Schraik, Daniel; Wang, Di; Hovi, Aarne; Rautiainen, Miina

Quantifying stand-level clumping of boreal, hemiboreal and temperate European forest stands using terrestrial laser scanning

Published in:
Agricultural and Forest Meteorology

DOI:
10.1016/j.agrformet.2023.109564

Published: 15/08/2023

Document Version
Publisher's PDF, also known as Version of record

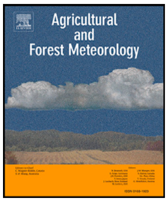
Please cite the original version:
Schraik, D., Wang, D., Hovi, A., & Rautiainen, M. (2023). Quantifying stand-level clumping of boreal, hemiboreal and temperate European forest stands using terrestrial laser scanning. *Agricultural and Forest Meteorology*, *339*, Article 109564. https://doi.org/10.1016/j.agrformet.2023.109564

Agricultural and Forest Meteorology 339 (2023) 109564

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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



Quantifying stand-level clumping of boreal, hemiboreal and temperate European forest stands using terrestrial laser scanning

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

[a] *Aalto University, School of Engineering, Department of Built Environment, PO Box 14100, 00076 Aalto, Finland*
[b] *Department of Remote Sensing Science and Technology, School of Electronic Engineering, Xidian University, Xi'an 710077, China*

ARTICLE INFO

Keywords:
Clumping
Silhouette area
Leaf area index
Terrestrial laser scanning
Spectral invariants

ABSTRACT

Clumping is critical for quantifying the radiation regime of forest canopies, but challenging to measure. We developed a method to measure clumping index (CI) of forest stands using voxel-based estimates of leaf area density from terrestrial lidar data. Our method uses the principle of the silhouette to total area ratio (STAR), a commonly used shoot clumping correction approach. We adapted the concept to forest stands, and derived that STAR at canopy scale ($\mathrm{STAR_f}$) is no longer simply a clumping index, but a summary variable for forest structure in general. CI can be calculated from $\mathrm{STAR_f}$ when leaf area index is known.

We measured CI and $\mathrm{STAR_f}$ of 38 forest stands in Finland, Estonia, and Czechia to study the natural range of these variables, their relationships to other forest variables, and to Landsat 8 OLI surface reflectance. CI did not include clumping below voxel scale (20 cm), and ranged from 0.6 to 0.9, with the lowest values (i.e., the most clumped canopies) in conifer forests and temperate oak forests, and the highest CI values (i.e., the most random canopies) in boreal broadleaved forests. CI was closely correlated with surface reflectance in conifer forests, which may be explained by contradicting influence of clumping that decreases canopy reflectance, but increases visibility of the forest floor.

From the viewpoint of forest reflectance modeling, STAR is a useful variable due to its close relationship with the photon recollision probability, i.e., the probability that a photon will interact with a canopy element after being scattered from another canopy element. The photon recollision probability is used to model the influence of forest structure on reflectance. Our method provides a physically-based means of measuring $\mathrm{STAR_f}$, and thus the photon recollision probability, hence contributing to the development of new methods for interpreting forest canopy structure from optical remote sensing data.

1. Introduction

Clumping and leaf area index are the most important canopy structural parameters that determine light interception, primary production, albedo, and photosynthesis rates of forests (Parker, 2020). Also the optical remote sensing signal of a forest is strongly influenced by clumping (Chen et al., 2012). For example, clumping has been shown to explain the difference in surface reflectance between coniferous and broadleaved forests (Rautiainen and Stenberg, 2005). The recognition of the role of clumping in remote sensing and vegetation ecology has led to an interest in mapping clumping over large areas (e.g. He et al., 2012; Jiao et al., 2018; Wei et al., 2019) and across time (He et al., 2016).

Clumping quantifies the dispersion patterns of leaves in a canopy. If leaf positions are random and independent of one another, the canopy is said to exhibit no clumping, whereas if leaves in a canopy tend to be close to one another, the canopy is said to exhibit clumping. An obvious example of clumping is the clumping of needles into shoots in coniferous forests. Clumping is typically quantified by the clumping index (CI), which modifies the leaf area index (LAI) in Beer's law (Nilson, 1999) into an effective LAI, that is, the equivalent LAI of a canopy with randomly distributed elements (Chen and Black, 1992). The effective LAI is typically smaller than the true LAI, as forest canopies are generally clumped (Fang, 2021).

Clumping can be defined at several different levels (Silva et al., 2008). A common example is the above mentioned shoot clumping, which is based on the silhouette to total area ratio (STAR) of conifer shoots (e.g. Therezien et al., 2007; Oker-Blom and Smolander, 1988; Smolander and Stenberg, 2003). The shoot clumping index Ω_{sh} is calculated from shoot STAR as $\Omega_{sh} = 4\mathrm{STAR_{sh}}$. In other words, the shoot CI is the ratio of spherically averaged shoot area to spherically

---

* Corresponding authors.
E-mail addresses: daniel.schraik@aalto.fi (D. Schraik), diw@xidian.edu.cn (D. Wang).

https://doi.org/10.1016/j.agrformet.2023.109564
Received 29 November 2022; Received in revised form 21 April 2023; Accepted 8 June 2023

averaged needle area (Stenberg et al., 1993). The factor 4 is used with $\mathrm{STAR_{sh}}$ to transform the total needle area into spherically averaged area. When accounting for shoot clumping in Beer's law, the implicit assumption becomes that the shoots, not the needles, are randomly distributed in the canopy. One can combine the shoot CI with a canopy element CI to quantify clumping at all scales.

In standard gap-fraction-based measurements of leaf area index (LAI), using hemispherical photography, LAI-2200 Plant Canopy Analyzer, TRAC (Tracing Radiation and Architecture of Canopies) or similar instruments, clumping is commonly corrected on two levels; the shoot level and at the stand level (for a review, see Fang, 2021). Stand level clumping is quantified based on logarithmic gap fraction averaging (Lang and Xiang, 1986), gap size distribution (Chen and Cihlar, 1995), a combination of the two (Leblanc et al., 2005), or using a fractal approach (Li and Mu, 2021; Lai et al., 2022).

The increased availability of terrestrial laser scanning (TLS) data and associated processing methods has enabled a new way of quantifying canopy structure, moving away from gap fraction measurements towards explicit three-dimensional models of forest canopies. For example, TLS has been used for estimating leaf area density (e.g. Béland et al., 2014; Pimont et al., 2018), but also clumping at branch scale (Béland and Baldocchi, 2020), and crown scale (Schraik et al., 2021a). In Schraik et al. (2021a) we used the concept of STAR to measure clumping of tree crowns with TLS data.

In this paper, we continue to develop the STAR measurement approach, moving away from individual tree crowns to quantify clumping of entire forest canopies. In adapting STAR to stand level, we need to consider that, unlike a shoot or an individual tree, a forest canopy has no concrete horizontal boundaries, and is usually described as if it were horizontally infinite. Thus, it is convenient to define forest stand STAR through canopy interceptance and leaf area index, as will be explained next. Smolander and Stenberg (2003) derived the relationship between $\mathrm{STAR_{sh}}$ and photon recollision probability for shoots as $p_{sh} = 1-4\mathrm{STAR_{sh}}$. Stenberg (2007) showed that at stand level, the photon recollision probability can be calculated from the diffuse interceptance i_D and the LAI L as $p = 1 - \frac{i_D}{L}$. Denoting the spherically averaged silhouette area of a stand as $\mathrm{SA} = \int_0^{\pi/2}[1-t(\theta)]A_C sin\theta cos\theta d\theta$, and the diffuse interceptance as $i_D = 1-2\int_0^{\pi/2} t(\theta) sin\theta cos\theta d\theta$ (Stenberg, 2007), we get the relationship between the stand silhouette to total area ratio $\mathrm{STAR_f}$, i_D, and LAI at stand level as

$$\mathrm{STAR_f} = \frac{\mathrm{SA}}{A_L} = \frac{\int_0^{\pi/2}[1-t(\theta)]A_C sin\theta cos\theta d\theta}{2A_C L} = \frac{1-2\int_0^{\pi/2} t(\theta) sin\theta cos\theta d\theta}{4L} = \frac{i_D}{4L}, \tag{1}$$

with the spherically averaged silhouette area SA, total (i.e. two-sided) leaf area A_L, and the stand area A_C. Thus, we see that $\frac{i_D}{L}$ is the stand level equivalent to $4\mathrm{STAR_f}$. To avoid confusion, hereafter we write $\mathrm{STAR_{sh}}$ for instances concerning specifically shoot STAR, $\mathrm{STAR_f}$ specifically for stand or forest STAR, and STAR without a scale specification when we mean the concept of silhouette to total area ratio in general.

The advantage of using STAR to quantify clumping lies in its direct relationship to the photon recollision probability p (Smolander and Stenberg, 2003). The photon recollision probability originates from p-theory, which states that only the structural parameter p is needed, together with canopy element spectra, to explain radiative transfer in a vegetation canopy (Knyazikhin et al., 1998). Smolander and Stenberg (2003) interpreted p as the photon recollision probability, and Rautiainen and Stenberg (2005) created the PARAS forest reflectance model around p. PARAS models have since been used in different formulations, and have been applied in several remote sensing studies at different spatial scales (e.g. Yáñez Rausell et al., 2015; Majasalmi et al., 2014; Hadi et al., 2017; Schraik et al., 2019; Hovi et al., 2020, 2022). Most forest reflectance models require clumping as an input, since LAI alone is insufficient to model the radiative transfer process in forest vegetation. Since PARAS relies on the photon recollision probability p, it is possible to directly use STAR to parameterize canopy scattering properties. For accurate estimation of p, information on both shoot STAR and a higher-level CI are needed. The method recently proposed in Schraik et al. (2021a) enables direct measurement of STAR from terrestrial laser scanning (TLS) data at levels above shoot scale. TLS point clouds provide explicit three-dimensional spatial information. This level of detail allows, for the first time, direct estimation of STAR at stand level ($\mathrm{STAR_f}$) which has not been possible with passive optical methods, as they lack three-dimensional explicit information, thus requiring to make assumptions about the missing dimensions. $\mathrm{STAR_f}$ is closely related to the photon recollision probability, and is thus of particular interest for quantifying forest canopy structure in forest radiation regime modeling. $\mathrm{STAR_f}$ summarizes the three canopy structure variables that are critical drivers of the forest radiation regime; leaf area index, clumping index, and leaf angle distribution. Our method to quantify $\mathrm{STAR_f}$ relies on a spatially explicit quantification of leaf area density. Thus, the method can be used to simultaneously estimate clumping index, since it explicitly quantifies both the LAI and the directional interceptance. The latter can be used to calculate the effective LAI, and the CI is simply the ratio of effective to true LAI. The analysis of the CI is needed by a wider research community beyond photon recollision probability modeling.

In this study, we quantify $\mathrm{STAR_f}$ and clumping index from TLS point clouds across different European forests ranging from boreal to temperate biomes. We examine the relationship of $\mathrm{STAR_f}$ with commonly measured forest variables, and of both $\mathrm{STAR_f}$ and clumping index with Landsat 8 OLI surface reflectance data. The relationship between $\mathrm{STAR_f}$ and forest variables may provide insights for estimating $\mathrm{STAR_f}$ indirectly, since current methods for measuring $\mathrm{STAR_f}$ are laborious and difficult to scale over large areas. We further evaluate the reliability of $\mathrm{STAR_f}$ measurements by comparing its leaf area component to indirect measurements of leaf area.

2. Materials and methods

2.1. Study sites and stand description

We measured a total of about 2100 trees in 38 plots in four study sites. The sites were located in Lanžhot (48°41'N, 16°57'E, 9 plots) and Bílý Kříž (49°30'N, 18°32'E, 6 plots) in the Czech Republic, Järvselja in Estonia (58°17'N, 27°19'E, 9 plots), and Hyytiälä in Finland (61°51'N, 24°18'E, 16 plots) (Fig. 1). The plots in Lanžhot (Fig. 1A) are diverse temperate broadleaved floodplain forests, with several oak species, ash, aspen, and hornbeam as stand-dominating species, field maple, linden as occasionally abundant understory species, and occasional individuals of alder, elm, and wild pear. Bílý Kříž (Fig. 1B) is a temperate mountain forest dominated by Norway spruce, with a small portion of European beech. Hyytiälä (Fig. 1C) is a boreal site dominated by Scots pine, Norway spruce, and birch. Järvselja (Fig. 1D) is a hemi-boreal mixed broadleaved and conifer site. The main species are alder, Norway spruce, Scots pine, aspen, and birch. The terrain in all sites was mostly flat, except for Bílý Kříž, which was moderately mountainous with slopes in our plots reaching up to 10°. The measurements in Hyytiälä were made in June and July 2019, in Lanžhot and Bílý Kříž in August and September 2019, and in Järvselja during June and July 2020. At the time of the field campaigns, all sites were in peak growing season. In addition to the terrestrial laser scanning and hemispherical photographic measurements described below (Sections 2.2 and 2.5), we measured standard forest inventory variables in all plots. These included diameter at breast height (DBH, 1.3 m above ground) for all trees in a plot that exceeded a DBH limit (5 cm in young forest, 8 cm in mature forest), and tree height for the two median trees among the 10% thickest trees in a plot. In a single young pine plot in Hyytiälä, the DBH limit was reduced to 2.5 cm, as most trees had less than 5 cm DBH. Basal area and stem number per hectare were calculated from the above measurements. Table 1 contains a summary of forest inventory variables for each site and species. Further description of the sites and photographs of selected stands are available in Hovi et al. (2022).

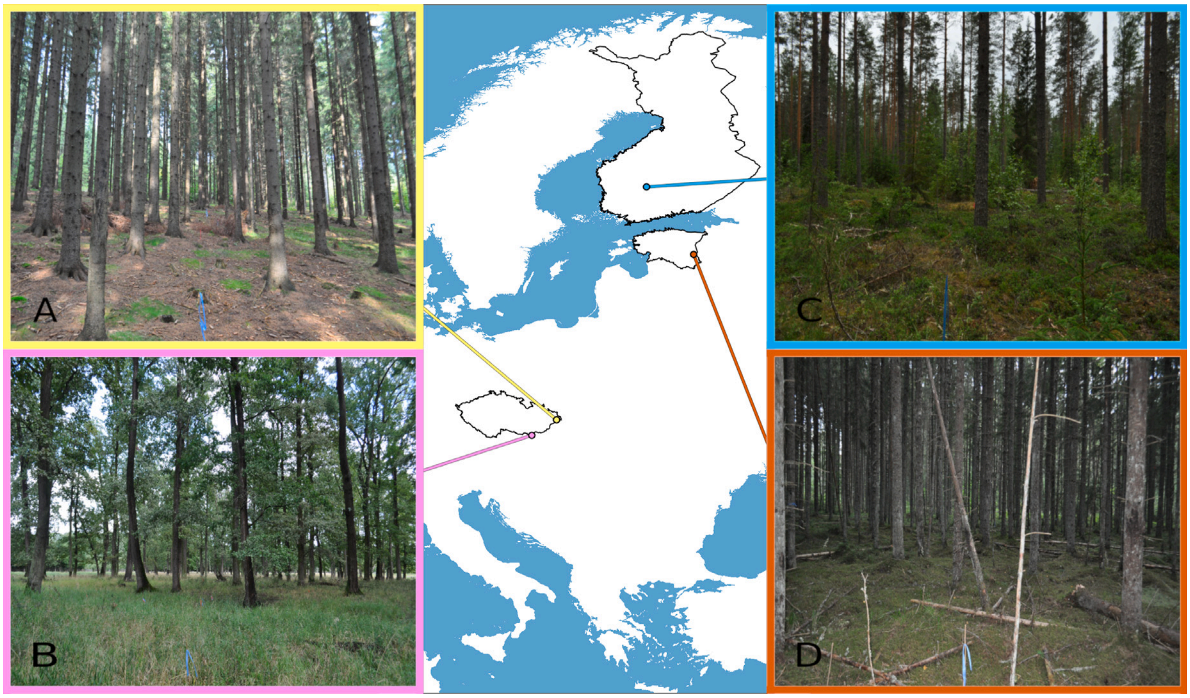


Fig. 1. Location of the study sites. The pink (B) and yellow (A) sites denote Bílý Kříž and Lanžhot in the Czech Republic, blue (C) denotes Hyytiälä in Finland, and red (D) denotes Järvselja in Estonia.

Table 1
Summary of forest variables per site and dominant species. The values express the mean of the forest stand variable, and the full range is given in parentheses if applicable.

Site	Species	Plots	Effective LAI [m^2m^{-2}]		Mean DBH [cm]		Stem number [ha^{-1}]		Basal area [m^2ha^{-1}]		Tree height [m]	
Hyytiälä	*Pinus sylvestris* L. (Scots pine)											
		6	1.77	(0.45, 3.04)	18.4	(6.8, 30.3)	1266	(272, 2144)	21.5	(4.6, 33.0)	17.4	(6.8, 25.0)
	Picea abies (L.) H. Karst. (Norway spruce)											
		6	2.88	(2.15, 3.4)	31.1	(21.9, 44.8)	498	(368, 672)	31.6	(22.0, 46.0)	27.4	(21.9, 34.3)
	Betula sp. (birch; *B. pendula* Roth, *B. pubescens* Ehrh.)											
		4	2.08	(1.46, 2.33)	15.9	(12.0, 20.0)	1256	(416, 1904)	17.7	(11.5, 22.9)	20.1	(17.2, 23.2)
Järvselja	*Pinus sylvestris* L. (Scots pine)											
		2	1.13	(0.8, 1.46)	33.3	(26.8, 39.9)	432	(112, 752)	23.9	(13.5, 34.3)	28.4	(26.8, 30.0)
	Alnus glutinosa (L.) Gaertn. (European black alder)											
		2	4.68	(3.05, 6.32)	13.2	(13.1, 13.4)	1680	(1184, 2176)	19.4	(12.3, 26.6)	17.1	(16.5, 17.8)
	Populus tremula L. (European aspen)											
		2	4.44	(3.93, 4.94)	24.8	(12.4, 37.1)	1584	(944, 2224)	35.9	(20.8, 51.1)	28.3	(17.9, 38.6)
	Picea abies (L.) H. Karst. (Norway spruce)											
		1	4.05		21.6		1328		37.7		28.3	
	Betula sp. (birch; *B. pendula* Roth, *B. pubescens* Ehrh.)											
		2	1.99	(1.84, 2.13)	17.1	(11.6, 22.6)	928	(576, 1280)	15.3	(9.8, 20.9)	23.1	(15.8, 30.4)
Bílý Kříž	*Picea abies* (L.) H. Karst. (Norway spruce)											
		4	3.35	(2.76, 4.11)	37.9	(21.5, 50.7)	576	(416, 768)	50.2	(24.0, 65.9)	32.8	(20.6, 42.7)
Lanžhot	*Quercus* sp. (oak; *Q. robur* L., *Q. petraea* (Matt.) Liebl.)											
		6	3.45	(2.11, 5.04)	38.6	(14.8, 72.4)	592	(96, 1552)	27.2	(14.5, 42.2)	28.4	(18.5, 38.8)
	Fraxinus sp. (ash; *F. excelsior* L., *Fraxinus angustifolia* Vahl)											
		1	4.78		61.6		544		60.1		40.0	
	Populus tremula L. (European aspen)											
		1	3.63		64.6		336		40.2		33.3	
	Carpinus betulus L. (European hornbeam)											
		1	5.33		30.4		560		29.4		32.4	

2.2. *Terrestrial laser scanning* (TLS) *measurements*

In each plot, TLS scans were done at 16 grid points, forming a square grid of 30 by 30 m (4 by 4 grid points with 10 m spacing, see Fig. 2). We used a Leica P40 ScanStation, which has a vertical field of view of 305 degrees (centered at the zenith). The laser beam has a wavelength of 1550 nm, a diameter of 6 mm at the instrument, and a beam divergence of 0.23 mrad. We used a scan resolution equal to the beam divergence (0.23 mrad, or 2.3 mm at 10 m distance). The scans were taken only during calm wind (less than 4 m/s) and dry conditions. The scans were made at heights between 1.4 and 1.8 m above ground. Most scans in Hyytiälä were carried out during the night until late morning because during the daytime the wind conditions, despite occasionally having low average wind speeds, frequently featured strong gusts, to which our scans were sensitive due to the scan time of about 6 min per scan. We used a total of 25 polystyrene sphere targets, mounted on 1.5 m tall sticks, that were spread out regularly throughout the plot in a 5 by 5 grid with 10 m spacing so that the spheres were placed in between the TLS measurement points, with additional spheres around the edge of the TLS grid. The sphere positions were adapted in case that grid points fell into a tree crown or other potentially occluding object.

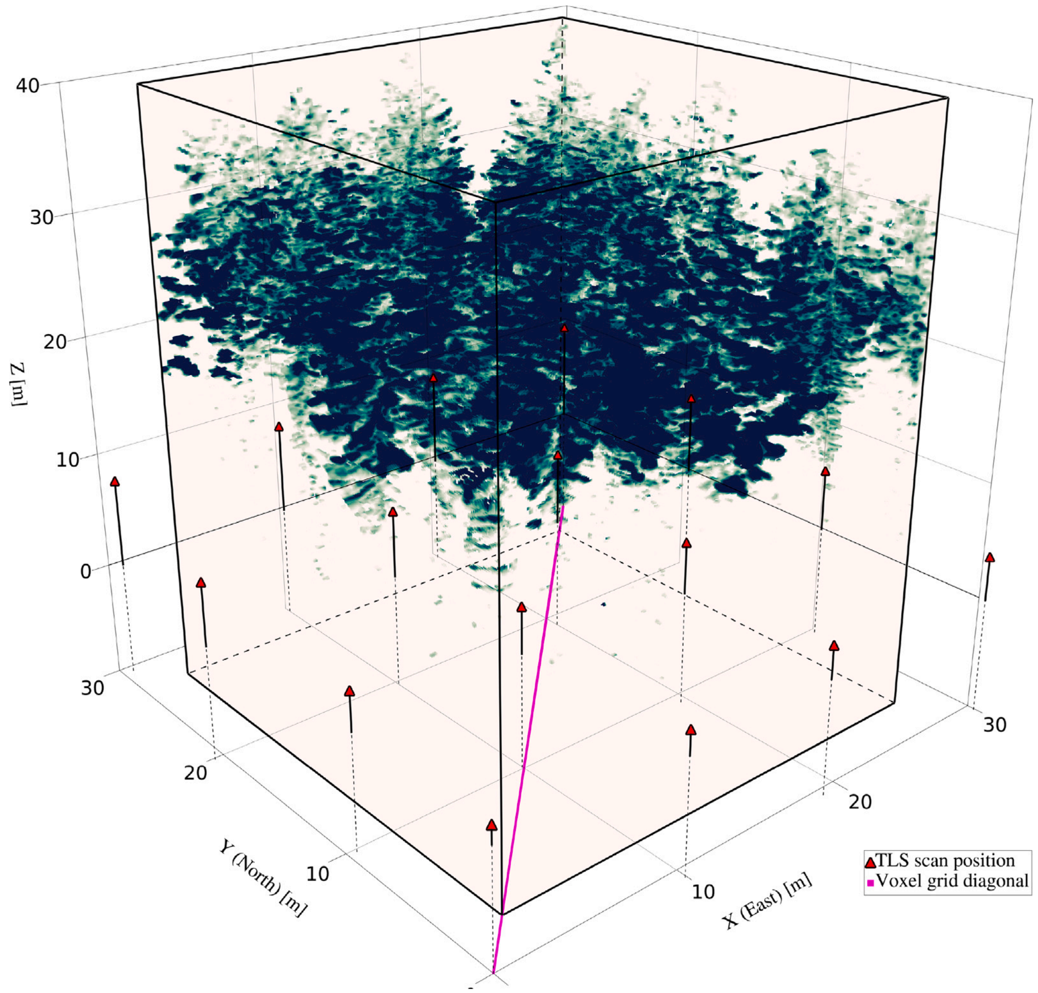


Fig. 2. Visualization of the TLS measurement setup in a plot. The 16 TLS scanning positions were spaced by 10 m. The magenta diagonal line, which runs between the SW and NE corner TLS positions, was used to center the voxel grid's bounding box (in black) in the plot.

2.3. TLS preprocessing

2.3.1. Co-registration in Cyclone

The co-registration of each plot's 16 scans was carried out in Leica Cyclone (version 9.4.0). The points belonging to the sphere targets were selected manually from the point cloud, and we used Cyclone's sphere fitting algorithm to calculate the position of the sphere targets. We ensured the error for individual targets was below 5 cm, and the translation and rotation tolerance for co-registration of TLS scans was 0.9 cm and 1°. Only one plot (in Järvselja) had three scans for which registration accuracy could not be calculated, because despite each of these scans having four to five visible targets, the number of targets shared with other stations was less than three, and hence it was not possible to extract accuracy metrics from Cyclone. Visual inspection of small objects (e.g. branches, small trees) indicated that the co-registration for those scans had errors in the order of millimeters, hence we did not discard those scans from our analysis. The co-registered scans were exported at full resolution, with the individual scans intact (using the e57 file format). In addition, we cropped the point clouds approximately 5 m outside their plot boundaries and downsampled them to an average point spacing of 2 cm using a proprietary method in Cyclone. These cropped and downsampled point clouds were used in tree segmentation and leaf–wood separation (Section 2.3.2). The lower resolution point clouds were needed to reduce computational costs of the leaf–wood separation to a manageable level, and the classification was upscaled to full resolution as described in Section 2.3.2. For estimating silhouette to total area ratio (Section 2.4), the full resolution and uncropped point cloud data were used.

2.3.2. Tree segmentation and leaf–wood separation

Our method for estimating leaf area assumes that vegetation elements are randomly distributed within voxels. While this assumption can work reasonably well for leaves, it leads to bias for woody elements (Pimont et al., 2019). Therefore, we separated leaf from wood points and estimated leaf area only. The leaf–wood separation algorithm, while capable of processing whole plots, required tree segmentation because the plot point clouds were too large to process at once.

The point cloud was first normalized by a Digital Terrain Model (DTM) to estimate the point heights above ground. The DTM was created by a hierarchical approach based on robust moving plane interpolation in a coarse to fine fashion (Kraus and Pfeifer, 1998; Wang et al., 2016). Individual tree segmentation was achieved by using a newly developed graph pathing algorithm that operates on a point cloud graph (Wang et al., 2021). This method first generated a compact point cloud graph by merging a K-nearest neighbors (KNN) graph with the Delaunay triangulation. Each point (i.e., a node in the graph) is pathed to its lowest neighboring point first to locate root nodes that represent potential tree lower stem positions. Such root nodes were further pruned and merged so that the stems for individual trees were identified. The pathing and root information in combination resembled a clustering mechanism that segmented points into groups that represented individual trees. The readers are referred to Wang et al. (2021) for a full description of the graph pathing method.

In this study, we used the well-established LeWoS tool (Wang et al., 2020a,b) to accomplish the task of leaf–wood separation. LeWoS is a fully unsupervised method based on the geometric characteristics of points and their spatial distribution. This leaf–wood separation method has been shown to be effective at classifying trunks and large branches (Wang et al., 2020a), displaying rather balanced accuracies for wood and leaf points (Wang et al., 2020a; Wu et al., 2020; Hui et al., 2021). LeWoS first applied a recursive segmentation to consolidate points into overly segmented clusters. Semantic information of each cluster was exploited to estimate its probability of being leaf or wood type. Such cluster level probability distributions were mapped back to the point level and the leaf–wood labels were further smoothed by a regularization step that takes into account spatial autocorrelation, i.e. that points near each other have the same label. LeWoS required one input parameter, feature threshold, that is crucial for the processing. We used a value of 0.125 as recommended by the original publication, as this value was the most robust based on the comprehensive sensitivity analysis done in Wang et al. (2020a). The leaf labels were upscaled to the original point cloud resolution using a nearest neighbor search. Each point in the full-resolution point cloud was classified as a leaf point if the nearest point in the downsampled point cloud was a leaf point. If a point of the full-resolution point cloud had a nearest neighbor that was a wood point, or if the distance to the nearest neighbor was more than 5 cm, it was classified as a non-leaf hit and did not contribute to the leaf area estimation in Section 2.4. For points that have an equal probability for being classified as leaf or wood, LeWoS takes into account spatial autocorrelation with nearby points to assign a final, regularized label. After classifying leaf and wood points, no manual intervention was done on the data.

2.3.3. Empty pulse recovery

By default, Leica Cyclone only exports non-empty points, i.e., pulses that triggered a return. For accurate estimation of the attenuation coefficient in a voxel grid, it is important to recover the direction of pulses which did not trigger a return. In a nutshell, we did this by arranging the pulses in a grid, calculating the horizontal and vertical resolutions, and filling in the missing points in the grid. This procedure was adapted for full-field-of-view scans from Schraik et al. (2021a).

The e57 file format, as Leica Cyclone exports it, contains information on the row and column ID of each point in the point cloud. We converted the point cloud from Cartesian to spherical coordinates, and processed azimuth directions (columns) and zenith directions (rows) in a similar way, but for azimuth angles we restricted the points to near-horizontal directions (within 0.2 radians) to avoid the increasing azimuth angle errors near the zenith. For each column ID number, we calculated the mean azimuth angle of all points of that column ID number. Then, we calculated the median difference between adjacent columns, which is an estimate for the azimuth resolution that is robust to outliers (i.e. empty columns). We calculated the azimuth angle for any column ID number that did not contain any points by adding the azimuth resolution to the neighboring column ID number. This process was applied similarly to the zenith angles, but the instrument design required us to calculate and assign empty pulses for the “left” and “right” sides of the scanner separately. These left and right sides correspond to the two faces of the scanner, since a single rotation of the scanner’s mirror scans the entire vertical field of view (one entire column) before rotating horizontally to scan the next column. The two opposite sides of the zenith are not perfectly aligned, and not distinguishing the two would result in an error of several resolution steps. Therefore, we estimated the zenith angles for each side of the zenith separately. Finally, after having assigned a nominal azimuth and zenith angle to every column and row ID number, respectively, we filled in row–column ID combinations that were not present in the data with empty pulses with an arbitrarily large range (200 m) and converted the point cloud back to Cartesian coordinates.

2.4. Silhouette to total area estimation

2.4.1. Overview

The following section describes our new method to estimate the silhouette to total area ratio at forest stand level ($\mathrm{STAR_f}$). We based our method on already existing methods to estimate leaf area density from TLS point clouds. $\mathrm{STAR_f}$ of each stand was estimated in two steps (Fig. 3). First, we established a voxel grid on the plot and estimated the attenuation coefficient for each voxel in the grid by ray tracing. The leaf area density, and consequently the LAI of the forest stand, was obtained from that attenuation coefficient. Then, the voxel-wise attenuation coefficient was used for a second ray tracing step, which we used to calculate the diffuse interceptance (i_D), and finally $\mathrm{STAR_f}$ was obtained as

$$\mathrm{STAR_f} = i_D/(4L). \tag{2}$$

The leaf area density estimation followed Pimont et al. (2018), and in addition we applied our newly developed method to correct for partial hits in the lidar data (Schraik et al., 2021b). The method of leaf area density estimation is based on the assumption that leaves are randomly distributed inside each voxel (i.e., the number of leaves in a voxel is Poisson distributed). Therefore, leaf area estimates do not account for clumping inside voxels, which concerns shoot clumping in conifer stands, but also clumping at branch scale, which can be significant for broadleaved tree species (Béland and Baldocchi, 2020) and potentially also for conifer species. The computations were carried out on the Puhti server of the Finnish IT Center for Science (CSC).

2.4.2. Leaf area estimation

We defined a voxel grid as a square cuboid with faces parallel to the point clouds’ coordinate system axes, which in turn were approximately aligned with geographic North (Fig. 2). Horizontally, the cuboid had side length of 25 by 25 m, and vertically it extended from −10 m (w. r. t. to the Z-coordinate from the ground below the first TLS scan) 50 m upwards. The cuboid was centered between the first and thirteenth TLS scan position, which marked the Southwestern (SW) and Northeastern (NE) corners of the plots. The azimuth direction of the line between the SW and NE corners of the plots was assumed to be exactly 45°. This square cuboid formed the bounding box of the voxel grid, i.e., a grid of cubic cells, with a voxel side length of 0.2 m. We chose the voxel size of 0.2 m based on previous findings which showed that within-voxel clumping increases significantly at voxel sizes larger than 20 cm (Schraik et al., 2021a). At 10 to 20 cm voxel size, voxels are more likely to be either gaps or filled with leaves, as this matched the typical gap size between branches of Norway spruce. At voxel sizes over 20 cm, most voxels contain a mixture of larger gaps and foliage, which considerably increased within-voxel clumping (Schraik et al., 2021a).

The TLS point clouds, which included the recovered empty pulses, were converted into rays, i.e., the TLS data from a plot were separated into single scans, and the scanner location, or starting coordinates of the rays, were noted. We estimated the attenuation coefficient for each voxel and for each scan using the modified contact frequency (MCF) method as described in Pimont et al. (2018). We used the voxel traversal algorithm by Amanatides and Woo (1987) for ray tracing. For each voxel, the MCF approximates the attenuation coefficient by measuring the path length of each ray that traversed the voxel, and noting whether or not the ray has been intercepted inside the voxel. The basic principle of estimating attenuation from lidar point clouds is to calculate the ratio of the fraction of hits in a voxel $I = \frac{N_i}{N}$ and the mean path length $\bar{z}$ that rays traveled within the voxel as (Pimont et al., 2018)

$$\lambda \approx \frac{I}{\bar{z}} = \frac{\sum \mathbf{1}_{z<\delta}}{\bar{z}\sum \mathbf{1}} = \frac{\sum \mathbf{1}_{z<\delta}}{\sum z}, \tag{3}$$

where $\mathbf{1}_{z<\delta}$ is an indicator function that takes the value 1 when a pulse is intercepted inside a voxel, and 0 otherwise. The subscript is denoted

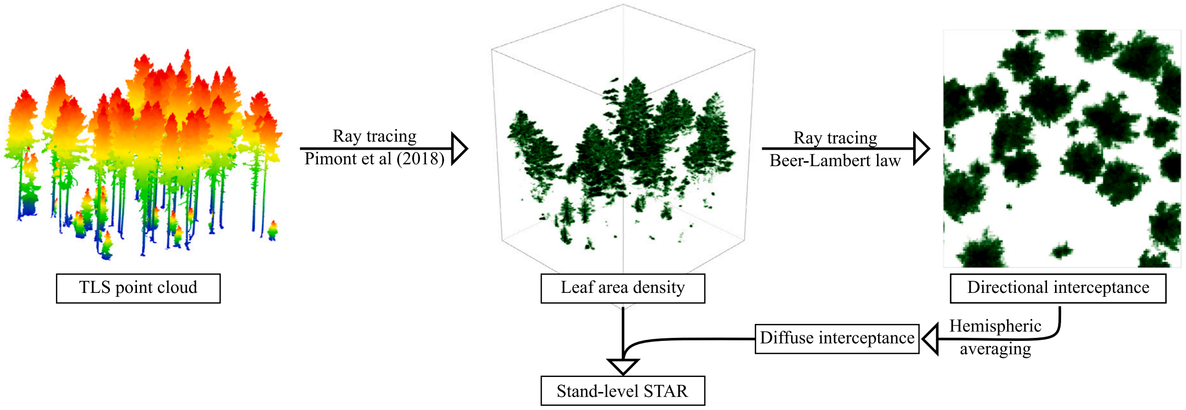


Fig. 3. Schematic of our proposed method to estimate STAR_f. The TLS point clouds are used to estimate leaf area density on a voxel grid. The leaf area index of a plot is the sum of leaf area density, multiplied by the voxel volume and divided by the plot area. In a subsequent ray tracing step, directional interceptance is calculated based on the leaf area density. The directional interceptance is averaged over the hemisphere to obtain diffuse interceptance. Finally, STAR_f is calculated as the ratio of diffuse interceptance to four times the leaf area index.

by the free path length z being smaller than the theoretical path length to the end of the voxel δ. The sum of $\mathbf{1}_{z<\delta}$ is the number of pulses that are intercepted in a voxel, and similarly, the sum of $\mathbf{1}$ is the number of all pulses that enter the voxel.

Lidar data, however, has some properties that introduce theoretical biases when applying the MCF method. First, laser beams decay after the first contact and thus do not always explore the entire voxel. This is typically considered by assuming that the unexplored volume of a voxel is statistically represented by the explored volume of the voxel. Secondly, the size of attenuating elements is finite in real vegetation, which would lead to considerable bias because Eq. (3) assumes infinitesimal element size. Thirdly, the number of beams that traverse a voxel is finite and sometimes even very small, which can cause biases. Pimont et al. (2018, section 3.5 therein) addressed the second and third mentioned biases by replacing the path length z by the effective path length z_e (accounting for finite element size), and by introducing a term to correct attenuation coefficient when the number of beams is small as

$$\lambda = \frac{1}{\sum z_e}\left(\sum \mathbf{1}_{z<\delta} - \frac{\sum_{\text{hits}} z_e}{\sum z_e}\right), \text{and} \tag{4}$$

$$z_e = -\frac{log(1-\lambda_1 z)}{\lambda_1}, \tag{5}$$

where $\mathbf{1}_{z<\delta}$ is an indicator function that takes the value 1 if a ray is intercepted in the voxel, and 0 otherwise, z_e is the effective free path length, calculated from the free path length z and the attenuation coefficient for a single scattering element λ_1, which is a constant factor that depends on leaf and voxel size. We set λ_1 to 0.0382, which corresponds to an expected leaf size of about 3 cm^2 at 0.2 m voxel size. This value of 3 cm^2 is the average projection area of a Norway spruce shoot, based on previous measurements (Schraik et al., 2021a), and we assumed that 3 cm^2 is reasonably close to the leaf area of all species in our data, since to our understanding, λ_1 only causes significant changes if its order of magnitude changes.

2.4.3. Partial hit correction

An additional bias source for MCF in TLS is the beam size effect, that is, the fact that the laser beam has a finite size in the order of several millimeters or centimeters, which causes targets not centered in the beam to trigger a return. This violates the assumption in the MCF that the probes have an infinitesimal size, and leads to overestimation of attenuation coefficient and leaf area (Soma et al., 2018; Schraik et al., 2021a). We used the partial hit correction method proposed by Schraik et al. (2021b), which estimates the fraction of each TLS return that is covered by vegetation targets, also called the per-pulse cover fraction. This correction is based on applying a one-inflated beta regression model to the return intensity, which estimates the per-pulse cover fraction for each leaf or needle hit. Schraik et al. (2021b) reported regression coefficients for Norway spruce, Scots pine, and silver birch. We used the regression coefficients for silver birch in all broadleaved plots in this study.

We applied the per-pulse cover fraction correction as a weight in Eq. (4) to express that the path length of a pulse that partially hit a target is the per-pulse cover fraction-weighted average of the intercepted path length, and the theoretical path length to the edge of the voxel. This correction assumes that a partial hit will travel up to the boundary of the voxel without being intercepted by any other target, which is below the longitudinal object separation threshold, since our TLS instrument's pulse is about 4 ns long, which corresponds to about 60 cm in scan range. Therefore, objects which are separated by less than 60 cm become indistinguishable from each other. For simplicity, we assumed that a fraction of the pulse (as indicated by the per-pulse cover fraction) is intercepted in the voxel, and the remaining fraction travels to the nearest edge of the voxel. Combining the per-pulse cover fraction weight w_{cf} with Eqs. (4) and (5), we get

$$\lambda = \frac{1}{\sum z_e}\left(\sum w_{cf}\mathbf{1}_{z<\delta} - \frac{\sum_{\text{hits}} z_e}{\sum z_e}\right), \text{and} \tag{6}$$

$$z_e = -\frac{log(1-\lambda_1(z w_{cf} + \delta(1-w_{cf})))}{\lambda_1}, \tag{7}$$

where the number of hits is modified by the per-pulse cover fraction to account for partial hits, and the effective free path length is an interpolation between the free path length z and the theoretical path length δ that the ray would travel if it were not intercepted. The leaf area density for each scan l and voxel j was calculated as

$$L_{j,l} = \frac{\lambda_{j,l}}{G(\theta_{j,l})}, \tag{8}$$

with G, the projection area of unit leaf area at the zenith angle $\theta_{j,l}$ between the scanner position and the voxel. We describe the parameterization of $G(\theta_{j,l})$ in Section 2.4.4. Note that leaf area estimates do not include clumping at scales below the voxel level, therefore the leaf area estimates are not corrected for shoot clumping in conifers or branch scale clumping in all tree species. This was done to maintain comparability with effective LAI in hemispheric photographs (Section 2.5).

Finally, the leaf area density was averaged over all 16 scans per plot, weighted by the number of rays per scan n_l that traversed each voxel as

$$\overline{L_j} = \frac{\sum_{l=1}^{16} L_{j,l} n_l}{\sum_{l=1}^{16} n_l}. \tag{9}$$

Table 2
Leaf inclination angle statistics for parameterizing the G-function. St.D. stands for standard deviation. For Norway spruce and Scots pine, we used leaf angle distribution functions from de Wit (1965), parameterized for the beta distribution by Weiss et al. (2004) because no direct measurements were available. For other species, we directly used measured mean and standard deviation.

Tree species	Mean	St.D.	De Wit type (if used)	Reference
Norway spruce	26.8	32.54	Planophile	Janoutová et al. (2019), Weiss et al. (2004)
Scots pine	57.3	61.22	Spherical	Stenberg et al. (1993), Weiss et al. (2004)
Birch	57.88	17.49	–	Raabe et al. (2015)
Alder	51.07	21.05	–	Pisek et al. (2013)
Aspen	41.75	23.13	–	Pisek et al. (2013)
Ash	50.56	17.51	–	Pisek et al. (2022)
Oak	35.80	19.44	–	Pisek et al. (2013)
Hornbeam	38.19	16.88	–	Pisek (unpublished data)

We calculated the LAI per plot as the sum of leaf area density over all voxels in the plot, multiplied with the voxel volume V, and divided by the stand area $A_C = 25 \times 25 m^2$ as

$$L = \sum_{j=1}^{N} \overline{L_j} V \frac{1}{A_C}. \tag{10}$$

2.4.4. G-function and leaf angle distribution

The directional projection area of unit leaf area (G) was parameterized with the two-parameter beta distribution (Wang et al., 2007), which required mean and standard deviation of leaf inclination angles. We obtained those leaf angle statistics from the literature (Table 2), and parameterized a single G-function for each plot based on the dominant species. For each voxel and each TLS position, we calculated the zenith angle θ between the TLS position and the voxel midpoint. Then, using the beta distribution, we parameterized the probability density of leaf inclination angles $f(\theta_L)$. To obtain $G(\theta)$, we integrated the projection area of all leaf inclination angles as (Wang et al., 2007)

$$G(\theta) = \int_0^{\pi/2} A(\theta, \theta_L) f(\theta_L) \mathrm{d}\theta_L \tag{11}$$

$$A(\theta, \theta_L) = \begin{cases} \cos(\theta)\cos(\theta_L) & \text{if } |\cot(\theta)\cot(\theta_L)| > 1, \text{ and} \\ \cos(\theta)\cos(\theta_L)\left(1 + \frac{2}{\pi}(\tan(\Psi) - \Psi)\right) & \text{, otherwise,} \end{cases} \tag{12}$$

with $\Psi = \arccos(\cot(\theta)\cot(\theta_L))$. For all broadleaved species, leaf angle statistics were available in the literature, and therefore we used the species-specific parameters directly in the beta distribution. As of writing, there were, to the best of our knowledge, no scientific studies that published shoot angle distribution parameters on Norway spruce and Scots pine. Janoutová et al. (2019, Fig. 2 therein) shows shoot angle measurements of Norway spruce, which average at around 27°, but there was no information available from which standard deviation could be deduced. The mean shoot angle suggests a planophile shoot angle distribution for spruce (Janoutová et al., 2019). Scots pine was reported to exhibit a spherical shoot angle distribution (Stenberg et al., 1993). Thus, for Scots pine and Norway spruce we used the spherical and planophile leaf angle distribution types of de Wit (1965), which we obtained by using the parameters from Weiss et al. (2004) (Table 2). Note that for conifer species, we used the shoot angle distribution, as the shoot can be considered the basic scattering element in conifer canopies (Rautiainen and Stenberg, 2005).

2.4.5. Estimation of diffuse interceptance

The voxel grid containing the averaged leaf area density was used to create synthetic orthogonal projection images, where each pixel corresponded to a ray traversing through the voxel grid, and the pixel value corresponded to the intercepted energy i according to the Beer–Lambert law in each voxel as

$$i = 1 - \prod_{j=1}^{M} exp^{-\overline{\lambda}_j \delta_j} = 1 - \prod_{j=1}^{M} exp^{-G(\theta)\overline{L}_j \delta_j}, \tag{13}$$

with the attenuation coefficient $\overline{\lambda}_j$, path length δ, leaf area density L and leaf projection function G for each voxel j that a ray traversed at zenith angle θ, and M denoting the total number of voxels that the ray traversed. The rays entered the voxel grid from the top and were evenly spaced in a grid with 9.9 cm spacing. On stand level, $\mathrm{STAR_f}$ must be calculated from a horizontally infinite canopy, therefore, if rays would exit the voxel grid at any horizontal boundary, they were set to re-enter the voxel grid at the same position but on the opposite side of the voxel grid. The pixel value (between 0 and 1) was interpreted as the fraction of the pixel that is covered by vegetation targets, and the directional interceptance was calculated as the mean of the pixel values.

We obtained the diffuse interceptance (i_D) of each stand as the numerical integral of the directional interceptance over the hemisphere, calculated at 100 evenly spaced angles throughout the hemisphere (10 azimuth and zenith angle steps, respectively).

2.5. Comparison measurements of leaf area index

We obtained measurements of effective LAI using 16 hemispheric photographs (HP) per plot. The HPs were thresholded using the method by Nobis and Hunziker (2005), and divided into zenith angle rings, from which the gap fractions T_k were calculated as the sum of (white) sky pixels to total number of pixels per ring. The effective LAI was obtained as

$$L_e = -2 \sum_{k=1}^{5} \ln(T_k) \cos\theta_k W_k, \tag{14}$$

with the weights $W_k = \sin\theta_k d\theta_k$ (normalized to $\sum_{k=1}^{5} W_k = 1$) for five concentric zenith rings that had median (and range) of zenith angles θ_k as 10.7° (0-15°), 23.7° (15-30°), 38.1° (30-45°), 52.8° (45-60°), and 66.6° (60-73°). These angles match closely those used by the LAI-2000 and LAI-2200 instruments, and the calculations were done using weights similar to the LAI-2200 instrument (LI-COR, Inc., 2012).

2.6. Forest variables

We examined the relationship of $\mathrm{STAR_f}$ with eight forest variables. The tree height, mean DBH, basal area and number of trees per hectare were described in Section 2.1, and effective LAI was described in Section 2.5. In addition, we calculated crown length from the TLS data, and diffuse non-interceptance and vertical canopy gap fraction from hemispherical photographs. For the crown length, we first calculated the vertical leaf area density profile, and then calculated the mean crown length as the sum of voxel slices which contained more than 25% of the leaf area relative to the slice with the maximum leaf area. 25% was chosen so that the estimate would be robust to noise and the rare case that tree stems were not excluded by the leaf–wood separation. The resulting crown length likely differs from crown lengths measured in the field, but we expected it to nevertheless be a suitable proxy for actual crown length. The vertical canopy gap fraction was the uppermost zenith ring's gap fraction from the hemispherical photographs, which we used as a proxy for vertical canopy cover.

Table 3
Observed $\mathrm{STAR_f}$ by dominant species and site, summarized by mean, standard deviation, and number of plots (n). Species which occurred in multiple sites were summarized altogether. Species with only one plot do not have a standard deviation.

Species	Site	Mean	St.dev.	n
Pinus sylvestris	all	0.086	0.016	8
	Hyytiälä	0.087	0.018	6
	Järvselja	0.082	0.007	2
Picea abies	all	0.082	0.013	11
	Hyytiälä	0.072	0.007	6
	Järvselja	0.095	–	1
	Bílý Kříž	0.092	0.009	4
Betula sp.	all	0.109	0.034	6
	Hyytiälä	0.125	0.029	4
	Järvselja	0.076	0.010	2
Alnus glutinosa	Järvselja	0.071	0.007	2
Populus tremula	all	0.081	0.014	3
	Järvselja	0.088	0.010	2
	Lanžhot	0.068	–	1
Quercus sp.	Lanžhot	0.068	0.002	6
Fraxinus excelsior	Lanžhot	0.054	–	1
Carpinus betulus	Lanžhot	0.066	–	1

In addition, we calculated the diffuse non-interceptance (DIFN) from hemispheric photographs as DIFN $= \sum_{k=1}^{5} T_k W\prime_k$, with the weights $W\prime_k = \sin\theta_k \cos\theta_k d\theta_k$ (normalized to $\sum_{k=1}^{5} W\prime_k = 1$). The data for this analysis, containing values for $\mathrm{STAR_f}$, the above mentioned forest variables, and species shares, can be found in supplementary file S1.

2.7. Satellite data

For each of our study areas, we obtained a cloud-free Landsat 8 OLI image that was temporally closest to the field campaign. The images were acquired on 31th Aug 2019 for Bílý Kříž and Lanžhot, 28th Jun 2019 for Hyytiälä, and 25th Jun 2020 for Järvselja. We used Collection 2 Level-2 surface reflectance (Earth Resources Observation And Science (EROS) Center, 2013) downloaded from U.S. Geological Survey (2000) for each plot. We used spectral bands 2–7 of OLI, spanning the region of 450–2290 nm. For each plot, we used the value from a 30 m OLI pixel that intersected with the plot center. We compared single pixels to the average of two by two pixels, and found only negligible differences owing to the homogeneity of the stands and the fact that plots were located at least 30 m away from the nearest stand border. Accurate coordinates (sub-meter accuracy) for the plot centers were found by matching treetops in TLS data with those found in airborne laser scanning data collected from the study areas.

3. Results and discussion

3.1. Stand STAR

We calculated TLS-based $\mathrm{STAR_f}$ for 38 plots in four study areas in Finland, Estonia and Czech Republic. Overall, $\mathrm{STAR_f}$ ranged between 0.054 and 0.143, with an average of 0.08 and a standard deviation of 0.02 (Table 3). On average, conifers had a higher $\mathrm{STAR_f}$ than broadleaved trees, at stand level, however, these values do not include shoot level clumping, and therefore the overall STAR of spruce and pine may be similar or slightly lower than the broadleaved plot STAR values. On the other hand, broadleaved trees have been shown to exhibit branch scale clumping, which occurs below the voxel scale (Béland and Baldocchi, 2020). Conifer species likely also exhibit branch scale clumping to some degree. The lowest $\mathrm{STAR_f}$ values of around 0.054 to 0.069 were found throughout Lanžhot, where stands were dominated mostly by oak, one pure ash plot, and two mixed plots dominated by hornbeam and aspen, respectively.

Scots pine exhibited the highest $\mathrm{STAR_f}$ values, with slightly higher values in Hyytiälä (0.087) than in Järvselja (0.082). In addition, Hyytiälä pine plots exhibited the highest variation, with a relative standard deviation (RSD) of 21%. The high variation is greatly influenced by a single plot with mean tree height of about 8 m, which had the exceptionally high $\mathrm{STAR_f}$ of 0.12. Without this plot, the mean of Hyytiälä pine plots would have been 0.080, with a RSD of 6%. Mean $\mathrm{STAR_f}$ for pine decreased to 0.048 when taking into account shoot clumping, assuming a commonly used shoot clumping factor of 0.6.

Norway spruce exhibited considerably different $\mathrm{STAR_f}$ values in the different study sites. The boreal Hyytiälä site had a mean $\mathrm{STAR_f}$ of 0.072 (0.043 with shoot clumping), while the Järvselja and Bílý Kříž sites had 0.095 and 0.093 (0.057 and 0.056 with shoot clumping), respectively. This difference could stem from the lower LAI but also higher occlusion in the Estonian and Czech sites, as the spruce trees in Bílý Kříž were taller on average than those in Hyytiälä, which would lead to lower point cloud coverage, especially in the upper portions of the canopy. Another possible explanation is that it may be due to different crown morphology between boreal and temperate spruce trees.

In birch plots, $\mathrm{STAR_f}$ ranged from 0.069 to 0.143, which was the largest range in $\mathrm{STAR_f}$ values we observed. Among the four birch plots in Hyytiälä, three had exceptionally high $\mathrm{STAR_f}$ values with an average of 0.140. However, there was an exceptional birch plot in Hyytiälä that had $\mathrm{STAR_f}$ of 0.083, and was in this regard closer to the birch plots in Järvselja, which had a mean of 0.077 and 14% standard deviation. This exceptional birch plot had a denser spruce understory than the other birch plots in Hyytiälä, which contributed about 0.4 to the LAI, according to the TLS leaf area estimates. Some spruce understory was also present in other Hyytiälä birch plots, but there it only contributed up to 0.16 to LAI. Another possible explanation is that the other three birch plots were underestimated in their LAI compared to the LAI from hemispherical photography (Fig. 5), which would cause an increase in the $\mathrm{STAR_f}$ value. The reduced density of the birch plots with high $\mathrm{STAR_f}$ can be seen simply from the point counts of the downsampled point clouds (which were cropped to cover only the plot area), as they had about 212 to 331 thousand points, compared to the about 530 thousand points in other plots.

The most homogeneous results were found in the Lanžhot site. There, stands were dominated by oak, European ash, hornbeam, and aspen, with field maple or hornbeam mixed in. The most common species was oak, for which we observed a mean $\mathrm{STAR_f}$ of 0.068, with only 2.0% RSD. The average for all Lanžhot plots is slightly lower, at 0.066, with 7.1% RSD, which was mainly caused by the ash plot in the site, which had a low $\mathrm{STAR_f}$ of 0.054. Without the ash plot, the mean $\mathrm{STAR_f}$ for all Lanžhot plots was 0.067, with 2.5% RSD. Considering the tree heights ranging from 18.5 m to 40 m (and height roughly representing differences in stand age), $\mathrm{STAR_f}$ appeared to remain relatively constant throughout the plots in Lanžhot regardless of species or height. The younger oak stands had a rather shallow canopy with a high degree of homogeneity, while canopies in the older stands were dominated by few large and tall tree crowns. The Lanžhot plots with smaller LAI (to an observed minimum of 2.8) also had smaller clumping index (min. 0.61), whereas the plots with relatively high LAI (max. 3.4) had higher clumping indices (max. 0.91). The only exception was a clumped and dense ash stand (LAI=4.0, CI=0.66), which also had an exceptionally low $\mathrm{STAR_f}$=0.05 for Lanžhot. The other eight plots showed that stands with low LAI tended to also have lower clumping index in our data. What the Lanžhot plots show us is that clumping can occur in young broadleaved stands where the canopy is very shallow, where the canopy elements may well be randomly distributed horizontally, but vertically they are highly clumped.

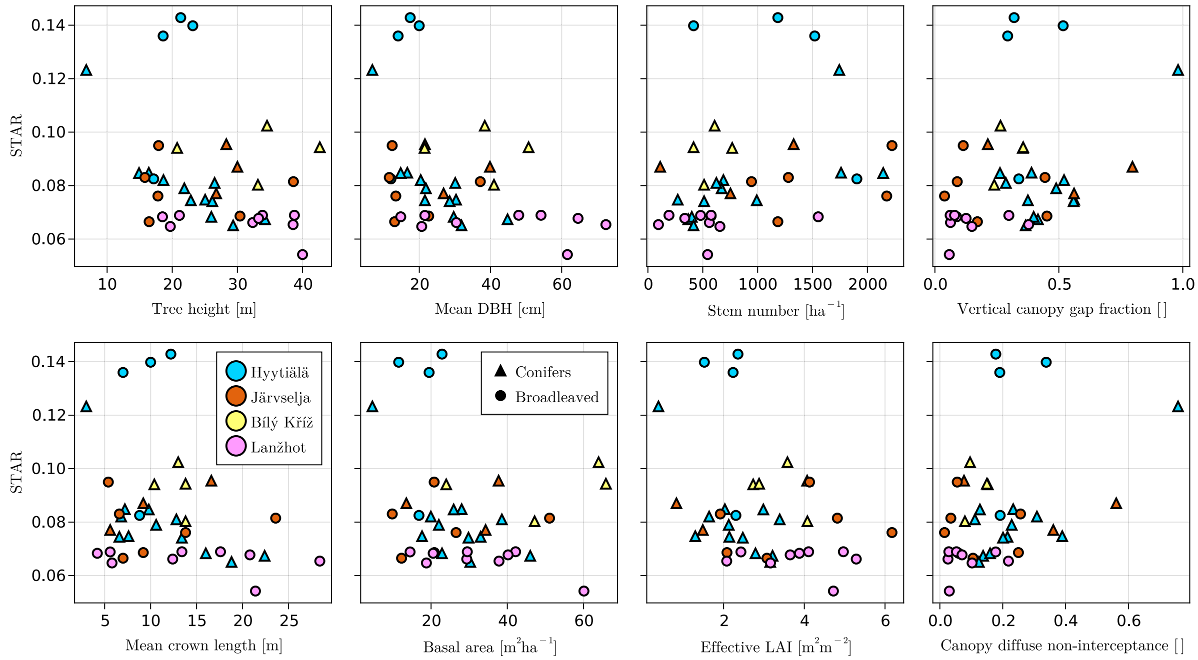


Fig. 4. Relationship between stand STAR$_f$ and different forest variables which were measured in the field. Triangle symbols represent conifer plots (Norway spruce and Scots pine), and circle symbols are broadleaved plots. The colors refer to the study sites.

3.2. Forest variables and STAR$_f$

We examined the relationship between STAR$_f$ and eight easily measurable forest variables to determine whether STAR$_f$ may be estimated indirectly. Since STAR$_f$ is determined by the LAI, clumping index, and leaf angle distribution, which are in turn usually species-specific variables and closely related to other forest descriptors such as stand age or biomass, we expected to potentially find relationships that would allow regression of STAR$_f$ using forest variables that are more commonly measured or do not require expensive equipment or data collection and processing. Such an indirect estimate would ease applying knowledge about STAR$_f$ to new study sites, or to map STAR$_f$ over large areas. The relationship between STAR$_f$ and forest variables in our data indicates that conifers have a seemingly distinct trend towards decreasing STAR$_f$ values with increasing tree height, crown length, mean stem diameter, basal area and decreasing number of trees per hectare (Fig. 4). This is contrary to broadleaved plots, which, especially in Lanžhot, indicate a STAR$_f$ that is nearly independent from stand characteristics and has generally little variability. STAR$_f$ in broadleaved plots in Järvselja and Hyytiälä exhibited more variability, but also showed no clear relationship with other forest variables. However, the broadleaved plots in our data were diverse, spreading over six major tree species, with several species only represented in one or two plots.

The clearest candidate for estimating STAR$_f$ would be LAI, since it is itself a factor in calculating STAR$_f$. However, to obtain LAI from optical measurements made in the field, one would require an accurate measurement of the clumping index, which can be difficult to obtain directly (Fang, 2021). We found the highest absolute Pearson correlation coefficients in mean DBH (−0.40), diffuse non-interceptance (0.38), effective LAI (−0.36), tree height (−0.34), and vertical gap fraction (0.33). Crown length, basal area, and stem number had slightly lower absolute correlation coefficients, which ranged from 0.22 to 0.31. Given our limited sample size, and the plot selection criteria (which imposed an upper limit to canopy density to avoid excessive occlusion in TLS point clouds), we note that more field measurements are required to accurately examine the relationship between STAR$_f$ and forest variables, especially for broadleaved-dominated forests. The fact that the effective LAI is not the single most correlated forest variable, as we would have expected, is likely because it was measured using a different technology, i.e., hemispherical photography rather than TLS. We conclude that although it seems possible to indirectly estimate STAR$_f$ using variables commonly measured in forest inventory or hemispherical photography, our results indicate that forest structure may be best quantified using lidar point clouds.

3.3. Comparison to hemispherical photography

We compared the TLS-derived LAI to the effective LAI we obtained through hemispherical photography because neither LAI accounts for shoot clumping, and thus, in principle, should be similar if there were no higher level clumping. Although TLS and other active sensing techniques provide new ways to explicitly and potentially more accurately quantify forest structure in 3D, hemispheric photography remains one of the best established and most efficient and accessible means of measuring canopy structure (Chianucci, 2019). In the absence of destructive measurements, hemispherical photography provided a valuable means of validating our TLS-based leaf area estimates.

In Hyytiälä, LAI of Scots pine stands showed good agreement between TLS and HP, with a mean difference (MD) of −0.13 and a root mean square deviation (RMSD) of 0.45. In contrast, the LAI of pine stands in Järvselja were slightly overestimated by TLS (MD=-0.53), which may be due to the sparser pine stands in this site. The Norway spruce plots showed a contrasting trend between Hyytiälä and the other sites, with the Hyytiälä LAI being only slightly underestimated by the TLS (MD=0.09), whereas in the Czech Republic and Estonia the TLS considerable underestimated LAI (MD=1.29 and 2.02, respectively). The TLS-derived LAI in Lanžhot appeared to be overestimated in plots with low LAI, whereas plots with average and high LAI were underestimated. This may partially explain the low variation in STAR$_f$ in Lanžhot plots. Although we cannot exclude a possible influence of occlusion in these plots, only two plots in Lanžhot showed non-negligible percentage (0.1%–1%) of voxels which were traversed by less than five rays, a limit which has been shown to result in sufficiently unbiased estimates of leaf area density (Pimont et al., 2019). Possible reasons for the TLS-leaf area estimates' lack of agreement with hemispherical photography estimates could be the partial hit correction using the per-pulse cover

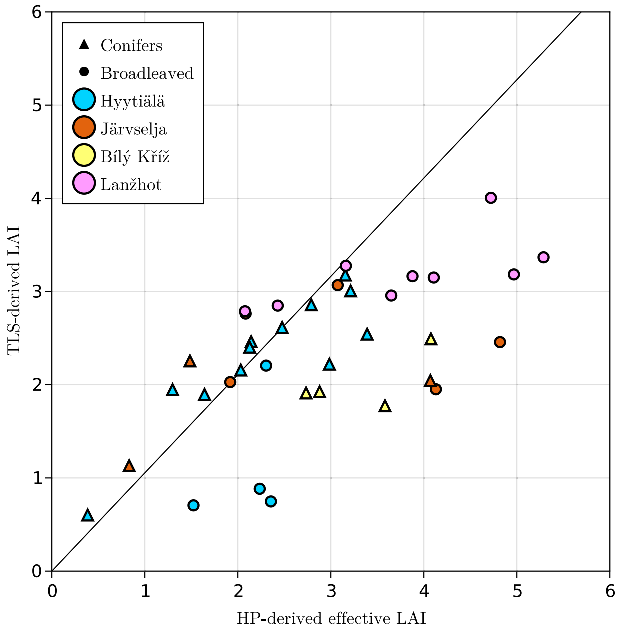


Fig. 5. Comparison between TLS-derived leaf area index (LAI) and HP-derived effective plant area index. The TLS-derived values represent leaf area index, i.e. they do not contain woody elements, and they account for clumping at scales above 20 cm. The HP-derived values contain woody elements, as it was not possible to separate them from leaves, and they do not account for clumping.

fraction, for which coefficients were only available for birch (Schraik et al., 2021b). Since the per-pulse cover fraction correction incorporates both structural and spectral properties of the lidar measurement into a regression model, it is possible that the different shoot structure and leaf optical properties of oak and other species in Lanžhot were not very suitable for this correction without separate measurements of the regression coefficients for those species. In addition, the leaf angle distribution may have influenced the LAI estimates in some of the Lanžhot plots. We relied on leaf angle parameters published in the literature, however, for some species it was not possible to obtain parameters that match the region as well as the season, as is the case with hornbeam, for which we relied on measurements that were done on site, but early in the vegetation period, whereas TLS measurements were done late in the vegetation period. In Järvselja, the TLS-derived LAI reached an apparent maximum at a value around 3, possibly due to occlusion, and did not increase anymore, while the HP-derived LAI indicated higher values (Fig. 5).

One likely factor explaining differences between LAI from hemispheric photographs and TLS is the exclusion of woody components in the TLS-based estimates, whereas the hemispheric photograph-LAI includes wood area, and is thus actually a plant area index. Separating leaf and wood was not possible from hemispherical photos because most of the stands were too dense, rendering the vegetation pixels very dark (pixel values close to zero), and thus there was insufficient dynamic range for separating leaf and wood pixels based on their color. We excluded wood points from our TLS data because the presence of wood points in a voxel can cause considerable bias to LAI estimates (Pimont et al., 2019). A solution to this would be to model the wood area separately using quantitative structure models (e.g. Raumonen et al., 2013), and including the tree surface area in both the estimation of total leaf area while accounting for the wood volume reduction of voxels (Pimont et al., 2019) as well as in the ray tracing step for calculating diffuse interceptance. We expect $\mathrm{STAR_f}$ values would decrease if wood area were taken into account, since the contribution of wood area to the total plant area is greater than to the diffuse interceptance, as woody parts tend to be shaded by leaves. The woody components can make up between 5% and 35% of the total plant area (Gower et al., 1999). Considering this range of wood fraction, the leaf area index estimates from hemispheric photography would be even smaller than the TLS estimates. Such an increased difference could indicate that the TLS method to estimate leaf area is closer to the true leaf area index than hemispheric photography because the TLS method is robust against clumping at scales above the voxel size.

3.4. Limitations and uncertainty of the proposed method

Our method is not limited to voxel-based quantification of leaf area. Any three-dimensional reconstruction would be suitable for our method, provided it retains the spatial distribution of canopy elements. The use of TLS data in our study, similarly, was a matter of pragmatic choice that was done because TLS provides extremely high accuracy data and resolution. However, lidar scanning from unoccupied aerial vehicles or mobile lidar systems, may, in principle, provide sufficient accuracy to produce a good estimate of leaf area density, and thus $\mathrm{STAR_f}$. We hypothesize that a key factor influencing the potential of close range aerial and mobile lidar systems is the noise in point coordinates, and whether a negative effect of increased positional noise on leaf area density estimation can be mitigated by scanning at high resolution. Additionally, one needs to consider the beam footprint of any lidar system and account for effects of beam size in leaf area density estimation. Future research should investigate suitable acquisition parameters and partial hit correction for these lidar systems and their effects on leaf area density, $\mathrm{STAR_f}$ and CI estimation.

The estimates of leaf area density do not include clumping below the voxel scale (20 cm). This limitation is inherent to current lidar measurement and processing methods, as leaves and needles cannot, to date, be reconstructed from point cloud data. The relatively large size of terrestrial lidar beams, which are the most accurate and small footprint available, compared to many broad leaves, and virtually all conifer needle-leaves, make geometrically explicit reconstruction impossible. An alternative to using voxel-based leaf area density estimates could be indirect reconstruction. With this approach, one could avoid relying on the Beer–Lambert law for within voxel attenuation to calculate directional transmittance, and thus relax the assumptions of random distribution of infinitely small foliage within voxels. Several attempts at simulation, for example, based on point density (Janoutová et al., 2019) or position within a tree quantitative structure model (Åkerblom et al., 2018), could be combined with a voxel-based leaf area density method to achieve an indirect reconstruction method.

Another weakness of TLS point clouds is the potential of occlusion in the upper parts of the canopy (Schneider et al., 2019). In our data, less than 1% of voxels were occluded, that is, traversed by less than 5 rays (Pimont et al., 2018). This may partly be due the exceptionally high scan resolution of 0.23 mrad that we used. However, uneven sampling of the upper canopy or the inner volume of large tree crowns, could have introduced estimation errors in our data.

Separating leaf and wood points was necessary for estimating leaf area density, as the distribution of woody elements in voxels cannot be considered random and therefore using wood points would introduce a significant bias (Pimont et al., 2019), as was mentioned in Section 3.3. The leaf–wood separation algorithm LeWoS has been shown to accurately classify trunks and large branches (Wang et al., 2020a), and overall providing balanced estimates of leaf and wood points (Wu et al., 2020; Hui et al., 2021). Locally, however, some random errors may have occurred within tree canopies, as our visual inspection indicated the occurrence of minor errors in less than 5% of trees.

Bias in the leaf area density estimates would propagate further into $\mathrm{STAR_f}$ and clumping index. However, $\mathrm{STAR_f}$ and clumping index may be somewhat less biased than the leaf area density estimates, as the former describe the spatial distribution of leaf area. Thus, a global bias in leaf area density may lead to only a small bias in clumping index

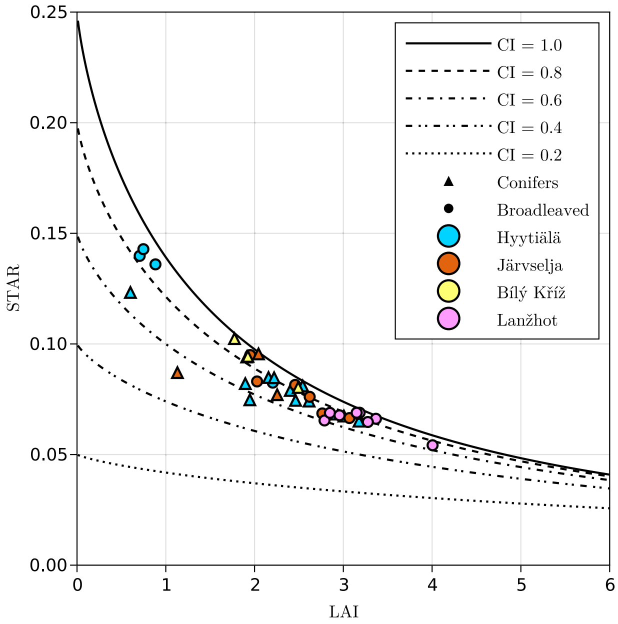


Fig. 6. Theoretical relationship between leaf area index, STAR, and clumping index (CI). The lines denote the theoretical STAR at different clumping index values, which were calculated from LAI using Beer's law with a fixed G-value of 0.5.

and STAR, as a bias in leaf area density and the propagated bias in directional interceptance partially cancel each other out.

3.5. Relationship between $STAR_f$, LAI and clumping index

Most plots had stand-level $\mathrm{STAR_f}$ values between about 0.06 and 0.10. This, however, does not directly mean that they all exhibit similar levels of clumping, as $\mathrm{STAR_f}$ is a summary variable of canopy structure and depends on leaf area index, clumping index, and the leaf angle distribution. At increasing leaf area index, $\mathrm{STAR_f}$ will decrease at constant clumping index. This is due to the relationship between $\mathrm{STAR_f}$ and the leaf area index $\mathrm{STAR_f} = \frac{i_D}{4L} = \frac{i_D \Omega}{4L_e}$, and the fact that i_D cannot exceed one, hence there is a theoretical maximum value of $\mathrm{STAR_f}$ at a given LAI. At constant $\mathrm{STAR_f}$, a stand with small LAI exhibits a higher degree of clumping than a stand with large LAI. For illustration, we analyzed the relationship between LAI, clumping index, and $\mathrm{STAR_f}$. We calculated the theoretical diffuse interceptance as a function of CI, LAI, and leaf angle distribution (Eqs. (1), and (12)) to obtain the relationship between CI, LAI and $\mathrm{STAR_f}$. Inverting this relationship numerically with CIs calculated in 0.01 increments, and LAI in 0.1 increments, and using k-nearest neighbor interpolation (k=4), we found CIs between 0.53 and 0.96 in our plots (Fig. 6, with lines indicating the theoretical $\mathrm{STAR_f}$ of a spherical leaf angle distribution).

The lowest average clumping indices were found in pine stands in Järvselja and spruce stands in Hyytiälä, with an average CI of 0.60 (range 0.54 to 0.67) and 0.63 (range 0.57 to 0.75), respectively. In broadleaved species, the Lanžhot plots were most clumped, with CIs ranging from 0.53 to 0.76, with oak exhibiting CIs throughout the range, while other species' plots had CIs over 0.64. In Järvselja, broadleaved stands were moderately clumped, at CIs between 0.70 and 0.79, while the spruce plot had a relatively high CI of 0.82. The spruce plots in Bílý Kříž were also moderately clumped with CIs between 0.7 and 0.77. The most random canopy was observed in birch plots in Hyytiälä, where the plots had CIs of 0.78, 0.85, 0.89, and 0.96. The plots with the high clumping indices over 0.8 correspond to the three birch plots discussed in Section 3.1.

Overall, we observed considerable levels of clumping throughout the sites in the Czech Republic, Estonia and Finland. The clumping indices reported here quantify clumping exclusively at levels above the voxel scale (i.e., 20 cm). Therefore, real clumping indices are likely even smaller, if one takes into account shoot level clumping for conifers, and branch level clumping (Béland and Baldocchi, 2020) for all species.

3.6. Influence of stand STAR and clumping index on forest reflectance

Forest structure has a considerable influence on the radiation regime by governing light interception and transmission through a canopy. $\mathrm{STAR_f}$ should indicate at least some degree of correlation with forest surface reflectance, as it quantifies forest structure in a single value per stand that depends on the LAI, clumping index, and the leaf angle distribution. Besides STAR, we also analyzed the correlation between clumping and reflectance, since the influence of stand scale clumping on reflectance has not yet been widely studied.

$\mathrm{STAR_f}$ decreases with increasing LAI and decreasing clumping index. Therefore, at a given LAI, a more clumped canopy will have the smaller $\mathrm{STAR_f}$ value, and at a given clumping index, the stand with higher LAI will have the smaller $\mathrm{STAR_f}$ value. With increasing LAI, the optical signal is increasingly dominated by the canopy relative to the forest floor (Hovi et al., 2022). Clumping influences the directional gap fractions as well as the photon recollision probability of a canopy at a given LAI. Even though a clumped canopy tends to be darker due to an increased photon recollision probability, it also tends to have higher transmittance compared to a canopy with randomly distributed elements (Pisek et al., 2010). This increasingly open canopy exposes more of the often bright forest floor to sunlight and the sensor, which might increase the overall surface reflectance despite the darker clumped canopy (Spanner et al., 1990).

$\mathrm{STAR_f}$ showed varying relationships with surface reflectance for different species and sites; Pearson correlation coefficients r ranged from −0.82 to 0.78. The wide range of r likely indicates the variable effect of the forest floor in more clumped stands. We can analyze the underlying cause for these fluctuating relationships in terms of the PARAS model (Rautiainen and Stenberg, 2005). This model is based on the photon recollision probability, which was used to derive $\mathrm{STAR_f}$ in this study. The PARAS model uses the gap fractions in illumination and view directions to quantify the contribution of the forest floor to reflectance. The canopy contribution is quantified by the interceptance in illumination direction and another term that depends only on the photon recollision probability as a structural variable. In this context, $\mathrm{STAR_f}$ is closely related to the photon recollision probability, i.e., it quantifies what happens to a photon once it has been intercepted by the canopy. In order to explain the influence of forest structure on reflectance, both $\mathrm{STAR_f}$ and directional gap fractions are needed. Clumping index, on the other hand, plays an important role for both $\mathrm{STAR_f}$ and directional gap fractions, and can therefore be expected to strongly influence forest reflectance.

As stated above, a clumped canopy becomes darker, but also transmits more light to the forest floor. We observed a clear relationship between clumping index and reflectance (Fig. 7). In conifer stands, the reflectance increased with decreasing clumping index, with most r values ranging from −0.75 to −0.79 (except $r = -0.52$ in the NIR band). In other words, clumped stands appear brighter than stands with randomly distributed elements. This is likely due to the above-mentioned two effects of clumping. The increased gaps in clumped stands, and the subsequently increased transmittance of forest floor-reflected radiation appear to outweigh the darkening of the canopy with increasing clumping. These results confirm earlier hypotheses on the effect of clumping on forest reflectance (Spanner et al., 1990; Rautiainen and Stenberg, 2005). On the other hand, a similar effect could not be found in broadleaved stands, where clumping index was slightly positively correlated with reflectance (with r of 0.1 to 0.28 in visible

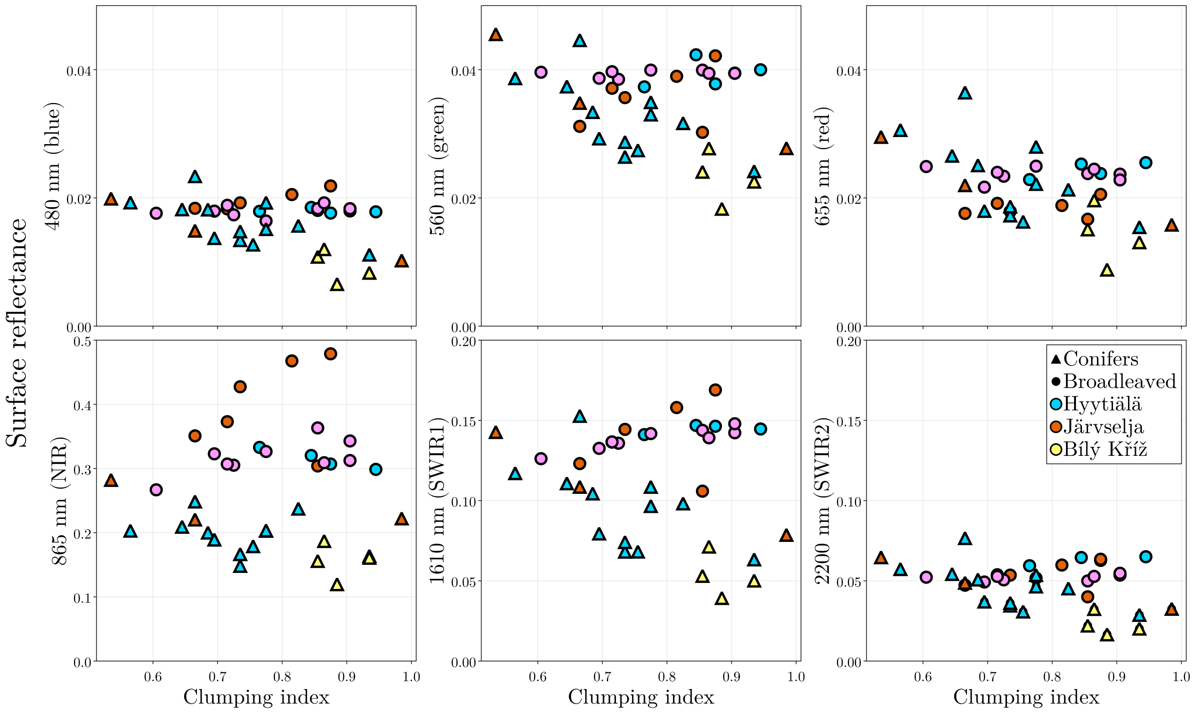


Fig. 7. Relationship between clumping index and Landsat 8 OLI surface reflectance observations by spectral bands. Triangle symbols represent conifer plots (Norway spruce and Scots pine), and circle symbols are broadleaved plots. The colors refer to the study sites.

and NIR bands, and 0.41 to 0.43 in SWIR). This correlation could be caused by differences in leaf optical properties, as different broadleaved species exhibited different ranges of clumping index. Another possible explanation is related to the connection between tree layer and understory vegetation: forest canopy leaf area index and the fractional cover of understory have been shown to be strongly related (Majasalmi and Rautiainen, 2020), and furthermore, variation in optical properties of forest floor and understory vegetation depend on canopy transmittance (Forsström et al., 2023). If the difference in reflectance between the broadleaved plots was purely due to canopy structure, and considering that broadleaved plots had higher average LAI than conifer plots, we would conclude that at a certain LAI (or canopy density), more clumping leads to lower reflectance, i.e. the clumping-induced increase in transmittance is smaller than the clumping-induced darkening of the canopy. However, further data on species-specific variation would be needed to confirm these results.

4. Conclusions

We developed, applied and assessed a method to estimate the silhouette to total area ratio of forest stands ($STAR_f$) and clumping index on an extensive point cloud dataset covering 38 plots in four study sites across the boreal, hemiboreal and temperate biomes in Europe. We showed that $STAR_f$ can range between values of 0.05 and 0.15 in four sites comprising temperate, hemiboreal and boreal biomes in the Czech Republic, Estonia and Finland. Contrary to the previous use of STAR as a clumping correction factor for shoots in conifer forests, stand level STAR quantifies the entire canopy structure, that is, leaf area index, clumping index, and leaf angle distribution. We found that the leaf area index is the strongest driver of $STAR_f$ variation, next to clumping index and leaf angle distribution, and explored the theoretical relationships between LAI, clumping, and $STAR_f$. As a canopy structure quantifier, STAR is closely related to the photon recollision probability.

This direct relationship between $STAR_f$ and the photon recollision probability makes $STAR_f$ a uniquely important forest structural variable whose influence on the forest radiation regime should be studied further. The connection between STAR and photon recollision probability manifests in the relationship observed between $STAR_f$ and surface reflectance in the Landsat 8 OLI bands. However, we found that clumping index, as a stand-alone quantifier of canopy structure, is more closely correlated to surface reflectance than $STAR_f$, with a lower clumping index generally resulting in higher surface reflectance.

We conclude that (1) the method we proposed can quantify both $STAR_f$ and clumping index, (2) the relationships between $STAR_f$ and forest variables indicate that regression using easier-to-measure forest variables is likely of limited accuracy, and (3) both $STAR_f$ and clumping index showed considerable correlation with Landsat 8 OLI surface reflectance, indicating that these variables are valuable for forest radiation regime modeling. We showed that forests throughout Europe exhibit considerable levels of clumping at stand scale, and our results suggest that this spatial heterogeneity is likely best quantified with close-range laser scanning to retain as fine scale information about clumping as possible.

CRediT authorship contribution statement

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

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

D.S., A.H. and M.R. 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. D.W. acknowledges funding from the National Key R&D Program of China (2021YFF0704600) and the National Natural Science Foundation of China (Grant 42101330). We thank Lucie Homolová, Růžena Janoutová, Jussi Juola, Mihkel Kaha, Bijay Karki, Lauri Korhonen, Mait Lang, Petr Lukeš, Titta Majasalmi, Jan Pisek, Andre Purret, Ville Ranta, and the staff of Hyytiälä forestry field station for their assistance in different stages of field work and data processing. We acknowledge the resources provided by the Finnish IT Center for Science (CSC) for enabling our computations.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.agrformet.2023.109564.

References

Åkerblom, M., Raumonen, P., Casella, E., Disney, M.I., Danson, F.M., Gaulton, R., Schofield, L.A., Kaasalainen, M., 2018. Non-intersecting leaf insertion algorithm for tree structure models. Interface Focus 8 (2), 20170045. http://dx.doi.org/10.1098/rsfs.2017.0045.

Amanatides, J., Woo, A., 1987. A fast voxel traversal algorithm for ray tracing. Proc. EuroGraphics 87.

Béland, M., Baldocchi, D., 2020. Is foliage clumping an outcome of resource limitations within forests? Agricult. Forest Meteorol. 295, 108185. http://dx.doi.org/10.1016/j.agrformet.2020.108185.

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.

Chen, J.M., Black, T.A., 1992. Defining leaf area index for non-flat leaves. Plant Cell Environ. 15 (4), 421–429. http://dx.doi.org/10.1111/j.1365-3040.1992.tb00992.x.

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.

Chen, J.M., Mo, G., Pisek, J., Liu, J., Deng, F., Ishizawa, M., Chan, D., 2012. Effects of foliage clumping on the estimation of global terrestrial gross primary productivity. Glob. Biogeochem. Cycles 26 (1), http://dx.doi.org/10.1029/2010gb003996.

Chianucci, F., 2019. An overview of in situ digital canopy photography in forestry. Can. J. Forest Res. 50 (3), 227–242. http://dx.doi.org/10.1139/cjfr-2019-0055.

de Wit, C.T., 1965. Photosynthesis of Leaf Canopies. Technical Report, Pudoc.

Earth Resources Observation And Science (EROS) Center, 2013. Collection-2 Landsat 8-9 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) Level-2 Science Products. U.S. Geological Survey, http://dx.doi.org/10.5066/P9OGBGM6, Accessed 11 December 2021.

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

Forsström, P., Hovi, A., Juola, J., Rautiainen, M., 2023. Links between light availability and spectral properties of forest floor in European forests. Agricult. Forest Meteorol. http://dx.doi.org/10.1016/j.agrformet.2023.109481, (submitted for publication).

Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote Sens. Environ. 70 (1), 29–51. http://dx.doi.org/10.1016/S0034-4257(99)00056-5.

Hadi, Pfeifer, M., Korhonen, L., Wheeler, C., Rautiainen, M., 2017. Forest canopy structure and reflectance in humid tropical Borneo: A physically-based interpretation using spectral invariants. Remote Sens. Environ. 201, 314–330. http://dx.doi.org/10.1016/j.rse.2017.09.018.

He, L., Chen, J., Pisek, J., Schaaf, C., Strahler, A., 2012. Global clumping index map derived from the MODIS BRDF product. Remote Sens. Environ. 119, 118–130. http://dx.doi.org/10.1016/j.rse.2011.12.008.

He, L., Liu, J., Chen, J., Croft, H., Wang, R., Sprintsin, M., Zheng, T., Ryu, Y., Pisek, J., Gonsamo, A., Deng, F., Zhang, Y., 2016. Inter- and intra-annual variations of clumping index derived from the MODIS BRDF product. Int. J. Appl. Earth Obs. Geoinf. 44, 53–60. http://dx.doi.org/10.1016/j.jag.2015.07.007.

Hovi, A., Forsström, P., Ghielmetti, G., Schaepman, M.E., Rautiainen, M., 2020. Empirical validation of photon recollision probability in single crowns of tree seedlings. ISPRS J. Photogramm. Remote Sens. 169, 57–72. http://dx.doi.org/10.1016/j.isprsjprs.2020.08.027.

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.

Hui, Z., Jin, S., Xia, Y., Wang, L., Ziggah, Y.Y., Cheng, P., 2021. Wood and leaf separation from terrestrial LiDAR point clouds based on mode points evolution. ISPRS J. Photogramm. Remote Sens. 178, 219–239. http://dx.doi.org/10.1016/j.isprsjprs.2021.06.012.

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.

Knyazikhin, Y., Martonchik, J.V., Myneni, R.B., Diner, D.J., Running, S.W., 1998. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. J. Geophys. Res.: Atmos. 103 (D24), 32257–32275. http://dx.doi.org/10.1029/98jd02462.

Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 53 (4), 193–203. http://dx.doi.org/10.1016/S0924-2716(98)00009-4.

Lai, Y., Mu, X., Li, W., Zou, J., Bian, Y., Zhou, K., Hu, R., Li, L., Xie, D., Yan, G., 2022. Correcting for the clumping effect in leaf area index calculations using one-dimensional fractal dimension. Remote Sens. Environ. 281, 113259. http://dx.doi.org/10.1016/j.rse.2022.113259.

Lang, A., Xiang, L., 1986. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. Agricult. Forest Meteorol. 37 (3), 229–243. http://dx.doi.org/10.1016/0168-1923(86)90033-x.

Leblanc, S.G., Chen, J.M., Fernandes, R., Deering, D.W., Conley, A., 2005. Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. Agricult. Forest Meteorol. 129 (3–4), 187–207. http://dx.doi.org/10.1016/j.agrformet.2004.09.006.

Li, W., Mu, X., 2021. Using fractal dimension to correct clumping effect in leaf area index measurement by digital cover photography. Agricult. Forest Meteorol. 311, 108695. http://dx.doi.org/10.1016/j.agrformet.2021.108695.

LI-COR, Inc., 2012. LAI-2200 plant canopy analyzer instruction manual. URL https://www.licor.com/documents/6n3conpja6uj9aq1ruyn.

Majasalmi, T., Rautiainen, M., 2020. The impact of tree canopy structure on understory variation in a boreal forest. Forest Ecol. Manag. 466, 118100. http://dx.doi.org/10.1016/j.foreco.2020.118100.

Majasalmi, T., Rautiainen, M., Stenberg, P., 2014. Modeled and measured fPAR in a boreal forest: Validation and application of a new model. Agricult. Forest Meteorol. 189–190, 118–124. http://dx.doi.org/10.1016/j.agrformet.2014.01.015.

Nilson, T., 1999. Inversion of gap frequency data in forest stands. Agricult. Forest Meteorol. 98–99, 437–448. http://dx.doi.org/10.1016/s0168-1923(99)00114-8.

Nobis, M., Hunziker, U., 2005. Automatic thresholding for hemispherical canopy-photographs based on edge detection. Agricult. Forest Meteorol. 128, 243–250. http://dx.doi.org/10.1016/j.agrformet.2004.10.002.

Oker-Blom, P., Smolander, H., 1988. The ratio of shoot silhouette area to total needle area in Scots pine. For. Sci. 34 (4), 894–906.

Parker, G.G., 2020. Tamm review: Leaf area index (LAI) is both a determinant and a consequence of important processes in vegetation canopies. Forest Ecol. Manag. 477, 118496. http://dx.doi.org/10.1016/j.foreco.2020.118496.

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.

Pimont, F., Soma, M., Dupuy, J.L., 2019. Accounting for wood, foliage properties, and laser effective footprint in estimations of leaf area density from multiview-LiDAR data. Remote Sens. 11 (13), 1580. http://dx.doi.org/10.3390/rs11131580.

Pisek, J., 2023. Mean and standard deviation of *Carpinus betulus* leaf angles from measurements in Lanžhot, April 2017 (unpublished data).

Pisek, J., Chen, J., Alikas, K., Deng, F., 2010. Impacts of including forest understory brightness and foliage clumping information from multiangular measurements on leaf area index mapping over North America. J. Geophys. Res. Biogeosciences 115 (3), http://dx.doi.org/10.1029/2009JG001138.

Pisek, J., Diaz-Pines, E., Matteucci, G., Noe, S., Rebmann, C., 2022. On the leaf inclination angle distribution as a plant trait for the most abundant broadleaf tree species in Europe. Agricult. Forest Meteorol. 323, 109030. http://dx.doi.org/10.1016/j.agrformet.2022.109030.

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 Meteorol. 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 Meteorol. 214–215, 2–11. http://dx.doi.org/10.1016/j.agrformet.2015.07.008.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P., 2013. Fast automatic precision tree models from terrestrial laser scanner data. Remote Sens. 5 (2), 491–520. http://dx.doi.org/10.3390/rs5020491.

Yáñez Rausell, L., Malenovský, Z., Rautiainen, M., Clevers, J.G.P.W., Lukeš, P., Hanuš, J., Schaepman, M.E., 2015. Estimation of spruce needle-leaf chlorophyll content based on DART and PARAS canopy reflectance models. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (4), 1534–1544. http://dx.doi.org/10.1109/JSTARS.2015.2400418.

Rautiainen, M., Stenberg, P., 2005. Application of photon recollision probability in coniferous canopy reflectance simulations. Remote Sens. Environ. 96 (1), 98–107. http://dx.doi.org/10.1016/j.rse.2005.02.009.

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 Meteorol. 268, 249–257. http://dx.doi.org/10.1016/j.agrformet.2019.01.033.

Schraik, D., Hovi, A., Rautiainen, M., 2021a. Crown level clumping in Norway spruce from terrestrial laser scanning measurements. Agricult. Forest Meteorol. 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., Varvia, P., Korhonen, L., Rautiainen, M., 2019. Bayesian inversion of a forest reflectance model using Sentinel-2 and landsat 8 satellite images. J. Quant. Spectrosc. Radiat. Transfer 233, 1–12. http://dx.doi.org/10.1016/j.jqsrt.2019.05.013.

Silva, D.D., Boudon, F., Godin, C., Sinoquet, H., 2008. Multiscale framework for modeling and analyzing light interception by trees. Multiscale Model. Simul. 7 (2), 910–933. http://dx.doi.org/10.1137/08071394x.

Smolander, S., Stenberg, P., 2003. A method to account for shoot scale clumping in coniferous canopy reflectance models. Remote Sens. Environ. 88 (4), 363–373. http://dx.doi.org/10.1016/j.rse.2003.06.003.

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.

Spanner, M.A., Pierce, L.L., Peterson, D.L., Running, S.W., 1990. Remote sensing of temperate coniferous forest leaf area index: The influence of canopy closure, understory vegetation and background reflectance. Int. J. Remote Sens. 11 (1), 95–111. http://dx.doi.org/10.1080/01431169008955002.

Stenberg, P., 2007. Simple analytical formula for calculating average photon recollision probability in vegetation canopies. Remote Sens. Environ. 109 (2), 221–224. http://dx.doi.org/10.1016/j.rse.2006.12.014.

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, Vol. 191. Studia Forestalia Suecica, p. 94.

Therezien, M., Palmroth, S., Brady, R., Oren, R., 2007. Estimation of light interception properties of conifer shoots by an improved photographic method and a 3D model of shoot structure. Tree Physiol. 27 (10), 1375–1387. http://dx.doi.org/10.1093/treephys/27.10.1375.

U.S. Geological Survey, 2000. Earth explorer. URL https://earthexplorer.usgs.gov/. (Accessed 18 December 2020).

Wang, D., Hollaus, M., Puttonen, E., Pfeifer, N., 2016. Automatic and self-adaptive stem reconstruction in landslide-affected forests. Remote Sens. 8 (12), http://dx.doi.org/10.3390/rs8120974.

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 Meteorol. 143 (1), 106–122. http://dx.doi.org/10.1016/j.agrformet.2006.12.003.

Wang, D., Liang, X., Mofack, G.I., Martin-Ducup, O., 2021. Individual tree extraction from terrestrial laser scanning data via graph pathing. For. Ecosyst. 8 (1), http://dx.doi.org/10.1186/s40663-021-00340-w.

Wang, D., Momo Takoudjou, S., Casella, E., 2020a. LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. Methods Ecol. Evol. 11 (3), 376–389. http://dx.doi.org/10.1111/2041-210X.13342.

Wang, D., Momo Takoudjou, S., Casella, E., 2020b. LeWoS Github release. https://github.com/dwang520/LeWoS. (Accessed on 05 April 2023).

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.

Weiss, M., Baret, F., Smith, G., Jonckheere, I., Coppin, P., 2004. Review of methods for in situ leaf area index (LAI) determination. Agricult. Forest Meteorol. 121 (1–2), 37–53. http://dx.doi.org/10.1016/j.agrformet.2003.08.001.

Wu, B., Zheng, G., Chen, Y., 2020. An improved convolution neural network-based model for classifying foliage and woody components from terrestrial laser scanning data. Remote Sens. 12 (6), 1010. http://dx.doi.org/10.3390/rs12061010.