic position index measured at a 4 km scale, and SRI 135m is the solar radiation index measured at a 135 m scale. Number within each panel shows the normalized importance of each predictor in the model (‘‘imp. =”). Variance explained by model shown as a pseudo R2 (‘‘RSQ =”). Solid lines show trends in RdNBR in response to each predictor (left scale) while histograms show the distributions of values for each predictor (right scale). Model trends where there are few samples at extreme values for predictors should be treated with caution. Relationships between each predictor and the response (RdNBR) calculated as partial dependence plots where the values for each predictor are varied throughout their range while values for all other predictors are held to their mean values. As a result, partial dependence plots do not show interactions between predictors.

The random forest models explained 39% to 48% of the variation in canopy area in tree cluster strata >8 m, but only 29% of the variation for the 2-8 m cluster stratum. Increased area in the >32 m and 16-32 m strata was related to predictors associated with increased AET and likely reflect the ability to grow taller trees with more mesic growing conditions. Increased area in the 8-16 m stratum, however, was associated with predictors indicating decreased AET and decreased Deficit and likely reflects that trees in this strata dominate in areas with poorer growing conditions. Area in the 2-8 m stratum shows no clear trends with the predictors, and the poor predictability may reflect either no clear biophysical relationships and or the limited ability of tree segmentation algorithms to detect TAOs in this stratum.

The random forest modeling found that the predictability (44%) of the number of clusters per clump (our cluster ratio that normalized this value to the area in canopy) was similar to that for the area in strata >8 m. This again suggests only a moderate relationship between the biophysical environment and the complexity of tree arrangement reflected in tree cluster patterns within a clump.

The relationship between the biophysical environment and the number of tree clumps was relatively weak (34% of variance explained). The relationships between the predictors related to water balance and the number of tree clumps showed abrupt transitions suggesting some unimodal relationships between the environment and clumping patterns.

The ability to predict the dominant elements of local stand structure (dominant tree height, area in openings/canopy, and overall complexity) was higher for than for canopy area by tree cluster stratum. This suggests a moderately strong relationship between the biophysical environment and overall patterns of stand structure. The lower predictability of area in each tree cluster stratum may reflect the results of local stochastic disturbances that create locally random patterns of tree retention and mixed age regeneration. Similarly, the overall poor predictability of the number of tree clusters and clumps may reflect local processes dominating how the area in canopy is divided into distinct clumps when the canopy is not continuous. While ecologists have developed a good understanding of gap formation and subsequent regeneration (which leads to

our tree cluster patterns), we are not aware of any work that examines the processes that determine the number of tree clusters and clumps. The methods and patterns presented in this work may form the basis for developing hypotheses to form the basis of future work both in the field and using spatially explicit forest gap and tree regeneration models.

3.2 Relationships of Key Forest Structures

We found that the area in openings was related to patterns of the number of openings >10 m^2, number of tree clumps, and the tree cluster ratio (which is the number of tree clusters divided by the total canopy area) (Fig. 8). The majority of the forested areas in the park had less than 50% of their area in openings. At very low (<10%) and high (>50%) area in openings, the number of openings >10 m^2 typically was less than five per 0.81 ha area. At the low end, this would reflect nearly continuous canopy while at the high end this would reflect nearly continuous openings. Between approximately 10% and 30% opening area, the number and range of openings >10 m^2 sharply increased suggesting a pattern of many small openings rather than the formation of larger openings as canopy area modestly decreased.

We found that the number of tree clumps (one or more tree clusters separated from other clusters by an opening) was low and flat up until approximately 20% area in openings and then increased approximately monotonically up to 50% area in openings. Beyond this latter area in openings, there was a wide diversity in tree clump numbers but the strongest trend was to fewer tree clumps. This wide range of tree clump numbers for >50% area in openings suggests multiple processes at work that may either reflect different biophysical environments or a wide range of localized stochastic processes.

The number of tree clusters (our cluster ratio that normalized this value to the area in canopy) reflected the complexity of tree clumps divided into groups of trees with different heights that likely reflect within tree clump gap-phase regeneration. The cluster ratio similarly showed an approximately monotonic increase with increasing area in openings up to approximately 50% area in openings. At higher percentages of opening area, the cluster ratio trended toward lower values indicating a gradual trend toward less complex tree clumps.

Taken together, these results show that stands are most complex in terms of the number openings, tree clumps, and tree clusters peaks when the area in openings is approximately 20-40%. Stands were least complex when openings are less than 10% or greater than 50% of the area. The lower complexity for stands with openings >50% was expected since these generally represent areas with poor growing conditions (high elevation or xeric) and therefore the ability to grow complex stands would be less.

The lower complexity for opening areas less than 10% was surprising. Chronic small-scale disturbances could have broken the nearly continuous canopies into numerous gaps and tree clusters representing regeneration patches of different heights and ages. One possible explanation is that the locations with <10% openings tend to be those with be best growing conditions (e.g., highest AET) and trees quickly grow to reach the tallest stratum we measured, >32 m, as suggested by the strong relationship between growing conditions and area in tree clusters >32 m. However, the area in the tree clusters >48 m was so low (<2% over the entire forested area of the park) that it's not clear that using higher strata breaks would reveal past

patterns of gap-phase regeneration.

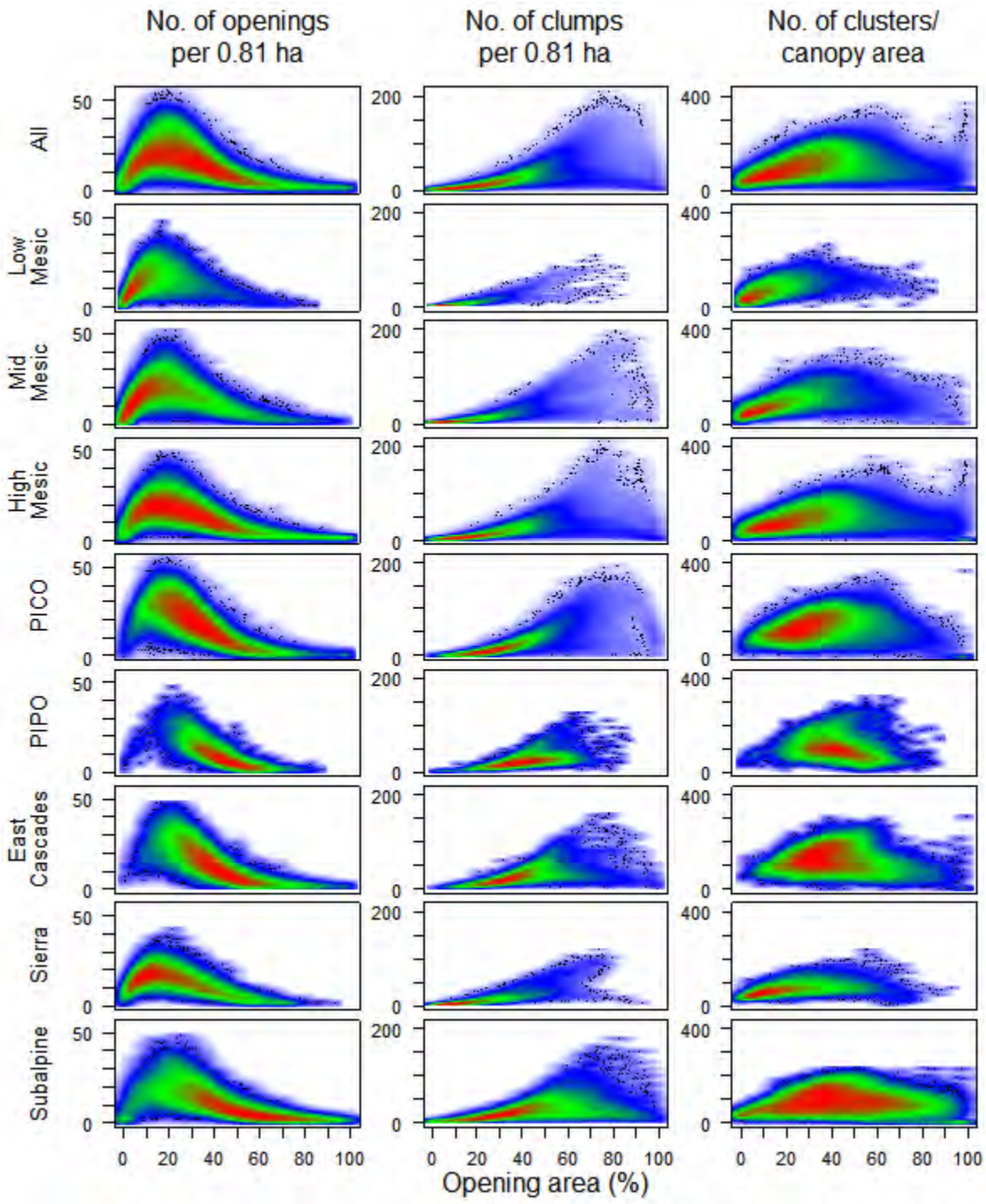


Figure 8. Changes in the number of openings >10 m^2 and tree clumps per 0.81 ha sample area and the number of tree clusters per area of canopy in each sample ('cluster ratio') by the percentage of area in openings.

The trend towards correlated increases in complexity for openings, tree clumps, and tree clusters at moderate (20-40%) opening areas suggests underlying ecological processes. Stands with this range of open area are widespread across the park and therefore represent a range of ecological conditions and possibly processes. Considerable work has looked at the effects of chronic disturbances on gap-phase regeneration with implications on the resulting complex patterns of tree clusters. Other work has looked at the effects of variations of fire intensity patterns on tree mortality with resulting spatial heterogeneity of openings, tree clumps, and regeneration. However the long exclusion of fire and limited time for regeneration to advance for recent fires may limit the effect of this process on the patterns we observed. Insect mortality currently is widespread among the monospecific lodgepole pine stands in the park with surviving trees frequently found in small clumps or individual trees. These stands, however, tend to have nearly uniform heights and their mortality patterns would not tend to increase the ratio of tree clusters. It may be that intermediate growing conditions, such as those associated with our samples with 20-40% open area, more frequently experience or are more sensitive to processes that create local stand complexity similar to the effects of intermediate disturbance on local species richness.

A fundamental problem with identifying and quantifying the effects of ecological processes on stand scale structural complexity is that few studies have quantified spatial patterns and only over relatively small areas. Methods such as ours could be used to identify broader patterns and form hypotheses that could then guide stratified field sampling to elucidate underlying processes.

3.3 Structure classes and their patterns

We identified nine statistically distinct forest structure classes using the area of openings (CSM <2 m) and area of TAOs in four height strata (2-8, 8-16, 16-32, >32 m) within 90×90 m areas across the park (Fig. 9, 10 ,11, 12). Eight of the structure classes were in statistically similar pairs of classes and we labeled these pairs with sequential numbers and separated numbering for pairs by units of ten.

We found that the structure classes could be interpreted based on the dominant tree cluster height strata, the vertical distribution of canopy area across height strata, and the horizontal arrangement of canopy and openings (Table 3). In our analysis, we used both the class area metrics as well as several auxiliary metrics not used to define the classes that measure aspects of structure: the number of tree clumps, the number of openings >10 m^2, the ratio of tree clumps to canopy area, the total length of edges between patches, dominant tree height estimated by the 95th percentile of LiDAR return heights >2 m, and canopy cover >2 m estimated from the LiDAR return cloud.

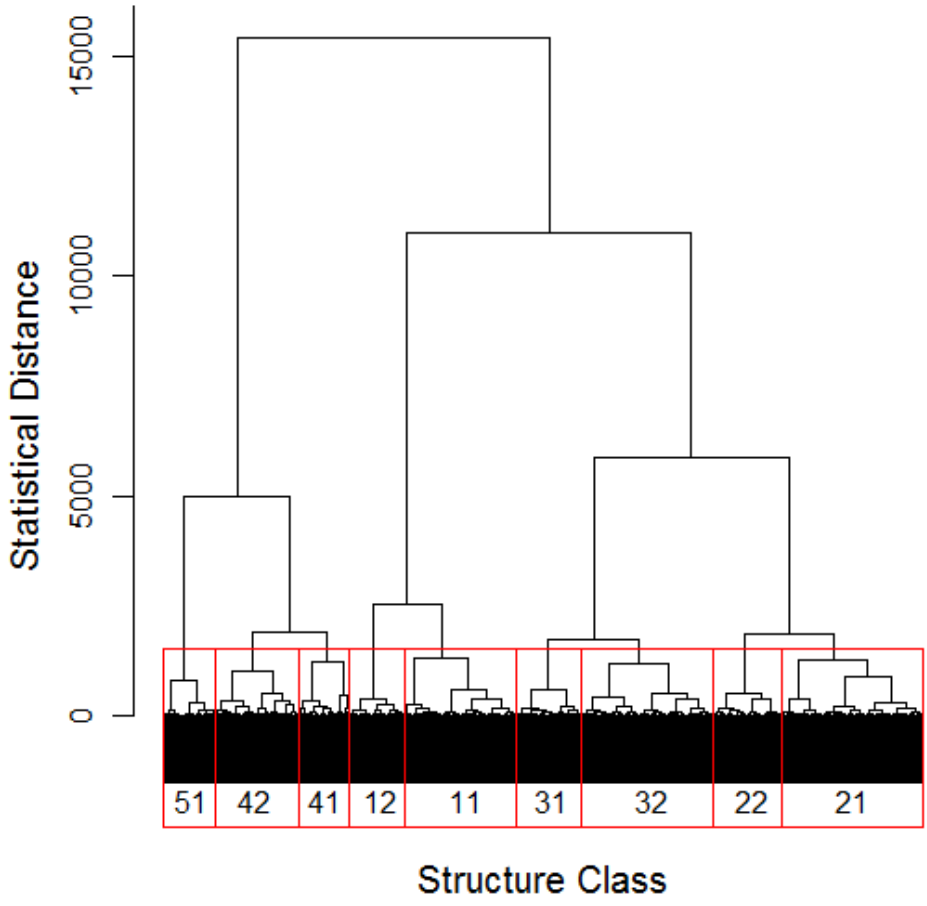


Figure 9. Dendogram showing statistical distance between structure classes identified in this study. Structure classes are identified by numbers with classes that are statistically very similar given adjacent numbers (e.g., 11 and 12 or 21 and 22). Structural classes that are less statistically similar have numbers separated by units of 10 (e.g., 11 vs 21). Structure classes having taller trees were assigned lower numbers within units of 10 and within statistically similar pairs the class with higher canopy cover having the lower number.

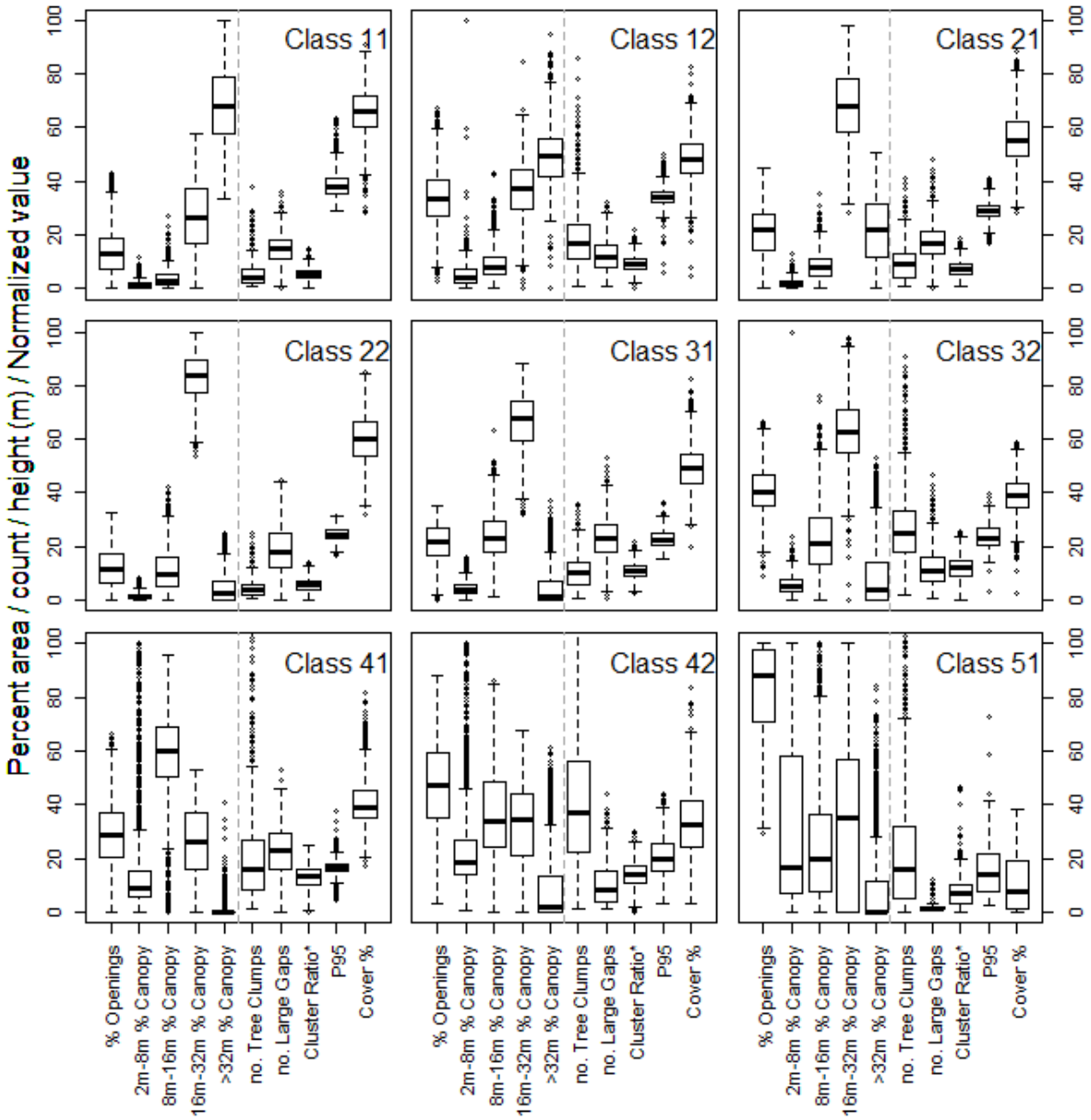


Figure 10. Characteristics of structure classes identified in in the Park. Scale is either percentage, meters in height, count of tree clumps or large gaps, or normalized cluster ratio. Boxplots show ranges of values for metrics used to define each class (left of dotted line in each panel) and metrics that provide additional information on each of the structure classs but which were not used to define the structure classes (right of dotted line in each panel). Open area is the percentage of each 0.81 ha sample with no returns >2 m in height while canopy strata areas are percentage of canopy area in each sample. All metrics were calculated from the canopy surface model except the 95th percentile height for returns >2 m (p95) and the cover percentage which were calculated from the LiDAR return data. The cluster ratio is the number of tree clusters normalized to the area of canopy in each sample and was normalized to a relative range of 0-100

with actual ranges of 0 to 400 clusters. Bold lines show median values; the bottom and top of the boxes show the 25th and 75th percentile values; the upper and lower whiskers show either minimum and maximum values or 1.5 times the interquartile range (approximately two standard deviations), whichever is nearer to the mean; and circles show outliers.

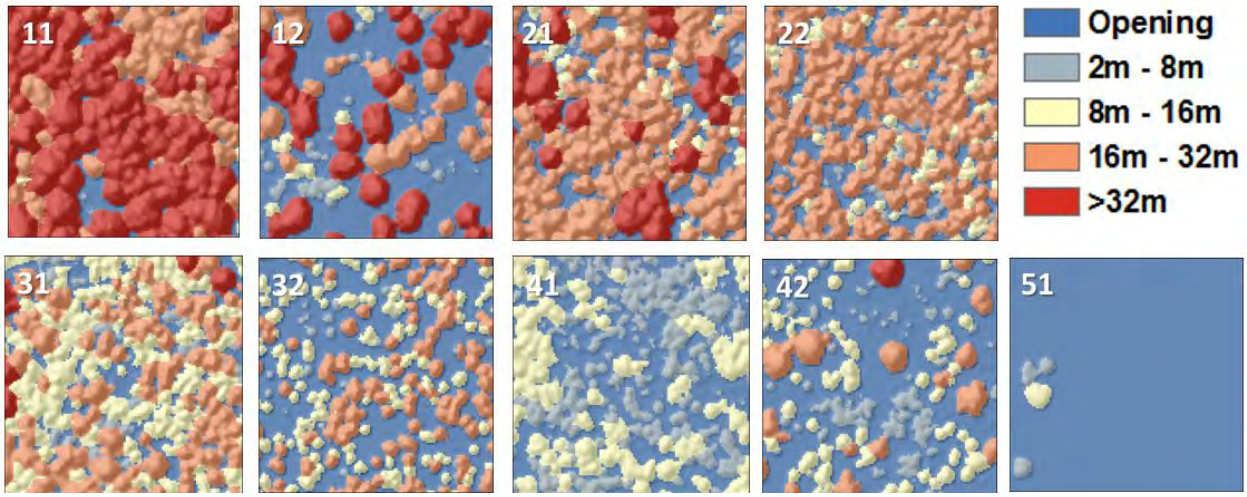


Figure 11. Examples of each of the structure classes identified in this study (identified by numbers within each panel) with each 0.81 ha sample area selected to represent the approximate statistical centroid of structure metrics for each structure class.

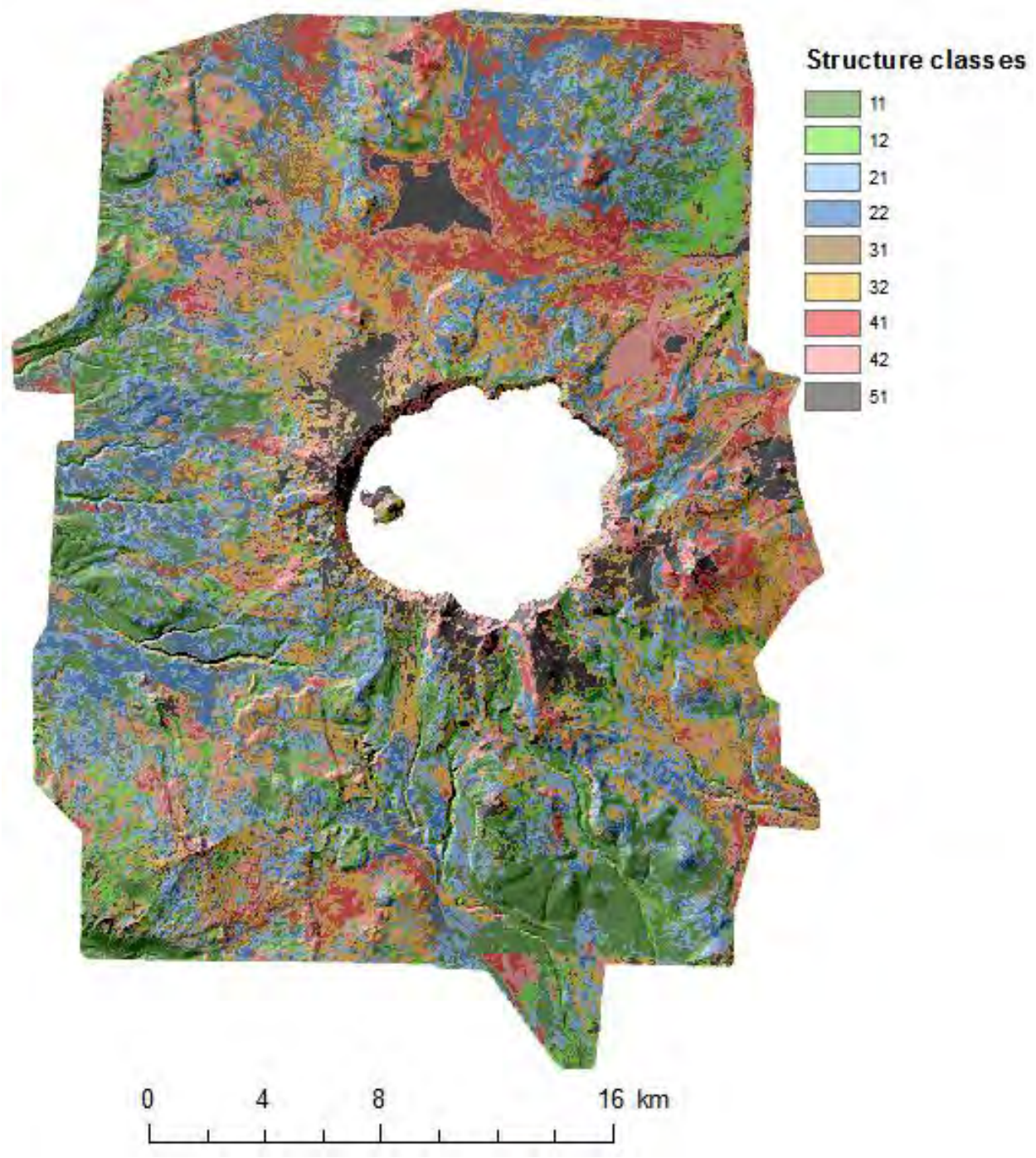


Figure 12. Map of structure classes identified in this study.

Table 3. Key characteristics for the structure classes identified from patterns of canopy and openings across the park.

Structure class	Dominant TAO heights	TAO vertical structure	TAO horizontal structure
11	Tall	Top story	Canopy-gap
12	Tall	Top story	Clump-open
21	Mid	Top story	Canopy-gap
22	Mid	Top story	Canopy-gap
31	Mid	Top story	Clump-open
32	Mid	Multistory	Clump-open
41	Short	Multistory	Open-clump
42	Short	Multistory	Open-clump
51	Short	Multistory	Open-clump

We found that the structure classes divided into three groups based on the dominant tree cluster height strata. The tall classes (classes 11, 12) had the majority of their canopy associated with TAOs in the >32 m and 16-32 m strata. The mid-height classes (classes 21, 22, 31, 32) had the majority of their TAO canopies in the 16-32 m stratum. The short classes (classes 41, 42, 51) had the majority of the TAO canopies in strata < 16 m. Ranges for the 95th percentile height of returns >2 m for each class followed the trends in the distribution of canopy area for the TAO height strata.

The structure classes divided into two groups based on the vertical distribution of canopy area. The top story classes (classes 11, 12, 21, 22, 31) had almost all their canopy area associated with TAOs in the higher strata with little area in the lower strata. The multistory classes (classes 32, 41, 42, 51) had canopy more evenly distributed across height strata with substantial canopy area in lower strata as well as higher strata. Ranges for the clump ratio corresponded to the division of classes into top story and multistory groups.

We also found that the structure classes divided into three groups based on the horizontal arrangement of canopy and openings. As the area in openings increased (and area with canopy decreased), the number of tree clumps increased and the number openings >10 m^2 decreased. Kane et al. (2014) observed similar trends at Yosemite National Park and found three broad patterns of tree clumps and openings. The canopy-gap pattern had nearly continuous canopy with enclosed gaps (typical canopy cover >2 m was >60%); the clump-open pattern had interspersed tree clumps and openings in a similar proportion of area (typical canopy cover >2 m

was 40-60%); and the open pattern were primarily a nearly continuous opening with enclosed tree clumps and individual trees (typical canopy cover >2 was <40%). Our Crater Lake structure classes also were divided between the canopy-gap (classes 11, 21, 22), clump-open (classes 12, 31, 32), and open (classes 42, 51) patterns.

We found that none of the structure classes was uniquely associated with a forest zone (Fig. 13). Conversely, all forest zones had mixtures of structure classes, but one to three structure classes typically were dominant for each zone. The lower and more mesic zones (mesic low, mesic mid, and Sierra Mixed) were dominated by the tall structure classes (11, 12) with significant proportions of their area in the mid-height class (21). Forest zones that represent harsher climate extremes (subalpine, lodgepole pine, ponderosa pine, and East Cascades Mixed) were dominated by mixtures of the mid to shorter height and more open structure classes (31, 32, 41, 42, 51). Only the mesic high forest zone included substantial area across a wide range of structure classes, which may represent its large area that includes a wider range of ecotones than the other forest zones.

We observed several emergent properties from our structure classes. Classification routines such as the hierarchical classifier we used attempt to create classes with the narrowest range of values each metric possible and we observed this for the class area metrics used for the classification. We were surprised to find that the ranges of values for the auxiliary metrics used to interpret the classes showed similarly tight ranges of values for each class (Fig. 10). It appears that given a particular mix of area among the height strata that the number of tree clumps, tree clusters, and openings as well as the dominant tree height (P95 height) and canopy cover are correlated emergent properties. This suggests that there are consistent processes that underlie how forest structure is organized at fine scales (our 0.81 ha areas) across broad biophysical gradients. This would be a fruitful area to explore using spatially explicit forest gap and tree regeneration models.

A second emergent property was that when examined on a map our structure classes across the park, the classes were found largely found in relatively large contiguous patches rather than in fine-scale interspersion (Fig. 12). Where the structure classes were interspersed, it was generally by pairs of classes that were statistically similar. This result has implications for the scale at which stochastic chronic disturbance patches operate. If stochastic disturbance was driving tree clump/cluster and opening patterns at the scale of our 0.81 ha analysis area, then we'd expect to see smaller patches of structure classes possibly approaching a 'salt and pepper' texture of highly intermixed structure classes. The distribution of structure classes into relatively large patches suggests that the scale of these disturbances typically is significantly smaller than our analysis area. In addition, the ubiquitous and spatially synchronous effects of fire exclusion allowed widespread infill (Forrestel and Stephens, 2011) that may have homogenized local structure patterns.

Smith and Urban (1988) used a spatially explicit forest gap and tree regeneration model to explore this question. They found that at small scales (0.01 ha) the state of the structure at each location through time was essentially random. However, the variability between patches (0.1-0.16 ha) was minimal. Their modeling suggests that the sum of fine scale chronic disturbances rapidly average out spatially to produce repeating patterns of tree clusters and openings at slightly larger areas. Our large contiguous patches of structure classes are consistent with this pattern.

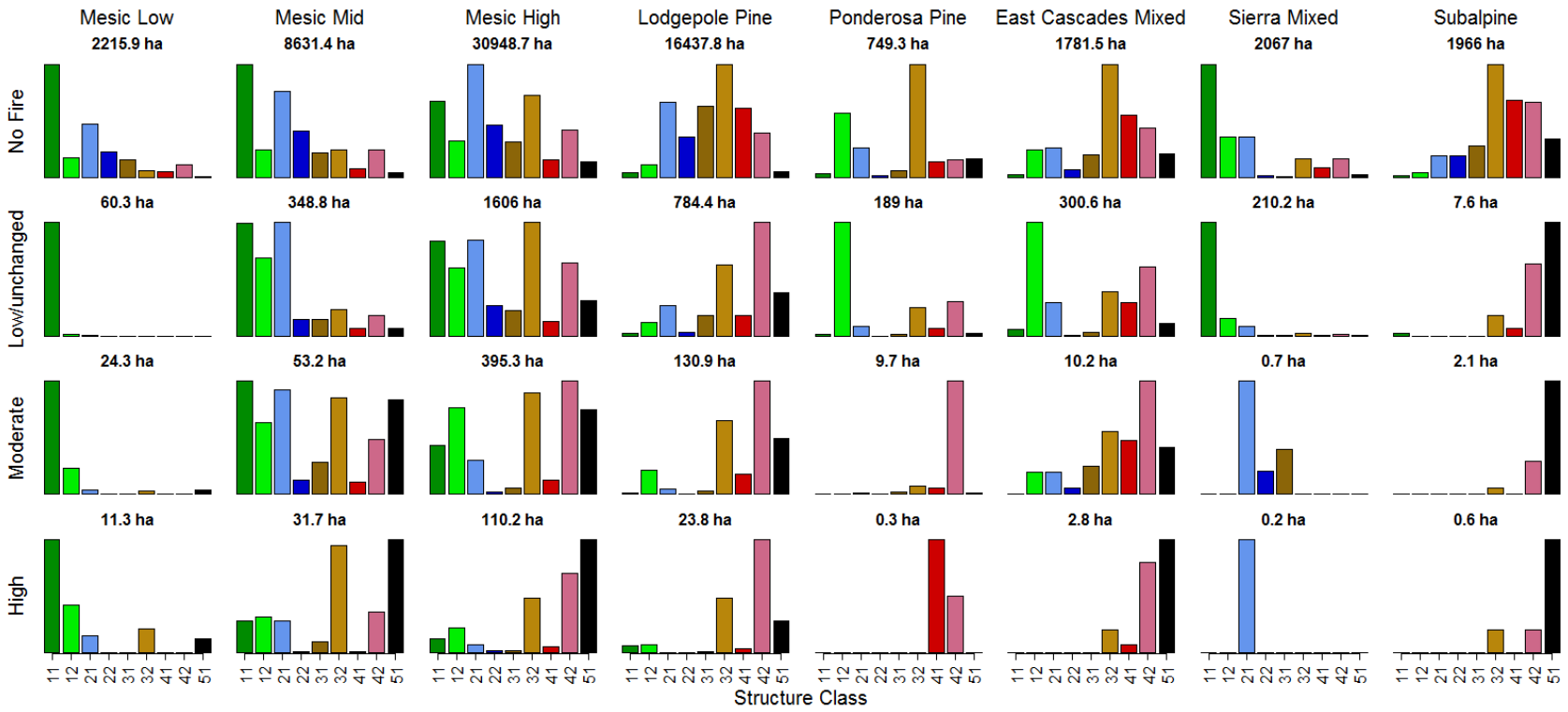


Figure 13. Frequency of structure classes by forest zone for areas outside of all fire perimeters ('No Fire') and by dNBR classified burn severity. No forest zone had more than xx% percent of its area burn, making it difficult to be certain of the trends in changes in forest structure with fire.

3.4 Fire

Only 6.2% of the park had burned between 1984, the first year for which dNBR burn severity measurements are available, and the time of the LiDAR acquisition in 2010 with the percentage ranging from <1% for the subalpine forest zone to 20.9% for the lodgepole zone. Just 14.5% of the burned area burned with a moderate burn severity and 0.42% with a high burn severity (Table 4). Because percentage of area burned is so low, especially for moderate and high burn severities, it is risky to identify more general trends of how fire severity relates to forest structure (Fig. 13). In addition, dNBR estimates mortality only for the year following a fire, while park managers report that substantial mortality occurs in subsequent years (C. Farris, personal communication). As a result, the reported burn severity may under estimate the eventual effects of fire on forest structure.

We found that burn severity for areas that burned was weakly correlated (44% predictability) with the biophysical environment. Burn severity was most strongly related to increases in AET, which would be associated with increased production of fuels and also to increases in Deficit, which would be associated with drier fuels during the burning season.

Table 4. Proportion of forested area burned with different estimated burn severities (dNBR) by forest zone and for all forested areas.

	Burn severity		
Forest zone	Low/No Change	Moderate	High
Mesic Low	2.6%	1.1%	0.5%
Mesic Mid	3.8%	0.6%	0.3%
Mesic High	4.9%	1.2%	0.3%
Lodgepole	4.5%	0.8%	0.1%
Ponderosa pine	19.9%	1.0%	0.0%
East Cascades mixed	14.3%	0.5%	0.1%
Sierra mixed	9.2%	0.0%	0.0%
Subalpine	0.4%	0.1%	0.0%
All	5.1%	0.9%	0.3%

4.0 GROUND PLOTS

We used existing NPS field data to compare forest structure metrics derived from LiDAR and field plots. A total of 68 0.1 ha long-term fire effects monitoring plots using the National Park Service Fire Monitoring (FMH) protocol (National_Park_Service, 2001) were used from four vegetation types (n = 23 mixed conifer; n = 13 lodgepole pine; n = 14 ponderosa pine; n =

14 mountain hemlock). The closest measurement to the LiDAR acquisition date (2010) was used. Additionally, 20 0.1 ha research plots were utilized from lodgepole pine and ponderosa pine stands where tree data were collected in June 2010 using the same FMH protocol. The diameter at breast height, species, status (dead/live) were collected within each plot. GPS location data for the field plots was subsequently collected and post processed by the Park Service to derive plot locations.

We created linear regression models for tree density and basal area per hectare using Lidar metrics as the explanatory variables, and plot data as the response variables. We considered a variety of height- and cover-related Lidar variables, and used the regsubset function of the leaps R package to consider all possible subsets of those variables as predictors. We restricted the analysis to models with three or fewer predictors, to avoid overfitting. The results are summarized in table X. All terms were significant ($p < 0.05$). Additionally, the standard four plots produced by R's plot() function were examined for each of the regressions. Response variables were transformed by raising to the exponent found by the powerTransform function in the car package; this eliminated heteroscedasticity as detected by the Breusch-Pagan test, implemented as bptest in the lmtest package.

Both models were moderately successful (Table 5) with RMSE errors as percentage of mean values of 0.24 to 0.27. Both over predicted most values for moderate actual values (Fig. 14).

Table 5. Results of regression modelling. RMSE is presented in the units of the (transformed) response variable. RMSE % is RMSE as a percentage of the mean of the transformed response.

Modelled variable	Model	Adj. R^2	RMSE	RMSE %
Live tree density	TreeDensity^0.22 = 18.18 -6.84*Elev.CSV +0.24*Elev.Kurtosis -0.18*CVR>mean	0.689	0.92	0.275
Live basal area per ha	BasalArea^0.26= 5.01 -2.95*Elev.CV +0.10*Elev.kurtosis +2.48*Elev.L.skewness	0.769	0.49	0.242

Elev.CV - Coefficient of variation of LiDAR return heights >2 m
Elev.Kurtosis - Kurtosis of LiDAR return heights >2 m
CVR>mean - Cover calculated as returns above mean/all first returns
Elev.L.skewness - L moment skewness of LiDAR returns >2 m

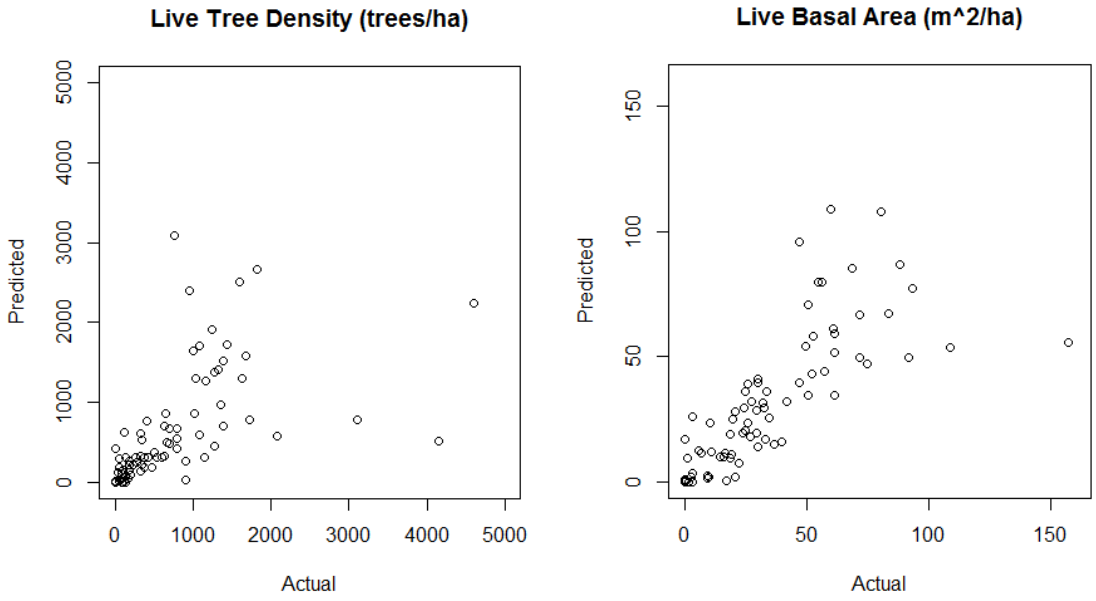


Figure 14. Scatter plots showing actual versus predicted values from regression modeling for tree density and basal area.

5.0 DELIVERED LIDAR-DERIVED METRICS

5.1 Brief Introduction to Key Metrics

LiDAR datasets are collections of georeferenced points described in three-dimensional space. These point clouds are processed to produce combinations of metrics in order to describe different characteristics of stand structure, much like how combinations of field plot data are used to generate metrics to describe stand characteristics

The program used to produce the metrics in this package is called FUSION. The FUSION program was developed by Robert J. McGaughey at the USFS Pacific Northwest Research Station. FUSION output metrics are stored as rasters, which are grids of the area being described. The finest resolution of a raster is a single cell. Within each cell, a single value is assigned which described the overall characteristics of the returns that fall within that cell. The resolution of a raster indicates the size of its cells, and is included in the file name. The resolution of the majority of the rasters in this package is 30 meters (98.424 feet). Therefore, the values of a single cell summarize the conditions within a 30 by 30 meter area.

Area-based LiDAR data is collected from above the canopy surface. Dense canopies reduce the number of pulses that are able to penetrate below the outer canopy surface. As a result, canopy layers below the outer surface are measured with lower fidelity than for the outer surface.

In order to ensure that the canopy metrics reflect only the structure of the trees within

each raster grid cell, the majority of metrics in this package were calculated using only returns greater than 2 meters in height above the surface. This excludes returns from the ground and low-lying vegetation.

Metrics labeled with elev_p95 represent the 95th percentile height value for all returns measured within the cell above the height cutoff. Elev_p95 can be used as a surrogate for dominant tree height. It is similar to maximum height but less sensitive to anomalously high returns (Kane et al., 2010).

Similarly, metrics labeled with elev_p25 are calculated in the same way as metrics labeled with elev_p95 but indicate the height value of the 25th percentile height of returns greater than 2 meters. This metric has been shown to be a strong surrogate for crown base height (i.e. height to live crown) (Erdody and Moskal, 2010).

Cover metrics describe estimates of canopy closure at the resolution of the grid cell size. Values range from 0 to 100, and represent the percentage of all returns within a cell that are above a threshold cutoff height (McGaughey, 2014). Cover may be calculated using all returns or only first returns.

Metrics labeled elev_ave indicate the mean height value for all returns measured within the cell above the height cutoff. Average height of LiDAR returns is often useful in modeling applications.

Strata metrics are metrics that describe the characteristics of a given cell within the upper and lower height cutoffs of a given stratum. Strata cover metrics describe the proportion of all returns within a given stratum of a cell relative to all returns at or below the upper cutoff of the stratum within the cell above the height cutoff (McGaughey, 2014). Strata cover is a measurement of canopy closure at a given strata, and can be used to describe the vertical distribution of the canopy.

Rumple is a measure of the rugosity of the canopy surface and is the ratio of the outer canopy surface area divided by the underlying ground surface ratio (Kane et al., 2010; Parker et al., 2004). Rumple indicates the degree of canopy complexity and therefore stand structure. A value of 1 represents a completely flat surface (ground only with no vegetation) and increasing values indicate increasing canopy complexity.

Topographic metrics, or topo metrics, are metrics that describe the topography of an area, irrespective of the vegetative cover. While these metrics are reported at a 30 meter resolution, the values of each cell are calculated for values found within a reference window centered on the cell, which may be larger than the cell itself. Changes in the size of this window affect the values of the resulting topographic metrics. For example, information about micro topography might be lost when using a larger window, while landscape level patterns might be lost when using a smaller window. The radius of the window used in calculating topo metrics is reported at the beginning of the file name. topo_slope is a metric which describes the percent slope within the grid cell. topo_aspect indicates the dominant aspect of a cell in degrees. topo_curvature is a measure of overall curvature, combining profile curvature (along the slope), and plan curvature (across the slope) (Zevenbergen and Thorne, 1987).

5.2 Illustrations of LiDAR metrics

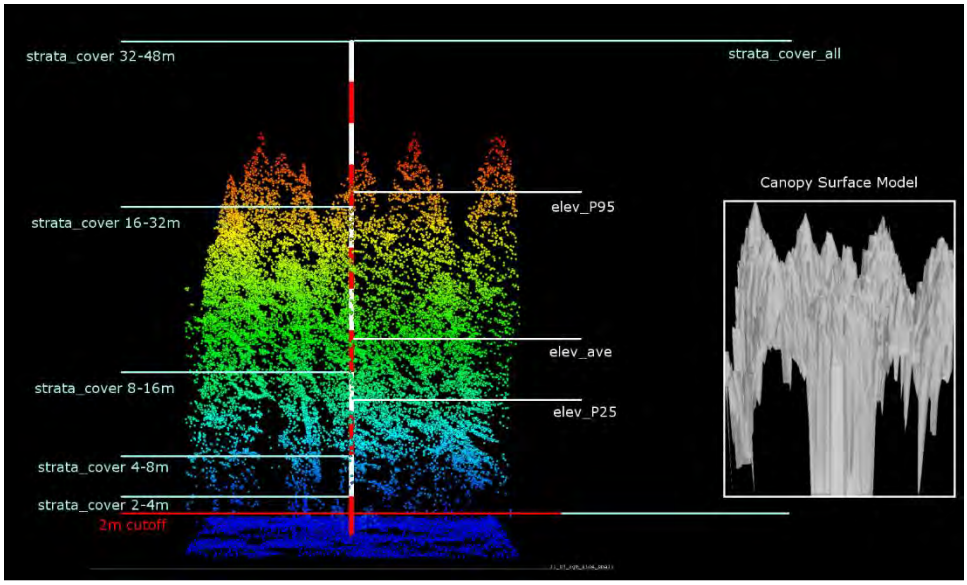
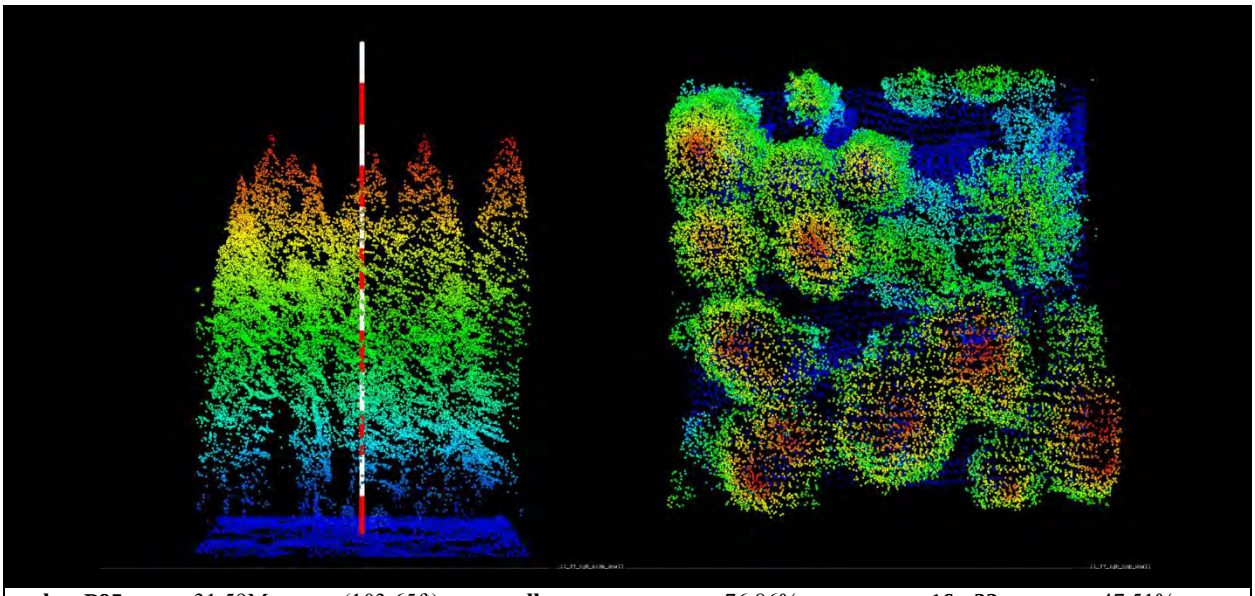


Figure 15. Illustration of key vertical structure metrics mapped across Crater Lake National Park. All of these metrics were calculated from the LiDAR return data except for rumple which was calculated from a 1 m canopy surface model. Cover and height return values were calculated for returns >2 m above the ground. Metric names shown are the same as the names of the corresponding rasters supplied with this final report.

It is often useful to visualize samples from the point cloud in order to better understand how the metrics being used describe the areas being studied. The image above show the returns measured within the area covered by a single pixel (30×30 m area) of the metric rasters. The range pole is marked in four meter increments.

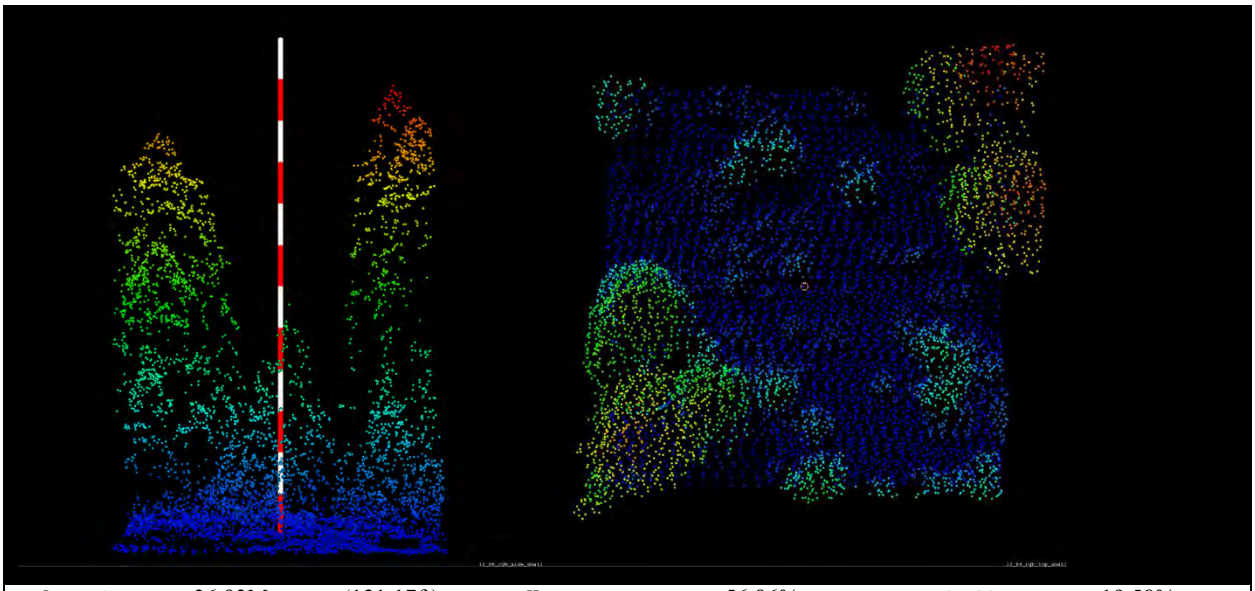
The locations of several of the metrics described earlier are marked on this image. Elev_P95, which is a surrogate for crown height, is 31.59 meters (103.65 feet) for this sample. Elev_P25, which is a surrogate for crown base height, is 13.07 meters (42.88 feet). Elev_ave is 19.22 meters (63.06 feet). Note that the returns below the 2 meter cutoff are excluded for these metrics. The ranges of the strata cover metrics are marked on the left side of the image. They, from bottom to top are 4.01%, 17.59%, 42.76%, 47.51%, and 3.3% respectively. Total cover for this pixel is 76.86%. Note that since strata cover is calculated irrespective to the pixels above a given strata, the cumulative value of strata cover will not necessarily equate to total cover. Returns below the cutoff are included in cover calculations. Rumple, which is calculated as the ratio of the canopy surface model (pictured in the left of the above image) divided by the underlying ground surface ratio (Van R. Kane et al., 2010; Parker et al., 2004). The rumple value for this pixel is 3.25.

Below are example pixels from each canopy structure class, with a selection of corresponding metric values. This selection includes the metrics that were used to generate the structure classes, as well as several which were not. Pay attention to the way in which these metrics vary between classes, and observe how the values correspond to the stand conditions.



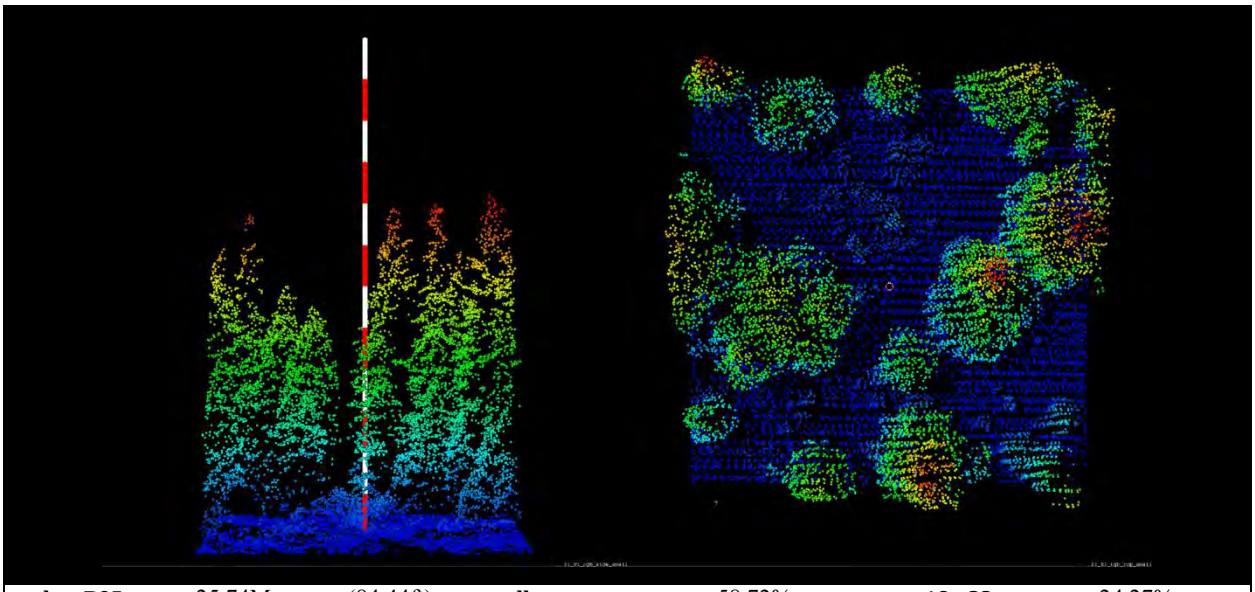
elev_P95	31.59M	(103.65ft)	**all_cover**	76.86%	**cover16to32**	47.51%
elev_P25	13.07M	(42.88ft)	**cover_2to4**	4.01%	**cover_32to48**	3.30%
elev_ave	19.22M	(63.06ft)	**cover_4to8**	17.59%	**cover_48to64**	0%
rumple	3.2483		**cover_8to16**	42.76%	**cover_64plus**	0%

Figure 16-a. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



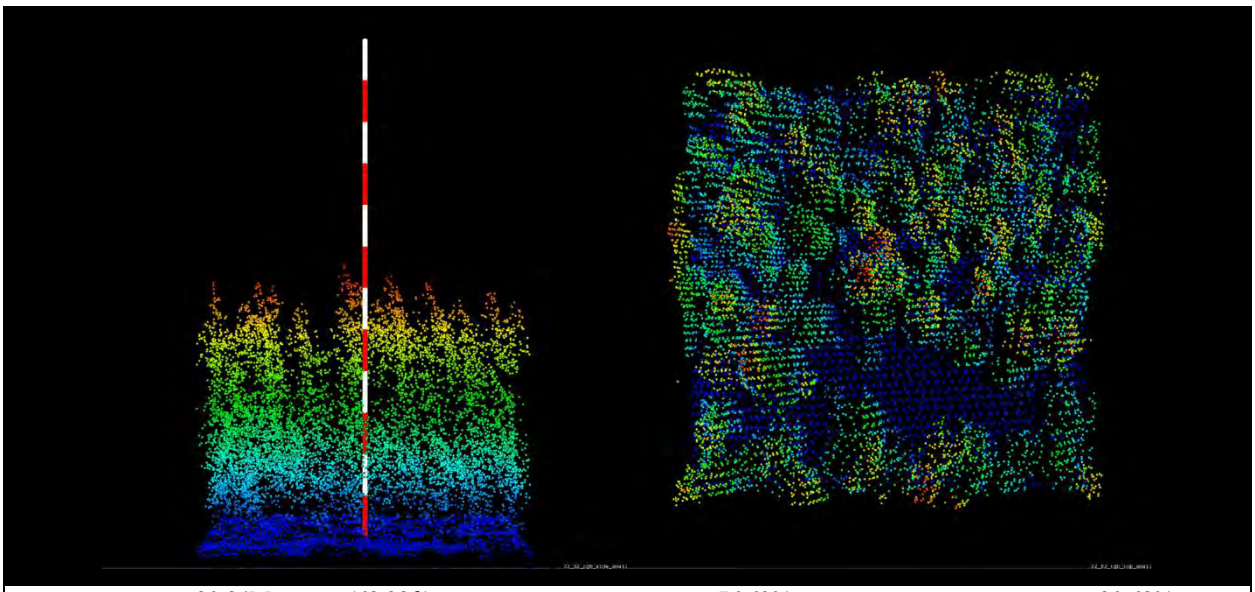
elev_P95	36.93M	(121.17ft)	**all_cover**	56.86%	**cover16to32**	19.58%
elev_P25	5.7M	(18.7ft)	**cover_2to4**	20.09%	**cover_32to48**	9.70%
elev_ave	16.78M	(55.06ft)	**cover_4to8**	17.11%	**cover_48to64**	0%
rumple	3.9702		**cover_8to16**	17.86%	**cover_64plus**	0%

Figure 16-b. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



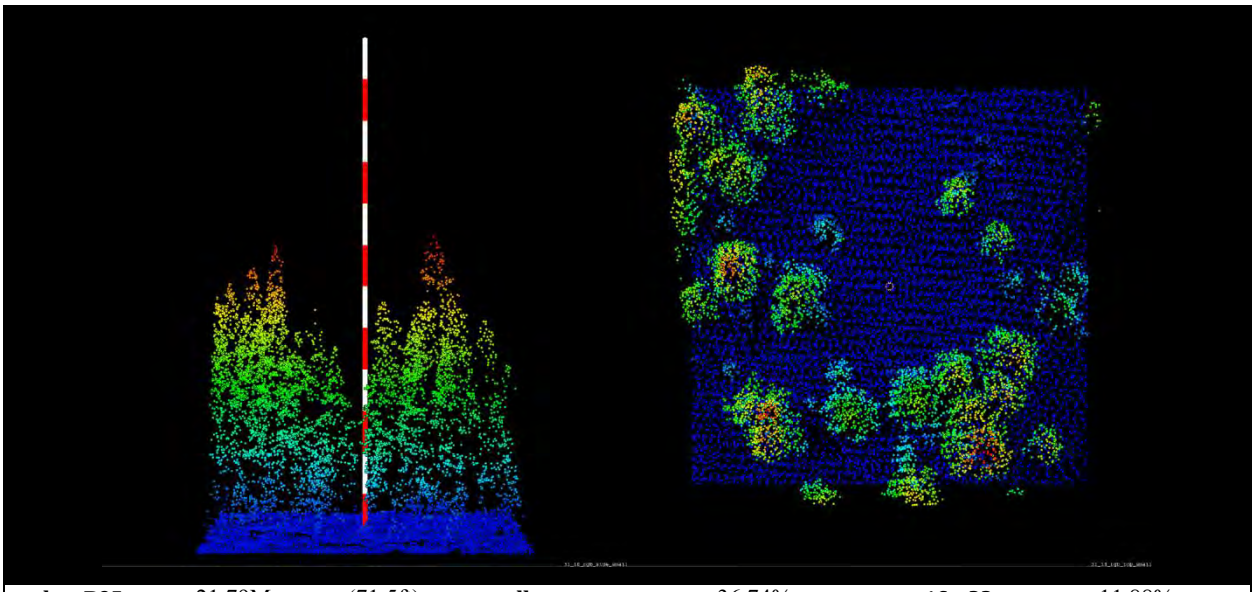
elev_P95	25.74M	(84.44ft)	**all_cover**	58.72%	**cover16to32**	24.27%
elev_P25	8.55M	(28.04ft)	**cover_2to4**	11.35%	**cover_32to48**	0.05%
elev_ave	13.92M	(45.66ft)	**cover_4to8**	17.41%	**cover_48to64**	0%
rumple	3.5408		**cover_8to16**	30.51%	**cover_64plus**	0%

Figure 16-c. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



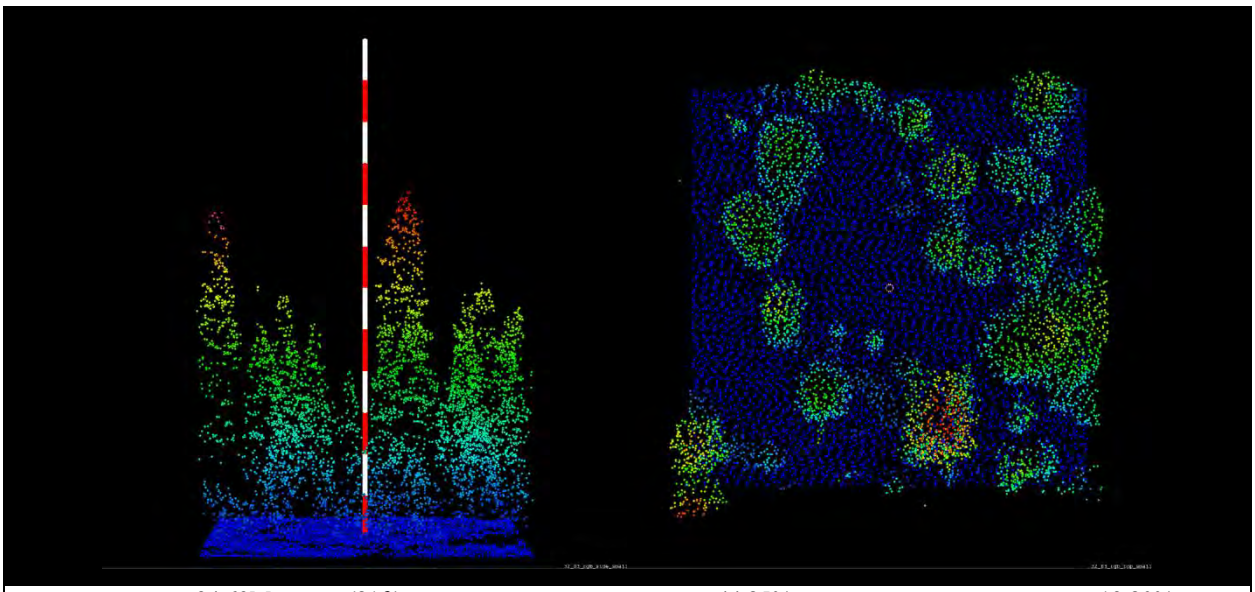
elev_P95	20.84M	(68.38ft)	**all_cover**	76.69%	**cover16to32**	20.68%
elev_P25	6.63M	(21.76ft)	**cover_2to4**	18.21%	**cover_32to48**	0%
elev_ave	11.27M	(36.96ft)	**cover_4to8**	46.11%	**cover_48to64**	0%
rumple	2.9732		**cover_8to16**	35.88%	**cover_64plus**	0%

Figure 16-d. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



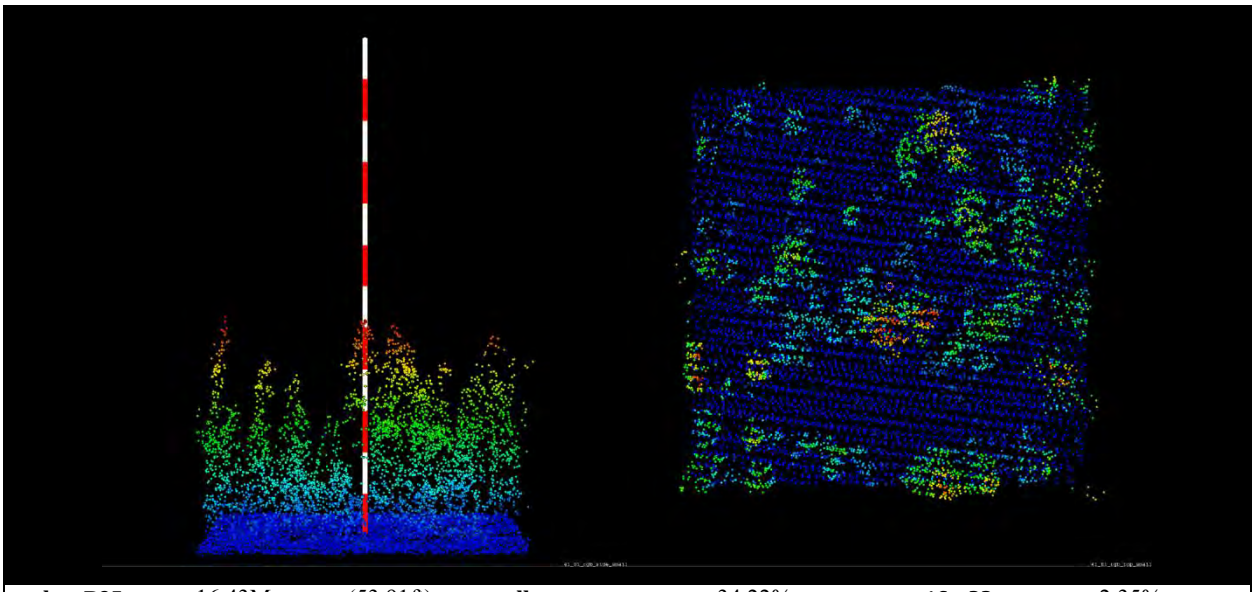
elev_P95	21.79M	(71.5ft)	**all_cover**	36.74%	**cover16to32**	11.88%
elev_P25	8.52M	(27.96ft)	**cover_2to4**	3.68%	**cover_32to48**	0%
elev_ave	13.01M	(42.69ft)	**cover_4to8**	8.30%	**cover_48to64**	0%
rumple	3.0045		**cover_8to16**	19.38%	**cover_64plus**	0%

Figure 16-e. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



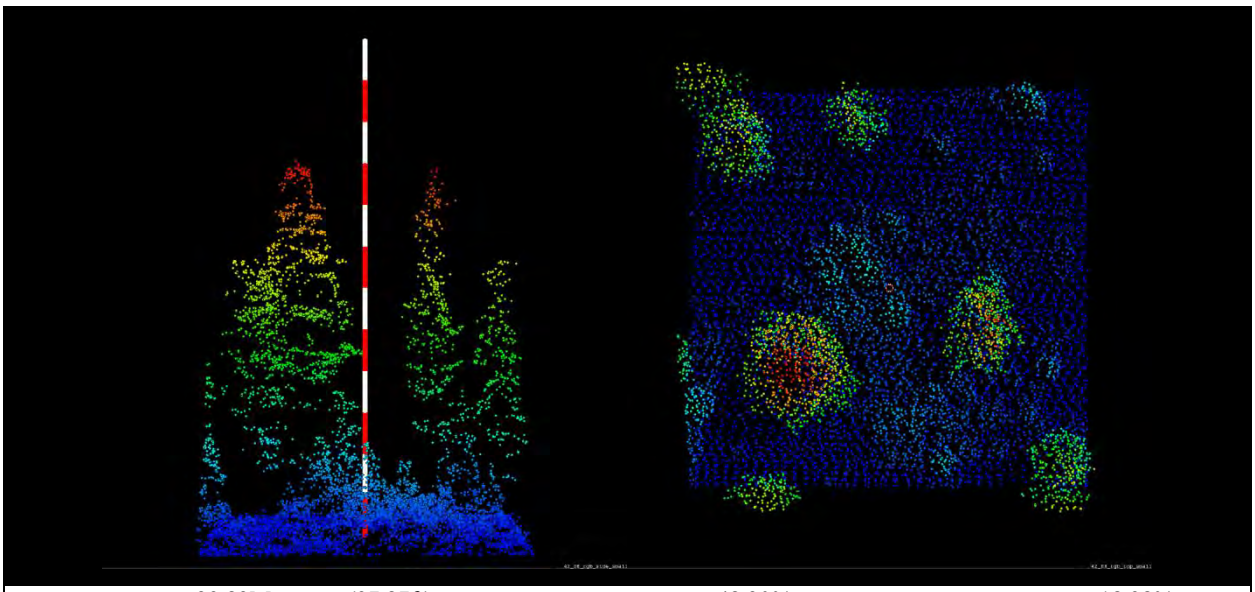
elev_P95	24.69M	(81ft)	**all_cover**	44.85%	**cover16to32**	13.80%
elev_P25	7.37M	(24.18ft)	**cover_2to4**	6.79%	**cover_32to48**	0.07%
elev_ave	12.66M	(41.53ft)	**cover_4to8**	13.15%	**cover_48to64**	0%
rumple	3.2913		**cover_8to16**	22.02%	**cover_64plus**	0%

Figure 16-f. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



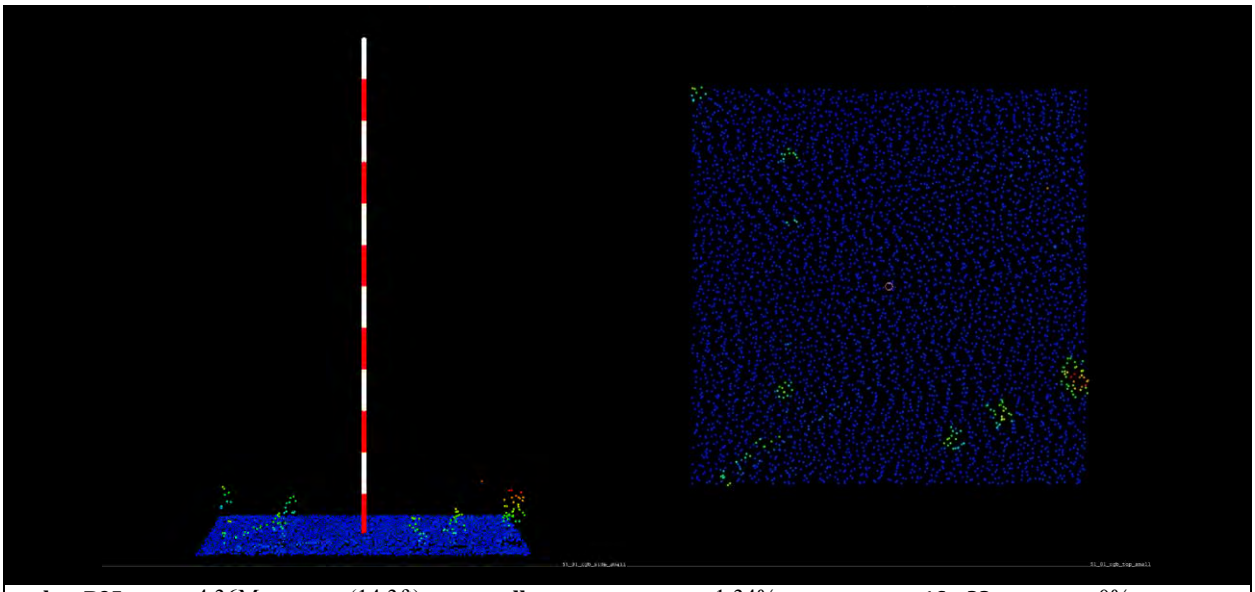
elev_P95	16.43M	(53.91ft)	**all_cover**	34.22%	**cover16to32**	2.35%
elev_P25	4.38M	(14.38ft)	**cover_2to4**	10.55%	**cover_32to48**	0%
elev_ave	8.08M	(26.51ft)	**cover_4to8**	13.81%	**cover_48to64**	0%
rumple	2.7149		**cover_8to16**	14.54%	**cover_64plus**	0%

Figure 16-g. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



elev_P95	29.83M	(97.87ft)	**all_cover**	48.30%	**cover16to32**	18.98%
elev_P25	3.35M	(10.97ft)	**cover_2to4**	26.72%	**cover_32to48**	1.21%
elev_ave	12.24M	(40.15ft)	**cover_4to8**	12.56%	**cover_48to64**	0%
rumple	3.3456		**cover_8to16**	9.52%	**cover_64plus**	0%

Figure 16-h. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.



elev_P95	4.36M	(14.3ft)	**all_cover**	1.34%	**cover16to32**	0%
elev_P25	2.38M	(7.8ft)	**cover_2to4**	1.19%	**cover_32to48**	0%
elev_ave	2.94M	(9.64ft)	**cover_4to8**	0.16%	**cover_48to64**	0%
rumple	1.0326		**cover_8to16**	0%	**cover_64plus**	0%

Figure 16-i. Example of LiDAR structure metrics for 30×30 m area. Stripes on the range pole are 4 m in length.

5.2 Listing of All Supplied Metrics

This document describes the LiDAR metric layers provided. They were calculated using the FUSION software with additional post processing done for some layers.

These layers and this document are intended for use by our research collaborators and presume knowledge of LiDAR processing and the types of metrics produced by the FUSION software as well as an understanding of raster GIS data.

Details on the calculation of the different metrics can be found in the FUSION software manual available at http://forsys.cfr.washington.edu/fusion/fusionlatest.html within the install file.

To process the Crater Lake LiDAR data, an unreleased tool written by the author of the FUSION software (R. McGaughey, Pacific Northwest Research Station) was used to split the data into smaller chunks for processing efficiency. The tool divided the LiDAR vendor-supplied LAS tiles into blocks that were further subdivided into new tiles. The scripts used to process each tile or block are provided in a directory with the rasters of the calculated metrics.

5.2.1 Unit Equivalents

Heights and elevations within all products are in meters.

Names of directories and file names contain the raster resolution and for some metrics the ranges over which the metrics were calculated. These are given in either meters with ‘pt’ or ‘p’

representing the decimal place in the number. A table with foot equivalents is provided below:

METERS	FEET
0	0
0.5	1.6404
0.75	2.4606
1	3.2808
1.5	4.9212
2	6.5616
4	13.1232
5	16.404
8	26.2464
10	32.808
15	49.212
16	52.4928
30	98.424

METERS	FEET
32	104.9856
45	147.636
48	157.4784
60	196.848
64	209.9712
120	393.696
135	442.908
200	656.16
270	885.816
500	1640.4
1000	3280.8
2000	6561.6
4000	13123.2

5.2.2 Quick Summary of Directory Contents

Canopy surface models and products derived from canopy surface model

- CanopyMetrics_30METERS (contains the metrics calculated using canopy surface model)
- CanopyHeight_0p75METERS(contains the 0.75 meter resolution canopy height model). Because of the high resolution, results are provided in blocks rather than for the entire area.
- CanopyHeight_1METER (contains the 1 meter resolution canopy height model)
- CanopyHeight_1p5METERS (contains the 1.5 meter resolution canopy height model)
- CanopyHeight_2METERS (contains the 2 meter resolution canopy height model)

Vegetation metrics derived from the LiDAR return data

- Metrics_30METERS (contains the metrics calculated using the LiDAR point cloud)
- StrataCoverMetrics_30METERS (contains the normalized cover metrics calculated on those returns that fall within each height stratum)
- StrataMetrics_30METERS (contains the metrics calculated on those returns that fall within each height stratum)
- Intensity_1p5METERS (contains the 1.5 meter resolution mean intensity metric)

Topographic metrics derived from the LiDAR digital surface model and USGS DEM

- MASKED_NormalizedTPIMetrics_30METERS (contains the 30 meter resolution topographic position index metric normalized to the ground model, masked to exclude areas calculated solely using the USGS ground model)
- MASKED_TopoMetrics_30METERS (contains the 30 meter resolution topographic metrics, masked to exclude areas calculated solely using the USGS ground model)
- NormalizedTPIMetrics_30METERS(contains the 30 meter resolution topographic position index metric normalized to the ground model)
- TopoMetrics_30METERS (contains the 30 meter resolution topographic metrics)

Files related to processing

- Layout_shapefiles (contains the shapefiles showing the layout of the acquisition)
- Fusion processing scripts

Tree-approximate objects and related metrics

- Segments_0p75METERS (contains the tree-approximate objects)
- FRAGSTATSMetrics_90METERS (contains FRAGSTATS)

5.2.3 Processing Index Files

Directory:

- Layout_shapefiles

Processing of areas was done by blocks of tiles. For most metrics, the results were merged into a single raster representing the entire study area. For certain high resolution products, the results are presented in either blocks or tiles.

Shape file indices to the blocks and tiles are in the directory 'Layout_shapefiles'. The shape files included are:

crla_DeliveryTiles.shp

crla_ProcessingBlocks.shp

crla_ProcessingTiles.shp

5.2.4 Canopy Metrics Calculated From LiDAR Return Data

Directory:

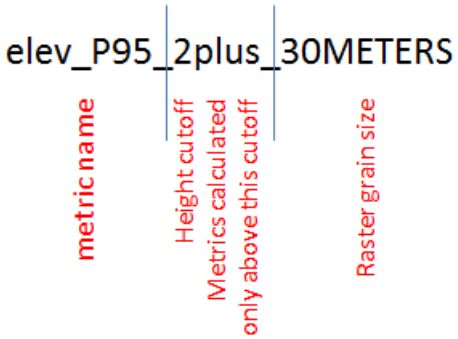
- Metrics_30METERS

We calculate a number of metrics based on the distribution of LiDAR heights normalized to the ground surface and LiDAR intensity values:

elev* -- metrics calculated based on height

int* -- metrics calculated based on intensity

We calculate each metric twice, once using all returns and once using only the value of the first return in each pulse. For the latter, the raster name is preceded by 'FIRST_RETURNS*'.



All metrics are calculated using 30 m grid cells.

In addition to the height and intensity metrics, FUSION also calculates the number of LiDAR returns within each 30 m grid cell:

pulsecnt_30METERS.img – count of pulses (equal to the count of first returns)

Count of all returns

all_cnt_2plus_30METERS.img

all_cnt_30METERS.img

all_cnt_30METERS.img.aux.xml

all_cnt_above2_30METERS.img

all_cnt_above_mean_30METERS.img

all_cnt_above_mode_30METERS.img

Count of first returns

1st_cnt_above2_30METERS.img

1st_cnt_above_mean_30METERS.img

1st_cnt_above_mode_30METERS.img

Count of returns by return number

r1_cnt_2plus_30METERS.img

r2_cnt_2plus_30METERS.img

r3_cnt_2plus_30METERS.img

r4_cnt_2plus_30METERS.img

r5_cnt_2plus_30METERS.img

Canopy cover calculated above different height cutoffs (calculated based only on height)

1st_cover_above2_30METERS.img

1st_cover_above_mean_30METERS.img

1st_cover_above_mode_30METERS.img

all_1st_cover_above2_30METERS.img

all_1st_cover_above_mean_30METERS.img

all_1st_cover_above_mode_30METERS.img

all_cover_above2_30METERS.img

all_cover_above_mean_30METERS.img

all_cover_above_mode_30METERS.img

Statistical measures of LiDAR. The metrics calculated using return height and intensity are the same except for canopy cover as noted below, and the naming conventions are the same except for the prefix ('elev' or 'int'). Only the height metrics are shown below.

elev_AAD_2plus_30METERS.img (Average Absolute Deviation)

elev_canopy_relief_ratio_30METERS.img

elev_CV_2plus_30METERS.img

elev_IQ_2plus_30METERS.img

elev_kurtosis_2plus_30METERS.img

elev_max_2plus_30METERS.img

elev_min_2plus_30METERS.img

elev_mode_2plus_30METERS.img

elev_ave_2plus_30METERS.img

elev_cubic_mean_30METERS.img

elev_quadratic_mean_30METERS.img

elev_skewness_2plus_30METERS.img

elev_stddev_2plus_30METERS.img

elev_variance_2plus_30METERS.img

elev_MAD_median_30METERS.img (Median of the absolute deviations from the overall median)

elev_MAD_mode_30METERS.img (Median of the absolute deviations from the overall mode)

Metrics calculated on the L-moments

elev_L1_2plus_30METERS.img

elev_L2_2plus_30METERS.img

elev_L3_plus_30METERS.img

elev_L4_2plus_30METERS.img

elev_LCV_2plus_30METERS.img

elev_Lkurtosis_2plus_30METERS.img

elev_Lskewness_2plus_30METERS.img

Percentile values

elev_P01_2plus_30METERS.img

elev_P05_2plus_30METERS.img

elev_P10_2plus_30METERS.img

elev_P20_2plus_30METERS.img

elev_P25_2plus_30METERS.img

elev_P30_2plus_30METERS.img

elev_P40_2plus_30METERS.img

elev_P50_2plus_30METERS.img

elev_P60_2plus_30METERS.img

elev_P70_2plus_30METERS.img

elev_P75_2plus_30METERS.img

elev_P80_2plus_30METERS.img

elev_P90_2plus_30METERS.img

elev_P95_2plus_30METERS.img

elev_P99_2plus_30METERS.img

5.2.5 Strata Metrics

Directories:

- StrataMetrics_30METERS
- StrataCoverMetrics_30METERS

Strata metrics were calculated using breaks of 0.5, 1, 2, 4, 8, 16, 32, 48, and >64 m using the FUSION software. Canopy cover for a subset of those strata are in the StrataCoverMetrics_30METERS folder.

The original FUSION strata metrics calculated by the FUSION software are in the StrataMetrics_30METERS. The following metrics were calculated for each strata (using the 0 to 2 m strata raster names as an example):

strata_0to0p5M_CV_30METERS.img
strata_0to0p5M_kurtosis_30METERS.img
strata_0to0p5M_max_30METERS.img
strata_0to0p5M_mean_30METERS.img
strata_0to0p5M_median_30METERS.img
strata_0to0p5M_min_30METERS.img
strata_0to0p5M_mode_30METERS.img
strata_0to0p5M_return_proportion_30METERS.img
strata_0to0p5M_skewness_30METERS.img
strata_0to0p5M_stddev_30METERS.img
strata_0to0p5M_total_return_cnt_30METERS.img

The following strata cover metrics were derived from the strata point cloud data:

strata_0p5to1M_cover_30METERS.img
strata_2to4M_cover_30METERS.img
strata_4to8M_cover_30METERS.img
strata_8to16M_cover_30METERS.img
strata_16to32M_cover_30METERS.img
strata_32to48M_cover_30METERS.img
strata_48to64M_cover_30METERS.img
strata_64M_plus_cover_30METERS.img

5.2.6 Canopy Height Models

Directories:

- CanopyHeight_0p75METERS
- CanopyHeight_1METERS
- CanopyHeight_1p5METERS
- CanopyHeight_2METERS

These are models of the top of canopy height, normalized to height above the ground. Each grid cell records the maximum height of the highest LiDAR return in each grid cell.

Because of the fine resolution (1, 1.5, and 2 m), the resulting files are too large to be merged for larger acquisitions. As a result, the models are broken down by processing block.

Canopy height models are organized in directories by resolution. Within each directory, there are several versions of canopy height models. Examples for 1 m canopy height models:

BLOCKn_CHM_filled_3x_smoothed_1METERS.img – Model is smoothed using a 3x3 smoothing algorithm

BLOCKn_CHM_filled_not_smoothed_1METERS.img – No smoothing is done. It is common for small areas at the scale of a few meters to not have any LiDAR returns. Values for these areas are 'filled in' by extrapolating values from surrounding grid cells with LiDAR returns.

Two specialized canopy height models are also created that might be useful for exploring features close to the ground surface. Examples for 1 m canopy height models:

BLOCKn_CHM_lt0p25m_1METERS.img – Canopy height model for returns less the 0.25 m.

BLOCKn_CHM_ltCoverCutoff_1METERS.img – Canopy height model for returns less than the height cutoff used for calculating canopy cover, which is 2 m in our processing.

5.2.7 Metrics calculated from the canopy height models

Directory:

- CanopyMetrics_30METERS

Metrics are calculated over 30 m grid cells. Rasters are merged to cover the entire acquisition area.

Metrics:

canopy_30METERS_average_height.img – Mean height of the canopy height model

canopy_30METERS_maximum_height.img – Maximum height of the canopy height model

canopy_30METERS_stddev_height.img – Standard deviation of the canopy height model

canopy_30METERS_rumple.img – Canopy height complexity (rugosity) calculated as the area of the canopy height model divided by the area of the ground surface (30×30 m).

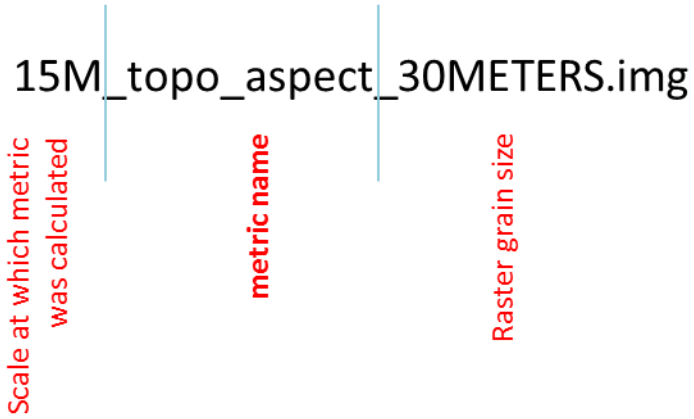
canopy_30METERS_FPV.img – Filled potential volume of the canopy height model. This is the proportion of the volume defined by the maximum canopy height and represents the proportion of the volume beneath the canopy height model's surface.

5.2.8 Topographic metrics

Directory:

- MASKED_TopoMetrics_30METERS

We calculated several topographic metrics using the LiDAR-derived ground surface model. (At the edges of the LiDAR acquisition, the USGS 10 m DEM is used to prevent abrupt edge effects in the calculation of these metrics.) Because the USGS DEMs extend beyond the LiDAR area, we include copies of topographic metrics masked to the area of the LiDAR. All the metrics are calculated over multiple window sizes to allow users to explore the effects of topographic scale on their topic of interest. Each file name includes the scale over which the metric was calculated and at the end of the name the grid cell size of the raster. The value in each grid cell reflects the topographic measurement made centered on that grid cell.



The metrics calculated at each scale are:

aspect -- degrees azimuth, 0 degrees at 12 o'clock increasing clockwise

curvature – integrated measurement of the surface curvature incorporating both plan and profile curvatures (see below)

elevation – in the units of the LiDAR acquisition projection

plancurv – plan curvature (across the slope)

profilecurv - profile curvature (along the slope)

sri – solar radiation index which combines information about the aspect, slope, and latitude into a single index that describes the amount of solar radiation theoretically striking an arbitrarily oriented surface during the hour surrounding noon on the equinox. The FUSION software calculates sri using a single latitude for the entire study area. We do post processing (see above) to recalculate SRI using the latitude for each grid cell.

tpi – Jenness topographic position index. We calculated slope positions using annuli of 100 m, 250 m, 500 m, 1000 m, and 2000 m radii using an algorithm that replicates the Topographic Position Index (TPI; Jenness, 2006; Weiss, 2001). More negative TPI values indicate a position towards a valley bottom, values near zero indicate flat areas or mid-slope, and more positive values indicate a hill or ridge top.

5.2.9 Tree Approximate Objects (TAOs)

Directory:

- Segments_0p75METERS

We used the 0.75 m filled, 3×3 smoothed canopy height model to identify tree approximate objects using a watershed segmentation algorithm. Our own analysis and the results of many other studies show that tree identification using segmentation algorithms generally varies between <50% to as high as ~80% depending on the degree of canopy closure and the variation in tree heights at any given location. Trees over topped by dominant trees are almost never identified.

The objects we identify should be thought of as tree clusters that may represent a single dominant tree or a closely spaced cluster of trees. We emphasize the 'approximate objects'. We believe these layers should be used to study tree clumping and are unlikely to be useful for generating tree lists.

To allow post processing without creating memory issues, the acquisition area is broken into rasters based on processing blocks and tiles (e.g., BLOCKn_C0000n_R0000n).

The following files are created for each tile:

BLOCKn_C0000n_R0000n_segments_HighPoints.shp – shape file showing high point for each TAO

BLOCKn_C0000n_R0000n_segments_Basin_Map.img – raster showing canopy area identified with each TAO. Each TAO has a unique identifying number within each tile, but identifiers are not unique across tiles.

BLOCKn_C0000n_R0000n_segments_Max_Height – raster with the maximum height of each TAO

5.2.10 Landscape (FRAGSTATS) Metrics

Directory:

- FRAGSTATSMetrics_90METERS

Because characteristic clump and opening patterns typically are expressed at scales from <0.4 ha to 1 ha (Larson and Churchill, 2012), we used 90×90 m (0.81 ha) areas to measure patterns. The analysis used a gliding window approach in which the TAOs and openings in a 0.81 ha area surrounding the center of a 30×30 m grid cell were analyzed. The center point was then moved to the next grid cell and the analysis was repeated, resulting in analyses that included openings and TAOs that overlapped with those used for adjacent grid cells. We did this so that the analysis grain matched the 30×30 m grain of the data sets used in this study and commonly used by the Park Service. The values recorded for each 30×30 m grid cell, therefore, can be thought of as the tree clump-opening context centered on a 30 m grid.

We used a custom Python routine to process the CSMs and TAOs by tile, compute landscape metrics for 90×90 m areas centered on each 30×30 m grid cell, and assemble single 30×30 m rasters covering the entire study area for each metric. The landscape metrics were computed using the same formulas as those used by the FRAGSTATS software as checked by computing the metrics using both our software and the FRAGSTATS software on a set of samples and ensuring that the results were identical.

Formulas for calculating the FRAGSTATS metrics can be found in the FRAGSTATS documentation (McGarigal et al., 2012).

Each file name includes an abbreviation for the metric, the stratum height breaks (if appropriate to the metric), and '90m' to reflect the length of one edge of the grid cell over which the measurements were done.

Class area metrics

CA.Gap_90m.img – area in openings (gaps) with no returns <2 m in height

CA.2to8_90m.img

CA.8to16_90m.img

CA.16to32_90m.img

CA.gt32_90m.img

CA.Canopy_90m.img – area in canopy with at least one return ≥2 m in height

Patch number, where a patch can be an opening (gap), tree cluster, or tree clump

Npatch.Gap.lrgpatch_90m.img – number of openings (gaps) ≥10 m^{-2}

Npatch.Gap_90m.img – number of openings (gaps) of any size

Ncluster_90m.img – tree cluster number

Npatch.2to8_90m.img – tree cluster number

Npatch.8to16_90m.img – tree cluster number

Npatch.16to32_90m.img – tree cluster number

Npatch.gt32_90m.img – tree cluster number

Npatch.Canopy_90m.img – tree clump number

True_ClusterRatio_90m.img – number of clusters divided by the area in canopy

Total edge

TE.Allstrata_90m.img

TE.Canopy_90m.img

TE.Gapcanopy_90m.img

TE.Gap_90m.img

TE.2to8_90m.img

TE.8to16_90m.img

TE.16to32_90m.img

TE.gt32_90m.img

REFERENCES

Agee, J.K., 1993. Fire ecology of Pacific Northwest forest. Island Press, Washington D.C.

Attiwill, P.M., 1994. The disturbance of forest ecosystems: the ecological basis for conservative management. For. Ecol. Manage. 63, 247–300. doi:10.1016/0378-1127(94)90114-7

Bekker, M.F., Taylor, A.H., 2010. Fire disturbance, forest structure, and stand dynamics in montane forests of the southern Cascades, Thousand Lakes Wilderness, California, USA. ECOSCIENCE 17, 59–72. doi:10.2980/17-1-3247

Bekker, M.F., Taylor, A.H., 2001. Gradient analysis of fire regimes in montane forests of the southern Cascade range, Thousand Lakes Wilderness, California, USA. PLANT Ecol. 155, 15–28. doi:10.1023/A:1013263212092

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T., Solberg, S., 2010. Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. Remote Sens. Environ. 114, 911–924. doi:10.1016/j.rse.2009.12.004

Breiman, L., 2001. Random Forests. Eur. J. Math. 45, 5–32. doi:doi:10.1023/A:1010933404324

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and regression trees. Chapman & Hall, New York.

Breiman, T., Cutler, A., Classification, D., 2012. Package " randomForest ."

Comer, P., Faber-Lanendoen, D., Evans, R., Gawler, S., Josse, C., Kittel, G., Menard, S., Pyne, M., Reid, M., Schulz, K., Snow, K., Teague, J., 2003. Ecological Systems of the United States: A Working Classification of U.S. Terrestrial System. Arlington, VA.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792. doi:10.1890/07-0539.1

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. Int. J. Climatol. 28, 2031–2064. doi:10.1002/joc.1688

Dingman, S.L., 2002. Physical Hydrology. Prentice Hall, Upper Saddle River, NJ.

Erdody, T.L., Moskal, L.M., 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sens. Environ. 114, 725–737. doi:10.1016/j.rse.2009.11.002

Forrestel, A., Stephens, S.L., 2011. Mixed Severity Fire Regimes & Forest Structure at Crater Lake National Park, Oregon, USA. Draft Final Project Report: CESU Cooperative Agreement H8C07030001. Berkeley, CA.

Franklin, J.F., Pelt, R. Van, Van Pelt, R., 2004. Spatial aspects of structural complexity in old-growth forests. J. For. 102, 22–28.

Franklin, J.F., Spies, T. a., Pelt, R. Van, Carey, A.B., Thornburgh, D. a., Berg, D.R., Lindenmayer, D.B., Harmon, M.E., Keeton, W.S., Shaw, D.C., Bible, K., Chen, J., 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. For. Ecol. Manage. 155, 399–423. doi:10.1016/S0378-1127(01)00575-8

Franklin, J.F., Van Pelt, R., 2004. Spatial aspects of structural complexity in old-growth forests. J. For. 102, 22–28.

Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., Hastie, T. J., R. J. Tibshirani, and J.H.F., Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., Hastie, T. J., R. J. Tibshirani, and J.H.F., 2001. The elements of statistical learning: data mining, inference and prediction, The elements of statistical learning: data mining, inference, and prediction. Springer Series in Statistics. Springer, New York.

Heyerdahl, E.K., Loehman, R. a, Falk, D. a, 2014. Mixed-severity fire in lodgepole pine dominated forests: are historical regimes sustainable on Oregon' s Pumice Plateau, USA? Can. J. For. Res. Can. Rech. For. 44, 593–603. doi:10.1139/cjfr2013-0413

Jenness, J., 2006. Topographic Position Index (tpi_jen. avx) extension for ArcView 3. x, v. 1.3 a. Jenness Enterprises.

Jeronimo, S., 2015. LiDAR individual tree detection for assessing structurally diverse forest landscapes. University of Washington.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., Wu, J.C., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sens. 4, 950–974. doi:10.3390/rs4040950

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. Can. J. For. Res. 40, 774–787. doi:10.1139/X10-064

Kane, V.R., Gersonde, R.F., Lutz, J.A., McGaughey, R.J., Bakker, J.D., Franklin, J.F., 2011. Patch dynamics and the development of structural and spatial heterogeneity in Pacific Northwest forests. Can. J. For. Res. 41, 2276–2291. doi:10.1139/x11-128

Kane, V.R., Lutz, J.A., Alina Cansler, C., Povak, N.A., Churchill, D.J., Smith, D.F., Kane, J.T., North, M.P., 2015. Water balance and topography predict fire and forest structure patterns. For. Ecol. Manage. 338, 1–13. doi:10.1016/j.foreco.2014.10.038

Kane, V.R., McGaughey, R.J., Bakker, J.D., Gersonde, R.F., Lutz, J.A., Franklin, J.F., 2010. Comparisons between field- and LiDAR-based measures of stand structural complexity. Can. J. For. Res. 40, 761–773. doi:10.1139/X10-024

Kane, V.R., North, M.P., Lutz, J.A., Churchill, D.J., Roberts, S.L., Smith, D.F., McGaughey, R.J., Kane, J.T., Brooks, M.L., 2014. Assessing fire effects on forest spatial structure using a fusion of Landsat

and airborne LiDAR data in Yosemite National Park. Remote Sens. Environ. 151, 89–101. doi:10.1016/j.rse.2013.07.041

Keating, K.A., Gogan, P.J.P., Vore, J.M., Irby, L.R., 2007. A Simple Solar Radiation Index for Wildlife Habitat Studies. J. Wildl. Manage. 71, 1344–1348. doi:10.2193/2006-359

Key, C.H., Benson, N.C., 2006. Landscape assessment: ground measure of severity, the Composite Burn Index, and remote sensing of severity, the Normalized Burn Ratio., 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. Rocky Mountain Research Station, Fort Collins, Colorado, USA, pp. 1–51.

Larson, A.J., Churchill, D., 2012. Tree spatial patterns in fire-frequent forests of western North America, including mechanisms of pattern formation and implications for designing fuel reduction and restoration treatments. For. Ecol. Manage. 267, 74–92. doi:10.1016/j.foreco.2011.11.038

Lefsky, M.A., Hudak, A.T., Cohen, W.B., Acker, S.A., 2005. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. Remote Sens. Environ. 95, 517–531. doi:10.1016/j.rse.2005.01.004

Lutz, J.A., van Wagtendonk, J.W., Franklin, J.F., 2010. Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. J. Biogeogr. 37, 936–950. doi:10.1111/j.13652699.2009.02268.x

McGarigal, K., Cushman, S.A., Ene, E., 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. [WWW Document]. URL http://www.umass.edu/landeco/research/fragstats/fragstats.html

McGaughey, R.J., 2014. FUSION/LDV: Software for LiDAR Data Analysis and Visualization Version 3.42.

McNeil, R.C., 1975. Vegetation and Fire History of a Ponderosa Pine- White Fir Forest in Crater Lake National Park. Oregon State University.

Miller, C., Urban, D.L., 1999. A model of surface fire, climate and forest pattern in the Sierra Nevada, California. Ecol. Modell. 114, 113–135. doi:10.1016/S0304-3800(98)00119-7

Muscolo, A., Bagnato, S., Sidari, M., Mercurio, R., 2014. A review of the roles of forest canopy gaps. J. For. Res. 25, 725–736. doi:10.1007/s11676-014-0521-7

National_Park_Service, 2001. Fire monitoring handbook. Boise, ID.

Parker, G.G., Harmon, M.E., Lefsky, M. a., Chen, J., Pelt, R. Van, Weis, S.B., Thomas, S.C., Winner, W.E., Shaw, D.C., Frankling, J.F., 2004. Three-dimensional Structure of an Old-growth Pseudotsuga-Tsuga Canopy and Its Implications for Radiation Balance, Microclimate, and Gas Exchange. Ecosystems 7, 440–453. doi:10.1007/s10021-004-0136-5

R Core Team, 2014. R Language Definition V. 3.1.1.

Shuffield, C.D., 2010. Overstory composition and stand structure shifts within inter-mixed ponderosa pine and lodgepole pine stands of the south-central Oregon pumice zone. Oregon State University.

Smith, T.M., Urban, D.L., 1988. Scale and resolution of forest structural pattern. Vegetatio 74, 143–150. doi:doi:10.1007/BF00044739

Taylor, A.H., 2000. Fire regimes and forest changes in mid and upper montane forests of the southern Cascades, Lassen Volcanic National Park, California, USA. J. Biogeogr. 27, 87–104. doi:10.1046/j.13652699.2000.00353.x

Taylor, A.H., Solem, M.N., 2001. Fire Regimes and Stand Dynamics in an Upper Montane Forest Landscape in the Southern Cascades, Caribou Wilderness, California. J. Torrey Bot. Soc. 128, 350. doi:10.2307/3088667

Taylor, H.W., 1995. Forest expansion and climate-change in the mountain hemlock (Tsuga-Mertensiana) zone, Lassen Volcanic National Park, California, USA. Arct. Alp. Res. 27, 207–216. doi:10.2307/1551951

Thornthwaite, C.W., Mather, J.R., 1955. The Water Balance. Publ. Climatol. 8, 1–104.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., Maltamo, M., 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry 85, 27–40. doi:10.1093/forestry/cpr051

Vincent, L., Soille, P., 1991. Watersheds in digital spaces - an efficient algorigthm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13, 583–598.

Weiss, A., 2001. Topographic position and landforms analysis, in: Poster Presentation, ESRI User Conference, San Diego, CA.

Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth Surf. Process. Landforms 12, 47–56. doi:10.1002/esp.3290120107