Impact of the sampling design on the quality of ground-based LiDAR datasets

Dimitry Van der Zande[1], Inge Jonckheere[1], Jan Stuckens[1], Willem Verstraeten[1] & Pol Coppin[1]

[1]Katholieke Universiteit Leuven, Biosystems Departement, M3-BIORES, Celestijnenlaan 200, BE-3001, Leuven, Belgium dimitry.vanderzande@biw.kuleuven.be

Abstract

This research was undertaken to study the influence of the sampling design and laser beam density of ground-based LiDAR measurements on the quality of laser datasets in terms of shadowing. The generation of virtual forest stands by means of stochastic L-systems as tree descriptors were opted for based on the study frame. The dynamic plant modeler and plant nursery natFX (Bionatics, CIRAD, Montpellier, France) was used to simulate full grown forest stands of two tree species (i.e. *Fagus sylvatic* and *Platanus acerifolia*). Next, hemispherical laser measurements with different laser beam densities were simulated according to three different sampling patterns (i.e. single, diamond, and corners) inside these virtual forest stands through the use of ray-tracing technology. An adjusted sampling design has proven its effectiveness since an average decrease of 27.27% in shadowing in comparison with a single measurement was obtained. This contrasts with an average decrease of 13.64% by increase of the laser beams density by a factor 25. Afterwards, contact frequency values, calculated from the virtual laser data sets, were utilized to successfully model the shadowed parts of the canopy demonstrating the potential of ground-based laser scans to capture the 3D leaf distribution inside a full grown forest stand.

Keywords: ground-based LiDAR, virtual forest stands, sampling design, shadow effect

1. Introduction

One of the challenges of Light Detection And Ranging (LiDAR) research in forestry is the quantification of the 3D structure of forest canopies and their components (individual tree crowns) in an accurate and comprehensive manner. The forest canopy is a unique part of the forest ecosystem which fulfills the important role of cycling material and energy through photosynthesis and transpiration, maintaining forest microclimates, and providing habitats for various species (e.g. Erdelen, 1984; Fitzjarrald and Moore, 1995). The description of the structure of this interface between vegetation and atmosphere plays a key role in the understanding of biophysical processes at different levels.

Ground-based LiDAR systems offer unique opportunities in terms of viewing angles and point densities needed to model canopy structure in high detail. The static setup of a ground-based laser scan, in comparison with a moving airborne platform, allows comprehensive beam coverage of the area of interest. Several studies on ground-based LiDAR systems have used one or a combination of single range imagery to make individual tree measurements or plot level summaries (e.g. Radtke and Bolstad, 2001; Parker *et al.*, 2004; Watt and Donoghue, 2005). Hitherto, applications exploring the use of scans acquired from multiple viewpoints to assess the spatial distribution of canopy structure are rare (Henning and Radtke, 2006, Takeda *et al*., 2007, Hosoi and Omasa, 2006). The biomass profile, being the vertical distribution of phyto-elements (leaf, stem, twig, etc.) density above the ground, is the most commonly used parameter to describe

the biomass distribution in the measured forest stands. This vertical structure is often represented by the leaf area density (LAD) per height bin, where LAD is defined as the total one-sided leaf area per unit layer volume (Weiss *et al*. 2004) for a vertical stratified canopy. The LAI/LAD calculation from multi angular laser data can be based on the inclined point quadrat method by Warren-Wilson (1960, 1963) or on gap fraction inversion, a methodology also used to determine LAI of a forest stand using hemispherical photographs. Hosoi and Omasa (2006) described a methodology for voxel based 3D modeling of the LAD by calculating the contact frequency of the laser beams in an arbitrary volume.

The level of detail in which the LAD distribution can generally be determined strongly depends on the laser beam distribution inside the forest canopy. This varies with the distance to the laser device and the leaf density of the measured space which causes shadowing (Van der Zande et al., 2006). In case of a first return laser device, the laser measurement variability along angular viewing differences, or shadowing, is caused by the physical laser pulse/object interactions and is an intrinsic characteristic of the laser device. This is due to reflection of an emitted laser pulse by the first object it encounters. Spatial information of the vegetative elements located behind the target/object is therefore not available due to the shadow effect. Consequently, these background objects have to be measured from different angles to obtain comprehensive laser coverage. When quantifying the distribution of the vegetative elements inside a canopy, this shadowing needs to be minimized, firstly to secure a certain accuracy of the contact frequency measurements per volume of choice and secondly to minimize the 'blind spots' or areas of which no information could be gathered.

To enable an accurate study of the interaction of laser beams with a complex organized object like a canopy, a detailed description of the 3D organization of the vegetative elements in the canopy is required. The lack of detailed, consistent, and precise reference information of forest structure hinders this approach. A solution is offered by simulation techniques which enable the reconstruction of forest stands in a virtual environment. Structural aspects such as leaf surface distribution are calculated directly during the simulation process, resulting in the generation of accurate reference data of the virtual forest stand. LiDAR range images can subsequently be acquired using ray tracing algorithms. Ray tracing algorithms are based on tracing the path of a ray of light through a scene as it interacts with objects in an environment and therefore strongly resembles the LiDAR principles. This technique allows detailed studies of light beam/canopy interactions and their effect on the quality of the LiDAR measurements which is only limited by the degree of complexity of the reconstructed forest stand. From remote sensing modeling point of view, it is important to obtain realistic descriptions of the forest stand, which complies with three main criteria: (1) a description based on architectural growth processes capable of simulating various tree species over various conditions (age, density, environment, etc.); (2) a description based on experimental data; and (3) a description capable of providing realistic 3D trees. The AMAP model (Atelier de Modelisation et d'Architecture des Plantes), developed by CIRAD (Montpellier, France), meets these criteria making it a valuable tool for simulating forest stands for quality testing of LiDAR measurement protocols. The AMAP model is a growth simulation software that respects the plant's genetic coding and reconstitutes the tree morphology and natural esthetics in synthetic 3D images. These canopy models have gained acceptance as a research tool in forestry and have led to increasingly convincing visualizations due to recent developments in information technology, more specifically in the field of simulation technologies (Jonckheere *et al*, 2006).

This research studies the influence of the sampling design and laser beam density of ground-based LiDAR measurements on the quality of the collected laser datasets in terms of shadowing. A total of three virtual forest stands were generated, varying in tree species (i.e.*Fagus sylvatica*, *Platanus acerifolia*) and thus structural built-up. Next, ray-tracing technology allowed for the simulation of hemispherical laser measurements, with varying laser beam densities, inside these virtual forest

stands. These individual laser measurements were simulated in three different sampling patterns (i.e. single, diamond, and corners) to determine the optimal sampling design guarantying a minimal shadow effect.

2. Materials and Method

2.1. 3-D canopy simulation

Two homogeneous forest plots with fully grown broadleaf trees (*Fagus sylvatica, Platanus acerifolia*) were simulated in a virtual environment using the dynamic plant modeler and plant nursery natFX (Bionatics, CIRAD, Montpellier, France) that interfaces with an architectural plant model called AMAP (CIRAD, Montpellier, France). A forest plot consisted of four individual trees of the same specie placed in a regular pattern in a 3D volume with a ground surface of 15 m side and a height of 30 m. The total LAI of the forest plots was calculated during the simulation process and was 4.50 for the Fagus stand and 5.15 for the Platanus stand. The architectural differences between the species (Fig. 1) resulted in structural diversity between the forest plots enabling a robust analysis of the factors influencing shadowing and thus the quality of the LiDAR datasets.

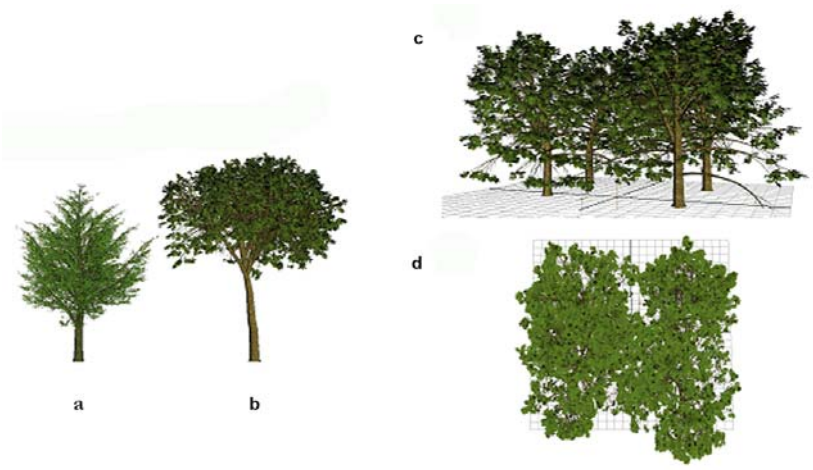

Fig. 1. Fully grown broadleaf trees of three species were simulated in a virtual environment using the Bionatics plant nursery, natFX: (a) *Fagus sylvatica*, (b) *Platanus acerifolia*. A forest plot consisted of four individual trees of the same specie placed in a regular pattern (c,d).

2.2. Laser measurement simulation

Tracking of the virtual laser beams through the canopy was done using the physically based rendering theory (pbrt, Pharr & Humphreys, 2004) as ray-tracer algorithm. Viewing rays can be shot into the scene to see whether they interact with any of the objects in the scene (Pharr and Humphreys, 2004), which strongly resemble the LiDAR principles. Specifications of the commercially available Laser Measurement System 200 (LMS200, Sick AG, Germany) were used as a standard template to characterize simulated laser beams (i.e. wavelength and beam divergence) and hemispherical measurement pattern (GMP). The LMS200 is a non-contact optical active sensor which scans its direct surroundings in a 2D pattern. By mounting it on a dynamic measurement platform like a rotating table, a 3D hemispherical measurement pattern is enabled. The full description of the characteristics of the laser device and measurement platform can be found in Van der Zande et al. (2006). To reproduce the hemispherical GMP for the virtual laser beams, the environmental camera of the pbrt-package was selected. This camera traces rays in all directions around a specific point in the scene (i.e. laser measurement location). Each beam

was described by its polar coordinates (φ,α), where the zenith angle (α) ranged from 0 to π /2 and the azimuth angle (φ) from 0 to 2π. The zenith resolution was fixed at 0.25° while the azimuth resolution was set at 0.1° and exceptionally at 0.02° for the central standard measurements. Different laser beam densities were acquired by adjusting the azimuth resolution of a laser scan by considering portions of the emitted laser beams. Each laser beam was traced up to the first vegetative element (represented as a collection of triangles) it interacted with. As the LMS200 was used as a template, at least 10% of the emitted light energy had to be reflected in order for the virtual laser system to register a distance measurement (Sick, AG).

2.3. Sampling design

By measuring a canopy scene from different directions the probability that a certain vegetative element (e.g. leaf) is reached by at least one laser beam increases since a more comprehensive laser beam coverage of the measured object is obtained. In this study, three different sampling strategies were investigated: (1) single, (2) diamond and (3) corners (Fig.2).

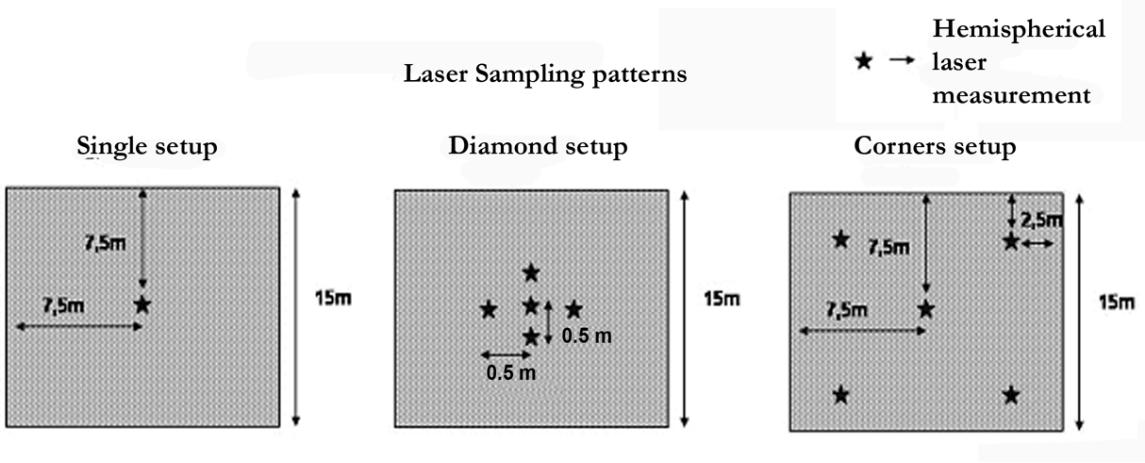

Fig. 2. Illustration of the three different sampling designs tested in this study: single, diamond and corners. The last two setups consisted of five individual hemispherical laser measurements positioned in a specific geometric format while the central setup function as a reference measurement.

The last two sampling designs consisted of five individual hemispherical laser measurements positioned in a specific geometric format (Fig.2) in an effort to minimize the shadow effect. By altering the azimuth resolution of the individual hemispherical measurement, different laser beams densities were available per measurement setup (Table 1). For comparison purposes, the azimuth resolution of the central single measurement was chosen so that the total amount of emitted laser beams matched that of the combination of the five separate laser scans of the diamond and corners setups.

Table 1. Alternation of the azimuth resolution of the hemispherical laser scans resulted in different laser beams densities per measurement

# laser beams (millions)	Single (azimuth angle)	Diamond & Corners (azimuth angle)
6.48	0.02°	0.10°
3.24	0.04°	0.20°
1.30	0.10°	0.50°
0.65	0.20°	1.00°
0.26	0.50°	2.50°

2.4. Laser data processing

A single hemispherical laser measurement consisted of 259,200 up to 6,480,000 separate distance measurements depending on the resolution set up (Table 1). Each laser beam was characterized by a zenith angle, an azimuth angle and a beam divergence. The voxel-based contact frequency (Hosoi and Omasa, 2007) was generated in three steps from the virtual LiDAR datasets:

1) Registration: The five datasets of the diamond and corners setup, with their own coordinate system, were registered into a single comprehensive laser data set or 3D point cloud using a standard translation and rotation algorithm based on their know relative positions;
2) Voxelization: The 3D space considered was arbitrarily subdivided into 'small' cubic voxels of 10 cm in side. This resulted in voxel arrays of 150 x 150 x 250 voxels. Following the methodology of Hosoi and Omasa (2006), the 'small' voxels where characterized depending on beam/voxel interaction. For voxels with at least one intercepted laser beam attribute 1 was assigned. Attribute 2 was assigned to voxels that were intersected by laser beams without interception. Attribute 3 was granted to voxels that were not touched by any laser beam.
3) Contact frequency calculations based on LiDAR measurements: The 'small' voxels were grouped into 'large' voxels of 100 cm in side consisting of 1000 'small' voxels each. The contact frequency (CF) for each 'large' voxel was calculated as follows:

$$CF(\theta) = n_I/(n_I+n_P) \quad (1)$$

With n_I the number of 'small' voxels with attribute 1 and n_P the number of voxels with attribute 2. The contact frequency was calculated for each 'large' voxel and could then be extrapolated to the small voxels which could not be reached by any beam (i.e. voxels with attribute 3). Only the laser beams exiting a voxel in the opposite side as from which they entered were considered as passing beams. This ensured a minimal traveling path of 10 cm through the voxel for an accurate contact frequency calculation.

3. Results

3.1. Relative shadow effect

The extent of the shadow effect was investigated by determining the number of effectively 'filled' voxels that were hit by at least one laser beam that interacted with the leaf material (i.e. voxels with attribute 1). The relative shadow effect (RS) was calculated as the proportion of 'filled' voxels which were not seen by the laser system, these filled voxels were incorrectly given the attribute 2 or 3 instead of attribute 1. Fig. 3 shows the RS-values for each of the different sampling patterns in the two forest stands and this for five resolution settings.

The average shadow effect for a single laser scan ranged from 68.74% to 82.38% depending on the beam density. This means that an increase of the number of laser beams with a factor of six reduced the shadow effect by 13.64%. The logarithmic character of the decrease of RS with increasing number of laser beams demonstrated that tackling the shadowing problem by sheer hardware improvement would not be cost-efficient. These results support the need for adjusted sampling designs of multiple laser scans from different locations to improve the probability that voxels would be reached by at least one laser beam. In the case of the diamond set-up the average RS-values ranged from 46.67% to 74.4%. This decrease of 27.75% due to increasing resolution settings differed significantly from the decrease of 13.64% in the case of a single measurement. These results also show that the use of five low resolution scans instead of one high resolution scan, with a similar amount of total laser beams, significantly reduced the

shadow effect with 27.75%. Considering the RS-values, the corners sampling design showed results similar to the diamond set-up. Fig. 4.b illustrates the direct comparison of the vertical distribution of the filled voxels (attribute 1) derived from the LiDAR datasets with the reference profile and this for the different sampling designs. The shadow effect becomes visible as the measured profiles show an underestimation compared to the reference profile. This underestimation is considerably higher in case of the single LiDAR measurement compared to the diamond and corners setups which is consistent with previous results (Fig. 4). The general shape of the reference profile is observable in the three measured profiles showing the potential of laser systems, like the LMS200, to capture essential structure information which could be used to model the actual leaf distribution.

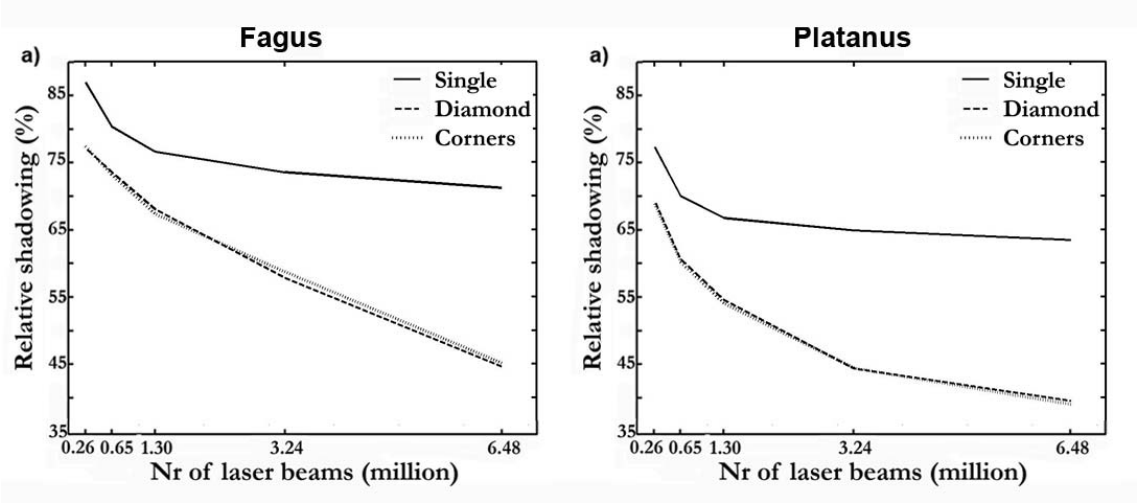

Fig. 3. The relative shadowing (RS) decreased significantly as a function of the different sampling setups for the two virtual forest stands ((a) Fagus and (b) Platanus). The diamond and corners setup were compared to the single design.

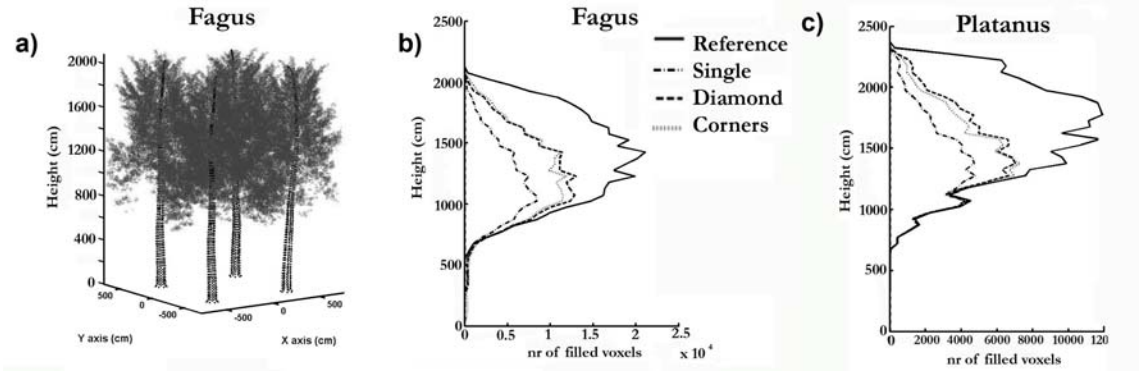

Fig. 4. Direct comparison of the vertical distribution of the filled voxels (attribute 1) derived from the LiDAR datasets with the reference profile for the different sampling designs (Fagus (a,b), and Platanus (c)).

3.2. 3D distribution of the shadow effect

The 3D distribution of the measured filled voxels, and thus also the 3D shadowing, were visualized as a collection of horizontal slices for the different sampling patterns in the Fagus stand (Fig. 5). The availability of this detailed 3D description of the shadow effect enables a more thorough analysis of the actual laser beam/canopy interaction. Where the study of the

relative shadowing did not reveal a significant difference between the diamond and corners setup, this 3D study exposed that the central area suffered mostly from shadowing and that this is more profoundly present in the corners measurements. The peripheral areas on the other hand, are less affected by the shadowing using this last set-up.

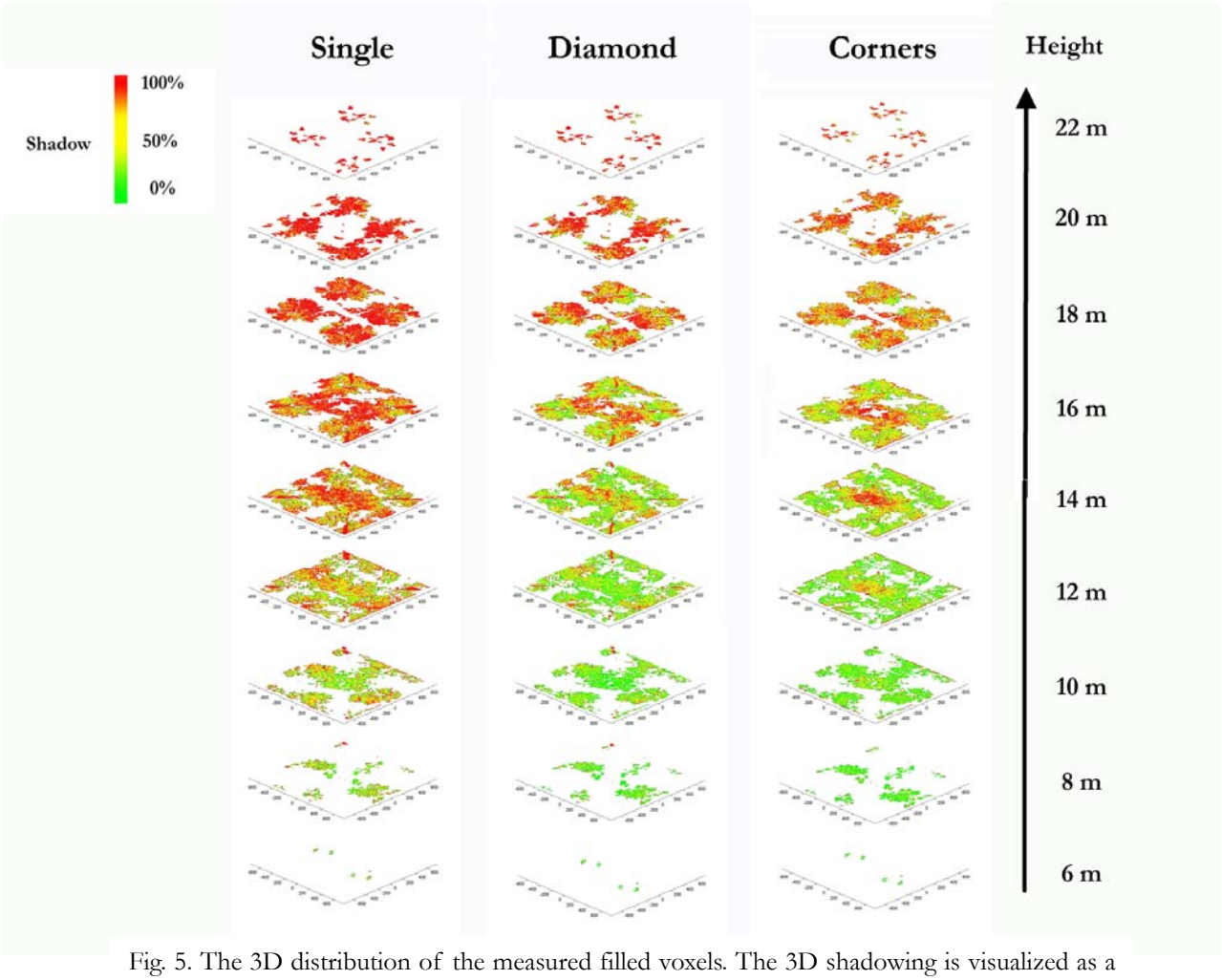

Fig. 5. The 3D distribution of the measured filled voxels. The 3D shadowing is visualized as a collection of horizontal slices for the different sampling patterns in the Fagus stand.

3.3. Contact frequency calculations for modelling purposes

Whereas an adjusted sampling design shows great potential in minimizing this effect, values of up to 55% are still detected. This means that for an accurate structural description of the actual leaf distribution, at least 55% of the canopy needs to be modeled using the information extracted from the measured parts of the canopy. Up to this point only the voxels with attribute 1 were used for the analysis enabling the shadow effect mapping. In order to correct for the shadow effect, more information needs to be extracted from the LiDAR datasets. This is done by including the empty voxels (voxels with attribute 2) in the further analysis. The method presented by Hosoi and Omasa (2006) introduced the calculation of contact frequencies per unit of volume from the ratio calculated using Eq. 1. This ratio is then extrapolated to the voxels in that unit of volume for which no information is available (i.e. voxels with attribute 3). This methodology was used to study the potential capacity of a LiDAR scanner to estimate the real leaf distribution in forest canopies. The large voxels that had a coverage of 100% were analyzed as an accuracy assessment for the estimation of the leaf density inside those large voxels. For

the diamond setup in the Fagus stand, a total of 1807 voxels (37% of total) were available for analysis. A linear regression between the calculated contact frequency and the actual leaf density resulted in a R^2-value of 0.97. The linear regression model showed an underestimation of 22% which could only be caused by leaf/laser beam interactions since every small voxel in the considered large voxels had been scanned and no direct shadow effect was present. The large voxels were grouped according to their laser coverage and on each of these groups this accuracy assessment described above was repeated. Table 2 presents the linear relations between the contact frequencies, calculated from the diamond and corners LiDAR datasets, and reference datasets. The decreasing laser coverage results in decreasing slopes of the linear regressions. This indicates that the degree of underestimation of the leaf densities from contact frequencies increased with shadowing. While these underestimations reach values up to 89.00%, the R^2 values indicated that even with low laser coverage an accurate estimation of the leaf density is possible.

Table 2. The slope and R^2-values of the linear relations between the contact frequencies, calculated from the diamond and corners LiDAR datasets, and reference datasets are presented.

	Fagus				Platanus			
	Diamond		Corners		Diamond		Corners	
Laser Coverage	slope	R^2	slope	R^2	slope	R^2	slope	R^2
Full	0.78	0.97	0.80	0.98	0.90	0.94	1.00	0.99
80%-100%	0.45	0.86	0.52	0.89	0.59	0.82	0.85	0.81
60%-80%	0.21	0.85	0.29	0.90	0.39	0.78	0.74	0.84
40%-60%	0.15	0.86	0.21	0.85	0.39	0.78	0.71	0.83
20%-40%	0.09	0.74	0.12	0.71	0.37	0.69	0.90	0.60
0%-20%	***	***	***	***	***	***	1.38	0.39

Using these findings, two sets of correction factors per laser coverage class were extracted enabling the correction of the leaf density estimation based on measured contact frequencies. Fig. 6 demonstrates the corrected profiles in comparison with the reference profile. This proved the potential of the ground-based LiDAR technology to measure complex structure of objects such as forest canopies. Even when shadowing and leaf/laser beam interactions are responsible for the fact that almost 55% of the leaves could not be measured, the LiDAR datasets still contain enough information to accurately describe the distribution of vegetative elements in a 3D space.

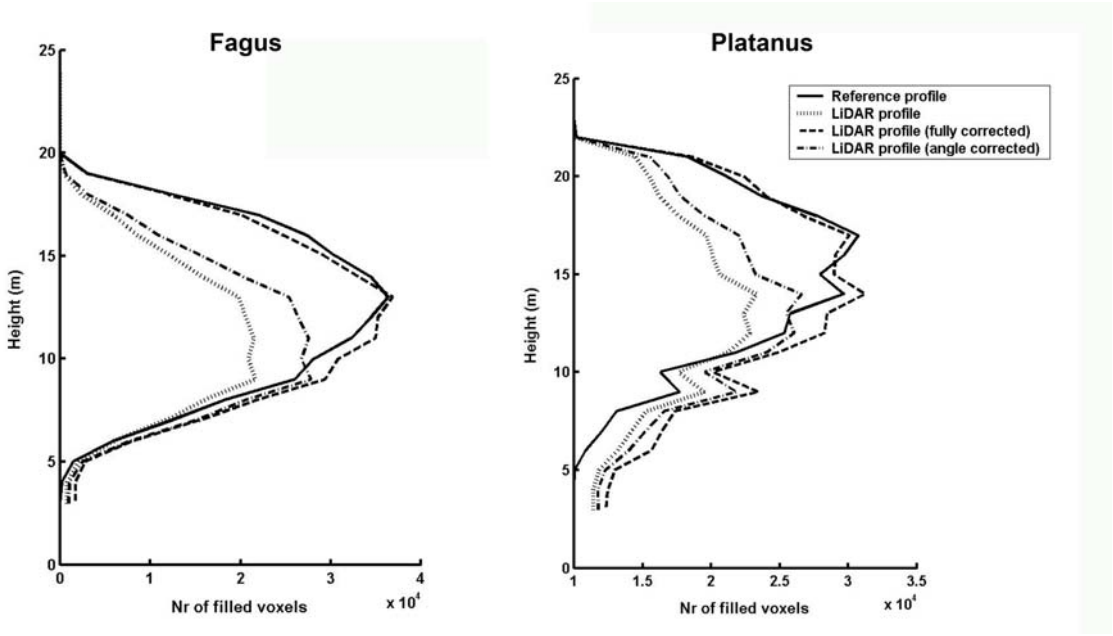

Fig. 6. Comparison of the corrected (angle corrections and full corrections) with the reference profile.

4. Discussion

This paper explored the potential of simulation and ray tracing techniques for structural algorithm development for LiDAR datasets. The results showed substantial improvements in the quality of the datasets when measuring a forest stand from different locations in comparison with a single measurement. This allowed laser beams to enter the canopy under a variety of angles and directions and that increased the probability of a laser beam to penetrate the canopy deeper than it would be possible when measuring from a single location. This caused a decrease of the shadow effect, as it also enabled more accurate density estimates of the shadowed parts of the canopy. In order to maximize the quality of the LiDAR datasets, these results suggest a combination of the diamond and corners sampling setups. However, an increase in the number of separate laser scans would in reality imply more labor and time consuming field campaigns, a factor of unimportance in virtual LiDAR studies. The registration of the separate laser scans to one comprehensive scan was errorless in this controlled environment while under real circumstances the combined registration error could negatively affect the results. Hence, a delicate balance should be pursued between the number of measurement positions and the shadow effect.

Acknowledgements

We would like to acknowledge the useful technical information about the LMS200 by Sick, A.G. (Germany), and their willingness to contribute to this project. Funding support for this project has been provided by FWO Flanders (Project no. G.0407.05). Finally, credit is due to Cas Van der Zande for creating the right atmosphere for the writing of this paper.

References

Erdelen, M., 1984. Bird communities and vegetation structure. I. Correlations and comparisons of simple and diversity indices. *Oecologia*, 61, 277–284.

Fitzjarrald, D.R., and Moore, K.E., 1995. Physical mechanisms of heat and mass exchange between forest and the atmosphere. In: M. Lowman, and N. Nadkarni (Eds.). *Forest Canopies*. Academic Press, California, USA, 45–72.

Henning, J.G., and Radtke P.J., 2006. Ground-based laser imaging for assessing three-dimensional forest canopy structure. *Photogrammetric Engineering and Remote Sensing*, 72 (12), 1349-1358.

Hosoi, F., and Omasa, K., 2006. Voxel-Based 3-D Modeling of Individual Trees for Estimating Leaf Area Density Using High-Resolution Portable Scanning Lidar. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (12), 3610-3618.

Jonckheere, I., Nackaerts, K., Muys, B., Van Aardt, J., and Coppin, P., 2006. A fractal dimension based modelling approach for studying the effect of leaf distribution on LAI retrievel in forest canopies. *Ecological Modelling*, 197 (1-2), 179-195.

Parker, G.G., 1995, Structure and microclimate of forest canopies. In: M. Lowman, and N. Nadkarni (Eds.). *Forest Canopies*. Academic Press, California, USA, 73–106.

Parker, G.G., Harding, J.H., and Berger, M.L., 2004. A portable LIDAR system for rapid determination of forest canopy structure. *Journal of Applied Ecology*, 41, 755–767.

Phar, M. and Humphreys, G. (Eds.), 2004. *Physically Based Rendering: from Theory to Implementation*. Morgan Kaufmann Publishers, San Fransisco, USA.

Radtke, P.J. and Bolstad, P.V., 2001. Laser quadrat sampling for estimating foliage-height profiles in broad-leaved forests. *Canadian Journal of Forest Research,* 31, 410–418.

Takeda, T., Oguma, H., Sano, T., Yone, Y., and Fujinuma, Y., 2007. Estimating the plant area density of a Japanese larch (Larix kaempferi Sarg.) plantation using a ground-based laser scanner. *Agricultural and Forest Meteorology*. In press.

Van der Zande, D., Hoet, W., Jonckheere, I., van Aardt J., and Coppin, P., 2006. Influence of measurement set-up of ground-based LiDAR for derivation of tree structure. *Agricultural and Forest Meteorology,* 141 (2), 147-160.

Warren-Wilson J., 1960. Inclined point quadrats. *New Phytologist*, 59, 1-8.

Warren-Wilson J., 1963. Estimation of foliage denseness and foliage angle by inclined point quadrats. *Australian Journal of Botany*, 11, 95-105.

Watt, P. and Donoghue, D., 2005. Measuring forest structure with terrestrial laser scanning. *International Journal of Remote Sensing*, 26 (7), 1437–1446.

Weiss, M., Baret, F., Smith, G.J., Jonckheere, I., and Coppin, P., 2004. Review of methods for in situ leaf area index (LAI) determination. Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology*, 121 (1-2), 37-53.