

# Estimation of woodland canopy structure with terrestrial LiDAR: expanded methods

John L. Godlee

13th August 2021

## **1 1 Introduction**

This chapter provides expanded field and analytical methods for the study of tree canopy structure in southern African woodlands, presented in brief in Chapter 5. The study aimed to understand the effects of tree species diversity and stand structure on tree canopy structural complexity, using terrestrial LiDAR. Firstly, I provide technical details on the field setup for the terrestrial LiDAR and the hemispherical photography used to validate terrestrial LiDAR canopy closure estimates. Secondly, I describe the processing chain used to extract canopy complexity metrics from the terrestrial LiDAR point clouds. Thirdly, I describe in further detail the behaviour and suitability of the different canopy complexity metrics and stand structural metrics used in the study.

## **11 2 Terrestrial LIDAR field setup**

Within each 1 ha (100x100 m) square plot, nine 10 m diameter circular subplots were laid out in a grid, with 35 m between subplot centre points (Figure 1). These subplots constitute the basic sampling unit of the study. Within each subplot, a Leica HDS6100 phase-shift Terrestrial Laser Scanner (TLS) was used to capture woodland canopy structure. The number and position of scan locations within a subplot was determined by the arrangement and density of canopy material in the subplot. Scan positions were arranged to minimise shadows within the canopy, and to maximise canopy penetration. Between one and five scans were recorded per subplot, across all plots. Further information on the field setup of the TLS is presented in Table 1.

Five Leica 6" (15.24 cm) diameter planar tilt-and-turn cross-pattern reflective targets were used in each subplot to align scans (Figure 2). The five targets were located roughly in a quincunx pattern, with one target at the subplot centre and the remaining four targets arranged in a cross pattern around the edges of the subplot, ensuring that all scans could see all five targets. To facilitate alignment of scans among subplots, the location of each target in real space was recorded using a Leica VIVA GS10 GNSS (Global Navigation Satellite Systems) unit (Figure 3). The GNSS was set up in a Post-Processing Kinematic (PPK) configuration with a base-station located ~100 m from the edge of each 1 ha plot with an unobstructed view of the sky hemisphere where possible. The location of each target was measured for at least 4 minutes to minimise measurement error (Figure 4).

## **30 3 Hemispherical photography field setup**

In order to validate TLS canopy closure estimates, at the centre of each subplot a single photograph was taken with a full-frame DSLR camera, equipped with a circular fisheye lens. Further information on the hemispherical photography setup is presented in Table 2.

The fisheye lens had an equisolid (equal area) projection, with a projection function given by:

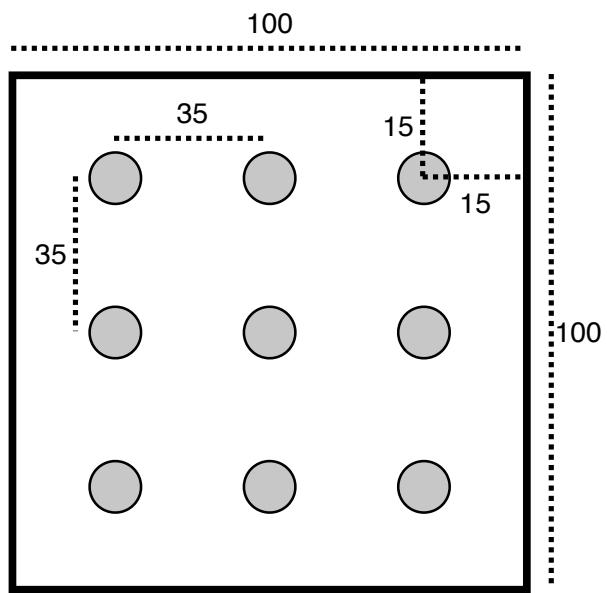


Figure 1: The layout of 10 m diameter subplots within each 1 ha plot. Each subplot is situated inside a 15 m buffer from the plot edge, with 35 m between subplot centres. Subplots are arranged in a 3x3 grid. All distances are in metres.

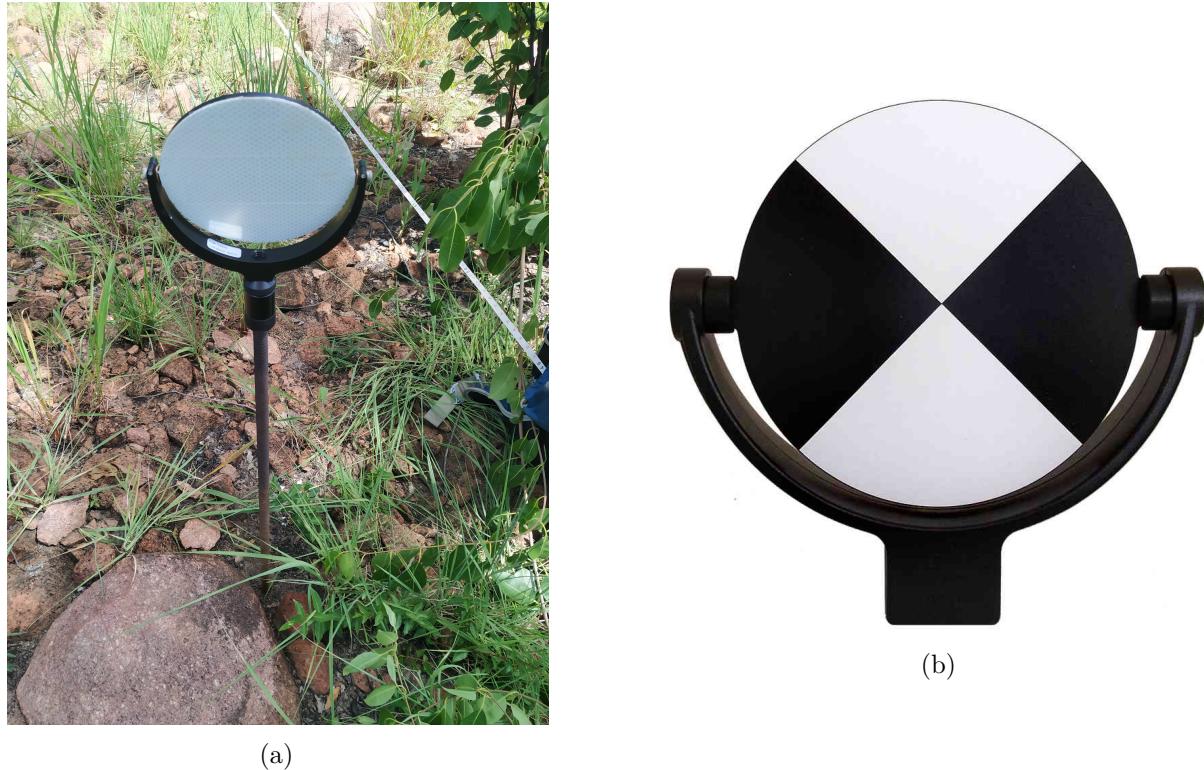


Figure 2: Example of a Leica 6" diameter reflective target, (a) in situ mounted on a length of threaded bar, and (b) showing the cross pattern face of the target.



Figure 3: A Leica VIVA GS10 GNSS unit in the field, showing the antenna atop an aluminium pole, attached to the base station on the ground, and the rover terminal in the hand of a research assistant.

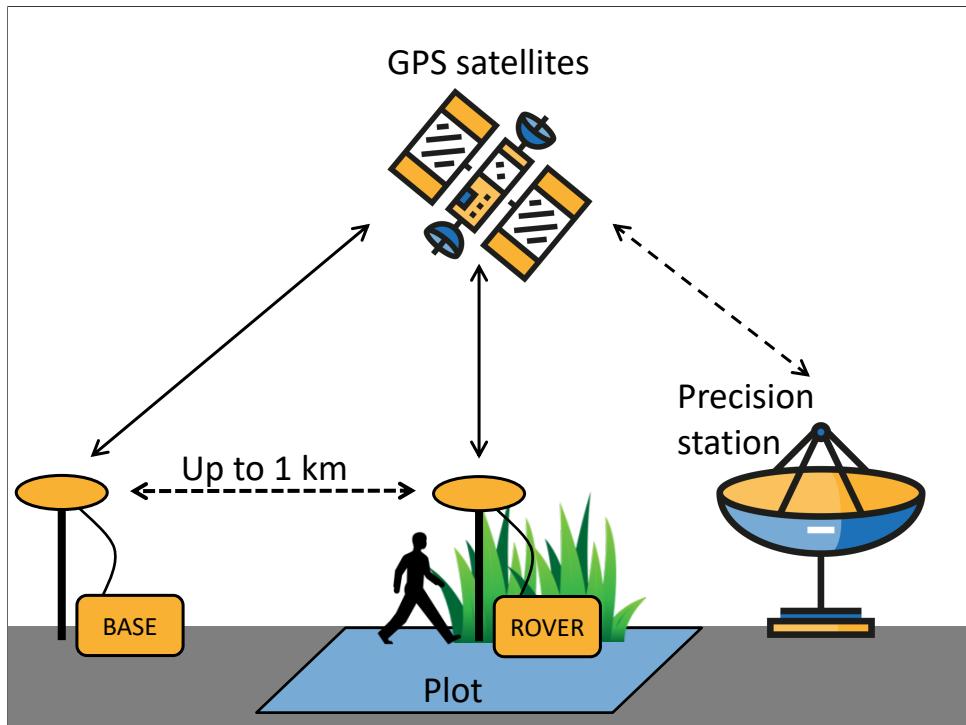


Figure 4: Schematic diagram of the GNSS PPK configuration used to precisely locate targets in real space. The base station is located in an area with a full unobstructed view of the sky hemisphere, up to ~1 km from the plot, and is left in the same location for the duration of the data collection, recording its location once per second. The rover is moved around inside the plot to record the location of each target, for >4 minutes at each target. The rover and the base station both utilise GPS and GLONASS satellites to record their position. After data collection, a two stage validation technique is used to improve the precision of the recorded positions, firstly using the base station, and secondly using the TrimbleRTX service which utilises highly precise distributed regional stations.

Table 1: Description of scan settings used for each scan.

Setting	Value
TLS model	Leica HDS6100
Wavelength	650-690 nm
Spot size at exit	3 mm
Beam divergence	0.22 mrad
Range	79 m @90%; 50 m @18% albedo
Azimuth range	0-360°
Zenith range	0-155°
Increment	0.018°
Point spacing over 25 m	7.9 mm
Pixels per line	20,000
Lines	10,000
Compressed file size	~800 MB
Duration of scan	6 minutes 44 seconds

$$R = 2f \sin\left(\frac{\theta}{2}\right) \quad (1)$$

35 Where  $R$  is the radial position of a point on the image,  $f$  is the focal length of the lens, and  $\theta$  is  
 36 the angle in radians of incident light on the lens. Equisolid lenses are preferred for hemispherical  
 37 photography because they maintain an equal area for each pixel, i.e. a pixel projected through  
 38 the lens has the same solid angle irrespective of the incident light angle, meaning that canopy  
 39 closure estimations are not biased towards any part of the sky hemisphere (Herbert, 1987).

40 Photographs were taken facing directly to zenith using a camera-mounted spirit level, with  
 41 the top of the camera body facing magnetic north, at a height of 1.3 m or above understorey  
 42 vegetation, whichever was higher. Photographs were captured under uniform light conditions as  
 43 much as possible, either under overcast skies or early in the day before direct sunlight could be  
 44 seen on the photograph, to minimise lens flare, which can preclude accurate differentiation of  
 45 plant material and sky, and ‘blooming’, a phenomenon where light ‘bleeds’ into dark areas of  
 46 the image in highly contrasting light conditions (Frazer et al., 2001).

47 ImageJ (Fiji version 2.1.0/1.53c) was used to binarise hemispherical photographs, to separate  
 48 plant material from sky (Schneider et al., 2012). Images were binarised using the Huang algorithm  
 49 (Huang & Wang, 1995) using only the blue channel of the image, under the assumption that  
 50 plant material reflects little blue light, while the sky reflects much more (Brusa & Bunker,  
 51 2014). Images were saved as PNG files at the original pixel resolution, with a circular image of  
 52 4016x4016 pixels.

## 53 4 Terrestrial LiDAR processing

### 54 4.1 Scan alignment and registration

55 Point clouds within a subplot were aligned using the reflective targets as anchor points. Point  
 56 cloud alignment was conducted in Leica Cyclone (version 9.1) (Leica Camera AG, 2009).  
 57 Reflective targets were manually located within each point cloud, then the precise centre of  
 58 each target was identified automatically by Cyclone. Anchor points were discarded if they had

Table 2: Description of camera settings used for hemispherical photographs. Note that shutter speed and ISO are deliberately variable within sensible thresholds to allow adjustments for ambient light conditions.

Setting	Value
Camera model	Nikon D750
Lens model	Sigma 8 mm f/3.5 EX DG Circular Fisheye
Pixel pitch	5.95 $\mu\text{m}$
Sensor resolution	24.3 MP
Shutter speed	>1/60s
Aperture	5-7
ISO	100-200
Exposure compensation	-0.7 (Brusa & Bunker, 2014)
Focus	$\infty$ (Hu & Zhu, 2009; Frazer et al., 2001)
Image size	Large Fine JPEG - circular image 4016x4016 px
Orientation	Landscape

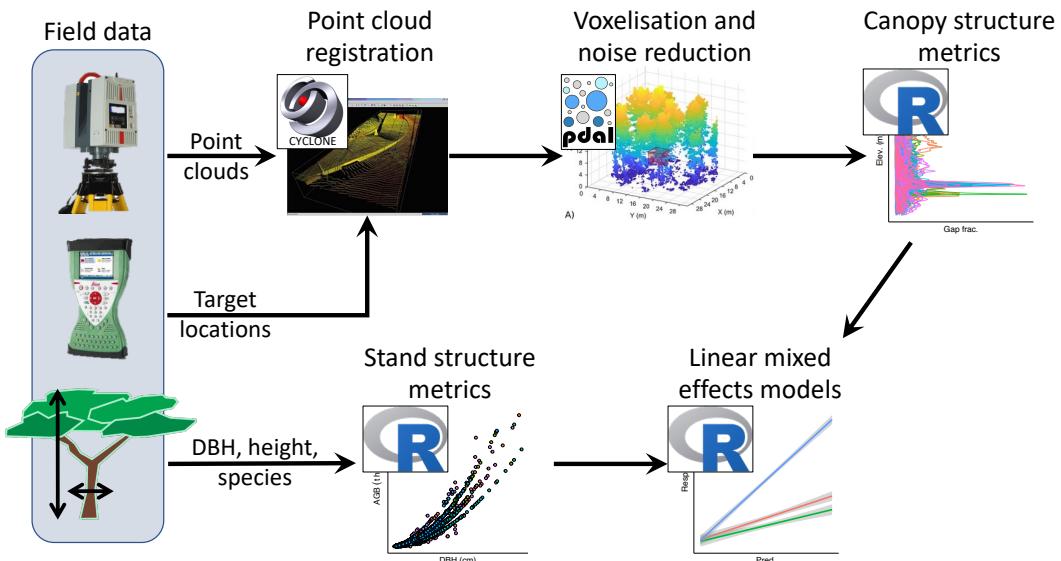


Figure 5: Schematic diagram summarising the data processing and analysis workflow for the TLS data. Processing steps are labelled according to the principal software used during that step.

Table 3: The five closes regional base stations to each site, used by TrimbleRTX to refine GNSS measurements

Site	Receiver	Distance (km)
Bicuar	JAVAD TRE 3	863
	JAVAD TRE 3 DELTA	1448
	TRIMBLE NETRS	1630
	TRIMBLE NETRS	1648
	JAVAD TRE G3TH DELTA	1796
	SEPT POLARX5	686
Mtarure	TRIMBLE ALLOY	795
	LEICA GRX1200GGPRO	903
	JAVAD TRE 3 DELTA	1119
	JAVAD TRE 3 DELTA	1315

59 a location uncertainty of >3 cm. After alignment, subplot point clouds were exported from  
 60 Cyclone as PTX files for further processing.

61 GNSS target locations were used to register point clouds in real space. The TrimbleRTX GNSS  
 62 post-processing service was used to improve the precision of target locations recorded with GNSS,  
 63 using distributed regional stations to validate the rover and base station GNSS measurements  
 64 (Table 3) (Chen et al., 2011). Following point cloud registration, subplot point clouds were  
 65 combined to a plot-level point cloud.

66 PTX files were converted to compressed LAZ files using PDAL (PDAL Contributors, 2018), to  
 67 reduce file size and speed up further processing. Code 1 contains the code used to transform  
 68 PTX to LAZ.

```

1 # Get file name without extension
2 noext=${1%.ptx}
3
4 # Find the PTX scan array dimension header material
5 lines=$(grep -E -n '^.{1,10}$' $1 |
6   cut -f1 -d: |
7   awk 'NR%2!=0' |
8   tr '\n' ' ' |
9   sed 's/^([0-9])\s//g')
10
11 # Split PTX file into individual scans
12 csplit -f "$noext" -b "%d.ptx" $1 $lines
13
14 $ Find split files
15 ptxsplit=$(find . -type f -regex ".*/${noext}_[0-9].ptx")
16
17 # For each file:
18 for j in ${ptxsplit}; do
19   jnoext="${j%.ptx}"
20   matrix=$(head -n 10 $j | tail -4 | sed -r 's/0\s+?$/0.0/g' | dos2unix)
21   pdal pipeline ptx_laz.json --readers.text.filename=$j \
22     --filters.transformation.matrix="${matrix}" \
23     --writers.las.filename=${jnoext}.laz
24 done
25
26 # List LAZ files
27 lazsplit=$(find . -type f -regex ".*/${noext}_[0-9].laz")
28
29 # Merge LAZ files
30 pdal merge ${lazsplit} ${noext}.laz

```

Code 1: The processing chain used to convert Leica Cyclone PTX files to LAZ files, using PDAL, POSIX shell scripting, and common UNIX utilities. The `ptx_laz.json` JSON pipeline is shown in Code 2.

```

1 [
2   {
3     "type" : "readers.text",
4     "filename" : "input.txt",
5     "header" : "X Y Z I",
6     "skip" : 10
7   },
8   {
9     "type" : "filters.transformation",
10    "matrix" : "0 -1 0 1 1 0 0 2 0 0 1 3 0 0 0 1"
11  },
12  {
13    "type" : "writers.las",
14    "compression" : "true",
15    "minor_version" : "2",
16    "dataformat_id" : "0",
17    "forward" : "all",
18    "filename" : "output.laz"
19  }
20 ]

```

Code 2: The JSON pipeline used in Code 1 to convert PTX files to LAZ files, and applying a rotation matrix.

71 **4.2 Voxelisation**

72 Point clouds were voxelised to different voxel sizes depending on the application of the data. 5  
73  $\text{cm}^3$  cubic voxels were used for subplot height profile estimation, while  $50 \text{ cm}^3$  voxels were used  
74 for whole plot canopy rugosity. Variation in voxel size reflects the spatial scale of each analysis,  
75 and is bounded by the beam divergence of the scanner over longer distances (Grau et al., 2017).  
76 Choosing voxels that are too small can result in pock-marked representations of surfaces that  
77 are especially problematic when calculating larger scale canopy structure metrics such as canopy  
78 top roughness, while voxels that are too large can result in an over-estimation of plant volume  
79 when estimating canopy foliage density at the subplot scale, especially when foliage is clumped  
80 (Seidel et al., 2012; Cifuentes et al., 2014). Voxels were classified as ‘filled’ if they intersected  
81 one or more points.

82 **4.3 Noise reduction**

83 Outlier detection and noise reduction of point clouds was conducted in PDAL, using the  
84 “statistical method” (sensu Rusu et al. 2008) of `filters.outlier`, with  $k = 8$  (mean number  
85 of neighbours), and  $m = 1.96$  (outlier distance threshold multiplier, here approximating a 95%  
86 confidence interval):

$$\begin{aligned}\bar{\mu} &= \frac{1}{N} \sum_{i=1}^N \mu_i \\ \sigma &= \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\mu_i - \bar{\mu})^2} \\ t &= \mu + m\sigma \\ \text{with } outlier_i &= \begin{cases} \text{true,} & \text{if } \mu_i \geq t \\ \text{false,} & \text{otherwise} \end{cases}\end{aligned}\tag{2}$$

87 Where  $\mu_i$  is the mean distance from point  $i$  to all  $k$  nearest neighbour points,  $N$  is the number  
88 of points in the scene,  $\bar{\mu}$  is the mean distance to nearest neighbour points,  $\sigma$  is the standard  
89 deviation of these mean distances,  $t$  is the threshold distance used to define an outlier and  
90  $outlier_i$  is the condition of a point in the scene being identified as an outlier.

91 **4.4 Foliage density profiles**

92 To calculate subplot foliage density profiles, the  $5 \text{ cm}^3$  voxelised point cloud was first cropped to  
93 a 10 m diameter cylinder of infinite height. Ground points were identified using `filters.pmf`  
94 (Progressive Morphological Filter - PMF) in PDAL (sensu Zhang et al. 2003), and the height  
95 above ground of all points was calculated using `filters.hag_nn` (Nearest Neighbour) in PDAL.  
96 Points below ground level and above the 99.9th percentile of height were excluded from further  
97 analyses. Height profile points were exported to XYZ coordinates then imported into R for  
98 further processing.

99 In R, foliage density was calculated in 5 cm layers as the proportion of filled  $5 \text{ cm}^3$  voxels. A  
100 loess model with a span of 0.1 was fitted to the foliage density values in each layer to estimate  
101 the foliage density profile (Figure 7). The foliage density profile was further filtered to only tree  
102 canopy material, by discarding all points below the first local minimum in the foliage density  
103 profile above 1.3 m, using a rolling window of 50 cm.

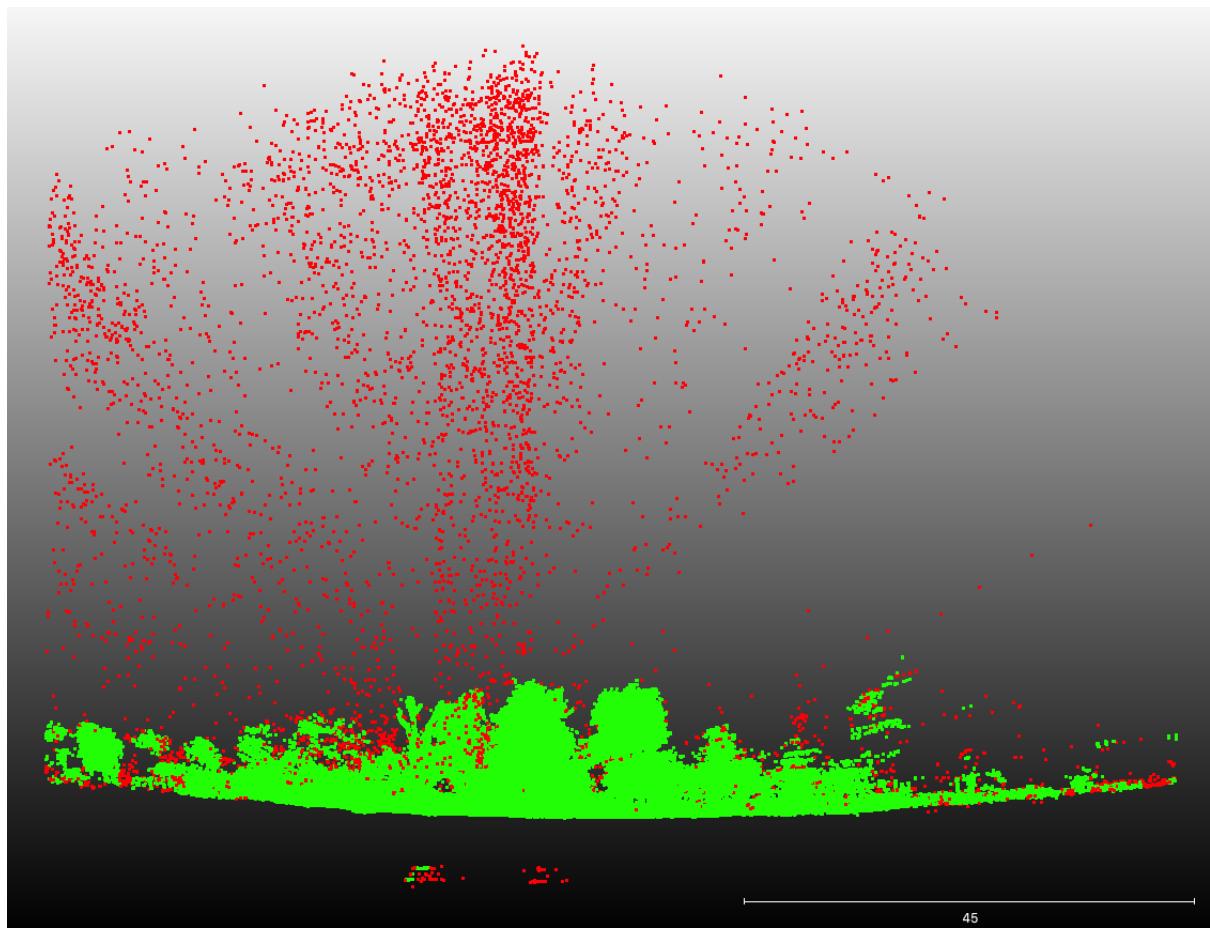


Figure 6: A 2 m deep cross section of a subplot point cloud showing the efficacy of the noise reduction and voxelisation process. Red points are excluded by the process, while green points are preserved for further analysis.

<sup>104</sup> Multiple statistics were extracted from the foliage density profile for use in statistical analyses.  
<sup>105</sup> Total canopy foliage density was calculated as the area under the curve of the canopy foliage  
<sup>106</sup> density profile, using trapezoid estimation. The Effective Number of Layers (ENL) in the  
<sup>107</sup> foliage density profile was used to estimate canopy structural complexity, using the true-numbers  
<sup>108</sup> equivalent Shannon entropy of foliage density among 50 cm layers (*sensu* Ehbrecht et al. 2016):

$$\text{ENL} = \exp\left(-\sum_{i=1}^N p_i \ln p_i\right) \quad (3)$$

<sup>109</sup> Where  $N$  is the number of 50 cm bins in the height profile, and  $p_i$  is the proportion of filled  
<sup>110</sup> voxels in layer  $i$  (foliage density). While Ehbrecht et al. (2016) used 1 m layers, their study was  
<sup>111</sup> conducted in temperate deciduous forest where the maximum height of the sampled forest stands  
<sup>112</sup> was 40 m, whereas the maximum canopy height in this study was only 22 m. Both Ehbrecht  
<sup>113</sup> et al. (2016) and Montes et al. (2004) assert that layer thickness is largely arbitrary, but should  
<sup>114</sup> be determined with respect to the variability within the canopy, thus in the sparse and highly  
<sup>115</sup> variable savanna tree canopies measured in this study, narrower layers were chosen.

<sup>116</sup> To describe the uniformity of the foliage density distribution through the canopy, a linear model  
<sup>117</sup> of foliage density with height was fitted. Under a completely even distribution of foliage material  
<sup>118</sup> through the canopy, the residuals of the linear model tend to zero, while clumping causes  
<sup>119</sup> deviations from this uniform distribution and increases the sum of residuals.

<sup>120</sup> Maximum canopy height has been used in other studies to describe canopy structural complexity  
<sup>121</sup> (Scheuermann et al., 2018). At the small spatial scale of the subplots used in this study however,  
<sup>122</sup> there proved to be too much stochastic variation in canopy height among subplots due to  
<sup>123</sup> the distribution of individual trees to make this statistic informative as a measure of canopy  
<sup>124</sup> complexity. Canopy height was calculated later at the plot level, however. Similarly, previous  
<sup>125</sup> studies have used the number of local maxima in the foliage density profile to estimate canopy  
<sup>126</sup> structural complexity (Wilkes et al., 2016). However, this metric covaries with ENL and total  
<sup>127</sup> foliage density, while ENL is uncorrelated with foliage density (Ehbrecht et al., 2016).

## <sup>128</sup> 4.5 Canopy closure

<sup>129</sup> Subplot canopy closure, i.e. the proportion of the sky hemisphere occluded by plant material,  
<sup>130</sup> a.k.a. gap fraction or site factor (Jennings, 1999), was measured by simulating a hemispherical  
<sup>131</sup> image at the centre of the subplot using the point cloud data from all scans per subplot. The  
<sup>132</sup> point cloud was first cropped to a 20 m diameter cylinder around the subplot centre using  
<sup>133</sup> PDAL. Points below 1.3 m and within a 50 cm sphere around the subplot centre at 1.3 m height  
<sup>134</sup> were discarded, to prevent the simulated hemispherical image being occluded by understorey  
<sup>135</sup> vegetation. POV-Ray was used to simulate the hemispherical image using ray-tracing (Persistence  
<sup>136</sup> of Vision Pty. Ltd., 2004). Filled voxels were represented in POV-Ray as non-reflective black  
<sup>137</sup> cubes filling the 5 cm<sup>3</sup> voxel volume, with a white uniform sky box and no light source. POV-Ray  
<sup>138</sup> produced an image with identical qualities to that of the real hemispherical photograph, using a  
<sup>139</sup> fisheye lens with an equisolid projection and a view angle of 180°, located at the subplot centre  
<sup>140</sup> at 1.3 m above the ground, with the top of the camera facing magnetic north and the camera  
<sup>141</sup> facing directly to zenith, producing a circular image of 4016x4016 pixels.

<sup>142</sup> Hemiphot (ter Steege, 2018) was used to estimate closure from both the hemispherical photo-  
<sup>143</sup> graphs and the TLS POV-Ray simulation. Images were cropped to a circle of 60° zenith angle.  
<sup>144</sup> It can be supposed that below 60°, under most woodland canopies, variation in tree canopy  
<sup>145</sup> density does little to affect sunlight penetration, due to the greater depth of canopy at these  
<sup>146</sup> angles (Jupp et al., 2008). Hemiphot calculates canopy closure in 90 evenly sized concentric  
<sup>147</sup> rings. To obtain the total closure of a circular image:

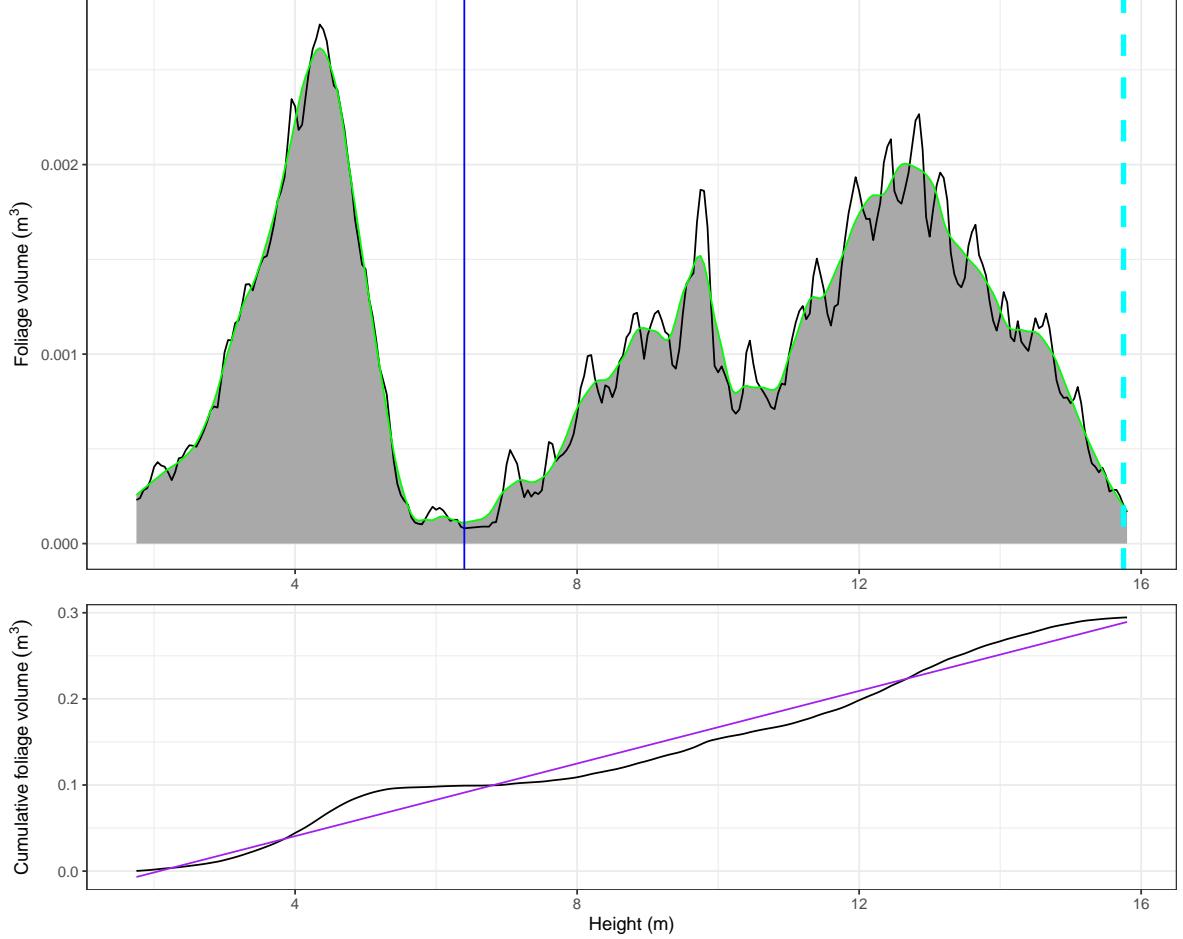


Figure 7: Subplot foliage volume height profile (top) and cumulative foliage volume profile (bottom) for a subplot in Bicuar National Park, Angola, to illustrate some of the canopy structure metrics extracted from each height profile. Starting with the top panel: the blue solid line represents the first local minimum above 1.3 m, used to define the base of the tree canopy. The dashed cyan line shows the 99.9th percentile of canopy height, used here as a measure of canopy top height across the subplot and in plot-level canopy surface modelling. The black trace shows the foliage density height profile, and the green trace shows the loess model fitted to the data, with the area under the canopy shaded grey. The bottom panel: the black trace shows the cumulative foliage volume through the canopy, taken from the loess fit in the top panel. The purple line shows the line of best fit of a linear model through this data. Not illustrated is the Effective Number of Layers (ENL) metric.

$$C_\alpha = 1 - G_{\text{tot}} = \sum_{\alpha=0.5}^{\alpha=89.5} (G_\alpha A_\alpha / A_{\text{tot}}) \quad (4)$$

148 Where  $G_\alpha$  is the fraction of unfilled pixels in ring  $\alpha$ ,  $A_\alpha$  is the sky area of the ring segment, and  
149  $A_{\text{tot}}$  is the total sky area of the hemisphere.

150 Canopy closure estimates from the TLS were validated using estimates from hemispherical  
151 photography. A Pearson's correlation analysis showed that both methods were highly correlated  
152 ( $r(195)=0.89$ ,  $p<0.001$ ). TLS estimates of closure were almost exclusively higher than hemispherical  
153 photography estimates, except in a few subplots with particularly low canopy closure. At  
154 higher canopy closure the over-estimation of canopy closure by TLS was larger (Figure 8). This  
155 finding is in agreement with previous studies which have found that the magnitude of TLS  
156 canopy closure over-estimation depends on gap size distribution, where a site with greater canopy  
157 cover and a gap fraction dominated by small within crown gaps will have a larger over-estimate  
158 than a more open site with a gap fraction dominated by large between crown gaps (Seidel et al.,  
159 2012). A linear mixed model which accounted for the nested sampling of subplots within plots  
160 was used to identify if sites differed significantly in their relationship between hemispherical  
161 photography and TLS estimates of canopy closure. There was no significant difference in model  
162 fixed effect slope between plots in Bicuar National Park, Angola, and those in Mtarure, Tanzania  
163 ( $\beta(173)=0.13\pm0.0098$ ,  $p=0.18$ ).

#### 164 4.6 Whole plot canopy metrics

165 The canopy height of each 1 ha plot was estimated using unified point clouds from all subplots.  
166 The unified point cloud was voxelised to  $10 \text{ cm}^3$ , and the 99th percentile of height from each  $10 \text{ cm}^2$   
167 column was taken as the canopy height. Maximum height was not used as this occasionally  
168 constituted a severe outlier which skewed further canopy surface model smoothing. The point  
169 cloud was then cropped to the plot boundaries, located using PPK GNSS similar to the TLS  
170 targets. A pit-filling algorithm described in Khosravipour et al. (2014) was used to smooth  
171 the canopy surface model, at a resolution of 50 cm, by removing gaps within trees caused by  
172 incomplete penetration of the LiDAR beam (Figure 10).

173 Mean canopy height across the plot and the coefficient of variation of canopy height were  
174 extracted from the canopy surface model for use in statistical analyses. The coefficient of  
175 variation of canopy height describes canopy structural diversity measured by the heterogeneity  
176 of the canopy surface. Other studies in closed canopy temperate and boreal forests have used  
177 metrics similar to the Topographic Roughness Index to measure canopy surface heterogeneity, by  
178 comparing canopy height to that of neighbouring pixels in the canopy height model (Weligepolage  
179 et al., 2012; Herrero-Huerta et al., 2020). In this study however, the sparse nature of the tree  
180 canopies meant that these metrics were overly influenced by canopy density and the edges  
181 of individual tree canopies. Canopy rugosity ( $R_c$ ) was also calculated to describe structural  
182 complexity across the entire canopy profile, rather than just the canopy surface, sensu Hardiman  
183 et al. (2011).  $R_c$  first calculates the standard deviation of foliage density in  $50 \text{ cm}^2$  columns  
184 across the plot ( $\sigma G_z$ ), then calculates the standard deviation of those standard deviations:

$$R_c = \sigma(\sigma G_z)_x \quad (5)$$

185 Where  $G_z$  is the vertical height axis  $z$ ,  $x$  is the horizontal axis, and  $\sigma$  is the standard deviation.  
186 Finally, plot-level canopy closure was calculated as the mean of subplot TLS canopy closure  
187 estimates.

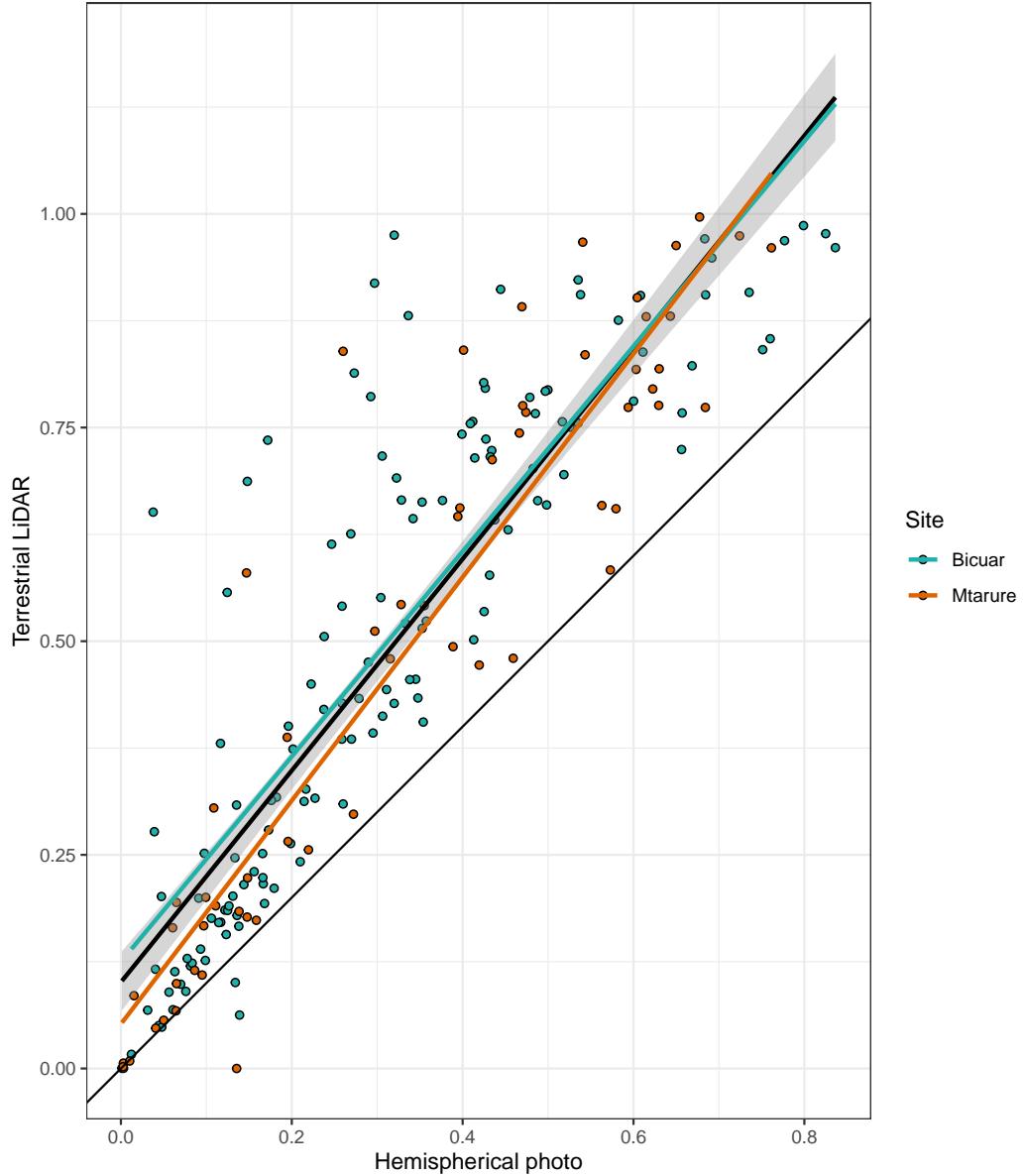
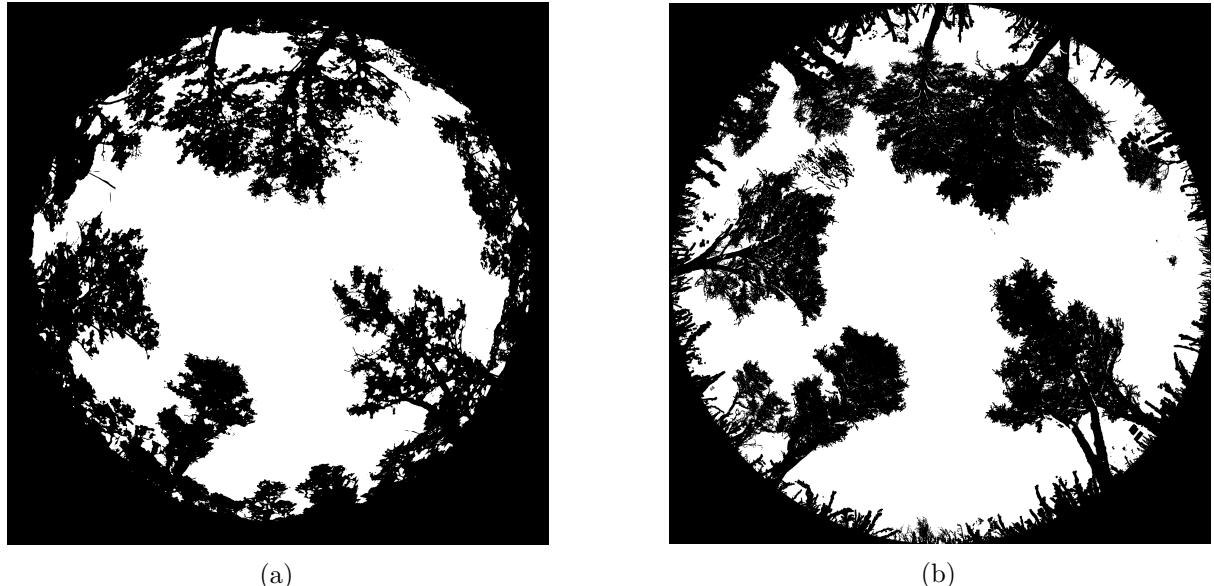


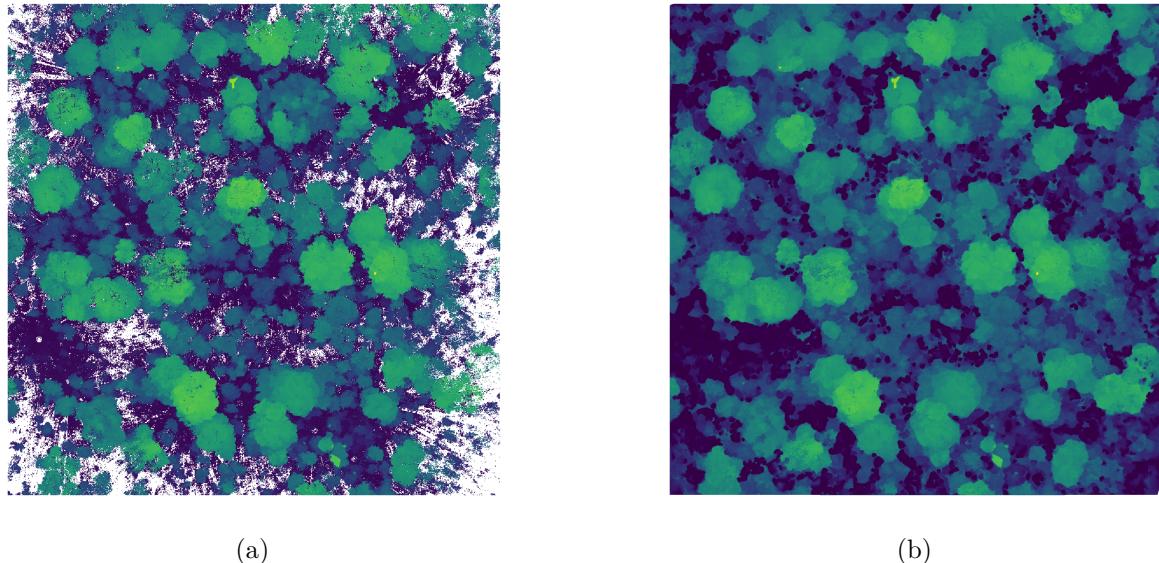
Figure 8: Comparison of canopy closure estimation from TLS and hemispherical photography. The thick black line of best fit is a linear model of all points  $\pm 1$  standard error, while the coloured lines are site specific linear models. The thin black line shows the 1:1 fit.



(a)

(b)

Figure 9: Comparison of hemispherical images for a subplot in Bicuar National Park, Angola. (a) A hemispherical photograph, and (b) a multi-scan point cloud modelled as cubic voxels with POV-Ray. The hemispherical photograph (left) shows some blooming, especially in the tree on the bottom right of the image, where light is seen ‘bleeding’ through the darker canopy material, causing an under-estimation in canopy closure. Note also that while there are minor variations in image height between TLS and hemispherical photos, this does not affect canopy closure estimates, as images were first cropped to 60°.



(a)

(b)

Figure 10: Top-down view of a 1 ha plot in Bicuar National Park. (a) The point cloud after voxelisation, noise reduction, and taking the 99th percentile of stem height in each 5 cm vertical bin. (b) The same point cloud after pit filling to generate a smooth canopy height profile. Points are coloured according to point height from the ground.

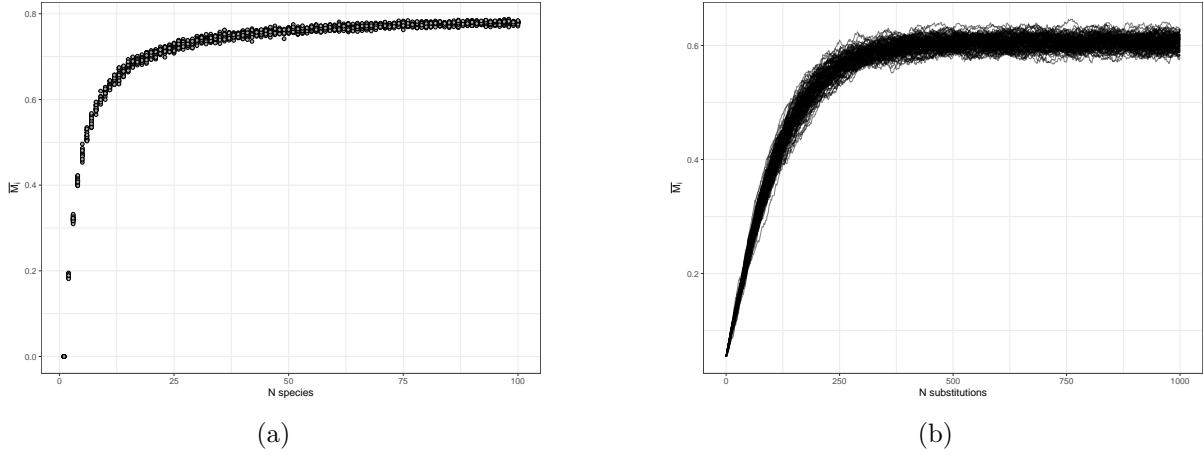


Figure 11: The behaviour of the spatial mingling index ( $M_i$ ) with increasing number of species (a), and increasing spatial mixing of species (b). The left panel was generated by randomly assigning different numbers of species, in equal proportions, to an evenly spaced grid of individuals. 20 replicates were conducted for each number of species. The right panel was generated by randomly swapping pairs of individuals in a plot with 9 species arranged in mono-specific square blocks in an evenly spaced grid. Each line shows a single replicate, where individuals were swapped in an additive fashion, with 100 total.

## 188 5 Stand structure metrics

### 189 5.1 Spatial mingling of species

190 The spatial mingling index ( $M_i$ ) is a spatially explicit estimate of the degree to which species  
 191 are spatially mixed within a plot. Here,  $M$  was calculated at the plot level as the mean of  $M_i$   
 192 according to von Gadow & Hui (2002), with the adjustment for potential neighbourhood species  
 193 pool suggested by Hui et al. (2011):

$$M = \bar{M}_i$$

$$M_i = \frac{S_i}{n_{\max}} \frac{1}{k} \sum_{j=1}^k v_j \quad (6)$$

with  $v_j = \begin{cases} 0, & \text{neighbour } j \text{ same species as reference } i \\ 1, & \text{otherwise} \end{cases}$

194 Where  $k$  is the number of nearest neighbours considered for each reference tree,  $S_i$  is the number  
 195 of species found among the  $k$  nearest neighbours of tree  $i$ ,  $n_{\max}$  is the potential number of  
 196 species in the neighbourhood, i.e.  $k + 1$ , and  $N$  is the total number of trees in the plot. The  
 197 conventional value of  $k = 4$  was used here (von Gadow & Hui, 2002; Hui & Albert, 2004; Hui  
 198 et al., 2007). The value of  $M_i$  increases with greater mixing of species, and all else being equal  
 199 will increase with number of species within the plot (Figure 11).

### 200 5.2 Spatial clustering of stems

201 The winkelmann ( $W$ ) was calculated to estimate the degree of spatial uniformity in stem spatial  
 202 distribution. Here,  $W$  was calculated at the plot level as the mean of  $W_i$ ) according to von  
 203 Gadow & Hui (2002):

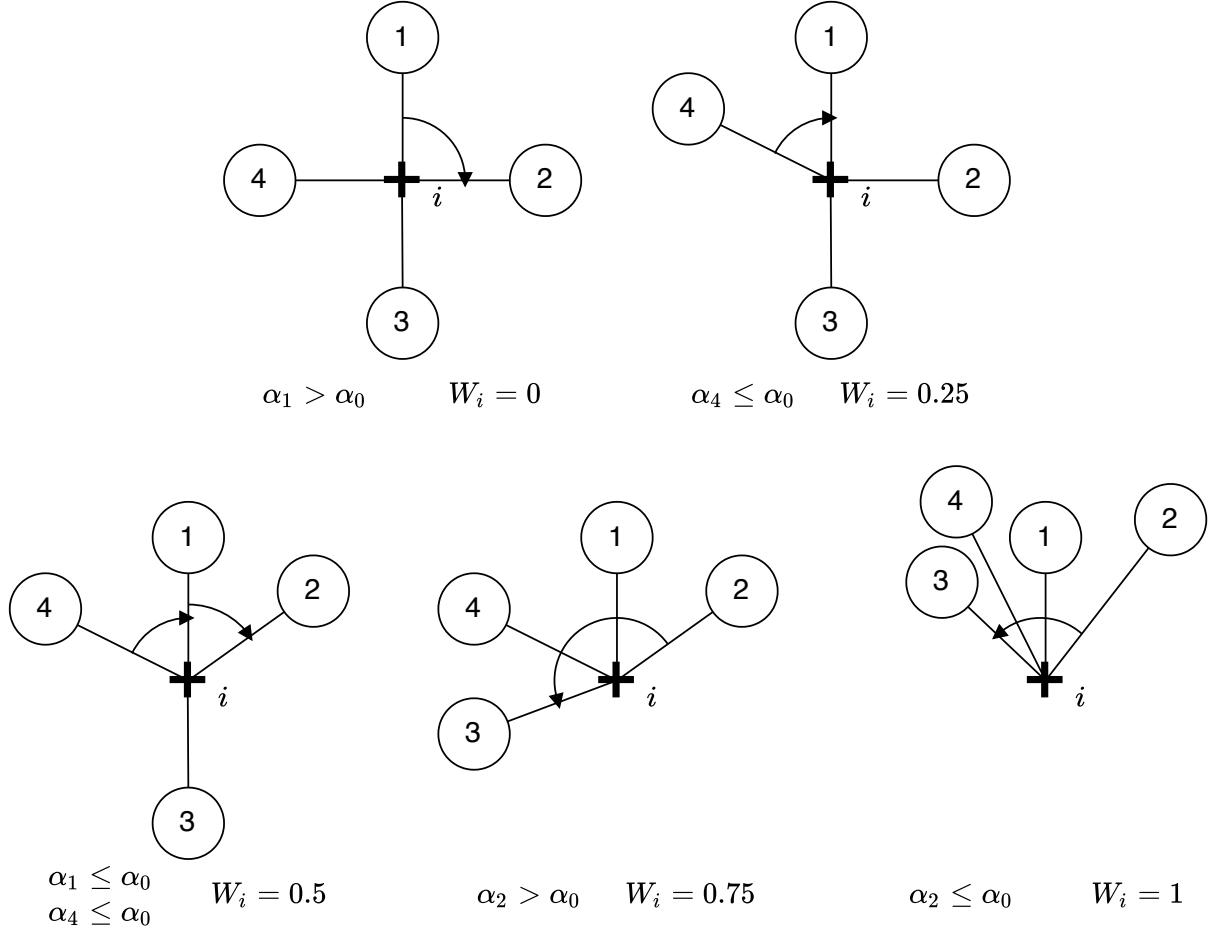


Figure 12: Possible values of  $W_i$  at a sample point  $i$ , denoted by a cross. Neighbours are represented as circles numbered sequentially from 1 to 4, where  $k = 4$ . The angles of arrows in each example are given below, along with the winkelmass for that example.

$$\begin{aligned}
 W &= \overline{W}_i \\
 W_i &= \frac{1}{k} \sum_{j=1}^k v_j \\
 \text{with } v_j &= \begin{cases} 0, & \alpha_j \leq \alpha_0 \\ 1, & \text{otherwise} \end{cases}
 \end{aligned} \tag{7}$$

204 Where  $k$  is the number of neighbours considered,  $\alpha_j$  is the angle between consecutive neighbours  
 205 and  $\alpha_0$  is the critical angle, where  $\alpha_0 = 72^\circ$  (Hui & Gadow, 2002). Figure 12 demonstrates  
 206 how the value of  $W_i$  varies according to spatial distribution of neighbours. The value of  
 207 the winkelmass increases with increasing spatial clumping (decreasing spatial regularity) of  
 208 individuals (Figure 13), and in a plot with random tree distribution will increase as more  
 209 neighbours are considered (Figure 14).

### 210 5.3 Subplot canopy crowding

211 An adapted version of the Iterative Hegyi Index ( $H_i$ ) was used to estimate tree spatial structure  
 212 in subplots (Hegyi, 1974). The adapted formula used here allows the index to be based on a point

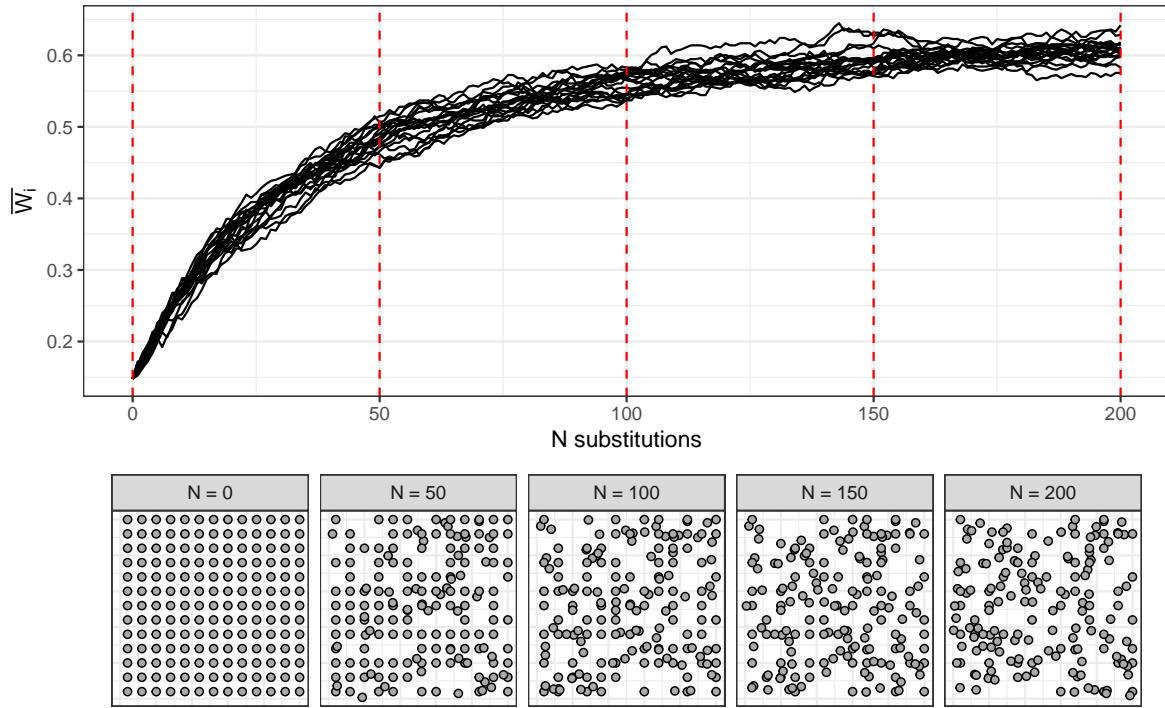


Figure 13: Variation in winkelmass with increasing spatial irregularity of individuals. The top panel shows variation of winkelmass in 20 plots as individuals are sequentially moved to a random location within the plot. Red dotted lines correspond to the panels below which show the spatial distribution of individuals after a given number of random individual movements.

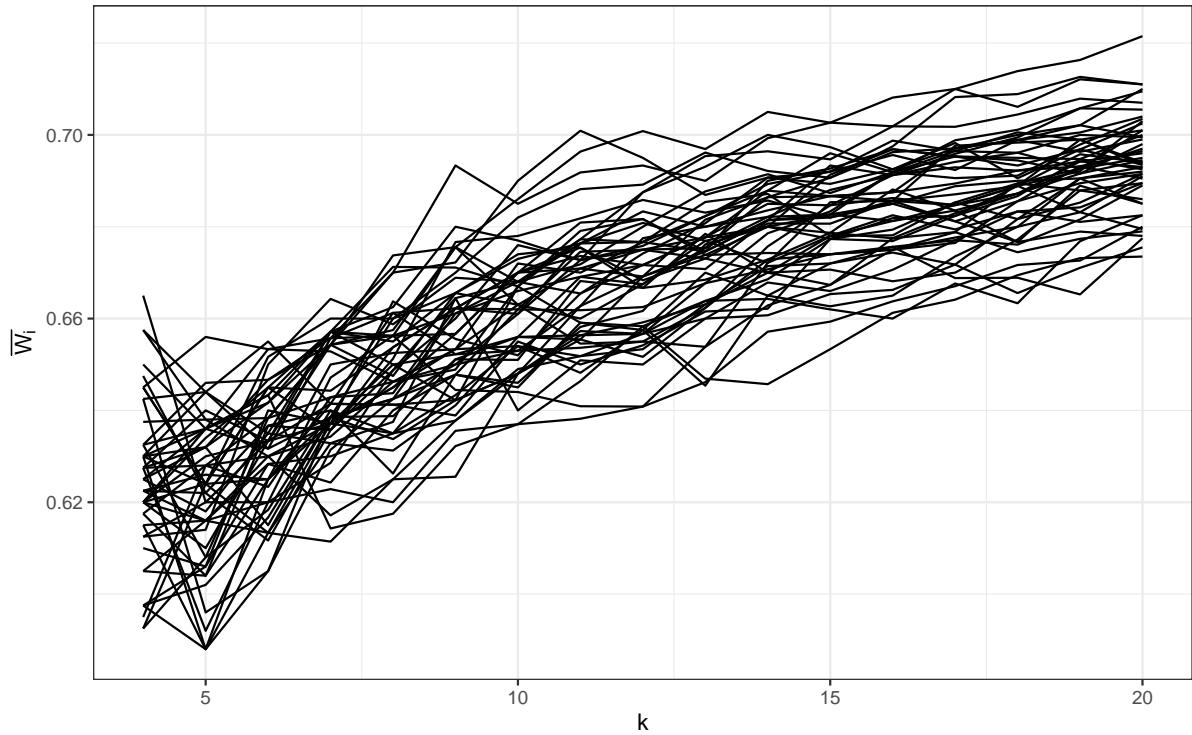


Figure 14: Variation in winkelmass with increasing number of neighbours  $k$  considered in the calculation. 50 replicate plots were used, each with 100 individuals randomly distributed in space.

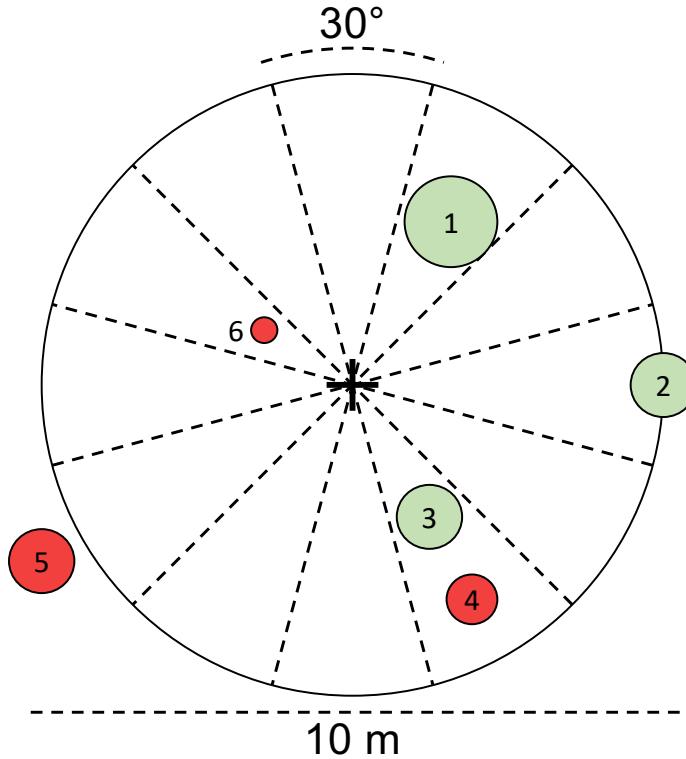


Figure 15: Schematic diagram demonstrating use of the Iterative Hegyi Index to assess crowding within each subplot. The 10 metre diameter subplot is divided into 12 equally sized sectors. Within each sector, the nearest stem of sufficient size ( $>5$  cm diameter) to the subplot centre is recorded (e.g. 1). All stems with any canopy material inside the subplot are valid (e.g. 2). Stem 4 is not valid as it is behind stem 3. Stem 5 is invalid as all its canopy is outside the subplot. Stem 6 is too small to be recorded.

213 rather than a focal tree, transforming it from a tree-centric competition index to a point-centric  
 214 crowding index:

$$H_i = \log \sum_{j=1}^n \left( \frac{1}{L_{ij}} D_j \right) \quad (8)$$

215 Where  $n$  is the number of stems with canopy material within the subplot,  $D_j$  is the stem diameter  
 216 of stem  $j$  and  $L_{ij}$  is the distance of stem  $j$  from the subplot centre  $i$ .  $H_i$  uses an iterative method  
 217 for choosing active canopy occupants at the subplot centre, where the nearest individual to the  
 218 subplot centre from each of 12 equally sized sectors is classified as the active stem (Figure 15).  
 219  $H_i$  was preferred over stem density to describe stem crowding in subplots because it is sensitive  
 220 to how close a stem is to the subplot centre, which will affect canopy closure even if the number  
 221 of stems in the subplot remains the same.

## 222 References

- 223 Brusa, A. & D. E. Bunker (2014). ‘Increasing the precision of canopy closure estimates from  
 224 hemispherical photography: Blue channel analysis and under-exposure’. In: *Agricultural and  
 225 Forest Meteorology* 195–196, pp. 102–107. DOI: 10.1016/j.agrformet.2014.05.001.

- 226 Chen, X., T. Allison, W. Cao, K. Ferguson, S. Grunig, V. Gomez, A. Kipka, J. Kohler, H.  
 227 Landau, R. Leandro et al. (2011). *Trimble RTX, an innovative new approach for network*  
 228 *RTK*. Tech. rep. Portland OR, USA: International Technical Meeting of the Satellite Division  
 229 of the Institute of Navigation, ION GNSS, pp. 2214–2219.
- 230 Cifuentes, R., D. V. der Zande, J. Farifteh, C. Salas & P. Coppin (2014). ‘Effects of voxel size  
 231 and sampling setup on the estimation of forest canopy gap fraction from terrestrial laser  
 232 scanning data’. In: *Agricultural and Forest Meteorology* 194, pp. 230–240. DOI: 10.1016/j.  
 233 agrformet.2014.04.013.
- 234 Ehbrecht, M., P. Schall, J. Juchheim, C. Ammer & D. Seidel (2016). ‘Effective number of layers:  
 235 A new measure for quantifying three-dimensional stand structure based on sampling with  
 236 terrestrial LiDAR’. In: *Forest Ecology and Management* 380, pp. 212–223. DOI: 10.1016/j.  
 237 foreco.2016.09.003.
- 238 Frazer, G. W., R. A. Fournier, J. Trofymow & R. J. Hall (2001). ‘A comparison of digital and  
 239 film fisheye photography for analysis of forest canopy structure and gap light transmission’. In:  
 240 *Agricultural and Forest Meteorology* 109.4, pp. 249–263. DOI: 10.1016/s0168-1923(01)00274-  
 241 x.
- 242 Grau, E., S. Durrieu, R. Fournier, J.-P. Gastellu-Etchegorry & T. Yin (2017). ‘Estimation  
 243 of 3D vegetation density with Terrestrial Laser Scanning data using voxels. A sensitivity  
 244 analysis of influencing parameters’. In: *Remote Sensing of Environment* 191, pp. 373–388. DOI:  
 245 10.1016/j.rse.2017.01.032.
- 246 Hardiman, B. S., G. Bohrer, C. M. Gough, C. S. Vogel & P. S. Curtis (2011). ‘The role of canopy  
 247 structural complexity in wood net primary production of a maturing northern deciduous  
 248 forest’. In: *Ecology* 92.9, pp. 1818–1827. DOI: 10.1890/10-2192.1.
- 249 Hegyi, F. (1974). ‘A simulation model for managing jack-pine stands’. In: *Royal College of  
 250 Forestry, editor*. Stockholm, Sweden: Royal College of Forestry, pp. 74–90.
- 251 Herbert, T. J. (1987). ‘Area projections of fisheye photographic lenses’. In: *Agricultural and  
 252 Forest Meteorology* 39.2-3, pp. 215–223. DOI: 10.1016/0168-1923(87)90039-6.
- 253 Herrero-Huerta, M., A. Bucksch, E. Puttonen & K. M. Rainey (2020). ‘Canopy Roughness: A  
 254 New Phenotypic Trait to Estimate Aboveground Biomass from Unmanned Aerial System’. In:  
 255 *Plant Phenomics* 2020, pp. 1–10. DOI: 10.34133/2020/6735967. URL: <https://doi.org/10.34133%2F2020%2F6735967>.
- 256 Hu, L. & J. Zhu (2009). ‘Determination of the tridimensional shape of canopy gaps using two  
 257 hemispherical photographs’. In: *Agricultural and Forest Meteorology* 149.5, pp. 862–872. DOI:  
 258 10.1016/j.agrformet.2008.11.008.
- 259 Huang, L. & M. J. Wang (1995). ‘Image thresholding by minimizing the measures of fuzziness’.  
 260 In: *Pattern Recognition* 28.1, pp. 41–51. DOI: 0031-3203/95.
- 261 Hui, G. & M. Albert (2004). ‘Stichprobensimulationen zur Schätzung nachbarschaftsbezogener  
 262 Strukturparameter in Waldbeständen [Simulation studies for estimating neighborhood-based  
 263 structural parameters in forest stands]’. In: *Allgemeine Forst und Jagdzeitung* 175, pp. 10–11.
- 264 Hui, G., K. von Gadow, Y. Hu & H. Xu (2007). *Structure-based forest management*. Beijing,  
 265 China: China Forestry Publishing House.
- 266 Hui, G., X. Zhao, Z. Zhao & K. von Gadow (2011). ‘Evaluating tree species spatial diversity  
 267 based on neighbourhood relationships’. In: *Forest Science* 57.4, pp. 292–300. DOI: 10.1093/  
 268 forestscience/57.4.292.
- 269 Hui, G. & K. Gadow (2002). ‘Das Winkelmaß-Theoretische Überlegungen zum optimalen  
 270 Standardwinkel’. In: *Allgemeine Forst U. Jagdzeitung* 173 (9), pp. 173–177.
- 271 Jennings, S. (1999). ‘Assessing forest canopies and understorey illumination: canopy closure,  
 272 canopy cover and other measures’. In: *Forestry* 72.1, pp. 59–74. DOI: 10.1093/forestry/72.  
 273 1.59.

- 275 Jupp, D. L. B., D. S. Culvenor, J. L. Lovell, G. J. Newnham, A. H. Strahler & C. E. Woodcock  
 276 (2008). ‘Estimating forest LAI profiles and structural parameters using a ground-based laser  
 277 called ‘Echidna’. In: *Tree Physiology* 29, pp. 171–181. DOI: 10.1093/treephys/tpn022.
- 278 Khosravipour, A., A. K. Skidmore, M. Isenburg, T. Wang & Y. A. Hussin (2014). ‘Generating  
 279 Pit-free Canopy Height Models from Airborne LiDAR’. In: *Photogrammetric Engineering &*  
 280 *Remote Sensing* 80.9, pp. 863–872. DOI: 10.14358/pers.80.9.863.
- 281 Leica Camera AG (2009). *Leica Cyclone*. Version 9.1.
- 282 Montes, F., I. Cañellas, M. del Río, R. Calama & G. Montero (2004). ‘The effects of thinning  
 283 on the structural diversity of coppice forests’. In: *Annals of Forest Science* 61.8, pp. 771–779.  
 284 DOI: 10.1051/forest:2004074.
- 285 PDAL Contributors (2018). *PDAL Point Data Abstraction Library*. DOI: 10.5281/zenodo .  
 286 2556738.
- 287 Persistence of Vision Pty. Ltd. (2004). *Persistence of Vision Raytracer (Version 3.7)*. [Computer  
 288 software].
- 289 Rusu, R. B., Z. C. Marton, N. Blodow, M. Dolha & M. Beetz (2008). ‘Towards 3D Point cloud  
 290 based object maps for household environments’. In: *Robotics and Autonomous Systems* 56.11,  
 291 pp. 927–941. DOI: 10.1016/j.robot.2008.08.005.
- 292 Scheuermann, C. M., L. E. Nave, R. T. Fahey, K. J. Nadelhoffer & C. M. Gough (2018).  
 293 ‘Effects of canopy structure and species diversity on primary production in upper Great  
 294 Lakes forests’. In: *Oecologia* 188.2, pp. 405–415. DOI: 10.1007/s00442-018-4236-x. URL:  
 295 <https://doi.org/10.1007%2Fs00442-018-4236-x>.
- 296 Schneider, C. A., W. S. Rasband & K. W. Eliceiri (2012). ‘NIH Image to ImageJ: 25 years of  
 297 image analysis’. In: *Nature Methods* 9.7, pp. 671–675. DOI: 10.1038/nmeth.2089.
- 298 Seidel, D., S. Fleck & C. Leuschner (2012). ‘Analyzing forest canopies with ground-based  
 299 laser scanning: A comparison with hemispherical photography’. In: *Agricultural and Forest  
 300 Meteorology* 154-155, pp. 1–8. DOI: 10.1016/j.agrformet.2011.10.006.
- 301 ter Steege, H. (2018). *Hemiphot.R: Free R scripts to analyse hemispherical photographs for  
 302 canopy openness, leaf area index and photosynthetic active radiation under forest canopies*.  
 303 Unpublished report. Leiden, The Netherlands: Naturalis Biodiversity Center. URL: <https://github.com/Naturalis/Hemiphot>.
- 305 von Gadow, K. & G. Hui (2002). ‘Characterising forest spatial structure and diversity’. In:  
 306 *Proceedings of the IUFRO International workshop ‘Sustainable forestry in temperate regions’*.  
 307 Ed. by L. Bjoerk. Lund, Sweden, pp. 20–30.
- 308 Weligepolage, K., A. Gieske & Z. Su (2012). ‘Surface roughness analysis of a conifer forest canopy  
 309 with airborne and terrestrial laser scanning techniques’. In: *International Journal of Applied  
 310 Earth Observation and Geoinformation* 14.1, pp. 192–203. DOI: 10.1016/j.jag.2011.08.014.  
 311 URL: <https://doi.org/10.1016%2Fj.jag.2011.08.014>.
- 312 Wilkes, P., S. D. Jones, L. Suarez, A. Haywood, A. Mellor, W. Woodgate, M. Soto-Berelov &  
 313 A. K. Skidmore (2016). ‘Using discrete-return airborne laser scanning to quantify number of  
 314 canopy strata across diverse forest types’. In: *Methods in Ecology and Evolution* 7.6. Ed. by  
 315 S. McMahon, pp. 700–712. DOI: 10.1111/2041-210x.12510. URL: <https://doi.org/10.1111%2F2041-210x.12510>.
- 317 Zhang, K., S.-C. Chen, D. Whitman, M.-L. Shyu, J. Yan & C. Zhang (2003). ‘A progressive  
 318 morphological filter for removing nonground measurements from airborne LiDAR data’. In:  
 319 *IEEE Transactions on Geoscience and Remote Sensing* 41.4, pp. 872–882. DOI: 10.1109/tgrs.  
 320 2003.810682.