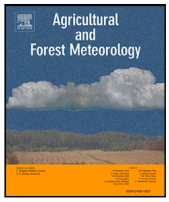
Agricultural and Forest Meteorology 358 (2024) 110218

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet



Questioning voxel grids: Semi-continuous sampling of leaf area density using airborne waveform lidar in boreal and hemiboreal conifer and broadleaved forests

Daniel Schraik [a,b], Aarne Hovi [a,*], Miina Rautiainen [a]

[a] Aalto University, School of Engineering, Department of Built Environment, PO Box 14100, Aalto, 00076, Finland
[b] Natural Resources Institute Finland, Latokartanonkaari 9, Helsinki, 00790, Finland

ARTICLE INFO

Keywords:
Airborne laser scanning
Plant area density
Forest canopy structure
Ray tracing
Clumping

ABSTRACT

Plant area density measurements provide spatially explicit information about the density and distribution of canopy elements. This information is needed for modeling of the forest radiation regime, climate and for other ecological applications. Terrestrial laser scanning (TLS) provides detailed information about canopy structure, but it cannot be used for monitoring large areas. Airborne laser scanning (ALS) uses similar methods to measure plant area density, but due to the larger beam footprints, the scale at which this information can be obtained is coarser than with TLS. The volumetric nature of the ALS measurement poses unique geometric challenges to plant area measurement methods, as assuming an infinitesimal beam size may lead to large errors. Further, the use of voxel grids with ALS measurements may increase errors in plant area measurements, as these grids require discrete spatial allocation of information.

In this study, we apply a spatial weighting technique to ray-traced measurements of plant area from ALS data. This spatial weighting scheme allows continuous allocation of trajectory information of ALS pulses, avoiding discontinuity introduced by voxel grids.

Our data consisted of high density ALS waveform data (over 40 points/m^2) in 33 plots across two study sites in Finland and Estonia. We compared the plant area index (PAI) obtained through this new measurement method to PAI measurements from hemispheric photography (HP) and TLS, and to ALS with a voxel grid. We found PAI, measured at agrid spacing of 0.6 m, correspond best to HP and TLS measurements. Occlusion severely biased PAI at 0.2 m spacing. With increasing grid spacing, PAI estimates become increasingly biased because of clumping effects at small scales. Continuously sampled PAI measurements corresponded closer to reference values than voxel-based PAIs, indicating that a spatially weighted approach avoids bias from partitioning the volumetric ALS beams into voxels.

1. Introduction

Plant area density measurements provide spatially explicit information about the density and distribution of canopy elements. This information is needed for modeling of the forest radiation regime, climate and microclimate, and for other ecological applications. Terrestrial laser scanning has seen significant advances in method development and measurements (e.g. Béland et al., 2014; Pimont et al., 2018; Schraik et al., 2023). TLS is among the most attractive remote sensing techniques for plant area density measurement due to its high resolution measurements with small beam sizes. However, spatial coverage is, relative to ALS, extremely limited by the high workload of TLS campaigns (e.g Calders et al., 2018).

Airborne laser scanning provides this coverage over large areas, and indeed has been used a number of times to measure PAD (e.g. Kamoske et al., 2019; Yin et al., 2022). The main limitation of ALS comes from its typical measurement range, causing large beam footprints in the order of tens of centimeters that limit the scale at which three-dimensional estimates can be obtained. Usually, a voxel-based approach for measuring PAD is used in both ALS and TLS where lidar beams are traced through the canopy, and PAD is estimated from the interceptance information and path lengths that rays travel through each voxel. While in TLS voxel side lengths can be as small as tens of centimeters (Schraik et al., 2021a), ALS studies have used voxels with side length of 0.5 meter or more (e.g. Yin et al., 2022). The relatively large voxel size in ALS-based PAD measurements is necessary when using a ray tracing approach.

* Corresponding author.
E-mail address: aarne.hovi@aalto.fi (A. Hovi).

https://doi.org/10.1016/j.agrformet.2024.110218
Received 7 March 2024; Received in revised form 4 July 2024; Accepted 3 September 2024
Available online 11 September 2024

The voxel size is a user-chosen parameter, and it poses a trade-off between occlusion and spatial detail. In small voxels, biases may occur as fewer rays traverse the voxel and thus insufficiently explore the volume (Pimont et al., 2018). On the other hand, in large voxels, spatial detail is lost as canopy structure below the voxel scale is not quantified.

Recently, volumetric ray tracing approaches have been developed, like the path volume method by Yin et al. (2022). In this method, the footprint area of a beam is accounted for as the central beam trajectory vector is traced through a voxel grid. This avoids errors from partial hits due to the large beam footprint, however, if the beam traverses near a voxel boundary, the entire attenuation is still assigned only to the voxel which the central beam vector traverses, and not to both adjacent voxels. To advance volumetric ray tracing, the next step would be to assign a hit across the volume from which it originated, rather than a single point or interval along the beam's trajectory vector. We suspect that in cases where beams do not traverse a voxel near the voxel center, assigning the entire hit as being intercepted within the focal voxel may lead to errors, as the attenuating media may be distributed arbitrarily across the beam cross-section. An exaggerated example would be a beam traversing very close, and parallel to the boundary between two voxels. In this case, the entire attenuation would be assigned to either voxel, while the attenuating vegetation elements could be distributed somewhat evenly between the voxels. Ignoring the continuous nature of the underlying spatial sampling process could thus lead to errors in estimating plant area density.

In this study, we explore an inverse distance weighted sampling approach to PAD measurement from ALS as an alternative to voxel grid sampling. In this approach, spatial sampling is done on a point grid, and each beam's waveform locations near the grid point is assigned a non-overlapping weight, that is, for every point in space, the sum of weights for all nearby grid points is 1. We used this approach to measure PAD from very high density ALS waveform data in 33 plots in boreal and hemiboreal study sites and compare the results to PAD from hemispheric photography and TLS. We also compare our continuous grid sampling approach to a standard voxel sampling approach. Finally, we analyzed how fine measurement scales are possible until occlusion limits sufficient sampling, given the high density of our point cloud data of over 40 points per m^2.

2. Materials and methods

2.1. Study sites

Our two study sites were located in Hyytiälä, Finland, and Järvselja, Estonia. Of our 33 plots, 20 were in Hyytiälä and 13 in Järvselja. Hyytiälä is in the boreal region, and stands were dominated by Norway spruce (*Picea abies* (L.) H. Karst.), Scots pine (*Pinus sylvestris* L.) or silver and downy birch (*Betula pendula* Roth, *B. pubescens* Ehrh.). Mean diameter at breast height and tree height ranged from 7 to 45 cm (mean 23 cm) and 7 to 35 m (mean 22 m), respectively. Stands were mostly single species, except for three mixed species plots, and even-aged managed forests. Järvselja is in the hemiboreal region with predominantly mixed species stands. Tree species included Scots pine, Norway spruce, silver and downy birch, European alder (*Alnus glutinosa* (L.) Gaertn.), European aspen (*Populus tremula* L.), littleleaf linden (*Tilia cordata* Mill.), goat willow (*Salix caprea* L.), and European ash (*Fraxinus excelsior* L.). Diameter at breast height and tree height ranged from 4 to 40 cm (mean 17 cm) and 4 to 39 m (mean 19 m), respectively. Stands were even-aged and managed, and species composition contained single species, a single dominant species with minor occurence of other species, to even mixtures between two or three tree species. The field measurements in Finland were done during June and July 2019, with a coincident airborne campaign on 13th July. In Estonia, the airborne campaign was carried out two days after the Finnish campaign, but the field campaign was conducted in July 2020.

Field measurements consisted of terrestrial lidar scans (TLS) and hemispheric photographs (HP) covering a 30 × 30 m area with the same plot center, and the analysis was done on a 25 × 25 m area. Both TLS and HP measurements were taken on 16 locations in a square grid pattern with a spacing of 10 m. Leaf area index (with woody elements removed) from TLS data was obtained with ray tracing a standard voxel grid with 20 cm side length. Each individual TLS scan produced a voxel grid of leaf area density, all of which were averaged to produce a single voxel grid that was aggregated to obtain leaf area index. HPs were used to calculate effective PAI (Chen and Cihlar, 1995) from angular gap fraction using five zenith angle rings. TLS and HP methods and the study sites are described in more detail in Schraik et al. (2023) and Hovi et al. (2022) (see Fig. 3 therein for a schematic of the measurements within each plot).

Airborne lidar waveform (ALS) data was collected by a Riegl LMS Q-780. The instrument's laser has a wavelength of 1064 nm, beam divergence of 0.25 mrad ($1/e^2$ width), and a sample spacing of 1 ns. The point density in both study sites was 40–60 points per m^2 in the field plots. The field plots' coordinates were determined by combining TLS and ALS data to sub-meter accuracy (Hovi et al., 2022). The TLS measurements of leaf area density were obtained in plot-wise local coordinate systems. These local coordinate systems used the diagonal of the (North-aligned) grid points. Due to the grid points being marked in the field using a compass, there may be small translation (approximately less than 1 m) and rotation (less than 5 degrees) differences between the TLS and ALS plots which prevent a comparison of the two 3D PAD measurements at the voxel level. The plots were located at least 30 m from the nearest stand border, therefore we expect any differences to be marginal and unsystematic.

In the following, we occasionally present results separately for the two study sites, but we note that this distinction does not indicate that the sites are necessarily different from each other. The plots in Järvselja contained more broadleaf-dominated stands with more species diversity than the plots in Hyytiälä, which is more conifer dominated. The plots were selected to cover a wide range of forest structure and composition, not to provide statistically representative samples of either site.

2.2. Method overview

We estimated plant area index (PAI) for all plots from airborne waveform lidar data using spatially explicit estimates of plant area density (PAD). PAI and PAD are half of the total surface area of canopy elements per unit ground area and per unit volume, respectively. PAD is estimated for subvolumes of the canopy, and the sum of the total plant area divided by the plot area is the PAI.

The PAD for a given volume is calculated from the path lengths z_i that N beams travel through the volume, and the fractions of the beams $\delta\mathrm{CF}_i$ that are intercepted within the volume as Pimont et al. (2018), Schraik et al. (2021b, 2023)

$$\mathrm{PAD} = \frac{\sum_{i=1}^{N} \delta\mathrm{CF}_i}{\sum_{i=1}^{N} z_i}. \tag{1}$$

For waveform data, δCF at a given range r is proportional to the return intensity I after deconvolution with the system pulse as

$$\delta\mathrm{CF}(r) = \frac{I'(r)}{1 - \int_0^r I'(x)\delta x}, \tag{2}$$

$$I'(r) = \frac{I(r)}{\int_0^R I(x)\delta x}, \tag{3}$$

where the denominator in Eq. (2) denotes the correction for attenuation until range r of the normalized intensity $I\prime$, and Eq. (3) denotes the intensity normalization over all returns in the waveform. We used a Gaussian decomposition to obtain a set of returns with their position and variance. The system pulse was measured by the instrument, and due to its Gaussian properties, deconvolution was done by subtracting

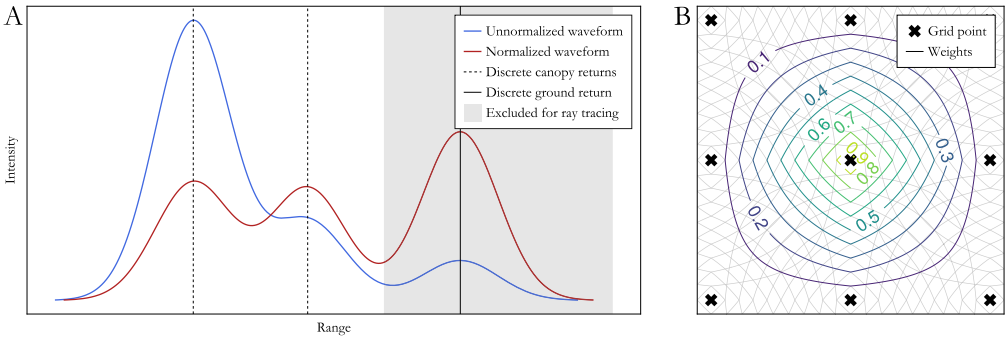


Fig. 1. Schematic of the waveform normalization and attenuation correction (A) and the spatial weighting scheme (B). The intensity is divided by the inverse of the cumulative intensity to correct for attenuation, and subsequently normalized to one. This process is done for all returns, including ground returns, however ground returns are excluded in subsequent ray tracing. For rays traveling near a grid point, a weight is applied for each point on the ray based on the distance to the grid point.

the variance of the system pulse from each return's variance. Finally, we calculated for each ray the distance at which it reached a height of 1.5 m above ground. The ground was represented by a digital elevation model that was created with the same data. The threshold at 1.5 m above ground level was used to exclude ground returns in each beam's CF distribution.

2.2.1. Spatial sampling scheme

Instead of discretely partitioning the canopy volume into a voxel grid, we defined a point grid with parameters similar to a voxel grid, i.e., cubic point spacing. We estimated the PAD around each point using an inverse distance weighting approach along each beam's waveform. Each point P_i along the ALS beam quantified a local change in return intensity, which is proportional to the fraction of the beam footprint that intercepts energy. This fraction is equivalent to the cover fraction, and therefore can be used to locally estimate PAD (Eq. (1)). We estimated PAD for each grid point using a weighted integral instead of voxel volume intersection. For a given point G, the weight w at G was 1, and decreased linearly along each spatial dimension to 0 at one grid resolution step. The total weight was the product of the three one-dimensional weights. The weight function was based on the absolute distance in three dimensions between G and P_i, therefore weights were assigned in all 6 directions around G (Fig. 1).

To account for increasing occlusion as lidar beams travel into the canopy, we assigned an additional weight w_o for each beam that quantifies its interceptance until reaching the grid point G, calculated from the CF distribution. This allowed assigning more importance to beams that have not been intercepted over beams that lost most of their energy on earlier hits.

Finally, the PAD for a grid point G was calculated as the weighted average (w.r.t. occlusion and spatial weights) of all beams as

$$\mathrm{PAD} = \frac{\sum_{i=1}^{N} \mathrm{PAD}_i w_{o,i}}{\sum_{i=1}^{N} w_{o,i}}, \text{and} \tag{4}$$

$$\mathrm{PAD_i} = \frac{\int_0^M \delta \mathrm{CF}(r) w(r,G) \mathrm{d}r}{0.5 \int_0^M \delta w(r,G) \mathrm{d}r}, \tag{5}$$

with the distance along the ray r used to calculate the distance between G and a point on the ray P_i. The denominator in Eq. (5) denotes the weighted path length around the point G. The factor 0.5 in the denominator comes from the average leaf area projection function (Wang et al., 2007). For simplicity, we assumed a spherical leaf angle distribution. Note that if the function w is defined so that it returns 1 when a ray at distance r is within half a grid resolution step and 0 elsewhere, one obtains a voxel grid sampling approach.

To obtain PAI of a plot, the PAD around each grid point was multiplied by the volume occupancy of that grid point, which is equivalent to integrating w over the region where $w > 0$. Since we used a linear weight for each spatial dimension, the volume occupancy was simply $V = (R/2)^3$. PAI for a square plot with 25 m side length is then calculated from PAD of each grid point g as

$$\mathrm{PAI} = \frac{\sum_{g=1}^{N} \mathrm{PAD}_g V}{25^2}. \tag{6}$$

We compared the PAI estimates from ALS at different grid spacings (0.2 m, 0.4 m, 0.6 m, 0.8 m and 1.0 m) to effective PAI obtained from hemispheric photography. Note that we use the terms grid spacing, scale and resolution synonymously in the presentation and discussion of our results. We also compared PAI estimates from our spatially weighted method to PAI obtained using a conventional voxel grid at grid spacings 0.4 m and 0.8 m. Finally, we compared the spatially weighted PAI to estimates from HPs (Hovi et al., 2022) and leaf area index from TLS (Schraik et al., 2023).

In the framework of this study, we created a software package to read PulseWaves files in Julia. This package, called "PulseWavesIO.jl", along with functionality for waveform data pre-processing and the code used to estimate PAD from airborne lidar waveform data was published on Github (Schraik, 2023).

3. Results and discussion

3.1. ALS and HP plant area index

We estimated PAI (hereafter ALS-PAI) from ALS data using a novel method that uses waveform information and a spatial weighting scheme rather than a discrete voxel grid. We found that our spatially weighted method results in ALS-PAI values that were mostly larger than HP-PAIe. The HP-PAIe is an effective PAI which does not account for clumping, whereas the ALS-PAI represents a PAI corrected for clumping at scales above the grid spacing. Therefore, the ratio of HP-PAIe to ALS-PAI corresponds to a clumping index. These clumping index (CI) values ranged from 1.19 at a grid spacing of 20 cm, 0.76 at 40 cm, and 0.86 to 0.94 at 60 to 100 cm (Fig. 2). Standard deviation of CI were consistent throughout the scales with 0.22 to 0.33. Forest canopies generally exhibit CIs below 1.0, indicating that leaves are more aggregated relative to a uniform random spatial distribution (Fang, 2021). In general, average CIs from satellite observations for evergreen needleleaf and broadleaf forests during peak growing season have been reported to be about 0.55 and 0.66–0.70, respectively (Jiao et al., 2018; Wei et al., 2019). In our two study sites, we previously reported TLS-based CIs ranging from 0.54 to 0.99 (Schraik et al., 2023).

Compared to these literature values, the CIs we found were slightly higher because they only consider clumping at scales higher than the grid spacing instead of clumping of canopy elements (leaves or shoots). We observed that changes in CI decrease with increasing scale between 60 cm to 1 m (CI = 0.86–0.94), whereas a relatively high degree of clumping seems to occur when changing from 60 cm to 40 cm (CI 0.86 to 0.76). Clumping indices at larger scales are the product of an

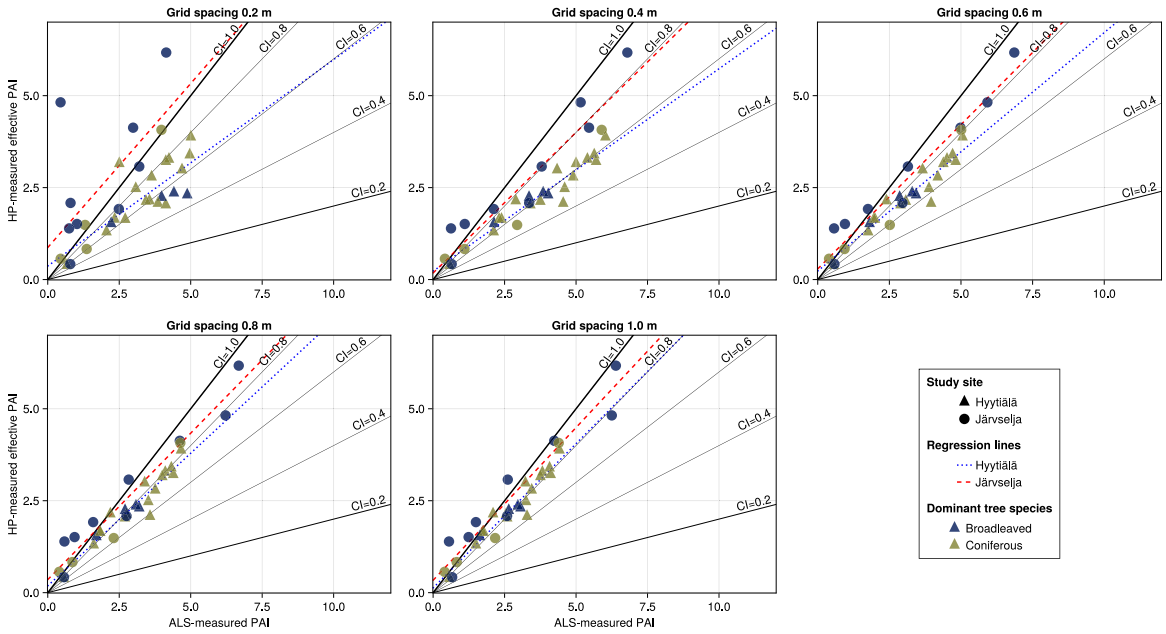


Fig. 2. ALS-estimated PAI compared to HP-estimated effective PAI. Several black lines indicate different levels of clumping index that correspond to the ratio of ALS-PAI to HP-PAIe. Two linear regression lines show the average clumping index for each study site.

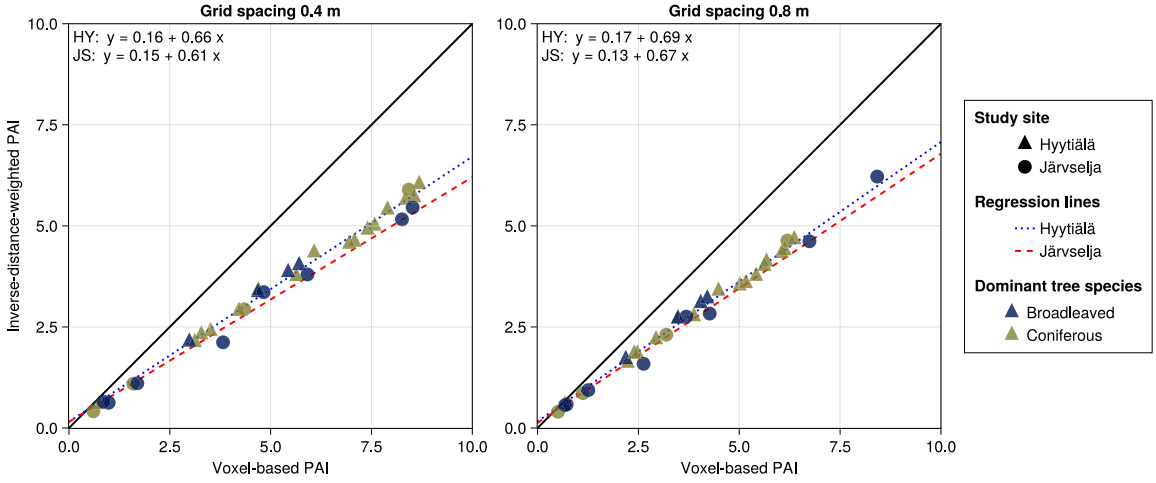


Fig. 3. Comparison of spatially weighted to voxel-based plant area index.

arbitrary number of finer scale CIs. For example, if one measured CI exclusively at shoot scale, crown scale, and stand scale, their product would result in an overall clumping index. When measuring CI with lidar point clouds, one can only measure a single CI down to the given spatial resolution. Since the CI at 40 cm is considerably smaller than at 60–100 cm, we can conclude that considerable clumping occurs at the scale between 40 and 60 cm (with a $CI_{40-60cm} \approx 0.88$ over a scale difference of 20 cm). Conversely, between 60 and 100 cm, only small changes in CI occur ($CI_{60-100cm} \approx 0.91$ over a scale difference of 40 cm). Thus, we observed that clumping at the scale of 40–60 cm is approximately twice as large as at 60–100 cm.

We estimated leaf area density assuming a spherical leaf angle distribution. This was done for simplicity, but it can lead to errors depending on how far the real leaf angle distribution deviates from the assumption. For Scots pine, a spherical leaf angle distribution appears likely valid (Stenberg et al., 1993). Norway spruce in the Czech Republic has been shown to have a planophile leaf angle distribution (Janoutová et al., 2019), but crown structure of boreal and hemiboreal Norway spruce can differ from Central Europe and it is thus unclear whether Northern European spruces would also be planophile. Broadleaved species are highly diverse, and the species contained in our study sites could be characterized as plagiophile with mean leaf angle between 40° and 60° (Pisek et al., 2013; Raabe et al., 2015). The plagiophile leaf angle distribution differs the least from the spherical. On the other hand, a planophile canopy would lead to a positive error in leaf area density when assuming a spherical leaf angle distribution, whereas an erectophile leaf angle distribution would lead to a negative error. Research in leaf angle measurements using terrestrial laser scanning is ongoing (e.g. Stovall et al., 2021) which should increase availability of reference data that can be used to improve our approach in future applications.

The grid resolution poses a trade-off between spatial detail and estimation bias. In coarse resolutions, grid cells are well sampled, but the level of detail is small and therefore clumping is insufficiently quantified. At fine resolutions, sampling by lidar beams becomes increasingly sparse, which typically leads to positive bias in PAI estimates (Pimont et al., 2018). To our knowledge, there is no method to optimize the spatial scale, and thus it remains a user's choice and requires a sensitivity analysis.

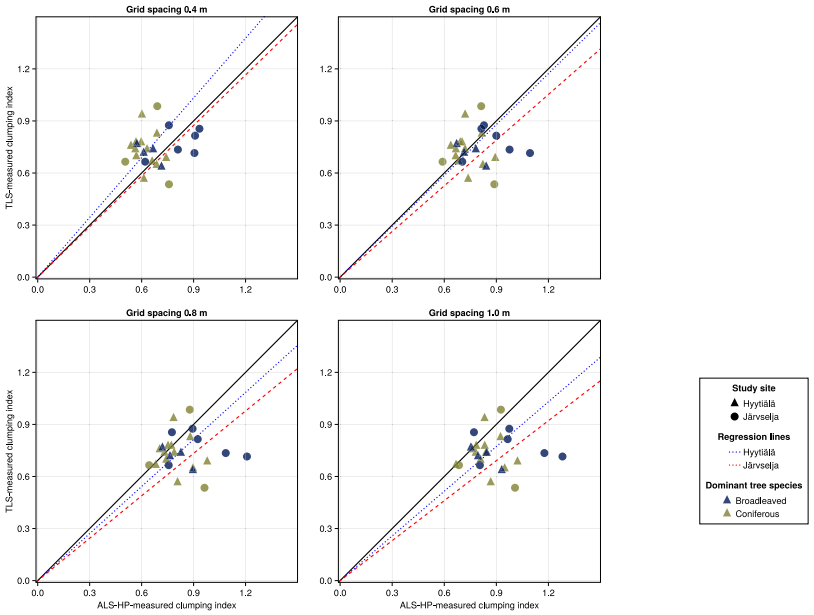


Fig. 4. Clumping index from ALS, based on the ratio of HP-PAIe to ALS-PAI, compared to TLS-measured clumping index. ALS clumping index quantifies clumping at scales above the grid spacing, TLS is at scales above 20 cm.

We consider the PAI estimates of the 20 cm grid to show clear signs of bias, highlighted by the fact that the fraction of grid points had more than 5 ALS beams traversing within their spatial domain was highly variable, ranging from 36% to 99% (average 90%). The values presented in Fig. 2 were calculated with occluded grid points (less than 5 rays per grid point) removed, which resulted in very low ALS-PAI values for some plots. The average CI at 20 cm was 1.19, which is unusually high, and typically one would expect that PAI increases with decreasing scale (Schraik et al., 2021a). The high CI can be attributed to the removal of occluded grid points. If these occluded grid points were included, ALS-PAI values at 20 cm grid spacing would have been higher than at larger grid spacings, as in occluded volumes, the interceptance of a single or a few rays can lead to extreme bias in PAI estimation.

In contrast, at 40 cm 99% (85%–100%) of grid points, and over 93% at larger scales, were sufficiently explored. While the point cloud data we used was very dense (over 40 points per square meter), occlusion, especially in the lower canopy, was too high for the 20 cm grid to consider the PAI estimates reliable. At grids with spacing of 40 cm or more, PAI estimates can be considered unaffected by occlusion bias. However, as the grid spacing increases to 1 m, PAI estimates decrease because canopy clumping is quantified to a lesser degree. Thus, we argue that the optimal scale for PAI estimation with ALS data may be around 40 to 60 cm, where spatial detail is high, with a low bias from occlusion.

3.2. Inverse distance weighted and voxel grid PAI

We compared PAIs estimated from the spatially weighted approach to standard voxel grids at grid spacings (and voxel sizes) of 40 and 80 cm. The voxel-PAIs were consistently higher than inverse-distance-weighted (IDW) PAIs (Fig. 3). The ratio of IDW-PAI to voxel-PAI was approximately 0.7, which was similar across the entire range of PAI values. This ratio varied little across different scales, with 0.68 at 40 cm, and 0.74 at 80 cm, both with standard deviations of around 0.05. These results indicate that using a spatial weighting scheme produces lower PAI values than voxel grids. Voxel-PAI often has a tendency to overestimate PAI (Soma et al., 2018; Schraik et al., 2021a). This is partly due discrepancy between the assumed infinitesimal footprint size in ray tracing (Pimont et al., 2018) and the real footprint size in the order of several millimeters to centimeters (Schraik et al., 2021b). Particularly problematic are cases in which leaf area is concentrated in near an edge of a voxel, as this leads to very high within-voxel clumping index and thus a negative error on PAI estimates (Schraik et al., 2021a). In this context, a lower PAI estimate may indicate that such errors are reduced with our method, although further investigation is needed to confirm this fact, particularly using simulated data.

Any volumetric estimate of PAI like ours or from the standard voxel grid is sensitive to occlusion. A common approach is to flag and interpolate occluded voxels to compensate for negative error in PAI estimates (Soma et al., 2020). In this study we did not interpolate because our goal was to compare our approach to voxel grids using the exact same data, which results in the same occlusion error for both methods. Interpolation would likely increase the ALS-PAI estimates, therefore clumping index values would be lower than reported above.

3.3. ALS and TLS clumping index

Previously, in Schraik et al. (2023), we estimated leaf area (with returns from woody elements removed) and clumping index from terrestrial laser scanning on the same plots. Here, we found that the CIs between ALS and TLS correspond most closely at the ALS scale of 60 cm (note that the ALS-CI is calculated as the ratio of HP-PAIe to ALS-PAI), and excluded the 20 cm scale from further analysis due to occlusion. With the TLS data, 20 cm voxels were used to partition the canopy space. We found the closest correspondence in ALS-CI and TLS-CI at a grid spacing of 60 cm, with smaller spacing producing lower CIs than TLS, and higher spacing resulting in higher CIs (Fig. 4). In a paired sample t-test, the 60 cm grid spacing was the only resolution with a p-value well above conventional thresholds at 0.26 (with a 95% confidence interval of the t-statistic of $[-0.1, 0.03]$). The 40 cm spacing had a p-value of 0.04, indicating that the difference between ALS-CI and TLS-CI is likely significantly different from zero (confidence interval indicates that TLS-CI is 0.003 to 0.12 larger than ALS-CI). At 80 and 100 cm grid spacing, p-values decreased further to 0.01 and < 0.01, with confidence intervals of $[-0.16, -0.02]$ and $[-0.21, -0.07]$, respectively. The increase of CI with increasing resolution is due to clumping being incorporated to within the elementary volume of a grid cell, within which it cannot be measured with ray tracing but is rather assumed to have a CI of 1 (Schraik et al., 2021a). In this data, we found the closest correspondence between ALS-CI and TLS-CI in the resolution

of 60 cm, which appeared in our data to form a compromise between avoiding both occlusion and fine enough scale to quantify clumping.

The bulk of the plots are in close agreement with high precision, with only six plots (four pine and two broadleaved plots) deviating from this general trend. The ALS-PAI in these six plots ranged from 0.9 to 3.2 at 60 cm grid spacing, which is 95% (83%–106%) of the respective TLS-LAI. Considering that the TLS-LAI was close to the HP-PAIe (Schraik et al., 2023), these six plots had the lowest ALS-PAI estimates, as the remaining plots at 60 cm grid spacing were 194% (94%–440%) of the TLS-LAI. Despite these large differences in PAI/LAI, the close correspondence of the CI between the two data sources shows that the quantification of CI with lidar can be very precise despite potential discrepancies in absolute leaf or plant area estimates.

In contrast to the TLS measurement, we did not separate leaf and wood returns from the ALS data, thus we are comparing CIs that came from a plant area index from this study to a leaf area index from Schraik et al. (2023). Quantifying the potential difference between a PAI-CI and an LAI-CI is not trivial. The conversion from PAI to LAI requires information on the wood area index (Yan et al., 2019), however, due to mutual occlusion of leaf and wood elements, such a simple conversion factor is not available for PAI-CI and LAI-CI.

Besides these conceptual differences in terms of plant architecture, there are also sensing technique differences that must be considered when comparing LAI-CI from TLS and PAI-CI from ALS data. From the aerial perspective, the upper part of tree crowns is explored most, and penetration to the lower canopy is relatively low. Thus, the relative contribution of woody elements to the total PAI may be less than when measured in situ, which would reduce the gap between the TLS and ALS measured CIs. Finally, the two perspectives differ in their relative exploration of the canopy volume, with TLS exploring the lower canopy more, while ALS obtains better sampling of the upper canopy (Schneider et al., 2019).

3.4. Conclusions

We applied a spatial weighting approach to estimating plant area using ray tracing and airborne laser scanning data. We showed that the commonly used voxel-based approach may not be ideal for lidar data with large footprints such as airborne laser scanning data. By sampling plant area around grid points with an inverse distance weighting scheme that uses waveform information, we showed that PAI estimates match more closely to measurements from hemispheric photography and terrestrial laser scanning. The derived clumping indices obtained from this weighting scheme also correspond more closely to values typically found in forests than those obtained from a voxel grid.

Through this new approach, we offer a new perspective on partitioning the canopy volume for spatially explicit measurements of plant or leaf area with lidar that could be used with state-of-the-art methods such as the path volume for airborne lidar data (Yin et al., 2022), and it could also be adapted to TLS data for potentially improved plant area estimation accuracy. While we used full waveform data in our study, our method can be applied to discrete return lidar data as well, given that the waveform can be reconstructed. That is, the discrete point cloud contains information on the underlying waveform, for example returns' standard deviation, and that the system pulse (i.e., the outgoing waveform convolved with the receiver's response) is known.

Three-dimensional measurements of canopy plant area, and in particular the clumping index, are gaining increasing importance in modeling forest reflectance, micro- and macroclimate, and for other ecological applications. These applications require easy-to-use and accurate measurement methods to study the influence of canopy structure on its environment. Our results show that partitioning canopies into voxels can result in overestimation of plant area index with airborne lidar waveform data, whereas a spatial weighting approach can produce more realistic plant area predictions. With our novel spatial sampling scheme, airborne lidar could become the best tool for generating three-dimensional canopy structure information over large areas to inform forest radiation regime, climate and ecological models.

CRediT authorship contribution statement

Daniel Schraik: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Aarne Hovi:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Miina Rautiainen:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We thank CzechGlobe's Department of Airborne Activities for their support with data preprocessing. This study received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 771049). The article reflects only the authors' view and the Agency is not responsible for any use that may be made of the information it contains. We acknowledge the resources provided by the Finnish IT Center for Science (CSC) for our computations.

References

Béland, M., Widlowski, J.-L., Fournier, R.A., 2014. A model for deriving voxel-level tree leaf area density estimates from ground-based LiDAR. Environ. Model. Softw. 51, 184–189. http://dx.doi.org/10.1016/j.envsoft.2013.09.034.

Calders, K., Origo, N., Burt, A., Disney, M., Nightingale, J., Raumonen, P., Åkerblom, M., Malhi, Y., Lewis, P., 2018. Realistic forest stand reconstruction from terrestrial LiDAR for radiative transfer modelling. Remote Sens. 10 (6), 933. http://dx.doi.org/10.3390/rs10060933.

Chen, J., Cihlar, J., 1995. Quantifying the effect of canopy architecture on optical measurements of leaf area index using two gap size analysis methods. IEEE Trans. Geosci. Remote Sens. 33 (3), 777–787. http://dx.doi.org/10.1109/36.387593.

Fang, H., 2021. Canopy clumping index (CI): A review of methods, characteristics, and applications. Agricult. Forest. Meterol. 303, 108374. http://dx.doi.org/10.1016/j.agrformet.2021.108374.

Hovi, A., Schraik, D., Hanuš, J., Homolová, L., Juola, J., Lang, M., Lukeš, P., Pisek, J., Rautiainen, M., 2022. Assessment of a photon recollision probability based forest reflectance model in European boreal and temperate forests. Remote Sens. Environ. 269, 112804. http://dx.doi.org/10.1016/j.rse.2021.112804.

Janoutová, R., Homolová, L., Malenovský, Z., Hanuš, J., Lauret, N., Gastellu-Etchegorry, J.-P., 2019. Influence of 3D spruce tree representation on accuracy of airborne and satellite forest reflectance simulated in DART. Forests 10 (3), 292. http://dx.doi.org/10.3390/f10030292.

Jiao, Z., Dong, Y., Schaaf, C.B., Chen, J.M., Román, M., Wang, Z., Zhang, H., Ding, A., Erb, A., Hill, M.J., Zhang, X., Strahler, A., 2018. An algorithm for the retrieval of the Clumping Index (CI) from the MODIS BRDF product using an adjusted version of the kernel-driven BRDF model. Remote Sens. Environ. 209, 594–611. http://dx.doi.org/10.1016/j.rse.2018.02.041.

Kamoske, A.G., Dahlin, K.M., Stark, S.C., Serbin, S.P., 2019. Leaf area density from airborne LiDAR: Comparing sensors and resolutions in a temperate broadleaf forest ecosystem. Forest Ecol. Manag. 433, 364–375. http://dx.doi.org/10.1016/j.foreco.2018.11.017.

Pimont, F., Allard, D., Soma, M., Dupuy, J.-L., 2018. Estimators and confidence intervals for plant area density at voxel scale with T-LiDAR. Remote Sens. Environ. 215, 343–370. http://dx.doi.org/10.1016/j.rse.2018.06.024.

Pisek, J., Sonnentag, O., Richardson, A.D., Mõttus, M., 2013. Is the spherical leaf inclination angle distribution a valid assumption for temperate and boreal broadleaf tree species? Agricult. Forest. Meterol. 169, 186–194. http://dx.doi.org/10.1016/j.agrformet.2012.10.011.

Raabe, K., Pisek, J., Sonnentag, O., Annuk, K., 2015. Variations of leaf inclination angle distribution with height over the growing season and light exposure for eight broadleaf tree species. Agricult. Forest. Meterol. 214–215, 2–11. http://dx.doi.org/10.1016/j.agrformet.2015.07.008.

Schneider, F.D., Kükenbrink, D., Schaepman, M.E., Schimel, D.S., Morsdorf, F., 2019. Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR. Agricult. Forest. Meterol. 268, 249–257. http://dx.doi.org/10.1016/j.agrformet.2019.01.033.

Schraik, D., 2023. PulseWavesIO.jl, a Julia package for reading and processing the PulseWaves LiDAR data forma. http://dx.doi.org/10.5281/zenodo.10418231, Zenodo.

Schraik, D., Hovi, A., Rautiainen, M., 2021a. Crown level clumping in Norway spruce from terrestrial laser scanning measurements. Agricult. Forest. Meterol. 296, 108238. http://dx.doi.org/10.1016/j.agrformet.2020.108238.

Schraik, D., Hovi, A., Rautiainen, M., 2021b. Estimating cover fraction from TLS return intensity in coniferous and broadleaved tree shoots. Silva Fennica 55 (4), http://dx.doi.org/10.14214/sf.10533.

Schraik, D., Wang, D., Hovi, A., Rautiainen, M., 2023. Quantifying stand-level clumping of boreal, hemiboreal and temperate European forest stands using terrestrial laser scanning. Agricult. Forest. Meterol. 339, 109564. http://dx.doi.org/10.1016/j.agrformet.2023.109564.

Soma, M., Pimont, F., Allard, D., Fournier, R., Dupuy, J.-L., 2020. Mitigating occlusion effects in Leaf Area Density estimates from terrestrial LiDAR through a specific kriging method. Remote Sens. Environ. 245, 111836. http://dx.doi.org/10.1016/j.rse.2020.111836.

Soma, M., Pimont, F., Durrieu, S., Dupuy, J.-L., 2018. Enhanced measurements of Leaf Area Density with T-LiDAR: Evaluating and calibrating the effects of vegetation heterogeneity and scanner properties. Remote Sens. 10 (10), 1580. http://dx.doi.org/10.3390/rs10101580.

Stenberg, P., Smolander, H., Kellomäki, S., 1993. Description of crown structure for lightinterception models: Angular and spatial distribution of shoots in young Scots pine. In: Linder, S., Kellomäki, S. (Eds.), Management of Structure and Productivity of Boreal and Subalpine Forests. In: Studia Forestalia Suecica, vol. 191, Studia Forestalia Suecica, p. 94pp.

Stovall, A.E.L., Masters, B., Fatoyinbo, L., Yang, X., 2021. TLSL<scp>e</scp>AF: Automatic leaf angle estimates from single-scan terrestrial laser scanning. New Phytol. 232 (4), 1876–1892. http://dx.doi.org/10.1111/nph.17548.

Wang, W.-M., Li, Z.-L., Su, H.-B., 2007. Comparison of leaf angle distribution functions: Effects on extinction coefficient and fraction of sunlit foliage. Agricult. Forest. Meterol. 143 (1), 106–122. http://dx.doi.org/10.1016/j.agrformet.2006.12.003.

Wei, S., Fang, H., Schaaf, C.B., He, L., Chen, J.M., 2019. Global 500 m clumping index product derived from MODIS BRDF data (2001–2017). Remote Sens. Environ. 232, 111296. http://dx.doi.org/10.1016/j.rse.2019.111296.

Yan, G., Hu, R., Luo, J., Weiss, M., Jiang, H., Mu, X., Xie, D., Zhang, W., 2019. Review of indirect optical measurements of leaf area index: Recent advances, challenges, and perspectives. Agricult. Forest. Meterol. 265, 390–411. http://dx.doi.org/10.1016/j.agrformet.2018.11.033.

Yin, T., Cook, B.D., Morton, D.C., 2022. Three-dimensional estimation of deciduous forest canopy structure and leaf area using multi-directional, leaf-on and leaf-off airborne LiDAR data. Agricult. Forest. Meterol. 314, 108781. http://dx.doi.org/10.1016/j.agrformet.2021.108781.