Using LiDAR to Assess Forest Structure and Fuels at the Whiskeytown National Recreation Area, Oregon Cave National Monument and Preserve, and Redwood National and State Parks

Miles LeFevre, Van Kane

UW Research Group,
School of Environmental and Forest Sciences,
University of Washington, Seattle, WA

Final Report to the National Park Service

Task Agreement P15AC01626
Cooperative Agreement H8W07110001

This research project was conducted through the
Pacific Northwest Cooperative Ecosystem Studies Unit

Contents

0. Executive Summary

This report describes the analysis, findings, and products of a study funded by the National Park Service (NPS), and carried out by the Forest Restoration and Resilience at Multiple Scales research group of the University of Washington (UW). The study was concerned with the analysis of Airborne Light Detection and Ranging (LiDAR) data acquisitions covering the Whiskeytown National Recreation Area, Oregon Caves National Monument site, and Redwood National and State Parks.

Primary objectives of the study were to (1) use airborne LiDAR to map forest and woodland structure within the boundaries of the LiDAR data; (2) measure the heterogeneity of forest and woodland structure and to define structure classes for the study areas; (3) use LiDAR-derived ground models to create maps of topography and slope position to aid in mapping topographic features that might influence fire severity; and (4) to associate forest structure classes with fuel models currently mapped across the study area, with elevation, and with topographic features and slope position.

The Oregon Caves National Monument site dataset delivered to the UW by NPS consisted of a 2 m ground surface model and 2 m canopy surface model. Our analysis was limited by the resolution of the dataset provided and the lack of a LiDAR point dataset. This prohibited the identification of fine scale and sub canopy structural patterns and characteristics. We performed a hierarchical classification using the proportions of area occupied by each height strata present in each grid cell. The results of this analysis were a set of 8 structural classes characterizing the heterogeneity of vertical forest structure present in the Oregon Caves National Monument site study area. This product is provided as a classified raster.

The Whiskeytown National Recreation Area and Redwood National and State Parks datasets delivered to the UW by NPS included LiDAR point datasets. We processed these data using the FUSION, FRAGSTATS and R software packages. The results of this analysis were two sets of rasters describing vertical and horizontal forest structure across the study areas. We used these datasets to perform hierarchical classifications, producing two sets of structural classes characterizing the heterogeneity of vertical and horizontal forest structure present for each of the study areas. These products are provided as two sets of directories containing rasters for the metrics and structure classes.

Additional analysis was carried out for the Whiskeytown National Recreation Area dataset. Whiskeytown National Recreation Area managers provided the UW with plot level data describing fuel loadings across the study area. Our analysis was limited by a lack of accuracy in plot locations provided for this dataset, as well as a relatively small sample size. We developed regression and classification models for each fuel loading metric provided using LiDAR, SURGGO, and PRISM datasets as predictor metrics. Models were predicted across the study area using raster datasets of predictor metrics. The results of this analysis vary in model strength. These products are provided as a set of rasters.

1. Introduction

LiDAR is increasingly being utilized in landscape analysis. Novel methods of LiDAR analysis can help land managers to answer site specific questions and increase management effectiveness. The Whiskeytown National Recreation Area (WHIS), Oregon Caves National Monument (ORCA), and Redwood National and State Parks (REDW) are managed for recreational activities. With high rates of visitor traffic and close proximities to urban areas, managers require tools to maintain aesthetic value and mitigate the risk of large scale fires.

In 2004, LiDAR data were collected for the ORCA study area. ORCA LiDAR data delivered to the UW group consisted of a ground surface model and canopy surface model. In 2010 and 2011, LiDAR data were collected for the WHIS. WHIS data consisted of a ground surface model and a tiled las point cloud. The National Park Service funded a study to explore methods for describing forest conditions using LiDAR data. Major objectives of this study were the development of fine grain models of forest structural conditions and ground fuel loadings.

We processed LiDAR data for the WHIS and REDW to produce a set of metrics significant to forest structure and local managers. Hierarchical clustering produced a set of distinct structural classes for all three study areas. Random forest modeling and prediction produced a rasterized description of the classes across the study area. For the WHIS, an additional set of metrics were selected with regards to fuel loading. Random forest models were developed using these metrics to predict fuel loadings for field surveyed plots and across the study area.

2. Methods

2.1 Study Area

The WHIS is a 42,500 ac area centered on Whiskeytown Lake, a 3,200 ac reservoir. The 54,000 ac study area includes the 42,500 ac WHIS as well as a region north of the WHIS. The study area possesses a Hot-summer Mediterranean climate, characterized by hot dry summers and mild wet winters (Kottek et al. 2006). The majority of the 63.2 in (1606 mm) of precipitation that falls on the study area occurs as rain between November and March, with seasonal mean highs and lows of 0.2 in (5.7 mm) in July and 11.3 (288 mm) in December. A total man annual precipitation of 1.6 in (41.1 mm) falls as snow. Seasonal fluctuations in temperature are moderate with seasonal mean highs and lows of 90.5°F (32.5 °C) in July and 33.3°F (0.76°C) in January (PRISM Climate Group 2015).

The WHIS is geographically unique as being a transitional zone between multiple distinct physiographic features. The region is defined by its proximity to the Klamath, Cascade, and Costal mountain ranges. Soils present in the park include Mariposa-Maymen, Auburn-Brandy, Chawanakee-Corbett, Goulding, and Josephine-Sites associations (Stuart and Lee 2003).

The WHIS was historically mined for gold, and was the site of a significant homesteading presence. The region has and continues to experience high frequencies of fire disturbance.

Vegetation is varied spatially. At low elevations, species present include *Arctostaphylos viscida, Heteromeles arbutifolia, Toxicodendron diversilobum, Quercus kelloggii, Quercus chrysolepi,* as well as *Pinus sabiniana* and *Pinus attenuata*. Southern slopes are often chaparral dominated shrub land. At higher elevations, species present include *Pseudotsuga menziesii*, *Lithocarpus densijlorus var. densijlorus*, *Pinus ponderosa*, and *Pinus lambertiana* (Stuart and Lee 2003).

2.2 Data Collection

2.2.1 LiDAR

Watershed Sciences, Inc. (Corvallis, OR) collected LiDAR data of 17,526 acre core area of WHIS south of the reservoir on July 8th and 9th, 2010. Data were collected using dual-mounted Leica ALS50 Phase II sensors in a Cessna Caravan 208B. The acquisition was flown at 1100 m above ground level and captured a scan angle of ±14 degrees, resulting in an average pulse density of 15.86 points/m^2.

A subsequent acquisition of the remaining 41,042 acre area of the WHIS study area was flown on July 20th and 21st 2011. Data were again collected using dual-mounted Leica ALS50 Phase II sensors in a Cessna Caravan 208B. The acquisition was flown at 1200 m above ground level and captured a scan angle of ±12 degrees, resulting in an average pulse density of 13.20 points/m^2.

Deliverables of both acquisitions from Watershed Sciences included all returns and ground returns stored in LAS v.1.2 format, 1 m digital elevation models stored in ESRI GRID format, as well as vector data describing the acquisition process.

LiDAR data for the REDW and ORCA study areas were delivered to the UW group without documentation from the vendor.

2.2.2 Fuels Data

From 2000 to 2007 WHIS technicians installed a set of 59 plots across the study area. 52 plots were installed following the standard fire monitoring handbook protocol. Plots were 20 m by 50 m, and oriented at a random azimuth. Four fuels transects were established along at fixed intervals along the centerline, running 50 ft and oriented on a random azimuth. The remaining 7 plots were established as circular plots with one 50 ft transect, emanating at a random azimuth from plot center. For all plots transect data was summarized by plot. Plot centers were located with consumer grade handheld Garmin GPS devices, which frequency have error rates of >10 m under forest canopies. These data as well as LANDFIRE 40 fuel models were delivered to the UW by WHIS managers.

2.2.3 Climate Metrics

The PRISM Climate group is a project operating through the Oregon State University (PRISM Climate Group 2015). The group uses climate observations from various sources in order to develop spatial climate datasets which describe temporal climatic patterns. We obtained PRISM data using the ClimateWNA tool, produced by the Centre for Forest Conservation Genetics (CFCG). ClimateWNA uses bilinear interpolation and elevation adjustment to downscale the baseline climate data, which is provided by PRISM at an 800 m by 800 m scale (Wang et al. 2012). The acquired PRISM data sets describe the 30 year normal values for the plots installed in 2015, as well as the entire study area. These values include annual precipitation, monthly mean temperatures, and annual mean temperature.

Topographic data were combined with PRISM data as well as USGS soil survey data to generate layers which estimate the annual actual evapotranspiration (AET) and annual climatic water deficit (Deficit) of the study area (United States Department of Agriculture, Soil Survey Staff, and Natural Resources Conservation Service 2016). AET can be considered the evaporative water loss from a site covered by a hypothetical standard crop, constrained by the current water availability. Deficit can be considered the difference between AET and the potential evapotranspiration, or AET without water availability constraints. We used methods outlined by Lutz et al. (2010) in performing these calculations, which are based on established methods (Stephenson 1990, 1998; Thornthwaite and Mather 1955).

2.3 Analysis

2.3.1 FUSION Metrics

We processed the LiDAR point data for the WHIS and REDW study areas. Data were processed using the USDA Forest Service's FUSION software package (Mcgaughey 2015) (version 3.51 http://forsys.cfr.washington.edu/fusion.html). All vegetation analyses were done with LiDAR return elevations normalized to height above ground to reflect height of vegetation canopy above ground. We produced a set of vertical forest structure metrics, which were calculated and reported using a 30 m by 30 m cell size. Each acquisition was processed separately and products for each study area were subsequently merged to produce a single file for each metric across each study area. Metrics for the distribution of the canopy profile were calculated as the heights at which a percentile of returns (e.g., 95^{th} percentile height) >2 m occurred. We measured the structural heterogeneity of the forest canopy with a measure of canopy rugosity, rumple, which was calculated from a canopy surface model (CSM) created using the maximum return height within each 1 m by 1 m grid cell smoothed with a 3 by 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 m, 2 m to 8 m, 8 m to 16 m, 16 m to 32 m, and >32 m. For canopy cover >2 m, the 95^{th} 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.

2.3.2 Tree Approximate Object Identification

We analyzed the WHIS and REDW datasets using individual tree detection. 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 (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 reflect their presence.

We carried out the segmentation using TreeSeg, a tool developed for the FUSION LiDAR software package. 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 (Figure 1). 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 study areas. 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 (Jeronimo 2015).

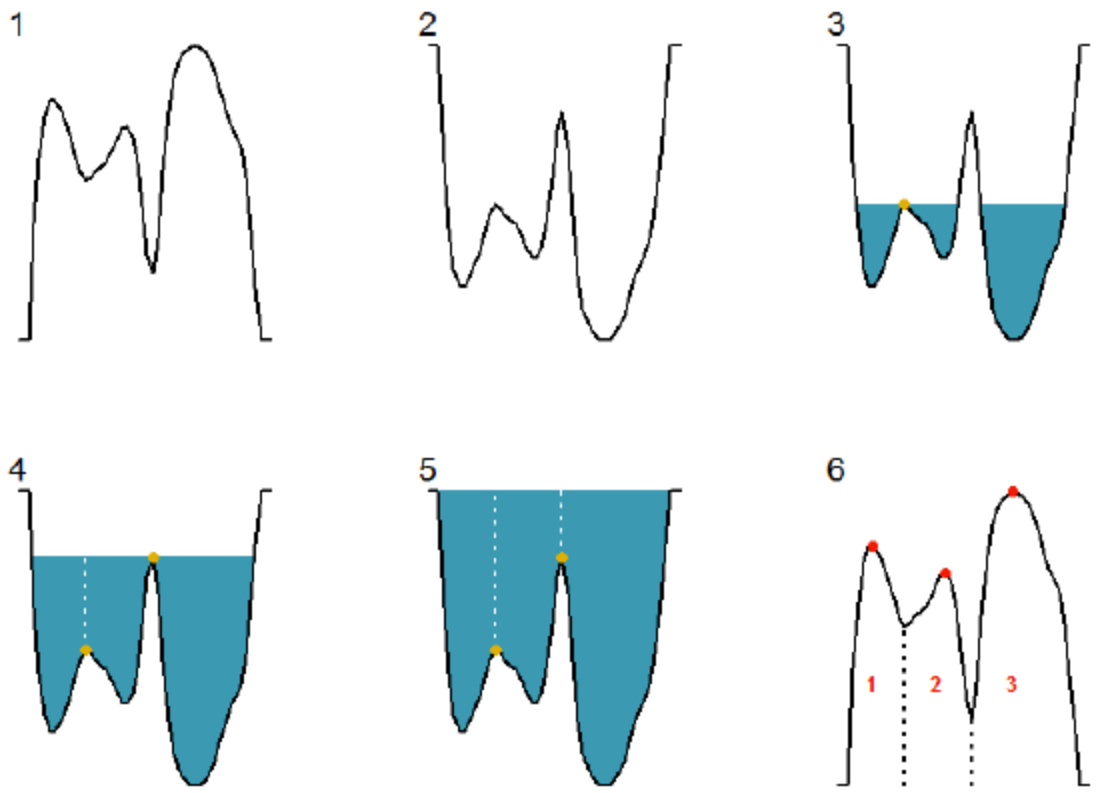


Figure 1. Overview of the watershed segmentation algorithm, presented here as a 2D concept, but actually performed in 3D. (1) The canopy surface model is draped over the LiDAR point cloud. (2) The canopy surface model is inverted (3-5). The surface is imagined to be made of a permeable material, and is slowly lowered into water. Any time two separate pools come into contact (yellow 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).

Because high resolution CSMs covering large study areas can be too large for many programs to load and process, the CSMs for the study areas were divided into tiles. For each tile, we identified TAO's and openings using a 0.75 m by 0.75 m 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) an ESRI ArcGIS shapefile showing the location of each TAO maximum height, its maximum height, and the area assigned to each TAO.

2.3.3 Tree Clump and Opening Metrics

We analyzed the WHIS and REDW datasets for tree clump and opening patterns. We reclassified the 0.75 m by 0.75 m rasters that recorded the maximum height for each TAO and openings

(section 2.3.2) into height strata: <2 m, 2 m to 8 m, 8 m to 16 m, 16 m to 32 m, >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 2).

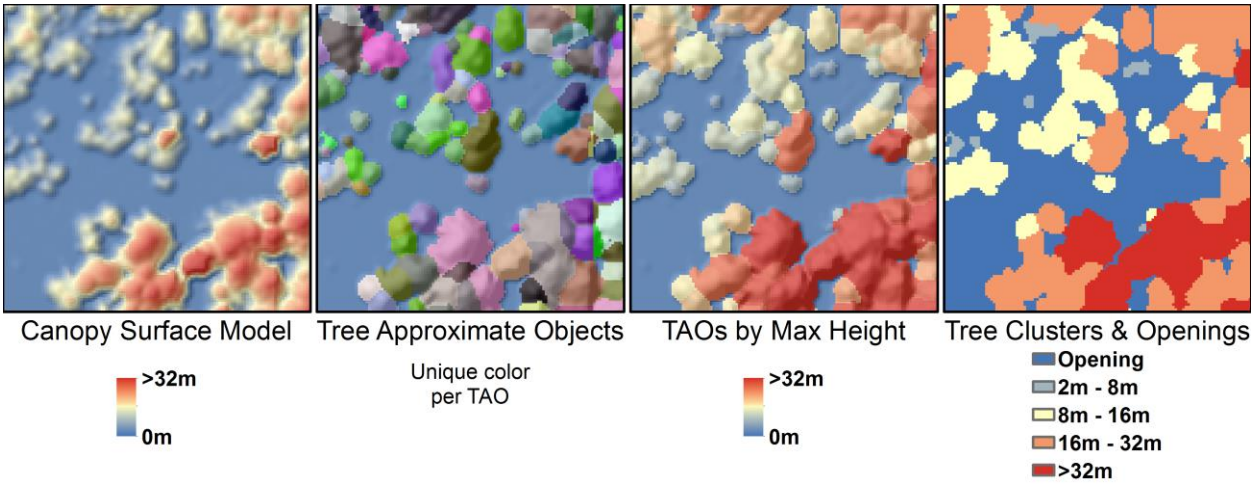


Figure 2. Identification and measurement of tree clusters, tree clumps, and openings from LiDAR-derived canopy surface models was done in four steps: (1) create a 0.75 m resolution canopy surface model, (2) identify immediately dominant tree approximate objects (TAOs) using a watershed segmentation, (3) assign the area of each TAO the maximum height of that TAO, (4) classify each TAO by height strata (2 to 8, 8 to 16, 16 to 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.

The result of this processing was to create a classified raster for each 90 m by 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, Cushman, and Ene 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 used a custom Python routine to process the CSMs and TAOs by tile, compute landscape metrics for 90 m by 90 m areas centered on each 30 m by 30 m grid cell, and assemble single 30 m by 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.

2.3.4 Forest Structural Classification

We classified forest structure for the WHIS, ORCA, and REDW datasets. For the WHIS and REDW datasets, a set of metrics were selected to describe forest structural conditions. Metrics were subset based on site conditions and local knowledge. We further subset descriptors by identification of highly correlated metrics though PCA values. For the ORCA dataset, proportions of pixel area in classified height strata were used in place of FUSION and FRAGSTATS derived metrics.

To account for collinearity between the metrics, we used the principle components analysis (PCA) (Legendre and Legendre 2012) axes of variation to define the structure classes. We used hierarchical clustering to split the sample set into 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 3.4.0) (R Core Team 2017) for this analysis. We used the classified random sample of up to 30,000 grid cells as training data to classify the vertical structure of all grid cells within the study areas using the random forest algorithm (Breiman 2001) predict function in the R statistical randomForest package.

2.3.5 Whiskeytown Ground Fuels Modeling

Ground fuels analysis and modeling was carried out for the WHIS study area. A cursory exploratory analysis of the relationship between derived structural classes and existing LANDFIRE 40 classification was performed. Histograms of structure classes by LANDFIRE classes and vice versa were generated.

We used the fuels plot locations provided by WHIS managers to clip 15, 30, and 60 meter radius circles from the LiDAR point data. We processed these data using the FUSION CanopyModel, TreeSeg, and CloudMetrics tools to create a set of metrics specific to the areas of the fuels plots. We replicated the methods described in section 2.3.3 to produce a set of metrics describing the proportion of land area covered by each canopy height class as well as the number of unique patches. We assigned to each plot the values of the overlapping 30 m by 30 m grid cells from the metric rasters for topographic and climatic variables.

We used random forest models for both regression with continuous response variables and classification with categorical response variables. We used the randomForest and partialPlot functions in the randomForest package for the R statistical package (release 3.4.0) (R Core Team

2017) to develop and analyze our models. For each fuel loading metric reported, continuous models were generated, as well as classified "high/low" models predicting if plot values were above or below the median value for each metric.

A parsimonious set of predictor metrics were selected for each model using forward stepwise selection. We used the VSURF packages to select from the full set of metrics a parsimonious set of predictors for each fuel loading response variable for both regression and classification (Genuer, Poggi, and Tuleau-malot 2015). Although VSURF attempts to avoid selecting redundant metrics, we were able to apply practical knowledge of the selected metrics to further reduce our set of independent variables. The final set of metrics for each model was a priori selections from VSURF results.

We used the corresponding 30 m by 30 m rasters of each predictor metric to predict fuel loading values across the extent of the study area. We generated two predicted models using both the regression and classification models for each fuel loading response metric.

3. Results

3.1 Redwood National and State Parks

3.1.1 Forest Structure Classes

We identified seven statistically distinct forest structure classes using the area of openings (CSM <2 m) and area of TAOs in five height strata (2 m to 8, 8 m to 16 m, 16 m to 32 m, 32 m to 48 m, and >48 m) within 90 m by 90 m areas across the study area. Six of the structure classes were in statistically similar pairs of classes (Figure 3; Figure 4).

We found that structure classes could be sorted into three distinct groups with regards to dominant canopy height (Table 1). The tallest classes (2 and 3) had the largest proportions of canopy area above 48 m. A second class of height strata (1 and 5) had moderate dominant canopy height values, and the majority of their canopy area in height strata below 32m. The third height strata (4, 6 and 7) had the lowest canopy max height values, and the majority of their canopy area in height strata < 16m.

Table 1. Key characteristics for the structure classes identified from patterns of canopy and openings across the Redwoods study area

Class	Dominant Height	Stratification	Density
1	Moderate	Single Story	High
2	Tall	Multi Story	High
3	Tall	Single Story	High
4	Short	Multi Story	Moderate
5	Moderate	Multi Story	Moderate
6	Short	Bottom Loaded	Low
7	Short	Single Story	Low

Structure classes could be sorted into groups depending on the distribution of canopy area across height strata (Figure 5). Multi strata structure classes (2, 4, and 5) had significant components of canopy area across three or more distinct height classes. These structure classes tended to have moderate to high cover. Single strata structure classes (1, 3, and 7) had the majority of their canopy area in one or two adjacent height strata. These structure classes tended to have very high or very low canopy cover. One strata structure class (6) had significant canopy area in the strata of 8 m to16 m and 2 m to 8 m. This structure class had a negative correlation between proportional canopy strata cover and canopy strata height.

Structure class 1 was characterized by having the highest overall cover (Figure 6). The majority of its canopy was occupied by trees in the height strata of 16 m and 32 m. Structure class 5 was statistically most similar to structure class 1, with the major distinction of having a lower overall cover and more canopy area in lower height strata.

Structure class 2 had the second highest overall cover, and had high proportions of canopy cover in canopy strata from 16 m to above 48 m. Structure class 3 was statistically most similar to structure class 2, with the distinction of having more area in large gaps, and the highest proportion of canopy area above 48 m.

Structure class 4 was characterized by having a moderate proportion of canopy gaps and a majority of canopy area between 2 m and 32 m. Structure class 5 was statistically most similar to structure class 4, with the major distinctions of having less canopy area between 8 m and 16 m and more canopy area between 16 m and 32 m.

Structure class 7 had the highest proportion of canopy gaps, often reaching 100 percent. This structure class likely indicates non forested areas.

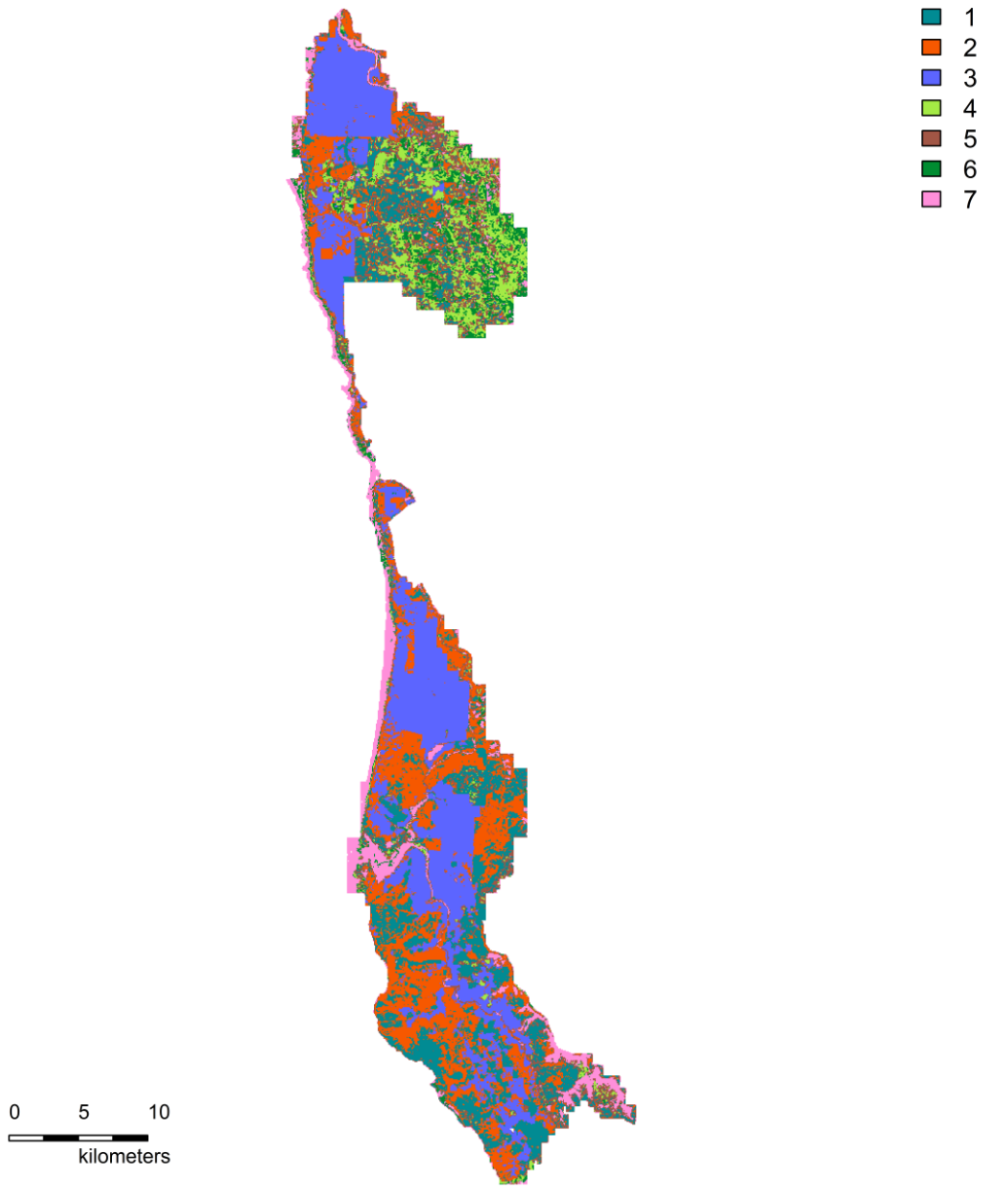


Figure 3. Map of structure classes identified for the Redwoods study Area.

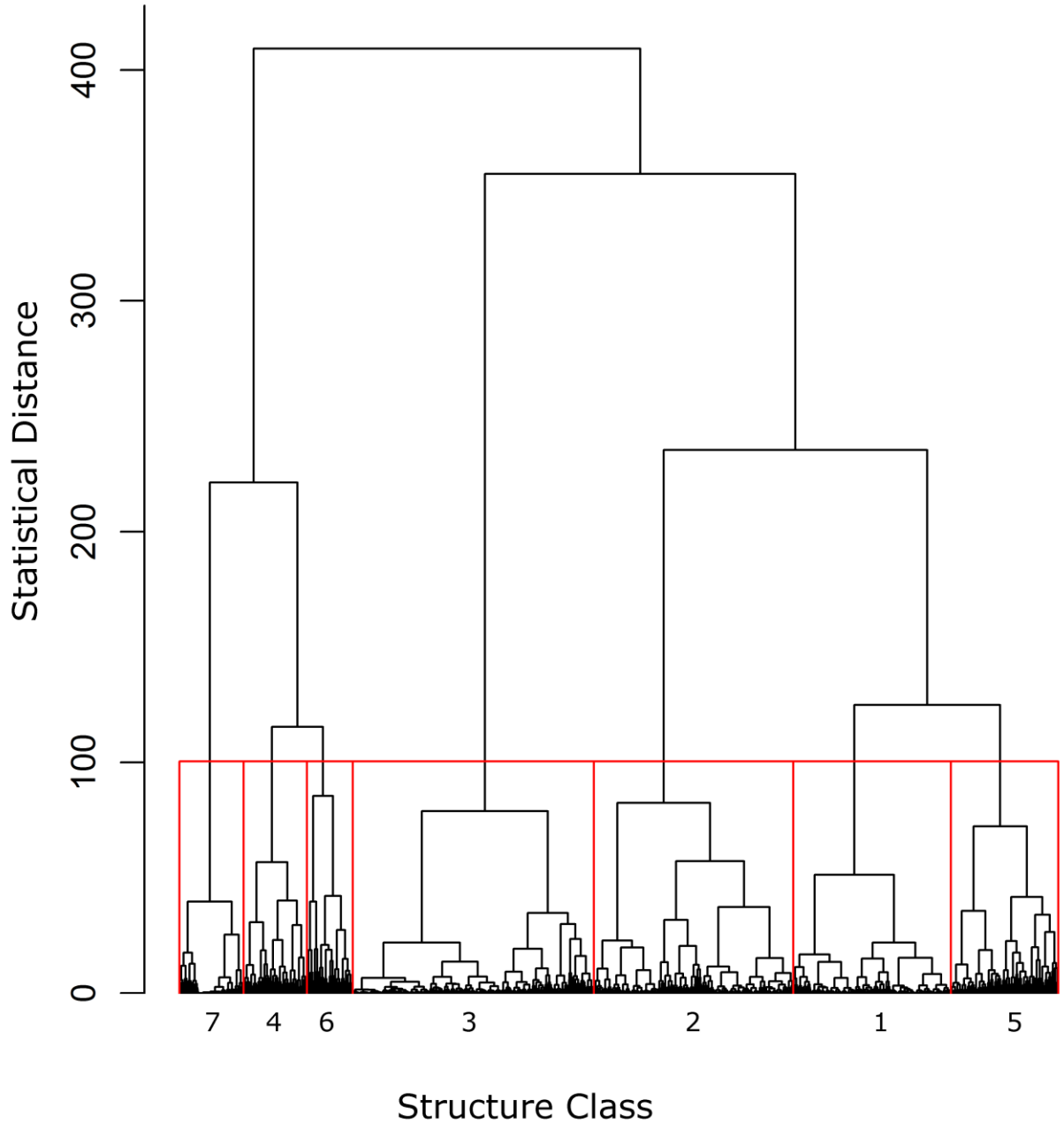


Figure 4. Dendogram showing statistical distance between structure classes identified in the Redwoods study area. Clustering was implemented using Ward's (1963) clustering criterion. Clustering is additive, where classes which join lower town on the tree are more statistically significant than classes that join higher on the tree.

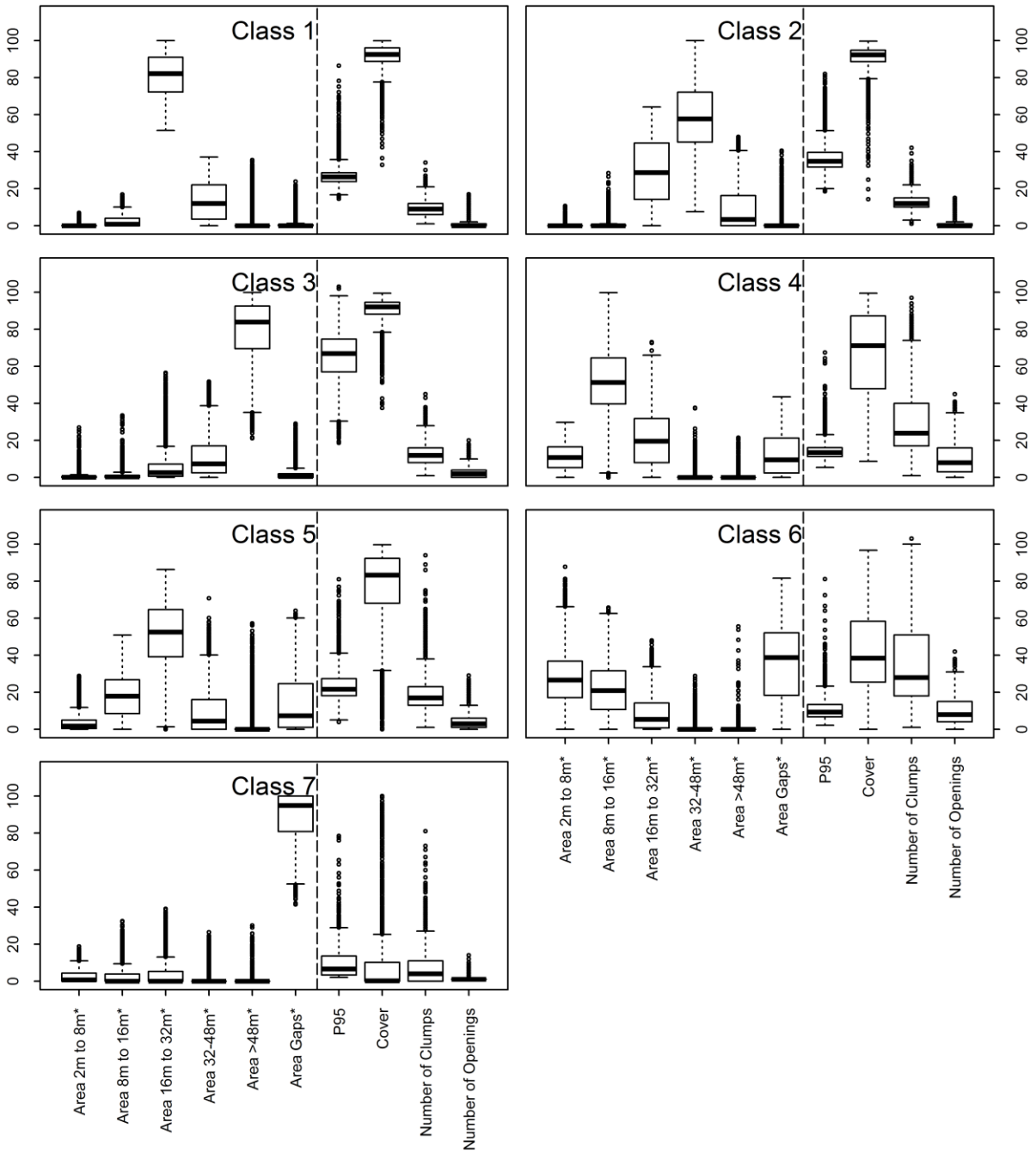


Figure 5. Characteristics of structure classes identified in the Redwoods study area. Scale is either percentage, meters in height, or count of large tree clumps or large gaps. Boxplots show ranges of values for metrics used to define each class (left of dotted line in each panel) and metrics that provide additional information on each of the structure class but which were not used to define the structure classes (right of dotted line in each panel). Area Gaps 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.

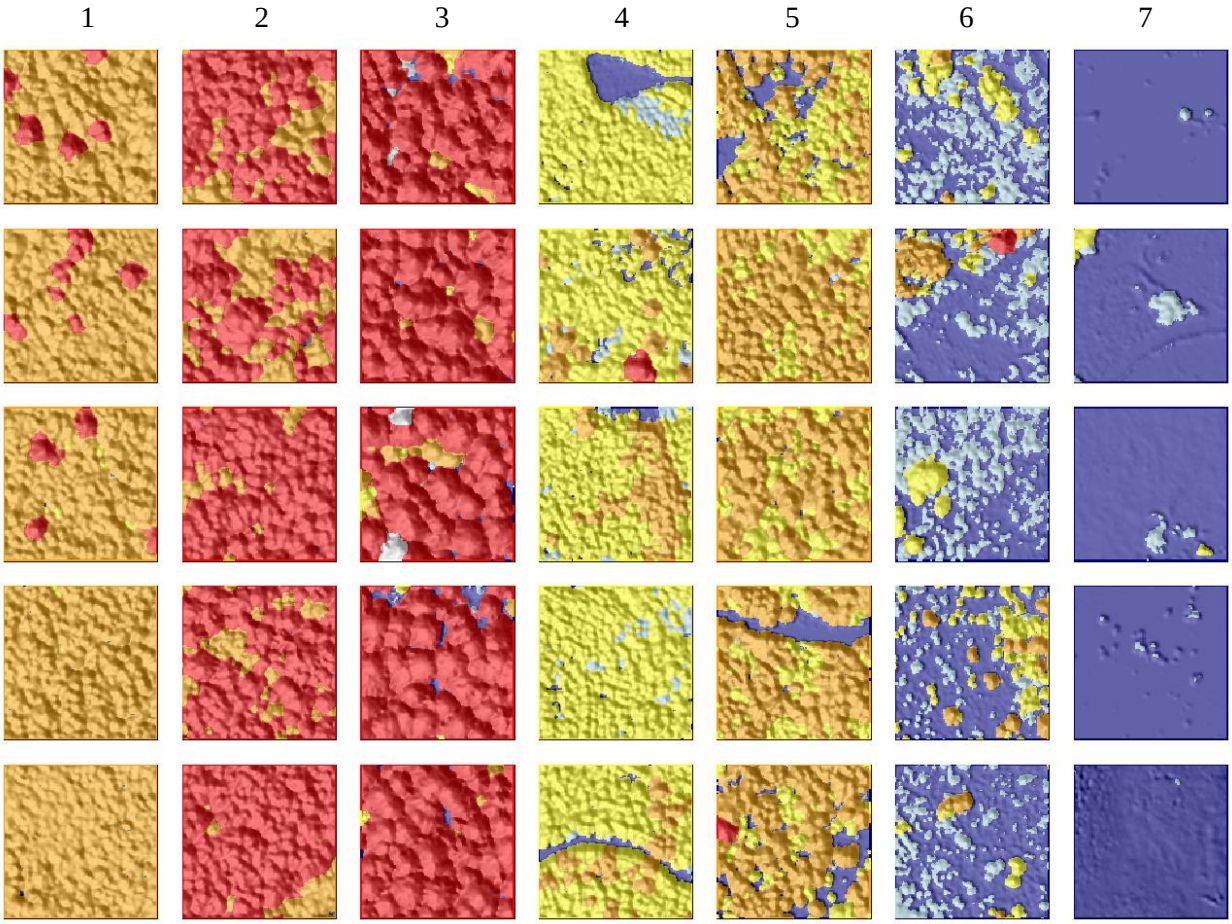


Figure 6. Examples of LiDAR structure metrics 30 m by 30 m areas. Each column contains 5 randomly selected pixels representing one of the 7 structure classes defined for the Redwoods study area. Color corresponds to canopy height, averaged across TAO extent. Warmer colors correspond to taller canopy areas while cooler colors correspond to shorter canopy areas.

3.2 Oregon Caves National Monument

3.2.1 Forest Structure Classes

We identified eight statistically distinct forest structure classes using the area of openings (CSM <2 m) and area of TAOs in five height strata (2 m to 8, 8 m to 16 m, 16 m to 32 m, 32 m to 48 m, and >48 m) within 90 m by 90 m areas across the study area (Figure 7). Six of the structure classes were in statistically similar pairs of classes (Figure 8).

We found that structure classes could be sorted into three distinct groups with regard to dominant canopy height (Table 2). The tallest classes (6, 7, and 8) had the largest proportions of canopy area above 16 m. Classes 7 and 8 had high proportions of canopy area above 32 m, and class 8 had the highest proportion of class area above 48 m. A second class of height strata (3, 4, and 5) had moderate dominant canopy height values, and the majority of their canopy area in height strata below 32 m. The third height strata (1 and 2) had the lowest canopy max height values, and the majority of their canopy area in height strata < 8 m.

Structure classes could be sorted into groups depending on the distribution of canopy area across height strata (Figure 9). Multi strata structure classes (5, 6, 7, and 8) had an even distribution of canopy area across three or more distinct height classes. These structure classes tended to have moderate cover. Single strata structure classes (1, 2, and 3) had the majority of their canopy area in one or two adjacent height strata. Structure classes 4, had significant canopy area in the strata of 8m to16m and 2m to 8m. This structure class had a negative correlation between proportional canopy strata cover and canopy strata height.

Table 2. Key characteristics for the structure classes identified from patterns of canopy and openings across the Whiskeytown study area

Class	Dominant Height	Stratification	Density
1	Short	Single Story	Low
2	Short	Single Story	Moderate
3	Moderate	Single Story	High
4	Moderate	Bottom Loaded	Low
5	Moderate	Multi Story	Moderate
6	Tall	Multi Story	Moderate
7	Tall	Multi Story	High
8	Very Tall	Multi Story	High

Structure class 1 had the highest proportion of open area and the lowest dominant canopy height (Figure 10). This structure class is likely to be non-forested. Structure class 2 had the highest proportion of canopy area from 2 m to 8 m. Structure class 4 was most statistically similar to structure class 2, with the dissimilarity of having a more even distribution of canopy area between 2 m and 16 m. Structure classes 1, 2, and 4 were statistically dissimilar from all other structural classes.

Structure class 3 had much less open area than structure classes 1, 2, and 4, and had the highest proportion of canopy area from 8 m to 16 m. Structure class 5 was most statistically similar to structure class 3, with the noted exception of having a higher proportion of open space.

Structure class 6 had the highest proportion of canopy area from 16 m to 32 m and a significant proportion of canopy area from 32 m to 48 m. Structure class 7 was most statistically similar to structure class 6, but had a much higher proportion of canopy area from 32 m to 48 m. Structure class 8 was similar to structure classes 6 and 7, but had a much higher proportion of canopy area above 48 m.

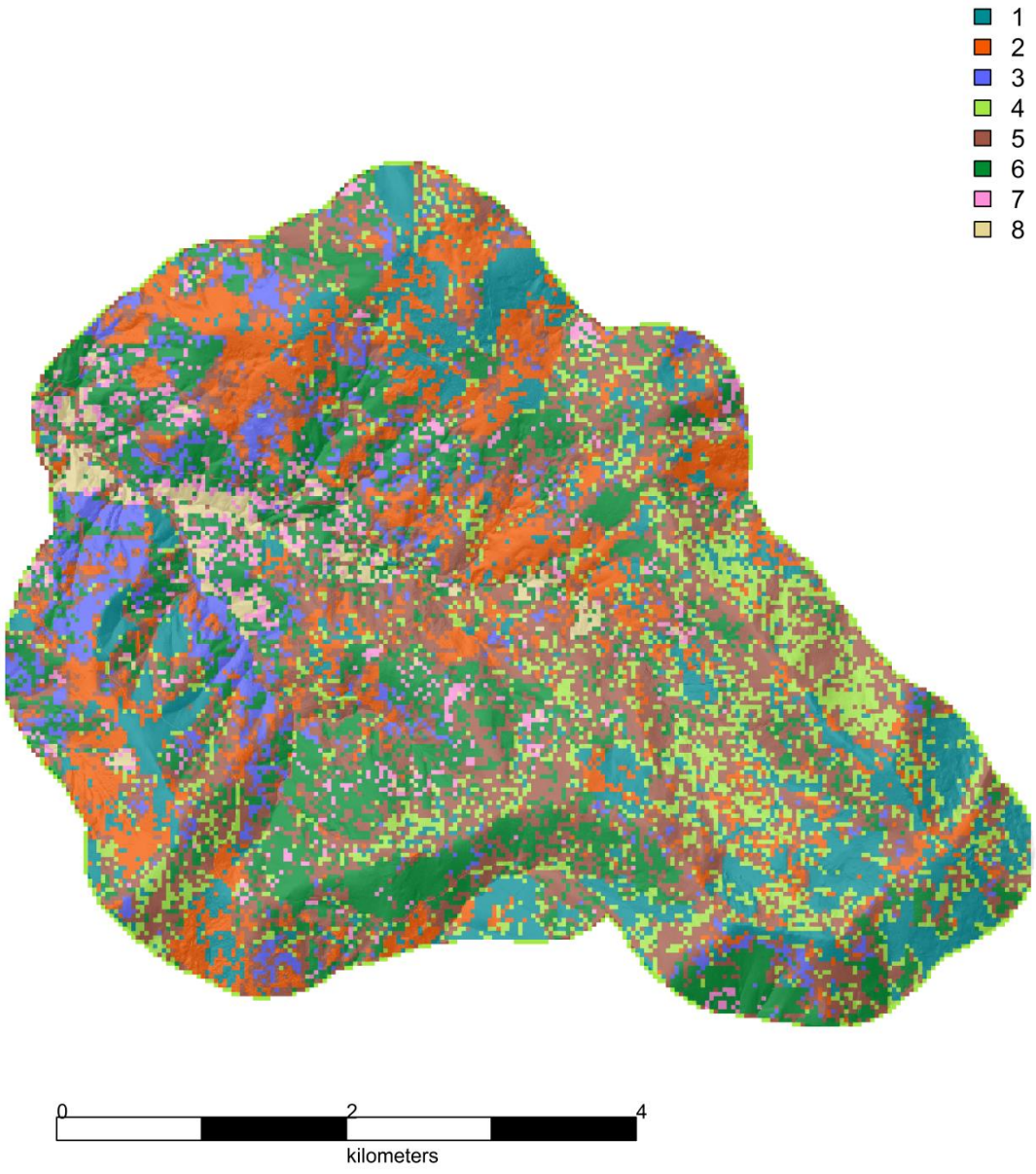


Figure 7. Map of structure classes identified for the Oregon Caves study area.

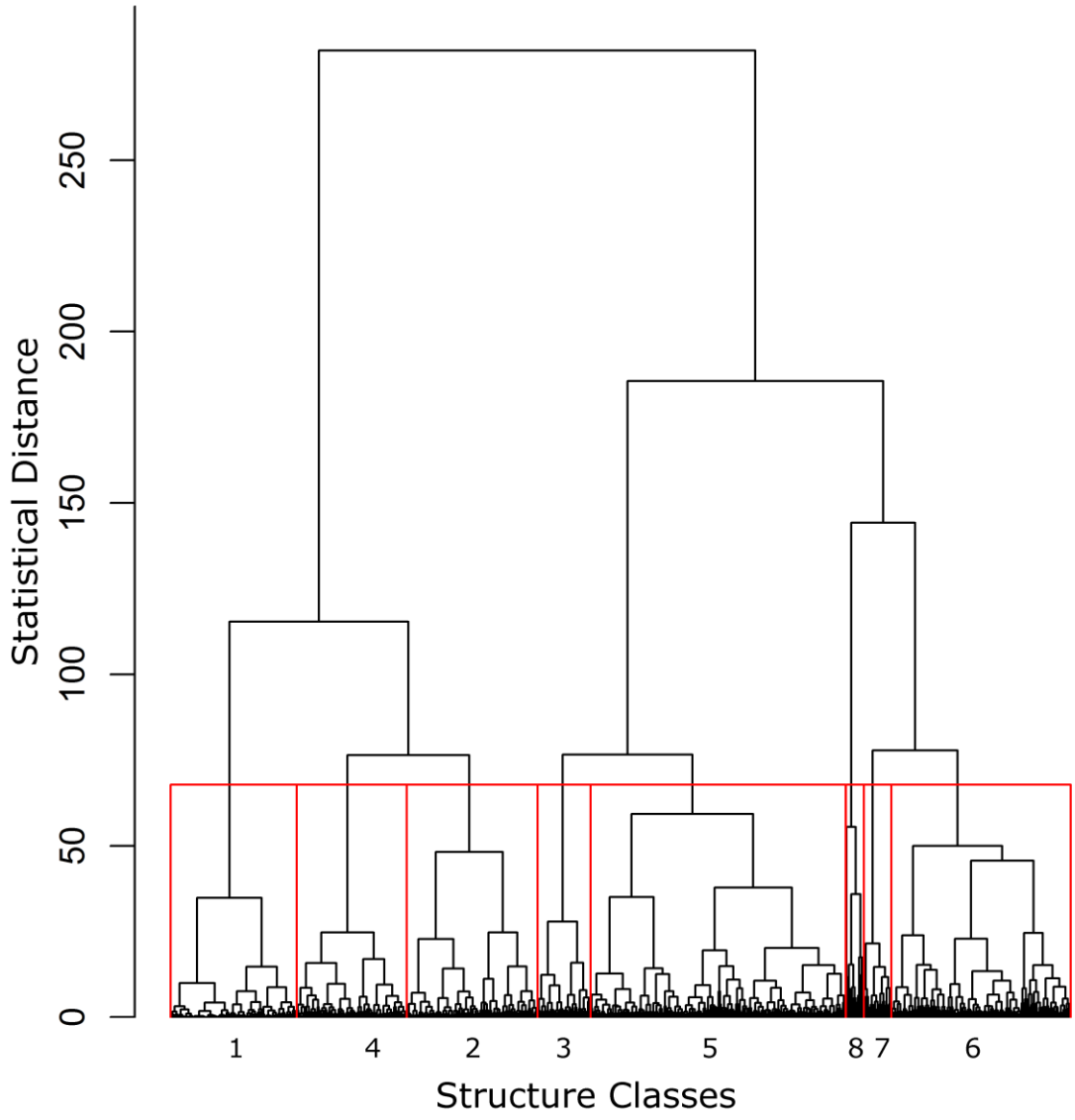


Figure 8. Dendogram showing statistical distance between structure classes identified in the Oregon Caves study area. Clustering was implemented using Ward's (1963) clustering criterion. Clustering is additive, where classes which join lower town on the tree are more statistically significant than classes that join higher on the tree.

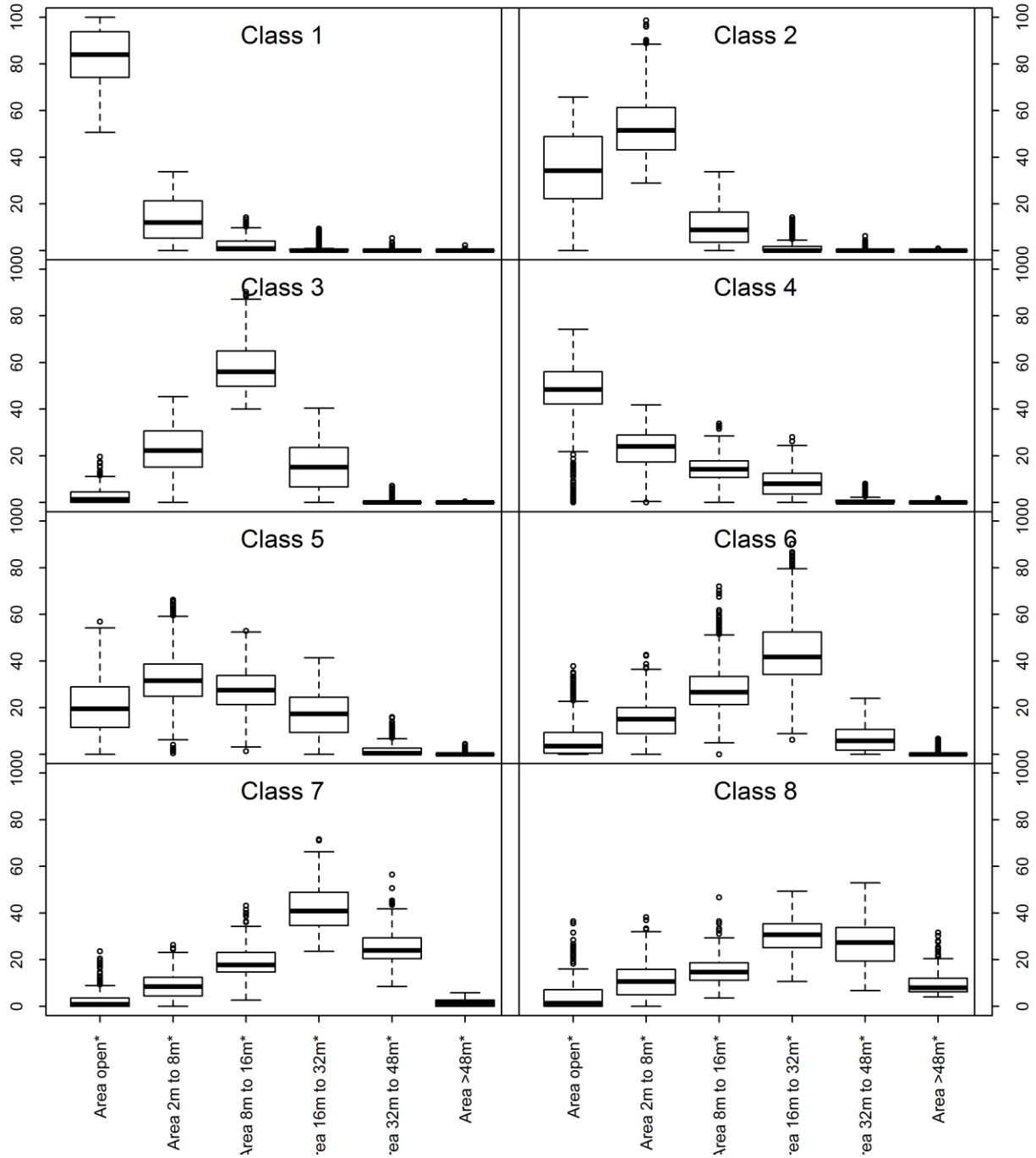


Figure 9. Characteristics of structure classes identified in in the Oregon Caves study area. Scale is either percentage, meters in height, or count of large tree clumps or large gaps. Boxplots show ranges of values for metrics used to define each class. Area open is the percentage of each 0.81 ha sample 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. 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.

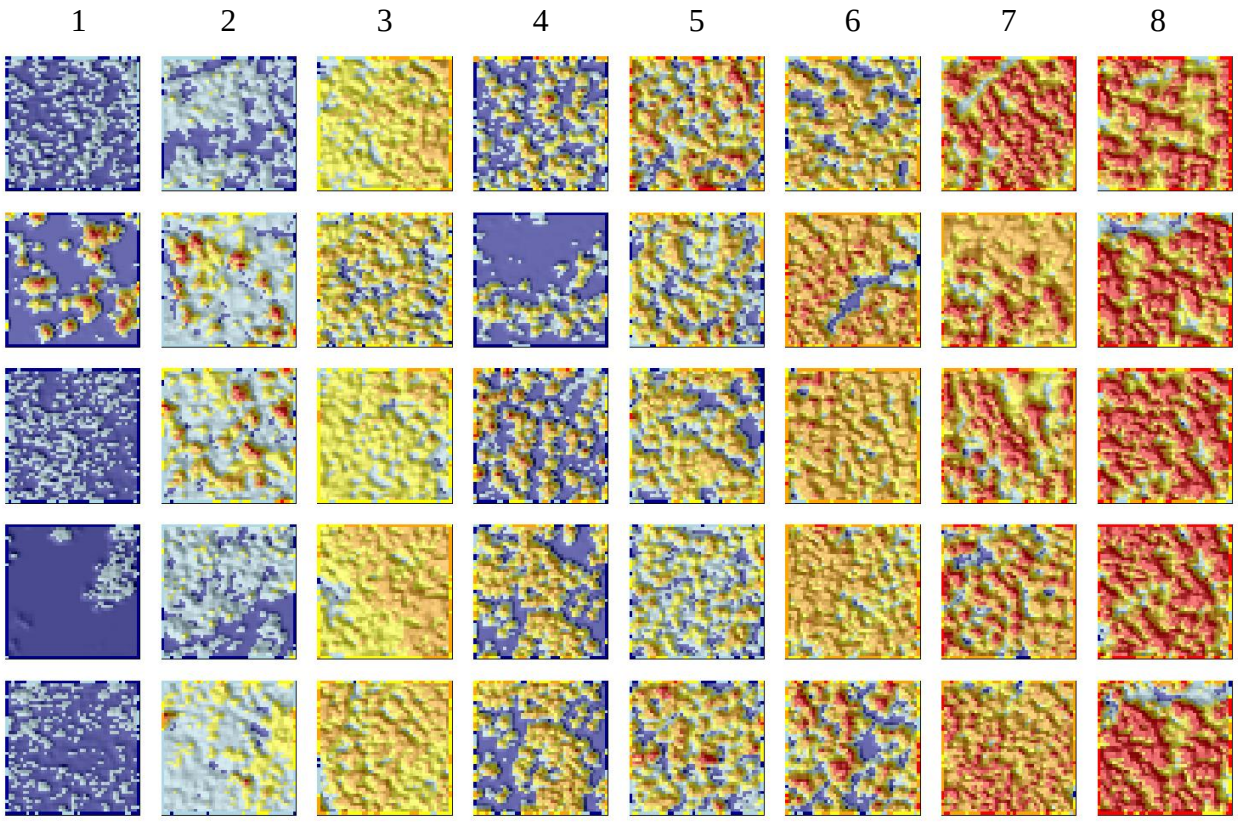


Figure 10. Examples of LiDAR structure metrics 30 m by 30 m areas. Each column contains 5 randomly selected pixels representing one of the 8 structure classes defined for the Oregon Caves study area. Color corresponds to canopy height, averaged across TAO extent. Warmer colors correspond to taller canopy areas while cooler colors correspond to shorter canopy areas.

3.3 Whiskeytown National Recreation Area

3.3.1 Forest Structure Classes

We identified nine statistically distinct forest structure classes using percent cover in two strata (1 m to 2 m and 2 m to 4 m), the area of openings (CSM <2 m) and area of TAOs in four height strata (2 m to 8 m, 8 m to 16 m, 16 m to 32 m, and >32 m) within 90 m 90 m areas across the study area (Figure 11). Eight of the structure classes were in statistically similar pairs of classes (Figure 12).

We found that structure classes could be sorted into three distinct groups with regard to canopy max height (Table 3). The tallest class (1 and 2) had the highest canopy max height values, and the majority of their canopy area in height strata >16 m. A second class of height strata (3, 4, 5, and 7) had moderate max height values and the majority of their canopy area in height strata between 8 m and 32 m. The third height strata (6, 8, and 9) had the lowest canopy max height values and the majority of their canopy area in height strata < 16 m. These breaks tracked well with canopy cover; structure classes with higher max height values tended to have higher levels of canopy cover.

Structure classes could be sorted into groups depending on the distribution of canopy area across height strata (Figure 13). Multi strata structure classes (4 and 7) had an even distribution of canopy area across three or more distinct height classes. These structure classes tended to have moderate cover. Single strata structure classes (1, 2, 3, and 8) had the majority of their canopy area in one or two adjacent height strata. These structure classes tended to have very high or very low canopy cover. Two strata structure classes (5 and 6), had significant canopy area in the strata of 8 m to16 m and 2 m to 8 m. These classes differed from one another in the proportion of canopy area in each height strata.

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

Class	Dominant Height	Stratification	Density
1	Tall	Single Story	High
2	Tall	Single Story	High
3	Moderate	Single Story	High
4	Moderate	Multi Story	High
5	Moderate	Bottom Loaded	Moderate
6	Short	Bottom Loaded	Moderate
7	Moderate	Multi Story	Low
8	Short	Single Story	Low
9	Short	Zero	None

Structure class 1 was characterized as having the highest overall canopy cover and proportion of canopy area in the range of 16 m to 32 m (Figure 14). This structure class was top loaded and single story, with very little canopy area below 16 m and very little cover below 4 m.

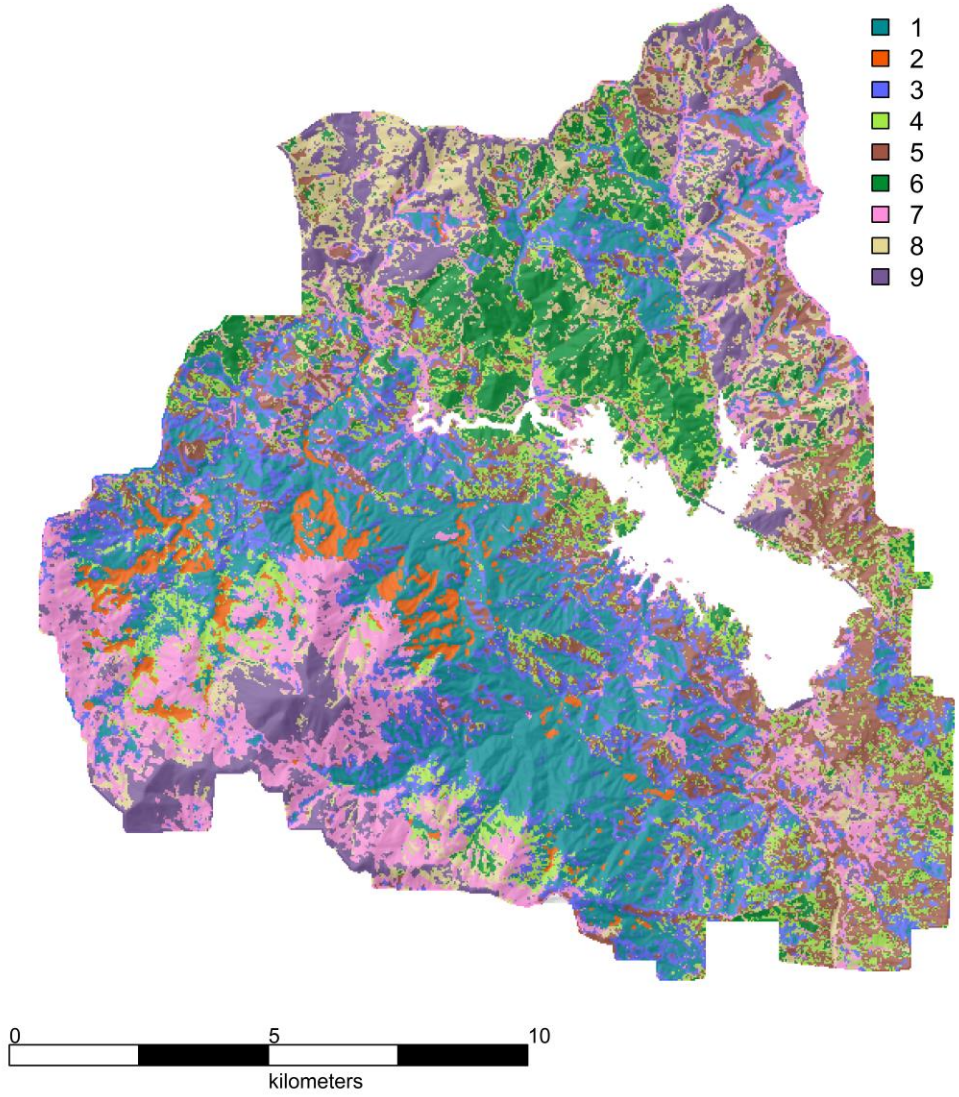


Figure 11. Map of structure classes identified for the Whiskeytown study area.

Structure class 2 was statistically similar to structure class 2, with the primary exception of having the highest proportion of canopy area above 48 m. Both structure classes 1 and 2 had very high proportions of canopy area above 32 m.

Structure class 3 had the second highest proportion of canopy area in the strata of 16 m to 32 meters, and a very high proportion of canopy area in the height strata from 8 m to 16 m. Structure class 3 was found to have a small proportion of canopy area between 32 m and 48 m. Structure class 7 was statistically similar to structure class 3, with the primary distinction of having a higher proportion of canopy gaps. Gaps in structure class 7 tended to be highly connected.

Structure class 4 was found to have the most even distribution of canopy area across strata from 2 m to 32 m. This structure class had higher levels of cover below 4 m. Structure class 5 was statistically similar to structure class 4, while having more cover from 8 m to 16 m (the highest of any class), less cover below 4 meters, and little to no cover above 32 m.

Structure class 6 was distinct in having the highest level of cover below 4 m, as well as the highest canopy area in the range of 2 m to 8 m. Structure class 6 was among the lowest canopy max height levels, and among the highest number of gaps. Structure class 6 was characterized by sinuous gaps that tended to snake and connect through the canopy.

Structure class 8 was the most open forested structure class observed. This structure class had very little canopy area in strata above 16 m, and was dominated by open space, with a high number of canopy clumps. Structure class 9 was most statistically similar to structure class 8. Structure class 9 had the lowest canopy area, often at or near 0%. This structure class is likely bare earth in most cases.

Several properties were observed with regards to the spatial distribution of structure classes across the study area. Structure classes with higher canopy cover tended to aggregate into larger patches. Structure class 1 had the largest continuous patch (1951 ha) and highest mean patch size (4.0 ha). Structure class 2, which represented only 3% of the landscape, was primarily interspersed with structure class 1 in large continuous patches. Together structure classes 1 and 2 represented 20% of the total landscape. These structure classes were generally found at elevations between 700 m and 900 m. Structure class 9, which describes primarily non-forested areas, had the second largest mean patch size at 3.2 ha. This structure class was generally found at high elevations >1000 m within the study area. Structure classes 4, and 5, and 6 tended to be interspersed as smaller patches. These structure classes were generally found at lower elevations below 700 m. Structure classes 6 and 7 had high standard deviations in patch size.

Examples of LiDAR point cloud data with accompanying FUSION metrics for each structure class are provided in figure 15.

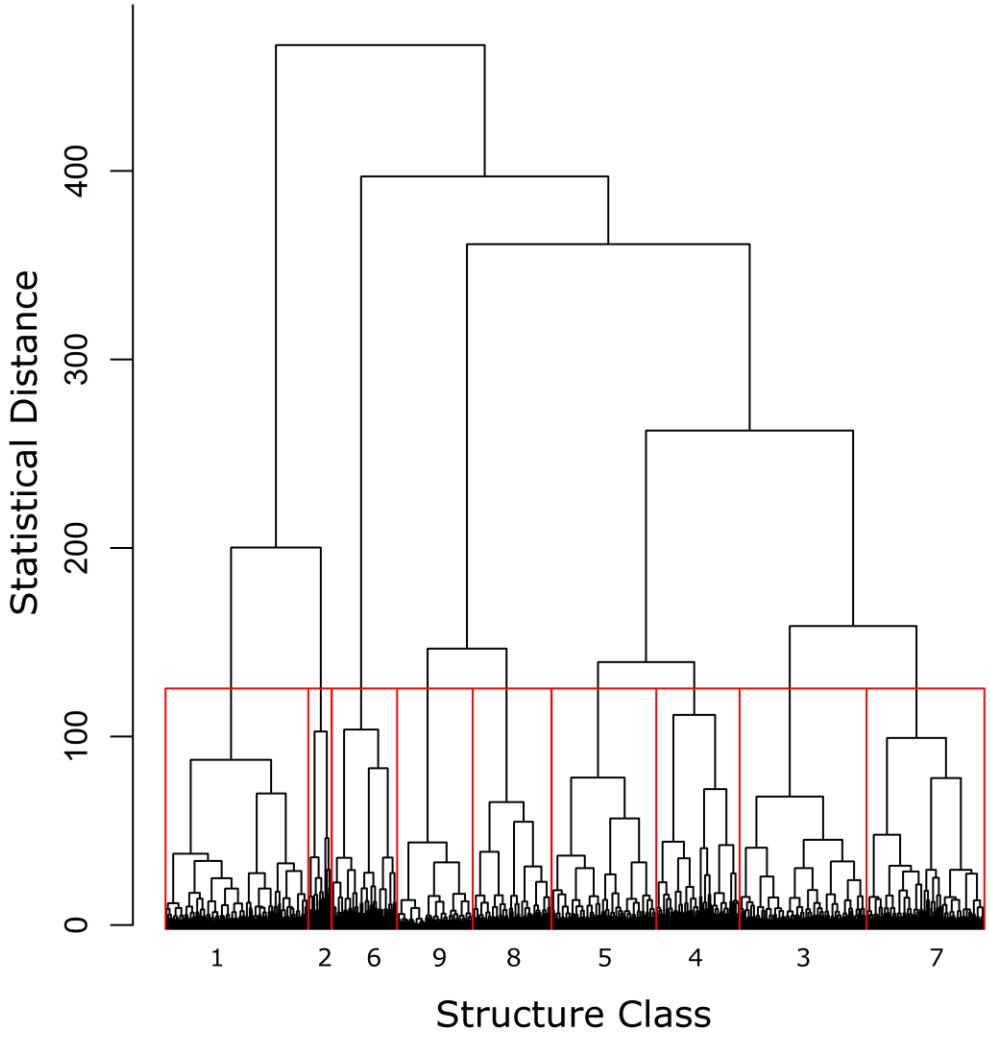


Figure 12. Dendogram showing statistical distance between structure classes identified in the Whiskeytown study area. Clustering was implemented using Ward's (1963) clustering criterion. Clustering is additive, where classes which join lower town on the tree are more statistically significant than classes that join higher on the tree.

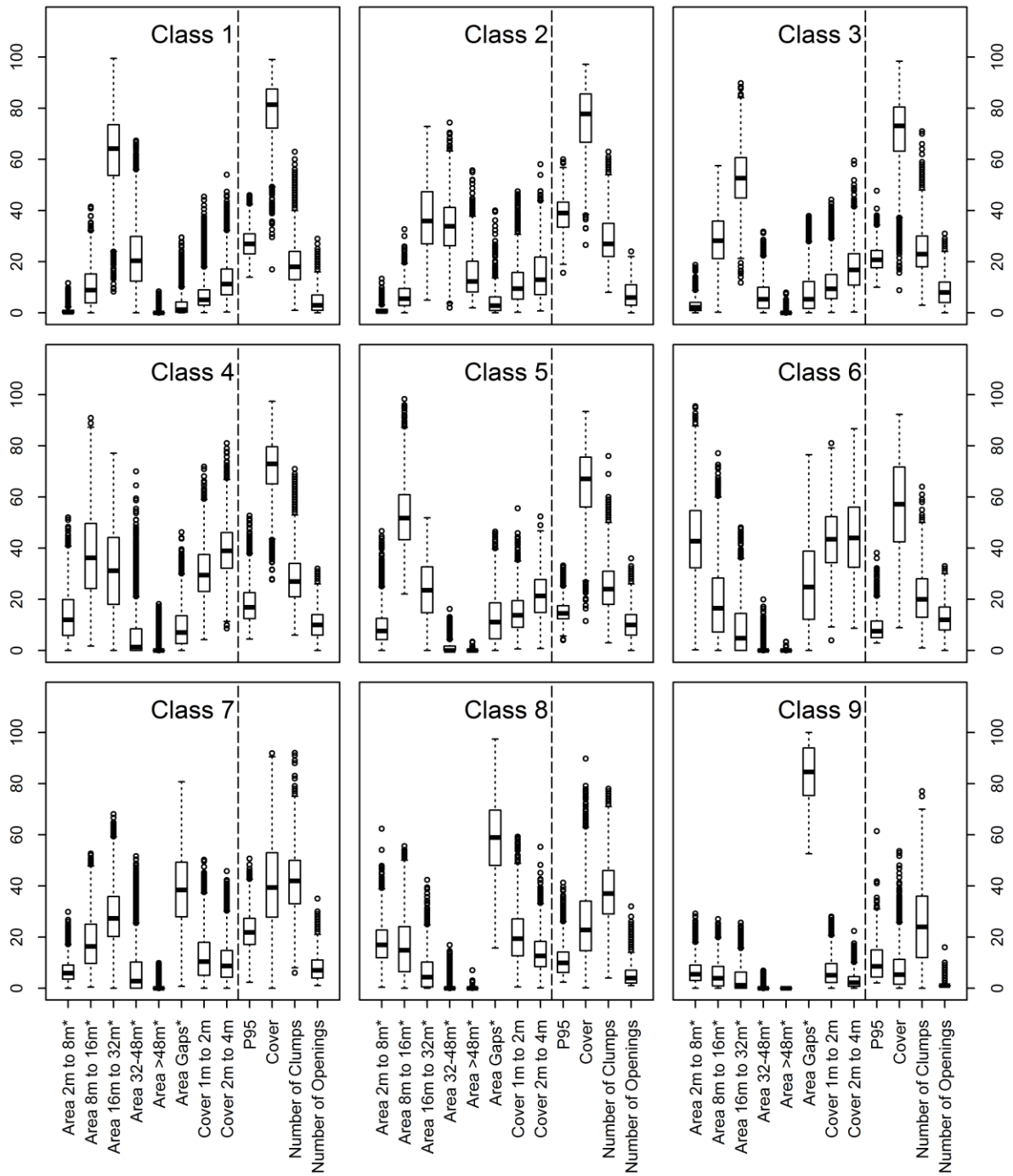


Figure 13. Characteristics of structure classes identified in in the Whiskeytown study area. Scale is either percentage, meters in height, or count of large tree clumps or large gaps. Boxplots show ranges of values for metrics used to define each class (left of dotted line in each panel) and metrics that provide additional information on each of the structure class but which were not used to define the structure classes (right of dotted line in each panel). Area Gaps 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.

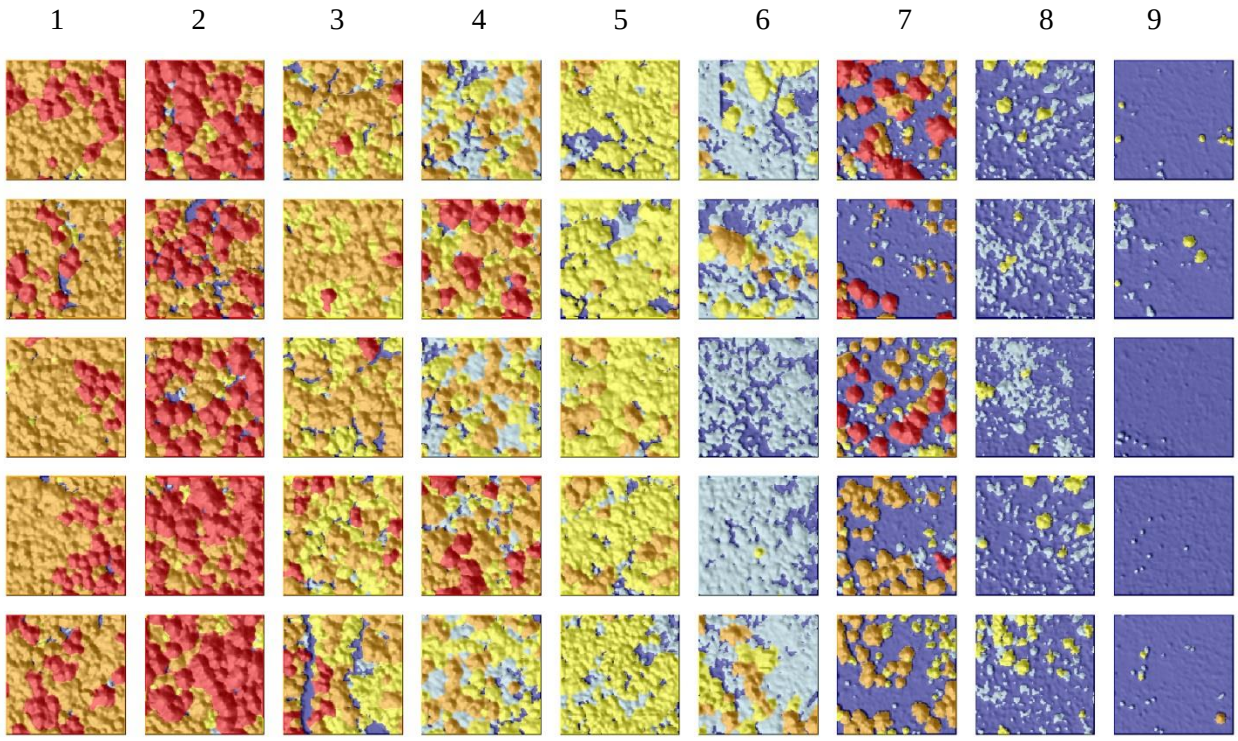
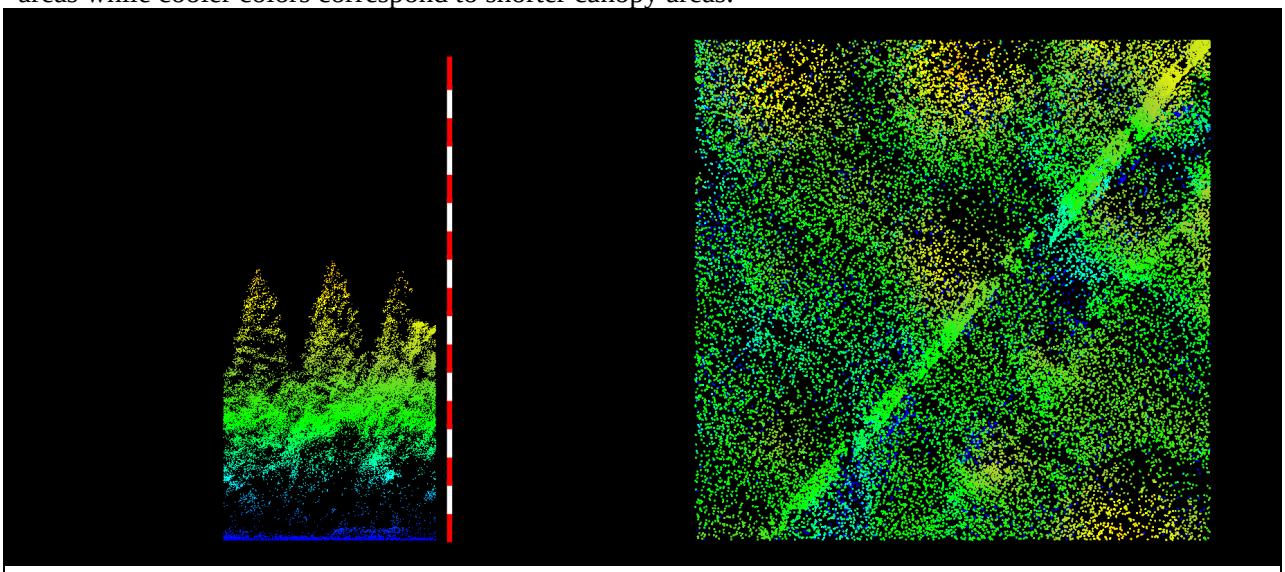
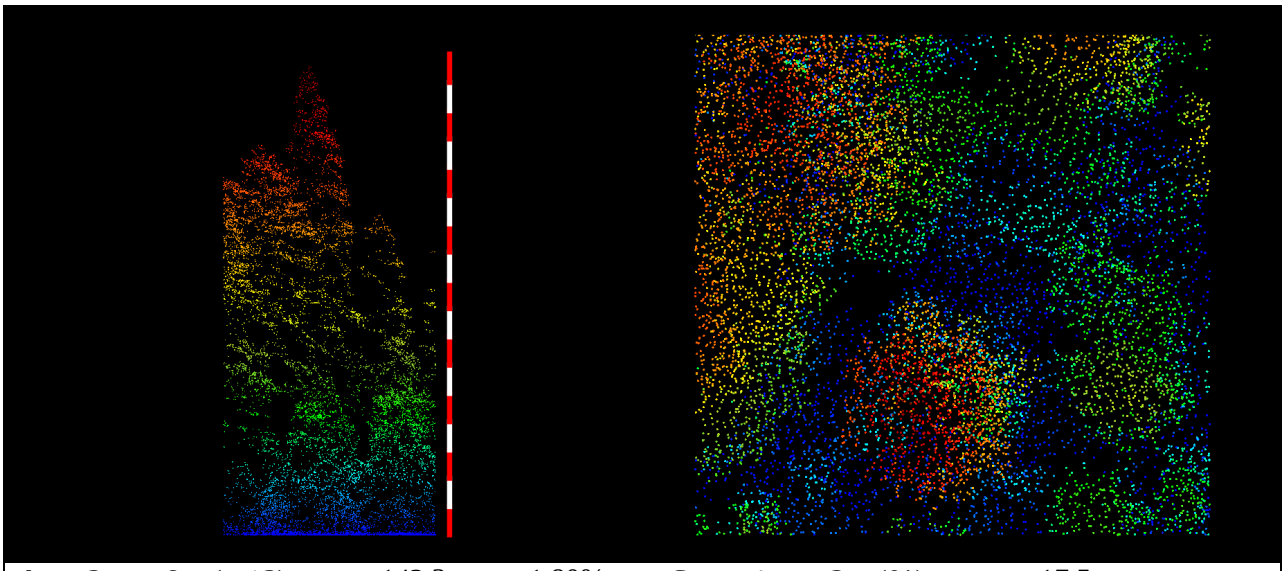


Figure 14. Examples of LiDAR structure metrics 30 m by30 m areas. Each column contains 5 randomly selected pixels representing one of the 9 structure classes defined for the Whiskeytown study area. Color corresponds to canopy height, averaged across TAO extent. Warmer colors correspond to taller canopy areas while cooler colors correspond to shorter canopy areas.



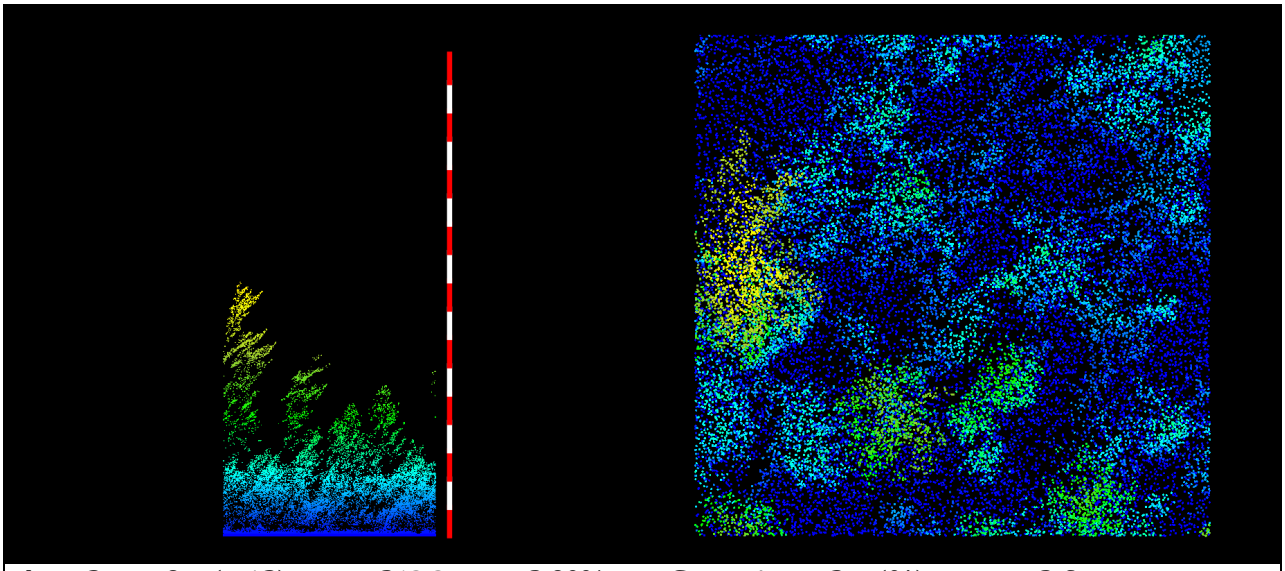
Area 2m to 8m (m^2)	0	0%	**Cover 1m to 2m (%)**	11.6
Area 8m to 16m (m^2)	2.8	0%	**Cover 2m to 4m (%)**	5.4
Area 16m to 32m (m^2)	5499.6	67.90%	**P95 (m)**	30
Area 32-48m (m^2)	2512.1	31%	**Cover (%)**	92.5
Area >48m (m^2)	0	0%	**Number of Clumps**	11
Area Gaps (m^2)	61.9	0.80%	**Number of Openings**	3

Figure 15-a. Example of LiDAR structure metrics of structure class 1 for 30 by 30 m area. Stripes on the range pole are 4 m in length.



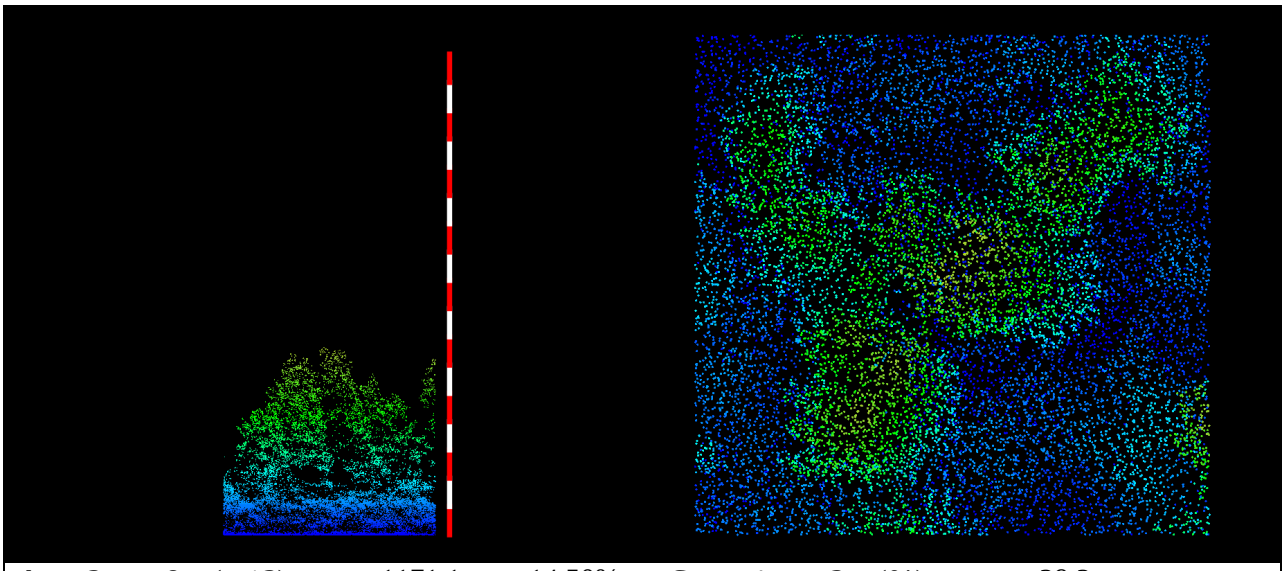
Area 2m to 8m (m^2)	142.3	1.80%	**Cover 1m to 2m (%)**	17.5
Area 8m to 16m (m^2)	374.1	4.60%	**Cover 2m to 4m (%)**	22.1
Area 16m to 32m (m^2)	1663.3	20.50%	**P95 (m)**	53.4
Area 32-48m (m^2)	1932.2	23.90%	**Cover (%)**	82.8
Area >48m (m^2)	3432.9	42.40%	**Number of Clumps**	59
Area Gaps (m^2)	421.9	5.20%	**Number of Openings**	13

Figure 15-b. Example of LiDAR structure metrics of structure class 2 for 30 m by 30 m area. Stripes on the range pole are 4 m in length.



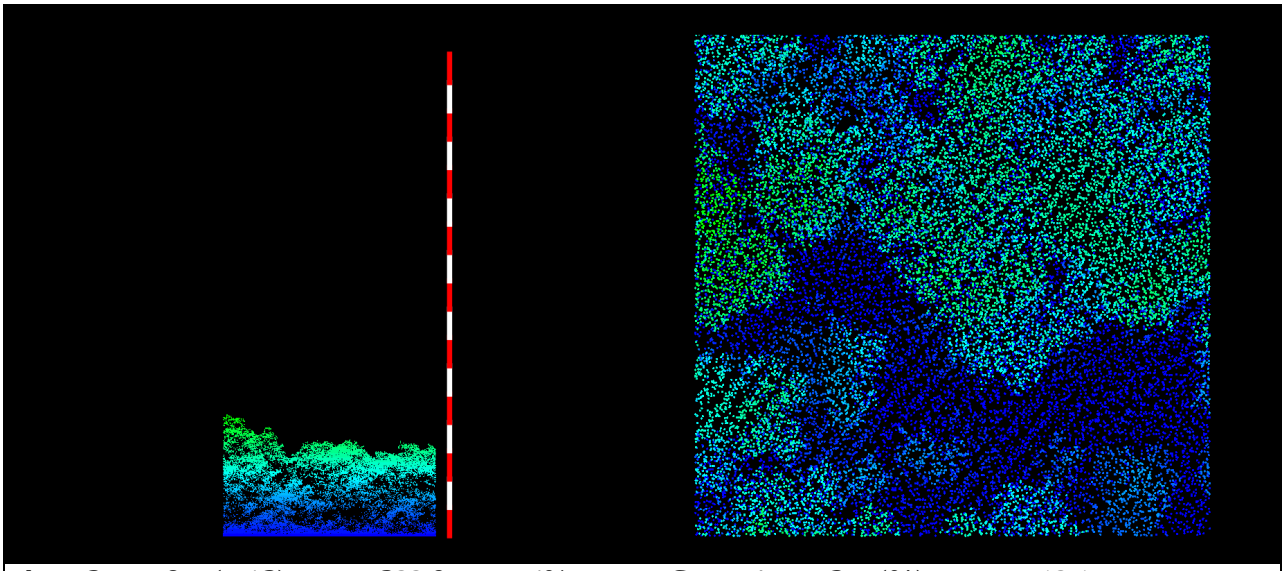
Area 2m to 8m (m^2)	312.2	3.90%	**Cover 1m to 2m (%)**	3.6
Area 8m to 16m (m^2)	3378.4	41.70%	**Cover 2m to 4m (%)**	11.4
Area 16m to 32m (m^2)	3584.8	44.30%	**P95 (m)**	27.3
Area 32-48m (m^2)	241.9	3%	**Cover (%)**	47.2
Area >48m (m^2)	0	0%	**Number of Clumps**	38
Area Gaps (m^2)	471.9	5.80%	**Number of Openings**	16

Figure 15-c. Example of LiDAR structure metrics of structure class 3 for 30 by 30 m area. Stripes on the range pole are 4 m in length.



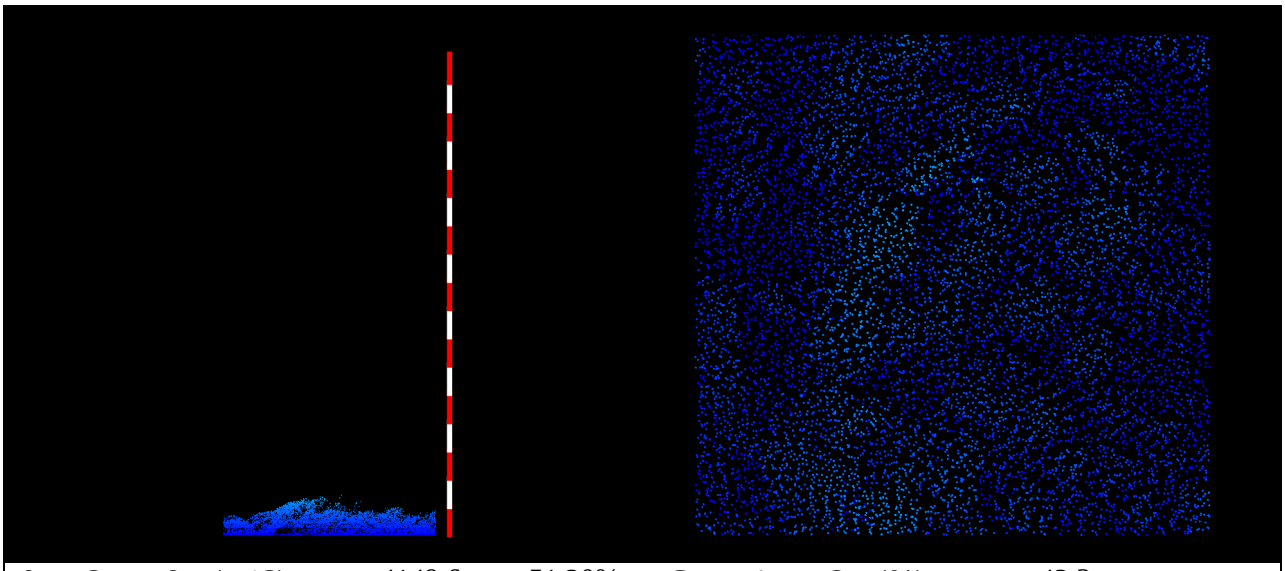
Area 2m to 8m (m^2)	1171.1	14.50%	**Cover 1m to 2m (%)**	29.2
Area 8m to 16m (m^2)	2957.1	36.50%	**Cover 2m to 4m (%)**	40.8
Area 16m to 32m (m^2)	3443.1	42.50%	**P95 (m)**	21.5
Area 32-48m (m^2)	82.1	1%	**Cover (%)**	77.4
Area >48m (m^2)	0	0%	**Number of Clumps**	42
Area Gaps (m^2)	357.8	4.40%	**Number of Openings**	7

Figure 15-d. Example of LiDAR structure metrics of structure class 4 for 30 m by 30 m area. Stripes on the range pole are 4 m in length.



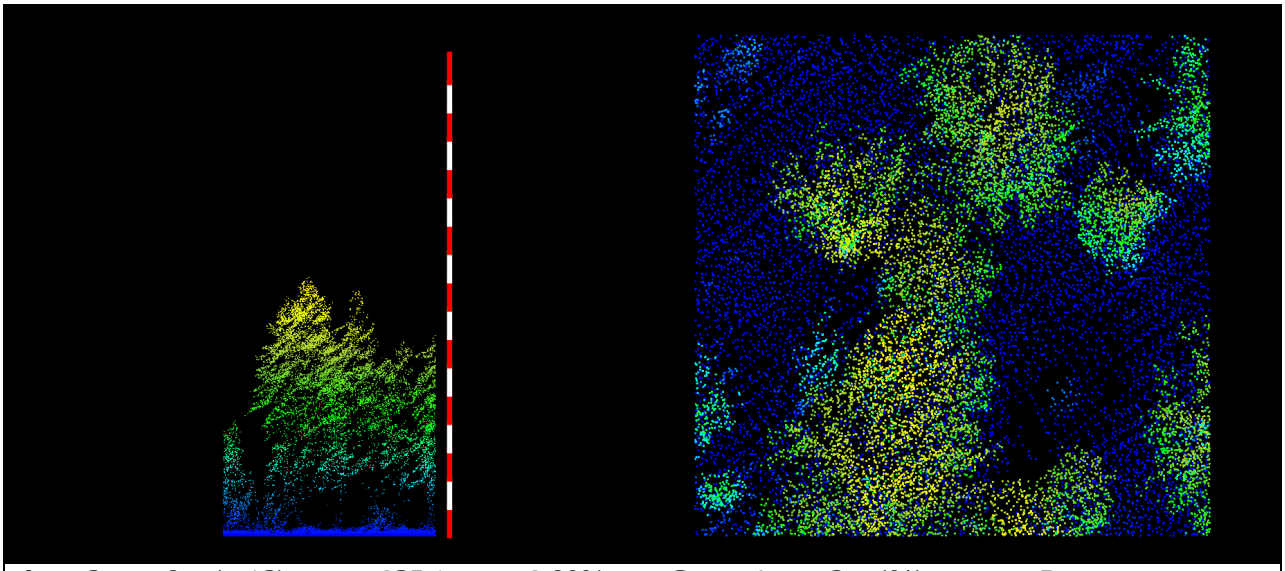
Area 2m to 8m (m^2)	322.9	4%	**Cover 1m to 2m (%)**	12.1
Area 8m to 16m (m^2)	5614.9	69.30%	**Cover 2m to 4m (%)**	11.5
Area 16m to 32m (m^2)	1621.7	20%	**P95 (m)**	13.2
Area 32-48m (m^2)	83.8	1%	**Cover (%)**	56.7
Area >48m (m^2)	0	0%	**Number of Clumps**	22
Area Gaps (m^2)	401.6	5%	**Number of Openings**	7

Figure 15-e. Example of LiDAR structure metrics of structure class 5 for 30 m by 30 m area. Stripes on the range pole are 4 m in length.



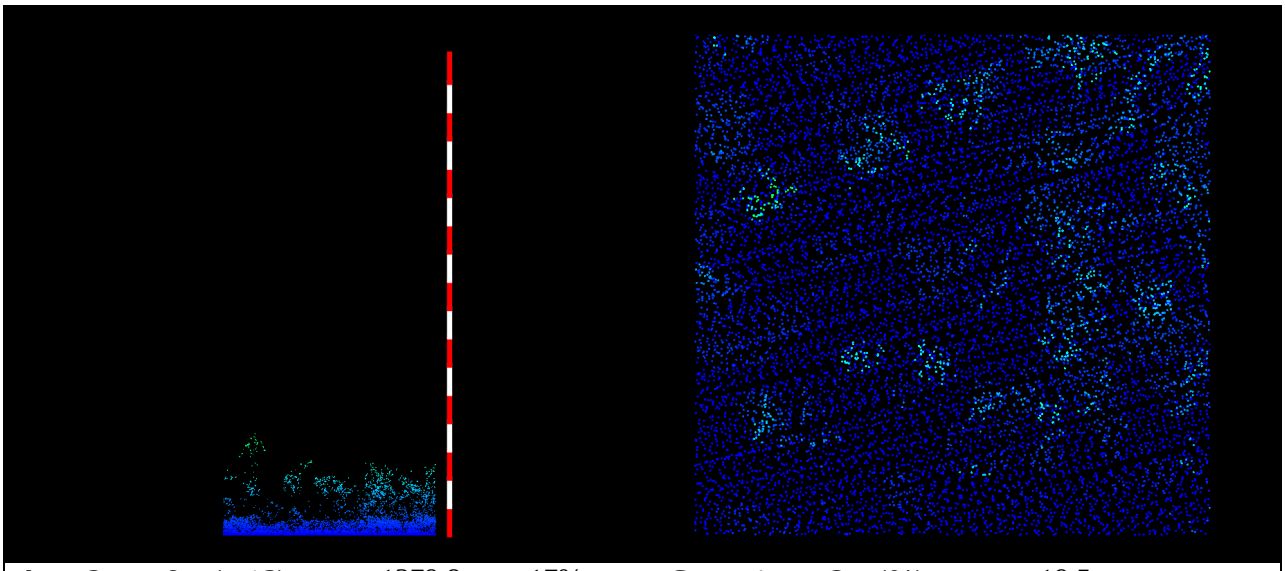
Area 2m to 8m (m^2)	4149.6	51.20%	**Cover 1m to 2m (%)**	42.3
Area 8m to 16m (m^2)	1307.8	16.10%	**Cover 2m to 4m (%)**	30.8
Area 16m to 32m (m^2)	0	0%	**P95 (m)**	4.4
Area 32-48m (m^2)	0	0%	**Cover (%)**	33.4
Area >48m (m^2)	0	0%	**Number of Clumps**	30
Area Gaps (m^2)	2503.7	30.90%	**Number of Openings**	14

Figure 15-f. Example of LiDAR structure metrics of structure class 6 for 30 by 30 m area. Stripes on the range pole are 4 m in length.



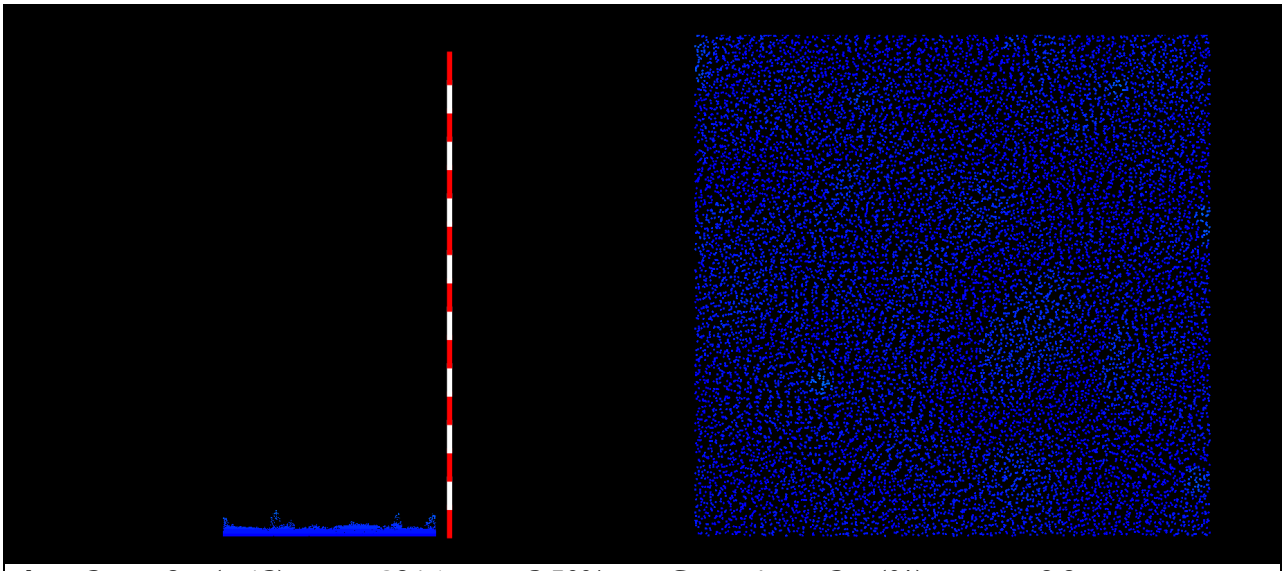
Area 2m to 8m (m^2)	235.1	2.90%	**Cover 1m to 2m (%)**	5
Area 8m to 16m (m^2)	599.1	7.40%	**Cover 2m to 4m (%)**	3.8
Area 16m to 32m (m^2)	2720.2	33.60%	**P95 (m)**	31.6
Area 32-48m (m^2)	1661.1	20.50%	**Cover (%)**	56.3
Area >48m (m^2)	0	0%	**Number of Clumps**	64
Area Gaps (m^2)	2850.8	35.20%	**Number of Openings**	7

Figure 15-g. Example of LiDAR structure metrics of structure class 7 for 30 m by 30 m area. Stripes on the range pole are 4 m in length.



Area 2m to 8m (m^2)	1379.8	17%	**Cover 1m to 2m (%)**	19.5
Area 8m to 16m (m^2)	2013.8	24.90%	**Cover 2m to 4m (%)**	7.2
Area 16m to 32m (m^2)	181.1	2.20%	**P95 (m)**	8.9
Area 32-48m (m^2)	0	0%	**Cover (%)**	14.7
Area >48m (m^2)	0	0%	**Number of Clumps**	100
Area Gaps (m^2)	4459.5	55.10%	**Number of Openings**	7

Figure 15-h. Example of LiDAR structure metrics of structure class 8 for 30 m by 30 m area. Stripes on the range pole are 4 m in length.



Area 2m to 8m (m^2)	284.1	3.50%	**Cover 1m to 2m (%)**	8.8
Area 8m to 16m (m^2)	151.3	1.90%	**Cover 2m to 4m (%)**	0.5
Area 16m to 32m (m^2)	24.2	0.30%	**P95 (m)**	3.2
Area 32-48m (m^2)	0	0%	**Cover (%)**	0.5
Area >48m (m^2)	0	0%	**Number of Clumps**	40
Area Gaps (m^2)	7640.4	94.30%	**Number of Openings**	1

Figure 15-i. Example of LiDAR structure metrics of structure class 9 for 30 m by 30 m area. Stripes on the range pole are 4 m in length.

3.3.2 Ground Fuels

3.3.2.1 Structure Class relationship to LANDFIRE

We performed a cursory exploration of the relationship between our derived structure classes and LANDFIRE 40 fuel models. Preliminary analysis included the creation of histograms describing the distribution of LANDFIRE classes with respect to each structure class, as well as those describing the distribution of structure classes with respect to each LANDIFRE class (Figure 16). In addition, correlations were calculated for each metric. This preliminary analysis did not yield significant results, and no further exploration was pursued regarding this relationship.

3.3.2.2 Ground Fuel Modeling

Models generated using different radii were found to be similar, with 15m radius plots preforming slightly better. This radius was selected for final model creation, having the added benefit of more closely describing field plot areas.

Random forest regressions of continuous values for individual surface fuel values had R^2 values ranging between 7.7% and 56.2% (Table 4). Models describing smaller diameter fuels (i.e. litter, one hour and 10 hour) and rotten thousand hour fuel performed poorly ($R^2 < 34.0$). Models describing larger diameter fuels (i.e. hundred hour and thousand hour sound) and duff preformed significantly better ($R^2 > 53.3$). Models summarizing fuel conditions (i.e. total fine wood, total wood, total organic layer depth) preformed moderately ($R^2 > 46.3$).

Random forest classifications of fuel loading above or below median values for individual surface fuel values had out of bag error rates ranging between 32.2% and 11.9%. Classification models generally preformed best when describing organic layer variables ($25.42 > \%\ error > 11.85$) and worse when describing wood fuels ($32.2 > \%\ error > 16.95$). Kappa scores for classification models were found to be negatively correlated with out of bag error rates, ranging between 0.36 and 0.76. Cohen's Kappa (K) is a measure of agreement for categorical terms. It quantifies the level of agreement between datasets, while taking into account the probability of agreement by random chance. A value of 1 indicates total agreement, while values below 0.4 indicate fair to poor agreement.

Regression model response values tended to have smaller distributions below the median values, with several outliers being much higher than the median value for each response metric (Figure 17). The relationship between observed and predicted fuel loading values had higher dispersion at higher values in most models. These findings indicate that regression and classification model values are likely to be more accurate or descriptive when below the median value.

Regression models were found to rely heavily on FUSION and FRAGSTATS metrics, describing the vertical and horizontal distribution of the canopy. Twelve and eight models used at least one FUSION and FRAGSTATS metrics respectively, while only three models used a climate or topographic metric. The most commonly used metric was the 60^{th} percentile elevation

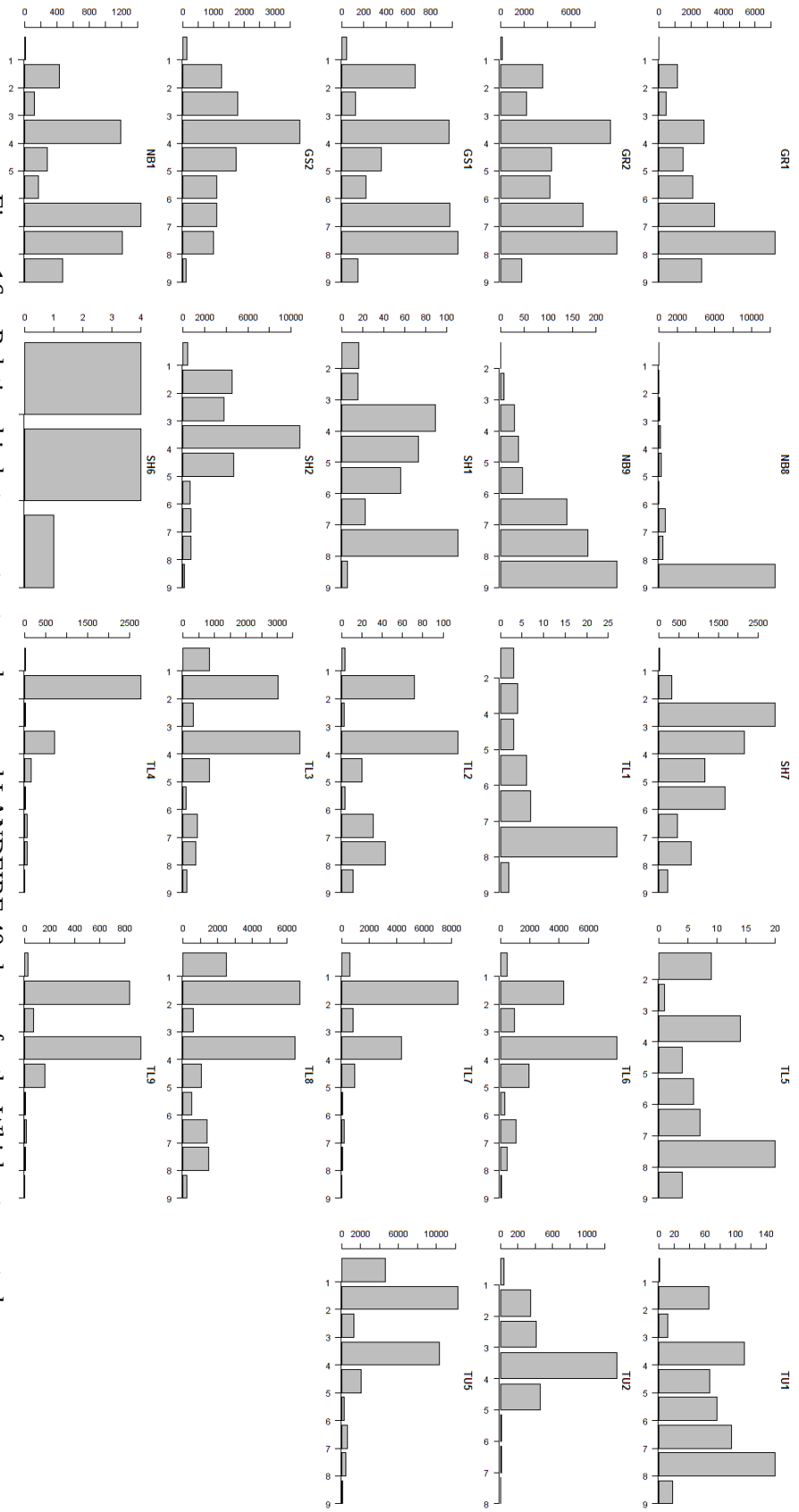


Figure 16-a. Relationship between structure classes and LANDFIRE 40 classes for the Whiskeytown study area.

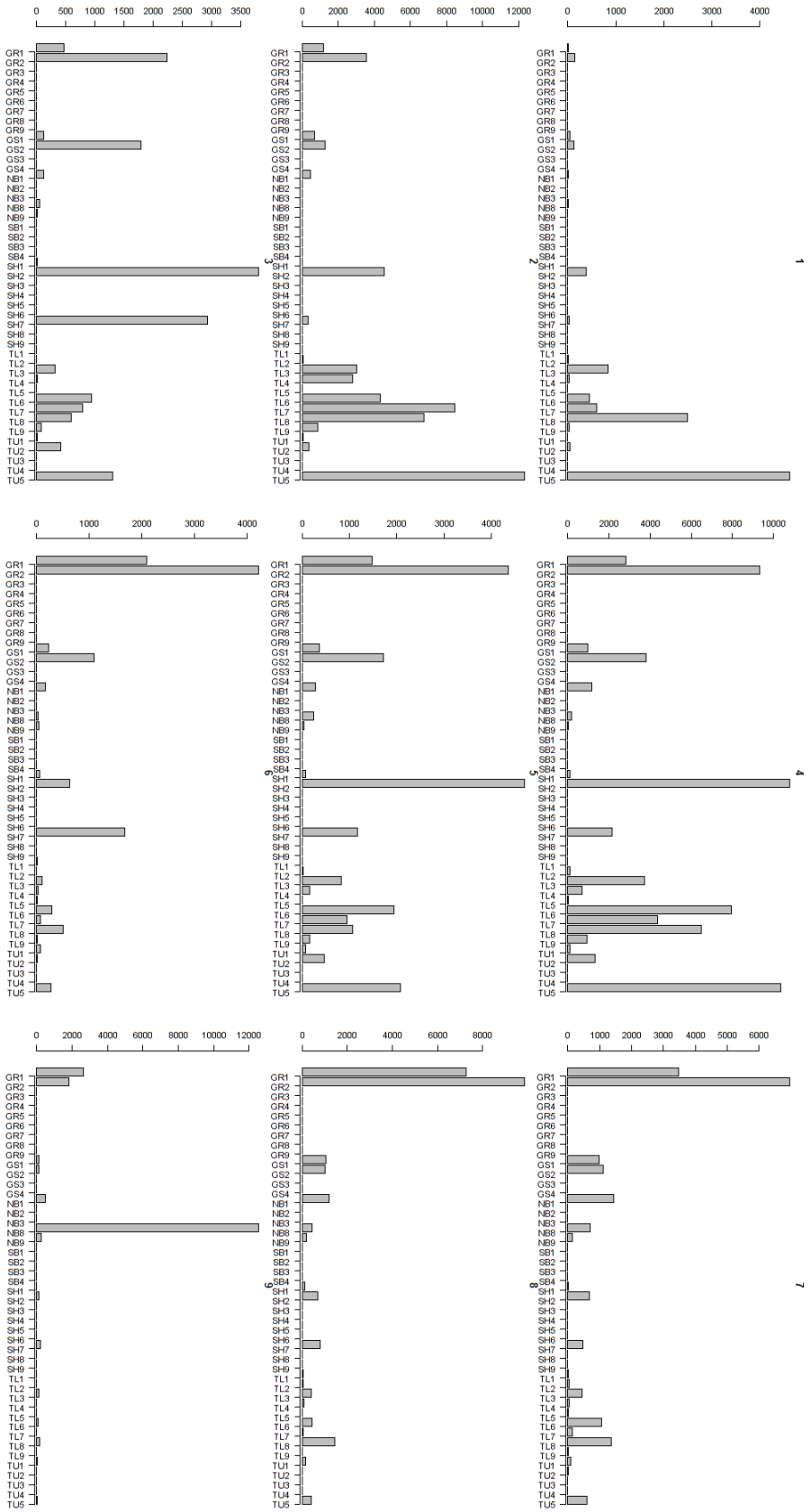


Figure 16-b. Relationship between structure classes and LANDFIRE 40 classes for the Whiskeytown study area

of vertical point distribution, followed by the proportion of land area covered by the 32 m to 48 m canopy strata.

Classification models were found to use a more even distribution of metrics across data sources. Ten models used at least one FUSION derived metric, while FRAGSTATS, topographic, and climate metrics were used at least once in six or seven models each. None of the models relied on a single data source during metric selection, and all metric sources were used in tandem in at least one model.

The mapped values for the climate datasets reflect the underlying approximate 800 m by 800 m resolution of the data. As a result, several maps of predicted fuel models that used climate metrics as predictors were found to have a blocky appearance reflecting the grain of the PRISM datasets (e.g. the lower left region of Figure 17-b-e). This influence had a more pronounced effect on classification models, whereas fine scale patterns were still detectable in affected regression models.

The sample size of training plots used in model creation was smaller than a past similar study (Kane et al. 2014). Additionally, plots were not well distributed across the study area. Plots existed on 7 of the 9 identified structure classes, with 75% of plots existing on only 2 structure classes. The range of values for predictor metrics extended outside the range of values observed in training plots (Figure 18). In several cases models were predicting values and classes outside the range of values used in model creation. These limitations had pronounced effects in the values predicted by several classification models. While predicting above or below median values should produce 1:1 distributions, in several cases classification models were dominated by high values. If used, these classification models should be combined with additional reliable datasets and interpreted carefully.

Table 4. Summary of fuel loading models accuracies. Depth values are reported in cm. All other values are reported in Tons/Acre

		Regression	Classification	
Metric	Median Value	R^2	Out of bag error	Kappa
One Hour Fuels	0.11	25.26	18.64	0.627
Ten Hour Fuels	1.02	34	25.42	0.492
Hundred Hour Fuels	1.85	56.18	25.42	0.491
All Fine Wood	3.69	49.38	32.2	0.357
Sound Thousand Hour Fuels	3.41	58.38	20.34	0.593
Rotten Thousand Hour Fuels	0.29	7.66	25.42	0.491
Total Wood	9.72	52.39	16.95	0.661
Duff	3.53	53.31	11.86	0.762
Litter	4.57	34.69	25.42	0.491
Total All Fuel	23.00	46.38	18.64	0.627
Duff Depth	0.35	54.2	11.86	0.762
Litter Depth	0.92	35.58	23.73	0.524
Total Depth	1.79	50.11	18.64	0.626

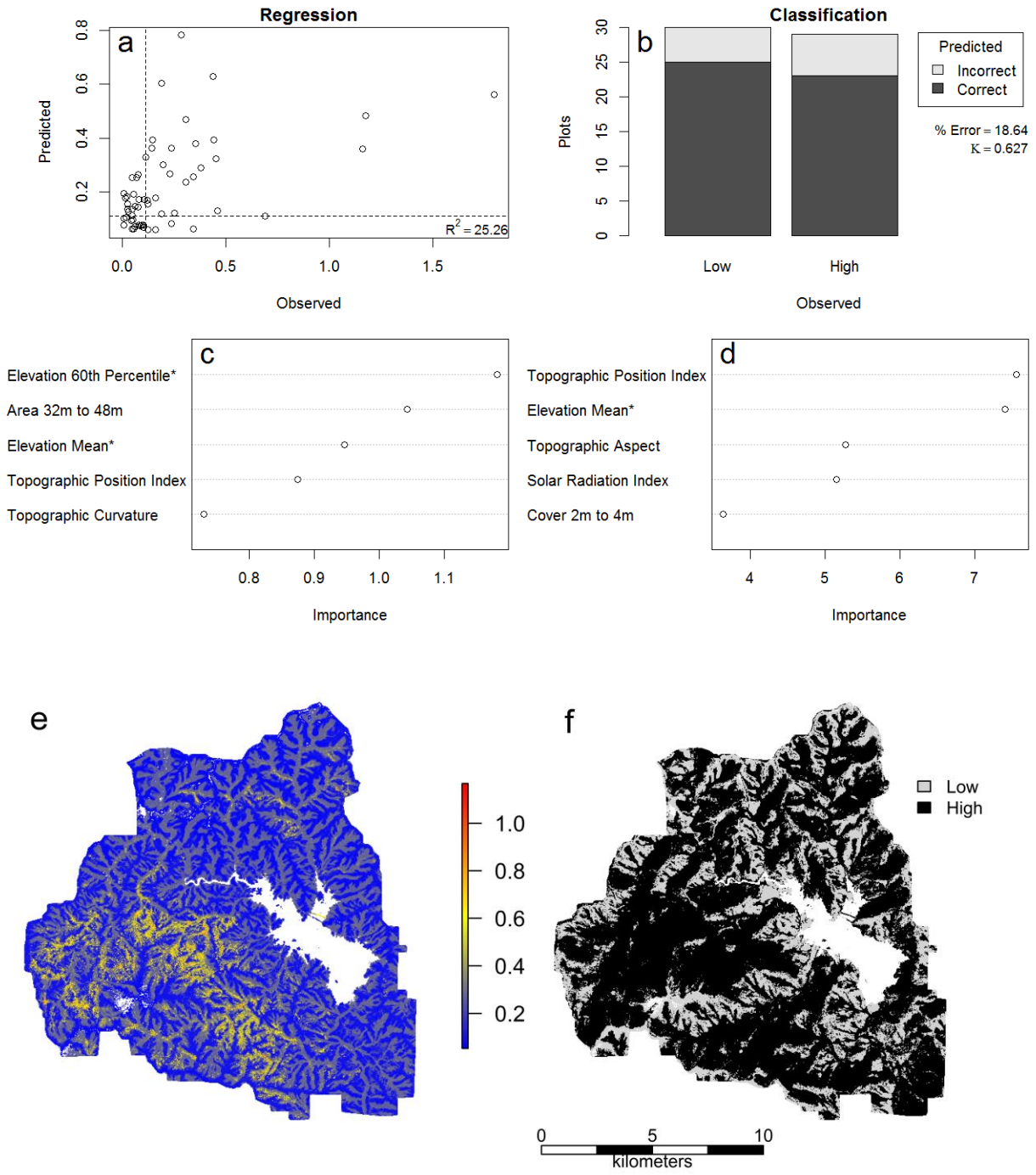


Figure 17-a. Models predicting one hour fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for one hour fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

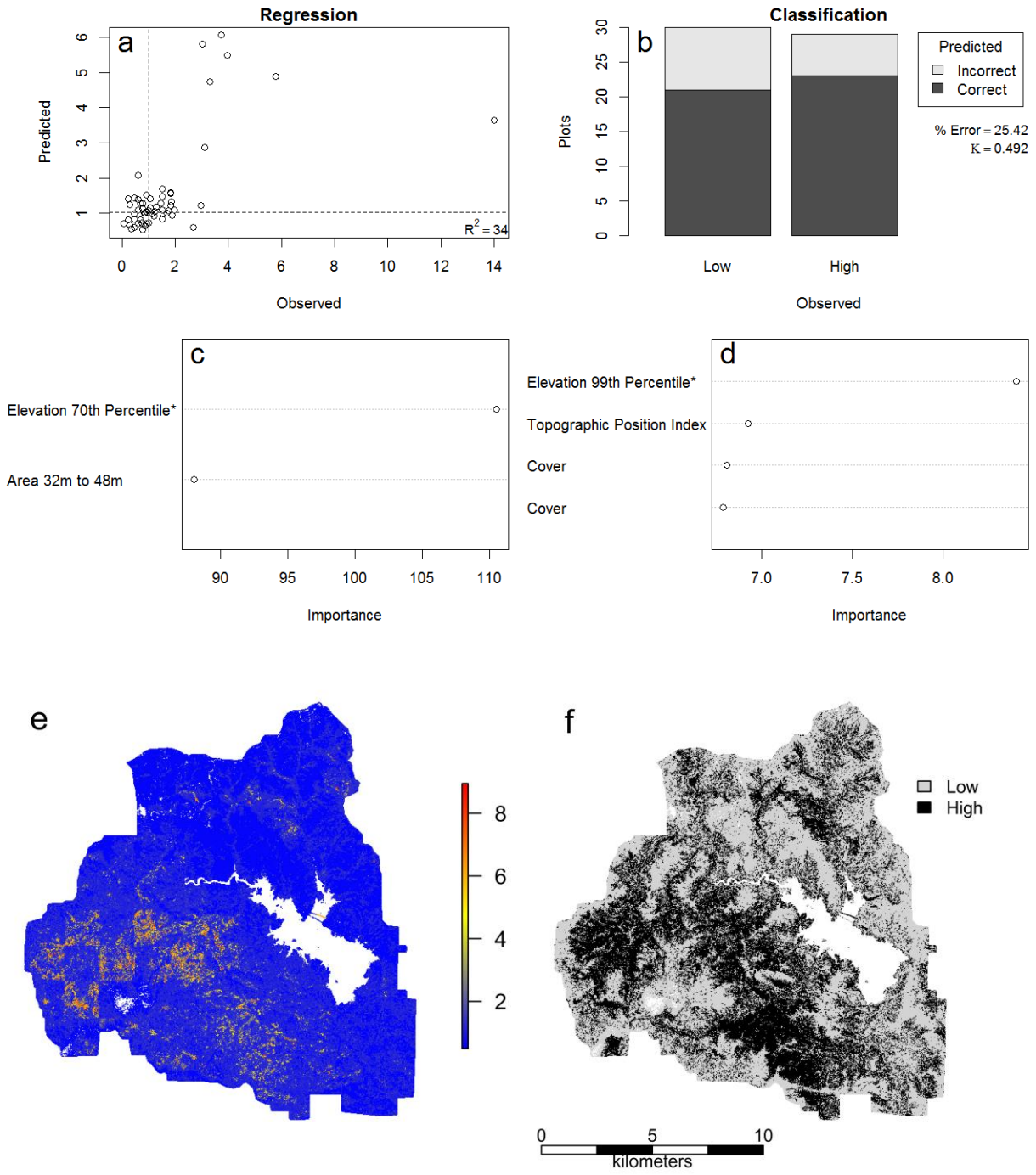


Figure 17-b. Models predicting ten hour fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for ten hour fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

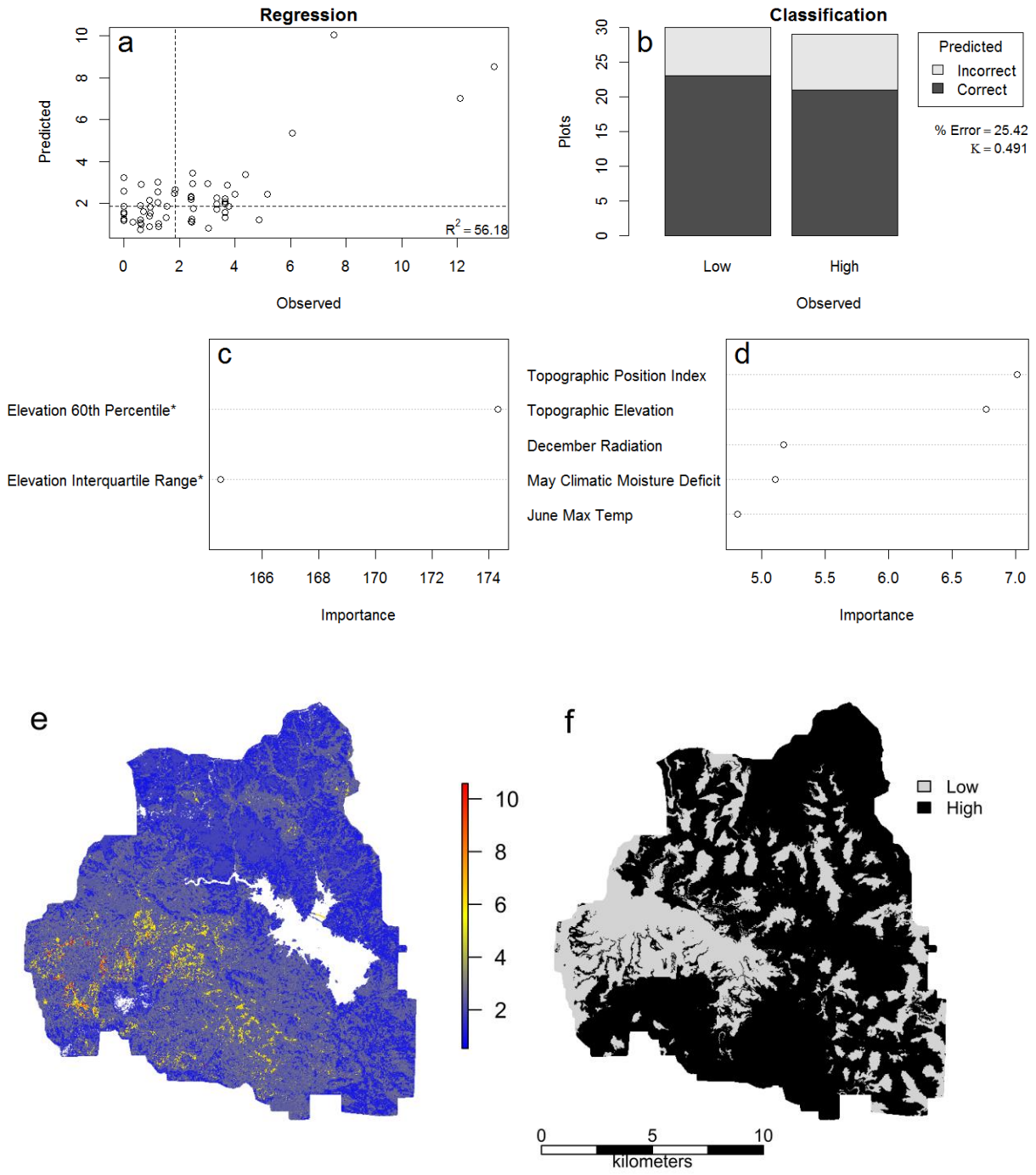


Figure 17-c. Models predicting hundred hour fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for hundred hour fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

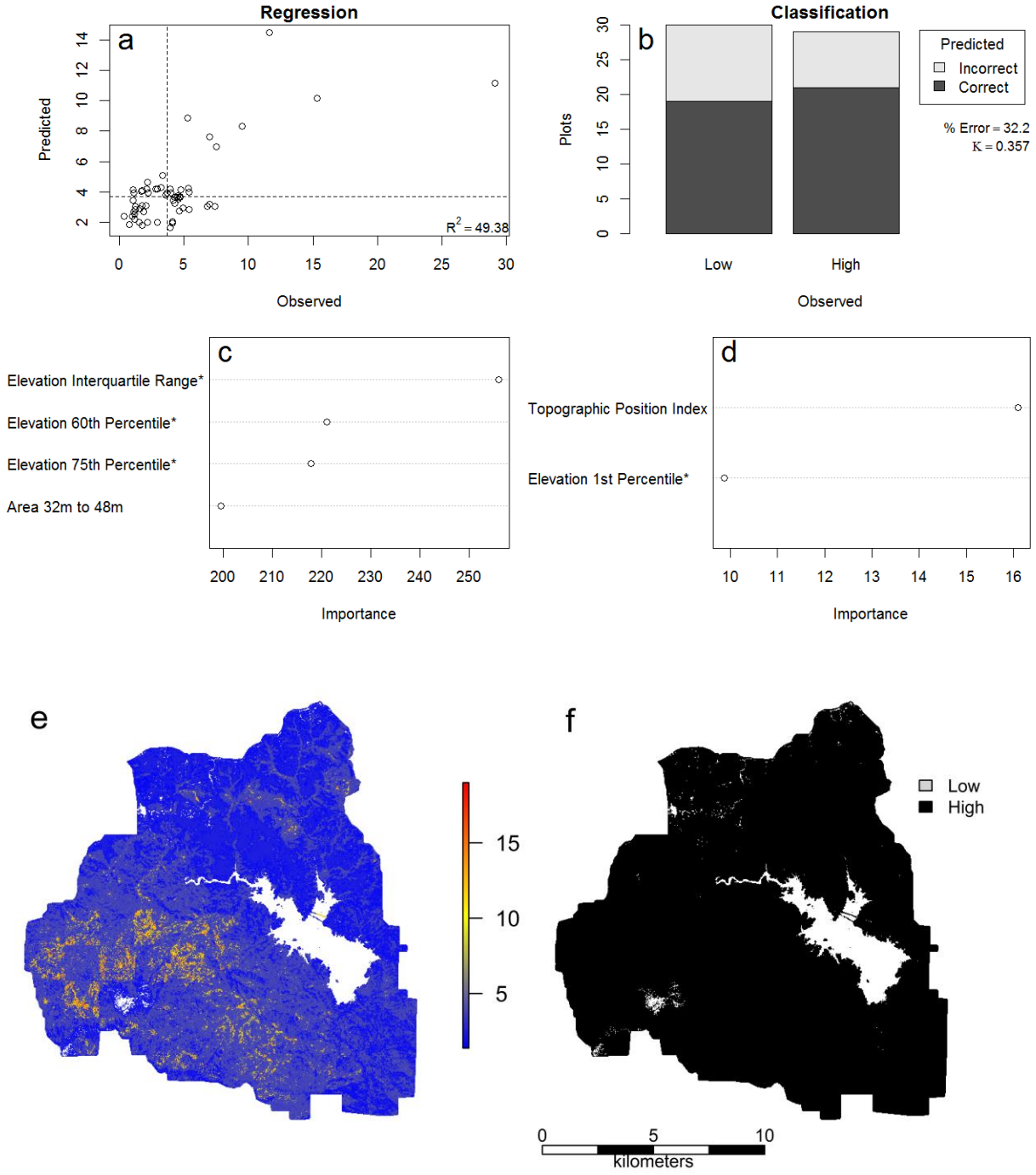


Figure 17-d. Models predicting all fine fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for fine fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

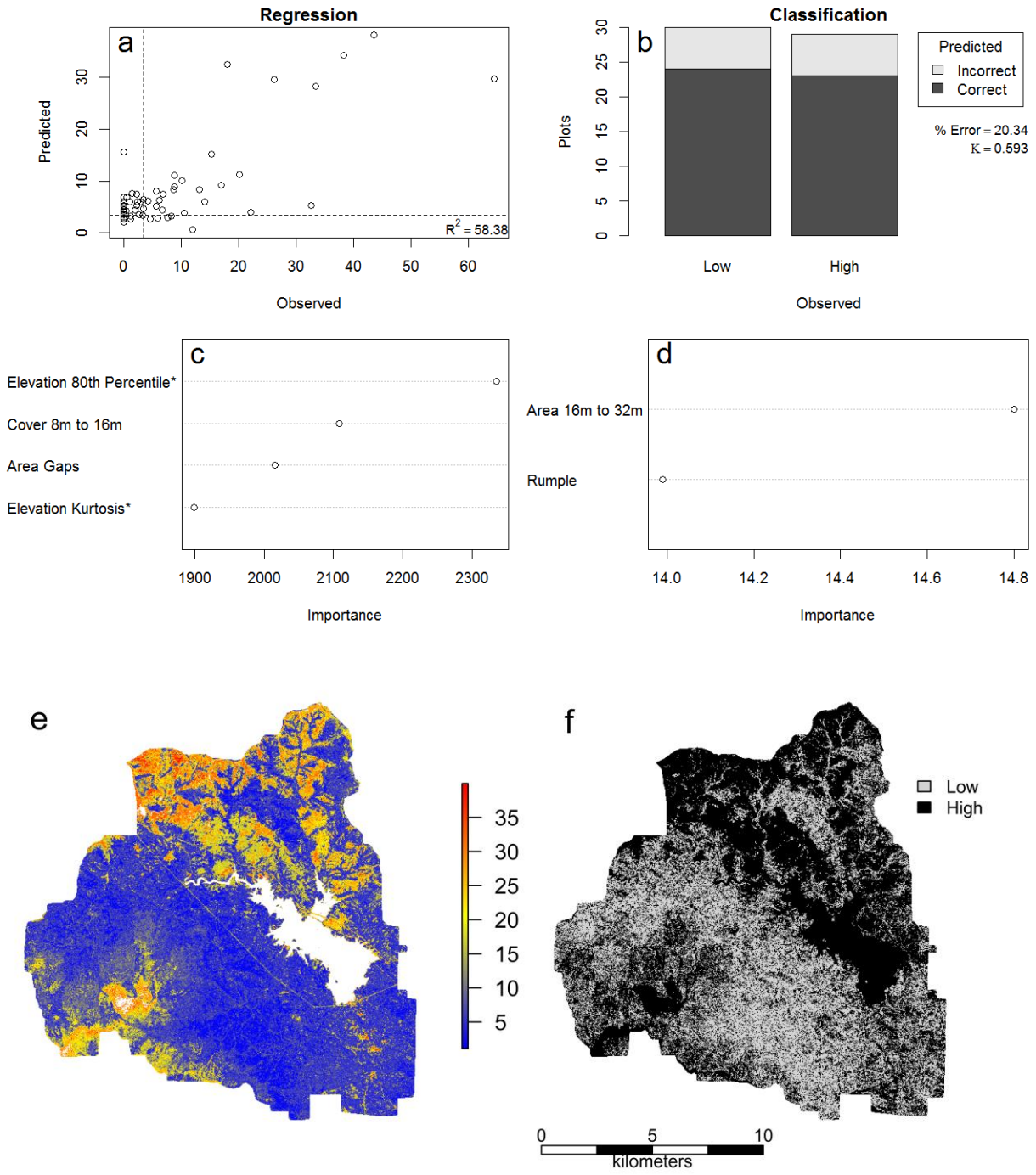


Figure 17-e. Models predicting sound thousand hour fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for sound thousand hour fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

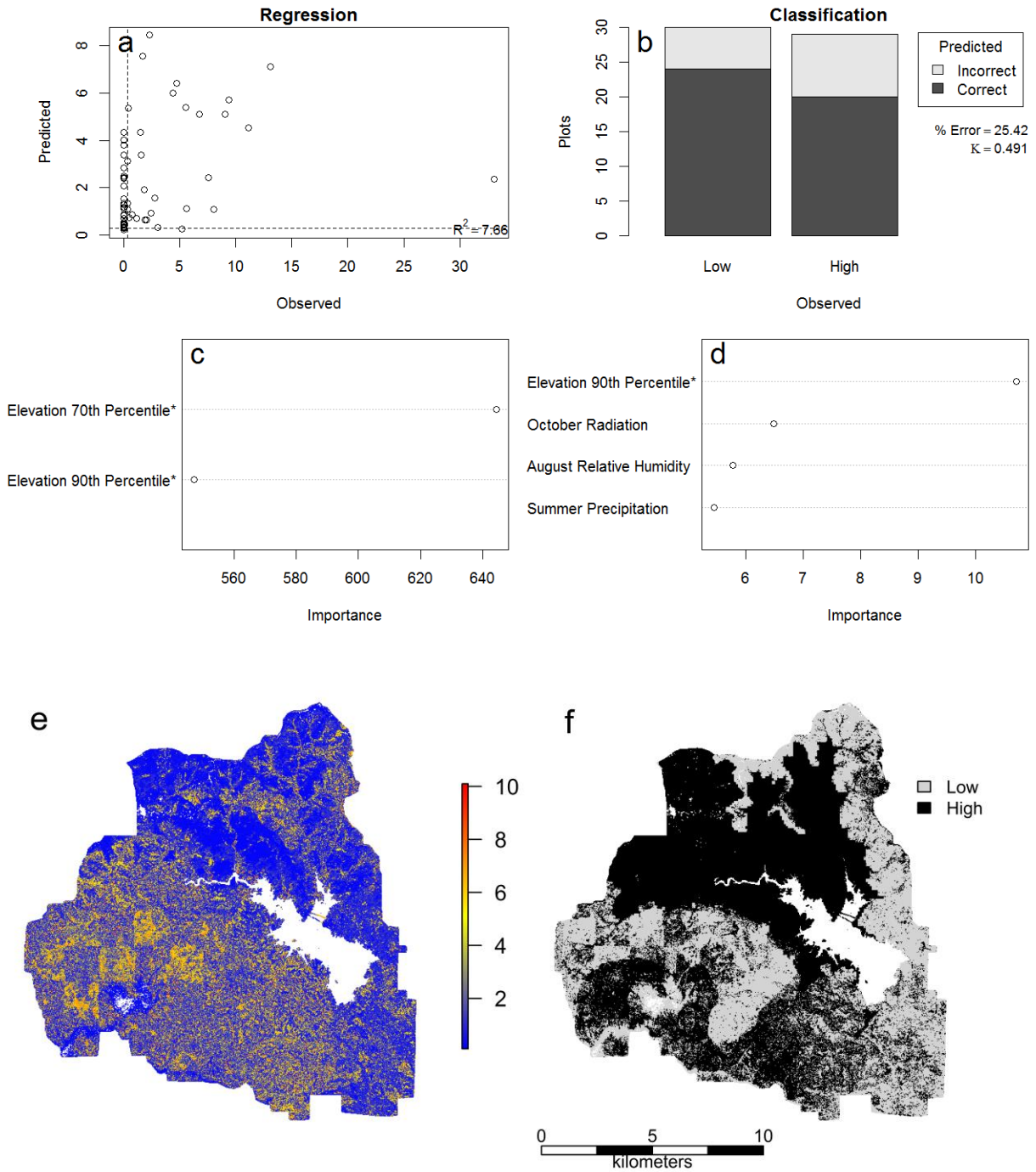


Figure 17-f. Models predicting rotten thousand hour fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for rotten thousand hour fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

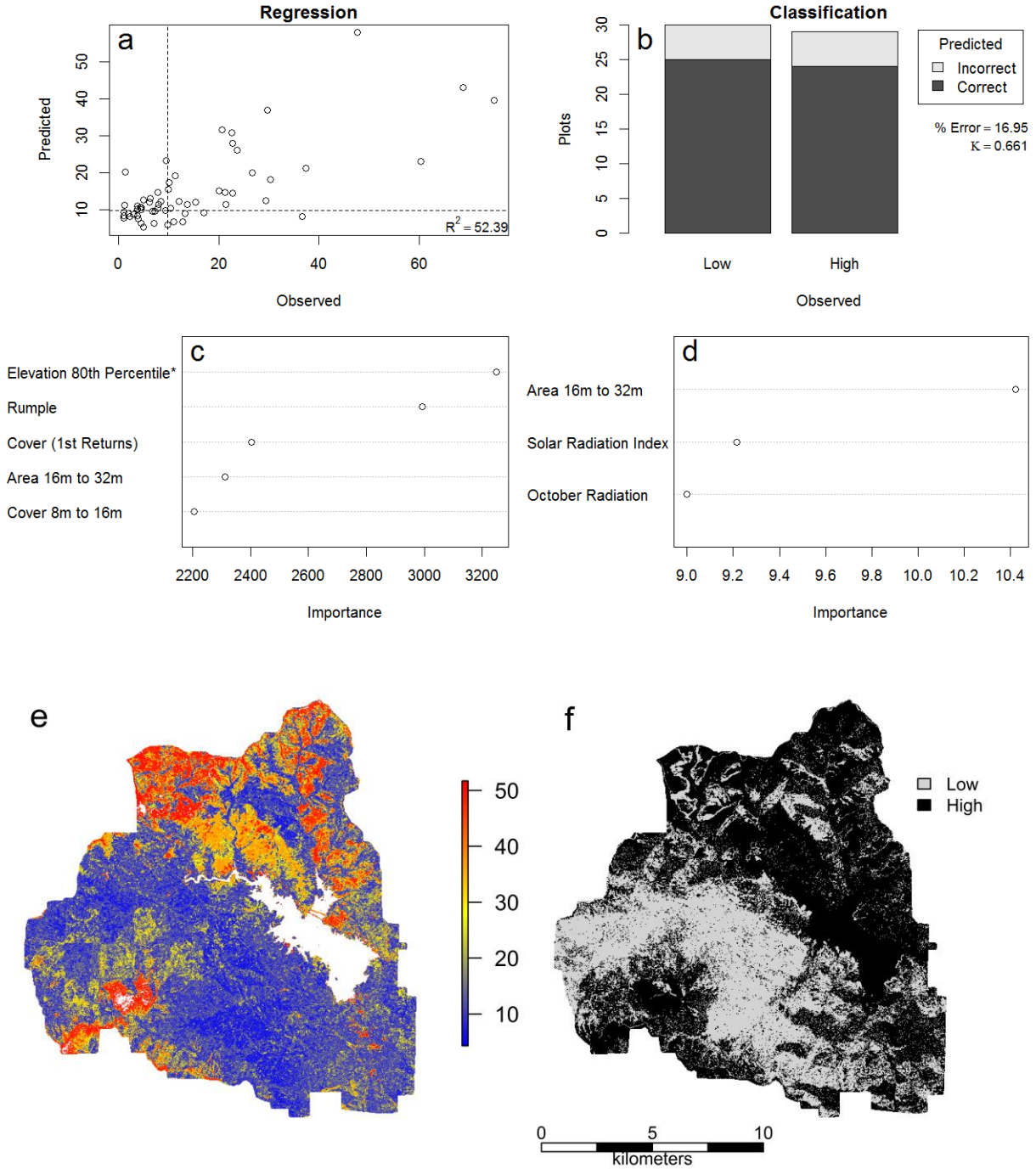


Figure 17-g. Models predicting total wood fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for wood fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

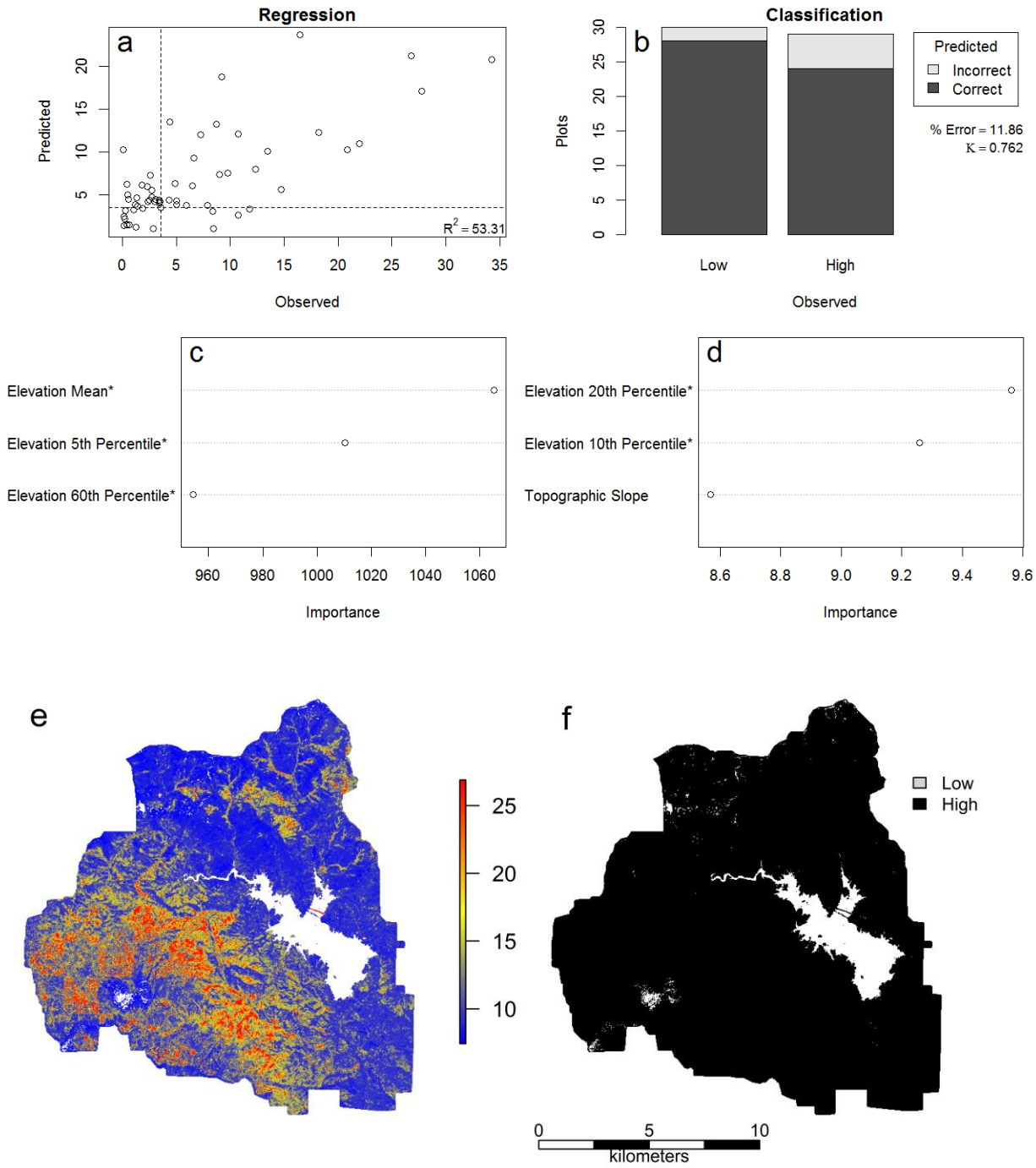


Figure 17-h. Models predicting duff fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for duff fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

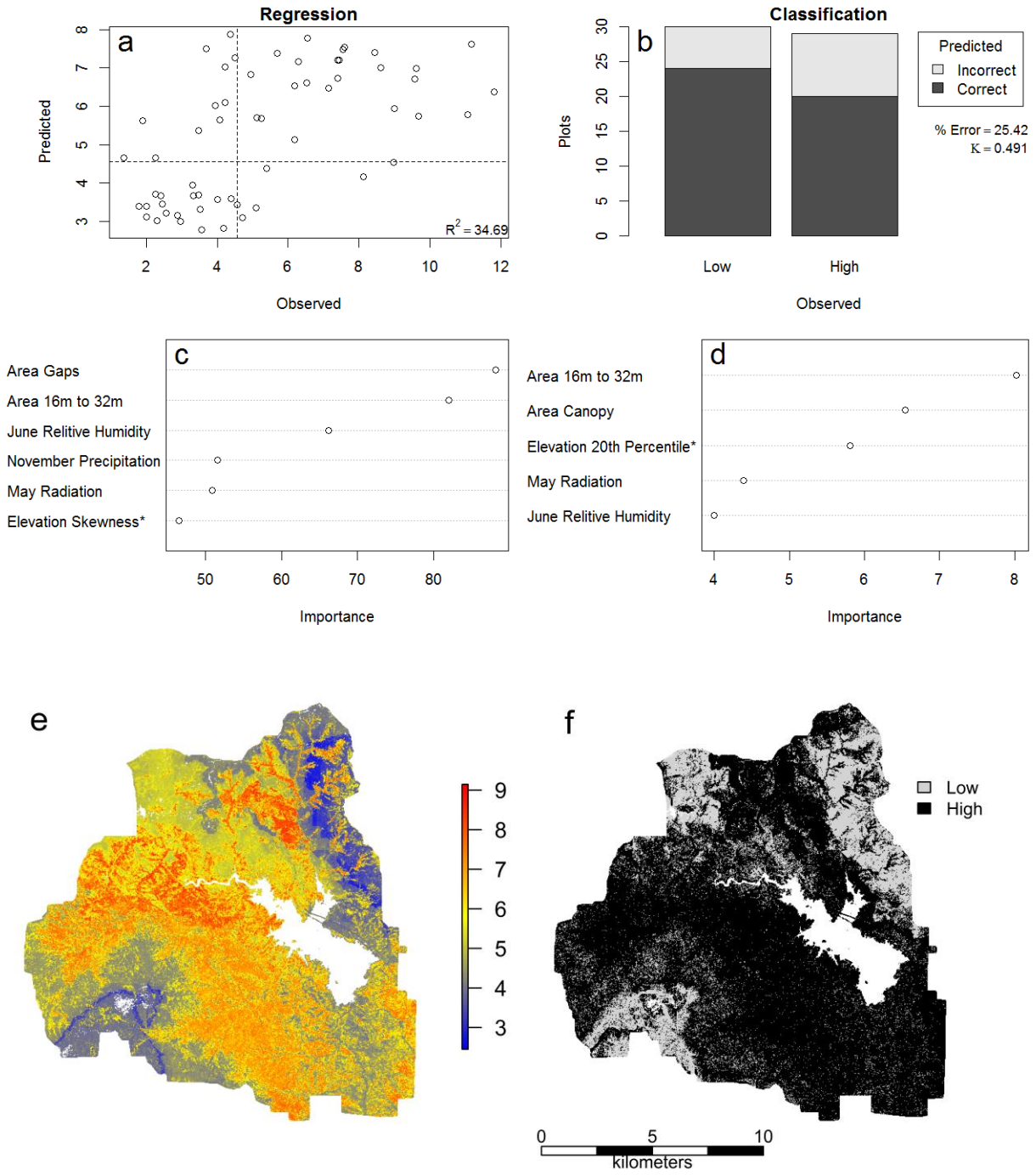


Figure 17-i. Models predicting litter fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for litter hour fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

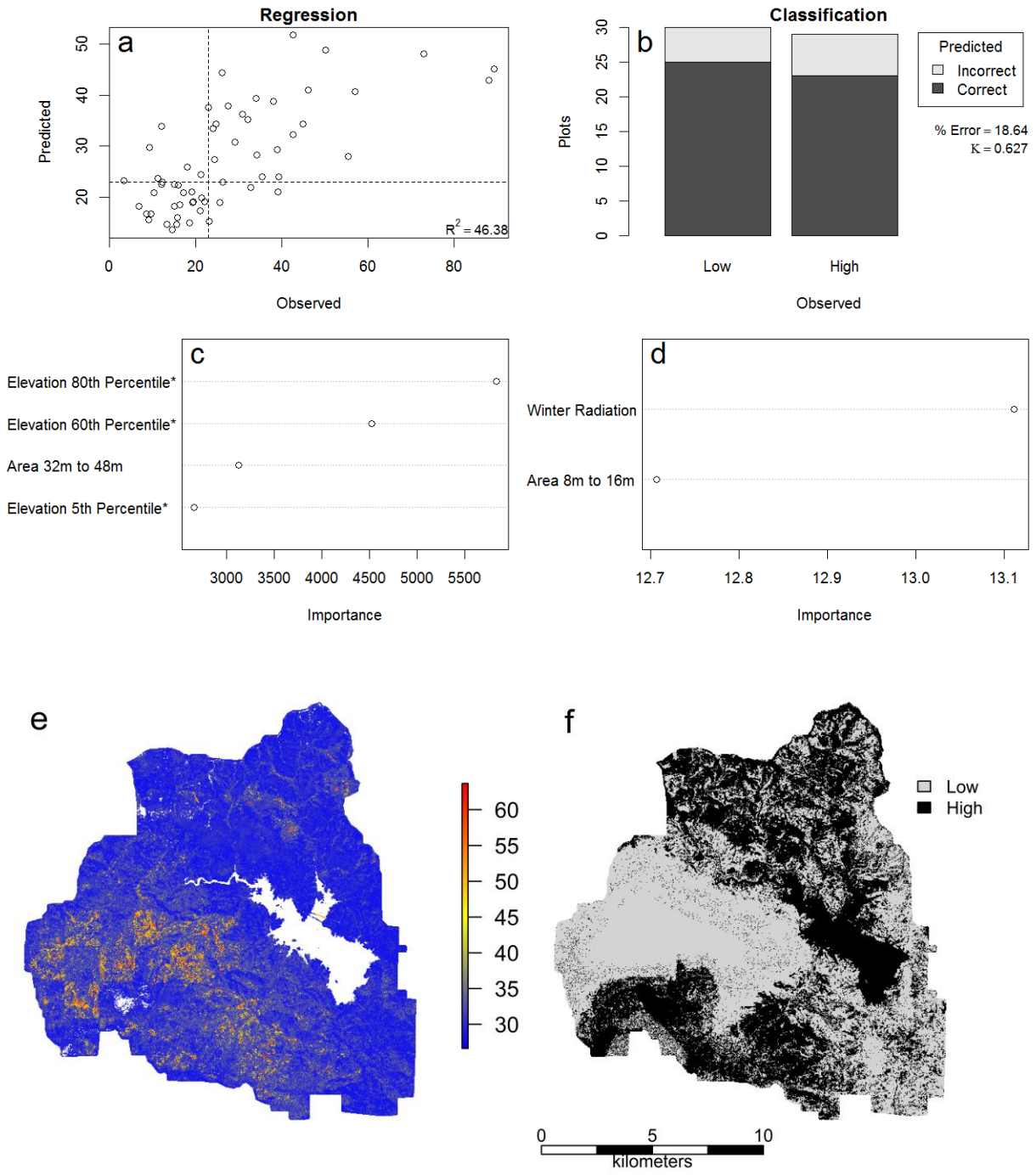


Figure 17-j. Models predicting total all fuel loadings. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel loading values for all fuels, representing (e) tons per acre and (f) above or below the median value observed in training plots.

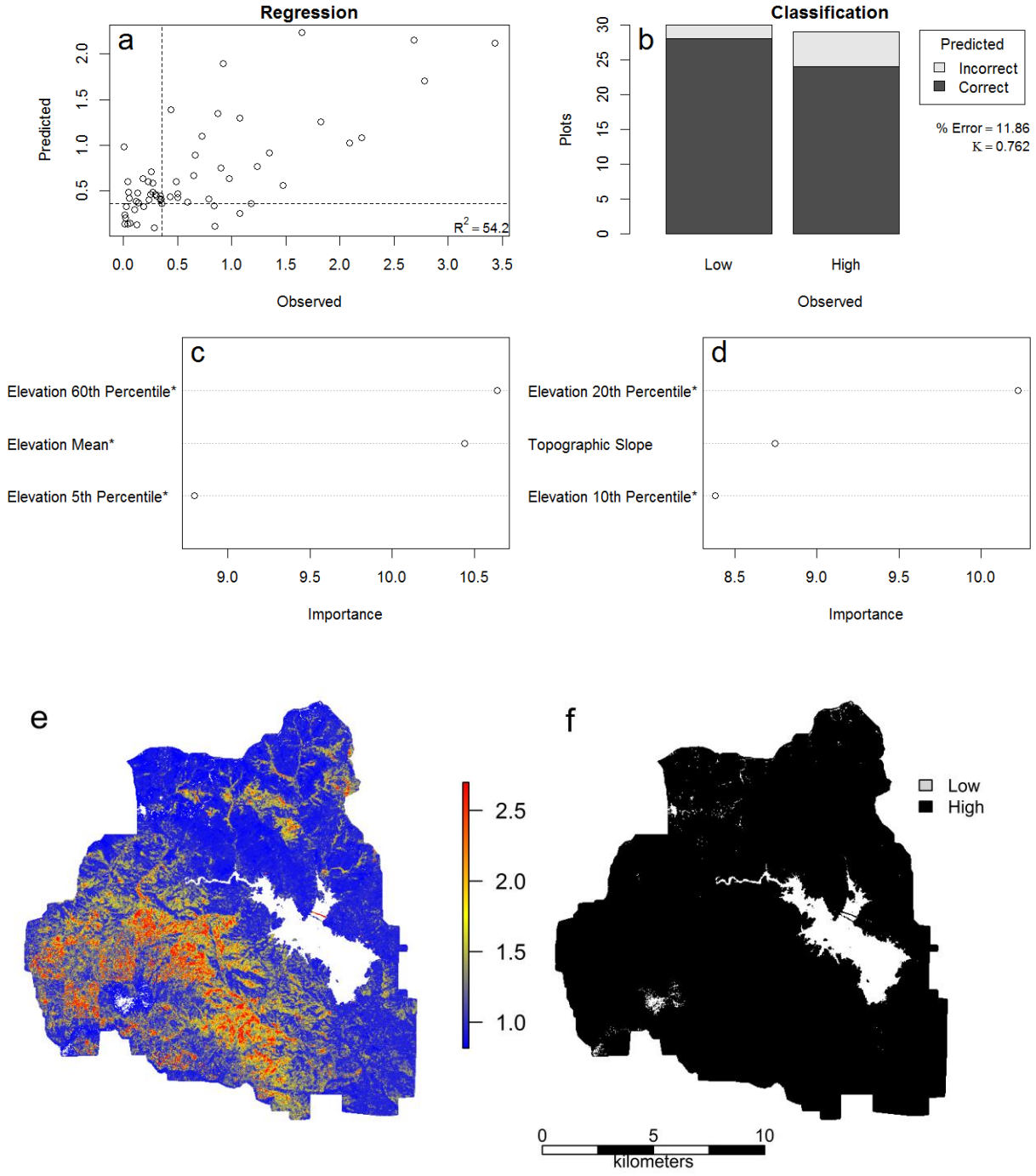


Figure 17-k. Models predicting duff depth. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel depth values for duff, representing (e) depth in cm and (f) above or below the median value observed in training plots.

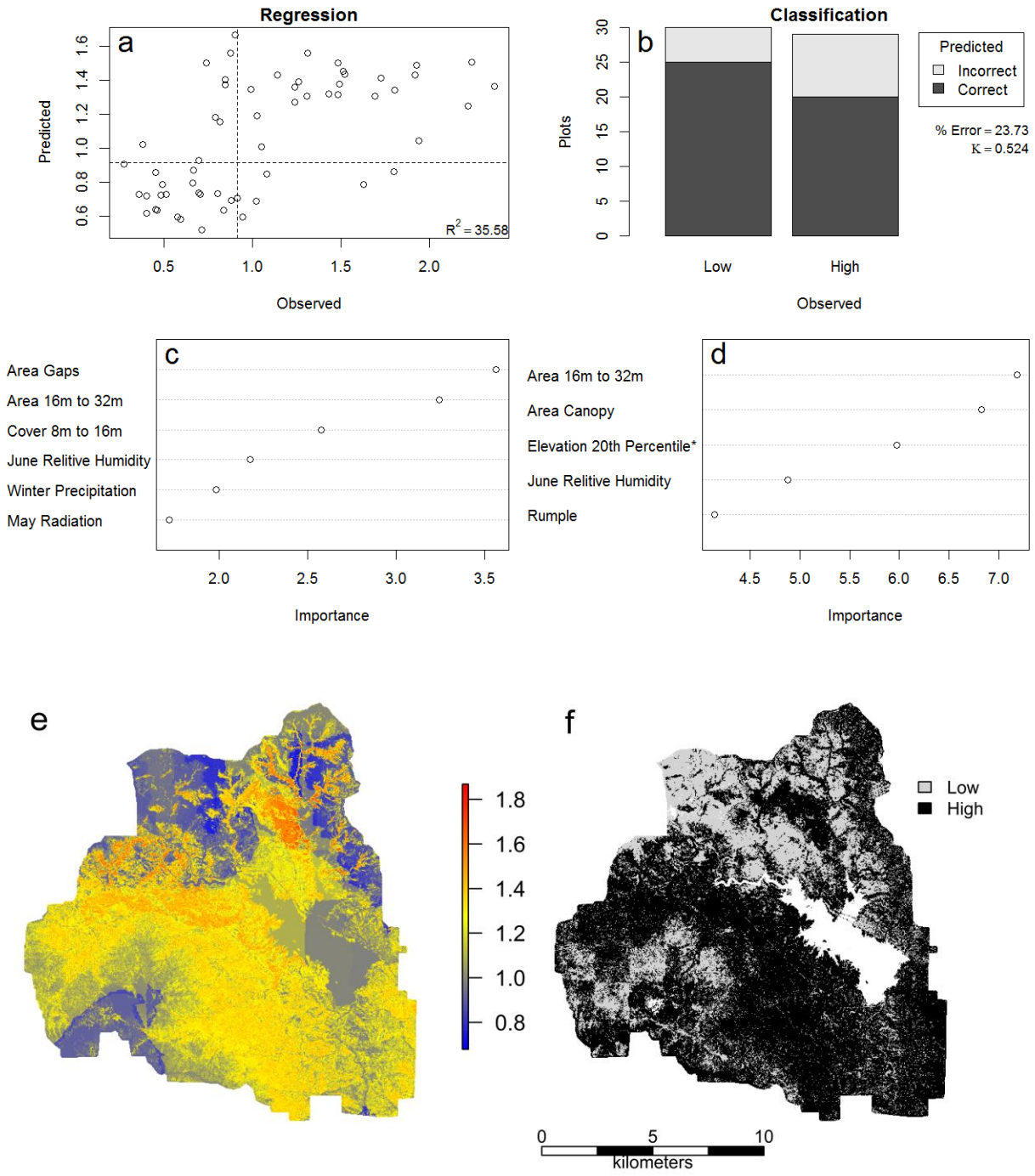


Figure 17-k. Models predicting litter depth. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel depth values for litter, representing (e) depth in cm and (f) above or below the median value observed in training plots.

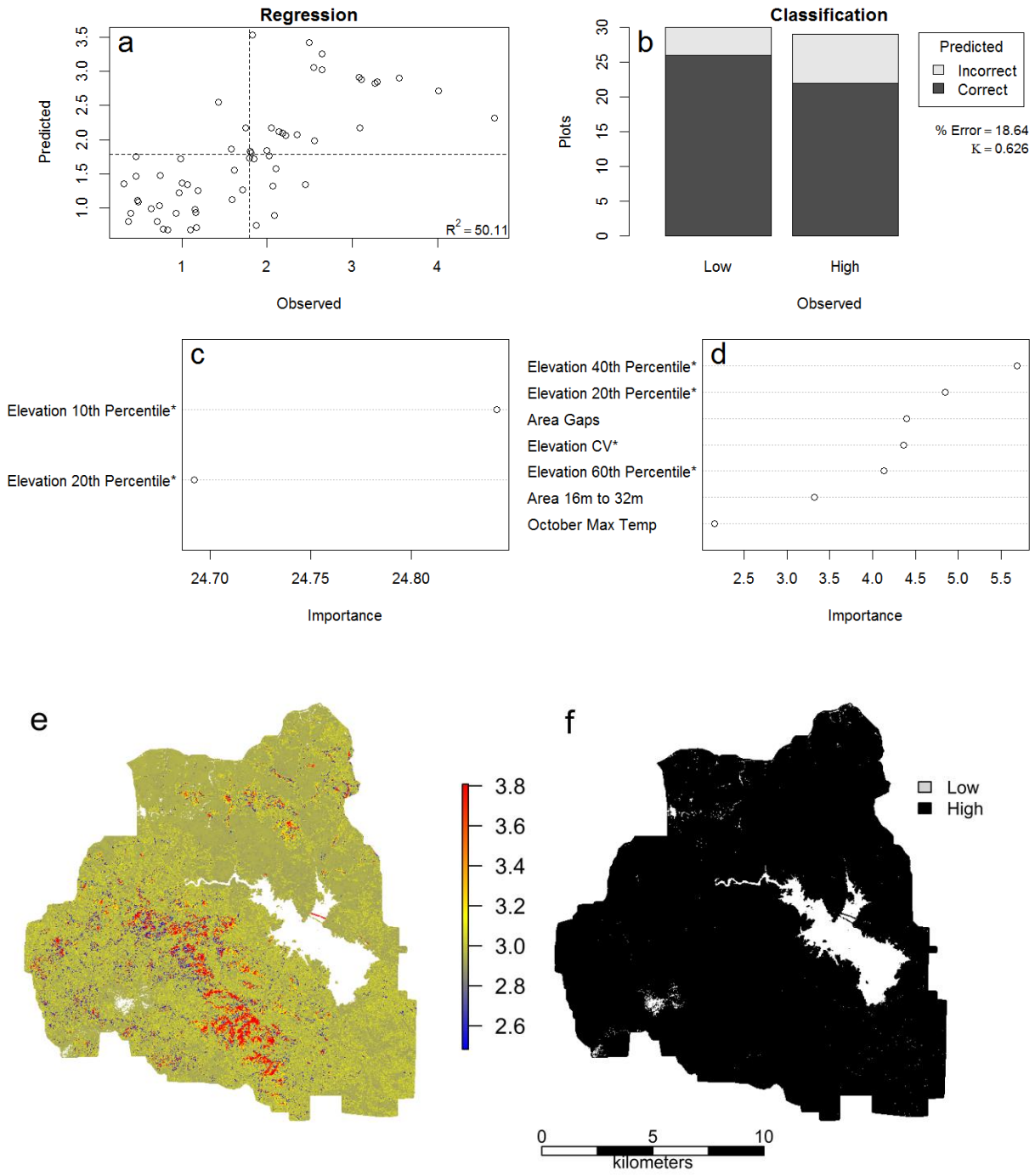


Figure 17-l. Models predicting total depth. Plots on the left of the figure describe the regression model, while plots on the right of the figure describe the classification model. Plots (a) and (b) describe the relationship between predicted and observed response values. Plots (c) and (d) describe the relative importance of the metrics used in building said models. (*) indicates elevation from forest floor. Plots (e) and (f) are maps of predicted fuel depth values for litter and duff, representing (e) depth in cm and (f) above or below the median value observed in training plots.

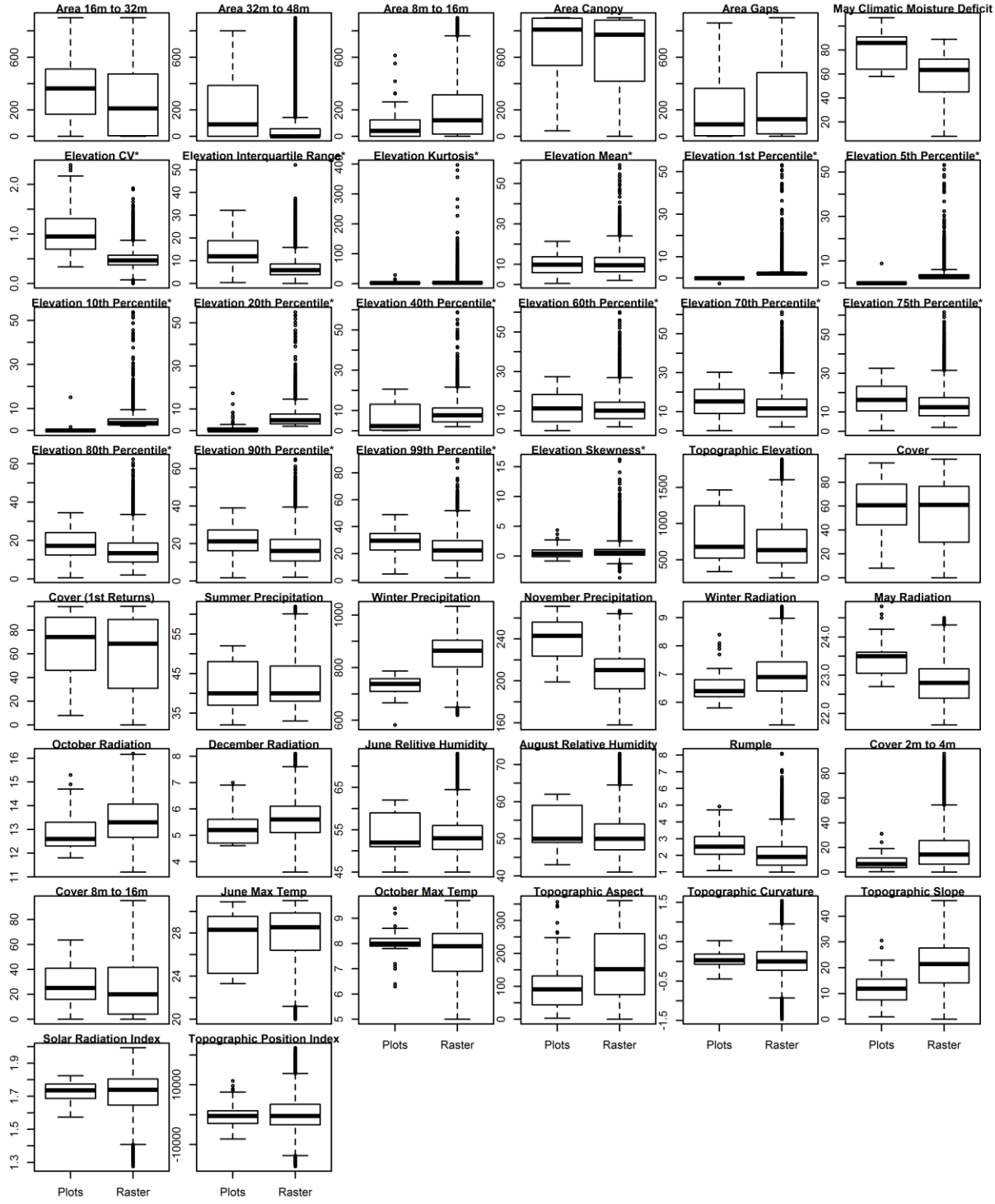


Figure 18. Distributions of plot level and study area level predictor metrics. Ranges for study area level data often exceed that of plot level data, causing models to predict outside the range of training data.

4. Delivered Products

This section describes the structure and content of the products delivered with this report. Note that not all of these products were produced for all study areas, and may not be present for your study area. Refer to sections 2.3 and 3. For a description of metrics produced for each study area.

4.1 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 m by 30 m (98.424 ft). Therefore, the values of a single cell summarize the conditions within a 30 mby 30 m area (Figure 19).

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 >2 m 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 m. 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 2015). 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 within the cell at or below the upper cutoff of the stratum and above the height cutoff (Mcgaughey 2015). 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 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 (Figure 20). 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).

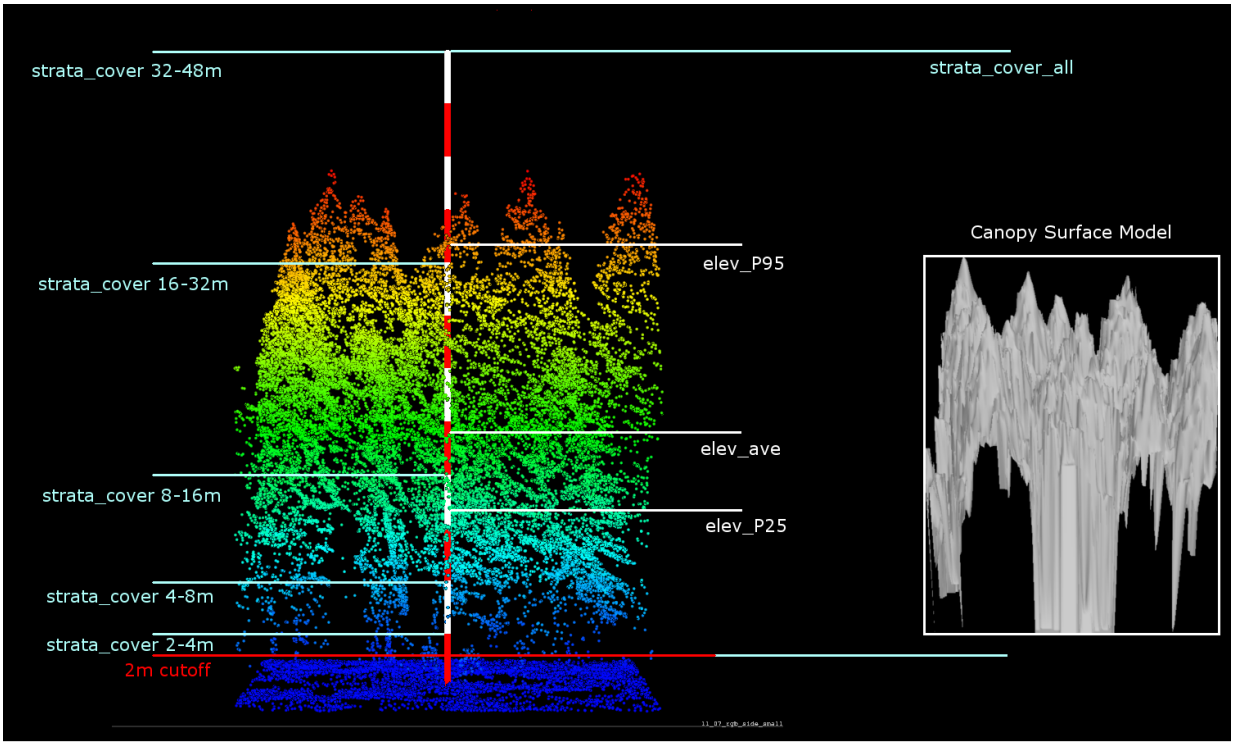


Figure 19. Illustration of key metrics. All metrics were calculated from LiDAR return data except for rumple which was calculated from a 1 m canopy surface model. Cover and height return values were calculated for returns >2 m above the ground. Metric names shown correspond to rasters supplied with this final report.

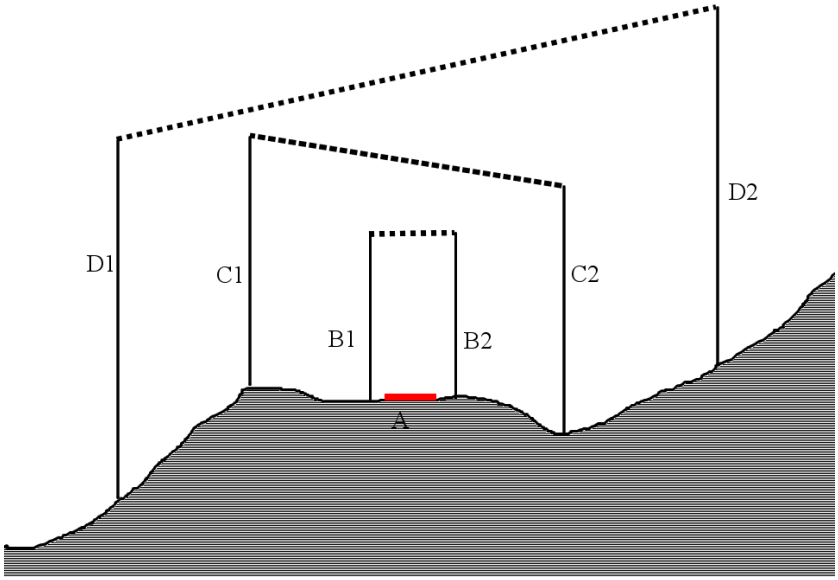


Figure 20. Illustration of the effects of window size on topography metric calculations. A profile view of an example landscape is shown above. Pixel (A) is calculated at three different scales. At the scale of the distance between lines B1 and B2, the slope for Cell A would be close to zero. At the 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.

4.2 Listing of All Supplied FUSION Metrics

Directory:

- FUSION_Data

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. Please contact us with any questions.

4.2.1 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 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

4.2.2 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_1p0METERS (contains the 1 meter resolution canopy height model)
- CanopyHeight_1p5METERS (contains the 1.5 meter resolution canopy height model)
- CanopyHeight_2p0METERS (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

- 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)

Tree-approximate objects and related metrics

- Segments_0p75METERS (contains the tree-approximate objects)

4.2.3 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'. Separate files exist for both the 2010 and 2011 acquisitions. The shape files included are:

DeliveryTiles.shp

ProcessingBlocks.shp

ProcessingTiles.shp

4.2.4 Canopy Metrics Calculated From LiDAR Return Data

Directory:

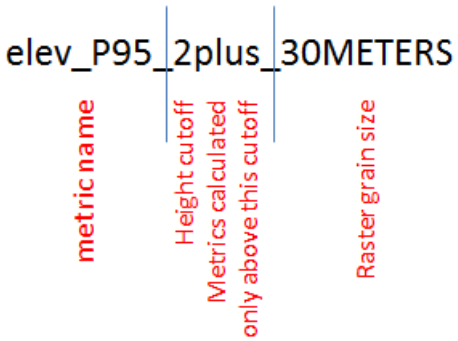
- Metrics_30METERS

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

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

4.2.5 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 covers 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

4.2.6 Canopy Height Models

Directories:

- CanopyHeight_0p75METERS
- CanopyHeight_1p0METERS
- CanopyHeight_1p5METERS
- CanopyHeight_2p0METERS

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.

4.2.7 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.

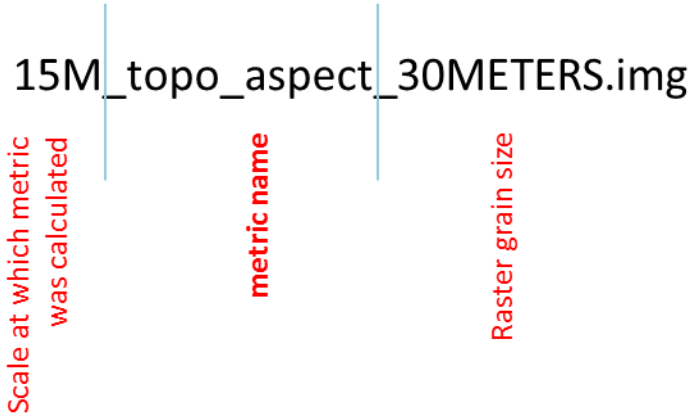
4.2.8 Topographic metrics

Directory:

- TopoMetrics_30METERS
- NormalizedTPIMetrics_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.) 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.

4.2.9 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

4.3 Landscape (FRAGSTATS) Metrics

Directory:

- FRAGSTATS

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. See section 2.3.3 for more detail.

Formulas for calculating the FRAGSTATS metrics can be found in the FRAGSTATS documentation (McGarigal, Cushman, and Ene 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

4.4 Structure Classification

Directory:

- Structure_Classes
  - figures
  - Input_metrics
  - raster

A single layer is provided describing the distribution of structure classes across the study area. Structure classes and numbered sequentially, and correspond to the classes described in section 3.1.1, 3.2.1, and 3.3.1. This layer is found in the "raster" directory

Multiple predictor metrics were used to classify and predict structure classes (section 2.3.4). These metrics can be found in the "input_metrics" directory. Figures produced describing the structure classes can be found in the "figures" directory.

4.5 Fuel Model Layers

Directory:

- Fuel_Models
  - figures
  - rasters
    - classification
    - regressions

Two fuel model layers are provided for each fuel loading metric reported in the fuel loadings plot data (section 2.3.5). These models are sorted into two directories labeled "classifications" and "regressions." Layers are labeled with "CloudMetrics_15m" to indicate that

a 15 meter radius circle was used when clipping the point data to be run though the CloudMetrics tool to produce the independent variables used to produce these models. Models are subsequently labeled with the fuel loading metric which they describe. Several figures and tables describing the extent and accuracy of these models can be found in the "figures" directory.

4.6 Climate Data

Directory:

- Climate_Data
  - CWNA
  - AETDEF

Climate data produced by the ClimateWNA tool is presented here as a set of rasters for the study area. All metrics were calculated using a 90 meter window, and downscaled to 30 meters. A key to interpreting the metrics provided can be found at http://climatemodels.forestry.ubc.ca/climatewna/downloads/help.pdf. AET and Deficit layers produced using these climate metrics can be found in the "AETDEF" folder.

5. References

Breidenbach, Johannes et al. 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 Sensing of Environment* 114(4): 911–24.

Breiman, L. 2001. "Random Forests." *Machine learning* 45(1): 5–32.

Erdody, Todd L., and L. Monika Moskal. 2010. "Fusion of LiDAR and Imagery for Estimating Forest Canopy Fuels." *Remote Sensing of Environment* 114(4): 725–37.

Genuer, Robin, Jean-michel Poggi, and Christine Tuleau-malot. 2015. "VSURF : An R Package for Variable Selection Using Random Forests." *The R Journal* 7(December): 19–33.

Jeronimo, S. 2015. "LiDAR Individual Tree Detection for Assessing Structurally Diverse Forest Landscapes." *University of Washington*: 65.

Kaartinen, Harri et al. 2012. "An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning." *Remote Sensing* 4(4): 950–74.

Kane, Van R., Robert J. McGaughey, et al. 2010. "Comparisons between Field- and LiDAR-Based Measures of Stand Structural Complexity." *Canadian Journal of Forest Research* 40(4): 761–73.

Kane, Van R., Jonathan D. Bakker, et al. 2010. "Examining Conifer Canopy Structural Complexity across Forest Ages and Elevations with LiDAR Data." *Canadian Journal of Forest Research* 40: 774–87.

Kane, Van R. et al. 2014. "Assessing Fire Effects on Forest Spatial Structure Using a Fusion of Landsat and Airborne LiDAR Data in Yosemite National Park." *Remote Sensing of Environment* 151: 89–101.

Kottek, Markus et al. 2006. "World Map of the Köppen-Geiger Climate Classification Updated." *Meteorologische Zeitschrift* 15(3): 259–63.

Legendre, Pierre, and Louis Legendre. 2012. *Numerical Ecology*. 24th ed. Elsevier.

Lutz, James A, Jan W Van Wagtendonk, and Jerry F Franklin. 2010. "Climatic Water Deficit , Tree Species Ranges , and Climate Change in Yosemite National Park." : 936–50.

McGarigal, Kevin, Samuel A. Cushman, and E. Ene. 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."

Mcgaughey, Robert J. 2015. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*.

Parker, Geoffrey G. et al. 2004. "Three-Dimensional Structure of an Old-Growth Pseudotsuga-Tsuga Canopy and Its Implications for Radiation Balance, Microclimate, and Gas Exchange." *Ecosystems* 7(5): 440–53.

PRISM Climate Group. 2015. "Average Monthly and Annual Climate Conditions over the Most Recent Three Full Decades." *Oregon State University*.

R Core Team. 2017. "R: A Language and Environment for Statistical Computing." https://www.r-project.org/.

Stephenson, Nathan L. 1990. "Climatic Control of Vegetation Distribution : The Role of the Water Balance." *The American Naturalist* 135(5): 649–70.

———. 1998. "Actual Evapotranspiration and Deficit : Biologically Meaningful Correlates of Vegetation Distribution across Spatial Scales." *Journal of Biogeography* 25(5): 855–70.

Stuart, John D, and Christopher A Lee. 2003. "VEGETATION ALLIANCES AND ASSOCIATIONS OF THE WHISKEYTOWN NATIONAL RECREATION AREA." *Humboldt State University*.

Thornthwaite, C W, and J R Mather. 1955. "The Water Balance." *Publications in Climatology* 8(1): 1–104.

United States Department of Agriculture, Soil Survey Staff, and Natural Resources Conservation Service. 2016. "Soil Survey Geographic (SSURGO) Database."

Vauhkonen, Jari et al. 2012. "Comparative Testing of Single-Tree Detection Algorithms under Different Types of Forest." *Forestry* 85(1): 27–40.

Vincent, L, and P Soille. 1991. “Watershed in Digital Spaces: An Efficient Algorithm Based on Immersion Simulation.” *IEEE Trans. Pattern Anal. Mach. Intell.* 13(6): 583–98.

Wang, Tongli, Andreas Hamann, David L. Spittlehouse, and Trevor Q. Murdock. 2012. “ClimateWNA-High-Resolution Spatial Climate Data for Western North America.” *Journal of Applied Meteorology and Climatology* 51(1): 16–29.

Zevenbergen, LW, and CR Thorne. 1987. “Quantitative Analysis of Land Surface Topography.” *Earth Surface Processes and Landforms* 12: 47–56.