Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data.

Txomin Hermosilla *[a], Luis A. Ruiz [a], Alexandra N. Kazakova [b], Nicholas C. Coops [c], L. Monika Moskal [b]

* Corresponding author: txohergo@topo.upv.es

[a] Geo-Environmental Cartography and Remote Sensing Group, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[b] Remote Sensing and Geospatial Analysis Laboratory and Precision Forestry Cooperative, School of Forest Resources, College of the Environment, University of Washington, Seattle WA 98195-2100, United States

[c] Integrated Remote Sensing Studio, Department of Forest Resources Management, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4

Abstract

Precise knowledge of fuel conditions is important to predict fire hazards and to simulate fire growth and intensity across the landscape. We present a methodology to retrieve and map forest canopy fuel and other forest structural parameters using small-footprint full-waveform airborne Light Detection and Ranging (LiDAR) data. Full-waveform LiDAR sensors register the complete returned backscattered signal through time, and can describe physical properties of the intercepted objects. This study is undertaken in a mixed forest dominated by Douglas-fir, occasionally mixed with other conifers in northwest Oregon (United States). We extracted two sets of LiDAR metrics using pulse detection and waveform modeling and then constructed a number of predictive models using forward stepwise multiple linear regression. The resulting models explained approximately 80% of the variability for many of the canopy fuel and forest structure parameters: aboveground biomass (R^2=0.84), quadratic mean diameter (R^2=0.82), canopy height (R^2=0.79), canopy base height (R^2=0.78), and canopy fuel load (R^2=0.79). The lowest performing models included basal area (R^2=0.76), stand volume (R^2=0.73), canopy bulk density (R^2=0.67), and stand density index (R^2=0.66). Our results indicate that full-waveform LiDAR systems has promise to systematically characterize the structure and canopy fuel loads of forests, which may enable accurate fire behavior forecasting that in turn supports the development of prevention and planning policies.

Short summary

Knowledge of fuel conditions is critical to accurately forecast fire behavior. We present a methodology to estimate forest canopy fuel parameters using small-footprint, full-waveform, Light Detection and Ranging (LiDAR) airborne data in a mixed forest region of northwest Oregon (US).

1. Introduction

Wildland fire is a principal disturbance that influences and changes vegetation composition, structure and function by selectively favoring certain species and creating conditions for new species to invade. As a result, fire shapes the landscape mosaic and influences biogeochemical cycles such as carbon and nutrient cycles (Mooney et al., 1981; Agee, 1996; Flannigan et al 2000). Fire behavior depends on weather, topography and fuels. When the specific characteristics of each of these factors are known, the behavior of fire can often be predicted (Agee, 1996). As a result, a number of mathematical models have been developed to predict, quantify and map fire spread, such as FARSITE (Finney, 2004) and BehavePlus (Andrews, 2009).

An accurate knowledge of fuel conditions is key to predicting spatial fire hazard and to simulate fire growth and intensity across the landscape (Keane et al., 2001), as well as help develop prevention and planning strategies, as fuel constitutes a primary component of fire risk. Of the required characteristics for these mathematical models, fuel is also arguably the most complex, as it depends on the physical characteristics of both living and dead biomass, particularity in the crown. This in turn contributes to the spread, intensity and severity of wildland fire (Anderson 1982). Canopy fuel load, canopy height, canopy bulk density, and canopy base height are all characteristics of forest fuel and control crown fire spread and are thus common inputs into fire behavior models (Keane et al., 2001; Cruz et al., 2003). Canopy fuel load is the amount of fuel in the canopy layer potentially available for combustion. Canopy height influences wind trajectory and wind speed reduction (Finney, 2004) which affects the distance that embers can be lofted (Chuvieco et al. 2003). Canopy bulk density quantifies the fuel in the canopy layer per unit of volume and affects the rate of spread of fire from tree to tree (Chuvieco et al. 2003; Scott and Reinhardt, 2001). Canopy base height is the vertical distance between the ground and live canopy fuel layers and it is critical for determining if ground surface fires can reach tree crowns (Scott and Reinhardt, 2001; Cruz et al., 2003). In addition to canopy characteristics influencing the rate and spread of fire through the tree crowns, the forest canopy also plays a critical role in the type, amount, and distribution of understory vegetation which is also an important field component (Cruz et al., 2003).

Canopy fuel attributes, however, are difficult to estimate and map and are most often acquired through detailed field programs, undertaken by fire professionals, requiring significant field effort (Arroyo et al 2008). Due to the

66 difficulty and cost of acquiring these intensive field measurements, remote sensing is increasingly being applied to
67 retrieve data for these models (Chuvieco and Salas, 1996; Reich et al., 2004). Initially, optical remote sensing
68 imagery was the first technology employed for wildland fire hazard mapping (Chuvieco and Congalton; 1989) yet,
69 although this technology enables discrimination of fuel loads and types at broad spatial scales (Wilson et al., 1994;
70 Riaño et al., 2002; Chuvieco et al, 2002; Chuvieco et al, 2004; Falkowski et al., 2005), it is limited in its ability to
71 retrieve information about the vertical structure and distribution of the vegetation due to canopy obstruction (Keane
72 et al., 2001).
73
74 LiDAR (Light Detection And Ranging) offers significant new opportunities to map vegetation structure. LiDAR
75 acquires detailed three-dimensional data using a series of laser pulses, and their subsequent return from the surfaces
76 they strike (Hippenstiel and Brownson, 2012). In addition to estimating vegetation height, LiDAR enables a much
77 more complete description of the vertical structure of the vegetation (Wagner et al., 2008), and it has been
78 successfully used in a range of forest applications (van Leeuwen and Nieuwenhuis, 2010). The most common type
79 of LiDAR systems available to resource managers are small-footprint discrete systems, which digitize the return
80 pulse into a small number of three dimensional coordinates usually coinciding with the return of the first and last
81 energy components and some intermediate energy peaks (Hall et al., 2005). The past decade has seen an increased
82 use of discrete return data for applications such as forest structural variable estimation (Koetz et al., 2006, Kim et
83 al., 2009; Sumnall et al., 2012) and tree species classification (Reitberger et al., 2008; Neuenschwander et al, 2009;
84 Heinzel and Koch 2011). The small (0.2-3 m) footprint provided by these LiDAR systems makes the data highly
85 suitable to predicting and mapping fuel map attributes for fine scale simulations of fire growth and behavior (Keane
86 et al., 2001). As a result of this, LiDAR has been used to retrieve a range of forest fuel metrics such as canopy fuel
87 weight, canopy base height, canopy bulk density, canopy height, and crown dimension (Riaño et al., 2003, 2004,
88 2007; Morsdorf et al., 2004; Andersen et al., 2005; Skowronski et al, 2007; Erdody and Moskal, 2010; Skowronski
89 et al, 2011, Zhao et al., 2011; Peterson and Nelson 2011; González-Olabarria et al., 2012).
90
91 Since discrete LiDAR returns are unable to provide complete information along the full path traveled by the emitted
92 pulse, discrete LiDAR has some limitations when characterizing the structure of the canopy and its different
93 vegetation layers. In contrast, full-waveform LiDAR sensors are able to register the complete returned backscattered

94 signal through time. The analysis of the returned waveform should enable researchers to more fully describe the
95 physical properties of the intercepted objects since the amplitude of the waveform at any height is proportional to
96 the amount of reflective material intercepted at that height, the orientation of that material, and its reflectance (Hyde
97 et al., 2005). In this paper, we propose and evaluate a methodology to use small-footprint full-waveform LiDAR-
98 derived metrics to retrieve forest canopy fuel parameters using data collected over a Pacific Northwest conifer
99 forest.
100
101 **2. Study area and data**
102 Data acquired for this research came from the Panther Creek Cooperative Research Project (Flewelling and
103 McFadden, 2011), which is intended to develop a suite of LiDAR applications for forest managers and currently
104 involves over forty researchers and land managers representing federal, state and local agencies, landowners, a
105 LiDAR provider, universities, and consultants.
106
107 **2.1. Study area**
108 Panther Creek is a 2,258 ha forested watershed located in the east side of the coastal mountain range in Yamhill
109 County, Oregon (Figure 1). Elevation ranges from 100 to 700 m (see study area topography in Figure 2.b). The
110 dominant species is Douglas-fir (*Pseudotsuga menziesii*), which covers more than half of the total forested area,
111 occasionally mixed with other conifers, such as western hemlock (*Tsuga heterophylla*), western red cedar (*Thuja*
112 *plicata*) and grand fir (*Abies grandis*), or deciduous species such as red alder (*Alnus rubra*) and bigleaf maple (*Acer*
113 *macrophyllum*). see Figure 2.c). Other shrub vegetation can also be found in patches (dogwood, cascara, etc.). Tree
114 heights are up to 60 m. The ecoregion classification is "*Cascade mixed forest*". Management intensity throughout
115 the study area has been highly variable, with different planting densities, and both thinned and unthinned regimes.
116
117 **2.1. LiDAR data**
118 Full waveform LiDAR data were collected on July 15, 2010 by Watershed Sciences, Inc. using a Leica ALS60
119 sensor mounted in a Cessna Caravan 208B. The system acquired data at a 105 kHz pulse rate, flown at an average of
120 900 m above ground level, with a scanning angle of $\pm 14^{o}$ from nadir. The returned waveforms were recorded in 256
121 bins with a temporal sample spacing of 2 ns, and a beam footprint size of approximately 0.25 m. This configuration

yielded a pulse density of ≥8 points/m^2. The study area was surveyed with opposing flight line side-lap of ≥50% (≥100% overlap) to reduce laser shadowing and increase surface laser painting. Aircraft position was recorded with a frequency of 2 Hz by an onboard differential GPS unit. Aircraft altitude was measured 200 times per second (200 Hz) as pitch, roll, and yaw from an onboard inertial measurement unit (IMU). LiDAR data were distributed in LAS 1.3 format (ASPRS, 2010). In addition to the full-waveform data a Digital Terrain Model (DTM) was provided by Watershed Sciences, Inc. based on last return pulses, with a documented Root-Mean-Square Error (RMSE) error using 33 GPS ground control points of 0.19 m (see DTM in Figure 2.b).

2.2. Field data

A total of 84 circular plots with 16 m radius were located within the study site. Plots were located on Bureau of Land Management or City of McMinnville land or on land held by other large landholders (see Figure 2.a). A significant fraction of the plots were located to sample a wide range of canopy depths. Plot positions were determined to an accuracy lower than 0.3 m in horizontal and vertical locations using Trimble R-8 GNSS Receivers and Leica TPS 800 total stations. All trees (conifer and hardwood stems) with diameter at breast height (DBH) of 2.5 cm or greater were identified, numbered and tagged (Table 1).

From the field data two groups of plot-level parameters were computed. The first group describes the basic forest inventory and structural properties of the forest: aboveground biomass, basal area, quadratic mean diameter (QMD), stand density index (Reineke, 1933), and Volume. Aboveground biomass was estimated using allometric equations based on DBH and tree height. Species specific allometric equations derived by Standish et al. (1985) were used, with biomass computed for each individual component and summed (Gholz et al.1979).

The second set of derived parameters was related to canopy fuels and consisted of canopy height, canopy base height, canopy bulk density, and canopy fuel load. Canopy bulk density was computed assuming a uniform vertical distribution of fuels by dividing the canopy fuel load by the canopy depth (Reinhardt and Crookston, 2003). Canopy depth was estimated as the mean crown length over all trees on the plot (Cruz et al. 2003), where crown length is the difference between tree height and height to live crown. Canopy fuel load was computed as the total amount of

biomass in the canopy fuel layer per unit surface area. Table 2 summarizes the forest structural and canopy fuel parameters of the 84 plots in the Panther Creek study area.

3. Methods

While small-footprint full waveform airborne LiDAR systems are relatively new, large-footprint full-waveform LiDAR sensors, such as the airborne SLICER and LVIS or the spaceborne GLAS instrument, have been collecting data for over a decade. These data have been applied to vegetation (Lefsky et al., 1999, Drake et al., 2002; Harding and Carabajal 2005) as well as to some canopy fuel studies (Ashworth et al 2010, Garcia et al., 2012). Although these large footprint sensors are limited by their footprint size (10-70 m), which prevents detailed forest structure to be retrieved (Listopad et al., 2011), analysis of data acquired from these sensors has resulted in a number of approaches being developed for processing full-waveforms and deriving descriptive metrics. Two basic approaches are pulse detection and waveform modeling (Mallet and Bretar, 2009). Pulse detection extracts a large number of echoes from the waveforms themselves, creating a very dense point cloud. The second approach, denoted waveform modeling, involves undertaking a spatio-temporal analysis of the return waveform, which enables extraction of detailed geometric and radiometric information from each waveform. In our methodology, full-waveform data are initially pre-processed using both approaches. After pre-processing, a set of descriptive metrics is derived, which are used as input variables for the predictive models of the forest structure and canopy fuel parameters. Finally, the accuracy of these models is assessed, and maps representing the canopy fuel properties of the study area are produced (see methodology scheme in Figure 3).

3.1. Data pre-processing

An initial noise assessment was performed to suppress background noise within each waveform. Waveforms were then smoothed and any remaining noise removed using a Gaussian filter, with a kernel size defined by the Full Width at Half Maximum (Duong, 2010). Once the noise was suppressed, a local maxima peak detection filter was applied to each waveform to obtain the point cloud retrieving the height and intensity from the intensity peaks, which are produced when intercepting reflective material such as canopy (Vaughn et al., 2012). The DTM was used to define the ground and to normalize the point cloud, deducting the influence of the terrain from the height of the points.

The waveform modeling based metrics were derived from a transformed dataset consisting of synthetic composite waveforms that simulate the vertical profile of vegetation for a given location. This methodology, similar to the applied by Buddenbaum et al. (2013), aims to integrate the non-vertical waveforms registered from different flight trajectories, and partitions the vertical aboveground space into regular voxels (0.25 × 0.25 × 0.30 m) corresponding to the approximate footprint size and the distance traveled by the pulse in a single temporal sample. First, each waveform record was spatially located using the information contained in the raw LiDAR file files, and then each waveform record was normalized by computing the difference between its height and the terrain height at that position. Each intersected voxel, by a given waveform track, was then filled with the maximum waveform amplitude value registered within it. The maximum value was used in order to preserve the significance of the major returns of the real waveforms in the composited waveforms. The composite waveforms were retrieved by extracting the information contained in the vertical column of voxels over a specific location.

3.2. Descriptive metrics

3.2.1. Pulse detection based metrics

From the normalized point cloud, obtained from the waveform peak location, a set of per-plot statistical descriptors of the height distribution were calculated: mean, standard deviation, range, kurtosis, and skewness, including the 25^{th}, 50^{th}, 75^{th}, and 100^{th} height percentiles. In addition, an analogous set of descriptors was computed from the intensity values, which are related to spectral reflectance of the observed target (Ahokas et al., 2006). A 1 × 1 m canopy elevation model was also produced from the maximum height value. Then, the mean, standard deviation, range, kurtosis, and skewness were extracted for each plot. A summary of the variables used in the analysis is shown in Table 3.

3.2.3. Waveform modeling based metrics

In addition to the pulse based LiDAR metrics, a number of additional descriptors have been developed to retrieve forest structural parameters from the analysis of the composite LiDAR waveforms (see the comprehensive compilation performed by Duong; 2010)). Metrics derived included: height of median energy (HOME), waveform distance, height to median ratio, number of peaks, roughness of outermost canopy, front slope angle, return

205 waveform energy, and vertical distribution ratio (VDR). Several of these metrics were initially developed using
206 large-footprint waveform sensors. However, the use of the metrics at finer scales provided by small-footprint data
207 has been shown to be successful, and most of the variables can readily be used to describe within crown – rather
208 than within stand canopy – variations (Neuenschwander et al, 2009, Heinzel and Koch 2011; Höfle et al., 2012;
209 Neuenschwander 2012; GuangCai et al., 2012).HOME is defined as the distance from the ground to the waveform
210 centroid, which is the position where the return energy above the DTM is divided into two equal parts. Previous
211 research has shown that HOME describes the vertical arrangement of canopy elements and the degree of canopy
212 openness (Drake et al., 2002). Waveform distance is the distance from the waveform begin to the ground and it is
213 usually related with the tree height. Height to median ratio is computed as HOME divided by waveform distance.
214 This metric is sensitive to changes in the canopy (Drake et al., 2002). The number of peaks detected is related to the
215 number of height levels intersected by the waveform. The roughness of outermost canopy is defined as the distance
216 from the waveform beginning to the first peak, which is usually considered as the peak of canopy. This metric
217 describes the spatial organization of plant surfaces within the canopy. Front slope angle is the vertical angle of the
218 vector from waveform begin and the first peak, and it is related to the variability of the upper canopy (Ranson et al.,
219 2004). Return waveform energy represents the total received energy and describes the surface characteristics. The
220 plots were characterized by the mean and standard deviation of these metrics. Finally, VDR is computed as the
221 difference between the canopy height and HOME, divided by the canopy height (Neuenschwander, 2012). Table 4
222 compiles the metrics based on waveform modeling.
223
224 **3.3. Definition of predictive models**
225 To reduce the number of LiDAR variables a preliminary selection was performed using the Akaike Information
226 Criterion (AIC), which is a measure of relative model fit that numerically expresses the amount of information
227 provided by each variable, and statistically determines the number of parameters in an equation (Akaike, 1974).
228 Thus, variables were sequentially selected until the AIC was minimized, obtaining a reduced set of predictive
229 variables. These variables were then used in a forward stepwise multiple linear regression to determine the
230 independent variables that significantly (p-value < 0.05 required) contribute to the model. The *best* model was
231 determined by progressively adding independent variables that minimize the residual mean square, maximizing the
232 coefficient of determination R^2. To avoid unrealistic over-fitting of the regression results the criterion of including a

low number of predictive variables was followed, so a maximum of three independent variables were added into the multiple linear regression models. Once models were developed, a leave-one-out cross-validation technique was employed for evaluation purposes. Prediction models were assessed using adjusted coefficients of determination R^2, Root-Mean-Square Error (RMSE), normalized RMSE (nRMSE) computed as the RMSE divided by the range of observed values, and coefficient of variation (CV) defined as the ratio of RMSE and the mean of the observed values. Finally, maps representing predicted canopy fuel parameters were created by applying the models to the data for the whole study area.

4. Results

The models developed are shown in Table 5, and Figure 4 shows scatterplots of the plot-level field-based observed vs. LiDAR-based predicted variables with linear fits and prediction confidence intervals (α=0.05). Figure 5 shows the maps for two of the canopy fuel parameters estimated: canopy bulk density (Figure 5.a) and canopy base height (Figure 5.b), at a 30 m resolution.

Overall, the forest structure parameters are well predicted, with more than 80% of the variability of aboveground biomass (R^2=0.84) and quadratic mean diameter (R^2=0.82) explained. Predictive models for basal area (R^2=0.76) and volume (R^2=0.73) had a slightly lower performance. The lowest coefficient of determination was for stand density index (R^2=0.66). The nRMSE values are very similar for all the structure parameters, ranging from 0.09 to 0.12. The lowest CV was for quadratic mean diameter (17%), whereas the rest of parameters had CV's of 23% or greater. Models developed for most of the canopy fuel parameters also explained significant amounts of variance, for example canopy height (R^2=0.79), canopy base height (R^2=0.78) and canopy fuel load (R^2=0.79), all producing similar nRMSE values (0.09-0.10). The CV however is more variable, ranging from 0.17 to 0.25, the highest value corresponding to canopy base height, the model for which clearly produces an overestimation of this parameter, as shown in Figure 4.g. The lowest model performance is for canopy bulk density (R^2=0.67), which resulted in both the highest nRMSE (0.13) and CV (38%) values. Analyzing the scatterplot for this parameter (Figure 4.h) some outliers can be observed corresponding to significant under prediction of the highest observed canopy bulk density values.

Analyzing the nature of the predictive metrics used by the various multiple regression models we observe that three of them (volume, canopy base height, canopy bulk density) combine descriptors based on both the pulse detection and waveform modeling, and two of them (quadratic mean diameter, canopy height) just utilized waveform modeling based metrics. The remaining models simply used descriptors derived from the pulse detection. Among the pulse detection based metrics, metrics describing the average height computed both from the normalized point cloud and from the CHM, are the most frequently used by the models, followed by height percentiles (25 and 50) and descriptors of the shape of a probability distribution of the heights (kurtosis and skewness). According to these models, the intensity value does not play an important role in describing the forest structure, and it is not used for estimating any canopy fuel parameters. Among the waveform modeling based metrics the mean and standard deviation of HOME, waveform distance, and height to median ratio, were selected in the predictive models.

5. Discussion

A methodology for estimation of structure and canopy fuel load variables based on metrics obtained from small-footprint waveform LiDAR data has been described and evaluated. Analogous work has been reported using discrete LiDAR data with different degrees of success, in different ecological areas, forest structure, and species, such as Mediterranean (Riaño et al., 2004; Gonzalez-Olabarria et al., 2012), boreal (Peterson and Nelson, 2011), pine barrens in New Jersey (Skowronski et al., 2011), old growth pine in Texas (Zhao et al., 2011), or Eastern Cascade dry forests (Erdody and Moskal, 2010) and coastal coniferous forest (Andersen et al., 2005). The latter two are the geographically closest and likely the most similar to this study. In our case, however, markedly more variability existed within the canopy spatial distribution, having higher mean and standard deviation of the fuel parameters compared to the other studies. This variability, which may hinder define more accurate models, is due to the heterogeneous age class distribution, as a consequence of forest management activities such as thinning practices.

The canopy base height estimation model is similar (R^2=0.78) to that derived by Andersen et al., (2005) and Erdody and Moskal, (2010) (R^2=0.77), and higher than that reported by other authors (ranging between 0.54 (Gonzalez-Olabarria et al., 2012) and 0.72 (Zhao et al., 2011). The adjusted coefficient of determination for canopy height (R^2=0.79) is significantly lower than previous studies (Andersen et al., 2005; Erdody and Moskal, 2010; Zhao et al., 2011), who reported R^2 values in the range 0.89-0.98. The result obtained for canopy bulk density (R^2=0.67) is

288 comparable with those reported by Peterson and Nelson (2011) or Zhao et al. (2011), however, this parameter has
289 been estimated to a high degree of accuracy in other studies (Riaño et al., 2004; Andersen et al., 2005; Erdody and
290 Moskal, 2010; Skowronski et al., 2011), with R^2 values greater or equal to 0.80. Our study site has an average
291 canopy fuel load (48.8 t/ha) considerably higher than that reported for other sites. The predictive model obtained for
292 this parameter (R^2=0.79) explains more variability than those reported by Zhao et al., 2011 (R^2=0.47) and
293 Skowronski et al., 2011 (R^2=0.71), but slightly less than those provided by Andersen et al., 2005 and Erdody and
294 Moskal, 2010, with 0.86 and 0.87 respectively. This diversity in results is likely due to the intrinsic differences in
295 forest structure, and denotes the possible benefit of stratifying the forest before obtaining fuel estimation models.
296 However, this can be difficult in some areas with a mixed distribution of species and varying management practices
297 over the time.
298
299 The modeled canopy bulk density (Figure 5.a) and canopy base height (Figure 5.b) maps show the distribution of
300 these two fuel parameters over the study area. The canopy bulk density is highest in areas with tallest coniferous
301 stands that have not been harvested. Additionally, canopy bulk density, canopy fuel loading and canopy base height
302 increase along riparian buffers left intact during harvesting activities (see Figure 2.c). Patterns in canopy fuel
303 distributions are mostly driven by the stand structure parameters including stand height and stand density, which in
304 turn are largely determined by land management that varies markedly across the Panther Creek study area. For
305 example, the lowest canopy height and lowest canopy fuel parameters are all located within stands that have recently
306 been clear cut, including areas along streams and some riparian buffers. The absence of defined structure within
307 these stands hinders the correct estimation of the canopy fuel parameters. The results confirm that predictive models
308 performed more accurately in areas with young and mature trees (as opposed to regrowth), especially in stands
309 subject to thinned regimes, characterized by consistent stand height, moderate canopy density, and homogeneous
310 spatial distribution.
311
312 Full-waveform LiDAR data provide valuable information regarding the full path travelled by the laser pulse, which
313 allows to the derivation of a very dense point cloud from the waveform intensity peaks and, moreover, to model the
314 returned signal to retrieve spatio-temporal information. These results show that the use of these methodologies based
315 on small-footprint waveform LiDAR data are suitable to accurately describe forest canopy fuel properties, providing

316 input variables to mathematical models that predict, quantify and map fire spread. Thus, a correct and systematic
317 characterization of both, structure and canopy fuel loads of the forests, enables a more accurate fire behavior
318 forecast that supports the development of policies for wildland fire prevention and planning.
319
320 **Acknowledgements**
321 This paper was developed as a result of a visiting scholar grant to the lead author to visit University of British
322 Columbia, funded by the Erasmus Mundus Programme of the European Commission under the Transatlantic
323 Partnership for Excellence in Engineering - TEE Project, and a mobility grant BEST/2012/235 funded by the
324 Generalitat Valenciana. The authors wish to thank the Panther Creek Remote Sensing and Research cooperative
325 program for the data provided for this research, Jim Flewelling (Seattle Biometrics) and George McFadden (Bureau
326 of Land Management) for their help in data availability and preparation.

327

328 **References**

329 Agee J (1996) Fire ecology of Pacific Northwest forests. (Island Press: Washington)

330 Ahokas E, Kaasalainen S, Hyyppä J, Suomalainen J (2006) Calibration of the Optech ALTM 3100 laser scanner
331 intensity data using brightness targets. International Archives of Photogrammetry, Remote Sensing and Spatial
332 Information Sciences, **36** (Part 1), 6p.

333 Akaike H (1974) A new look at the statistical model identification. Automatic Control, IEEE Transactions on **19**,
334 716-723.

335 Andersen HE, McGaughey RJ, Reutebuch SE (2005) Estimating forest canopy fuel parameters using LIDAR data.
336 Remote sensing of Environment 94, 441-449.

337 Andrews PL (2009) BehavePlus fire modeling system, version 5.0: Variables. Gen. Tech. Rep., United States
338 Department of Agriculture Forest Service Rocky Mountain Research Station. RMRS-GTR-213WWW Revised.
339 (Fort Collins, CO)

340 Arroyo LA, Pascual C, Manzanera JA (2008) Fire models and methods to map fuel types: The role of remote
341 sensing. Forest Ecology and Management **256**, 1239-1252.

342 Ashworth A, Evans DL, Cooke WH, Londo A, Collins C, Neuenschwander A (2010) Predicting southeastern forest
343 canopy heights and fire fuel models using GLAS data. Photogrammetric engineering and remote sensing **76**, 915-
344 922.

345 Buddenbaum H., Seeling S., Hill J. (2013). Fusion of full-waveform lidar and imaging spectroscopy remote sensing
346 data for the characterization of forest stands. International Journal of Remote Sensing, **34**, 4511-4524.

347 Chuvieco E, Cocero D, Riano D, Martin P, Martınez-Vega J, de la Riva J, Pérez F (2004) Combining NDVI and
348 surface temperature for the estimation of live fuel moisture content in forest fire danger rating. Remote Sensing of
349 Environment **92**, 322-331.

350 Chuvieco E, Congalton RG (1989) Application of remote sensing and geographic information systems to forest fire
351 hazard mapping. Remote Sensing of Environment **29**, 147-159.

352 Chuvieco E, Riano D, Aguado I, Cocero D (2002) Estimation of fuel moisture content from multitemporal analysis
353 of Landsat Thematic Mapper reflectance data: applications in fire danger assessment. International Journal of
354 Remote Sensing **23**, 2145-2162.

Chuvieco E, Riano D, Van Wagtendok J, Morsdof F (2003) Fuel loads and fuel type mapping, in Wildland Fire Danger Estimation and Mapping: The Role of Remote Sensing Data. Singapore: World Scientific 4, 119–142.

Chuvieco E, Salas J (1996) Mapping the spatial distribution of forest fire danger using GIS. International Journal of Geographical Information Science **10**, 333-345.

Cruz MG, Alexander ME, Wakimoto RH (2003) Assessing canopy fuel stratum characteristics in crown fire prone fuel types of western North America. International Journal of Wildland Fire **12**, 39-50.

Drake JB, Dubayah RO, Clark DB, Knox RG, Blair JB, Hofton MA, Chazdon RL, Weishampel JF, Prince S (2002) Estimation of tropical forest structural characteristics using large-footprint lidar. Remote Sensing of Environment **79**, 305-319.

Duong HV (2010) Processing and application of ICESat large footprint full waveform laser range data. Delft University of Technology. (Delft NL)

Erdody TL, Moskal LM (2010) Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sensing of Environment **114**, 725-737.

Finney MA (2004) FARSITE: Fire Area Simulator--Model Development and Evaluation, United States Department of Agriculture Forest Service Rocky Mountain Research Station Research Paper. RMRS-RP-4 Revised March 1998, revised February. (Ogden, UT)

Falkowski MJ, Gessler PE, Morgan P, Hudak AT, Smith A (2005). Characterizing and mapping forest fire fuels using ASTER imagery and gradient modeling. Forest Ecology and Management **217**, 129-146.

Flannigan MD, Stocks BJ, Wotton BM (2000) Climate change and forest fires. Science of the total environment **262**, 221-229.

Flewelling JW, McFadden G (2011) Lidar Data and Cooperative Research at Panther Creek, Oregon. SilviLaser 2011, Oct. 16-20, 2011 - Hobart, Australia

García M, Popescu S, Riaño D, Zhao K, Neuenschwander A, Agca M, Chuvieco E (2012). Characterization of canopy fuels using ICESat/GLAS data. Remote Sensing of Environment **123**, 81-89.

Gholz HL (1979) Equations for estimating biomass and leaf area of plants in the Pacific Northwest. Oregon State University, Forest Research Laboratory, (Corvallis OR)

González-Olabarria JR, Rodríguez F, Fernández-Landa A, Mola-Yudego B (2012) Mapping fire risk in the Model Forest of Urbión (Spain) based on airborne LiDAR measurements. Forest Ecology and Management **282**, 149-156.

383 GuangCai X, Yong P, Zengyuana L, Dan Z, Luxia L (2012) Individual trees species classification using relative
384 calibrated full-waveform LiDAR data. SilviLaser, 16-19 September 2012 - Vancouver, Canada
385 Hall SA, Burke IC, Box DO, Kaufmann MR, Stoker JM (2005) Estimating stand structure using discrete-return
386 lidar: an example from low density, fire prone ponderosa pine forests. Forest Ecology and Management **208**, 189-
387 209.
388 Harding DJ, Carabajal CC (2005) ICESat waveform measurements of within-footprint topographic relief and
389 vegetation vertical structure. Geophysical Research Letters **32**, L21S10.
390 Heinzel J, Koch B (2011) Exploring full-waveform LiDAR parameters for tree species classification. International
391 Journal of Applied Earth Observation and Geoinformation **13**, 152-160.
392 Hippenstiel R, Brownson JR (2012) Computing Solar Energy Potential of Urban Areas using Airborne LIDAR and
393 Orthoimagery, The Pennsylvania State University. (State College PA)
394 Höfle B, Hollaus M, Hagenauer J (2012) Urban vegetation detection using radiometrically calibrated small-footprint
395 full-waveform airborne LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing **67**, 134-147.
396 Hyde P, Dubayah R, Peterson B, Blair JB, Hofton M, Hunsaker C, Walker W (2005) Mapping forest structure for
397 wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. Remote Sensing of Environment
398 **96**, 427-437.
399 Keane RE, Burgan R, Van Wagtendonk J (2001) Mapping wildland fuels for fire management across multiple
400 scales: Integrating remote sensing, GIS, and biophysical modeling. International Journal of Wildland Fire **10**, 301-
401 319.
402 Kim Y, Yang Z, Cohen WB, Pflugmacher D, Lauver CL, Vankat JL (2009) Distinguishing between live and dead
403 standing tree biomass on the North Rim of Grand Canyon National Park, USA small-footprint lidar data. Remote
404 Sensing of Environment 113, 2499-2510.
405 Koetz B, Morsdorf F, Sun G, Ranson KJ, Itten K, Allgower B (2006) Inversion of a lidar waveform model for forest
406 biophysical parameter estimation. Geoscience and Remote Sensing Letters, IEEE **3**, 49-53.
407 Lefsky MA, Cohen WB, Acker SA, Parker GG, Spies TA, Harding D (1999) Lidar remote sensing of the canopy
408 structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environment **70**,
409 339-361.

410 Listopad CM, Drake JB, Masters RE, Weishampel JF (2011) Portable and airborne small footprint LIDAR: Forest
411 canopy structure estimation of fire managed plots. Remote Sensing **3**, 1284-1307.
412 Mallet C, Bretar F (2009) Full-waveform topographic lidar: State-of-the-art. ISPRS Journal of Photogrammetry and
413 Remote Sensing **64**, 1-16.
414 Mooney HA., Bonnicksen TM, Christensen NL, Lotan JE, Reiners WA (1981) Fire regimes and ecosystem
415 properties. USDA Forest Service General Technical Report WO-26. (Washington DC).
416 Morsdorf F, Meier E, Kötz B, Itten KI, Dobbertin M, Allgöwer B (2004). LIDAR-based geometric reconstruction of
417 boreal type forest stands at single tree level for forest and wildland fire management. Remote Sensing of
418 Environment **92**, 353-362.
419 Neuenschwander A (2012) Mapping vegetation structure in a wooded savanna at Freeman Ranch, TX using airborne
420 waveform LiDAR. SilviLaser, 16-19 September 2012 –Vancouver, Canada
421 Neuenschwander AL, Magruder LA, Tyler M (2009) Landcover classification of small-footprint, full-waveform
422 lidar data. Journal of Applied Remote Sensing **3**, 033544-033544.
423 Peterson B, Nelson K (2011) Developing a regional canopy fuels assessment strategy using multi-scale lidar. In
424 Proceedings of SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest
425 Ecosystems, University of Tasmania, Australia, 16-20 October 2011. (pp. 1-8).
426 Ranson KJ, Sun G, Kovacs K, Kharuk VI (2004) Landcover attributes from ICESat GLAS data in central Siberia. In
427 Geoscience and Remote Sensing Symposium. IGARSS'04. Proceedings. 2004 IEEE International **2**, 753-756.
428 Reich RM, Lundquist JE, Bravo VA (2004) Spatial models for estimating fuel loads in the Black Hills, South
429 Dakota, USA. International Journal of Wildland Fire **13**, 119-129.
430 Reineke LH (1933) Perfecting a stand-density index for even-aged forests. Journal of Agricultural Research. **46**,
431 627-638.
432 Reinhardt ED, Crookston NL (2003) The fire and fuels extension to the forest vegetation simulator. US Department
433 of Agriculture, Forest Service, Rocky Mountain Research Station. RMRS-GTR-116, (Ugden UT)
434 Reitberger J, Krzystek P, Stilla U (2008) Analysis of full waveform LIDAR data for the classification of deciduous
435 and coniferous trees. International journal of remote sensing **29**, 1407-1431.
436 Riaño D, Chuvieco E, Salas J, Palacios-Orueta A, Bastarrika A (2002) Generation of fuel type maps from Landsat
437 TM images and ancillary data in Mediterranean ecosystems. Canadian Journal of Forest Research **32**, 1301-1315.

438 Riaño D, Chuvieco E, Ustin SL, Salas J, Rodríguez-Pérez JR, Ribeiro LM, Viegas DX, Moreno JM, Fernández H
439 (2007) Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color
440 infrared ortho imaging. International Journal of Wildland Fire **16**, 341-348.
441 Riaño D, Meier E, Allgöwer B, Chuvieco E, Ustin SL (2003) Modeling airborne laser scanning data for the spatial
442 generation of critical forest parameters in fire behavior modeling. Remote Sensing of Environment **86**, 177-186.
443 Riaño D, Chuvieco E, Condes S, Gonzalez-Matesand J,y Ustin SL, (2004) Generation of crown bulk density for
444 Pinus sylvestris L. from Lidar. Remote Sensing of Environment 92, 345–352 .
445 Skowronski N, Clark K, Nelson R, Hom J, Patterson M (2007) Remotely sensed measurements of forest structure
446 and fuel loads in the Pinelands of New Jersey. Remote Sensing of Environment **108**, 123-129.
447 Skowronski NS, Clark KL, Duveneck M, Hom J (2011) Three-dimensional canopy fuel loading predicted using
448 upward and downward sensing LiDAR systems. Remote Sensing of Environment **115**, 703-714.
449 Standish JT, Manning GH, Demaershalk JP (1985) Development of biomass equations for British Columbia tree
450 species. Canadian Forestry Service, Pacific Forest Research Center: Victoria, BC, Canada, Inf. Rep. BC-X-264,
451 (Victoria BC)
452 Sumnall MJ, Hill RA, Hinsley SA, (2012) The estimation of forest inventory parameters from small-footprint
453 waveform and discrete return airborne LiDAR data. SilviLaser, 16-19 September 2012 –Vancouver, Canada
454 The American Society for Photogrammetry & Remote Sensing LAS specification, version 1.3 – R11, October 24,
455 2010., 18 pp.
456 Van Leeuwen M, Nieuwenhuis, M (2010) Retrieval of forest structural parameters using LiDAR remote sensing.
457 European Journal of Forest Research **129**, 749-770.
458 Vaughn NR, Moskal LM, Turnblom EC (2012) Tree species detection accuracies using discrete point lidar and
459 airborne waveform lidar. Remote Sensing **4**, 377-403.

460 Wagner W, Hollaus M, Briese C, Ducic V (2008) 3D vegetation mapping using small-footprint full-waveform

461 airborne laser scanners. International Journal of Remote Sensing **29**, 1433-1452.

462 Wilson BA, Ow CF, Heathcott M, Milne D, McCaffrey TM, Ghitter G, Franklin SE (1994) Landsat MSS
463 classification of fire fuel types in Wood Buffalo National Park, northern Canada. Global Ecology and Biogeography
464 Letters **4**, 33-39.
465 Zhao K, Popescu S, Meng X, Pang Y, Agca M (2011) Characterizing forest canopy structure with lidar composite
466 metrics and machine learning. Remote Sensing of Environment **115**, 1978-1996.
467

Figure 1. Location map of Panther Creek study area in Oregon.

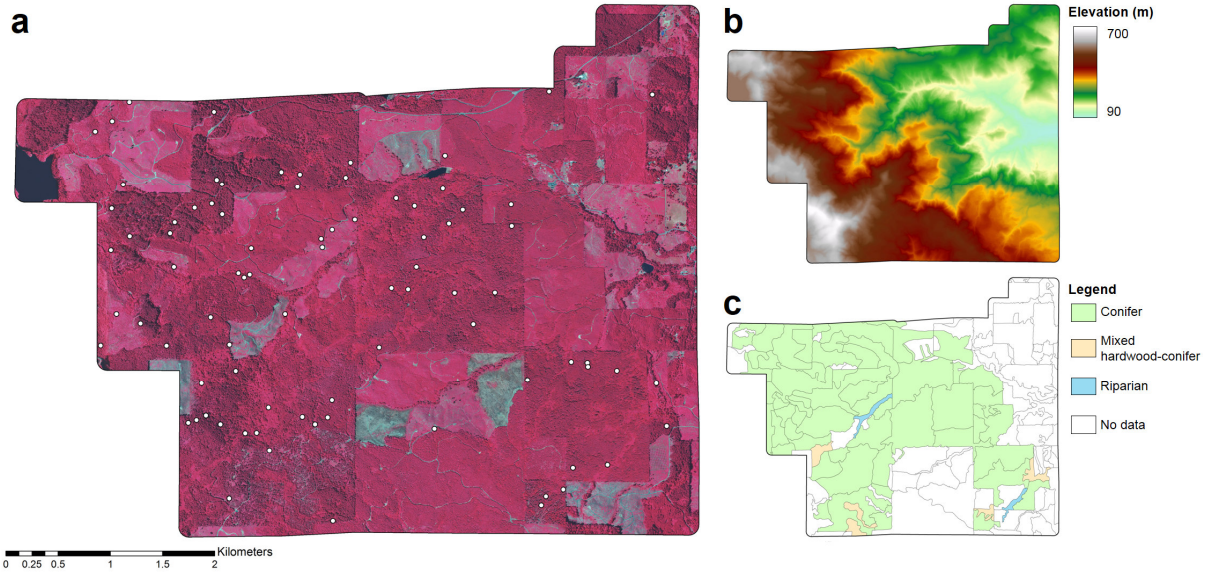
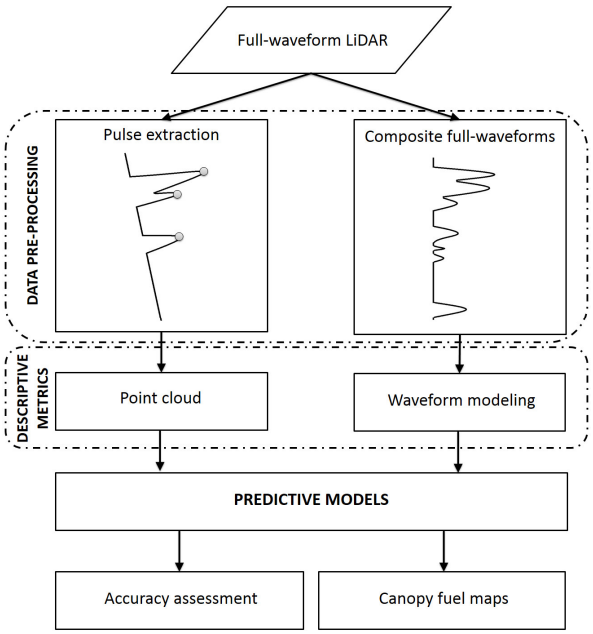


Figure 2. Graphical description of the study area: (a) color-infrared aerial ortho-image showing the location of the field plots, (b) digital terrain model, and (c) vegetation strata. Source: Prepared by the authors on the basis of data supplied by Panther Creek Cooperative Research Project.

476



477
478 Figure 3. Scheme of the proposed methodology.

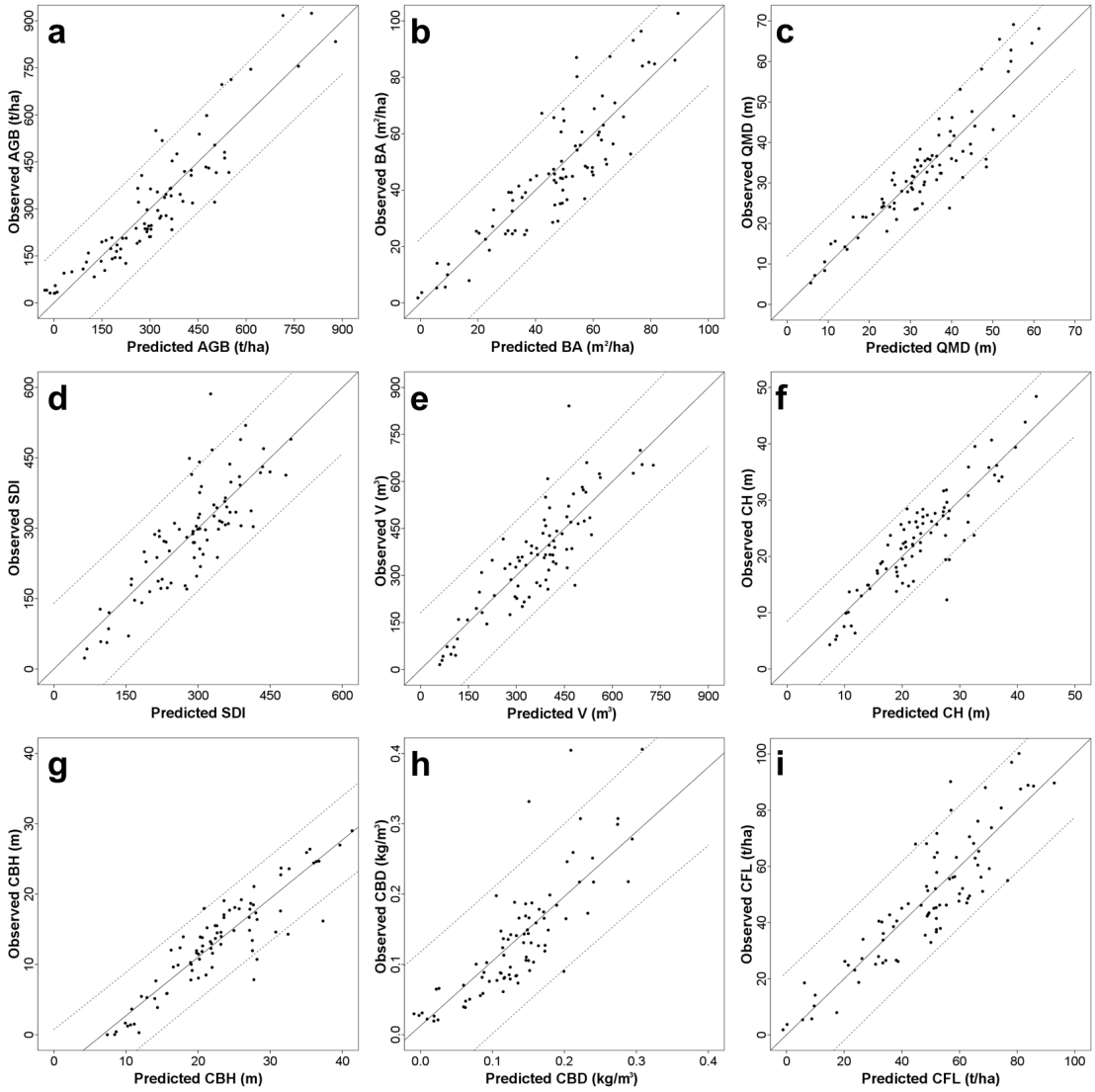


Figure 4. Plot-level observed vs. predicted values: a) aboveground biomass, b) basal area, c) quadratic mean diameter, d) stand density index, e) volume, f) canopy height, g) canopy base height, h) canopy bulk density, i) canopy fuel load. Solid line represents the linear fitting and dotted lines are the prediction confidence intervals (α=0.05).

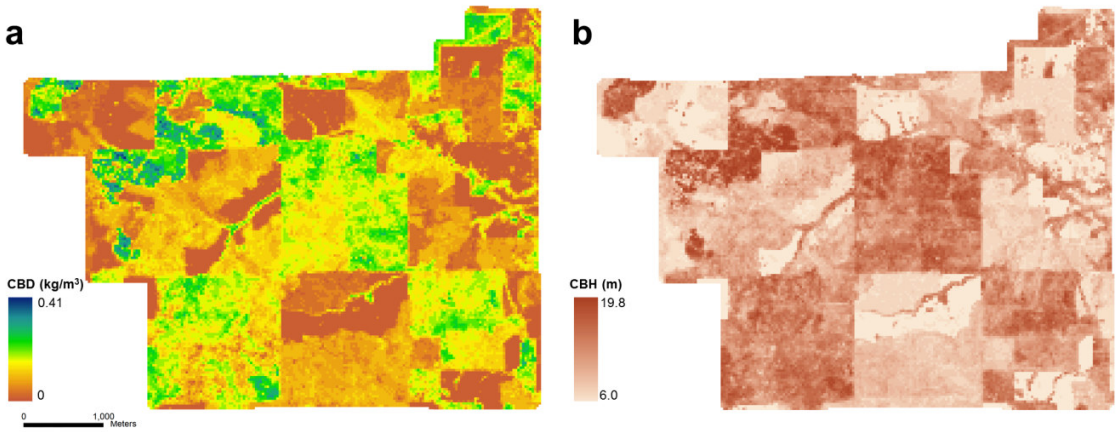


Figure 5. Maps resulting of applying the models to the study area for two of the canopy fuel parameters (a) canopy bulk density and (b) canopy base height.

Table 1. Field measurements collected for trees within the plots.

Live or estimated to be dead.
Diameter at breast height
Total Height to the highest green point.
Height-to-live-crown defined as: - Conifers: the point on the bole with live branches on 3 quadrants of the bole. - Hardwoods: the average height (on the bole) to live foliage.
Indicator for trees that are leaning over 10 degrees.
Indicator for trees with broken tops.
Indicator of trees forked above or below DBH.
Horizontal distance from plot center to tree face.
Azimuth from plot center to tree face.

491
492 Table 2. Summary statistics of per-plot forest structural and canopy fuel parameters in Panther Creek study area
493 (*n=84*).

Variable	Code (unit)	Min	Max	Mean	St. deviation
Aboveground biomass	AGB (t/ha)	30.8	924.2	310.4	202
Basal area	BA (m^2/ha)	1.8	102.7	46.4	22.8
Quadratic mean diameter	QMD (m)	5.3	69.1	33.2	13.6
Stand density index	SDI	23.3	586.3	288.3	115.6
Volume	V (m^3)	15.7	841.2	361.9	171.5
Canopy height	CH (m)	4.31	48.42	22.90	8.92
Canopy base height	CBH (m)	0	37.1	13.5	7.3
Canopy bulk density	CBD (kg/m^3)	0.019	0.406	0.136	0.084
Canopy fuel load	CFL (t/ha)	1.8	107.2	48.8	23.7

494
495 Table 3. Symbols used for the pulse detection based metrics

Symbol	Metric
CHM_{μ}	Average of the CHM height distribution
CHM_{σ}	Standard deviation of the CHM height distribution
CHM_{Range}	Range of the CHM height distribution
$CHM_{Kurtosis}$	Kurtosis of the CHM height distribution
$CHM_{Skewness}$	Skewness of the CHM height distribution
H_{μ}	Average of the normalized point cloud height values
H_{σ}	Standard deviation of the normalized point cloud height values
H_{Range}	Range of the normalized point cloud height values
$H_{Kurtosis}$	Kurtosis of the normalized point cloud height values
$H_{Skewness}$	Skewness of the normalized point cloud height values
H_{P25}	25^{th} percentile of the normalized point cloud height values
H_{P50}	50^{th} percentile of the normalized point cloud height values
H_{P75}	75^{th} percentile of the normalized point cloud height values
H_{P100}	100^{th} percentile of the normalized point cloud height values
I_{μ}	Average of the normalized point cloud intensity values
I_{σ}	Standard deviation of the normalized point cloud intensity values
I_{Range}	Range of the normalized point cloud intensity values
$I_{Kurtosis}$	Kurtosis of the normalized point cloud intensity values
$I_{Skewness}$	Skewness of the normalized point cloud intensity values
I_{P25}	25^{th} percentile of the normalized point cloud intensity values
I_{P50}	50^{th} percentile of the normalized point cloud intensity values
I_{P75}	75^{th} percentile of the normalized point cloud intensity values
I_{P100}	100^{th} percentile of the normalized point cloud intensity values

496
497 Table 4. Symbols used for the waveform modeling based metrics

Symbol	Metric
$HOME_{\mu}$	Average of height of median energy
$HOME_{\sigma}$	Standard deviation of height of median energy
WD_{μ}	Average of waveform distance
WD_{σ}	Standard deviation of waveform distance
$HTMR_{\mu}$	Average of height to median ratio
$HTMR_{\sigma}$	Standard deviation of height to median ratio
$NP_{\mu\sigma}$	Average of number of peaks
NP_{σ}	Standard deviation of number of peaks
$ROUGH_{\mu}$	Average of roughness of outermost canopy
$ROUGH_{\sigma}$	Standard deviation of roughness of outermost canopy
FS_{μ}	Average of front slope angle
FS_{σ}	Standard deviation of front slope angle
RWE_{μ}	Average of return waveform energy
RWE_{σ}	Standard deviation of return waveform energy
VDR	Vertical distribution ratio

498

499 Table 5. Predictive models and accuracy assessment results.

Variable (unit)	Model	Adjusted R^2	RSME	nRSME	CV
AGB (t/ha)	$- 73.48 + 20.828 \times H_{\mu} + 74.754 \times I_{Kurtosis} + 10.133 \times CHM_{Kurtosis}$	0.84	79.04	0.09	0.25
BA (m^2/ha)	$- 4.865 + 7.532 \times H_{\mu} - 2.906 \times H_{P50} - 1.303 \times CHM_{\mu}$	0.76	10.89	0.11	0.23
QMD (m)	$2.912 + 1.157 \times WD_{\mu} - 1.257 \times HTMR_{\mu} + 2.452 \times HTMR_{\sigma}$	0.82	5.64	0.09	0.17
SDI	$42.786 + 43.281 \times H_{\mu} - 19.246 \times H_{P50} - 7.310 \times CHM_{\mu}$	0.66	65.72	0.12	0.23
V (m^3)	$65.044 + 7.617 \times CHM_{Kurtosis} + 18.568 \times HOME_{\mu} - 9.466 \times HTMR_{\sigma}$	0.73	86.37	0.10	0.24
CH (m)	$3.323 + 1.069 \times HTMR_{\mu} - 1.509 \times HOME_{\sigma} + 3.465 \times WD_{\sigma}$	0.79	4.00	0.09	0.17
CBH (m)	$0.377 - 2.135 \times H_{Skewness} + 0.435 \times H_{P25} + 0.854 \times WD_{\sigma}$	0.78	3.34	0.09	0.25
CBD (kg/m^3)	$- 0.0157 - 0.008 \times CHM_{\sigma} + 0.005 \times HOME_{\mu} + 0.014 \times HOME_{\sigma}$	0.67	0.051	0.13	0.38
CFL (t/ha)	$- 5.5 + 7.847 \times H_{\mu} - 3.035 \times H_{P50} - 1.325 \times CHM_{\mu}$	0.79	860.63	0.10	0.22

500