esa ECOSPHERE

Accessible light detection and ranging: estimating large tree density for habitat identification

Heather A. Kramer,[1,4,†] Brandon M. Collins,[2,3] Claire V. Gallagher,[3] John J. Keane,[3] Scott L. Stephens,[1] and Maggi Kelly[1]

[1]*Ecosystem Sciences Division, Department of Environmental Science, Policy, and Management, University of California, 130 Mulford Hall, Berkeley, California 94720 USA*
[2]*Center for Fire Research and Outreach, University of California, Berkeley, California 94720 USA*
[3]*USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, California 95618 USA*

Citation: Kramer, H. A., B. M. Collins, C. V. Gallagher, J. J. Keane, S. L. Stephens, and M. Kelly. 2016. Accessible light detection and ranging: estimating large tree density for habitat identification. Ecosphere 7(12):e01593. 10.1002/ecs2.1593

Abstract. Large trees are important to a wide variety of wildlife, including many species of conservation concern, such as the California spotted owl (*Strix occidentalis occidentalis*). Light detection and ranging (LiDAR) has been successfully utilized to identify the density of large-diameter trees, either by segmenting the LiDAR point cloud into individual trees, or by building regression models between variables extracted from the LiDAR point cloud and field data. Neither of these methods is easily accessible for most land managers due to the reliance on specialized software, and much available LiDAR data are being underutilized due to the steep learning curve required for advanced processing using these programs. This study derived a simple, yet effective method for estimating the density of large-stemmed trees from the LiDAR canopy height model, a standard raster product derived from the LiDAR point cloud that is often delivered with the LiDAR and is easy to process by personnel trained in geographic information systems (GIS). Ground plots needed to be large (1 ha) to build a robust model, but the spatial accuracy of plot center was less crucial to model accuracy. We also showed that predicted large tree density is positively linked to California spotted owl nest sites.

Key words: California; canopy height; habitat; large tree; light detection and ranging; spotted owl; tree density.

Received 25 July 2016; accepted 29 September 2016. Corresponding Editor: Dawn M. Browning.
Copyright: © 2016 Kramer et al. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
[4] Present address: SILVIS Lab, Department of Forest and Wildlife Ecology, University of Wisconsin–Madison, 1630 Linden Drive, Madison, Wisconsin 53706 USA.
† **E-mail:** hakramer@wisc.edu

Introduction

Large trees are critical components of many temperate forest ecosystems (Franklin et al. 2002, Lutz et al. 2012). Large trees have features that directly provide habitat for wildlife (e.g., broken tops and cavities), in addition to indirectly providing habitat by contributing to greater complexity in forest structure. Both aspects of large trees have been shown to be important for wildlife species of conservation concern. Numerous studies have shown the California spotted owl's (*Strix occidentalis occidentalis*) (CSO) association with large, old-growth trees, and structurally complex stands used for nesting and roosting (Bias and Gutierrez 1992, Gutierrez et al. 1992, Moen and Gutierrez 1997, Keane 2014). CSO populations are declining in the Sierra Nevada, and they are currently under review for potential listing under the Endangered Species Act. Thus, information on the distribution and abundance of important large tree habitat elements and structurally complex forest stands is needed to inform assessment and development of

conservation strategies for CSOs, and more broadly, Sierra Nevada forest landscapes. In addition to CSOs, large tree habitat is important for numerous wildlife species including fishers (*Martes pennanti*), northern goshawks (*Accipiter gentilis*), woodpeckers, and others (Beier and Drennan 1997, Greenwald et al. 2005, Seavy et al. 2009, Hollenbeck et al. 2011, Zielinski 2014).

In the past, imagery from passive remote sensors, such as LANDSAT, was widely used to estimate the structure of forests (Forsman 1995, Hunter et al. 1995, Moen and Gutierrez 1997, McDermid et al. 2005). Light detection and ranging (LiDAR) is a form of active remote sensing that is better able to detect the height of vegetation than passive remote sensing, and therefore may be a very useful tool for identifying wildlife habitat (Lefsky et al. 2002, Vierling et al. 2008, Martinuzzi et al. 2009, Selvarajan et al. 2009, Kelly and Di Tommaso 2015, Kramer et al. 2016). This is particularly true in areas with tall (and likely large diameter) trees (Bergen et al. 2009, Wing et al. 2010, Garcıa-Feced et al. 2011, Ackers et al. 2015). Ackers et al. (2015) found that LiDAR was a better predictor than LANDSAT of Spotted Owl habitat, which depended heavily on large tree density and overall canopy height (Ackers et al. 2015).

Methods for estimating large tree density from LiDAR include individual tree segmentation and statistical modeling that utilizes one to many LiDAR-derived independent variables. Tree segmentation algorithms segment the canopies of individual tree crowns from the LiDAR point cloud (Popescu et al. 2003, Chen et al. 2006, Li et al. 2012, Jakubowski et al. 2013*a*). Using these algorithms, it is possible to estimate the location of dominant or isolated stems and predict the number of large trees in a given area; however, the accuracy of these predictions decreases with dense canopy cover, when stems are close together, and when trees are codominant, intermediate, or suppressed (Kaartinen et al. 2012, Swetnam and Falk 2014). Tree segmentation has been successfully implemented in wildlife studies (Garcıa-Feced et al. 2011), and while it requires fairly simple LiDAR input, it necessitates a multi-step process (where errors can be compounded) to calculate the density of stems in a given diameter class: (1) calculate each stem's location and maximum crown height, (2) back-calculate each tree's diameter, and (3) calculate the stem density of the diameter range of interest. Furthermore, this analysis also requires training data where all stems are mapped, which can be very labor intensive, especially for larger plots (>0.1 ha). Unfortunately, this work flow is not possible for most land managers due to the limited availability of LiDAR with high point densities (due to funding limitations or age of LiDAR acquisition), equipment, personnel for fieldwork, limited training in and access to tree segmentation software, and limited time to carry out complex processing tasks.

Statistical models using LiDAR to predict tree density can be more abstract, and while less accurate than predictions of tree height or basal area, are also able to make useful predictions (Hudak et al. 2006, Lee and Lucas 2007, Jakubowski et al. 2013*b*). While these models do not require the complex algorithms and high-density LiDAR needed by tree detection algorithms, they often utilize a suite of LiDAR-derived variables. These variables are not standard deliverables from LiDAR acquisitions and typically require expertise of a LiDAR specialist. Furthermore, once a model is created, it can be hard to understand the ecological underpinnings, and the model must be re-evaluated when moving between different areas or forest types. Simpler models (e.g., fewer variables, less intensive statistics) are uncommon in LiDAR applications for natural resources.

In this study, we explore regression-based approaches for accurately quantifying the density of large trees. While tree segmentation is another option, we intentionally avoided it to investigate less demanding approaches in terms of LiDAR point density, field plots, processing hardware and software, and expertise of processing personnel. Our specific research questions were as follows: (1) When estimating large tree density from LiDAR, what is the difference in predictive accuracy between (A) a multiple regression model comprised of many LiDAR-derived variables and (B) a simple linear regression model derived from the canopy height model (CHM)? (2) Does plot size or plot center accuracy influence the strength of this relationship? and (3) Can our LiDAR-derived large tree density estimates be used to identify important

structural habitat characteristics of California spotted owl (CSO) nest sites?

Methods

Study area

Our study was conducted in the Meadow Valley area of the Plumas National Forest, which is in the northern Sierra Nevada of California (centered at 39°55′ N, 121°03′ W; Fig. 1). With a Mediterranean climate, most of its 1050 mm/yr of precipitation falls during the winter (Ansley and Battles 1998). The 22,510-ha (55,623 ac) landscape is made up of forest, montane chaparral, and meadows, and falls between 1050 and 2150 m in elevation (Collins et al. 2013, Kramer et al. 2014). Mixed conifer tree species predominate, including ponderosa pine (*Pinus ponderosa*), Jeffrey pine (*Pinus jeffreyi*), sugar pine (*Pinus lambertiana*), Douglas-fir (*Pseudotsuga menziesii*), white fir (*Abies concolor*), incense-cedar (*Calocedrus decurrens*), and California black oak (*Quercus kelloggii*) (Schoenherr 1992, Barbour and Major 1995). At higher elevations, smaller pockets of red fir (*Abies magnifica*) and western white pine (*Pinus monticola*) can be found. Lower densities of lodgepole pine (*Pinus contorta*), western juniper (*Juniperus occidentalis*), California hazelnut (*Corylus cornuta*), dogwood (*Cornus* spp.), and willow (*Salix* spp.) also occur. Before fire suppression began in the early 1900s, the historic fire regime consisted of primarily low- to moderate-severity fires burning at 7- to 19-year intervals (Moody et al. 2006).

Many different fuel reduction treatments were implemented across this landscape between 1999 and 2008 as part of the Herger-Feinstein Quincy Library Group Pilot Project (Herger and Feinstein 1998). The fuel treatments occurred across approximately 20% of the landscape and were intended to mitigate potential for uncharacteristically large and severe wildfire while conserving critical habitat for CSO and other species (Moghaddas et al. 2010). Multiple nesting sites of CSO have also been located and monitored across this study area (Stephens et al. 2014).

Field data

The entire Meadow Valley study area was systematically surveyed for CSO nesting sites between 2002 and 2012 using standardized survey protocols to determine occupancy and reproductive status (Blakesley et al. 2010, Stephens et al. 2014). As part of these protocols, efforts were made to locate the specific nest tree used by breeding owl pairs each year. A total of 13 CSO nest tree locations were documented and sampled using the standard Forest Inventory and Analysis (FIA) protocol described below. Field plots were sampled between 2004 and 2009. Plots were centered on all 13 known CSO nest trees within the study area and at 132 CSO foraging locations. Foraging locations were estimated from 10 owls using standard radio-telemetry techniques (White and Garrott 1990, Kenward 2001); we conducted vegetation plots at a random subsample from 436 total foraging locations, with each owl sampled equally. In addition, any foraging location within a fuels treatment also received a vegetation plot. Error ellipses for radio-telemetry locations are dependent on distance to the animal, change in angle between bearings, and elapsed time between bearings; we sampled vegetation only at foraging locations in which the error ellipse was less than 1 ha, the size of the largest vegetation subplot (Gallagher 2010). A total of 145 plots were sampled (of which 134 were used for our study due to incomplete coverage by the LiDAR point cloud or inaccuracies between spatial and non-spatial datasets). Plot centers were recorded with a Trimble GeoExplorer3, with a reported accuracy of 2–5 m (actual accuracies for each plot were not recorded).

Standard FIA plot protocol was implemented to collect plot and subplot data (USDA Forest Service 2001). Additionally, trees over 76 cm (30 in) in diameter at breast height (dbh) were measured on a 1-ha plot that encompassed all subplots. Only a single subplot of each size was utilized to maintain subplot independence and control for slight inaccuracies in plot center coordinates due to lower accuracy GPS. These plots measured 1 ha (2.47 ac), 1/10th ha (0.247 ac), and 1/60th ha (0.041 ac), with plot radii of 56.41, 17.95, and 7.31 m (185.1, 58.9, and 24.0 ft), respectively, laid out concentrically around the recorded plot center. The plot and subplot arrangement that we used is illustrated in Fig. 2. In all plots, trees over 76 cm (30 in) in dbh were measured and considered for our analyses to reflect the harvesting regulations described in the

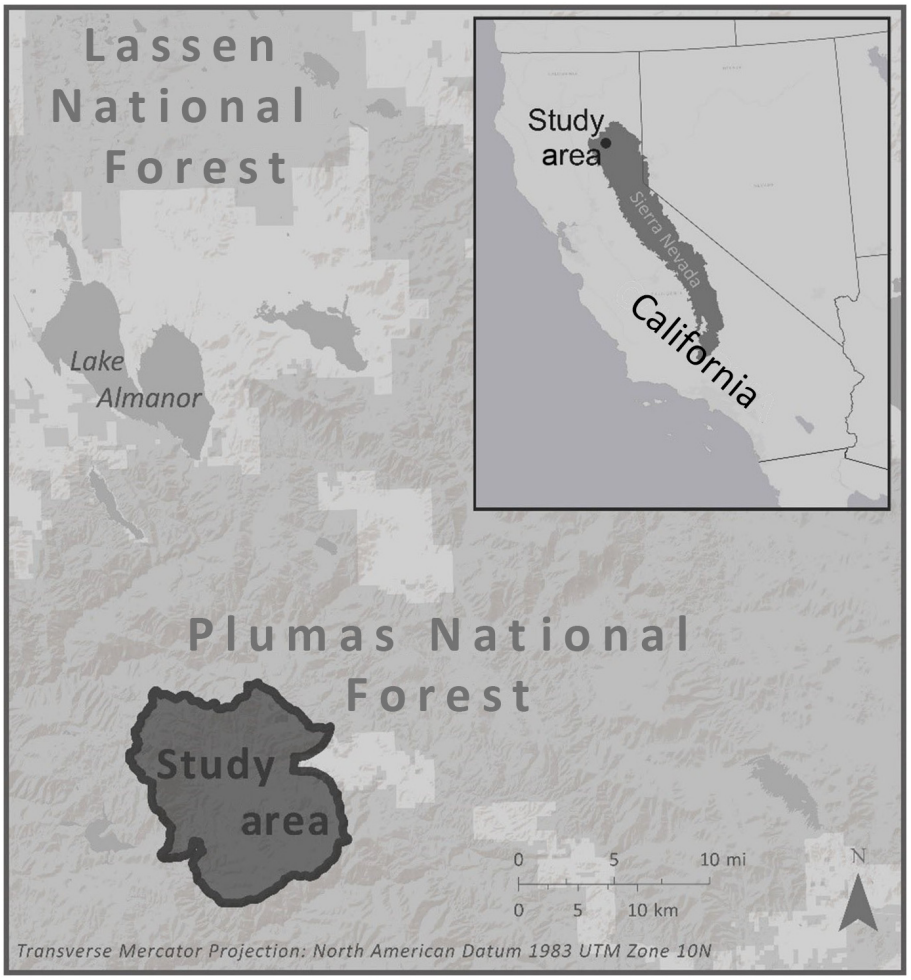

Fig. 1. Meadow Valley study area on the Plumas National Forest, California, measuring 22,510 ha.

US Forest Service 2001 Framework and the 2004 Sierra Nevada Forest Plan Amendment (US Forest Service 2001, 2004).

LiDAR data and processing

Watershed Sciences, Inc. collected aerial LiDAR over the Plumas and Lassen National Forests between 31 July 2009 and 11 August 2009. A Leica ALS50 Phase II laser system was used to collect LiDAR points utilizing a scan angle of 14° from nadir. A Leica RCD-105 39 megapixel digital camera was used to capture orthophotographs, which were processed with Leica's Calibration Post Processing software v.1.0.4. IPASCO v.1.3 (Heerbrugg, Switzerland) and the Leica Photogrammetry Suite v.9.2 were used to spatially place the photos. The vendor reported that average vertical and horizontal accuracy were 2.6 cm (1.02 in) and 7.2 cm (2.83 in), respectively, based on the mean divergence of points from ground survey point coordinates (3089 ground points were analyzed across four surveyed areas). An average point density of 4.68 points/m^2 (0.43 points/ft^2) was achieved. Although a variety of fuel reduction treatments were implemented on the landscape between field plot sampling and the LiDAR flight, no fuels were altered in the field plots.

The LiDAR point cloud was normalized, and variables were extracted for each 1-ha plot using LAStools (Isenburg 2011). Topographic and canopy structural variables were calculated (a total of 21 metrics; see Appendix S1: Table S1 for a detailed description of each metric). The CHM

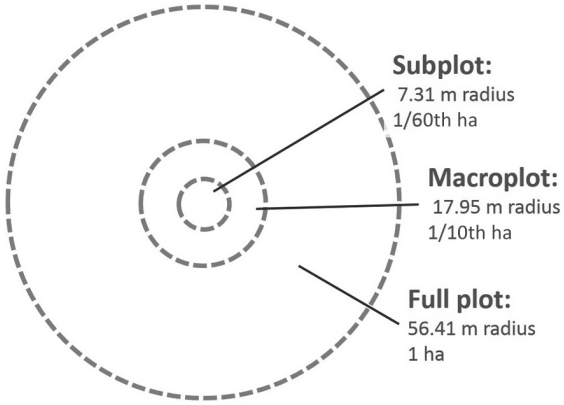

Fig. 2. Layout of the three concentric plot sizes that were used for this study. Note that all standard Forest Inventory and Analysis plots were collected, but only a single plot of each size was used to minimize spatial autocorrelation between macro- and subplots.

was generated at 2 m resolution using FUSION (McGaughey 2012). ArcGIS was then used to clip the CHM to each plot area (1 ha, 1/10th ha, 1/60th ha) and analyze these for "% tall cover." The "% tall cover" variable describes the proportion of the plot area with a canopy height over a given breakpoint. Breakpoints tested ranged from 26 to 38 m, at 2 m intervals (a total of seven metrics; see Appendix S1: Table S1). Fig. 3 shows the detailed work flow and illustrates how we used the LiDAR data to answer our key questions.

Statistical analysis

We used a combination of methods in the R software program (R Development Core Team 2008) to develop and evaluate linear regression models that estimated large diameter (>76 cm (30 in) dbh) tree density from LiDAR, without the use of tree segmentation. We chose the

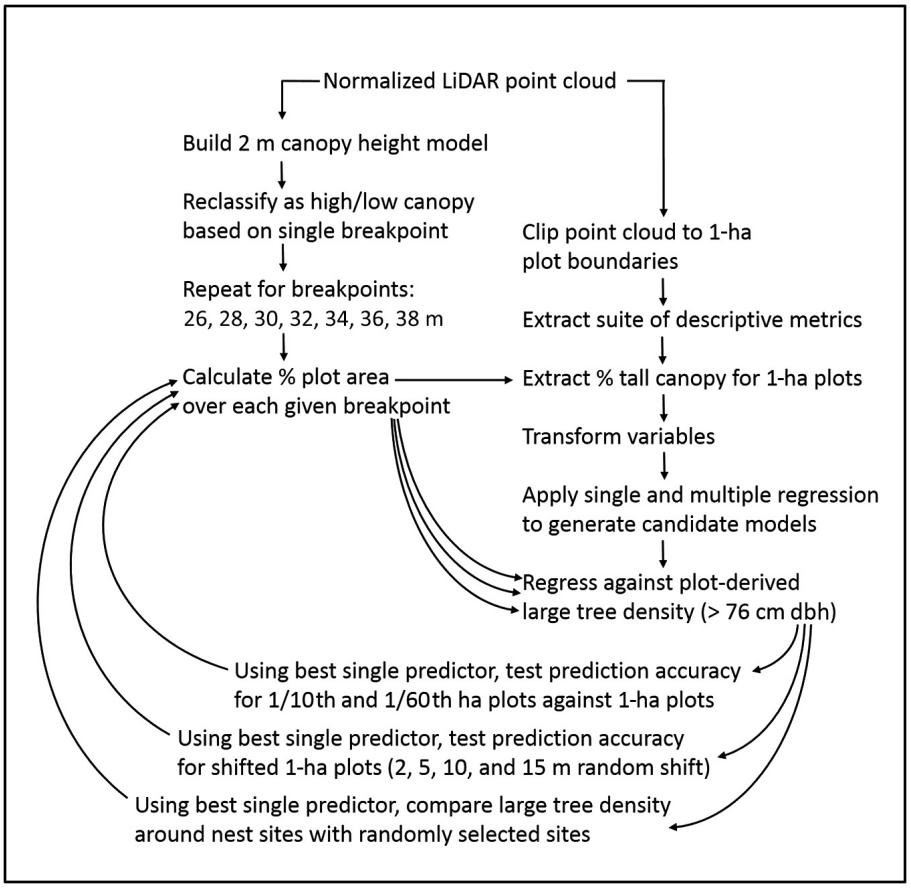

Fig. 3. Project work flow used to analyze light detection and ranging (LiDAR) and field data.

threshold of >76 cm (30 in) dbh to match the size threshold for trees collected in the largest, 1 ha plot size, as well as for both subplot sizes. This is also the maximum diameter limit guiding most forest management treatments specified in the 2001 Framework and the 2004 Sierra Nevada Forest Plan Amendment (US Forest Service 2001, 2004).

To avoid a non-normal response variable distribution according to the Shapiro–Wilk normality test (Shapiro and Wilk 1965), we transformed the field-based large tree count by taking its square root. We also used Q-Q plots to visually identify non-normal distributions of independent variables and bring their distributions closer to normality through transformation. A complete record of transformations is detailed in Appendix S1: Table S1. Applying these transformations increased the predictive ability of models and eliminated model heteroscedasticity, tested with the Breusch–Pagan test in the car package of R (Breusch and Pagan 1979, Fox et al. 2009).

Once variables were transformed, we calculated the best simple linear regression model (the model with the highest R^2 value), as well as the best multiple regression model. Because LiDAR-derived independent variables were highly correlated, an iterative model building approach was taken: (1) The best simple linear regression model was chosen based on the lowest Akaike's information criterion (AIC) score (Akaike 1987) using the leaps package in R (Lumley and Miller 2009). (2) Variables that were strongly correlated with the chosen independent variable (correlation >0.6 in either Pearson or Spearman correlations) were removed. (3) The process was repeated to find the next best independent variable. The final model was that with the lowest AIC and with all independent variables significant at $P < 0.05$. We recorded both AIC values and the 10-fold cross-validation error, calculated with the CVTools package in R (Alfons 2012).

We also derived predictive models for large tree density for the 1/10th ha (1012 m^2; 0.247 ac) and 1/60th ha (168 m^2, 0.041 ac) plots using the above methodology to test whether plot size influences the strength of the relationship. We wanted to evaluate whether the most reliable predictor changed and by how much the correlation coefficient degraded as plot size was reduced, especially as the plot size recommended for aerial LiDAR validation ranges between 300 and 600 m^2 (0.074 and 0.15 ac; Laes et al. 2011, Ruiz et al. 2014). Mascaro et al. (2011) examined the influence of plot size on model accuracy for predicting carbon density in a tropical forest and found that prediction accuracy scales with plot size due, in part, to decreased relative edge in larger plots. We predict that this relationship will be similar for estimating large tree density because in both cases LiDAR returns from primarily tree crowns are used to infer information about tree boles. Some of these factors are illustrated by Fig. 4, where two nearly identical plots have very different stem counts due to leaning stems, stems near the plot edge, and trees with uneven crowns.

Based on evaluation of model performance at the 1/60th ha, 1/10th ha, and 1-ha plot scales (see *Results*), we chose the simple linear regression model and the 1-ha plot for further analysis. This model was preferable as it had the highest adjusted R^2 value, uses the CHM (a commonly derived LiDAR product) as its base for prediction, and is simple enough to make logical sense for its predictions.

To address whether plot center accuracy influences the strength of this relationship, we shifted all 1-ha plots in a random direction (illustrated in

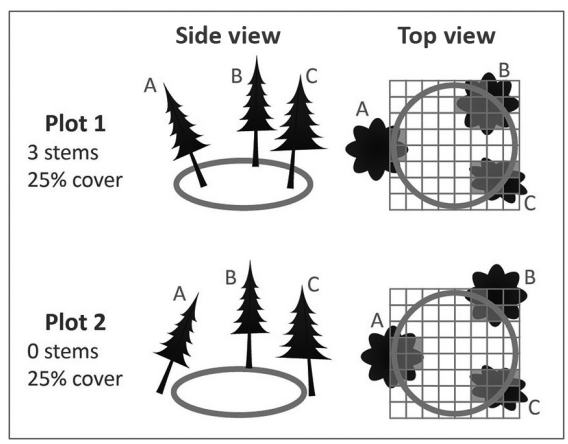

Fig. 4. Tree arrangements for two theoretical plots with 25% cover. Plot 1 contains three stems, while plot 2 contains none. Three tree characteristics that can lead to model inaccuracies, if they occur near the plot edge, are as follows: (A) stem lean, (B) stem near plot boundary, (C) uneven crown.

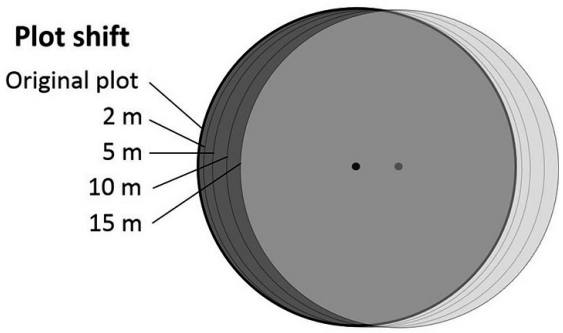

Fig. 5. Illustration of plot center shift of 2, 5, 10, and 15 m. The black dot and dark circle represent the original plot center and area, respectively, while the lighter dot and circle represent the location of the 15-m shifted plot. Other shifts are shown as outlines. At 15 m, this shift represents a highly inaccurate GPS point, yet with the 1 ha plot size, over 83% of the original plot is contained by the shifted plot.

Fig. 5) and recalculated the model coefficients, as well as the correlation coefficient. We repeated this shift and recalculation 100 times, and performed the analysis at shift distances of 2, 5, 10, and 15 m. For each shift distance, we compared the distribution of values for the model coefficients and correlation coefficient. We chose these distances because they are common values for the horizontal accuracy of many mid-range GPS units that are commonly used by a non-LiDAR-specific field crew when recording plot centers.

Large tree density centered on CSO nest trees was examined at multiple scales to evaluate the potential utility of using this variable to quantify habitat associations. Large tree density across the Meadow Valley study area was estimated and mapped using our single-regression model. Large tree density was extracted from this layer for areas within 50, 100, 500, and 1000 m of nest trees and at 100 randomly chosen points on the landscape for comparison. *T* tests were used to compare large tree density between owl and random sites at each spatial scale.

Results

Large diameter tree density

Due to high collinearity between independent variables (described in Appendix S1: Table S1), combined with these variables quickly becoming insignificant to the model at the $P < 0.05$ level, the multiple linear regression only contained two independent variables and was only slightly better able to predict large tree density than the single regression model, as shown in Table 1. Eq. 1 shows the best single regression linear model, where the independent variable was the percentage of the plot area where the CHM was over 32 m (CHM32). Eq. 2 shows the best multiple linear regression model, where the independent variables included CHM32 and the variance of point heights above 2 m (VAR).

$$\mathrm{Sqrt(trees/ha)} = 1.10 + 0.817 \quad \mathrm{sqrt(CHM32)} \tag{1}$$

$$\mathrm{Sqrt(trees/ha)} = 2.37 + 0.786 \quad \mathrm{sqrt(CHM32)} \quad 0.909 \times (\log(\mathrm{VAR}) + 1) \tag{2}$$

Both models had very similar AIC and cross-validated prediction error (Table 1). Both models had a cross-validation error below a single tree per ha, and an adjusted R^2 value of 0.77, which is accurate enough to be highly useful for managers.

Importance of plot size

Of the three plot sizes evaluated, the 1-ha plot was the best predictor of large tree density, with a model adjusted R^2 of 0.77 (Table 1). Model prediction accuracy decreased as plot size shrank. Even plots 1/10th ha in size (considered large by most managers and field crews) were poor predictors of large tree density, with the best model producing an adjusted R^2 of only 0.54 compared with 0.77 for the 1-ha plot. This decrease in

Table 1. Model statistics for the best single and multiple linear regression models based on the 1-ha plot size.

Variables included in model	AIC	10-fold cross-validation error†	Adjusted R^2
CHM32	347	0.88	0.77
CHM32, VAR	344	0.87	0.77

Notes: These include the predictor variables, AIC score, cross-validation error, and adjusted R^2 value for each model. CHM32 refers to the relative percentage of plot area where the CHM is over 32 m; VAR refers to variance among point heights. AIC, Akaike's information criterion; CHM, canopy height model.

† The cross-validation error is reported for the square root of trees/ha.

Table 2. Plot area with corresponding best linear model information for plot sizes <1 ha.

Plot type	Plot area	Explanatory variable	Adjusted R^2	P-value
Subplot	1/60th ha	COV32-34	0.21	<0.01
Macroplot	1/10th ha	CHM34	0.54	<0.01

Notes: COV32-34 represents the relative percent cover of all LiDAR returns between 32 and 34 m. CHM34 refers to the relative percentage of plot area where the CHM is over 34 m. CHM, canopy height model.

predictive accuracy based on plot size has been demonstrated by Mascaro et al. (2011). Even so, large tree density in both the 1-ha and 1/10th ha plots was best predicted by a CHM-derived variable. The best independent variable for each plot size, as well as coefficient of determination for each model for 1/10th and 1/60th ha plots are reported in Table 2. Note that while multiple regression was attempted, no variables beyond the first were significant.

To help explain the difference in model accuracy, Table 3 details the average number of large-diameter trees, as well as the proportional amount of edge to area, for the three plot sizes. These factors were likely both contributors to the poor predictive power of models built using smaller plot sizes. Table 3 shows that 1/60th and 1/10th ha plots had less than 0.5 and two large trees per plot, respectively, making prediction inherently difficult. Furthermore, the large ratio of plot edge to plot area likely contributed to model inaccuracy as well, with 1/60th ha plots having eight times as much relative edge than 1-ha plots (Table 3).

Importance of plot center accuracy

Plot centers were randomly shifted up to 15 m, but none of these shifts dramatically changed the single linear regression model accuracy or coefficient values. Boxplots displaying the model fit

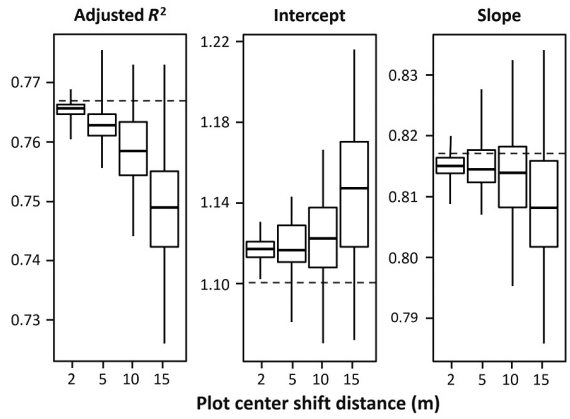

Fig. 6. The distribution of model values associated with predicting the large tree density using single linear regression (R^2, as well as the slope and intercept of the model), and how these changed on the 1-ha plot size as the coordinates of plot center were shifted 2, 5, 10, and 15 m in a random direction (n = 100). Original model values are shown as a gray dashed line.

and coefficients are shown in Fig. 6. Adjusted R^2 ranged between 0.72 and 0.78 for models built with shifted plot centers (the non-shifted model had an adjusted R^2 value of 0.77). Values for slope and intercept varied between 1.06 and 1.22, and 0.78 and 0.84, respectively (unshifted model values were 1.10 and 0.82, respectively).

Large trees around nest sites

Large diameter tree density was modeled for the entire study area and is shown with 13 CSO nest sites in Fig. 7. Smaller buffers around nest trees had a disproportionately high density of large trees, which dropped off gradually as the search radius from the nest tree increased, illustrated in Fig. 8. Within 50 m of nest trees, compared to random locations, large tree densities were significantly different, with 31 vs. 12 trees/ha, respectively. Area near nest sites remained

Table 3. Plot characteristics for the different plot sizes analyzed.

Plot area	Average large trees per plot	Radius (m)	Circumference (m)	Area (ha)	Edge:Area
1/60th ha	0.47	7.31	45	0.016	0.273
1/10th ha	1.74	17.95	112	0.101	0.111
1 ha	15.56	56.41	354	0.999	0.035

Notes: Large trees are >76 cm (30 in) in diameter at breast height. A lower edge to area ratio indicates less edge effect. Note that the 1/60th ha plot has almost eight times as much relative edge as the 1-ha plot.

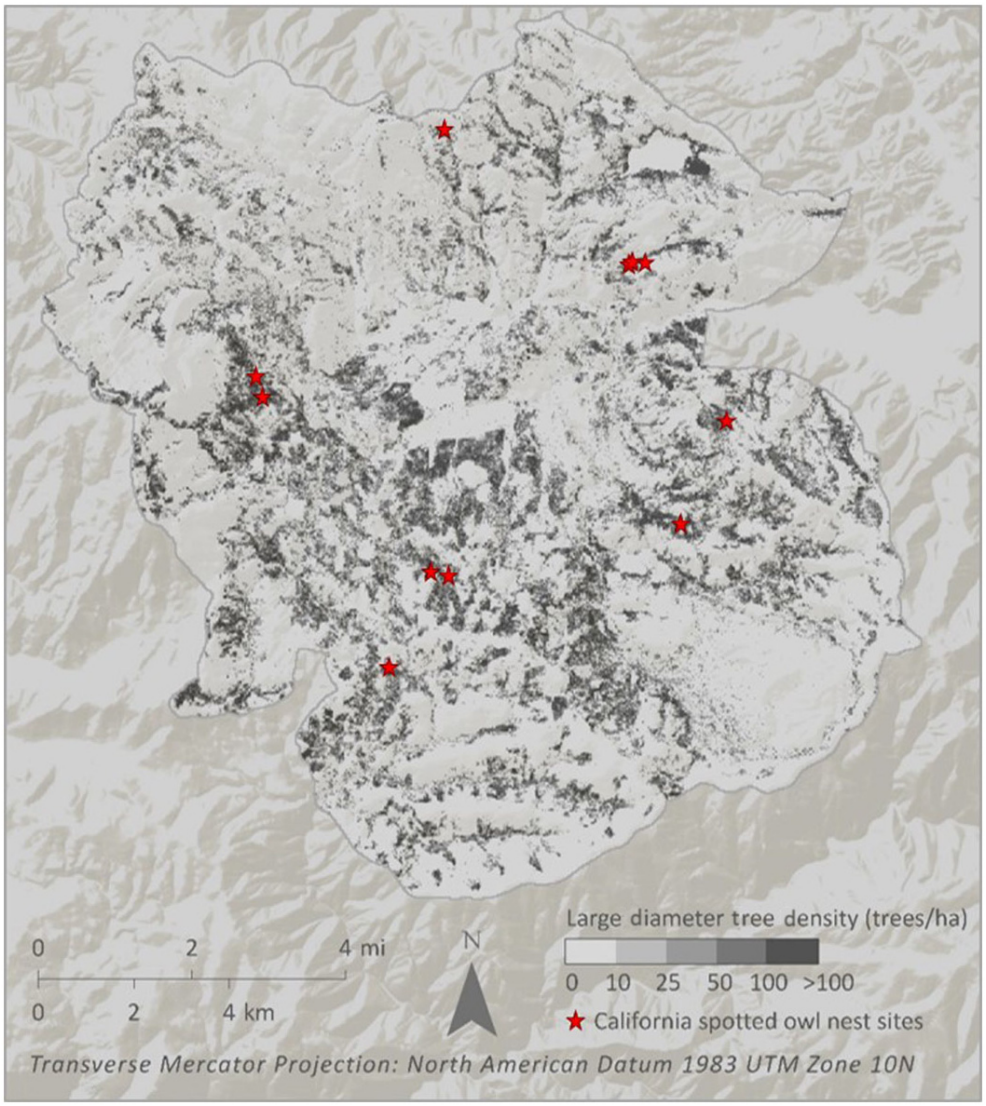

Fig. 7. Study area, showing large (over 76 cm (30 in) diameter at breast height) tree density and California spotted owl nest sites. Large tree density was derived from the canopy height model via single linear regression.

significantly different from random at distances of 50, 100, and 500 m. Only at a buffer distance of one km was large tree density no longer significantly different between CSO nest sites and random locations.

Discussion

Large trees are a critical habitat component for several wildlife species of concern and are presently lacking in many western forests relative to historical forest conditions (Franklin and Johnson 2012). In the case of the CSO, reductions in the number of large trees and structurally complex older forest stands due to past forest management and the more recent effects of wildland fire may be a contributing factor to current CSO population declines (Stephens et al. 2016*a*). Further, recent studies have documented high rates of large tree mortality due to interacting effects of drought, climate change, wildfire and insect activity (Lutz et al. 2009, Knapp et al. 2013, Dolanc et al. 2014). Thus, estimating large tree distribution and abundance is important for

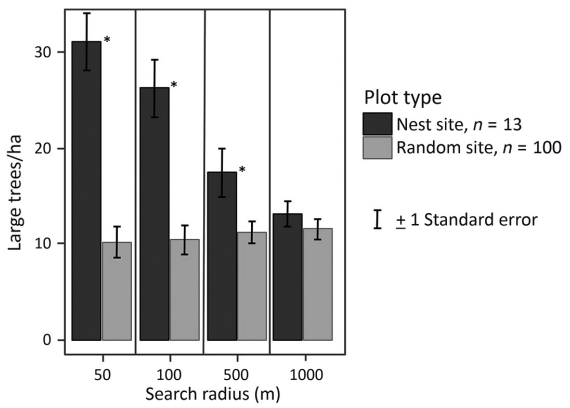

Fig. 8. Large (over 76 cm (30 in) diameter at breast height) tree density near 13 California spotted owl nest sites and 100 random sites over a range of search radii is shown. Plots with significantly different distributions (at $P < 0.01$) of large tree density are indicated by "*." Large tree density was derived from the canopy height model via single linear regression.

identifying and managing large tree habitat elements and older forest stands important to CSOs and other associated species. We show that aerial LiDAR can be successfully utilized to estimate the density of large trees, and that areas near CSO nest sites have a disproportionately high density of large trees, which corroborates the current understanding of preferred CSO habitat (Garcıa-Feced et al. 2011, Ackers et al. 2015). Both multiple and simple linear models accurately predicted the density of trees over 76 cm (30 in) dbh across 1-ha plot areas. However, the strength of this relationship decreased as plot size shrank to 1/10th ha and 1/60th ha, indicating that the plot size should be at least 1 ha when using the CHM to predict the density of large trees. Mascaro et al. (2011) found that for carbon density estimations, error associated with edge effect stabilized at a 1-ha plot size. We suspect that much of this difference in model accuracy was due to edge effect, illustrated by Fig. 4, and small sample size for smaller plots. Other factors that can decrease the accuracy of this estimate include trees that have non-standard crown:dbh relationships, including individuals with damaged crowns, sheared branches, or broken tops, and species with different ratios of crown area:stem diameter.

We also show that while shifting the placement of plot center (simulating decreased gps spatial accuracy of plot center) slightly decreases the accuracy of the large tree density model when 1-ha plots are used and marginally changes the model's coefficients, the model is still strong (adjusted R^2 never dropped below 0.72) with shifts in plot center of up to 15 m. This is likely due to the fact that with such a large plot size, much of the original plot area is included in the sampled LiDAR (83% of the original 1-ha plot is retained when the center is shifted 15 m). This indicates that for variables such as large tree density, which require large plots to accurately measure, a highly accurate (sub-meter) plot center may not be necessary. While most long-term plot networks do not utilize plot sizes as large as 1 ha, this study stands as a reminder that some datasets can still be useful to LiDAR validation, despite having less than ideal accuracy for plot centers.

Recently, many researchers have focused on producing a standardized LiDAR plot protocol (Laes et al. 2011, Ruiz et al. 2014), where recommended plot sizes range between 300 and 600 m^2 and nothing less than a mapping grade GPS is required. While this is an excellent step toward helping managers best prepare for maximizing the utility of new LiDAR acquisitions, it may also lead managers to assume that plots collected outside of these standards are useless for LiDAR analysis. This study suggests that these protocols might be less rigid for a variable such as large diameter tree density, where much larger plots are necessary to develop a robust predictive model. A similar relationship was found between plot size and prediction error by Mascaro et al. (2011), who show that relative prediction error decreases with increasing plot area for carbon density estimations in tropical forest. We suspect that in addition to large tree density and carbon density, this would also be the case for forest attributes such as basal area, where stem locations and characteristics are modeled from LiDAR returns that primarily represent the crowns of trees. We advise managers and researchers to critically examine the scale of the variable in which they are interested before deciding on an ideal plot design. This was an observational study based on one location, albeit a relatively large area, introducing the potential for locational bias in our analysis. Furthermore, because our study area is a single sample of the

landscape, results from this study should be carefully applied to other areas, and may require additional groundtruth analysis based on large plots from the new area. Regarding the plots themselves, placement was designed and implemented to sample CSO use areas and may not be representative of the entire landscape. Furthermore, plot center accuracy was not high by current standards (likely below 2 m, but possibly up to 5 m). However, our analysis showed that slight shifts in plot center did not influence results, so model inaccuracy from this source should be minimal.

Because this canopy-height-derived variable identifies the area of tall canopy, it assumes a link between canopy area and tree density, as well as a link between stem diameter and tree height. This means that our statistical model likely is unable to capture large trees when a significant portion of the tree tops are broken. However, to the extent that broken-top trees are still emergent in the dominant tree canopy or these large broken-top trees are often found near other large-diameter individuals without broken tops, our model would be expected to predict an accurate large tree density. Because tree species differ in the relationship between canopy height, canopy volume, and stem diameter, this model will need to be evaluated and perhaps recalibrated based on local knowledge of tree species and crown extents. However, for the purpose of a general prediction, this basic model performed surprisingly well, and is a simple and relatively accurate method for managers and researchers to evaluate large diameter stem density across the landscape in Sierra Nevada mixed conifer forests.

Future research

Additional studies to augment this research include carrying out similar analyses in different forest types. We suspect that these results will perform best in forests where the majority of large trees are coniferous, as these are identified by the CHM. Because of the variable quality of LiDAR available to land managers, investigation of the necessary point density to make accurate predictions is also essential.

Our work indicates that estimation of large tree density via LiDAR-derived CHM could be a useful method for identifying CSO nesting habitat. We encourage wildlife researchers to investigate the usefulness or improvement of this variable for modeling wildlife habitat for species associated with large tree habitat elements or forest stands with high densities of large trees, such as the CSO.

Our results show that older plots, which may not be ideal for traditional LiDAR-based derivations due to imprecise data or inaccurate plot center coordinates, could still be useful for other variables. We encourage others to explore their plot data and think critically about what can be compared to available LiDAR data.

Immediate implications for managers

Forest managers are challenged by the need to identify and manage large tree habitat across the Sierra Nevada. Information needs may range from identifying individual large trees that function as an important nesting/den habitat element within an area of generally younger, smaller forest, to identifying forest stands or patches with high densities of large trees across a landscape. We encourage thoughtful implementation of our methods to identify large trees and assess large tree density across landscapes. Such information on distribution and abundance of large trees can be used to identify areas of importance to associated wildlife species, such as the CSO, and to inform forest management options. Further, little to no information exists on large tree densities across the Sierra Nevada; thus, estimates of large tree density may be an important variable to incorporate into models of wildlife habitat. However, site productivity, dominant tree species composition, and management history are linked to the specific CHM threshold that is most appropriate, and we caution users to test multiple CHM cutoffs before finalizing their model. In other Sierra mixed conifer forests, we encourage managers with access to a LiDAR-derived CHM and a network of FIA plots that include the 1 ha plot size to derive a similar equation to predict the density of large trees. This would be a relatively simple project for anyone familiar with GIS and statistics, and could result in a highly useful layer for managers and wildlife ecologists. Large tree information could assist in the development of a long term management plan to conserve the CSO with an overall goal of increasing forest resilience to fires, insects, and drought (Stephens et al. 2016*b*).

Conclusions

Based on our method, managers can use the CHM, a common LiDAR deliverable, to accurately estimate large tree density, even when only lower density LiDAR (inappropriate for tree segmentation) is available. This can be accomplished without any specific LiDAR-processing hardware, software, or expertise, and does not require any LiDAR-specific plot protocol. This method demonstrates an excellent method for managers to put their LiDAR to practical use, although there are a few caveats.

We also show that older data traditionally labeled as "unusable" for LiDAR comparison, due to inaccurate plot center GPS coordinates, can provide valuable information for certain LiDAR-derived variables. For plots such as these to be successfully compared to LiDAR data, either (1) plots must be large enough to minimize GPS inaccuracy or (2) the variable must vary at a larger spatial scale than the potential inaccuracy of plot center.

Estimates of large tree density across the Sierra Nevada are lacking. Such information is needed to inform development of conservation and restoration strategies for CSOs and Sierra Nevada landscapes. Our methods provide an approach for generating this information in areas where LiDAR data are available.

Acknowledgments

We thank the Plumas National Forest for allowing us to use their LiDAR dataset. Thanks also to Robert McGaughey for assisting us with the LiDAR-processing program, FUSION. Much thanks to Martin Isenburg and LAStools, as well as the LASmoons grant, for use of the software program, LAStools. Marek Jakubowski contributed valuable advice on LiDAR-processing methodology. Joan Canfield and anonymous reviewers contributed comments that improved this manuscript. This material is based upon work supported by the Pacific Southwest Research Station, USDA Forest Service, and the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1106400. H. Kramer led the study. H. Kramer, B. Collins, S. Stephens, and M. Kelly designed the study. J. Keane and C. Gallagher designed, collected, and compiled plot and CSO data. H. Kramer led the data processing, analysis, and interpretation, with assistance from B. Collins, S. Stephens, and M. Kelly, J. Keane, and C. Gallagher. M. Kelly contributed expertise on LiDAR and spatial analysis. B. Collins and S. Stephens provided expertise in forest structure and management. J. Keane and C. Gallagher supplied insight on the CSO.

Literature Cited

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

Akaike, H. 1987. Factor analysis and AIC. Psychometrika 52:317–332.

Alfons, A. 2012. cvTools: cross-validation tools for regression models. R Package Version 0.3.2. https://cran.r-project.org/web/packages/cvTools/index.html

Ansley, J. S., and J. J. Battles. 1998. Forest composition, structure, and change in an old growth mixed conifer forest in the northern Sierra Nevada. Journal of the Torrey Botanical Society 125:297–308.

Barbour, M. G., and J. Major. 1995. Terrestrial vegetation of California: new expanded edition. California Native Plant Society, Davis, California, USA.

Beier, P., and J. E. Drennan. 1997. Forest structure and prey abundance in foraging areas of northern goshawks. Ecological Applications 7:564–571.

Bergen, K., S. Goetz, R. Dubayah, G. Henebry, C. Hunsaker, M. Imhoff, R. Nelson, G. Parker, and V. Radeloff. 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: review and implications for LiDAR and radar spaceborne missions. Journal of Geophysical Research: Biogeosciences 114:1–13.

Bias, M. A., and R. Gutierrez. 1992. Habitat associations of California spotted owls in the central Sierra Nevada. Journal of Wildlife Management 56:584–595.

Blakesley, J. A., M. E. Seamans, M. M. Conner, A. B. Franklin, G. C. White, R. Gutierrez, J. E. Hines, J. D. Nichols, T. E. Munton, and D. W. Shaw. 2010. Population dynamics of spotted owls in the Sierra Nevada, California. Wildlife Monographs 174: 1–36.

Breusch, T. S., and A. R. Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47:1287–1294.

Chen, Q., D. Baldocchi, P. Gong, and M. Kelly. 2006. Isolating individual trees in a savanna woodland using small footprint LiDAR data. Photogrammetric Engineering and Remote Sensing 72:923–932.

Collins, B. M., H. A. Kramer, K. Menning, C. Dillingham, D. Saah, P. A. Stine, and S. L. Stephens. 2013. Modeling hazardous fire potential within a completed fuel treatment network in the northern Sierra Nevada. Forest Ecology and Management 310:156–166.

Dolanc, C. R., H. D. Safford, J. H. Thorne, and S. Z. Dobrowski. 2014. Changing forest structure across the landscape of the Sierra Nevada, CA, USA, since the 1930s. Ecosphere 5:1–26.

Forsman, E. 1995. Appendix A: standardized protocols for gathering data on occupancy and reproduction in spotted owl demographic studies 1999. Pages 32–38 *in* J. Lint, B. Noon, R. Anthony, E. Forsman, M. Raphael, M. Collopy, and E. Starkey, editors. Northern spotted owl effectiveness monitoring plan for the Northwest Forest Plan. PNW-GTR-440. USDA Forest Service, Pacific Northwest Research Station, Corvallis, Oregon, USA.

Fox, J., D. Bates, D. Firth, M. Friendly, G. Gorjanc, S. Graves, R. Heiberger, G. Monette, H. Nilsson, and D. Ogle. 2009. CAR: companion to applied regression. R Package Version 1.2-16. http://cran.r-project.org/web/packages/car/index.html

Franklin, J. F., and K. N. Johnson. 2012. A restoration framework for federal forests in the Pacific Northwest. Journal of Forestry 110:429–439.

Franklin, J. F., T. A. Spies, R.Van Pelt, A. B. Carey, D. A. Thornburgh, D. R. Berg, D. B. Lindenmayer, M. E. Harmon, W. S. Keeton, and D. C. Shaw. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. Forest Ecology and Management 155:399–423.

Gallagher, C. V. 2010. Spotted owl home range and foraging patterns following fuels-reduction treatments in the northern Sierra Nevada, California. Thesis. University of California, Davis, California, USA.

Garcıa-Feced, C., D. J. Tempel, and M. Kelly. 2011. LiDAR as a tool to characterize wildlife habitat: California spotted owl nesting habitat as an example. Journal of Forestry 109:436–443.

Greenwald, D. N., D. C. Crocker-Bedford, L. Broberg, K. F. Suckling, and T. Tibbitts. 2005. A review of northern goshawk habitat selection in the home range and implications for forest management in the western United States. Wildlife Society Bulletin 33:120–128.

Gutierrez, R., K. McKelvey, B. Noon, G. Steger, D. Call, W. LaHaye, B. Bingham, and J. Senser. 1992. Habitat relations of the California spotted owl. Pages 79–98 *in* J. Verner, K. S. McKelvey, B. R. Noon, R. J. Gutierrez, G. I. Gould Jr., and T. W. Beck [tech. coordinators]. The California spotted owl: a technical assessment of its current status. PSW-GTR-133. USDA Forest Service, Pacific Southwest Research Station, Albany, California, USA.

Herger, W., and D. Feinstein. 1998. Herger-Feinstein Quincy Library Group Forest Recovery Act. Department of the Interior and Related Agencies Appropriations Act, Section 401. U.S. Congress, Washington, D.C., USA.

Hollenbeck, J. P., V. A. Saab, and R. W. Frenzel. 2011. Habitat suitability and nest survival of white-headed woodpeckers in unburned forests of Oregon. Journal of Wildlife Management 75:1061–1071.

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

Hunter, J. E., R. J. Gutierrez, and A. B. Franklin. 1995. Habitat configuration around spotted owl sites in northwestern California. Condor 97:684–693.

Isenburg, M. 2011. LAStools—efficient tools for LiDAR processing (Version 150925, academic). https://rapidlasso.com/lastools/

Jakubowski, M. K., Q. Guo, and M. Kelly. 2013*b*. Tradeoffs between LiDAR pulse density and forest measurement accuracy. Remote Sensing of Environment 130:245–253.

Jakubowski, M., W. Li, Q. Guo, and M. Kelly. 2013*a*. Delineating individual trees from LiDAR data: a comparison of vector- and raster-based segmentation approaches. Remote Sensing 5:4163–4186.

Kaartinen, H., J. Hyyppä, X. Yu, M. Vastaranta, H. Hyyppä, A. Kukko, M. Holopainen, C. Heipke, M. Hirschmugl, and F. Morsdorf. 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sensing 4:950–974.

Keane, J. J. 2014. California spotted owl: scientific considerations for forest planning. Pages 437–467 *in* J. W. Long, L. Quinn-Davidson, and C. N. Skinner, editors. Science synthesis to support socioecological resilience in the Sierra Nevada and southern Cascade Range. PSW-GTR-247. Chapter 7.2. USDA Forest Service, Pacific Southwest Research Station, Albany, California, USA.

Kelly, M., and S. Di Tommaso. 2015. Mapping forests with LiDAR provides flexible, accurate data with many uses. California Agriculture 69:14–20.

Kenward, R. E. 2001. A manual for wildlife radio tagging. Academic Press, London, UK.

Knapp, E. E., C. N. Skinner, M. P. North, and B. L. Estes. 2013. Long-term overstory and understory change following logging and fire exclusion in a

Sierra Nevada mixed-conifer forest. Forest Ecology and Management 310:903–914.

Kramer, H. A., B. M. Collins, M. Kelly, and S. L. Stephens. 2014. Quantifying ladder fuels: a new approach using LiDAR. Forests 5:1432–1453.

Kramer, H. A., B. M. Collins, F. K. Lake, M. K. Jakubowski, S. L. Stephens, and M. Kelly. 2016. Estimating ladder fuels: a new approach combining field photography with LiDAR. Remote Sensing 8:766.

Laes, D., S. Reutebuch, R. McGaughey, and B. Mitchell. 2011. Guidelines to estimate forest inventory parameters from LiDAR and field plot, companion document to the advanced LiDAR applications—forest inventory modeling class. Remote Sensing Applications Center, Salt Lake City, Utah, USA.

Lee, A. C., and R. M. Lucas. 2007. A LiDAR-derived canopy density model for tree stem and crown mapping in Australian forests. Remote Sensing of Environment 111:493–518.

Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding. 2002. LiDAR remote sensing for ecosystem studies. BioScience 52:19–30.

Li, W., Q. Guo, M. K. Jakubowski, and M. Kelly. 2012. A new method for segmenting individual trees from the LiDAR point cloud. Photogrammetric Engineering and Remote Sensing 78:75–84.

Lumley, T., and A. Miller. 2009. Leaps: regression subset selection. R Package Version 2.9. http://CRAN.R-project.org/package=leaps

Lutz, J. A., A. J. Larson, M. E. Swanson, and J. A. Freund. 2012. Ecological importance of large-diameter trees in a temperate mixed-conifer forest. PLoS ONE 7:1–15.

Lutz, J., J. Van Wagtendonk, and J. Franklin. 2009. Twentieth-century decline of large-diameter trees in Yosemite National Park, California, USA. Forest Ecology and Management 257:2296–2307.

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

Mascaro, J., M. Detto, G. P. Asner, and H. C. Muller-Landau. 2011. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. Remote Sensing of Environment 115:3770–3774.

McDermid, G. J., S. E. Franklin, and E. F. LeDrew. 2005. Remote sensing for large-area habitat mapping. Progress in Physical Geography 29:449–474.

McGaughey, R. 2012. FUSION/LDV: software for LiDAR data analysis and visualization, version 3.01. USDA Forest Service, Pacific Northwest Research Station, University of Washington, Seattle, Washington, USA.

Moen, C. A., and R. Gutierrez. 1997. California spotted owl habitat selection in the central Sierra Nevada. Journal of Wildlife Management 61:1281–1287.

Moghaddas, J. J., B. M. Collins, K. Menning, E. E. Y. Moghaddas, and S. L. Stephens. 2010. Fuel treatment effects on modeled landscape-level fire behavior in the northern Sierra Nevada. Canadian Journal of Forest Research 40:1751–1765.

Moody, T. J., J. Fites-Kaufman, and S. L. Stephens. 2006. Fire history and climate influences from forests in the northern Sierra Nevada, USA. Fire Ecology 2:115–141.

Popescu, S. C., R. H. Wynne, and R. F. Nelson. 2003. Measuring individual tree crown diameter with LiDAR and assessing its influence on estimating forest volume and biomass. Canadian Journal of Remote Sensing 29:564–577.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Ruiz, L. A., T. Hermosilla, F. Mauro, and M. Godino. 2014. Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates. Forests 5:936–951.

Schoenherr, A. A. 1992. A natural history of California. University of California Press, Berkeley, California, USA.

Seavy, N. E., J. H. Viers, and J. K. Wood. 2009. Riparian bird response to vegetation structure: a multiscale analysis using LiDAR measurements of canopy height. Ecological Applications 19:1848–1857.

Selvarajan, S., A. Mohamed, and T. White. 2009. Assessment of geospatial technologies for natural resource management in Florida. Journal of Forestry 107:242–249.

Shapiro, S. S., and M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). Biometrika 52:591–611.

Stephens, S. L., S. W. Bigelow, R. D. Burnett, B. M. Collins, C. V. Gallagher, J. Keane, D. A. Kelt, M. P. North, L. J. Roberts, and P. A. Stine. 2014. California spotted owl, songbird, and small mammal responses to landscape fuel treatments. BioScience 64:893–906.

Stephens, S. L., J. D. Miller, B. M. Collins, M. P. North, J. J. Keane, and S. L. Roberts. 2016*a*. Wildfire impacts on California spotted owl nesting habitat in the Sierra Nevada. Ecosphere 7:e01478.

Stephens, S. L., B. M. Collins, E. Biber, and P. Z. Fule. 2016*b*. U.S. federal fire and forest policy: emphasizing resilience in dry forests. Ecosphere 7: e01584.

Swetnam, T. L., and D. A. Falk. 2014. Application of metabolic scaling theory to reduce error in local

maxima tree segmentation from aerial LiDAR. Forest Ecology and Management 323:158–167.

US Forest Service. 2001. Sierra Nevada Forest Plan Amendment: Final Environmental Impact Statement. Volumes 1–6. USDA Forest Service Pacific Southwest Region, Vallejo, California, USA.

US Forest Service. 2004. Sierra Nevada Forest Plan Amendment, Final Supplemental Environmental Impact Statement. US Department of Agriculture Forest Service Region 5, Vallejo, California, USA.

USDA Forest Service. 2001. Forest Inventory and Analysis national core field guide, volume I: field data collection procedures for phase 2 plots, version 1.5. US Department of Agriculture, Forest Service, Washington, D.C., USA.

Vierling, K. T., L. A. Vierling, W. A. Gould, S. Martinuzzi, and R. M. Clawges. 2008. LiDAR: shedding new light on habitat characterization and modeling. Frontiers in Ecology and the Environment 6:90–98.

White, G., and R. Garrott. 1990. Analysis of wildlife radio-tracking data. Academic Press, San Diego, California, USA.

Wing, M. G., A. Eklund, and J. Sessions. 2010. Applying LiDAR technology for tree measurements in burned landscapes. International Journal of Wildland Fire 19:104–114.

Zielinski, W. J. 2014. The forest carnivores: marten and fisher. Pages 393–435 *in* J. W. Long, L. Quinn-Davidson, and C. N. Skinner, editors. Science synthesis to support socioecological resilience in the Sierra Nevada and southern Cascade Range. PSW-GTR-247. Chapter 7.1. USDA Forest Service, Pacific Southwest Research Station, Albany, California, USA.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1593/full