Effects of voxel size and sampling setup on the estimation of forest canopy gap fraction from terrestrial laser scanning data

Renato Cifuentes [a,*], Dimitry Van der Zande [b], Jamshid Farifteh [a], Christian Salas [c], Pol Coppin [a]

[a] M3-BIORES, Biosystems Department, Katholieke Universiteit Leuven, Willem de Croylaan 34, box 2470, BE-3001 Leuven, Belgium
[b] Management Unit of the North Sea Mathematical Models, Gulledelle 100, BE-1200 Brussels, Belgium
[c] Laboratorio de Análisis Cuantitativo de Recursos Naturales, Departamento de Ciencias Forestales, Universidad de La Frontera, PO Box 54-D, Temuco, Chile

A R T I C L E   I N F O

A B S T R A C T

Assessments of forest canopy structure remain a challenge and are most often conducted using indirect techniques limited to a two-dimensional perspective. Using terrestrial laser scanner (TLS) technology, a three-dimensional (3D) approach to study canopy structure was conducted by modeling forest scenes from three broadleaved forest stands with different canopy features. Field TLS data were collected from each stand using a phase based FARO® LS880 laser scanner on four sampling setups. The capability of TLS-derived data to represent canopy structure was evaluated by comparing gap fraction estimates from the 3D models with gap fraction values from digital hemispherical photographs (DHP). Firstly, the collected 3D point clouds were processed to obtain fully representative voxel-based models of the forest canopy. Secondly, ray tracing algorithms were applied on these models to simulate hemispherical views and estimate gap fraction. Finally, a sensitivity analysis was done using different voxel sizes and the four sampling setups on the simulations, in order to assess their impact on the gap fraction estimates derived from TLS data. Results of TLS-derived gap fraction showed that combining nine scans produced better results in all forest stand. Similarly, the dimension of voxels have a marked influence on these results. Voxel sizes of 1 cm, 2 cm and 4 cm were found to have less error when compared with real gap fraction values derived from DHP for young, intermediate and mature forest stands, respectively (RMSE ranging from 9% to 16%). However, substantial differences in gap fraction were observed and described at different zenith angles. These results suggest that specific TLS sampling setup and processing are required depending upon the forest type under analysis. Overall, this research indicates that phase based TLS data can be used for objective calculation of gap fraction.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The structure of the canopy controls the quality, quantity, temporal and spatial distribution of light and can be characterized at several levels of detail. Difficulties of canopy access, however, impose severe limitations for direct sampling and quantification of canopy structure. As a result, accurate representation of the spatial variation are rarely achieved (Parker, 1995). Forest canopy assessments, in general, have been largely performed using indirect methods, where characterization of forest canopy can be summarized by a factor or characteristic dimension. Commonly used variables, calculated and analyzed in order to understand canopy–radiation interactions, are leaf area index, canopy gap fraction, canopy cover or canopy closure (Korhonen et al., 2011; Parker, 1995). The field assessment of these variables are based on a variety of techniques and instruments (Bréda, 2003; Weiss et al., 2004; Welles and Cohen, 1996). One of the most used and validated techniques to study plant canopies is Digital Hemispherical Photography (DHP), where photographs acquired through a hemispherical 180° fisheye lens, from beneath the canopy and oriented towards zenith or placed above the canopy looking downwards, are used to calculate the fraction of diffuse and direct radiation reaching the camera viewpoint (Gonsamo et al., 2010; Jonckheere et al., 2005).

Over the past years there has been an increasing research demonstrating that light detection and ranging (LiDAR) technology can provide an effective monitoring approach in many applications and has gained wide attention in forest management and

* Corresponding author. Tel.: +32 16 329750; fax: +32 16 322966.
E-mail addresses: renato.cifuentes@biw.kuleuven.be, renegro33@hotmail.com (R. Cifuentes), dimitry.vanderzande@mumm.ac.be (D. Van der Zande), jamshid.farifteh@biw.kuleuven.be (J. Farifteh), christian.salas@ufrontera.cl (C. Salas), pol.coppin@biw.kuleuven.be (P. Coppin).

agriculture. Terrestrial laser scanner (TLS) data have made possible to quantify forest metrics (Dassot et al., 2011; Hilker et al., 2010; Moskal and Zheng, 2012; Yao et al., 2011) and represent 3D structure of forest canopies in a three-dimensional (3D) environment with a high level of detail (Hosoi and Omasa, 2007; Jupp et al., 2009; Lovell et al., 2003; Moorthy et al., 2011; Seidel et al., 2012) in a non-destructive, unequivocal, objective, and reproducible manner (Van Leeuwen et al., 2011). TLS use powerful highly collimated optical light i.e., laser light, as sensing carrier. The energy of such a laser beam interacting with the measured object is partially reflected back in the direction of the laser system where it is registered by a sensor and used to measure the distance between the sensor and the illuminated spot on the measured object(Dassot et al., 2011; Petrie and Toth, 2009a). Three types of TLS technologies are recognized in literature: (i) time-of-flight discrete-return scanners; (ii) time-of-flight waveform scanners; and (iii) continuous wave phase-shift scanners. Time-of-flight discrete-return instruments provide high accuracy at large range and resolve gaps well. This category of scanners has been the most often used for vegetation structure measurements (Côté et al., 2009). Time-of-flight wave-form scanners record the full time trace of energy that is returned after a laser pulse has been emitted from the instrument. Ranges can be resolved from the waveform by post-processing, similarly to the time-of-flight discrete-return instruments. The key difference is that, in the former, the full intensity trace is recorded for future analysis. Phase based scanners employ a constant wave laser with intensity modulated at a series of frequencies. The shifts in phase of the returned modulations are used to determine range. These type of scanners can sample at much higher frequencies than time-of-flight instruments. This comes at the cost of effective range (i.e., less than 100 m) and range averaging in cases where a single outgoing pulse is intercepted by multiple objects along the beam trajectory, making the data noisy (i.e., ghost points) around the edges of objects (Newnham et al., 2012). Indoor TLS measurements on artificial canopy-like structures and under controlled conditions, indicate that the ratio of ghost points to the number of valid scan points increases with range, mainly as a function of beam divergence (Cifuentes et al., unpublished result). Consequently, this ratio is also zenith angle dependent. That is, if objects being scanned are organized horizontally (e.g., forest canopy), the ratio of ghost points to the number of valid scan points increases with increase of zenith angle while the incident angle is kept constant. Moreover, the ambiguity of phase based scanning means that gaps are more difficult to identify. The influence that ghost points and gaps have on the quality of forest canopy TLS data is generally addressed through post-processing, with erroneous points removed based on the smoothness or range with respect to neighboring points. The limitation on range of phase based scanners can be minimized when measuring the 3D space placing the TLS at different locations, as described by Van der Zande et al. (2008); obtaining substantial improvements in the quality of the datasets. Originally proposed to overcome occlusion, this sampling setup allowed laser beams to enter the canopy at a variety of angles and directions and thus increased the probability of a laser beam penetrating deeper into the canopy. The process of merging multiple point clouds, in order to create a comprehensive dataset, is known as registration.

TLS data can be analyzed as point clouds or modelled using different techniques, such as the voxel-based approach (Hosoi and Omasa, 2007). Voxelization is required to normalize the raw point cloud data and give a determined volume to the elements in the study space. This procedure involves the specification of the voxel size, which has a marked impact on the results, as shown in previous studies on 3D modeling of trees and forest canopies using TLS data. On those studies, the authors have used a predefined voxel size based on the minimum size of the target element (Van der Zande et al., 2006) and, more recently, on TLS characteristics (Moskal and Zheng, 2012; Seidel et al., 2012). Both, registration and voxelization procedures are crucial in TLS processing chain, particularly for quantification of forest structure. It is, therefore, necessary to understand how the sampling setup and the selection of the voxel size influence the results. Additionally, complexity and features of forest stands can be considered important factors for voxel size and sampling setup definition. After correct processing of TLS point cloud data, it is possible to model the 3D space where leaves, branches, stems and all the elements contained within the instrument field of view are included.

In the present study, a voxel-based approach was adopted to model the 3D arrangement of forest canopy elements, using TLS data from three different forest stands. Different voxel sizes and sampling setups were used and tested in the processing chain to build the 3D models. Subsequently, gap fraction (GF) was derived from these models as descriptor of canopy structure. One of the main problems in forest canopy research is that results are very difficult to validate due to access-related issues (Barker and Pinard, 2001). Since digital hemispherical imagery is a well investigated and accepted standard for assessing canopy structure variability, DHP-derived canopy GF values were used as references for comparison purposes. The objectives of this study were, first, to evaluate the capabilities of TLS data to generate 3D models of different forest canopies and, second, to quantify the impact of voxel size and sampling setup assessed in terms of canopy structure variability, particularly in radiation interception calculated as GF (Hilker et al., 2010; Leblanc et al., 2005).

2. Materials and methods

2.1. Site description

The field measurements were conducted in Heverlee Forest (Heverlee–Meerdaal complex). The complex covers a total area of approximately 1890 ha and is the second largest forest complex in the Flanders province, Belgium (50°51′ N, 4°40′ E; 60 m a.s.l.). Three forest stands, in a generally flat terrain, were chosen for the data collection. The first stand (young forest) is a 2.0 ha plantation composed mainly by oak (*Quercus* sp. L.). The second stand (intermediate forest) is a 3.0 ha European beech (*Fagus sylvatica* L.) plantation on its intermediate stage of development, growing together with oak and birch (*Betula pendula* L.). The third stand (mature forest) is a mature 2.5 ha beech stand on the last stage of its cycle. In order to avoid unwanted elements within the field of view of the instruments, all the stands were selected where no understory was present. Average stand variables are given in Table 1.

2.2. Digital hemispherical photography acquisition and processing

The general sampling setup consisted of nine plots distributed on a 10 m × 10 m grid in each forest stand. (Weiss et al., 2004; Gonsamo and Pellikka, 2009a; Fig. 1). At each plot, TLS and DHP data collection were done between June 23rd and July 3rd, 2012.

DHPs of the canopy were acquired from each plot using a 6.1 MP Kodak DCS 660 digital camera (Eastman Kodak, New York, NY), with a 180° fisheye lens (8 mm, f/4, Sigma, Tokyo, Japan). Measurements were acquired in the center of each forest stand where the edge effects were minimal. The camera and lens were placed in the center of the plot on a tripod 1.3 m above the ground and leveled using a double axis spirit level. The camera was oriented towards zenith in such a way that the magnetic north was always located at the top of the image (Jonckheere et al., 2004). DHPs were taken before sunset (i.e., from 19h00 to 21h00), under

Table 1
Average stand variables at the study site in Heverlee Forest. N is density (trees/ha); G is basal area (m^2/ha); D_{gM} is quadratic mean diameter at breast height (cm); H_{gM} is mean tree height (m) and CBH$_{gM}$ is mean tree crown base height (m).

Stand	Dominant canopy species	N	G	D_{gM}	H_{gM}	CBH$_{gM}$
Young forest	Oak	2100	24.3	12.1	14.2	4.4
Intermediate forest	Beech	675	21.0	19.9	23.1	7.1
Mature forest	Beech	223	28.6	40.5	32.3	14.3

overcast sky conditions to ensure homogeneous illumination of the canopy and clear contrast between the canopy and the sky. For every plot, six DHPs were taken using variable aperture size and shutter speed. Then, a first visual analysis was done among the six images and the image having the best contrast between sky and canopy was selected for further processing.

2.2.1. Calculation of GF from DHP

The DHPs were processed and analyzed with Gap Light Analyzer 2.0 (Frazer et al., 1999) to extract GF values through the following sequence of steps. First, as absorption of leafy material is maximal and scattering of sky tends to be lowest in the blue channel, the blue channel was chosen for subsequent automatic thresholding. Then, to assure an efficient and consistent post-processing of hemispherical photographs (Jonckheere et al., 2005), the Ridler and Calvard (1978) automatic thresholding method was used. This technique was tested and found to be optimal and robust by Jonckheere et al. (2005) in a wide range of light and canopy structure conditions. It also offers a fast, reliable and objective use of hemispherical photographs for GF estimation in forest stands.

The DHPs were analyzed by selecting circular rings or zenith annuli at zenith angle range, over the whole range of azimuth angles (i.e., 0°–360°), which is suitable for canopy structure analysis because it identifies sequences of pixels representing foliage or gaps needed to measure gap size and fraction (Gonsamo and Pellikka, 2009b; Gonsamo et al., 2010; Leblanc et al., 2005). As defined by Ross (1981), GF in any particular direction is the fraction of sky not obstructed by canopy elements. Measurements of GF should integrate information over the entire sky hemisphere, however, the segment measured may vary with the instrument used. Several studies have proposed different zenith angle ranges (e.g., 55°–60°, 47°–68°, 30°–60°, 25°–65° and 10°–80°) for estimation of leaf area index from GF values, using various methods and

instruments (see Jonckheere et al., 2006; Leblanc et al., 2005; Weiss et al., 2004), where experimental results have suggested avoiding those zenith angles close to zenith and horizon. Nine equiangular zenith annuli were defined in the zenith angle range between 0° and 70°, as a result each annulus was 7.8° in width. The GF for a single circular ring (i.e., annulus) located at zenith angle θ was calculated as Frazer et al. (1997):

$$GF = \frac{Fs}{(Fs + Fv)} \tag{1}$$

where GF is the gap fraction for the annulus centered at θ, Fs is the number of sky pixels in the annulus centered at θ and Fv is the number of vegetation pixels (leaves and branches) in the annulus centered at θ. The central zenith angles for the nine annuli were 3.9°, 11.7°, 19.5°, 27.3°, 35.1°, 42.9°, 50.7°, 58.5° and 66.3°. DHP-derived GF was considered the reference for further comparisons, thus, hereafter referred to as real GF.

2.3. Terrestrial laser scanner data acquisition and processing

The TLS instrument used in this study was the FARO® LS880, which uses a continuous wave laser operating at $\lambda = 785$ nm, in the near infrared part of the spectrum, to measure the 3D position of points within a range of 76 m. Active sensor systems require the generation of a fairly large amount of energy to adequately illuminate targets, hence, TLS is limited by its own energy source and surely cannot detect anything beyond its maximum range because of hardware limitations. While operating, the beam emitted from the laser diode passes into the center of a rotating mirror that deflects it to a fixed angle of 90° giving the angular coverage of 320° on the vertical plane. Rotating motor allows the 360° azimuth scan. As in hemispherical photography acquisition, our measurements were acquired in the center of each forest stand where the edge effects were minimal. The instrument's vertical scanning range was
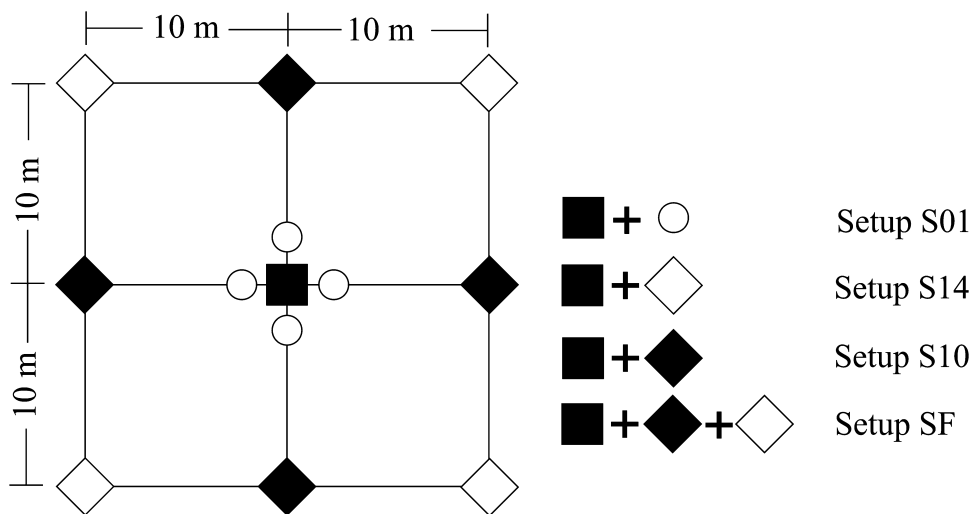
Fig. 1. Systematic data collection using a fixed grid of 10 by 10 m. Black square represent digital hemispherical photography (DHP) and central terrestrial laser scanner (TLS) measuring position. White circles represent TLS measuring position at 1 m from central TLS (combined in the S01 setup); white diamonds represent TLS measuring position at 14 m from central TLS (combined in the S14 setup); and black diamonds represent TLS measuring position at 10 m from central TLS (combined in the S10 setup). Setup SF is the combination of setups S14 and S10.

set to avoid recording unnecessary data from forest floor, while horizontal scanning range was set to hemispherical coverage. TLS measurements were done on a systematic geometrical pattern. In total thirteen scans were taken, one at the center of the plot and at the exact same position and height as the corresponding DHP, and two sets of scans were taken at different distances from the central one. A first set of four scans was located 1 m from the center along cardinal directions (N, S, E, and W), following the diamond-shape sampling setup proposed by Van der Zande et al. (2008). A second set of eight scans was located on the 10 m grid as shown in Fig. 1. The TLS data processing route comprised four steps: filtering, registration, voxelization and simulation; which are explained in the following sections.

2.3.1. Filtering

As specified in Section 1, gaps are more difficult to identify in phase based TLS data and ghost points can be created when the beam intercepts multiple objects. Given these limitations, erroneous measurements needed first to be removed from the point cloud before further processing. This was done using a similar range based kernel filter as the one used in Newnham et al. (2012). The point clouds were re-projected to a 2D format (i.e., a range map), where solid surfaces show a degree of continuity across their pixel values, while ghost points occurring at edges of these surfaces show noticeable changes in range. Thus, the filter, within a 3 by 3 pulse window, removes a point if more than 50% of its neighboring points are separated by more than 0.02 m in range.

2.3.2. Registration

Registering point cloud datasets recorded from different locations allows us to minimize the occlusion effect and improve substantially the quality of the data (Yao et al., 2011; Zhao et al., 2011). Registration, also called spatial overlap or fusion, was done using the FARO® Scene software. Success in the registration process of separate scans may be influenced by the dimension and shape of canopy elements (e.g., understory, stems, branches) and the way they are displayed in the measured 3D space, creating uncertainty in the registration process. To minimize the uncertainty and for better orientation and registration of scans, it is necessary to include a number of survey control points within the forest scene (Fassi et al., 2011). Hence, reference spheres with known dimensions and reflectivity were placed in each forest stand within the scanner field of view and range, in such a way that they can be visible from all TLS positions and recognizable as specific objects within the point cloud. The spheres were located in the scan and assigned an individual identification. FARO® Scene performs the registration, in which a scan position is determined, with corresponding scan references being matched to references in other scans. With reference spheres, the survey point (i.e., point extracted for registering the scan) lies in the center point of the sphere. Consequently, the point cloud datasets were registered into comprehensive point cloud data. TLS data were registered in four different setups, using the central scan as reference and (i) the four scans located 1 m away from it (hereafter called setup S01); (ii) the four scans located 10 m away from it (hereafter called setup S10; (iii) the four scans located 14 m away from it (hereafter called setup S14); and (iv) the four scans in (ii) combined with the four scans in (iii) (hereafter called setup SF), (Fig. 1). By registering, all points and reference spheres obtain a position and orientation in a global coordinate system. Because of the limitations of measurement accuracy, the local reference points from different point clouds generally obtain different global coordinates. The discrepancies give information about how the quality of the scans registration was. The tension of registration describes this discrepancy in the global coordinate system between the position and the orientation of the two corresponding reference objects in two scans, meaning that the lower the tension value the better the registration.

2.3.3. Voxelization

Both, single and registered point clouds may be composed of millions of points, each of them with its correspondent xyz-coordinates. These points are not suitable to simulate a hemispherical image because points do not have an area or a defined volume (Seidel et al., 2012), which is needed to reconstruct the objects and quantify the space that these objects are filling within the scanned space. In addition to that, and due to the fixed divergence of the laser beam, both in the horizontal and vertical plane, objects located within the field of view (i.e., not covered by other objects) and closer to the instrument will be represented by a higher density of points than those objects located within the field of view but further, resulting in an imbalanced representation of the measured 3D space. A voxel-based technique to overcome these problems is described in Hosoi and Omasa (2007), where the 3D space was divided into a finite number of cubic voxels and were classified depending on point/voxel interaction. Voxels with at least one reflected beam or hit in it were assigned a value of 1 (filled) and a value 0 (empty) was given to voxels that did not enclose any hit. In this way it was possible to homogenize and reduce the size of registered point clouds. Van der Zande et al. (2006) defined the size of the voxels accordingly to the minimum size of the object under study on a controlled environment using an artificial tree with squared-shape leaves of 5 cm in side. Therefore, cubic voxels of 5 cm × 5 cm × 5 cm were used finding under and overestimations on the amount of plant material due to occlusion and voxel size, respectively. The theoretical voxel size can also be determined based on the scanner range, beam divergence and beam diameter (Seidel et al., 2012). The latter approach was used in this study, meaning that for a given angular divergence of the beam, the greater the range, the larger the diameter of the area that is being covered at the object. The radius of such an area is given by the formula (Petrie and Toth, 2009b):

$$R = \frac{(Di \times BDv + BDm)}{2} \tag{2}$$

where R is the radius in m, Di is the distance to the scanner in m, BDv is the beam divergence in radians and BDm is the beam diameter at exit of the instrument in m. In the present study, the FARO® LS880 was set to emit the laser beam in fixed steps of 0.036° in the vertical and also in the horizontal plane. Each beam diverges by 0.25 mrad and the beam diameter at the exit of the instrument is 3 mm. Similarly to the increasing R with distance, the distance between two neighboring beams increases and can be calculated using the formula:

$$CD = (\tan \alpha \times Di) - 2R \tag{3}$$

where CD is the distance between two neighboring beams, α is the angular resolution in degrees, Di is the distance to the scanner in mm and R is the radius of the laser beam in mm at the distance Di obtained from Eq. (2). At the maximum range of the scanner (76 m) the distance between two neighboring beams (i.e., a maximum unsampled distance) is 25.75 mm. Thus, assuming that all beams reaching that distance give effective returns if an object fully obstruct the beam, the theoretical voxel size 2.6 or V2.6 (2.6 cm × 2.6 cm × 2.6 cm) was used, minimizing unsampled space for an homogeneous dataset between two continuous beams at the maximum range of the instrument. The applicability of the theoretically defined voxel size in the 3D simulation of forest canopies using real forest TLS data was tested and compared with five different voxel sizes: voxel size 1 or V1 (1 cm × 1 cm × 1 cm), voxel size 2 or V2 (2 cm × 2 cm × 2 cm), voxel size 3 or V3 (3 cm × 3 cm × 3 cm), voxel size 4 or V4 (4 cm × 4 cm × 4 cm) and voxel size 5 or V5

(5 cm × 5 cm × 5 cm). The effect of these voxel sizes on the TLS-derived GF estimates was calculated for each forest stand.

2.3.4. Simulation

Registered datasets of the different setups described in Section 2.3.2 and the six different voxel size were the inputs for simulations. These voxel-based representations of the 3D data were transformed into an hemispherical image to calculate the GF using the open source ray tracing software Persistence of Vision Raytracer or POV-Ray (Persistence of Vision Pty. Ltd., Williamstown, Victoria, Australia). POV-Ray was set to represent the voxels as dark-solid non-reflective cubic objects. As mentioned, all different voxel sizes described in the previous section were used in the process. White background and no source of light were included as additional features, in order to simulate an overcast sky condition and obtain the appropriate contrast in the image. Subsequently, a ray tracing algorithm using a virtual camera set in fisheye projection mode was applied to create a 2D image at very high resolution (i.e., 342 MP) in a Intel® Core™ i7 CPU, 2.80 GHz, 64-bit operating system, with 16.0 GB installed memory (RAM). That high resolution was chosen to circumvent any systematic error in the POV-Ray simulation and looking for a compromise between the smallest voxel size and computing time. This means that having 1 cm^3 as the smallest voxel size, a resolution of approximately 390 MP is required in order to detect these voxels at 76 m (maximum range of the scanner). However, that demands considerable processing power and time, therefore, we used a resolution of 342 MP, which allowed us to detect the smallest voxels at the whole range of distances within a reasonable time period. For comparison purposes this image was subsequently transformed to an equivalent resolution as the original DHP (i.e., 6.1 MP) using a bicubic interpolation method.

2.4. Calculation of GF from TLS and comparison

For determination of GF from registered scans and a direct comparison with real GF, the same Ridler and Calvard (1978) automatic thresholding method used in DHP processing (Section 2.2.1) was applied on the voxel-based hemispherical views created in Section 2.3.4. GF was subsequently calculated by means of Eq. (1), for each 7.8° annulus defined in the zenith angle range between 0° and 70°. Hemispherical images simulated using TLS data were first visually inspected and compared with DHP. The criteria for this qualitative evaluation was based on the performance of the algorithms to detect canopy elements throughout the full range of zenith angles, especially at the edges of the images (i.e., areas near the horizon). The quantitative comparison between simulated hemispherical images and DHP was made based on differences in GF values. In the interest of determining the influence of sampling setup and voxel size on GF estimates, different combinations of voxel sizes and TLS sampling setups were applied for image simulation. The root mean square error (RMSE) was used as measure of performance to evaluate the differences, in percentage, between the TLS-derived estimates of GF and the true GF values at different zenith angles.

3. Results

3.1. TLS-based simulations

After the range based kernel filtering, single point clouds were composed, on average, of 19 million points in the young and intermediate forests and 17 million points in the mature forest. Despite the use of a filter, ghost points occurrence could not be fully overcome. Since ghost points are difficult to detect visually from the perspective of the instrument, Fig. 2 depicts a 3D lateral view of a single point cloud, where ghost points are highlighted.

The registration process of the single point clouds for the sampling setups S01, S10, S14 and SF, was successfully completed. The mean tension value in registrations of 0.007 indicates a very low discrepancy in the global coordinate system between the position and the orientation of two corresponding reference spheres in two scans. In total, 12 registered point clouds were built combining five single point clouds in S01, S10 and S14; and combining nine single point clouds in SF; with an average registration error ranging from 2.56 to 5.9 mm. Registered point clouds in young forest were composed of 76, 95, 96 and 172 million points; in intermediate forest of 76, 95, 95 and 171 million points; and in mature forest of 67, 83, 84 and 152 million points; using S01, S14, S10 and SF, respectively. Registered point clouds were voxelized and used to simulate hemispherical views of each forest stand in POV-Ray (Fig. 3).

Visual inspection of simulated images confirms differences between the setups. S01 (Fig. 3(b)) shows high density of voxels at low zenith angles, while sparse coverage is seen at high zenith angles. S14 and S10 (Fig. 3(c) and (d)) present a rather homogeneous coverage throughout the range of zenith angles, however, the density of elements is still reduced when compared with DHP. SF (Fig. 3(e)) provides a greatly increased number of returns and yielded an enhanced representation of the elements within the analyzed space. Nevertheless, complete coverage of the sky hemisphere using TLS-derived data was not possible with either of the sampling setups, specially over areas near the horizon.

3.2. GF calculation and comparison

Real GF values and GF values derived from registered TLS point clouds are presented in Fig. 4

The general trend of GF values, relatively low and with small variations between forest stands, reflects the effect of silvicultural management, which is mainly focused to optimize production.

In the young forest (Fig. 4, left column), real GF shows a declining trend with extreme GF values of 30% at 4° (i.e., first annulus) and 7% at 66° (i.e., last annulus). For all setups, increasing voxel size steadily decreases GF along the range of zenith angles. Setup S01 has an opposite trend in GF estimation for all voxel sizes. Large underestimation of GF is seen at low zenith angles and different levels of overestimation is present between 65° and 70°, depending upon voxel size. Setups S14, S10 and SF show a better trend in the GF distribution for all voxel sizes. However, clear underestimation between 4° and 60° is still noticeable for most voxel size, except for V1. Voxel size V1 shows a better agreement when compared with real GF in setup SF.

The real GF distribution in the intermediate forest (Fig. 4, central column), shows a more stable trend between zenith angles 4° and 45°, with GF values between 16% and 23%. GF then decreases to 5% in the last annulus. S01 presents an increasing trend in GF values. Under and overestimation is seen at the lowest and highest zenith angles, respectively; and for all voxel sizes, as in the young forest. Setups S14, S10 and SF present an improved GF distribution, particularly between zenith angles 20° an 45°. Yet, GF at extreme annuli have visible discrepancy. Depending on setup, rather small voxel sizes, but larger than V1, showed better agreement when compared with real GF.

The mature forest stand (Fig. 4, right column) exhibits a different GF distribution in both real and TLS-derived, mainly at zenith angles below 15°. The high GF values here are, naturally, due to the presence of a large gap in the forest canopy. Above this zenith angle, a similar behavior to the previous forest stands is perceived on TLS-derived GF values. S01 has an increasing trend in GF distribution; S14, S10 and SF present a better agreement with real GF at zenith angles between 20° and 45°. Although there are substantial under and overestimation at extremes, greater voxel sizes appear
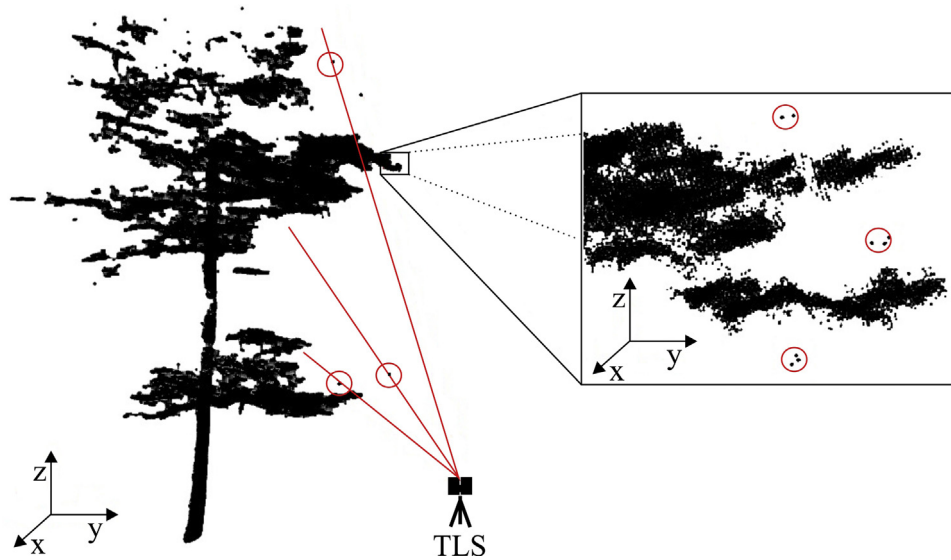
Fig. 2. Three dimensional (3D) point cloud of a tree. Scan points are in black. Example of ghost points are enclosed with red circles. Red lines represent laser beams being emitted by the terrestrial laser scanner (TLS). The zoom window shows a detailed view of ghost points surrounding edges of leaves. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to be closer to the real GF distribution. Small voxel sizes are prone to overestimate GF.

A quantitative comparison of the differences between real GF and TLS-derived GF while using several voxel sizes and different setups is shown in Fig. 5. Lines represent the different setups.

For every forest stand, the difference in GF at the extremes of the zenith angle range (i.e., 0° to 70°) is evident, for all setups and voxel sizes. Setup S01 is the one with higher differences in almost every situation. The rest of setups present, in general, a more similar shape. The zenith angles where the differences in GF are between −10% and 10% (i.e., between dotted lines in Fig. 5), is variable, depending upon setup and voxel size. In general terms, disregarding extremes and for the whole range of setups; low and steady differences in GF can be seen between zenith angles 20° and 45°
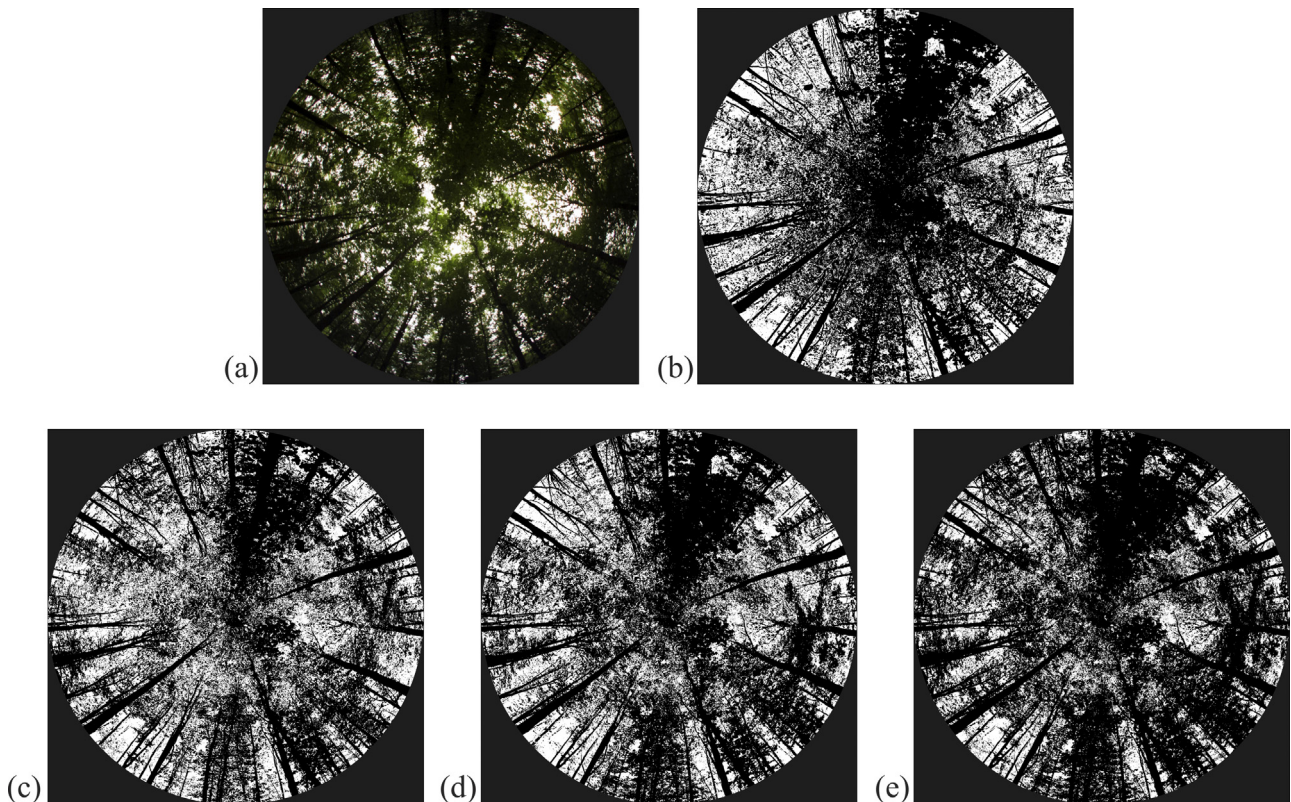
Fig. 3. Example of different types of images being analyzed for the intermediate forest and using voxel size 1 (V1). Digital hemispherical photograph (a); hemispherical view simulation in POV-Ray of registered point clouds 1 m from the center or S01 (b); 14 m from the center or S14 (c); 10 m from the center or S10 (d); and (c) and (d) combined or SF (e).
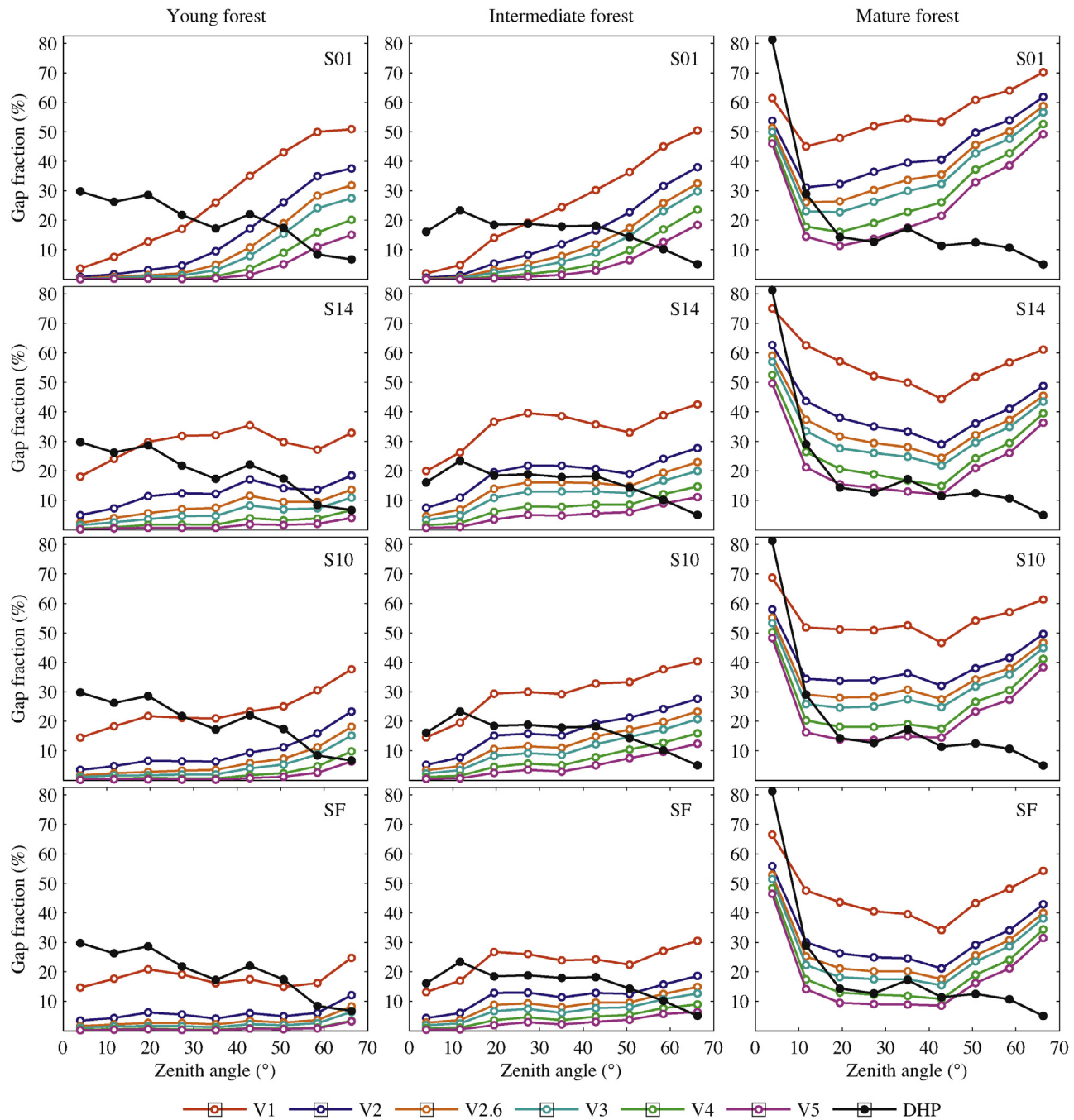
Fig. 4. Gap fraction (GF) as a function of zenith angle in young, intermediate and mature forest using setup of registered point clouds: 1 m from the center (S01); 14 m from the center (S14); 10 m from the center (S10); and S14 and S10 combined (SF). Real GF presented with solid circles; GF derived from terrestrial laser scanner (TLS) data are shown with open circles. V1: voxel size 1 cm; V2: voxel size 2 cm, V2.6: voxel size 2.6 cm; V3: voxel size 3 cm; V4: voxel size 4 cm; and V5: voxel size 5 cm.

and using V1 or V2 in young forest; from V1 to V3 in intermediate forest; and from V2.6 to V5 in mature forest.

The selection of a voxel size and a sampling setup seems essential for the retrieval of GF estimates derived from TLS. As depicted in Fig. 4, the RMSE of the differences in GF confirm that the lowest error on GF calculation from TLS-derived data in the young forest is using setup SF and V1; in the intermediate forests using SF or S14 with V2 and V2.6, respectively. Simulation with larger voxel size increases the RMSE. The mature forest, in turn, shows a decreasing trend in the RMSE while increasing voxel size, with a low error using SF with V4 or V5, or S14 withV5.

4. Discussion

Sampling setup definition for TLS data collection and voxel size determination for 3D data processing is complex but can lead to an efficient planning of field work and better representation of the forest canopies when forest stand characteristics are considered. In this study we found that both, voxel size and sampling setup have noticeable influence on results of GF derived by TLS data.

The visual inspection of simulated images and comparison with DHP resulted in the identification of areas that were not covered by TLS during data collection. This suggests that the laser hits low
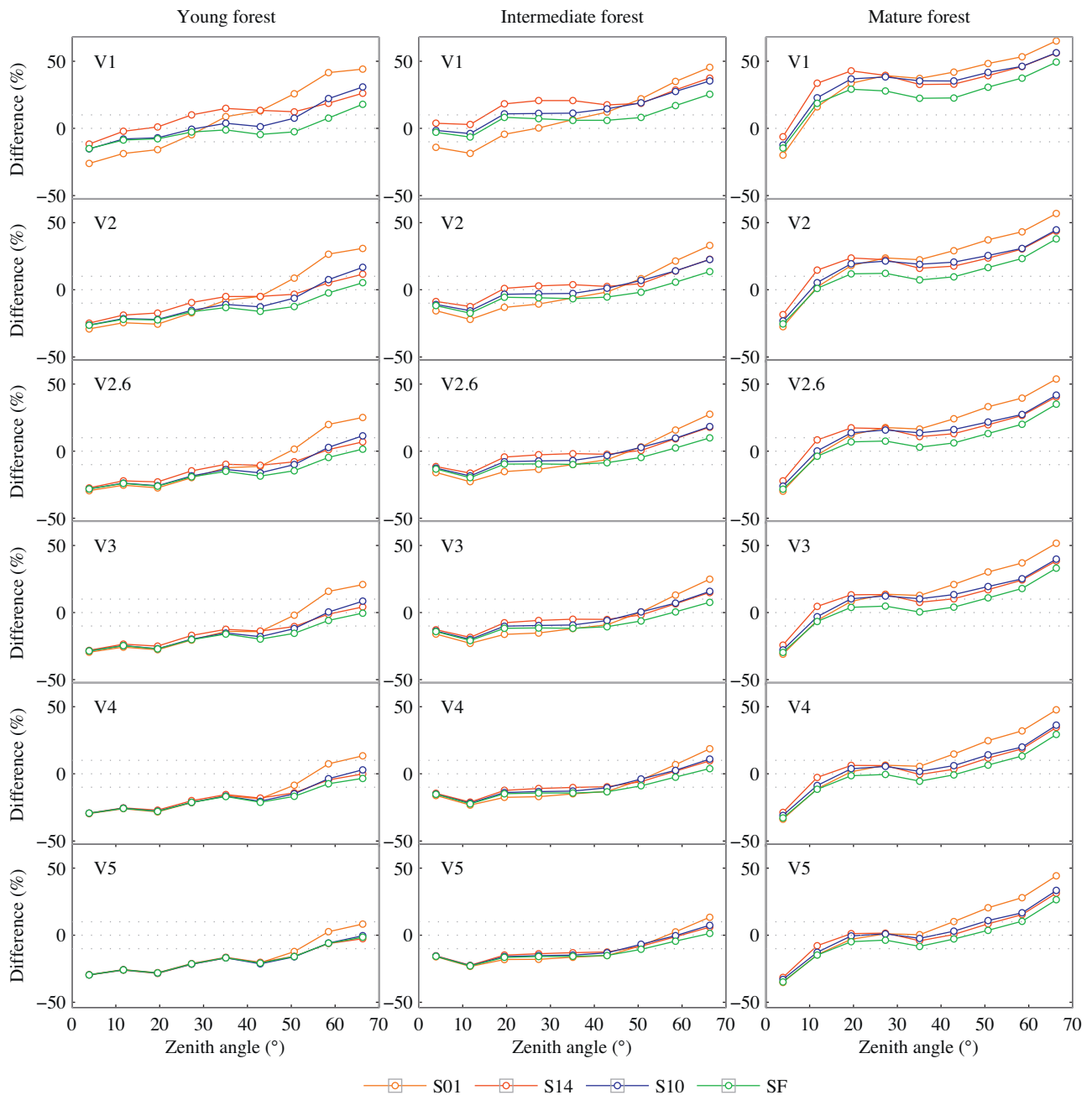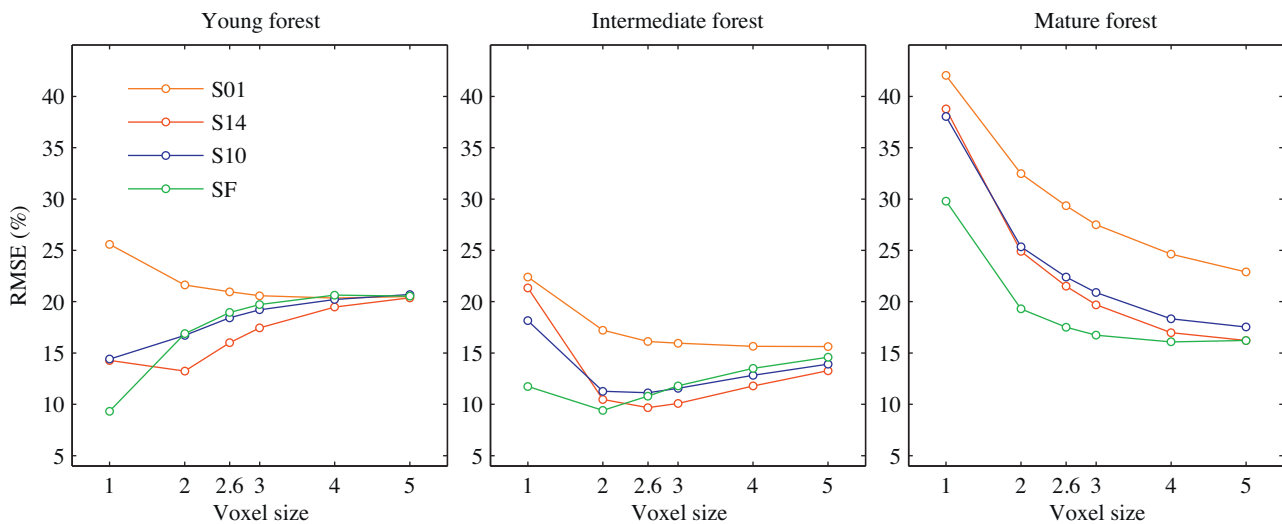
Fig. 5. Difference between terrestrial laser scanner (TLS)-derived gap fraction (GF) and real GF per voxel size in the young, intermediate and mature forest, using setup of registered point clouds in colored lines: 1 m from the center (S01); 14 m from the center (S14); 10 m from the center (S10); and S14 and S10 combined (SF). V1: voxel size 1 cm; V2: voxel size 2 cm, V2.6: voxel size 2.6 cm; V3: voxel size 3 cm; V4: voxel size 4 cm; and V5: voxel size 5 cm.

reflectance objects or that these objects were located beyond the maximum range of the instrument, so that the return intensities are too low for detection (Danson et al., 2007). The latter cause is more likely to occur in sparse and tall forest stands, as it was perceived in the mature forest of this study, but is less relevant and has lowest influence in closed forest environments, as in the young forest under analysis. It may be reasonable, as well, to consider that despite the reported maximum range of the scanner, its actual performance for representing canopy elements declines at ranges under 76 m.

Registered point clouds provide a greatly increased number of returns and yielded an enhanced representation of the elements within the plot. The objects located near the horizon are described in greater detail using SF than using any of the other setups (i.e.,

S01, S14 or S10), due to a higher number of single point clouds used for the registration (Hilker et al., 2012; Van der Zande et al., 2008). More scans, including some taken at greater distances from the area of interest, increase the amount of data in sparse forest and contribute to solve the problem of short zenith range (Seidel et al., 2012). Yet, this will dramatically increase the field work and size of the datasets and will make GF estimations not directly comparable with other instruments (Hilker et al., 2010). Conversely, estimates of other canopy variables (e.g., biomass and canopy cover) may be better calculated with multiple scans due to a high coverage of space and minimized occlusion effect. As mentioned, moving TLS instruments (including supporting bases and reference targets) for multiple scanning is labor intensive and time consuming. Mobile terrestrial laser scanning (MLS; Graham, 2010) for direct 3D

Fig. 6. Root mean square error (RMSE) of the gap fraction (GF) estimates derived from terrestrial laser scanner (TLS) data for each forest stand using different voxel size and registered point clouds: 1 m from the center (S01); 14 m from the center (S14); 10 m from the center (S10); and S14 and S10 combined (SF).

data collection can work as an alternate solution. More convenient for forestry research are unmanned aerial vehicles (UAV), which capture very high resolution data offering improved estimates of several metrics at tree level (Jaakkola et al., 2010; Lin et al., 2011), as well as multi-temporal monitoring of forest health and canopy closure (Wallace et al., 2012). Together with canopy measurements, UAV can be used to collect data at different heights below and within forest canopies, providing valuable 3D data on the location of foliage elements that are embedded in the canopy. In spite of their enhanced spatial coverage and higher measurement efficiency, MLS and UAV bear some limitations, such as platform stability influencing data georeferencing accuracy and sampling resolution due to velocity changes (Lin et al., 2013).

Concerning data quality, the filter applied to the point clouds does not assure ghost points-free datasets (Fig. 2). Balduzzi et al. (2011) report that extra noise might be added when reflective properties of objects and angle of incidence of the laser beam make difficult to the instrument to get enough energy when the laser is reflected. As previously indicated, these ghost points have a cumulative effect when registering multiple point clouds, creating false objects in the simulation and influencing the results of canopy measurements (e.g., bias in GF estimates). An attempt to correct for underestimation of GF caused by ghost points around the edges of canopy gaps is presented in Vaccari et al. (2013). Their approach examines the canopy perimeter for bias correction in GF retrieved by time-of-flight TLS and may be combined with other filtering techniques in order to improve phase based TLS data quality.

The analysis of radiation interception within simulated images (Fig. 4) is especially relevant to assess the effect of voxel size and sampling setup on TLS-derived estimates of GF. The variability of GF distribution indicates that the TLS-derived GF values need a thorough examination before selecting a specific voxel size and sampling setup for canopy assessments. Regardless the difference in GF, there is a clear increment in GF towards higher zenith angles. The behavior of this estimates from simulated hemispherical images gives clear indications of the natural constraints of phase based scanners.

In terms of canopy analysis using single TLS, Hilker et al. (2010) have found good relationships of leaf area index estimates derived from GF between DHP and TLS, calculated from zenith angles 0° to 60° and laser hits recorded at up to 120 m from their instrument. For canopy openness assessments, Seidel et al. (2012) used a range of 30° (zenith angles from 0° to 30°) in their analysis

due to the laser scanner limited range (79 m) and low visibility of some objects at higher zenith angles. The differences between real GF and TLS-derived GF of the present study (Fig. 5), suggest that with an appropriate selection of voxel size and sampling setup for point cloud registration, canopy assessments could be done within a defined zenith angle range from 20° to 45°, minimizing the impact of erroneous estimates of GF at high and low zenith angles (Gonsamo and Pellikka, 2009a; Jonckheere, in press). However, there is still an issue to be solved at zenith angles between 50° and 60° where the information of canopy structure is relevant for further analysis of e.g., leaf area index, leaf area distribution or foliage profiles (Leblanc et al., 2005; Weiss et al., 2004).

When using registered point clouds, a higher coverage of the area is expected (Van der Zande et al., 2011; Seidel et al., 2012), resulting in better estimates of GF (low RMSE) with specific voxel size per stand. Nevertheless, as displayed in Fig. 6, comparable RMSE can be seen on the intermediate forest using both setup SF (i.e., nine scans) and setup S14 (i.e., five scans) with voxel size 2 and 2.6, respectively. Likewise, similar RMSE can be seen on the mature forest using both setup SF and setup S14 with voxel size 5.

5. Concluding remarks

This study was a first step in an objective assessment of the quality of phase based TLS data and the effects of measurement setup and processing of point clouds (i.e., voxelization) on the analysis and comparison of GF as a forest canopy structure descriptor. The impact of voxel size and sampling setup for forest canopy assessments using TLS has been described and quantified. A voxel-based simulation and 3D modeling of three forest stands were used to derive GF with different voxel sizes and sampling setups. The results of this study highlighted the potential of TLS data retrieved from real broadleaved forest for 3D modeling of forest canopies at different degrees of complexity. Through an objective processing procedure, it was found that TLS-derived GF values are highly sensitive to voxel size and stand characteristics. In this study, the following combinations of TLS sampling setups and voxel sizes were found to be suitable for canopy GF assessment: SF with V1 in the young forest; SF with V2, or S14 with V2.6 in the intermediate forest; and SF with V4 or V5, or S14 with V4 or V5 in the mature forest. The findings presented in this paper demonstrate that a registered TLS point cloud, combining nine scans in the registration procedure (i.e., SF), is particularly advised to estimate

GF in young forest stands. It is not recommended, however, in intermediate and mature forest, taking into account the high costs involved and the limited improvement in the results. In spite of its high level of detail, phase based TLS datasets contain some erroneous measurements, that may be caused by reflective properties of the objects, the TLS ranging technology and/or external variables (e.g., wind, particles in the air, ambience light interference). The presence of such erroneous data within the point clouds affects the quality of the results and what is more, has a cumulative effect when registering point clouds from two or more scans. This may be one of the reasons explaining the underestimation of GF at low zenith angles and suggests a need to correctly identify these ghost points and filter them out from the 3D datasets as the very first step in the processing chain of phase based TLS data. Findings on the trend of the differences between zenith angles suggest a global consideration of these variables influencing GF assessments. Further research is under consideration, in order to develop filtering techniques able to correctly identify gaps (non-interceptions) and ghost points (partial interceptions) in phase based TLS data.

Finally, the advantages of TLS, such as having their own energy source, high scanning speed, capabilities for automation of routines and simulation sequences; plus the supplementary information that can be derived from 3D point cloud data (e.g., vertical profile, canopy height, canopy density); keep projecting the TLS as a valuable instrument in the domain of 3D forest canopy research.

Acknowledgments

References

Balduzzi, M., Van Der Zande, D., Stuckens, J., Verstraeten, W., Coppin, P., 2011. The properties of terrestrial laser system intensity for measuring leaf geometries: a case study with conference pear trees (*Pyrus Communis*). Sensors 11 (2), 1657–1681.

Barker, M.G., Pinard, M.A., 2001. Forest canopy research: sampling problems, and some solutions. Plant Ecol. 153, 23–38.

Bréda, N.J.J., 2003. Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. J. Exp. Bot. 54 (392), 2403–2417.

Côté, J.-F., Widlowski, J.L., Fournier, R.A., Verstraete, M.M., 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. Remote Sens. Environ. 113, 1067–1081.

Cifuentes R., Van der Zande D., Farifteh J. and Coppin P. Effects of object reflectivity, scanning settings and resolution on quality of terrestrial laser scanner data from a simulated forest canopy. Unpublished results.

Danson, F.M., Hetherington, D., Morsdorf, F., Koetz, B., Allgöwer, B., 2007. Forest canopy gap fraction from terrestrial laser scanning. IEEE Geosci. Remote Sens. Lett. 4, 157–160.

Dassot, M., Constant, T., Fournier, M., 2011. The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Ann. For. Sci. 68 (5), 959–974.

Fassi, F., Achille, C., Fregonese, L., 2011. Surveying and modelling the main spire of Milan cathedral using multiple data sources. Photogramm. Rec. 26 (136), 462–487.

Frazer, G.W., Trofymow, J.A., Lertzman, K.P., 1997. A method for estimating canopy openness, effective leaf area index, and photosynthetically active photon flux density using hemispherical photography and computerized image analysis techniques, Pacific Forestry Centre. In: Information Report, BC-X-373. Canadian Forestry Service, Victoria, BC.

Frazer, G.W., Canham, C.D., Lertzman, K.P., 1999. Gap light analyzer (GLA), version 2.0: imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs. In: Users Manual And Program Documentation. Simon Fraser UniversityBurnaby, Institute of Ecosystem Studies, British Columbia, Millbrook, New York, NY, Copyright © 1999.

Gonsamo, A., Pellikka, P., 2009a. The computation of foliage clumping index using hemispherical photography. Agric. For. Meteorol. 149, 1781–1787.

Gonsamo, A., Pellikka, P., 2009b. A new look at top-of-canopy gap fraction measurements from high-resolution airborne imagery. EARSeL eProceedings 8 (1), 64–74.

Gonsamo, A., Walter, J.M., Pellikka, P., 2010. Sampling gap fraction and size for estimating leaf area and clumping indices from hemispherical photographs. Can. J. For. Res. 40, 1588–1603.

Graham, L., 2010. Mobile mapping systems overview. Photogramm. Eng. Remote Sens. 76 (3), 222–228.

Hilker, T., Van Leeuwen, M., Coops, N.C., Wulder, M.A., Newnham, G.J., Jupp, D.L.B., Culvenor, D.S., 2010. Comparing canopy metrics derived from terrestrial and airborne laser scanning in a Douglas-fir dominated forest stand. Trees 24, 819–832.

Hilker, T., Coops, N.C., Culvenor, D.S., Newnham, G., Wulder, M.A., Bater, C.W., Siggins, A., 2012. A simple technique for co-registration of terrestrial LiDAR observations for forestry applications. Remote Sens. Lett. 3 (3), 239–247.

Hosoi, F., Omasa, K., 2007. Factors contributing to accuracy in the estimation of the woody canopy leaf area density profile using 3D portable lidar imaging. J. Exp. Bot. 58, 3463–3473.

Jaakkola, A., Hyyppä, J., Kukko, A., Yu, X., Kaartinen, H., Lehtomäki, M., Lin, Y., 2010. A low-cost multi-sensoral mobile mapping system and its feasibility for tree measurements. ISPRS J. Photogramm. Remote Sens. 65, 514–522.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F., 2004. Review of methods for in situ leaf area index determination. Part I: Theories, sensors and hemispherical photography. Agric. For. Meteorol. 121, 19–35.

Jonckheere, I., Nackaerts, K., Muys, B., Coppin, 2005. Assessment of automatic gap fraction estimation of forests from digital hemispherical photography. Agric. For. Meteorol. 132, 96–114.

Jonckheere, I., Nackaerts, K., Muys, B., van Aardt, J., Coppin, P., 2006. A fractal dimension-based modelling approach for studying the effect of leaf distribution on LAI retrieval in forest canopies. Ecol. Modell. 97, 179–195.

Jonckheere I., Macfarlane C. and Walter J.-M. N., Image analysis of hemispherical photographs. Algorithms and calculations. In: R. Fournier and R. Hall, (Eds.), Hemispherical Photography in Forest Science: Theory, Methods, Applications, 2014, Springer-Verlag. In press.

Jupp, D.L.B., Culvenor, D.S., Lovell, J.L., Newnham, G.J., Strahler, A.H., Woodcock, C.E., 2009. Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna®'. Tree Physiol. 29, 171–181.

Korhonen, L., Korpela, I., Heiskanen, J., Maltamo, M., 2011. Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. Remote Sens. Environ. 115, 1065–1080.

Leblanc, S.G., Chen, J.M., Fernandes, R., Deering, D., Conley, A., 2005. Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. Agric. For. Meteorol. 129, 187–207.

Lin, Y., Hyyppä, J., Jaakkola, A., 2011. Mini-UAV-Borne LIDAR for fine-scale mapping. IEEE Geosci. Remote Sens. Lett. 8 (3), 426–430.

Lin, Y., Hyyppä, J., Kukko, A., 2013. Stop-and-Go mode: sensor manipulation as essential as sensor development in terrestrial laser scanning. Sensors 13, 8140–8154.

Lovell, J.J., Jupp, D.L.P., Culvenor, D.S., Coops, N.C., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. Can. J. Remote Sens. 29 (5), 607–622.

Moorthy, I., Miller, J.R., Jimenez Berni, J.A., Zarco-Tejada, P., Hu, B., Chen, J., 2011. Field characterization of olive (*Olea europaea* L.) tree crown architecture using terrestrial laser scanning data. Agric. For. Meteorol. 151, 204–214.

Moskal, M., Zheng, G., 2012. Retrieving forest inventory variables with terrestrial laser scanning (TLS) in urban heterogeneous forest. Remote Sen. 4, 1–20.

Newnham, G., Armston, J., Muir, J., Goodwin, N., Tindall, D., Culvenor, D., Püschel, P., Nyström, M., Johansen, K., 2012. Evaluation of terrestrial laser scanners for measuring vegetation structure. In: CSIRO Sustainable Agriculture Flagship, Manuscript ID: EP124571, https://publications.csiro.au/rpr/pub?pid=csiro:EP124571 [Last accessed: 10/12/2013].

Parker, G., 1995. Structure and microclimate of forest canopies. In: Lowman, M., Nadkarni, N. (Eds.), Forest Canopies. Academic Press, San Diego, CA, pp. 73–106.

Petrie, G., Toth, C., 2009a. Terrestrial laser scanners. In: Shan, J., Toth, C. (Eds.), Topographic Laser Ranging and Processing, Principles and Processing. CRC Press Taylor and Francis Group, Boca Raton, FL, pp. 87–127.

Petrie, G., Toth, C., 2009b. Introduction to laser ranging, profiling, and scanning. In: Shan, J., Toth, C. (Eds.), Topographic Laser Ranging and Processing, Principles and Processing. CRC Press Taylor and Francis Group, Boca Raton, FL, pp. 1–27.

Ridler, T.W., Calvard, S., 1978. Picture thresholding using an iterative selection method. IEEE Trans. Syst., Man, Cybern. 8 (8), 630–632.

Ross, J., 1981. The Radiation Regime and Architecture of Plant Stands. Junk Publishers, The Hague.

Seidel, D., Fleck, S., Leuschner, C., 2012. Analyzing forest canopies with ground-based laser scanning: a comparison with hemispherical photography. Agric. For. Meteorol. 154–155, 1–8.

Vaccari, S., Van Leeuwen, M., Calders, K., Coops, N.C., Herold, M., 2013. Bias in lidar-based canopy gap fraction estimates. Remote Sens. Lett. 4 (4), 391–399.

Van der Zande, D., Hoet, W., Jonckheere, I., van Aardt, J., Coppin, P., 2006. Influence of measurement set-up of ground-based LiDAR for derivation of tree structure. Agric. For. Meteorol. 141, 147–160.

Van der Zande, D., Jonckheere, I., Stuckens, J., Verstraeten, W.W., Coppin, P., 2011. Sampling design of ground-based lidar measurements of forest canopy structure and its effect on shadowing. Can. J. Remote Sens. 34 (6), 526–538.

Van der Zande, D., Stucken, J., Verstraeten, W.W., Mereue, S., Muys, B., Coppin, P., 2011. 3D modeling of light interception in heterogeneous forest canopies using ground-based LiDAR data. Int. J. Appl. Earth Obs. Geoinf. 13, 792–800.

Van Leeuwen, M., Hilker, T., Coops, N.C., Frazer, G., Wulder, M.A., Newnham, G.J., Culvenor, D.S., 2011. Assessment of standing wood and fiber quality using ground and airborne laser scanning: a review. For. Ecol. Manag. 261, 1467–1478.

Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR system with application to forest inventory. Remote Sens. 4, 1519–1543.

Weiss, M., Baret, F., Smith, G.J., Jonckheere, I., Coppin, P., 2004. Review of methods for in situ leaf area index (LAI) determination—Part II: Estimation of LAI, errors and sampling. Agric. For. Meteorol. 121, 37–53.

Welles, J.M., Cohen, S., 1996. Canopy structure measurement by gap fraction analysis using commercial instrumentation. J. Exp. Bot. 47 (302), 1335–1342.

Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Lovell, J., Culvenor, D., Newnham, G., Ni-Meister, W., Schaaf, C., Woodcock, C., Wang, J., Li, X., Strahler, A., 2011. Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. Remote Sens. Environ. 115, 2965–2974.

Zhao, F., Yang, X., Schull, M., Roman-Colon, M., Yao, T., Wang, Z., Zhang, Q., Jupp, D.L.B., Lovell, J.L., Culvenor, D.S., Newnham, G.J., Richardson, A.D., Ni-Meister, W., Schaaf, C.L., Woodcock, C.E., Strahler, A.H., 2011. Measuring effective leaf area index, foliage profile, and stand height in New England forest stands using a full-waveform ground-based lidar. Remote Sens. Environ. 115, 2954–2964.