Canopy clumping appraisal using terrestrial and airborne laser scanning

Mariano García a,⁎, John Gajardo b,c, David Riaño a,d,e, Kaiguang Zhao f, Pilar Martín d,e, Susan Ustin a

a Center for Spatial Technologies and Remote Sensing (CSTARS), University of California, Davis, One Shields Avenue, 139 Veihmeyer Hall, Davis, CA 95616, USA
b Faculty of Forest Sciences, University of Talca, 2 Norte # 685, Talca, Chile
c Department of Geology, Geography and Environment, University of Alcalá, Alcalá de Henares, 28801 Madrid, Spain
d Institute of Economics, Geography and Demography, Spanish National Research Council (CSIC), Albacete 26–28, 28037 Madrid, Spain
e Associated Research Unit GEOLAB, University of Alcalá – Spanish National Research Council (CSIC), Alcalá de Henares, 28801/Albacete 26–28, 28037 Madrid, Spain
f School of Environment and Natural Resources, Ohio Agricultural Research and Development Center, The Ohio State University, Wooster, OH 44691, USA

A R T I C L E   I N F O

Article history:
Received 5 August 2014
Received in revised form 29 January 2015
Accepted 31 January 2015
Available online xxxx

Keywords:
Canopy clumping
Airborne laser scanner
Terrestrial laser scanner
Voxel
Leaf area index

A B S T R A C T

Accurate spatial information of canopy clumping degree (Ω) contributes to better understanding of the light regime within the canopy and the physiological processes associated with it. This paper evaluates the potential of terrestrial (TLS) and airborne laser scanning (ALS) to estimate Ω in different vegetation types after converting the point cloud into a 3-dimensional (3D) voxel-based model. Three methods are presented based on the spatial distribution of the returns (Standardized Morisita’s Index – SMI), the gap distribution (Pielou’s coefficient of segregation – PCS) and the gap size distribution (Chen & Cihlar’s clumping index – CCI). Compared to Ω values derived from hemispherical photographs (HPs), the CCI method outperformed PCS and SMI for both instruments, with a correlation value of 0.93 (vs. 0.79 – PCS and 0.65 – SMI) for oak trees using TLS; 0.83 (vs. 0.78 – PCS and 0.73 – SMI) for a shrub chaparral using ALS data; and 0.84 (vs. 0.81 – PCS and 0.50 – SMI) for a mixed Mediterranean forest using ALS data. Voxel size was an important parameter to estimate Ω showing statistically significant differences for the different resolutions tested. Voxel size had an opposite effect on SMI than that on PCS and CCI, with SMI providing better results for coarser voxel sizes, and PCS and CCI yielding higher accuracies for finer voxels. In the case of the TLS, the influence of the zenith angle was also evaluated by means of a Kruskal–Wallis test. CCI and PCS did not show significant differences among the zenith angles tested, but SMI did. The radius of the plot used to analyze ALS data significantly affected the correlations with HP, with the best results found at 13, 7 and 15 m for mixed Mediterranean forest and at 11, 10 and 5 m for shrubs for CCI, PCS and SMI, respectively. The methods presented have the potential to be operationally applied to other areas using TLS and ALS data, since they are not based on an empirical fit but on the analysis of the gap size in the canopy and the distribution of returns after voxelization.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The interactions between the vegetation and the atmosphere, the light radiation regimes through the vegetation canopies and the carbon and water cycles depend significantly on the canopy architecture (Chen & Cihlar, 1995; Chen, Menges, & Leblanc, 2005; Dickinson, Henderson-Sellers, Kennedy, & Wilson, 1986; Sellers et al., 1997). Leaf area index (LAI), defined as total one sided leaf area per unit ground surface area (Chen & Black, 1992) is a key variable to describe canopy architecture. Indirect optical methods like hemispherical photographs (HPs) estimate LAI by measuring the diffuse light that penetrates through the canopy (Chen et al., 2006), under the assumption of random distribution of the foliage elements. Nevertheless, real canopies are seldom randomly distributed, so the effective LAI (eLAI) is obtained instead (Chen, Black, & Adams, 1991):

\[ \text{eLAI} = \Omega \text{LAI} \] (1)

where Ω is the dispersion coefficient or clumping index, which accounts for the spatial distribution patterns of foliage (Nilson, 1971). Ω = 0 represents a completely clumped canopy, Ω = 1 a completely random one; and Ω > 1 represents a uniform canopy.

Ω also significantly alters the vegetation radiation environment, allows accurate separation of sunlit and shaded leaves in the canopy (Chen et al., 2005) and provides valuable information to explain the spatial variability of processes influenced by the light regime such as photosynthesis, evapotranspiration or the distribution of foliar nutrients within the canopies (Thomas, Noland, Treitz, & McCaughey, 2011).

Based on optical remotely sensed data, Chen et al. (2005) and He, Chen, Pisek, Schaaf, and Strahler (2012) generated global Ω products using multi-angular POLDER 1 and a combination of MODIS and POLDER 3 data, respectively. Nevertheless, the ability of these passive optical sensors to provide structural information is rather limited as they only provide information on the horizontal distribution of the integrated reflected energy from vegetation canopies. Instead, terrestrial (TLS), airborne (ALS) and satellite (SLS) laser scanning offer detailed information on both the horizontal and vertical distribution of vegetation...
canopy, although most studies have focused on the LAI estimation (García et al., 2012; Hopkinson et al., 2013; Hosoi & Omasa, 2006; Strahler et al., 2008) and few have addressed Ω. Using TLS, Moorthy et al. (2011) calculated Ω based on Chen and Cihlar’s (1995) approach by computing the ratio of the gap fraction between the measured crown and that of a simulated random canopy. The latter was simulated by randomly redistributing the laser returns within the boundary space of the canopy derived from the TLS data. Although no specific validation was performed on the Ω estimates, a clear improvement on LAI estimates was achieved in comparison to field estimates from the LAI-2000 Plant Canopy Analyser. Zhao et al. (2012) also used the gap size probability obtained by means of a Nominal Spatial Extent Index, derived from images of gap probabilities yielded by the TLS Echidna Validation Instrument, to estimate Ω. Their results showed good statistical agreement with those obtained with HP (R² = 0.866). Thomas et al. (2011) derived Ω from ALS data by regressing common height distribution metrics like height percentiles, and the gap fraction estimated from the proportion of ground returns to all returns.

Given the ecophysiological importance of having accurate information on the spatial distribution of Ω, this study aimed at implementing different methods to estimate Ω from TLS and ALS data based on the spatial distribution of the laser returns and of the gap size after converting the point cloud into a 3-dimensional (3D) voxel-based model. The study also assessed the impact of the voxel size; the effect of the zenith angles selected for the TLS; and the plot radius selected from the ALS data on the Ω estimates as compared to those obtained from HP.

2. Methods

2.1. Study areas

To provide useful insights on the robustness of the algorithms, the same methods to estimate Ω were applied to two study areas, Majadas and Jasper Ridge, with different canopy architecture and either TLS or ALS data, respectively. Majadas, a type of savannah called “dehesa”, is located in the Tiétar Valley, Cáceres, Extremadura, west of Spain (39°56′ 27″ N; 5°46′ 27″ W, Fig. 1-A). The area is a Holm oak (Quercus ilex, L.) wooded grassland managed for grazing with a Mediterranean climate characterized by a hot and dry summer and low precipitation (400–800 mm annually) concentrated in spring and autumn. The tree density is about 20 trees/ha, with an average height of 8 m and 0.4 m diameter at breast height.

Jasper Ridge, is located on the northeastern foothills of the Santa Cruz Mountains in Portola Valley, San Mateo County, California, USA (37°24′ N; 122°13′ 30″ W, Fig. 1-B). It is a biological preserve of about 500 ha used for scientific studies owned by Stanford University (http://jrbp.stanford.edu/, last accessed November 28th, 2014). This site also has a Mediterranean climate with an annual average precipitation of about 650 mm (Dahlin, Asner, & Field, 2011). The area is mainly covered by shrubland, evergreen and deciduous forest, herbaceous perennial wetlands and annual grasslands (García & Ustin, 2001). The shrub vegetation is composed of typical Chaparral species (see Casas, Riaño, Ustin, Dennison, & Salas, 2014) and the forest includes blue oak (Quercus douglasii Hook. & Arn.), California live oak (Quercus agrifolia Née); and valley oak (Quercus lobata Née).

2.2. Datasets

2.2.1. TLS data acquisition

Eight TLS point clouds, which were analyzed independently for Ω were acquired for four Holm oak trees at Majadas on October 9th, 2009, using a Leica HDS-6000 TLS (Leica Geosystems, Switzerland, www.leica-geosystems.com). Each tree, separated at least 7–8 m from each other, was scanned twice in diametrically opposite positions from the trunk. The instrument has a beam diameter of 3 mm at the exit and a beam divergence of 0.22 mrad that resulted in a footprint diameter of 5.2 mm at 10 m range. Scans were carried out on dry, clear days and under calm conditions which limited the noise/errors caused by wind gusts moving the leaves and branches of the crowns.

The TLS instrument was deployed under the canopy roughly in between the tree trunk and the edge of the canopy to capture as much of the crown as possible, while reducing the occlusion caused by the trunk. Each scan provided a horizontal perspective of 360° in azimuth (θ) and 154° in zenith (Φ), with an angular sampling interval of approximately 6.3 mm at a 10 m distance. The location of each scan was determined using differential GPS positioning and topographic (total station) methods.

The Cyclone software (Leica Geosystems, Switzerland, http://www.leica-geosystems.com) was used for post-processing and generating the point cloud of first returns. Following Seidel, Fleck, and Leuschner (2012), possible noisy returns were filtered out when their intensity was < 0.01 or > 1.0 or their range was > 80 m, or when they were separated by less than 1.6 mm from one another. Subsequently, the TLS point cloud from the tree crown was isolated from the rest of the TLS scene and exported as XYZ ASCII files.

Fig. 1. Location of the study areas: A) Majadas, Extremadura, Spain. B) Jasper Ridge, California, USA.

Please cite this article as: García, M., et al., Canopy clumping appraisal using terrestrial and airborne laser scanning, Remote Sensing of Environment (2015), http://dx.doi.org/10.1016/j.rse.2015.01.030
2.2.2. ALS data acquisition

ALS data was collected using the Carnegie Airborne Observatory (CAO) Beta sensor package (Asner et al., 2007), which was flown over Jasper Ridge in August of 2007 at an average 2700 m above ground level. The pulse repetition frequency was 33 kHz with a maximum scan angle of 36°, rendering a laser point spacing of 0.5 m and a footprint of 1.5 m. The ALS data was provided by CAO as discrete-return (up to 4) laser points. An exploratory analysis of the ALS data revealed the existence of laser returns with the exact same X, Y, and Z and intensity values which were removed for subsequent analyses. The filtered ALS dataset contained a final average laser pulse density of 4 points/m².

After verification of the ALS positional accuracy using building roofs (<50 cm in X, Y; and <10 cm in Z), the laser returns were classified into ground and non-ground using the algorithm developed by Streutker and Glenn (2006) and implemented in the BCAL LiDAR Tools software (BCAL). The height above the ground surface was calculated for all non-ground points and 5–20 m plot sizes at 1 m height intervals were extracted in order to identify the scale at which the relationship between HP and ALS data was the best.

2.2.3. Reference data

HP was used to validate the Ω estimates derived from TLS and ALS based on two of the methods implemented in CIMES (Gonsamo, Walter, & Pellikka, 2011), namely Chen and Cihlar’s (CCI) and Piélou’s coefficient of segregation (PCS).

A HP was taken immediately after each TLS acquisition in Majadas using a Canon camera EOS 400D with an APS-C sensor and Sigma 8 mm f/3.5 fisheye lens. The HP camera was attached to the tripod of the TLS with an adapter to minimize the impact of any possible mis-registration between them. The Field of View (FOV) of the fisheye lens was 180°, but the smaller camera image sensor cropped it to 120°, equivalent to a 0–60° zenith range. A second reference dataset was created by converting the TLS into pseudo HP (TLS-HP), which were also analyzed in CIMES as if they were conventional HP (Gajardo, 2014). In doing so, we evaluated the influence of the intrinsic differences between the 3D information provided by TLS and the 2D information yielded by HP to estimate Ω.

For Jasper Ridge, HPs were collected using a digital camera Nikon Coolpix 4300 with a Nikon FC-E8 fisheye lens converter with a maximum FOV of 180°. Thirteen forest and thirteen shrub plots were sampled in three different campaigns between 12–17 May 2006, 20–21 September 2006 and 13–14 August 2007. Similarly to Casas et al. (2014), who estimated LAI using the same HP dataset, the average Ω of the three dates was preferred to the closest date to the ALS flight in order to minimize the effect of differences in illumination conditions between plots at the time of HP acquisition. In addition, there were no statistical differences (p-values > 0.05) using the non-parametric Kruskal–Wallis (K–W) test between the Ω values derived for each plot in the three sampling dates. Five HPs were taken at the center of the plot and 2.5 m away in each of the four cardinal directions (N, S, E, W). All HP were collected under indirect sunlight conditions, to minimize glare from direct sunlight that could cause saturation of the image, also known as blooming effect, making canopy gaps larger than they really are (Leblanc, Chen, Fernandes, Deering, & Conley, 2005). The HPs were overexposed, by 2 F-stops below the automatic reference exposure to improve the contrast between sky and leaves. HP captured only the overstory vegetation using a tripod to elevate the camera approximately 0.7 m for forest plots, whereas for shrub plots it was placed at 0.2 m and at least 0.5 m from any vegetation, which given the small size of the leaves reduced the effects of leaf geometry on HP gap estimates. The lens was leveled to point toward zenith and was oriented so that the magnetic north was located at the top of the HP (Walter, Fournier, Soudani, & Meyer, 2003). Each plot center was geo-located using differential GPS positioning methods.

HPs were transformed into a binary image of canopy elements (black pixels) and gaps or sky (white pixels). First, the blue channel was extracted since in this band canopy elements appear darker; hence, the contrast between canopy objects and sky is greater (Gonsamo, Walter, & Pellikka, 2010; Leblanc et al., 2005). Subsequently, an edge enhancement 3 × 3 filter proposed by Kucharik, Norman, Murdock, and Gower (1997) was applied to enhance small gaps within dense foliage since digital HPs have poor sharpness in comparison with film cameras, for which many of the fine canopy structural details are poorly defined (Frazer, Fournier, Troyfomow, & Hall, 2001). The applied filter also facilitated finding an optimal threshold to create the binary image, which was selected automatically based on a clustering algorithm developed by Ridler and Calvard (1978).

2.3. Data voxelization

TLS and ALS laser returns were grouped in voxels to identify occupied/non-occupied ones. The effect of the voxel size on Ω was evaluated by choosing 5–25 cm voxels at 5 cm intervals for the TLS data (Fig. 2, A–C). Additionally, a 3 cm resolution was tested as it was found by García et al. (2011) to be the optimum resolution to estimate fractional cover from TLS data. Coarser resolutions of 30–100 cm at 10 cm intervals were tested for the ALS data due to its much lower laser pulse density (Fig. 2D–F).

The variability of Ω is a function of the zenith angle (Gonsamo & Pellikka, 2009). Thus, the effect of zenith angle on Ω was evaluated considering the following zenith ranges on the TLS data: 0–55°; 0–60°; 5–55°; 5–60°; 30–60° and 55–60°. Analyses were restricted to 60°, which was the maximum zenith angle for the HP in Majadas. Angles close to the zenith present high variability, affecting the estimation of the gap fraction (Walter et al., 2003). To assess this effect, two sets of data were analyzed excluding the 0–5° zenith angle range. The range 30–60° was selected because it provides little variation of Ω with view zenith angle and the mean Ω for this zenith range is similar to that of all zenith angles (Gonsamo & Pellikka, 2009; Leblanc et al., 2005). Finally, the 55–60° or hinge angle was selected because it has been recommended to estimate LAI, as at that zenith range the gap probability is nearly independent of the leaf angle distribution (Welles & Norman, 1991).

2.4. Standardized Morisita’s Index (SMI)

The SMI (Morisita, 1962) is based on a measure of dispersion of individuals in a population that is commonly used in ecological studies, Morisita’s index of dispersion (I_M) (Morisita, 1959):

\[
I_M = n \left[ \sum x_i^2 - \frac{X^2}{n} \right] / \left( \sum x_i - \frac{1}{X} \right)
\]

(2)

where \(n\) is the sample size (number of voxels) and \(X\) is the number of individuals (returns) within each voxel.

In order to compute the SMI it is necessary to determine two critical extreme values of \(I_M\), namely the uniform index (\(M_u\)) and the clumped index (\(M_c\)):

\[
M_u = \left[ \frac{X^2_{th} - n + \sum x_i}{\left( \sum x_i \right)^2 - 1} \right]
\]

(3)

\[
M_c = \left[ \frac{X^2_{th} - n + \sum x_i}{\left( \sum x_i \right)^2 - 1} \right]
\]

(4)

where \(X^2_{th}\) and \(X^2_{th}\) are the values of the Chi-square distribution with (\(n - 1\)) degrees of freedom that leaves 97.5% and 2.5% of the area to its...
right, respectively; $n$ is the sample size (number of voxels) and $x$ is the number of individuals within each voxel. Finally, the SMI is calculated as:

$$\text{SMI} = \begin{cases} 
0.5 + 0.5 \left( \frac{l_d - M_c}{n - M_c} \right), & l_d \geq M_c > 1 \\
0.5 \left( \frac{l_d - l}{M_c - l} \right), & M_c > l \geq 1 \\
-0.5 \left( \frac{M_c - l}{M_c - l} \right), & 1 > l > M_c \\
-0.5 + 0.5 \left( \frac{M_c - \mu}{\mu - M_c} \right), & 1 > \mu > l_d 
\end{cases}$$

The SMI ranges between $\pm 1$ and $\pm 0.5$ are the 95% confidence limits for random patterns. Nevertheless, for $\text{SMI} < 0$ the distribution tends to be uniform, for $\text{SMI} > 0$ the distribution tends to be clustered (clumped), whereas $\text{SMI} = 0$ represents a random one. Since these ranges are opposite to those used to estimate $\Omega$, they were transformed to represent $0 \leq \text{SMI} < 1$, which is a clumped distribution; $\text{SMI} = 1$ is a random canopy; and $\text{SMI} > 1$ is a uniform one. SMI is independent of both sample size and population density and is statistically testable, which makes it one of the most suitable dispersion indices available (Krebs, 1999).

SMI was computed for each height interval, defined by its voxel size, and weighted by the fractional cover of that interval and then integrated over the canopy. This weighting was applied to account for the effect of extremely high clumped values resulting from the canopy occlusion that causes a low proportion of returns reaching the upper parts of the canopy on the TLS or lower ones on the ALS data.

2.5. Pielou coefficient of segregation (PCS)

Pielou (1962) proposed a coefficient to estimate the degree of segregation between two species of plants growing together. The method records the sequences in which a given species of a two-species population occur in a sampling transect, and compare them to those expected from an unsegregated population of the same numerical

---

Fig. 2. TLS point cloud (A) and its voxelization at 5 (B) and 20 cm (C). ALS point cloud (D) and its voxelization at 30 (E) and 70 cm (F).
composition. This provides a measure of randomness for each species with respect to the other (Pielou, 1962), where \( a \) and \( b \) are respectively the probability of encountering species A and B for a two-species population, given that \( a + b = 1 \). For two randomly distributed populations; it follows with 95% probability that:

\[
\text{PCS} = \frac{\bar{a} + \bar{b}}{\mu_a + \frac{1}{\mu_b}} = 1 \pm 1.96 \sqrt{\frac{\sigma_a^2 + \sigma_b^2}{\mu_a^2}}
\]  

(6)

where \( \bar{a} \) and \( \bar{b} \) are the maximum likelihood estimates of \( A \) and \( B \), \( \mu_a \) and \( \mu_b \) are their mean length of occurrence, and \( \sigma_a^2 \) and \( \sigma_b^2 \) are their variances. If \( \text{PCS} < 1 \) the distribution of the species is clustered (clumped), if \( \text{PCS} > 1 \) the distribution is random and if \( \text{PCS} = 1 \) the distribution is uniform.

Walter et al. (2003) adjusted this method to HP, where the segregation of black and white pixels was computed on a circle with a constant zenith angle. Similarly we adapted the PCS method to the TLS data by computing the segregation of occupied/unoccupied voxels. As in the case of SMI, the PCS was applied for each height interval and integrated over the whole canopy. Due to the lower point density of ALS data and to diminish the effect of occlusion, the vertical stack of voxels for each location was converted to a single occupied/unoccupied voxel to produce a 2D binary image.

2.6. Chen & Cihlar’s clumping index (CCI)

CCI (Chen & Cihlar, 1995) estimates \( \Omega \) based on the gap size distribution. Although originally developed to be applied to the Tracing Radiation and Architecture of Canopies (TRAC) optical field instrument it has also been used with HP (Gonsamo & Pellikka, 2009; Walter et al., 2003). Clumped canopies present larger canopy gap fractions than random ones with the same LAI, as well as different gap size distributions. Thus, CCI removes large gaps iteratively to simulate a random canopy until no significant differences are found with the gap size cumulative distribution of the real canopy actually measured. In order to stop this iterative process, two threshold values were applied to the difference in gap fraction between the random and the real canopies, namely 0.1 and 0.05. Finally, the \( \Omega \) was estimated by comparing both gap size distributions:

\[
\Omega = \frac{\ln |F_m(0, 0)|}{\ln |F_{mr}(0, 0)|}
\]

(7)

where \( F_m(0, \theta) \) is the measured gap fraction and \( F_{mr}(0, \theta) \) is the gap fraction for an imaginary canopy with a random spatial distribution. The CCI equation proposed by Chen and Cihlar (1995) was subsequently modified by Leblanc (2002) introducing a compensation factor to account for the compactness of the canopy when large gaps are removed:

\[
\Omega = \frac{\ln |F_m(0, 0)|}{\ln |F_{mr}(0, 0)|} \left[ \frac{1 - F_{mr}(0, 0)}{1 - F_m(0, 0)} \right]
\]

(8)

This method was also adapted to the TLS and ALS data as described in Section 2.5.

3. Results

3.1. TLS \( \Omega \) estimates

SMI yielded moderate results when compared to HP-CCI, with better correlations for coarser voxel sizes (Fig. 3A). For all zenith angles, the 25 cm voxels yielded the best results whereas the finest voxels, 3 and 5 cm, gave the worst results with even negative correlations for all zenith ranges except for the hinge angle (55°–60°). Using the HP-PCS as reference (Fig. 3B), correlation values were approximately 20% higher for all resolutions and zenith ranges. TLS-HP-CCI also obtained higher correlations with SMI than HP-CCI and all of them were positive (Fig. 3C). This trend was observed for all voxel sizes and zenith angles except for 55°–60°, where correlations were similar for both cases. Compared to TLS-HP-PCS (Fig. 3D), correlation values were nearly 10% lower than for HP-CCI. Moreover, the highest correlation was found for 55°–60° followed by 5°–60°, whereas for the three previous reference methods the best zenith range was 5°–60°.

PCS obtained better, but still moderate, results than SMI compared to the HP-CCI (Fig. 3E). Contrary to SMI, the correlation improved for finer 3 and 5 cm voxels. Correlations between PCS and TLS-HP-CCI (Fig. 3G) were slightly lower for all cases, especially as voxel size increased and as before, the finer voxels performed the best. As for the zenith range, the results varied depending on the reference data used. The highest correlation with HP-CCI was obtained for the 0°–60°, although differences with 0°–55°, 5–55° and 5–60° were negligible regardless of the resolution. Instead, 55°–60° was the best for TLS-HP-CCI.

Surprisingly, when the PCS method was compared to HP-PCS and TLS-HP-PCS (Fig. 3F and H) the correlation values were approximately 15% lower than using HP-CCI and TLS-HP-CCI as reference data. The 10 cm voxels at 5–60° attained the best correlations for HP-PCS, whereas the 5 cm voxels at 55°–60° were best for TLS-HP-PCS.

The CCI method provided better results than SMI and PCS. The iterative gap removal with a 0.05 threshold offered better outcomes than 0.1, thus the results presented here refer to the former threshold value. Comparing HP-CCI (Fig. 3I), the best correlation was obtained for the 5 cm voxels at 55°–60°. Despite showing better correlation for the finer voxel sizes, the 10 cm voxels yielded slightly better results on an average basis for all zenith ranges. Using TLS-HP-CCI as reference data (Fig. 3J), the best correlations were found for the 5 cm voxels at the hinge angle range. Correlations were generally lower than for HP-CCI but at 55°–60°. Using the HP-PCS as reference (Fig. 3K), correlations were similar to HP-CCI for low zenith angles but decreased by 15% at 55°–60°, with the highest correlation found for 15 cm voxels at 0°–60°. Finally, compared to TLS-HP-PCS (Fig. 3L) the best correlation was obtained for the 5 cm voxels at 55°–60°, which was similar to the correlation found with HP-CCI and TLS-HP-CCI.

Table 1 shows the relationship obtained for each method at the best voxel size and zenith angle, based on the correlations found, along with the \( R^2 \) and Root Mean Square Error (RMSE). The CCI method provided the best agreement with the reference, followed by PCS and SMI. SMI also showed a large deviation from the 1:1 line whereas PCS and CCI showed a closer agreement with the 1:1 line when compared to CCI-reference data but large differences with the PCS-reference data. SMI provided lower \( \Omega \) values than HP-CCI and TLS-HP-CCI, whereas PCS and CCI also underestimated \( \Omega \) compared to HP-CCI but slightly overestimated \( \Omega \) compared to TLS-HP-CCI. All TLS methods gave \( \Omega \) values larger than those obtained by the references HP-PCS and TLS-HP-PCS.

Among the methods applied to TLS data, CCI offered higher \( \Omega \) values than PCS and SMI. This trend was observed for all zenith ranges considered. Still, a Mann–Whitney U test showed that differences between the values estimated by PCS and CCI applied to the TLS data were not statistically significant (p-value > 0.05) at any zenith range. Comparing SMI against CCI gave significant differences (p-value < 0.05) for all zenith ranges; whereas contrasted against PCS, SMI did not have significant differences except for the 30°–60° and 55°–60° zenith ranges.

3.2. Effect of the voxel size and zenith range on the \( \Omega \) estimation from TLS

The three methods tested showed a clear dependence on the voxel size used, although an opposite trend was observed for SMI compared to PCS and CCI. Thus, SMI obtained better results as voxel size increased, whereas the opposite occurred for PCS (Fig. 3). CCI showed a slightly different behavior, where the best correlation was for the 5–10 cm voxels at 55°–60°, decreasing for 3 cm and 15–25 cm, but the lower correlation
for the other zeniths remained relatively stable across different voxel sizes.

A K–W test applied to each method showed that differences in results obtained for each voxel size were statistically significant (p-value < 0.05). In addition, a multiple comparison test enabled to distinguish in general terms a small (3–5 cm) and a large (10–25 cm) voxel size for which the differences were not statistically significant. In the case of SMI, the 10–25 cm voxels also presented significant differences within them for all zeniths except for 55–60°. CCI also presented significant differences except within the 3–5 cm group.

A K–W test evaluated the effect of the zenith angle for the voxel size that provided the best results for each method tested. For CCI and PCS, no significant differences (p-value > 0.05) were found among Ω values at the zenith ranges evaluated and a multiple comparison test confirmed this fact. For the SMI method, differences among the Ω obtained at different zeniths were significant (p-value < 0.05). A multiple comparison test revealed differences between 55–60°, 30–60° and a third group encompassing all the other zeniths. In addition to the lack of statistical significant differences between zenith ranges, the higher influence of the voxel size was reflected in the fact that in many cases the best voxel size was found within the same or very similar zenith ranges.

3.3. ALS Ω estimates

SMI yielded moderate results for the ALS in the case of the shrubs when compared to HP-CCI and HP-PCS. The relationships generally worsen as plot radius and voxel size increased (Fig. 4A and B). Nevertheless, the correlation remained quite stable for 50 cm voxels throughout all plot radii analyzed. Whereas the smallest plot radius (5 m) rendered the highest correlation for most voxel sizes, the 7–10 m radii gave slightly higher correlations for finer voxels (30–40 cm). The same trends but slightly lower correlations were observed when comparing SMI to HP-PCS. Regarding the forests, poor results were generally achieved, and a more irregular pattern with plot size for all voxel sizes was observed (Fig. 4C and D). The 100 cm voxel size gave the highest correlation, which occurred for 11–16 m plot radii. The same best plot radii range, similar tendencies and lower correlations were observed when SMI was evaluated against HP-PCS instead of HP-CCI.

Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference data</th>
<th>R²</th>
<th>RMSE</th>
<th>Y = ax + b</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMI</td>
<td>HP-CCI</td>
<td>0.42</td>
<td>0.36</td>
<td>41.426−19.586</td>
</tr>
<tr>
<td></td>
<td>HP-PCS</td>
<td>0.74</td>
<td>0.27</td>
<td>24.681−11.943</td>
</tr>
<tr>
<td></td>
<td>TLS-HP-CCI</td>
<td>0.66</td>
<td>0.09</td>
<td>56.111−27.141</td>
</tr>
<tr>
<td></td>
<td>TLS-HP-PCS</td>
<td>0.41</td>
<td>0.36</td>
<td>2.6477−1.1456</td>
</tr>
<tr>
<td>PCS</td>
<td>HP-CCI</td>
<td>0.62</td>
<td>0.31</td>
<td>1.0770.260</td>
</tr>
<tr>
<td></td>
<td>HP-PCS</td>
<td>0.54</td>
<td>0.25</td>
<td>0.3820.046</td>
</tr>
<tr>
<td></td>
<td>TLS-HP-CCI</td>
<td>0.44</td>
<td>0.06</td>
<td>1.167−0.128</td>
</tr>
<tr>
<td></td>
<td>TLS-HP-PCS</td>
<td>0.42</td>
<td>0.45</td>
<td>0.431−0.124</td>
</tr>
<tr>
<td>CCI</td>
<td>HP-CCI</td>
<td>0.86</td>
<td>0.22</td>
<td>3.633−0.182</td>
</tr>
<tr>
<td></td>
<td>HP-PCS</td>
<td>0.66</td>
<td>0.09</td>
<td>0.2630.161</td>
</tr>
<tr>
<td></td>
<td>TLS-HP-CCI</td>
<td>0.85</td>
<td>0.07</td>
<td>0.8530.003</td>
</tr>
<tr>
<td></td>
<td>TLS-HP-PCS</td>
<td>0.85</td>
<td>0.48</td>
<td>0.341−0.082</td>
</tr>
</tbody>
</table>

Please cite this article as: García, M., et al., Canopy clumping appraisal using terrestrial and airborne laser scanning, Remote Sensing of Environment (2015), http://dx.doi.org/10.1016/j.rse.2015.01.030
Results improved moderately when applying the PCS to ALS data to shrub vegetation (Fig. 4E). Whereas smaller plot sizes provided poorer correlation for coarser voxels, finer voxels presented a more stable behavior across radii, with a small decrease for plot radii larger than 16 m. As happened for the TLS data, the correlations were about 10% lower when PCS estimates from ALS were compared to HP-PCS (Fig. 4F). Even more remarkable was that PCS from ALS yielded $\Omega > 1$, which is characteristic of a uniform canopy. This was especially observed for radii smaller than 12 m and finer voxels, but also true for a significant number of plots with larger radii and coarser resolutions. As for the forest (Fig. 4G), PCS in relation to HP-CCI yielded higher correlations than SMI, but still gave $\Omega > 1$ for most of the radii and voxel sizes evaluated. The same trends were observed when compared to HP-PCS (Fig. 4H). Correlations were approximately 17% lower, although for radii larger than 15 m the difference was smaller than 10%.

Contrary to TLS, CCI applied to ALS data produced higher correlations applying a 0.1 threshold for the iterative gap size removal. For shrubs, correlations with HP-CCI (Fig. 4I) presented a very irregular pattern for the finest 30 cm voxels, whereas the fine-medium 40–70 cm voxels showed a more stable behavior, with a correlation peak at the 7–10 m plot radii and a subsequent slight correlation decrease for larger radii. Results worsened considerably for coarser voxels, particularly for 100 cm, which presented negative correlations for most of the radii test- ed. CCI estimates for forest (Fig. 4K) showed a more irregular pattern for finer voxels, more stable trends for medium 40–50 cm voxels and the worst results for coarser voxels, particularly 90–100 cm voxels. The 50 cm voxels, which gave the highest correlation for radii between 10 m and 15 m, corresponded to the mean point spacing of the ALS data. Very similar results were obtained for the 40 cm voxels through the whole set of plot radii evaluated. The same tendencies according to voxel size and plot radii were obtained when ALS estimates were compared to HP-PCS, for both shrub and forest vegetation (Fig. 4 J and L), although lower correlations were achieved.

SMI yielded on average the lowest $\Omega$ values, i.e., the highest clumping, followed by CCI and then PCS. This trend was observed for all plot radii and for the two types of vegetation analyzed. For shrubs, SMI largely under estimated $\Omega$ compared to the reference HP-CCI values, although for the finer resolutions and some plot radii it gave $\Omega > 1$. CCI slightly over estimated $\Omega$ compared to HP-CCI estimates, whereas PCS consistently and significantly overestimated it, yielding $\Omega > 1$. Regarding forest vegetation, SMI $\Omega$ values were significantly lower than HP-CCI $\Omega$ values, though for smaller plot radii and finer voxels it yielded $\Omega > 1$; CCI provided slightly higher values whereas PCS estimates were significantly higher. Compared to HP-PCS, all methods overestimated $\Omega$ values for both vegetation types.

Table 2 shows the relationship obtained for each method and the HP-CCI estimates at the best voxel size and plot radii as well as the $R^2$ and the RMSE obtained. Based on these results, CCI performed the best for shrub and forest vegetation. Although PCS offered higher $R^2$ than SMI, PCS yielded the highest RMSE and more importantly, consistently provided $\Omega > 1$, thus highly overestimating $\Omega$ compared to the reference values. SMI and PCS showed a large deviation from the 1:1 line whereas CCI showed a closer agreement with the 1:1 when compared to CCI-reference data but large differences with PCS-reference data for both vegetation types.

3.4. Effect of the voxel size and plot radii on the $\Omega$ estimation from ALS

As was found for the TLS data, voxel size had a significant effect on $\Omega$ estimates from ALS data. SMI showed a correlation increase for coarser

![Image](image-url)
voxels and smaller plot radii for the forest plots, but correlations obtained for finer voxels in the case of shrubs remained quite stable for all radii, particularly at 50 cm. A K–W test showed that there were statistically significant differences (p-value < 0.05) as a function of voxel size for all radii and vegetation types. Moreover, a multiple comparison test enabled to generally distinguish a group for the 30–50 cm voxels and another >50 cm within which there were no significant differences. As for the effect of the plot radii, the K–W test showed that there were no significant differences for shrubs at different radii (p-value > 0.05) for the best voxel size (90 cm); although a multiple comparison test enabled three groups of Ω values for plot radii <8, 9–18 and >18 m. As for forest vegetation, there were significant differences (p-value < 0.05), and two groups were identified, <10 m and ≥10 m.

A confounding effect between voxel size and plot radius was also observed for PCS on both vegetation types. In this case, the voxel size had a larger effect causing a clear drop in the correlation values for the coarser voxels, although this effect was neutralized for the larger plots. Statistically significant differences for shrub and forest vegetation were found (p-value < 0.05) for different voxel sizes and plot radii. As for the voxel size, two groups were found for a threshold at 60 cm for shrubs and at 90 cm for forests. Regarding the plot radii, three groups were identified for radii <10, 10–17 and >17 m for shrubs; and <10, 10–15 and >15 m for forest. Likewise, the CCI was influenced by both factors (p-values < 0.05). A multiple comparison test for forest identified two groups for the <60 cm and ≥60 cm voxels, whereas for shrub vegetation a higher variability was observed identifying one group for voxels <50 cm and significant differences for the other resolutions tested. In regards to plot radii, the test identified two groups with thresholds at 11 m in the case of shrubs and 8 m in the case of forests.

4. Discussion

SMI consistently predicted the lowest Ω values from LiDAR data, regardless of the platform. It also showed a very small range of variation and hence, low sensitivity to different clumping levels. Since this algorithm was not included in the software used to process the reference data our results could be biased and it is difficult to reach a conclusion on its real performance. Further validation of SMI could be done by using simulated data based on radiative transfer models (RTM) with known clumping values or by using destructive sampling and gap fraction estimates, although the latter approach might be difficult to perform operationally.

For TLS data CCI yielded the highest Ω values followed by PCS whereas the opposite trend was found for ALS data. Other studies based on indirect methods including HP, TRAC and the Multiband Vegetation Imager (MVI), also reported higher Ω values using the CCI algorithm (Kucharik et al., 1997). Kucharik, Norman, and Gower (1999) argued that the higher CCI Ω values are due to the dependency of the canopy gap size distribution on the zenith angle. Walter et al. (2003) and Gonsamo and Pellikka (2009) also found that PCS systematically produced lower Ω values than CCI when they applied them to HP of simulated forest stands with known clumping degree of foliage and in situ HP from forest canopies with known architecture. In those studies CCI underestimated Ω for highly clumped canopies whereas PCS

---

Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>R²</th>
<th>RMSE</th>
<th>Y = aX + b</th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP-CCI</td>
<td>0.53</td>
<td>0.17</td>
<td>0.267</td>
<td>0.575</td>
<td></td>
</tr>
<tr>
<td>HP-PCS</td>
<td>0.46</td>
<td>0.48</td>
<td>0.151</td>
<td>0.130</td>
<td></td>
</tr>
<tr>
<td>HP-CCI</td>
<td>0.25</td>
<td>0.31</td>
<td>3.085</td>
<td>−0.739</td>
<td></td>
</tr>
<tr>
<td>HP-PCS</td>
<td>0.47</td>
<td>0.25</td>
<td>4.178</td>
<td>−1.830</td>
<td></td>
</tr>
</tbody>
</table>

---

overestimated $\Omega$ in poorly clumped canopies. The lower $\Omega$ value obtained by PCS was also observed in the HP derived $\Omega$ reference values of this research, providing values up to three times lower than the those of the CCI algorithm.

CCI outperformed PCS and SMI for both TLS and ALS. Our results applying CCI to TLS data were similar to those obtained by Zhao et al. (2012), who also computed $\Omega$ based on the CCI method and obtained an $R^2$ of 0.866 as compared to HP collected on 30 plots distributing across six coniferous sites in the Sierra National Forest, CA, over a similar zenith range (54°–63°). With regard the ALS data, the results obtained in this study for forest vegetation (Table 2) were slightly weaker than those achieved by Thomas et al. (2011) who reported a $R^2 = 0.81$ and RMSE $= 0.02$ for a power regression model developed over a mixed forest. Nevertheless, their models are site specific, whereas the approaches presented in this study have the potential to be applied to different environments and different vegetation types. Moreover, a linear relationship between HP and ALS estimates was obtained in this research while the relationships found by Thomas et al. (2011) despite being stronger, were not linear. No other approaches were found in the literature that applied SMI and PCS to estimate $\Omega$ from TLS or ALS data. In addition, the approaches are novel for their application to shrub vegetation and it is encouraging to find similar results to those obtained for forest vegetation. Validation of low resolution satellite products (Chen et al., 2005; He et al., 2012) using TRAC or HP measurements is not feasible and the use of spatially explicit estimates of $\Omega$ over relative large areas derived from ALS data would be of great value to up-scale and validate estimates using spaceborne sensors.

Correlations for PCS applied to TLS and ALS were surprisingly lower when compared to HP-PCS than HP-CCI. These results could be explained by the effect of the voxel size on the PCS algorithm. Walter et al. (2003) pointed out the high sensitivity of PCS to small structural variations and subtle differences in canopy clumping. Obviously the ability to capture small structural variations depends on the resolution of the image. Thus, while the HP had a spatial resolution of 0.8 cm at 10 m, the finest TLS voxel size was 3 cm. The effect of the different resolution between the HP and the TLS was also pointed out by Zhao et al. (2012). It should be noted that since the distance among returns will increase as a function of range, the voxel size must be sufficiently small to capture fine structural details but sufficiently large to avoid false gaps resulting from a voxel smaller than the scan resolution. Moreover, the binary approach applied has considered a voxel as fully occupied as long as it contained at least one return, thus reducing the level of detail captured by TLS. In the case of ALS data, the low density of the data prevented capturing small structural variations, which also made the effect of the difference in resolution between HP images and ALS data larger. This fact caused PCS applied to ALS data to consistently provide $\Omega > 1$. The lower point density of ALS data also affected CCI. The 0.1 threshold performed better than 0.05 because of the much lower resolution attainable from ALS data. Therefore, if the convergence threshold was set too low the removal of large gaps did not converge as the differences in gap fraction were higher due to the lower resolution.

Even though the same patterns were observed when TLS estimates were compared to the reference HP or TLS-HP, their correlation values varied significantly. A Wilcoxon test showed that differences between HP and TLS-HP $\Omega$ estimates were statistically significant ($p$-value $< 0.05$) regardless of the method used. It is remarkable that TLS $\Omega$ provided a much smaller RMSE but slightly lower $R^2$ against the reference TLS-HP than against HP estimates (Table 1). Despite HP being a widely accepted method to estimate canopy architectural parameters, it provides a 2D versus the 3D TLS and ALS representation of the canopy, so these inherent differences highlight the need for additional methods to comprehensively evaluate methods applied to 3D data. In this regard, data simulation might be a feasible way of providing reliable reference data on canopy architecture.

There are also factors related to the HP, such as images taken under sub-optimal illumination conditions (Zhang, Chen, & Miller, 2005) or the choice of image threshold applied to create a binary image, which could affect the gap fraction estimation. In this study an automatic threshold was used to minimize the analyst’s influence on the results, yet it is not exempt from errors. In addition, since the camera was not attached to the head of the scanner, any mis-registration between both sensors would translate into a slight displacement between their points of view, which could explain part of the differences found; especially for the lower zenith angles. View angle and lens distortions of the fisheye lens could also introduce slight differences in gap fraction estimates (Inoue, Yamamoto, Mizoue, & Kawahara, 2004) as the lens was not calibrated. Another source of discrepancy is the fact that gap segments measured on HP do not represent true physical sizes, varying with the distance from the sensor (Zhao et al., 2012). This gap size dependency on distance also affects TLS data as the distance between points and beam size increase with the distance from the sensor. Finally, the TLS data used consisted only on first returns, which worsens the effect of canopy occlusion of the canopy further away. Both, HP and TLS, are subject to similar occlusion issues, a problem that can only be minimized in the case of TLS by merging multiple scans, as for example Moorhoy et al. (2011).

In the case of ALS data, besides the much lower resolution attainable compared to HP, the different points of view of the two sensors also contribute to the differences found. Several studies have demonstrated how the distribution of returns is biased toward the upper canopy in ALS data undersampling the lower canopy whereas the opposite trend is found for TLS and consequently it affects the estimates of biophysical and structural variables derived from them (Chasmer, Hopkinson, & Treitz, 2004; Hilker, Van Leeuwen, Coops, Walder, & Newnham, 2010). Unfortunately, we did not have coincident TLS and ALS data for any of the study sites and so, we could not quantify the effect of the different points of view on the estimation of $\Omega$. This problem can also be extrapolated to the relation between ALS and HP. Demarez, Duthoit, Baret, Weiss, and Dedieu (2008) found significant differences in LAI and $\Omega$ estimated from HP using upward and downward perspectives over different crops. More importantly, ALS samples the canopy at near nadir view angles whereas HP does at multiple angles, which affects the gap distribution and therefore the $\Omega$.

Despite being derived from the same data and the same algorithms, differences between TLS and TLS-HP $\Omega$ can be explained by the deformations introduced in the TLS-HP after transforming the TLS point cloud into a polar projection (Maling, 1992). In addition, the TLS-HP was generated at a resolution equivalent to the angular resolution of the TLS data, which was finer than the finest voxel size. In addition, the approach to make a binary TLS-HP image was based on the proportion of returns versus missing returns within a pixel (Gajardo, 2014), whereas in our study a voxel was considered occupied as long as it contained one return. Another important factor is that different space segments were analyzed for each case. Segmentation for the TLS data was done in height intervals defined by the voxel size and subsequently integrated for the whole canopy, whereas HP or TLS-HP used CIMES, which segments the space into annular zenith angle rings.

In addition to differences intrinsic to the methods and data used to estimate $\Omega$, the voxel size, the zenith range and the plot radii to process the TLS and ALS data significantly affected the results. Voxel size represented the minimum unit at which the data were analyzed and, therefore, the scale at which $\Omega$ was estimated. SMI showed its worst result for finer voxel sizes, which could be due to the loss of the variability of the laser return distributions within each voxel. Instead, PCS and CCI are based on the gap size distribution therefore, the higher the resolution the greater the ability to capture smaller gaps.

Regarding the zenith range, the lack of statistically significant differences between zenith ranges for the CCI and PCS algorithms reinforces the relatively small $\Omega$ variation found by Leblanc et al. (2005) using CCI in the zenith range 30°–60°, who also pointed out the possibility of retrieving $\Omega$ outside that range since no radical changes were observed. Yet, for zenith angles larger than 75° increases of $\Omega$...
by approximately 20% can occur due to structural and optical effects (Chen et al., 2006).

In the case of ALS data, the effects of plot radii and voxel size were mixed and difficult to isolate. SMI showed very different patterns by plot radii for shrubs and forest (Fig. 4A–D). The higher irregularity for forest can be explained by the different distribution of returns within the canopy, which is affected by the variability of structure associated with the plot size (Fig. 5). In addition to changes in height and fractional cover for the different plot radii, the canopy return permeability of the different tree species that comprises the tree vegetation type of Jasper Ridge affected the distribution of returns within the canopy and therefore, within the voxel used for the SMI method. For PCS and CCI, it should be noted that although the point cloud was initially voxelized, due to the low density of the data, a binary image was subsequently created. Therefore, the effect of the canopy permeability is not as important as for the SMI, although it limits the possibility of providing 3D estimates. Whereas for CCI the effect of voxel size clearly affected the correlations found with HP estimates, PCS seemed to be more affected by voxel size and plot radius which resulted in £ 1 for small radii and coarse resolution for most of the plots. The overestimation of £ values compared to the reference data found for PCS decreased for the larger plot size, which could be related to the increase in plot variability and the distribution of voids/occupied voxels.

5. Conclusions

This research evaluated the potential of SMI, PCS and CCI to estimate £ from TLS and ALS data. Overall CCI performed the best compared to HP. PCS yielded smaller £ values for TLS, whereas consistently gave £ = 1 for ALS. SMI was not implemented to analyze the HP, which might have biased the results favoring the other two methods.

HP provides a 2D representation of the canopy and thus of the £ phenomenon whereas the methods applied over TLS have the potential to estimate £ in 3D and serve to validate ALS £, given also the lack of statistical significant differences found between zenith ranges. Nevertheless, differences in terms of point of view and point density would need to be considered. The low point density of the ALS prevented the application of PCS and CCI in 3D, as it was done for the TLS data. The use of high point density multiple-return or full waveform data could help to alleviate this lack of resolution of the ALS data. The algorithms tested demonstrated their robustness for TLS and ALS and for different vegetation types, so they have the potential to be applied operationally across sites and to validate spaceborne £ products.

Acknowledgments

This work was supported by the United States National Aeronautics and Space Administration grant #NNX11AF93G and #NNX09AN51G; by the European Union grant FP7-PEOPLE-2009-IRSES-246666; and by the Spanish Ministry of Economy and Competitiveness grants Biospec (CGL2008-02301/CL) and FLUyPEC (CGL2012-34383). J. Gajardo was supported by a CONICYT-Becas Chile Doctoral Fellowship from the Chilean Government. We thank the people of CSTARS at the University of California, Davis for their assistance during the field campaign and especially to Maria Santos for her contribution to all data collections. We thank G. Asner, D. Knapp, and the Carnegie Airborne Observatory team for collecting and providing the ALS data, which was funded by a grant to G. Asner from the Gordon and Betty Moore Foundation. The Carnegie Airborne Observatory was made possible by the Avatar Alliance Foundation, Margaret A. Cargill Foundation, John D. and Catherine T. MacArthur Foundation, Grantham Foundation for the Protection of the Environment, W.M. Keck Foundation, Gordon and Betty Moore Foundation, Mary Anne Nyburg Baker and G. Leonard Baker Jr., and William R. Hearst III. We also thank the team of Stanford University’s Jasper Ridge Biological Preserve for all their support throughout this work, very especially Nona Chiariello.

We also greatly appreciate the constructive comments of the 3 anonymous reviewers which significantly improved our paper.

References


BCAL LiDAR Tools ver 1.3.5. Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory (BCAL), Boise, Idaho. URL: http://bcal.geology.isu.edu/envitools.shtml.


Please cite this article as: García, M., et al., Canopy clumping appraisal using terrestrial and airborne laser scanning, Remote Sensing of Environment (2015), http://dx.doi.org/10.1016/j.rse.2015.01.030