



Estimating individual tree mid- and understory rank-size distributions from airborne laser scanning in semi-arid forests



Tyson L. Swetnam^{a,*}, Donald A. Falk^a, Ann M. Lynch^b, Stephen R. Yool^c

^a School of Natural Resources and Environment, The University of Arizona, Biological Sciences East, Tucson, AZ 85721, USA

^b United States Forest Service, Rocky Mountain Research Station, 1215 E. Lowell Street, Tucson, AZ, 85719, USA

^c School of Geography and Development, The University of Arizona, P.O. Box 210076 Tucson, AZ, 85721, USA

ARTICLE INFO

Article history:

Received 17 December 2013

Received in revised form 9 July 2014

Accepted 10 July 2014

Keywords:

ALS

Allometry

Scaling

Tapered Pareto distribution

Understory

Local maximum

ABSTRACT

Limitations inherent to airborne laser scanning (ALS) technology and the complex sorting and packing relationships of forests complicate accurate remote sensing of mid- and understory trees, especially in denser forest stands. Self-similarities in rank-sized individual tree distributions (ITD), e.g. bole diameter or height, are a well-understood property of natural, non-plantation, forests undergoing density dependent self-thinning and thus offer an approach to solving this problem. Alternately, semi-arid conifer forests of the southwestern USA that experience episodic wildfires and herbivory tend to exist as open stands compared to forests where disturbances are less common. We found the ITD for semi-arid forest plots with ALS-estimated canopy cover < 50% had a low rate of omission error for mid- and understory ALS trees making distribution fitting of the mid- and understory ITD unnecessary. In dense semi-arid forest plots (>50% canopy cover) the ITD correlated significantly with a tapered Pareto distribution, a power law probability distribution that is not heavy right-tailed. Two-sample Kolmogorov–Smirnov tests confirmed that observed vs. ALS-estimated overstory ITD parameters were not significantly different regardless of canopy cover. Therefore an overstory ITD derived from ALS is sufficient for fitting a continuous distribution function to estimate the ITD of the forest understory when the scale parameter is established *a priori*. Foresters and ecologists interested in measuring and modeling stand dynamics from ALS can use this approach to correct for stand density effects when developing ALS-derived single-tree inventories. Canopy cover can be used as a proxy for stand density when developing a combined ITD with area-based approaches for estimating understory in semi-arid forests.

Published by Elsevier B.V.

1. Introduction

Conventional forest inventories are based upon plot-based field surveys which can be both economically and physically impractical across large areas and in complex terrain. In contrast, today foresters with access to the right technology and skill sets have the capacity to census nearly every single tree at the landscape level using airborne laser scanning (ALS) technology – known also as Light Detection and Ranging (LiDAR) (van Leeuwen and Nieuwenhuis, 2010; Maltamo et al., 2014). Producing a complete forest inventory at the landscape level was considered an impossible undertaking before the advent of ALS. Despite their utility, a current weakness of existing ALS inventory techniques is the

inability to differentiate all trees in the stand, in particular understory trees (Maltamo et al., 2004; Hudak et al., 2009; Gatzliolis et al., 2010; Frazer et al., 2011). Differentiation is difficult in understory trees which tend to be interconnected and obscured beneath the overstory trees (Kaartinen et al., 2012; Wing et al., 2012). Hybrid techniques involving individual tree crown (ITC) isolations (Breidenbach and Astrup, 2014) and area based approaches (Lindberg et al., 2010; Maltamo and Gobakken, 2014) have been developed to obtain precision and accuracy comparable to field measurements. Most of these ALS isolation and extrapolation techniques were developed in productive temperate forests, which tend to exhibit density dependent competition for space. Applicability of hybrid techniques is uncertain in semi-arid forests, which are by contrast resource limited, have low-productivity, and experience episodic disturbances.

To estimate tree size, e.g. biomass, biometricians measure trees first by their primary size measures, e.g. diameter at breast height (DBH), tree height, and canopy volume, and then apply allometric

* Corresponding author. Address: Biological Sciences East, 1311 East Fourth Street, Tucson, AZ 85721, USA. Tel.: +1 (520) 247 2293.

E-mail address: tswetnam@email.arizona.edu (T.L. Swetnam).

equations related to each measure. At the stand scale the critical density of plants in space can be linked to the size of individuals (Enquist et al., 1998, 2009; Deng et al., 2012). This behavior, termed as Reineke's (1933) stand density index (SDI) or the Yoda et al. (1963) '–3/2 self-thinning rule' (White, 1981; Westoby, 1984), assumes that as each tree grows to fill space with its canopy and roots it competes for limited resources. Metabolic Scaling Theory (MST) (Enquist et al., 1998) predicts competitive thinning to be the main source of tree mortality in the absence of exogenous processes, resulting in ranked-size individual tree distributions (ITD) following power laws (Enquist et al., 1998, 2009; West et al., 2009; Deng et al., 2012). MST principles are based on thermodynamic and hydraulic laws which extend from the individual tree to the forest scale (Enquist et al., 1998, 1999; Enquist and Niklas, 2001, 2002; Niklas and Enquist, 2001; Niklas et al., 2003; West et al., 2009; Deng et al., 2012). The allometry of a tree becomes predictable in part because each tree must maintain a bole diameter-height proportionality based on its elastic buckling strength to remain upright and support its branches (McMahon, 1973; McMahon and Kronauer, 1976). Canopy and roots must also fill enough volume space to support the tree's metabolic demand (Enquist et al., 2009; West et al., 2009; Savage et al., 2010; Kempes et al., 2011). We assumed in the current study that semi-arid forest trees compete chiefly for plant available soil water rather than light, although both constraints are possible at cooler higher elevations on more mesic sites.

Our chief aim was to determine whether the direct inventory limitation of ALS can be addressed by considering density-dependent relationships predicted to exhibit scale-invariant power-law forms in their ITD (Enquist et al., 1998, 1999; Enquist and Niklas, 2001; Niklas and Enquist, 2001; Niklas et al., 2003; Deng et al., 2012). Many semi-arid stands are open structured, and there was also the question of whether the ITC isolation techniques are accurate in such stands. The chief question revolves around whether the observed semi-arid forest stands exhibit density dependence and if so would we be able to estimate the ITD of the understory trees using a scale-invariant function, i.e. a power law probability distribution, after assuming an understory is present? If confirmed, such a technique could improve the accuracy of forest inventories over large semi-arid landscapes, with attendant implications for estimating standing biomass, carbon sequestration, and species composition.

There are various techniques for identifying ITC from ALS data using canopy height models (CHM) (Lefsky et al., 1999, 2002; Zimble et al., 2003; Zhao et al., 2009; Kaartinen et al., 2012; Koch et al., 2014), including: (1) local maxima [with filtering] (Dralle and Rudemo, 1996; Hyyppä et al., 2001, 2005; Persson et al., 2002; Monnet et al., 2010), (2) variable area window (Popescu et al., 2002; Popescu and Wynne, 2004; Swetnam and Falk, 2014), (3) hierarchical inverse watersheds (Chen et al., 2006; Zhao and Popescu, 2007), and (4) spatial wavelet technique (Falkowski et al., 2006, 2008).

1.1. Scale invariant rank-size frequency distributions

MST predicts that density-dependent forest ITD scale as inverse square laws:

$$\Delta n_k \propto x_k^\alpha \quad (1)$$

where Δn_k is the number of individuals in a standardized area in a given size class or bin, k , x is any primary size measure, and the exponent $\alpha < 0$ is negative, which for the example of bole radius is $\alpha = -2$ (Enquist et al., 2009). For linear binning a continuous function is given as:

$$f(x) \equiv dn/dx \propto x^\alpha \quad (2)$$

Continuous ITD functions should be approximated using a probability distribution function because there is a minimum size limit $x > x_{\min}$ to the point at which the power law holds before the density diverges (as $x \rightarrow 0$) (Clauset et al., 2009); typically the x_{\min} is the size of the smallest individual tree. A Pareto distribution:

$S(x) = \Pr\{X \geq x\} = \left(\frac{x}{x_{\min}}\right)^{-\alpha}$, should be used with continuous data in this case (Newman, 2005; White et al., 2008; Clauset et al., 2009). In practical applications a power law Pareto is unlikely to be the best fit for a forest ITD because of divergence in both tails of the distribution due to stand dynamics, e.g. physiological tolerances and disturbance, and physiological limits to tree size. Exogenous impacts such as trampling, wind throw, wildfire and herbivory influence the ITD and result in both left- and a right-tail truncation of the canonical power law distribution (Seuront and Mitchell, 2008; Enquist et al., 2009; Swetnam, 2013). Examples of other continuous distributions that account for divergence in the right-tail include the tapered Pareto (Kagan and Schoenberg, 2001), truncated Pareto (White et al., 2008; Enquist et al., 2009), and negative exponential and stretched exponential functions (Clauset et al., 2009). In forestry applications ITD have been reported using a negative or stretched exponential function and in particular the Weibull distribution (Weibull, 1951; Dubey, 1967; Bailey and Dell, 1973; Maltamo et al., 2000, 2004). Truncated Weibull distributions tend to model the left tail of the distribution well (Maltamo et al., 2004); relatedly, both the two-parameter Weibull (with a shape parameter < 1) and Pareto distribution are heavy right tailed (Newman, 2005). In the right tail of a forest ITD we expect a decline in abundance at an exponential rate. For our analyses we applied the tapered Pareto distribution (Kagan and Schoenberg, 2001; Schoenberg and Patel, 2012), a power law distribution that accounts for the observed behavior of the right-tail in forest rank-size ITD.

2. Methods

Maltamo and Gobakken (2014) suggest the combination of an ALS derived ITD and an area-based approach to estimate the full ITD of a stand. Here we follow that approach, evaluating first whether a tapered Pareto distribution (Pareto, 1896; Kagan and Schoenberg, 2001) describes the ITD in a defined area of interest, and then testing whether the ITD of the understory shares the same scale-invariant parameters as the ALS derived overstory ITD. We grouped our plots by canopy cover percentage to determine whether the self-similar properties of the ITD are a function of density dependence.

2.1. Study areas

We incorporated both field datasets and ITC from ALS for two semi-arid conifer-dominated forests in the southwestern USA. The Valles Caldera National Preserve is located in the Jemez Mountains west of Santa Fe, New Mexico at 35.9°N, 106.5°W (Fig. 1). Elevations range from 2300 m above mean sea level (amsl) in Redondo Meadow to 3431 m amsl atop Redondo Peak. The Valles Caldera is a collapsed volcanic caldera with a rim approximately 19 km wide; within the caldera are resurgent domes over 200 m high, locally referred to as *cerros*. Ponderosa pine (*Pinus ponderosa*) is common in the lowest elevations (2100–2400 m amsl), with some limber pine (*Pinus flexilis*) and Douglas-fir (*Pseudotsuga menziesii*) on mesic sites. Gambel oak (*Quercus gambelii*) is common in post-fire seral stands along with ponderosa pine and quaking aspen (*Populus tremuloides*). North aspects tend to be dominated by Douglas-fir and white fir (*Abies concolor*), with some subalpine fir (*Abies lasiocarpa*) and Engelmann spruce (*Picea engelmannii*). The highest elevation sites are dominated by Engelmann spruce.

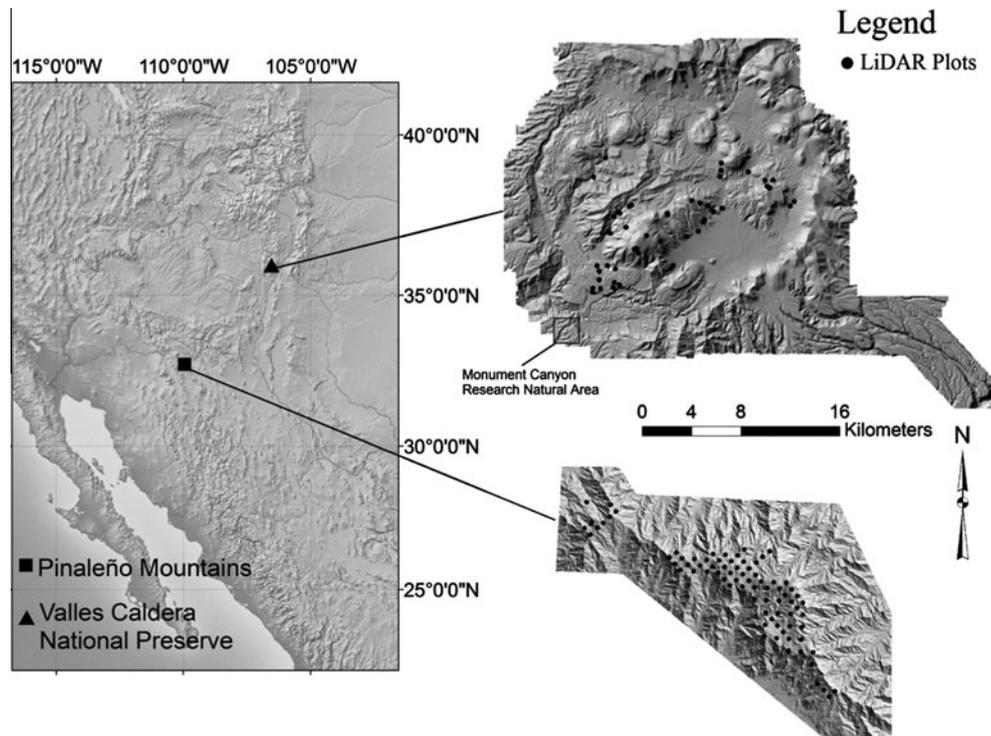


Fig. 1. ALS study areas in Arizona and New Mexico. The Valles Caldera have 48 0.1 ha randomly located forest calibration plots; the Pinaleno Mountains have 78 0.05 ha plots located in a gridded transect.

The Pinaleno Mountains are located south of Safford, Arizona at 32.7°N, 109.9°W rising to 3267 m amsl at High Peak (Fig. 1). The Pinaleno are characteristic of the Basin and Range province and are comprised of a complex of steeply sided canyons with relatively gentle high elevation uplands above 2700 m. Lower elevation (1830–2440 m amsl) forests of the Pinaleno are typical of the Madrean Sky Islands (Whittaker and Niering, 1975; Niering and Lowe, 1984). Forests above 2135 m amsl are distributed along gradients of elevation and aspect. The flora of the Pinaleno shares many species with the Jemez (Niering and Lowe, 1984; Muldavin et al., 2006). In the Pinaleno lower elevations with ponderosa pine and several oak species transition to dry mixed-conifer forests dominated by Douglas-fir, southwestern white pine, and ponderosa pine with minor components of white fir and aspen. Above 2750 m amsl mesic mixed-conifer forests are dominated by Douglas-fir and white fir with minor components of southwestern white pine, Engelmann spruce and corkbark fir (*A. lasiocarpa* var. *arizonica*). Engelmann spruce and corkbark fir are dominant species at the highest elevations, forming the southernmost spruce-fir forest in North America. Post-fire communities of quaking aspen are present throughout the upper elevations.

2.1.1. Plot observations

Plot centers were located with a differentially corrected global positioning system, which logged points continuously during tree surveys; absolute plot location after differential correction had sub-meter accuracy in the horizontal plane. Ground-based primary size observations collected in both study areas included DBH and maximum height. Tree heights were measured with a Nikon Forster 550 laser range finder (Valles Caldera) and Laser Technology Impulse 2000 Hypsometer (Pinaleno). Of 399 field surveyed tree heights identified from the Pinaleno plot data, the ITC cross-validation had Pearson's $r = 0.984$ and a mean square error (MSE) ± 81 cm (SM Fig. 1). Variability in the relative locations of tree boles at ground level vs. the point of the apical leader (tree lean) resulted

in locational differences in the CHM exceeding three meters horizontal in some cases.

Forty-eight 0.1-ha radial plots were sampled in the Valles Caldera in 2010 (Swetnam, 2013) (Fig. 1). A total of 1520 live and dead trees were measured, and 3952 trees were counted (including seedlings and saplings >15 cm tall). Within each 0.1 ha plot all trees >18.5 cm DBH were measured; within a 0.01 ha inner plot all trees >2.5 cm DBH were measured. The understory distribution was extrapolated from the 1/10th area by multiplying the frequency distribution by ten.

In the Pinaleno seventy-nine 0.05 ha radial plots were sampled in 2008–2009 (O'Connor, 2013; O'Connor et al., in press) (Fig. 1). A total of 2,862 trees were measured for DBH and height. Within each plot all trees >20.0 cm DBH were measured; in a random 1/3 subplot area all trees >2.5 cm DBH were measured. The understory distribution was extrapolated from the 1/3 plot area by multiplying the frequency distribution by three.

2.2. ALS acquisitions

ALS data for the Valles Caldera and Pinalenos share similar flight parameters and achieved comparable pulse return densities (Table 1). The discrete point cloud data were considered sufficiently dense [>8 points per meter square (ppsm)] to conduct rigorous assessments of canopy structure in complex terrain (Laes et al., 2008, 2009; Gatzliolis et al., 2010; Hudak et al., 2009). The pulse returns have a footprint of ~ 20 –50 cm at ground level assuming an average flight elevation of 1000 m above ground level and beam divergence of 0.20–0.5 milliradians (Table 1). Because the surface is not 100% illuminated by the laser pulses, we use the ratio of pulse returns to estimate canopy cover within the sample plots (cover% = # returns >2 m ground level/total # returns) (McGaughey, 2012).

CHMs (Lefsky et al., 1999) were generated at a consistent resolution (0.333 m) in USFS FUSION (McGaughey, 2012) (SM Fig. 2).

Table 1
Study area ALS flight-parameters.

Scan characteristic	Valles Caldera	Pinalaño
Vendor/provider	NCALM	Watershed sciences
Acquisition date	January 2010, July 2010	September 22–27, 2008
Scanner	Optech Gemini	Leica ALS50 Phase 2
Pulse rate	100 kHz	70–90 kHz
Scan rate	<100 Hz	52.2 Hz
Pulse returns	1–4, + 8-bit Intensity (0–255)	1–4, + 8-bit Intensity (0–255)
Scan angle	25 degrees	15 degrees
Divergence in milliradians (mrad)	0.25 mrad	0.22 mrad
Stated accuracy (vertical/horizontal)	7.0 cm/1.0 m	3.2 cm/1.0 m
Flight above ground level	~1000 m	800–1300 m
Flight line overlap	50% side lap	50% side lap
μ Pulses per square meter (ppsm)	Leaf-off (Snow): 8.86 ppsm Leaf-on, 7.36 ppsm	Leaf-on: 5.91 ppsm 0.98 ppsm
μ Bare ground spacing (ppsm)	1.11 ppsm	0.98 ppsm
Acquisition Area	Leaf-off: 72,648 acres (29,400 Ha) Leaf-on: 186,811 acres (75,600 Ha)	85,518 Ac (34,608 Ha)
Σ Pulse returns	Leaf-off: 2,541,885,987 Leaf-on: 7,754,915,628	2,892,925,979
Units	Meters	Meters
Projection, datum	WGS84, NAD83	WGS84, NAD83

For all plots in both data sets the mean pulse return density (μ) was calculated to be $9.3 < \mu < 25.1$ ppsm based on the density of all returns in each plot divided by the plot area. At these pulse return densities the CHM was assumed to have least one pulse return per pixel. No smoothing or median filters were applied to the initial CHMs. The plot extracted CHMs included some 'salt and pepper' NoData pixels (SM Fig. 2), suggesting that the density of the pulse return data was not uniform. However, these pixels had little effect on tree height estimation and tree identification because the Variable-Area Local Maxima (VLM) algorithm (Swetnam and Falk, 2014) incorporates a disk-shape dilation for each identified maximum.

ITC were identified from CHMs using the VLM algorithm (Swetnam and Falk, 2014) in MATLAB 2012b (Mathworks, 2012). ITC were determined by setting the expected minimum canopy radius as: $r_{can} = 0.1h$, where h was the local maximum height (Swetnam and Falk, 2014). For the frequency distribution analysis, we grouped plots by 10% canopy cover classes and linearly binned the distribution of observed tree height vs. the ALS derived ITC. We report the absolute error:

$$\Delta h = h_{VLM} - h_{obs} \quad (3)$$

where h_{VLM} are the height of the ITC and the observed tree heights h_{obs} . The relative error is given as:

$$\delta h = \frac{\Delta h}{h_{obs}} = \frac{h_{VLM} - h_{obs}}{h_{obs}} = \frac{h_{VLM}}{h_{obs}} - 1. \quad (4)$$

2.3. Fitting the ALS-derived continuous distribution function

Taubert et al. (2013) reviewed potential biases and uncertainties in fitting tree size data to power law, exponential, and Weibull distributions by either least-squares regression or maximum likelihood estimation (MLE). Pre-binning data can lead to symmetric variations around the true value of the distribution potentially biasing least-squares regression (White et al., 2008; Clauset et al., 2009; Taubert et al., 2013). The frequency distributions of trees in our data are related to the dynamic life histories of the stands and contain some observational errors. Here we were interested primarily in calculating the best fit ITD for the field observation to the ALS by the tapered Pareto (Kagan and Schoenberg, 2001; Schoenberg and Patel, 2012). Using MLE to fit only the overstory distribution would not account for the divergent behavior in the

left-tail of the distribution, resulting in a poorer fit than by least squares.

We linearly binned the height ITD data in 1 m width bins, based on the observation that bins smaller than 1 m resulted in an increasing variation among bins. For the tapered Pareto the cumulative distribution function (CDF) is given as:

$$F(h) = 1 - \left(\frac{\beta}{h}\right)^\alpha e^{\left(\frac{\beta-h}{\theta}\right)}, \quad h > h_{min} \quad (5)$$

where α is the shape parameter, β is the scale parameter typically known from the observed data (we use $\beta = 1$ m), and θ is the point at which the distribution begins an exponential decline away from the log-log linear behavior of the Pareto (Kagan and Schoenberg, 2001; Schoenberg and Patel, 2012). The Probability Density Function (PDF) is given as:

$$f(h) = \begin{cases} \left(\frac{\alpha}{h} + \frac{1}{\theta}\right) \left(\frac{\beta}{h}\right)^\alpha e^{\left(\frac{\beta-h}{\theta}\right)}, & h \geq h_{min} \\ 0, & h < h_{min} \end{cases} \quad (6)$$

The θ value is scale-dependent, decreasing for very large sample sizes because the upper limit to the distribution is finite. The tapered Pareto's left-tail decreases hyper-exponentially in the right-tail when it has a negative α parameter; for $\alpha = 0$ the right-tail of the distribution decreases exponentially (Kagan and Schoenberg, 2001), and sub-exponentially decreases when α is positive (Fig. 2).

We determined the α and θ parameters of the tapered Pareto (Eq. (6)) normalized with a constant, $c * f(h)$ for the full ITD by least-squares regression using the MATLAB 2012b Curve Fitting Tool (Mathworks, 2012). For our least-squares regressions we report the estimated α as $\hat{\alpha}$, the estimated θ as $\hat{\theta}$, and estimated c as \hat{c} . We set the scale-invariant α parameter of the ALS derived ITC to the *a posteriori* value of $\hat{\alpha}$ from the observed ITD. We also evaluate whether increasing degrees of canopy cover, a proxy value for increasing stand density, affect the $\hat{\alpha}$, $\hat{\theta}$ parameters.

2.4. Determining whether the ALS overstory ITD describes the understory ITD

In order to evaluate whether the tapered Pareto distribution of the ALS derived ITC for overstory share a self-similar distribution with the observed understory ITD we conducted a two sample Kolmogorov–Smirnov (K–S) test (Massey, 1951; Wang et al., 2003).

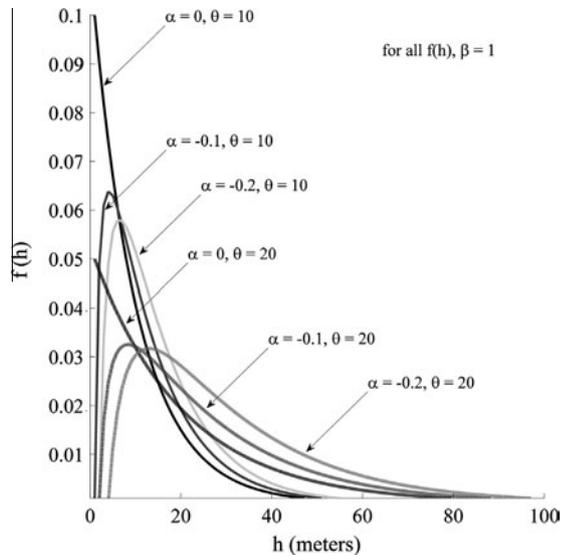


Fig. 2. Examples of the tapered Pareto PDF $f(x)$ with different shape (α) and (θ) taper parameters; the scale β parameter is set to 1 m for each. The x -axis is set to the range of known tree heights (in meters) but could be changed to any range of scales. When $\alpha = 0$, the distribution declines exponentially (black line); when $\alpha < 0$ the left-tail is truncated toward zero below x_{\min} . Increasing the taper parameter θ makes the right-tail of the distribution heavier (dashed lines). Co-varying both α and θ changes the skew of the left-tail of the distribution.

The two-sample K–S test is a nonparametric test for the equality of two sample distributions sharing the same theoretical distribution function (Massey, 1951; Wang et al., 2003). If the p -value of the test is small, we can conclude that the samples come from different distributions; if it is large we cannot reject them as being different. The K–S statistic is:

$$D = \max |F_1(h) - F_2(h)| \quad (7)$$

where F_1 and F_2 are the distribution functions of the first and second sample respectively. We tested whether the observed understory ITD is the same as the tapered Pareto empirical CDF derived from the ALS overstory ITD by conducting the two sample K–S tests of both study areas.

3. Results

3.1. ALS ITD compared to the observed ITD

The estimated range of canopy cover from ALS in the plots varied between ~5% and 97% (Table 2). The VLM was generally accurate at isolating all trees in the plots when cover percentage was <50% (Figs. 3 and 4). The VLM was also generally accurate at isolating trees >10 m height regardless of cover percentage (Figs. 3 and 4). The height at which the VLM accuracy is reduced in the understory co-varies with increasing canopy cover.

There was a non-linear increase in the number of shorter trees in the ITD with increasing canopy cover at both sites (Fig. 5). The Pinaleno had higher canopy cover on average, with 46 of 79 plots exceeding 60% cover, compared to 6 of 48 plots in the Valles Caldera (Table 2). Further, in plots with a larger sample size the ITD exhibited less variance between the linear bins, as might be expected from a larger area of interest.

Absolute error (Δh) in the number of trees observed by the ALS increased at both study sites below 10 m height (Fig. 6a). The relative error (δh) between the ALS and the observed for trees between 8 m and 30 m was $\sim 14.5\% \pm$ the observed number (Fig. 6b). Below 8 m height the rate of omission for both Δh and

δh increases rapidly as trees became increasingly difficult to discriminate with the ALS data using the VLM algorithm. Above 30 m height δh errors tend to be large because of the small number of samples, typically one or no observed trees; errors in these height classes are due to omission errors of trees where the VLM failed to isolate a canopy and commission errors from unrecorded trees, likely along the edges of the plot boundary. The observed MSE between the plot measured trees and the ALS was ± 81 cm (SM Fig. 1); for the 1 m linear bins (Figs. 3 and 4) it is likely that some of the error is caused when trees are binned incorrectly into a larger or smaller bin size.

3.2. Tapered Pareto distribution parameter estimation

For both study areas the ITD exhibited a close fit to the tapered Pareto (Table 2). The scale parameters of the observed ITD tapered Pareto PDFs [Eq. (6)] were similar across the same height range (1–38 m) in the two study areas ($\hat{\alpha} \cong -0.105 \pm 0.304$, $r^2 = 0.916$ in the Pinalenos and $\hat{\alpha} \cong -0.106 \pm 0.021$, $r^2 = 0.896$ in the Valles Caldera (Table 2)). Estimated $\hat{\alpha}$ values were also similar across canopy cover categories to those of the full ITD and ranged between -0.278 and -0.074 in the Valles Caldera, and -0.177 and -0.0001 in the Pinaleno (Table 2). The estimated taper parameter $\hat{\theta}$ for the full distribution was $\hat{\theta} \cong 9.5 \pm 4.8$ m (Valles Caldera) and $\hat{\theta} \cong 9.7 \pm 3.9$ m (Pinaleno) (Table 2, Fig. SM5). The estimated normalization constant \hat{c} was greater than the observed frequency of stems n for the actual plot data in almost all cases (Table 2). The \hat{c} consistently overestimated the actual stem frequency by $\sim 125\%$ ($n = 0.7992\hat{c} - 23.543$, $r^2 = 0.9997$).

3.3. Comparing the ALS tapered Pareto to the observed understory ITD with two-sample K–S tests of the empirical cumulative distribution functions

The VLM algorithm accurately estimated the frequency of almost all trees, including the understory, in stands with canopy cover < 50%. Therefore we limited our application of the tapered Pareto distribution fitting to the plots that under-predicted the understory, e.g. >50% canopy cover in the Valles Caldera (Figs. 3 and 7), and >70% canopy cover in the Pinaleno (Figs. 4 and 8). The VLM lost precision ~ 10 m height (Figs. 3 and 4), so we set a cut-off in the size distribution <10 m from which to fit the tapered Pareto (Figs. 7 and 8, SM Fig. 3). In both study areas, the ALS-derived tapered Pareto distribution function fit the observed overstory ITD > 10 m with r^2 values > 0.89 (Table 3).

There was no significant difference between the ALS overstory tapered Pareto and observed ITD empirical CDFs for either study area suggesting that tapered Pareto accurately reflects the empirical ITD (Table 4 and Fig. 9). The ALS tapered Pareto distribution vs. the observed understory ITDs were also not significantly different based on a two sample K–S test (Fig. 10 and Table 4).

4. Discussion

Our objective was to determine whether ITD derived from ALS can be used to generate tree inventories with accuracy similar to those of field-based methods. Our reasoning was based on the recognition that natural, non-plantation, semi-arid forest ITDs exhibit measurable self-similarity due to density-dependence. We modeled ITDs with a continuous power law distribution, the tapered Pareto, rather than an exponential function as has been employed in the past (Maltamo et al., 2004) because MST (Enquist et al., 1999, 2009; West et al., 2009) predicts that forest ITDs are power law distributed rather than exponentially distributed. We considered the tapered Pareto an optimal candidate for fitting

Table 2
Least squares regression of the tapered Pareto $c * f(h)$ of each full study area observed ITD, and by canopy cover (CC%) for the Valles Caldera (top) and Pinalaño (bottom). In all cases linear bins are 1 m wide and β was fixed at 1 m.

CC (%)	Plot n	Tree n	Range	r^2	RMSE	\hat{c}	$\hat{\alpha} \pm SE$	$\hat{\theta} (m) \pm SE$
Valles Caldera All	48	3909	1.0 m < h < 38 m	0.916	34.19	4962	-0.106 ± 0.021	9.38 ± 1.19
<40	11	636	1.0 m < h < 27 m	0.59	15.08	869	-0.278 ± 0.120	7.43 ± 2.02
40–50	14	697	1.0 m < h < 34 m	0.488	16.27	926	-0.085 ± 0.055	12.87 ± 6.31
50–60	17	1881	1.0 m < h < 38 m	0.901	23.5	2371	-0.123 ± 0.715	7.81 ± 5.35
60–70	6	695	1.0 m < h < 31 m	0.578	13.32	866	-0.074 ± 1.102	11.72 ± 17.52
Pinalaño All	79	3,941	1.0 m < h < 35 m	0.896	35.25	4921	-0.105 ± 0.304	9.77 ± 5.10
<40	13	266	1.0 m < h < 28 m	0.558	4.24	335	-0.001 ± 0.252	17.03 ± 6.70
40–50	11	346	2.0 m < h < 30 m	0.717	6.19	462	-0.115 ± 0.052	9.55 ± 3.10
50–60	9	566	2.0 m < h < 32 m	0.763	12.79	727	-0.171 ± 0.050	6.19 ± 8.90
60–70	7	300	3.0 m < h < 31 m	0.515	7.53	410	-0.093 ± 0.061	11.76 ± 5.31
70–80	13	598	3.0 m < h < 32 m	0.832	6.36	790	-0.177 ± 1.210	10.47 ± 11.41
80–90	19	1393	1.0 m < h < 35 m	0.860	15.98	1768	-0.114 ± 0.676	9.33 ± 6.20
90–97	7	472	3.0 m < h < 30 m	0.606	7.28	621	-0.064 ± 1.190	13.56 ± 19.80

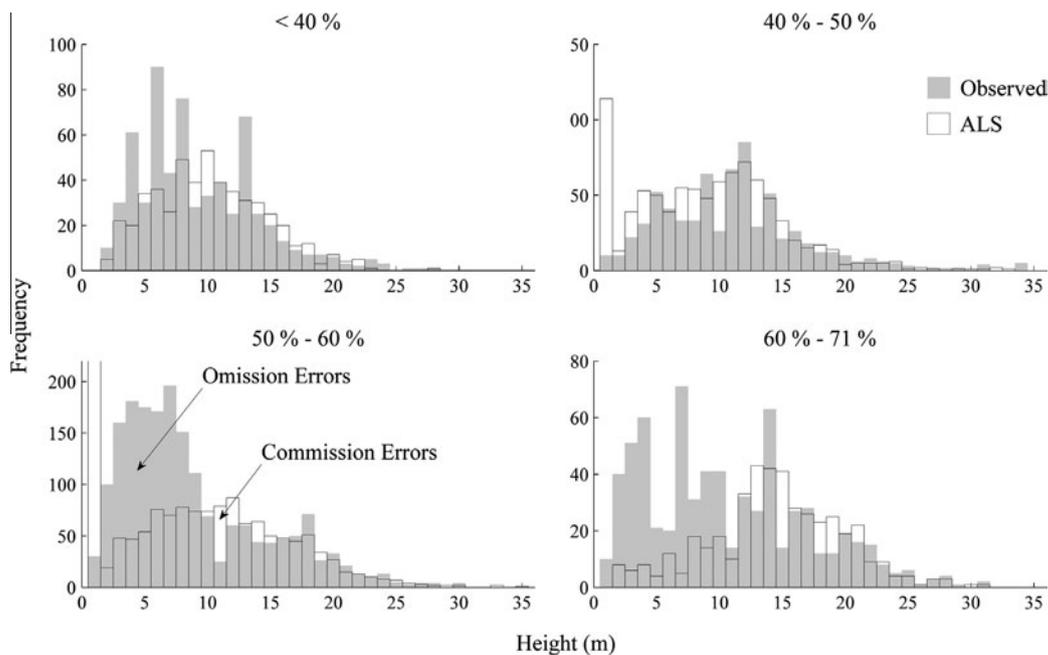


Fig. 3. Observed (gray bars) and ALS (open bars) height ITD for the Valles Caldera plots by 10% canopy cover classes.

the distribution because it is a continuous power-law distribution which includes an exponential decline in the right tail, and has been shown to be a good fit to other natural phenomena (Schoenberg and Patel 2012).

We were able to successfully isolate individual trees in the plots from the ALS when canopy cover was relatively low and/or trees were >10 m in height. When canopy cover is low there was no need to fit a distribution function to the understory; the observed frequency distribution was close to the observed. Open stands were a common feature in these semi-arid forests prior to modern fire suppression (Muldavin et al., 2006; O'Connor, 2013; O'Connor et al., in press), and are common feature of forests throughout much of the southwestern USA. These results should give managers confidence in the precision of ALS derived semi-arid forest inventories in stands with open structure.

For stands with high canopy cover and relatedly higher stem density, the ITD exhibited a close fit to the tapered Pareto. The scale parameters of the observed ITD were similar between study areas and canopy cover categories, suggesting the forests are exhibiting self-similarity in the ITD. For plots where canopy cover was high we used the tapered Pareto PDF of the ALS overstory ITD to estimate the PDF of the understory ITD, normalized by a

constant. The tapered Pareto PDF for both the field observed ITD and ALS-derived overstory ITD were not significantly different based on the two sample K-S tests. Further, the observed understory ITD and ALS tapered Pareto were not significantly different based on the two sample K-S tests. These results suggest that in plots where canopy cover is high, stand density is determined by plant functional type specific density dependencies. Thus, the ITDs exhibit measurable self-similarity for which any part of the distribution can be used to predict the scaling of the entire distribution.

4.1. Uncertainties

The MSE in field measured tree height vs. the ALS derived tree height was ± 81 cm. The binned distributions as presented in Figs. 3 and 4 may include some error where trees are incorrectly binned into either a larger or smaller 1 m-wide bin. Errors related to the binning are normally distributed (SM Fig. 1); based on the reported confidence intervals these differences in the least-squares regressions are not significantly different across the two studies. Due to observational uncertainty parameter values estimated from least squares with binning are equally unlikely to fit canonical MST predictions regardless of whether we used MLE (Taubert et al., 2013).

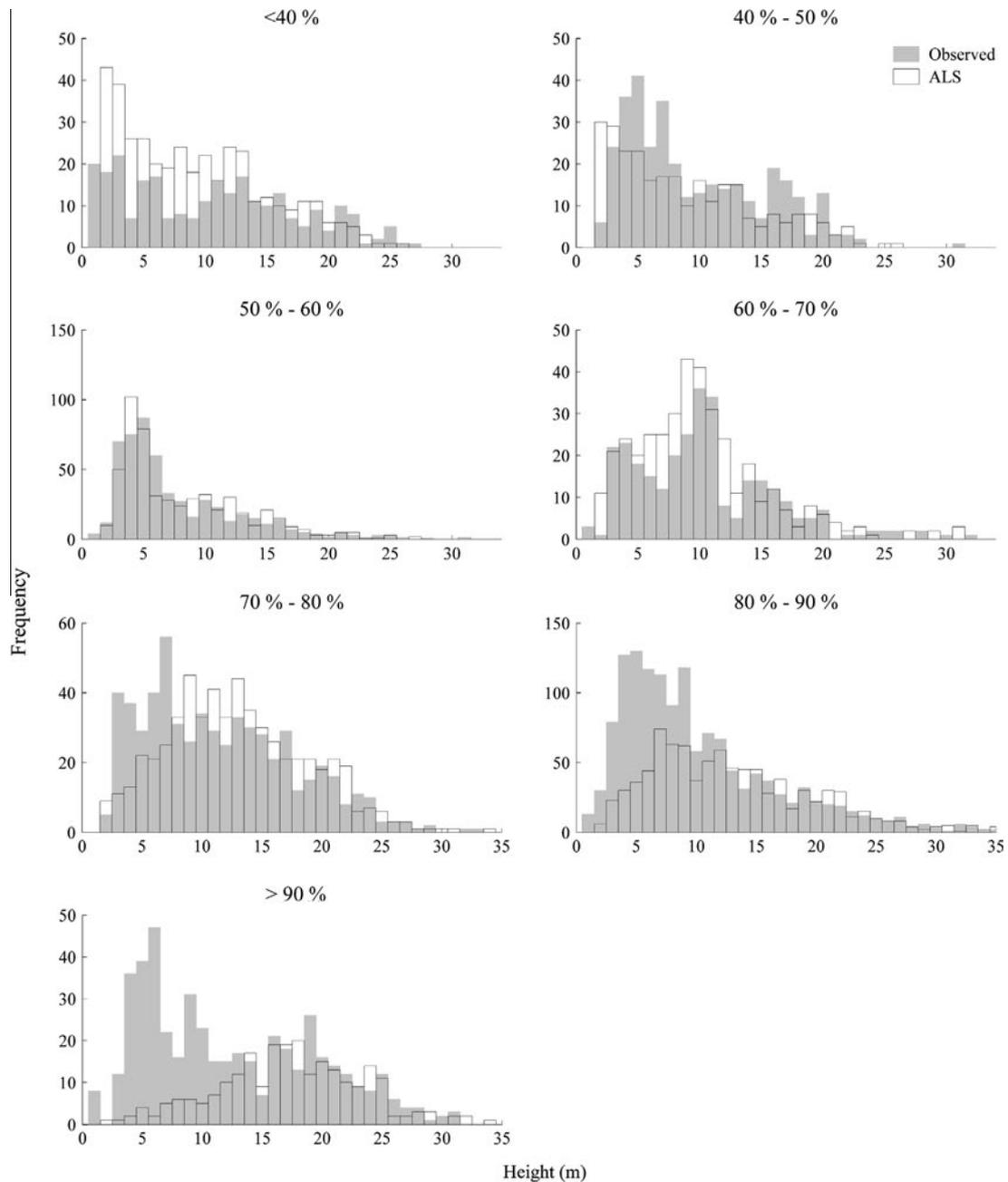


Fig. 4. Observed (gray bars) and ALS (open bars) height ITD for the Pinalaño plots by 10% canopy cover classes.

The values of the estimated scale parameter $\hat{\alpha}$ for our observed distributions tended to be slightly negative (Table 2 and SM Fig. 4). The negative sign reflects a cross-over: the point at which the distribution reverses from an increasing frequency to decreasing frequency with increased tree size. In other forest ITD studies this behavior has been modeled using Weibull distributions (Maltamo et al., 2004); here we modeled this behavior for the tapered Pareto for the first time. Uncertainty in the estimates of $\hat{\alpha}$ made it impossible to detect significant trends in the value of the parameter with changes in canopy cover (SM Fig. 4), but the finding that the sign of $\hat{\alpha}$ is negative given all the distributions decline to zero at or below h_{\min} is robust. There was also no significant difference in the values of $\hat{\theta}$ among canopy cover percentage groups (SM Fig. 5). This is likely due to the average sample area being similar across each group and our sample size being relatively small.

We did not attempt to fit $\hat{\alpha}$ for the ALS-derived overstory ITD [10 m < h < 38 m] because least-squares regression is likely to fit a positive $\hat{\alpha}$ parameter value if the distribution does not contain the cross-over in the left-tail. This would also be the case if we had used MLE (Taubert et al., 2013). An $\hat{\alpha} \geq 0$ would over-predict the frequency of the understory ITD. For the regression of the ALS derived ITD the observed negative $\hat{\alpha}$ was fixed *a priori* as was β ; distribution fitting was confined to \hat{c} and $\hat{\theta}$ of the overstory ITD. Additional supporting evidence for this approach was the finding that the shape parameter $\hat{\alpha}$ did not change significantly regardless of increasing canopy cover (Table 2, SM Fig. 4).

The understory distribution was extrapolated in both study areas because the effort to measure all understory trees in the field, particularly in dense plots where understory stem densities exceed 1000 trees per ha, is overwhelmingly time consuming and would

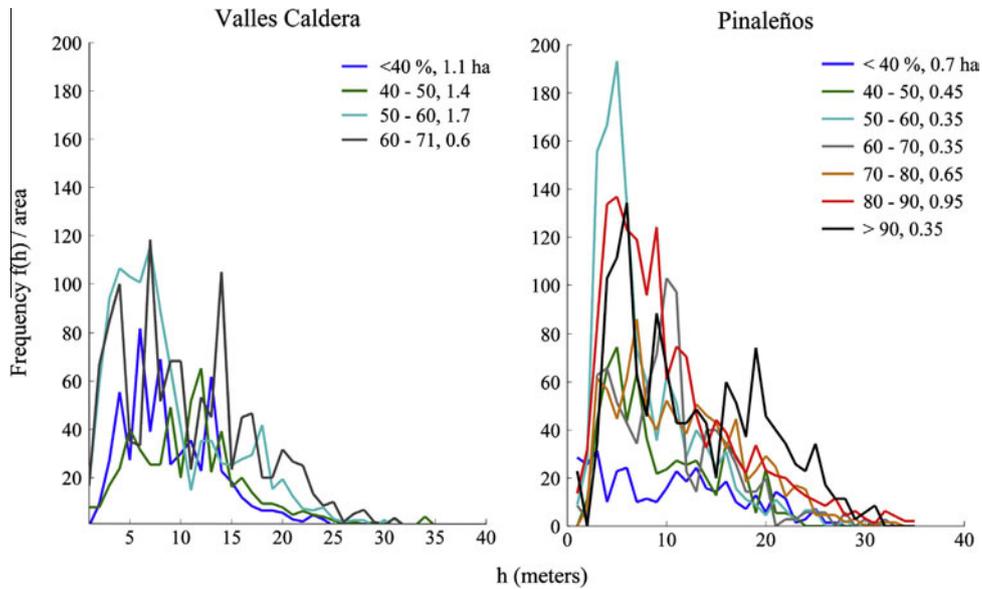


Fig. 5. Observed height ITD for both sites normalized per unit area and grouped by cover percentage in Valles Caldera (left) and Pinalteño (right).

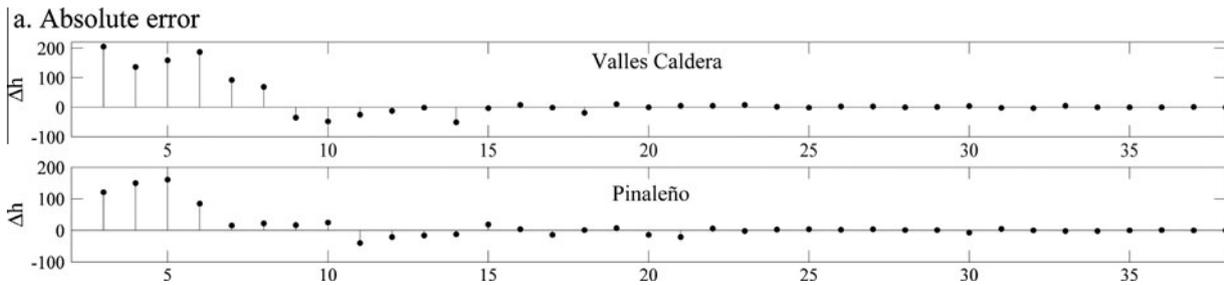


Fig. 6a. Absolute error (Δx) of the observed ITD vs. the ALS ITD for the Valles Caldera and Pinalteño. There are few errors across the distribution >8 m for each location. Trees <1 m are not shown as the VLM used a lower cutoff at 1 m height.

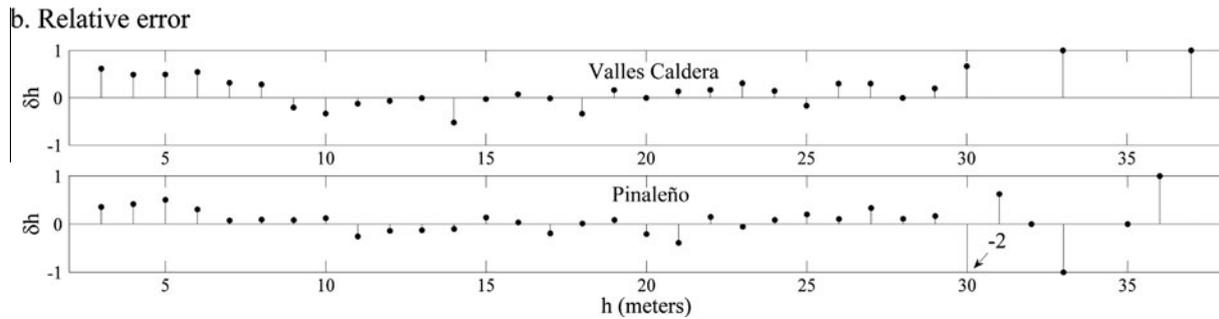


Fig. 6b. Relative error δx for the observed vs. VLM modeled tree size frequency for the Valles Caldera and Pinalteño. There are larger errors relative to the frequency of trees for both ends of the distribution. For trees <8 m there is an increase in the omission rate due to problems related to canopy spacing and the VLM. There are larger relative errors in trees >29 m height due to the absolute number of those trees being very small in the sample – a single error of omission or commission results in a relative error of 1.0 or greater; these are due to failure of the VLM to isolate a tree or to edge effects in the circular plots where unmeasured trees are included in the VLM.

dramatically reduce the spatial extent that the field crews are able to sample. The mean $\bar{h} \pm SD$ of the measured trees with DBH = 18.5 cm (the minimum cut-off for measurement in the full plot area) was $\bar{h} = 11 \pm 4$ m across the two study areas. The accuracy of the ALS derived ITD diverged from the observed ITD below 10 m height (Figs. 3–6a and 6b). This suggests that the trees >10 m in height isolated from the ALS by the VLM were likely to have been measured directly and not extrapolated from trees in the inner sub-plots. Further, it suggests that the field protocol for tree

measurements was sufficient for comparing the overstory distribution without introducing uncertainty errors related to the extrapolation of the understory distribution. Uncertainty in the extrapolated understory density is a particular issue in plots with open cover percentages. The spatial distribution of trees in open cover plots could be heterogeneous or asymmetric, e.g. 1/2 of the plot is densely forested while the other 1/2 is not forested at all. In such plots extrapolation of understory frequency could over or under predict the actual frequency distribution. For example, the

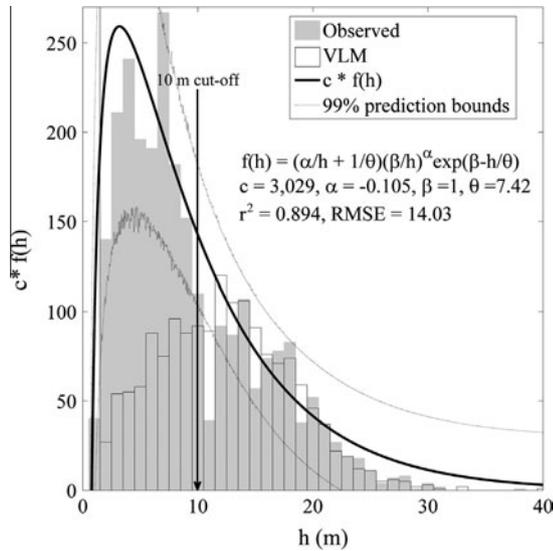


Fig. 7. Observed size-frequency distribution of trees in all plots >50% canopy cover in the Valles Caldera (gray). Histogram bars are 1 m wide linear bins. The VLM frequency (empty black) and the tapered Pareto (black line) fit using α and θ and normalization constant c determined by least squares, the β parameter was set at 1 m. The >10 m cut-off height (vertical dashed line) for the VLM was determined from the relative error rates of the validation data.

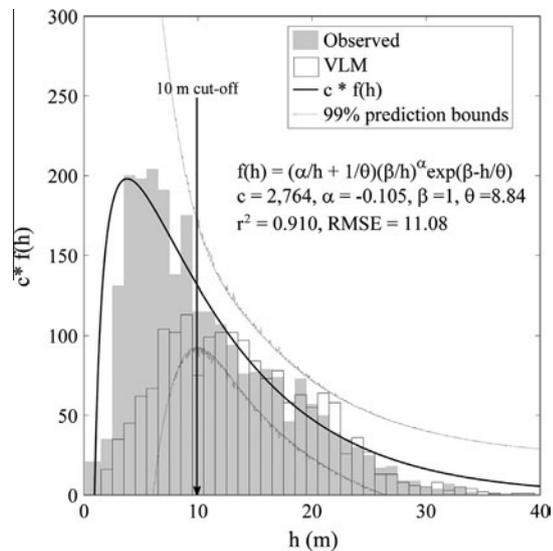


Fig. 8. Same as Fig. 7 but for the observed frequency of trees in all plots >70% canopy cover in the Pinaleno.

extrapolations of the understory ITD in the Valles Caldera (1/10th understory census) has greater uncertainty than in the Pinaleno (1/3rd understory census).

4.2. Limitations

The presence of understory trees directly beneath overstory canopies is a potential source of omission error when using a

CHM-based isolation technique. For some objectives, such as reconstruction of forest biomass estimates, the measurement uncertainty of tree size under normal growing conditions and other non-linear growth effects such as exceeding age and differing site indexes makes the omission of understory trees directly below an overstory tree a nominal issue, because total forest biomass tends to be dominated in the large trees. The problem is more significant for ecologists and foresters interested in determining population distributions, forest dynamics, biodiversity, and ecosystem status.

While each of the different CHM techniques developed to date offers strengths and weaknesses for census in various forest types, few demonstrate the ability to discriminate mid- or understory trees with a level of significance equivalent to a field inventory. Vector-based techniques which segment the point cloud are reported to have better accuracy in understory vs. the CHM techniques (Reitberger et al., 2009; Li et al., 2012; Yao et al., 2012). The vector approach appears to be a step forward in the census of forests, particularly with terrestrial laser scanning. However most ALS datasets have pulse densities that are too low to resolve features of the mid- to understory canopy architecture at landscape level.

Vauhkonen et al. (2012) and Kaartinen et al. (2012) compared the relative performances of different tree extraction techniques in different forest systems, e.g. European, Canadian, and USA plantation and natural forests, and found each technique performed best under the conditions for which it was developed. Our work was focused in semi-arid forests with varying levels of canopy closure as a result of topographic variation and landscape legacies. The ability to detect understory trees directly in a CHM is one distinctive feature of the two semi-arid forests we studied. Stands in semi-arid forests tend not to exhibit the multi-tiered canopy architecture found in more mesic forests, e.g. coastal Pacific temperate conifer forests. It was only in dense canopy cover where the enhancements of an ITC and area-based approach (Maltamo and Gobakken, 2014) to estimate the understory ITD < 10 m tall indirectly was necessary. Attention to changes in stand structure is important: if the forest is measured at space-appropriate scales, i.e. within similar high density canopy cover type stands, then the modeled ITD should fit the tapered Pareto that reflects the density dependence of the stand.

4.3. Significance and future research

Our findings address the problem identified in previous studies by Persson et al. (2002), Popescu and Wynne (2004), Chen et al. (2006), and Falkowski et al. (2006, 2008), which report a decrease in accuracy of isolation for the understory with increasing canopy cover. Notably, our study benefited from denser ALS data, a suggestion made in Falkowski et al. (2008), who reported a reduction in accuracy with increased tree clumping.

We suggest that a continuous distribution function, the tapered Pareto, is a valid alternative distribution to other stretched exponential functions previously reported in the literature (Maltamo et al., 2004; Maltamo and Gobakken, 2014). A benefit of using the tapered Pareto is that a change in the taper parameter affects the distribution in the right-tail and can thus reflect differential

Table 3

Least squares regression of the tapered Pareto to ALS overstory ITD [1 m linear bins] for the Pinaleno and Valles Caldera subset by plot's with high cover percentages. The $\hat{\alpha}$ (from the observed distribution) was fixed at -0.105 ; the β was fixed at 1 m.

CC%	Plot n	Tree n	Range	r ²	RMSE	\hat{c}	$\hat{\alpha}$	$\hat{\theta}$ (m)
Valles Caldera > 50%	23	2576	10 m < h < 38 m	0.894	14.03	3029	-0.105	7.424
Pinaleno > 70%	39	2463	10 m < h < 35 m	0.920	11.08	2764	-0.105	8.84

Table 4

Two-sample Kolmogorov–Smirnov (K–S) test results for: (1) the observed overstory CDF vs. the VLM derived tapered Pareto distribution (shown in Fig. 9), and (2) the observed understory CDF vs. the VLM derived ITD for the overstory fit by a tapered Pareto distribution (shown in Fig. 10). The *p*-values for all of the K–S tests are not significant; therefore we cannot reject them having different distributions.

	<i>p</i> -Value	K–S statistic
<i>Valles Caldera</i>		
Overstory Observed ITD vs. VLM tapered Pareto (10 m < <i>h</i> < 38 m)	0.153	0.296
Understory Observed ITD vs. VLM tapered Pareto (1 m < <i>h</i> < 10 m)	0.736	0.273
<i>Pinaleño</i>		
Overstory Observed ITD vs. VLM tapered Pareto (10 m < <i>h</i> < 35 m)	0.995	0.107
Understory Observed ITD vs. VLM tapered Pareto (1 m < <i>h</i> < 10 m)	0.111	0.500

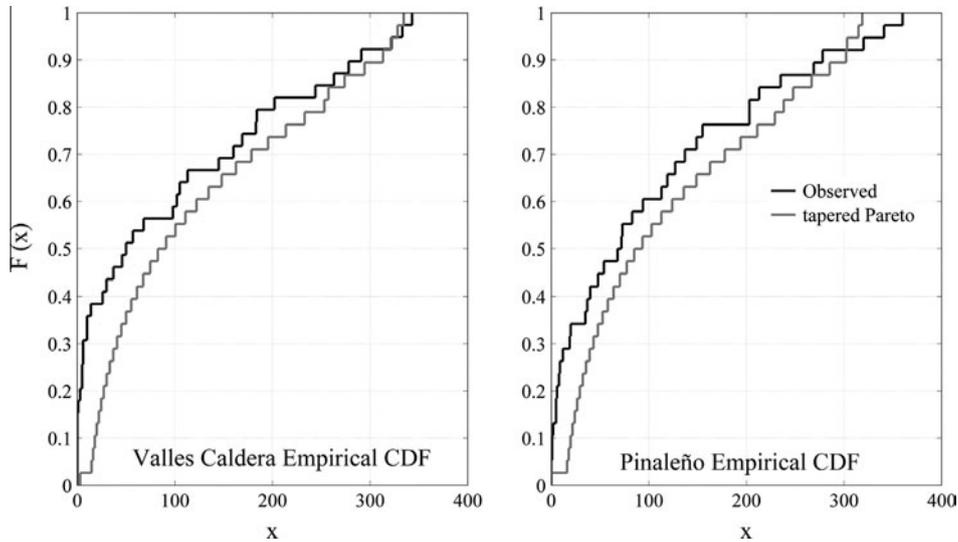


Fig. 9. Cumulative distribution function for the observed overstory height ITD vs. the ALS tapered Pareto distribution for Pinaleño and Valles Caldera.

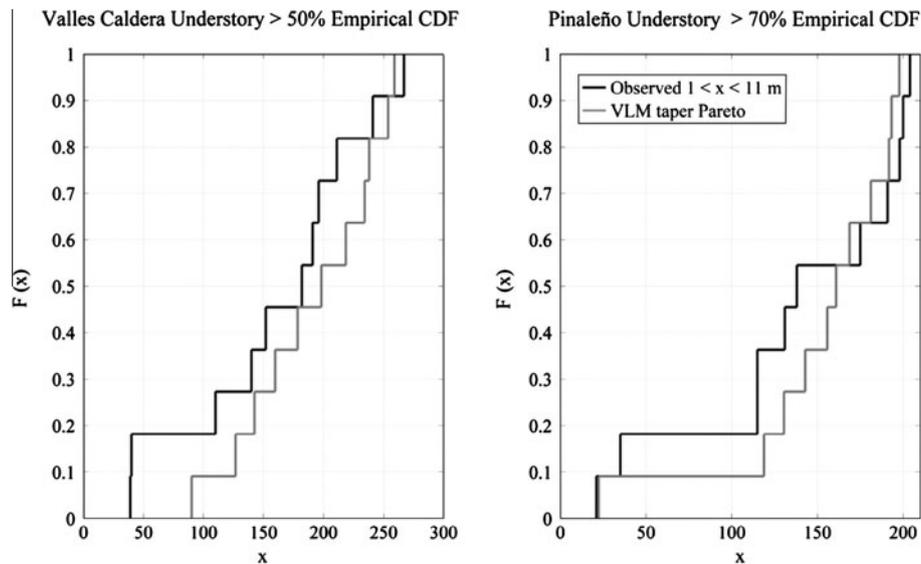


Fig. 10. Cumulative distribution functions for observed understory vs. VLM modeled distribution using only overstory trees >10 m in the Valles Caldera and Pinaleño.

rates of decline in abundance with increasing tree size classes, i.e. hyper-exponential, exponential, or sub-exponential declines. This is an important feature for analyzing very large distributions which become upper-truncated in the right-tail at large sample sizes, e.g. landscape level inventories.

In our observed data, the ITD of closed canopy cover was well approximated by the tapered Pareto (Figs. 9 and 10). These findings are consistent with Maltamo et al. (2004), who used a truncated Weibull distribution to predict the understory frequency from overstory ITD derived from ALS in Scandinavian pine stands.

Both the tapered Pareto and truncated Weibull distribution exhibit good fits for our data when the entire distribution is used (SM Fig. 6). In our example, however, the tapered Pareto had a slightly better fit in the understory when only the overstory ITD was used in a regression and the shape parameter was set to the known total ITD shape parameter value (SM Fig. 6). A thorough evaluation of the benefits and weaknesses of the truncated Weibull vs. the tapered Pareto would be a useful line of inquiry for future research.

5. Conclusions and applications

We were able to isolate individual trees in the ALS when canopy cover was relatively low and/or trees were >10 m in height. The height ITD fit a tapered Pareto function for both field observations and ALS-derived inventories, and the tapered Pareto scale parameters were similar between study areas, and canopy cover categories. We conclude that understory tree size distributions can be estimated based on overstory inventories when canopy cover is accounted for. These results are a further demonstration of the potential for ALS inventories using the combination of ITD and area-based approaches (Maltamo and Gobakken, 2014). Our findings extend this work into semi-arid forests where: (1) for open canopy cover the ALS derived ITD are characterized robustly across all size classes because of the open spacing between trees; (2) in closed canopy cover stands the ITD tends to exhibit density-dependent spatial sorting and packing well characterized by a tapered Pareto; and (3) the frequency of understory predicted by the overstory ALS derived ITD parameter values for a tapered Pareto distribution are not significantly different than the ITD for the observed understory. These results also support the MST prediction that density dependence among individuals drives self-similar rank-size frequency of the entire stand (Deng et al., 2012; Enquist and Niklas, 2001, 2002; Enquist et al., 2009; West et al., 2009).

These results may help forest ecologists refine their techniques for mapping rank-size ITD in semi-arid forests. Managers and scientists interested in generating virtual inventories of forests based on ALS can parameterize the ITD with local field observations to account for the differences in stand density indices of their forest types. For managers, ALS-derived inventories of forest stands can help in monitoring forest conditions and processes without expending scarce resources on field monitoring teams who can measure only tiny portions of complex landscapes. For ecologists, the benefit of having landscape-level inventories opens up new questions about the distribution of mass and energy flux utilization in forests. The scale of a complete landscape level ALS inventory removes the need for statistical extrapolation and expands our ability to understand pattern, structure, and distributions at both large spatial extents and fine spatial resolution.

Acknowledgements

We thank Christopher 'Kit' O'Connor and Craig Wilcox for providing the data for the Pinaleno. We also thank the several anonymous reviewers and the editor for their suggestions, which significantly improved the manuscript. The Pinaleno LiDAR Project was funded by the U.S. Forest Service, Coronado National Forest, Pinaleno Ecosystem Restoration Project; The University of Arizona; and the Nature Conservancy. The Valles Caldera LiDAR project was funded by the SCM-JRB-CZO (NSF Award#0724958). Additional support was provided by the U.S. Forest Service, Coronado N.F. A final thanks to Sherry Tune, USFS, for her vision and belief in the greatest good.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.foreco.2014.07.011>.

References

- Bailey, R.L., Dell, T.R., 1973. Quantifying diameter distributions with the Weibull function. *Forest Sci.* 19 (2), 97–104.
- Breidenbach, J., Astrup, R., 2014. The semi-individual tree crown approach. In: *Forestry Applications of Airborne Laser Scanning*. Springer, Netherlands, pp. 113–133.
- Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogramm. Eng. Remote Sens.* 72, 923–932.
- Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distribution in empirical data. *SIAM Rev. Soc. Ind. Appl. Math.* 51, 661–703.
- Deng, J., Zuo, W., Wang, Z., Fan, Z., Ji, M., Wang, G., Ran, J., Zhao, C., Liu, J., Niklas, K.J., Hammond, S.T., Brown, J.H., 2012. Insights into plant size-density relationships from models and agricultural crops. *Proc. Nat. Acad. Sci.* 109, 8600–8605.
- Dralle, K., Rudemo, M., 1996. Stem number estimation by kernel smoothing of aerial photos. *Can. J. For. Res.* 26, 1228–1236.
- Dubey, S.D., 1967. Some percentile estimators for Weibull parameters. *Technometrics* 9, 119–129.
- Enquist, B.J., Brown, J.H., West, G.B., 1998. Allometric scaling of plant energetics and population density. *Nature* 395, 163–165.
- Enquist, B.J., Niklas, K.J., 2001. Invariant scaling relations across tree-dominated communities. *Nature* 410, 655–660.
- Enquist, B.J., Niklas, K.J., 2002. Global allocation rules for patterns of biomass partitioning in seed plants. *Science* 295, 1517–1520.
- Enquist, B.J., West, G.B., Brown, J.H., 2009. Extensions and evaluations of a general quantitative theory of forest structure and dynamics. *Proc. Nat. Acad. Sci.* 106, 7046–7051.
- Enquist, B.J., West, G.B., Charnov, E.L., Brown, J.H., 1999. Allometric scaling of production and life-history variation in vascular plants. *Nature* 401, 907–911.
- Falkowski, M.J., Smith, A.M., Hudak, A.T., Gessler, P.E., Vierling, L.A., Crookston, N.L., 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. *Can. J. Remote Sens.* 32, 153–161.
- Falkowski, M.J., Smith, A.M., Gessler, P.E., Hudak, A.T., Vierling, L.A., Evans, J.S., 2008. The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. *Can. J. Remote Sens.* 34, S338–S350.
- Frazer, G., Magnussen, S., Wulder, M., Niemann, K., 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sens. Environ.* 115, 636–649.
- Gatzliolis, D., Fried, J.S., Monleon, V.S., 2010. Challenges to estimating tree height via LiDAR in closed-canopy forests: a parable from western Oregon. *For. Sci.* 56, 139–155.
- Hudak, A.T., Evans, J.S., Stuart Smith, A.M., 2009. LiDAR utility for natural resource managers. *Remote Sens.* 1 (4), 934–951.
- Hyyppä, J., Kelle, O., Lehto, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *Geosci. Remote Sens.* 39, 969–975.
- Hyyppä, J., Mielonen, T., Hyyppä, H., Maltamo, M., Yu, X., Honkavaara, E., Kaartinen, H., 2005. Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds. In: *Proceedings of ISPRS workshop laser scanning*. Proceedings of ISPRS Workshop Laser Scanning 2005, 12–14 September 2005, Enschede, Netherlands (Netherlands: GITC bv), XXXVI, 3/W19, pp. 144–149.
- Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., et al., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sens.* 4 (4), 950–974.
- Kagan, Yan Y., Schoenberg, F., 2001. Estimation of the upper cutoff parameter for the tapered Pareto distribution. *J. Appl. Probability*, 158–175.
- Kempes, C.P., West, G.B., Crowell, K., Girvan, M., 2011. Predicting maximum tree heights and other traits from allometric scaling and resource limitations. *PLoS ONE* 6, e20551.
- Koch, B., Kattenborn, T., Straub, C., Vauhkonen, J., 2014. Segmentation of forest to tree objects. In: Maltamo, M., Næsset, E., Vauhkonen, J. (tech. Eds.), *Forestry Applications of Airborne Laser Scanning*. Springer, Netherlands, pp. 89–112.
- Laes, D., Reutebuch, S., McGaughey, B., Maus, P., Mellin, T., Wilcox, C., Anhold, J., Finco, M., Brewer, K., 2008. Practical LiDAR acquisition considerations for forestry applications. RSAC-0111-BRIEF1. U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. Salt Lake City, UT, 7p.
- Laes, D., Mellin, T., Wilcox, C., Anhold, J., Maus, P., Falk, D.A., Koprowski, J., Drake, S., Dale, S., Fisk, H., Joria, P., Lynch, A.M., Alanen, M., 2009. Mapping vegetation structure in the Pinaleno Mountains using LiDAR. RSAC-0118-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 22p. 173–181.

- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sens. Environ.* 70, 339–361.
- Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. Lidar remote sensing for ecosystem studies: lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *Bioscience* 52, 19–30.
- Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. *Photogramm. Eng. Remote Sens.* 78, 75–84.
- Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., Olsson, H., 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *Int. J. Remote Sens.* 31, 1175–1192.
- Maltamo, M., Kangas, A., Uutera, J., Tornaiainen, T., Saramäki, J., 2000. Comparison of percentile based prediction methods and the Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. *For. Ecol. Manage.* 133 (3), 263–274.
- Maltamo, M., Erikäinen, K., Pitkänen, J., Hyyppä, J., Vehmas, M., 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sens. Environ.* 90 (3), 319–330.
- Maltamo, M., Gobakken, T., 2014. Predicting tree diameter distributions. In: Maltamo, M., Næsset, E., Vauhkonen, J. (tech. Eds.), *Forestry Applications of Airborne Laser Scanning*. Springer, Netherlands, pp. 177–191.
- Maltamo, M., Næsset, E., Vauhkonen, J. (tech. Eds.), 2014. *Forestry Applications of Airborne Laser Scanning*. Springer Netherlands, 460p.
- Massey Jr., F.J., 1951. The Kolmogorov–Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* 46, 68–78.
- Mathworks, 2012. MATLAB version 8.0.0783 (R2012B) Natick, Massachusetts: The Mathworks Inc.
- McGaughey, R., 2012. FUSION/LDV: Software for LIDAR Data Analysis and Visualization, Version 3.01. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, University of Washington. Available online at: <<http://forsys.cfr.washington.edu/fusion/latest.html>> (last accessed 24.08.12).
- McMahon, T., 1973. Size and shape in biology. *Science* 179, 1201–1204.
- McMahon, T.A., Kronauer, R.E., 1976. Tree structures: deducing the principle of mechanical design. *J. Theor. Biol.* 59, 443–466.
- Monnet, J., Mermin, É., Chanussot, J., Berger, F., 2010. Tree top detection using local maxima filtering: a parameter sensitivity analysis. In: 10 th International Conference on LiDAR Applications for Assessing Forest Ecosystems, 14–17 September 2010, Freiburg, Germany. *Silvilaser* 2010.
- Muldavin, E., Neville, P., Jackson, C., Neville, T., 2006. A vegetation map of the Valles Caldera National Preserve, New Mexico. Final report for National Park Service Award No. 1443-CA-1248-01-001 and Valles Caldera Trust Contract No. VCT-TO 0401.
- Newman, M.E., 2005. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* 46, 323–351.
- Niering, W.A., Lowe, C.H., 1984. Vegetation of the Santa Catalina Mountains: community types and dynamics. *Vegetatio* 58, 3–28.
- Niklas, K.J., Enquist, B.J., 2001. Invariant scaling relationships for interspecific plant biomass production rates and body size. *Proc. Nat. Acad. Sci.* 98, 2922–2927.
- Niklas, K.J., Midgley, J.J., Rand, R.H., 2003. Tree size frequency distributions, plant density, age and community disturbance. *Ecol. Lett.* 6, 405–411.
- O'Connor, C.D., 2013. Spatial and Temporal Dynamics of Disturbance Interactions Along an Ecological Gradient. Dissertation. University of Arizona. 204pp.
- O'Connor, C.D., Falk, D.A., Lynch, A.M., Swetnam, T.W., in press. Fire Severity, Size, and Climate Associations Diverge from Historical Precedent along an Ecological Gradient in the Pinaleno Mountains, Arizona, U.S.A. *Forest Ecology and Management*.
- Pareto, V., 1896. *Cours d'économie Politique*, reprinted in 1965 as a volume of *Oeuvres Complètes*. Droz, Geneva.
- Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogramm. Eng. Remote Sens.* 68, 925–932.
- Popescu, S.C., Wynne, R.H., 2004. Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogramm. Eng. Remote Sens.* 70, 589–604.
- Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Comput. Electron. Agric.* 37, 71–95.
- Reineke, L.H., 1933. Perfecting a Stand-Density Index for Even-Aged Forests. *Agric. Res.* 46, 627–638.
- Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. *Photogramm. Remote Sens.* 64 (6), 561–574.
- Savage, V., Bentley, L., Enquist, B., Sperry, J., Smith, D., Reich, P., von Allmen, E., 2010. Hydraulic trade-offs and space filling enable better predictions of vascular structure and function in plants. *Proc. Nat. Acad. Sci.* 107, 22722–22727.
- Schoenberg, F.P., Patel, R.D., 2012. Comparison of Pareto and tapered Pareto distributions for environmental phenomena. *Euro. Phys. J. Special Topics* 205, 159–166.
- Seuront, L., Mitchell, J.G., 2008. Towards a seascape typology. I. Zipf versus Pareto laws. *J. Mar. Syst.* 69 (3), 310–327.
- Swetnam, T.L., 2013. *Cordilleran Forest Scaling Dynamics and Disturbance Regimes Quantified by Aerial LiDAR*. Dissertation, University of Arizona, 276pp.
- Swetnam, T.L., Falk, D.A., 2014. A variable-area local maxima tool for segmenting individual trees from aerial LiDAR: allometric scaling rules to reduce error in individual tree segmentation. *For. Ecol. Manage.* 323, 158–167.
- Taubert, F., Hartig, F., Dobner, H.J., Huth, A., 2013. On the challenge of fitting tree size distributions in ecology. *PLoS ONE* 8 (2), e58036.
- Van Leeuwen, M., Nieuwenhuis, M., 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *Eur. J. For. Res.* 129, 749–770.
- Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., et al., 2012. Comparative testing of single-tree detection algorithms under different types of forest. *Forestry* 85 (1), 27–40.
- Wang, J., Tsang, W.W., Marsaglia, G., 2003. Evaluating Kolmogorov's distribution. *J. Stat. Software* 8(18).
- Weibull, W., 1951. A statistical distribution function of wide applicability. *J. Appl. Mech. – Trans. ASME* 18(3), 293–297.
- West, G.B., Enquist, B.J., Brown, J.H., 2009. A general quantitative theory of forest structure and dynamics. *Proc. Nat. Acad. Sci.* 106, 7040–7045.
- Westoby, M., 1984. The self-thinning rule. *Adv. Ecol. Res.* 14, 167–226.
- White, J., 1981. The allometric interpretation of the self-thinning rule. *J. Theor. Biol.* 89 (3), 475–500.
- White, E.P., Enquist, B.J., Green, J.L., 2008. On estimating the exponent of power-law frequency distributions. *Ecology* 89, 905–912.
- Whittaker, R.H., Niering, W.A., 1975. Vegetation of the Santa Catalina Mountains, Arizona. V. Biomass, production, and diversity along the elevation gradient. *Ecology*, 771–790.
- Wing, B.M., Ritchie, M.W., Boston, K., Cohen, W.B., Gitelman, A., Olsen, M.J., 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sens. Environ.* 124, 730–741.
- Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sens. Environ.* 123, 368–380.
- Yoda, K.T., Kira, H., Ogawa, H., Hozumi, K., 1963. Self-thinning in over-crowded pure stands under cultivated and natural conditions. (Intraspecific competition among higher plants. XI.). *J. Biol. Osaka City Univ.* 14, 107–129.
- Zhao, K., Popescu, S., 2007. Hierarchical watershed segmentation of canopy height model for multi-scale forest inventory. In: *Proceedings of the ISPRS working group "Laser Scanning"* pp. 436–442.
- Zhao, K., Popescu, S., Nelson, R., 2009. Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. *Remote Sens. Environ.* 113 (1), 182–196.
- Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., Gerard, P.D., 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sens. Environ.* 87 (2), 171–182.