



HAL
open science

Sensitivity of voxel-based estimations of leaf area density with terrestrial LiDAR to vegetation structure and sampling limitations: A simulation experiment

Maxime Soma, François Pimont, Jean-Luc Dupuy

► To cite this version:

Maxime Soma, François Pimont, Jean-Luc Dupuy. Sensitivity of voxel-based estimations of leaf area density with terrestrial LiDAR to vegetation structure and sampling limitations: A simulation experiment. *Remote Sensing of Environment*, 2021, 257, pp.112354. 10.1016/j.rse.2021.112354 . hal-03353461

HAL Id: hal-03353461

<https://hal.inrae.fr/hal-03353461>

Submitted on 9 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Sensitivity of voxel-based estimations of Leaf Area Density with Terrestrial LiDAR to vegetation**
2 **structure and sampling limitations: a simulation experiment**

3 Maxime Soma^{1,2*}, François Pimont¹, Jean-Luc Dupuy¹

4 ¹ UR 629 Ecologie des Forêts Méditerranéennes (URFM), INRAe, 84914 Avignon, France

5 ² UMR 114 EMMAH, UMT CAPTE, INRAe, 84914 Avignon, France

6 **Abstract**

7 The need for fine scale description of vegetation structure is increasing as Leaf Area Density
8 (LAD, m²/m³) becomes a critical parameter to understand ecosystem functioning and energy
9 and mass fluxes in heterogeneous ecosystems. Terrestrial Laser Scanning (TLS) has shown
10 great potential for retrieving the foliage area at stand, plant or voxel scales. Several sources of
11 measurement errors have been identified and corrected over the past years. However,
12 measurements remain sensitive to several factors, including, 1) voxel size and vegetation
13 structure within voxels, 2) heterogeneity in sampling from TLS instrument (occlusion and
14 shooting pattern), the consequences of which have been seldom analyzed at the scale of forest
15 plots. In the present paper, we aimed at disentangling biases and errors in plot-scale
16 measurements of LAD with TLS in a simulated vegetation scene. Two negative biases were
17 formerly attributed to (i) the unsampled voxels and to (ii) the subgrid vegetation heterogeneity
18 (i.e. clumping effect), and then quantified, thanks to a the simulation experiment providing
19 known LAD references at voxel scale, vegetation manipulations and unbiased point
20 estimators. We used confidence intervals to evaluate voxel-scale measurement accuracy.

21 We found that the unsampled voxel effect (i) led to underestimations with the “mean layer”
22 method –commonly used to fill unsampled voxels- for small voxels (0.1-0.2 m) and/or low
23 number of scans (<4). It was explained by the spatial correlations in vegetation, which
24 induced that dense voxels were more often occluded by dense neighbors than light voxels.
25 The distribution of the bias was heterogeneous in canopy, the bias being stronger at mid

26 canopy where occlusion started, but smaller in highly-occluded upper layers. This somehow
27 counterintuitive result was explained by a more random sampling of upper layers, but could
28 highly depend on vegetation structure.

29 The subgrid vegetation heterogeneity effect (ii) was confirmed to increase with voxel size,
30 yet, the magnitude of this bias -quantified with vegetation manipulation- was found to be
31 more homogeneously-distributed than the unsampled voxel effect.

32 Overall, we found that no scenario was unbiased. However, an intermediate voxel size (0.5m)
33 was the best option, because the relatively homogeneous subgrid effect could be handled with
34 a single correction factor and voxel-scale measurements errors were reasonable. On the
35 contrary, smaller voxels led to poor voxel-scale measurements and variable biases in
36 magnitude and spatial distribution with sampling design. However, more similar research in
37 other context is required to adapt these conclusions to other forest plots.

38

39 **Keywords:** Terrestrial LiDAR; Leaf Area Density (LAD); LAI; Occlusion; Voxel

40

41

42

43

44

45

46

47

48

49 **1. Introduction**

50

51 The tridimensional distribution of vegetation in forest ecosystems constitutes a complex
52 surface of exchange between canopies, ground and atmosphere, driving mass and energy fluxes
53 (Norman and Campbell, 1989). Characterizing its structure is key for monitoring health and carbon
54 storage of forests, but also for understanding and modelling tree ecophysiological processes. This
55 amount of vegetation can be described by the Leaf Area Density (LAD) distribution, i.e. the total leaf
56 area per unit of volume (Weiss et al., 2004). In the past decade, LiDAR technologies have been used
57 for measuring LAD at various scales. Airborne LiDAR can provide canopy profile estimates at coarse
58 grains and over large areas, which is very useful for forests monitoring, but estimates can hardly be
59 evaluated against ground references (Kamoske et al., 2019; Stark et al., 2012). Meanwhile, Terrestrial
60 laser scanning (TLS) has been used successfully for local measuring, such as the total leaf area of
61 individual trees (Béland et al., 2011; Hosoi et al., 2010; Hu et al., 2018) and the LAD profiles at plot
62 scale (Pimont et al., 2016; Schneider et al., 2019). The detailed 3D point clouds provided by TLS can
63 closely describe the scene in three dimensions, the density of points being related to the presence of
64 vegetation (Durrieu et al., 2008). LAD or PAD (Plant Area Density, i.e. vegetation area density,
65 indiscriminate between leaf and wood material) has been estimated from a variety of methods based
66 on the similar theoretical background related to the attenuation of light in uniform medium. Typically,
67 most commonly used methods are inherited from Leaf Area Index measurements, and are based on the
68 Beer's law theory (Yan et al., 2019), i.e. the exponential attenuation of the transmittance along the
69 path of the beam, which readily relates the gap fraction and the attenuation coefficient of the medium.
70 Path length distribution models fully exploit the 3D information contained in point clouds by
71 accounting for the path length of beam within canopy before its interception, which has been shown to
72 provide better estimates (Yan et al., 2019), whether integrated to Beer's law approaches (Zhao et al.,
73 2015) or directly arising from contact frequency method (Béland et al., 2011; Pimont et al., 2019,
74 2018). LAD estimates can be retrieved either for a given layer in order to yield LAD profiles (Hosoi
75 and Omasa, 2006; Zhao et al., 2015) or directly within each voxel (Béland et al., 2011; Soma et al.,
76 2018). Measuring LAD at voxel scale is particularly relevant for optical measurements in non-uniform

77 canopies context because it allows to discriminate gaps and within crown volumes, thus limiting
78 clumping effect (Yan et al., 2019). Pimont et al. (2018) provided a detailed comparison of voxel-based
79 methods for the estimation of the attenuation coefficient within a single voxel, which is linearly
80 related to LAD/PAD. They introduced a more straightforward approach based on Maximum
81 Likelihood Estimation (MLE) and suggested corrections when the beam number was low and the
82 vegetation element size was not very small with respect to voxel size, resulting in an unbiased
83 estimator. In general, most theoretical estimates compare reasonably well with reference
84 measurements (Bailey and Mahaffee, 2017; Béland et al., 2011; Hu et al., 2018; Pimont et al., 2016;
85 Soma et al., 2018). However, TLS-based estimations are also affected by several biases, arising both
86 from inversion method and instrument limitations (Yan et al., 2019).

87 With the voxel-based approach, LiDAR estimates of LAD/PAD generally decrease with voxel
88 size (Bailey and Mahaffee, 2017; Béland et al., 2014a; Soma et al., 2018). Although the effect of voxel
89 size has been repeatedly observed (Cifuentes et al., 2014; Huang and Pretzsch, 2010; Li et al., 2016),
90 few studies conducted a complete analysis of this effect (Béland et al., 2014a; Grau et al., 2017). Such
91 an effect has been attributed to vegetation clumping, which induces the presence of heterogeneities in
92 the vegetation structure inside the voxel, referred to as subgrid vegetation heterogeneity. The choice of
93 too large voxels regarding vegetation structure hampers a proper discretization of gaps and clumps,
94 leading to a negative bias in estimates as a consequence of Jensen inequality (Ruel and Ayres, 1999).
95 Additionally, for some instruments, some critical overestimations of LAD occur as distance from
96 scanner increases. This bias has been attributed to the increase in effective footprint of the beam
97 associated with beam divergence (Soma et al. 2018). Both effects can be accounted for through an
98 adequate modelling of the LiDAR signal (Béland et al., 2014b, 2011) or simple calibration
99 coefficients (Soma et al. 2018). Such calibrations have been included in the unbiased estimator
100 approach presented in Pimont et al. (2019), in a new formulation of the estimation of the LAD, which
101 also includes a rigorous incorporation of the wood element volume and multiview LiDAR data
102 (Pimont et al. 2019).

103 Another important TLS drawback at plot scale, is the spatial variation in sampling quality
104 within point clouds (Schneider et al., 2019; Wilkes et al., 2017). Indeed, the geometry of the shooting

105 pattern of most TLS instruments (such as FARO 130X) has a fixed angular step resolution. Hence, the
106 beam density decreases with distance to scanner. The second source of variation in beam sampling is
107 vegetation occlusion: when a beam is fully intercepted (single hit for monoecho TLS or the last return
108 in case of multiecho TLS), it cannot explore beyond the intercepting voxel. This implies that fewer
109 beams remain when vegetation is present between the scanner and the voxel of interest. Both geometry
110 and occlusion contribute to the occurrence of unsampled voxels, where no information or limited
111 information is available. Since the reliability of LAD estimation within a given voxel strongly depends
112 on the number of beams exploring each voxel (Pimont et al., 2018), random errors in individual voxels
113 are very large when beam number is low. Some point estimators are theoretically unbiased (Pimont et
114 al., 2019, 2018), so that one could expect that the random errors will cancel when averaged over a
115 large number of voxels, e.g. when vertical profiles are computed.

116 This assumption, however, might not be correct in the field because of the presence of
117 unsampled voxels in the context of heterogeneous canopy. Indeed, one could expect that unsampled
118 voxels would mostly be located in dense vegetation spots, where occlusion is the strongest. At plot
119 scale, this could result in a sampling bias, leading to a potential underestimation of the mean LAD.
120 Typically, the most commonly used methods for correcting for occlusion relies on the mean estimate
121 in layers, i.e. on estimates within explored areas, which might not be representative of occluded areas
122 in clumped vegetation context. To date, few studies have conducted an analysis of occlusion
123 distribution, magnitude and mechanisms at tree (Béland et al., 2011; Yun et al., 2019) or plot scale
124 (Morsdorf et al., 2018; Schneider et al., 2019). Moreover, these studies did not formally disentangle
125 the relative contributions of unsampled voxels from subgrid vegetation heterogeneity, since both vary
126 with voxel size. Also, most previous studies carried out at tree or forest-plot scales relied on point
127 estimates with theoretical biases, which complicates the analysis of biases at tree or plot scales.

128 To improve sampling quality, a widespread practice during data acquisition is to acquire
129 multiple scans across the scene in order to increase both absolute number of beams and variety of
130 viewpoints (Côté et al., 2011; Wilkes et al., 2017). Moreover, the number of scans used in the field
131 and the scanner resolution are generally limited given technical constraints, such as time to proceed or
132 storage capability (Wilkes et al., 2017). This implies some constraints on voxel size used for

133 estimations, since the fraction of non-explored voxels strongly rises as voxel size decreases (Béland et
134 al., 2014a). Scanning design and post-processing recommendations accounting for these sampling
135 limitations are generally dedicated to applications such as detection of stems or measurement of tree
136 geometrical features (e.g. diameter) (Gollob et al., 2019; Wilkes et al., 2017), but recommendations
137 for description of canopies features are still rare at plot scale (Schneider et al., 2019).

138 Improving our understanding of the impacts of both heterogeneity effect and occlusion relying
139 on actual data is very challenging because of the lack of accuracy in references, particularly in dense
140 and complex canopy structure, while the estimation of leaf projection factor or leaf fraction are
141 additional sources of errors. In this difficult context, a simulation framework is mandatory to provide a
142 perfectly controlled environment to test various configurations and methods of estimations and
143 disentangle multiple sources of biases (Morsdorf et al., 2018; Yan et al., 2019), as done in Grau et al.,
144 (2017). In the present study, we rely on numerical experiments to analyze the sensitivity of LAD
145 estimations to biases arising from (1) vegetation heterogeneity and (2) heterogeneity of the sampling
146 quality at plot scale, which depends on both vegetation structure at canopy scale and sampling design.
147 We performed these simulations with unbiased point estimators in the framework of a Mediterranean
148 canopy sampled with a single echo terrestrial LiDAR. Here, we expect to disentangle and better
149 understand the mechanisms affecting estimations in the field, to analyze their magnitude and how they
150 distribute in forest plots, in order to finally derive recommendations on sampling design and selection
151 of appropriate voxel sizes.

152 **2. Numerical experiments**

153 *2.1. Overview*

154 The aim of the numerical experiments was to disentangle biases, errors and limitations arising
155 from voxel size and number of scans for the estimation of LAD in heterogeneous vegetation. For this,
156 we first generated a high resolution reference LAD field that had to meet the following requirements:
157 (i) it had to be as representative as possible of natural vegetation, with dense clumps (corresponding to
158 tree crowns) surrounded by empty locations (ii) it should be realistic for a forest plot, in terms of both
159 LAI and LAD profile, as occlusion patterns are highly sensitive to vegetation attenuation coefficient.

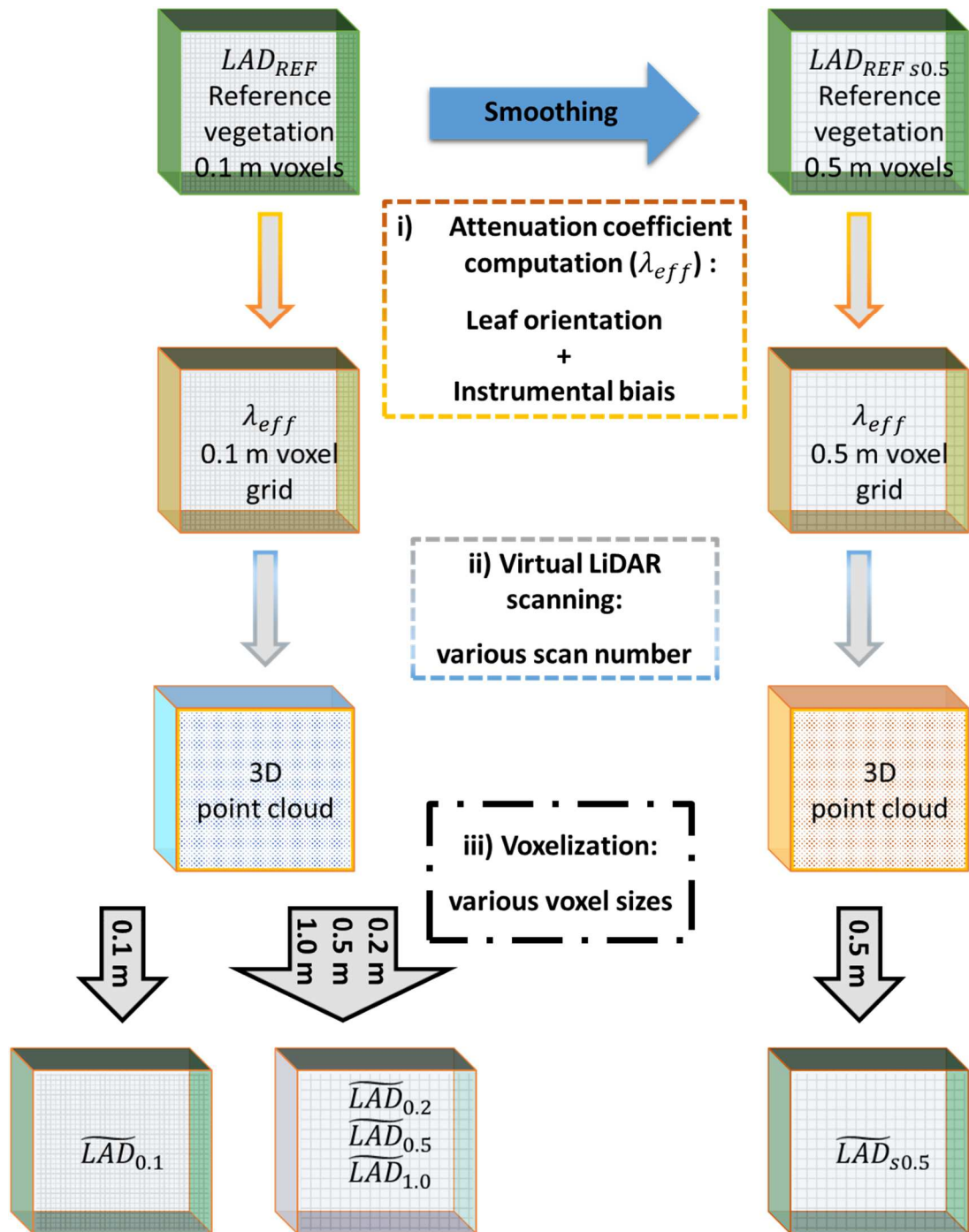
160 Generation of the LAD reference field (LAD_{REF}) is described in section 2.2. This field had a resolution
161 of 0.1 m. An additional smoother reference fields of the same vegetation scene ($LAD_{REF\ s0.5}$) was
162 generated by coarsening the resolution to 0.5 m.

163 Then, we simulated point clouds that would have been produced by a LiDAR scanning on
164 vegetation corresponding to reference LAD fields under different configurations (number and position
165 of scans). Following Soma et al. (2018) and Pimont et al. (2019), the effective attenuation coefficient
166 λ_{eff} in a given voxel, which is the rate at which beam density is attenuated by vegetation from a given
167 scan position, relates to LAD_{REF} as follows:

$$\lambda_{eff} = \frac{G}{FH} LAD_{REF} \quad (1)$$

168 where G is the leaf projection function, F is the leaf fraction and H is a calibration function
169 which accounts for both voxel size and distance to scanner. To simulate the influence of these factors
170 on LAD estimation, we computed the distribution of effective attenuation coefficient (λ_{eff}) for each
171 scan position, which corresponds to the perception of LAD_{REF} by the instrument. In order to remain
172 representative of typical field conditions, a maximum of 5 scans was performed in a 10 m X 10 m
173 virtual scene. Then for each scan position, the traversal algorithm was applied to the λ_{eff} -3D field to
174 simulate the perceived point clouds, following parameters arising from vegetation and scanner
175 properties (Eq. 1). The details of the point cloud simulations are presented in section 2.3.

176 Finally, LAD estimates were retrieved from these point clouds with different voxel sizes using a
177 traversal algorithm and calibrations provided by field experiments (Soma et al., 2018). The
178 information from multiple scans was combined using the formulation of Pimont et al., (2019) (section
179 2.4). The last section 2.5 describes how we analyzed bias and errors in these estimations. A workflow
180 diagram picturing the methodology used in this work is available in Figure 1.



181

182

183

184

185

186

Figure 1. Workflow diagram of simulated LAD scenes and used references. LAD_{REF} is the vegetation reference distribution at 0.1 m heterogeneity scale. $LAD_{REF\ s0.5}$ is the same reference distribution, but at 0.5 m heterogeneity scale. λ_{eff} is the attenuation coefficient. \overline{LAD} is the estimated vegetation distribution, with voxel sizes in subscripts.

187 2.2. *Vegetation scene*

188 A 3D virtual vegetation scene described from voxelized LAD_{REF} values was generated from a
189 clumped spatial distribution simulated with *RandomFields* R package (Schlather et al., 2015) in a grid
190 with 0.1 m resolution. The mean clump size, representative of tree crown diameter, was 4 m. This
191 virtual scene was cubic with an extent of 10 m in each grid dimension. The canopy height was thus
192 $h=10$ m. The voxel size was set to 0.1 m.

193 The vegetation scene included a realistic vertical profile, with little vegetation below 3 m, an
194 increase to a maximum of LAD values at 7 m, and a decrease until top canopy (see Appendix A). This
195 scene also included random occurrences of 1 m extent local decreases in LAD, with gaps (LAD equals
196 to zero) in 30% of the voxels. This resulted in a vegetation scene with a 70% cover fraction with
197 spatial aggregates corresponding to plant clumps (Figure 2). Such gap distribution was built to
198 analyze the effect of intermediate scale heterogeneity in tree crowns (see spatial correlation in
199 Appendix B). The LAI of the virtual scene was about 3.8, which corresponds to a mean LAD of 0.38
200 m^{-1} (the scene vertical extent was 10 m). Maximal LAD values reached $3.8 m^{-1}$. These vegetation
201 properties match the field measurements of a long-term experimental site characteristic of
202 Mediterranean ecosystems (Simioni et al., 2016). The same scene was formerly used in Soma et al.
203 2020, to evaluate the effect of a kriging method for unsampled voxels (Soma et al. 2020).

204 A leaf projection function was implemented to complete vegetation properties following:

$$G(\theta, z) = \frac{1}{2} + 0.4 \frac{z}{h} \cos(2\theta) \quad (2)$$

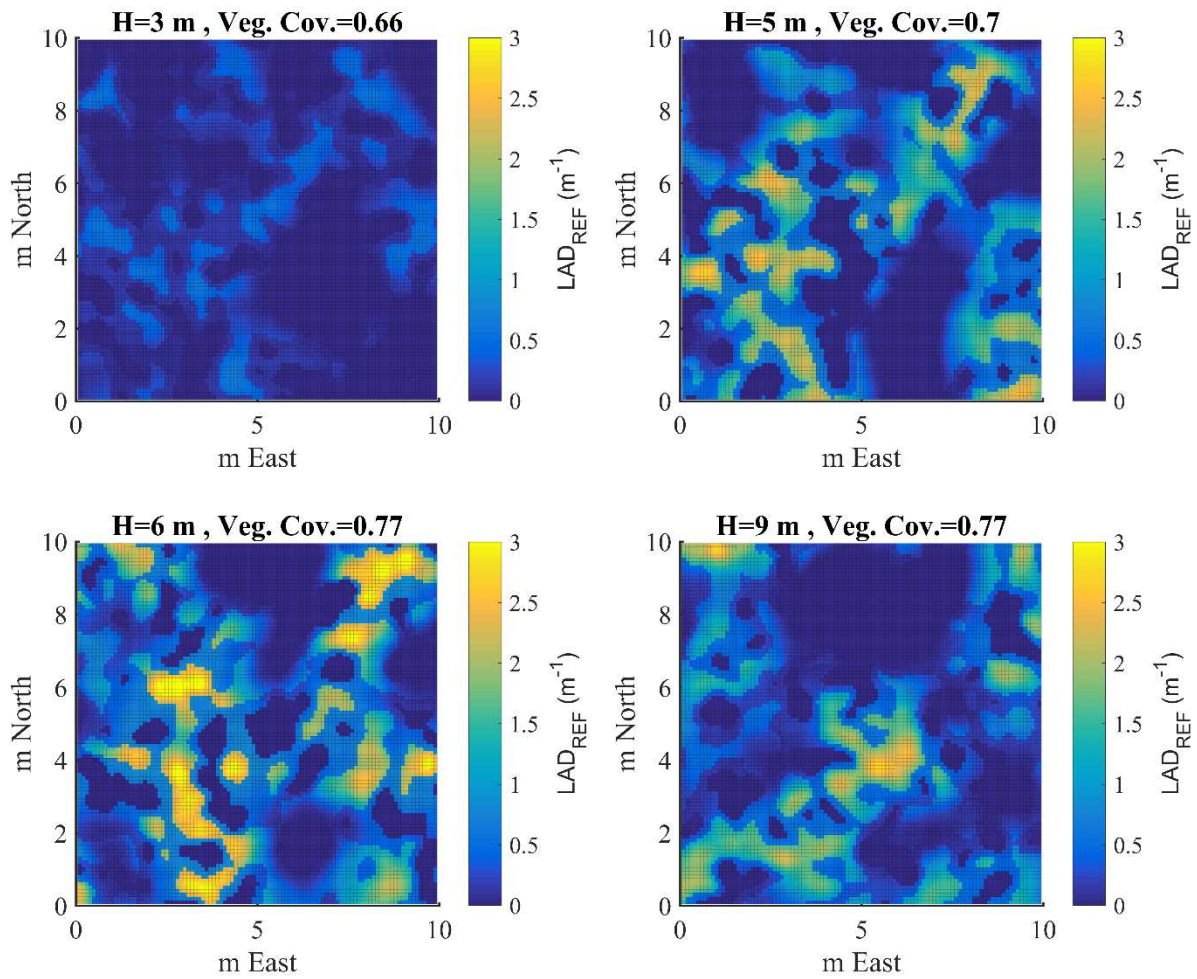
205 where θ is the angle between the beams and the vertical and ranges between 0 and π . According to
206 this setting leaves are planophile near the canopy top ($z \approx h$), with $G=0.9$ for vertical beams ($\theta \approx 0$
207 or $\theta \approx \pi$) and 0.1 for horizontal beams ($\theta \approx \frac{\pi}{2}$), and random near the ground ($z \approx 0$), with $G=0.5$.

208 At last, leaf fraction was parameterized to account for fine woody material and leaf association
209 along vertical axis following:

$$F(z) = \left(0.1 + 0.8 \frac{z}{h}\right)^2 \quad (3)$$

210 Leaf fraction was hence 0.9 at canopy top ($z \approx h$) and 0.1 near the ground ($z \approx 0$).

211
212



213

214 **Figure 2.** Reference horizontal distributions of LAD_{REF} at 4 different heights. Heights (H) of the
215 mapped layer and associated vegetation cover fractions are indicated in titles. Colors are scaled on
216 LAD_{REF} values from low to high densities (blue to yellow, respectively). A three dimensional view is
217 available in Appendix C.

218

219 2.3. Simulation of a virtual point cloud corresponding to LAD_{REF} field

220 Simulation of virtual point clouds were based on turbid media assumption, which states that the
221 probability of a beam to be intercepted increases exponentially with the optical depth (product of
222 attenuation coefficient and distance travelled). The locations in which individual laser beams are
223 intercepted are thus generated from the simulation of free paths (distance travelled before interception)

224 from random numbers following probabilistic laws which account for vegetation density, as done for
 225 example in Pimont et al. (2018) within a single voxel. Here the approach is generalized to a
 226 heterogeneous scene, corresponding to an idealized forest plot with a spatially-correlated LAD field.
 227 The process is iterated in the different voxels travelled along the beam pass, for all beams according to
 228 the shooting pattern of the instrument representative of TLS field sampling of a forest plots.

229

230 *2.3.1 Effective attenuation coefficient $\lambda_{eff,j}$ corresponding to a given scan j*

231 The reference vegetation scene is the voxelized LAD_{REF} distribution described above. For
 232 simplicity, we assume that the size of vegetation elements (Beer law's approximation) and the volume
 233 of voxel occupied by wood elements are negligible. This allows simple framework adequate for
 234 testing sampling effect at stand-scale, but both factors can be rigorously accounted for, as shown in
 235 Pimont et al. (2019). As detailed above (Eq. 1), the reference attenuation coefficient $\lambda_{eff,j}$ related to
 236 LAD_{REF} for a given scan j depends on leaf projection, leaf fraction, vegetation heterogeneity and
 237 scanner properties. Yet, the probability of interception of laser beams depends exclusively on $\lambda_{eff,j}$,
 238 not directly on LAD_{REF} . To proceed to a realistic point cloud simulation and account for various
 239 distances between voxels and scan positions, the $\lambda_{eff,j}$ has to be computed at each voxel for each
 240 viewpoint j according to Eq. 1.

241 Let (x_j, y_j, z_j) be the coordinates of the scanner corresponding to scan j and (x, y, z) the
 242 coordinates of the center of a voxel in the vegetation scene. Following Eq.1 and according to the
 243 vegetation definition above, the effective attenuation coefficient for both leaf and wood for scan j is:

244

245

$$\lambda_{eff,j}(x, y, z) = LAD_{ref}(x, y, z) \frac{G_j(x, y, z)}{F(z)H_j(x, y, z)} \quad (4)$$

246

247

248 A beam emitted from the scanner that would go in the direction of this point has the following
 249 projection function G (since $\cos(2\theta) = \cos(\theta)^2 - \sin(\theta)^2$):

250

$$G_j(x, y, z) = \frac{1}{2} + 0.4 \frac{z(z - z_j)^2 - (x - x_j)^2 - (y - y_j)^2}{h((x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2)} \quad (5)$$

251

252 We assumed that the distance effect (caused by an increase in effective footprint of the scanner, as
 253 identified in Soma et al. 2018)) have the following effect on the attenuation coefficient:

254

$$H_j(x, y, z) = 1 - 0.05 \sqrt{(x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2} \quad (6)$$

255

256 which expresses that vegetation area is overestimated by a factor 2 at a distance of 10 m to the scanner
 257 ($H_j=0.5$). The reference LAD_{REF} field used to compute $\lambda_{eff,j}$ in presented in Figure 2. at 0.10 m
 258 discretization level (see section 4.5 for details).

259

260 2.3.2 Point cloud simulations

261 From the effective attenuation coefficient distribution (Eq. 15), we can simulate a virtual point
 262 cloud based on free path distribution simulations, as explained below. We assumed that function in
 263 (Eq. 15) can be discretized in a 3D computational grid ($\lambda_{eff,j}$ assumed to be constant within a
 264 computation cell, its value being equals to (Eq. 17) with (x, y, z) from cell center coordinate) with
 265 vegetation elements much smaller than grid size (turbid medium assumption).

266 Five scans were simulated. For each scan configuration, scans positions were placed with respect to
 267 optimal scene sampling coverage, similarly as field experiment would be, with the potential exception
 268 of absence of dense vegetation under 3m. All scans were placed at 1 m from the ground. 4 scans were
 269 placed close to each corner of the plot and one scan at the center:

$$270 (x_1, y_1, z_1) = (7.5, 7.5, 1); (x_2, y_2, z_2) = (7.5, 2.5, 1); (x_3, y_3, z_3) = (2.5, 2.5, 1);$$

$$271 (x_4, y_4, z_4) = (2.5, 7.5, 1); (x_5, y_5, z_5) = (5, 5, 1).$$

272 We also parameterized their shooting pattern as follows. Their angular resolution was 0.05° over the
273 horizontal (ranging from 0 to 180°) and the vertical (ranging from 0 to 360°), so that each scan
274 contains 66 million beams, which is typical of the resolution used in the field (e.g. Pimont et al. 2015).
275 For each beam, we simulated its eventual hit position as follows: First, optical path (i.e. initial
276 potential to pass through vegetation) of each beam was randomly simulated according to Beer-
277 Lambert law (assuming infinitely small elements):

$$l = -\log(p) \tag{7}$$

279 with p a random number between $]0;1[$ (i.e. initial random chance to be intercepted by vegetation).

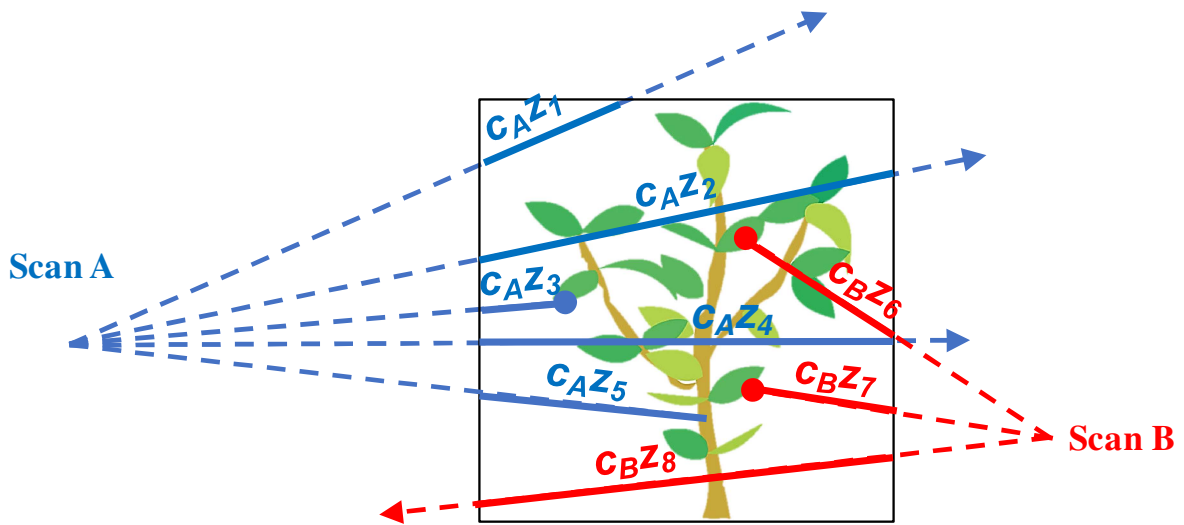
280
281 We then computed the trajectory of this beam within the computation grid, from its initial position at
282 scanner location, by computing the “amount” of optical path required to cross the next voxel. This
283 amount was calculated by multiplying the effective attenuation coefficient of this voxel by the length
284 of the segment corresponding to the intersection of the beam and the voxel. When the residual optical
285 path of the beam was shorter than this amount, a hit occurred within this voxel at a location
286 corresponding to this residual optical path. On the contrary, when the remaining optical path was
287 greater than this amount, it means that the beam travelled farther than the voxel. The process was
288 recursively applied to the next voxel, the “new” residual optical path corresponding to the remaining
289 of the previous one. The process ends in case of hit or when a beam reached the bounding box of the
290 computational grid. In this later case, the beam was never intercepted in the computational grid, thus
291 corresponding to a beam with no hit. This process is similar to the one used by (Pimont et al., 2009) to
292 simulate photons trajectories to compute the radiative transfer from a flame through a voxelized
293 heterogeneous vegetation scene with a MonteCarlo approach.

294
295 This process was repeated for each scan position. Hence, five virtual point clouds were
296 simulated in accordance with $\lambda_{eff,j}$, which accounts for both vegetation and instrument properties.

297

298 2.4. Computation of LAD estimates for simulated point clouds including unsampled voxels

299 For each viewpoint j , we used a traversal algorithm developed under Matlab software (The
 300 MathWorks, Inc., Natick, MA, USA). For each scan, intersections between each beam (i.e. line
 301 between hit and scan position) and the voxel grid were computed in order to retrieve the following
 302 metrics in each voxel: N_{ij} the number of hits in the voxel, $\sum z_j$ the sum of free paths, $\sum_{hit} z_j$ the sum
 303 of free paths corresponding to hits only.



304
 305 **Figure 3.** Scheme of the information provided by the traversal algorithm which is used to compute the
 306 unbiased estimator of LAD from multiview data from Scan A (in blue) and Scan B (in red): leaf hits
 307 (blue and red dots) and free paths (distances z travelled by the beams, blue and red lines) in the voxel.
 308 The dotted lines represent pulse trajectories. c_A and c_B represent the correcting factors for viewpoints A
 309 and B, which depend on distance to scanner and view angle (see Pimont et al., 2019 for details).

310
 311 From these quantities, we can define in each voxel the following “multi scan” quantities, using known
 312 values of G, H and F:

- 313 - $N_i = \sum_j N_{ij}$ the total number of hits
- 314 - $\sum \frac{G}{H} z = \sum_j \frac{G_j}{H_j} \sum z_j$ the total free path sum, with multiplicative correction factors.
- 315 - $\sum_{hits} \frac{G}{H} z = \sum_j \frac{G_j}{H_j} \sum_{hit} z_j$ the total “hit” free path sum, with multiplicative correction factors.

316 Following Pimont et al. (2019, Eq. 13), we used the following unbiased point estimator of LAD,
 317 provided that N was larger than 2:

$$\widetilde{LAD} = \frac{F}{\sum \frac{G}{H^z}} \left(Ni - \frac{\sum_{hits} \frac{G}{H^z}}{\sum \frac{G}{H^z}} \right) \quad (8)$$

318 In this formulation, $\frac{\sum_{hits} \frac{G}{H^z}}{\sum \frac{G}{H^z}}$ is a bias correction term ranging between 0 and 1, so that the estimator is
 319 simply $\frac{FNi}{\sum \frac{G}{H^z}}$ when Ni is large, which corresponds to the Maximum Likelihood Estimator (Pimont et al.
 320 2018). This formulation also accounts for leaf orientation, wood and instrument specifications (Pimont
 321 et al. 2019). This estimator generalizes the Modified Contact Frequency proposed by Béland et al.
 322 (2011) with the bias correction term to account for low sampling, leading to an unbiased point
 323 estimator provided that more than 2 beams are available (Pimont et al. 2018). The voxels with N
 324 smaller or equal to 1 were considered as unsampled, because no unbiased estimator has been
 325 suggested for N=1 (Pimont et al. 2019). Indeed, Eq. 8 leads to estimates of 0 when N=1, even when
 326 Ni=1, which is obviously negatively biased. As a result, voxels with N lower than 2 were considered
 327 as unsampled. For unsampled voxels, we used the “mean layer value” to estimate LAD, which was
 328 computed from the mean of \widetilde{LAD} (computed for N larger or equal than 2). This method for dealing
 329 with unsampled voxels is the most widely used (Béland et al., 2014b; Schneider et al., 2019).
 330 We also computed the radius of the 68% confidence interval for \widetilde{LAD} , which can be estimated by
 331 (Pimont et al. 2019):

$$\Delta \widetilde{LAD} = \frac{F \left(Ni - \frac{\sum_{hits} \frac{G}{H^z}}{\sum \frac{G}{H^z}} \right) + \frac{1}{2}}{\sqrt{FNi + \frac{1}{2} \sum \frac{G}{H^z} \left(1 + \frac{1}{N} \right)}} \quad (9)$$

332 with $N = \sum_j N_j$ the total number of beams entering a voxel.

333

334 2.5. Analysis of estimations

335

336 Table 1. Description of the different simulations. They differ both in the reference vegetation field
 337 used to generate the virtual point cloud (ray-tracing algorithm) and in the voxel size used for the
 338 estimation process.

<i>Reference vegetation used to generate the virtual point cloud (scale in m)</i>	<i>Voxel size used for estimation</i>	<i>Estimations</i>
LAD_{REF} (0.1 m)	0.1 m	$\overline{LAD}_{0.1}$
LAD_{REF} (0.1 m)	0.2 m	$\overline{LAD}_{0.2}$
LAD_{REF} (0.1 m)	0.5 m	$\overline{LAD}_{0.5}$
LAD_{REF} (0.1 m)	1.0 m	$\overline{LAD}_{1.0}$
$LAD_{REF\ s0.5}$ (0.5 m)	0.5 m	$\overline{LAD}_{s0.5}$

339 LAD_{REF} corresponds to the reference vegetation field defined at 0.1 m resolution.
 340 $LAD_{REF\ s0.5}$ corresponds to a manipulation of LAD_{REF} , which was smoothed at 0.5 m. As the scale of
 341 heterogeneity in these different reference fields differ, the subsequent virtual point clouds are also
 342 different.

343 2.5.1 Exploring bias arising from vegetation heterogeneity

344 It has been suggested that ignoring clumping effect within voxel could lead to strong LAD
 345 underestimation (Béland et al., 2014a; Soma et al., 2018). This bias depends on voxel size, large
 346 voxels being more prone to aggregate gaps and clumps within voxel volume. In the present study, we
 347 analysed the effect of voxel size by comparing estimates obtained with voxel sizes ranging from 0.10
 348 m to 1 m to reference values (section 5.4). This analysis relied on mean estimated vertical profiles
 349 (i.e. estimates were averaged by height layers). For each voxel size, the correction factor required to
 350 match reference value was derived at all heights. One of the drawback of a basic sensitivity analysis to
 351 voxel size is that it does not disentangle the effects of vegetation heterogeneity (i.e. Jensen's
 352 inequality) and sampling limitations (i.e. more or less beams exploring voxels). In order to minimize
 353 the role of sampling limitations in the voxel size analysis, we only used estimations derived from the
 354 maximum sampling level, i.e. 5 scans design. To unravel the effect of subgrid vegetation
 355 heterogeneity, LAD estimations with 0.5 m voxels were performed both on original LAD_{REF} and on

356 $LAD_{REF\ s0.5}$. By definition, $LAD_{REF\ s0.5}$ field was manipulated (smoothing) so that it does not include
 357 any vegetation heterogeneity below 0.5 m. This result is presented in section 3.3. Then the magnitude
 358 of vegetation heterogeneity effect is characterized in section 3.4 by comparing LAD estimations
 359 retrieved with tested voxel sizes on vegetation resolved at 0.1 m scale (i.e. $\overline{LAD}_{0.1}$ to $\overline{LAD}_{1.0}$
 360 compared to LAD_{REF}).

361 2.5.2 Consequences of sampling limitations

362 To provide insights regarding limitation associated with sampling in this context, we evaluated the
 363 sensitivity of estimates to the number of scans (different scan designs, from a single viewpoint to 5
 364 scans) in sections 3.4. and 3.5. The single scan configuration used the scanner in central position, the
 365 two scans configuration used opposite diagonal design (top left to bottom right), the three scans
 366 configuration used full diagonal design and the four scans configuration used corners positions and
 367 five scan design used all positions.

368 Additionally, in order to perform a cross analysis between voxel size and number of scans, each
 369 scan configuration was used at both 0.10 m and 0.5 m voxel sizes. To disentangle the effect of
 370 sampling from the effect of vegetation heterogeneity (see 2.5.1), we used $LAD_{REF\ s0.5}$ field for LAD
 371 estimations with 0.5 m voxel (i.e. $\overline{LAD}_{s0.5}$), similarly to section 2.5.1. Therefore, differences between
 372 estimates in section 3.5 cannot be attributed to a potential effect of vegetation heterogeneity effect, but
 373 only to sampling. Analyses of estimated vertical profiles were performed similarly to analysis of voxel
 374 size bias. For the computation of the vertical profile, we remind that the unsampled voxels were
 375 estimated by the “mean layer value”. One should notice that in practice, LAD profiles computed from
 376 all voxels (including unsampled voxels estimated with the “mean layer method”) provide identical
 377 profiles to those computed with means based on sampled voxels only.

378
 379 All configurations were tested against references. Evaluations were performed through the comparison
 380 of mean layer \overline{LAD} profile and mean absolute error (MAE). For a given set of voxels or layer i , MAE
 381 was computed according to:

$$382 \quad MAE_i = \frac{|\overline{LAD}_{REF,i} - \overline{LAD}_i|}{\overline{LAD}_{REF,i}} \quad (10)$$

383

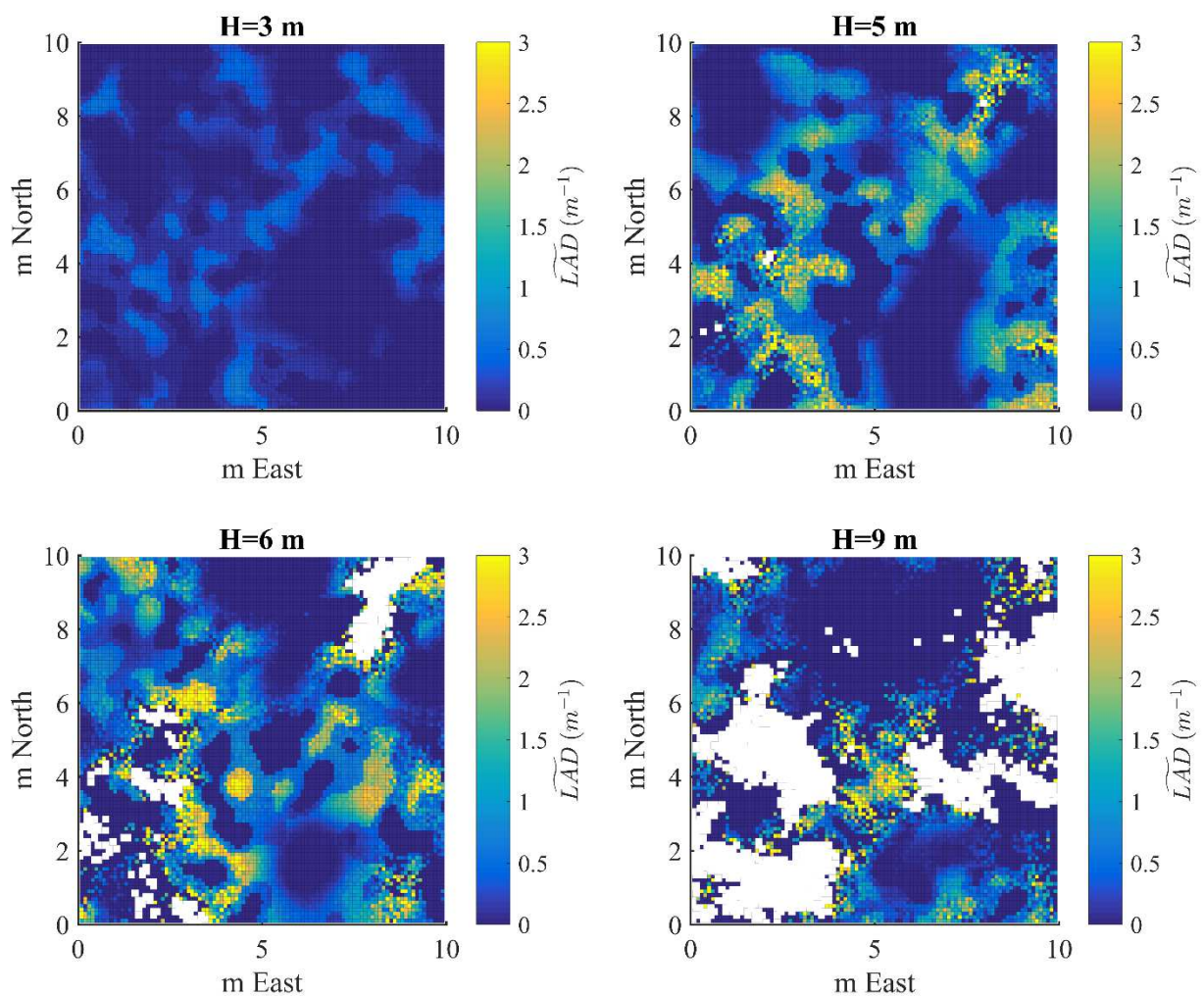
384 Additionally, we analyzed the distribution of 68%-confidence interval radius (Eq. 9) to evaluate
385 efficiency of local estimates for various scans and voxel sizes configurations.

386

387 3. Results

388 3.1. Overview of LAD patterns estimation

389



390

391 **Figure 4.** Estimated horizontal distribution of $\widehat{LAD}_{0.1}$ at four heights. Heights (H) of the mapped layer
392 and associated vegetation covers are indicated in titles. Colors are scaled on \widehat{LAD} values from low to
393 high densities (blue to yellow, respectively). These distributions can directly be compared to subplots

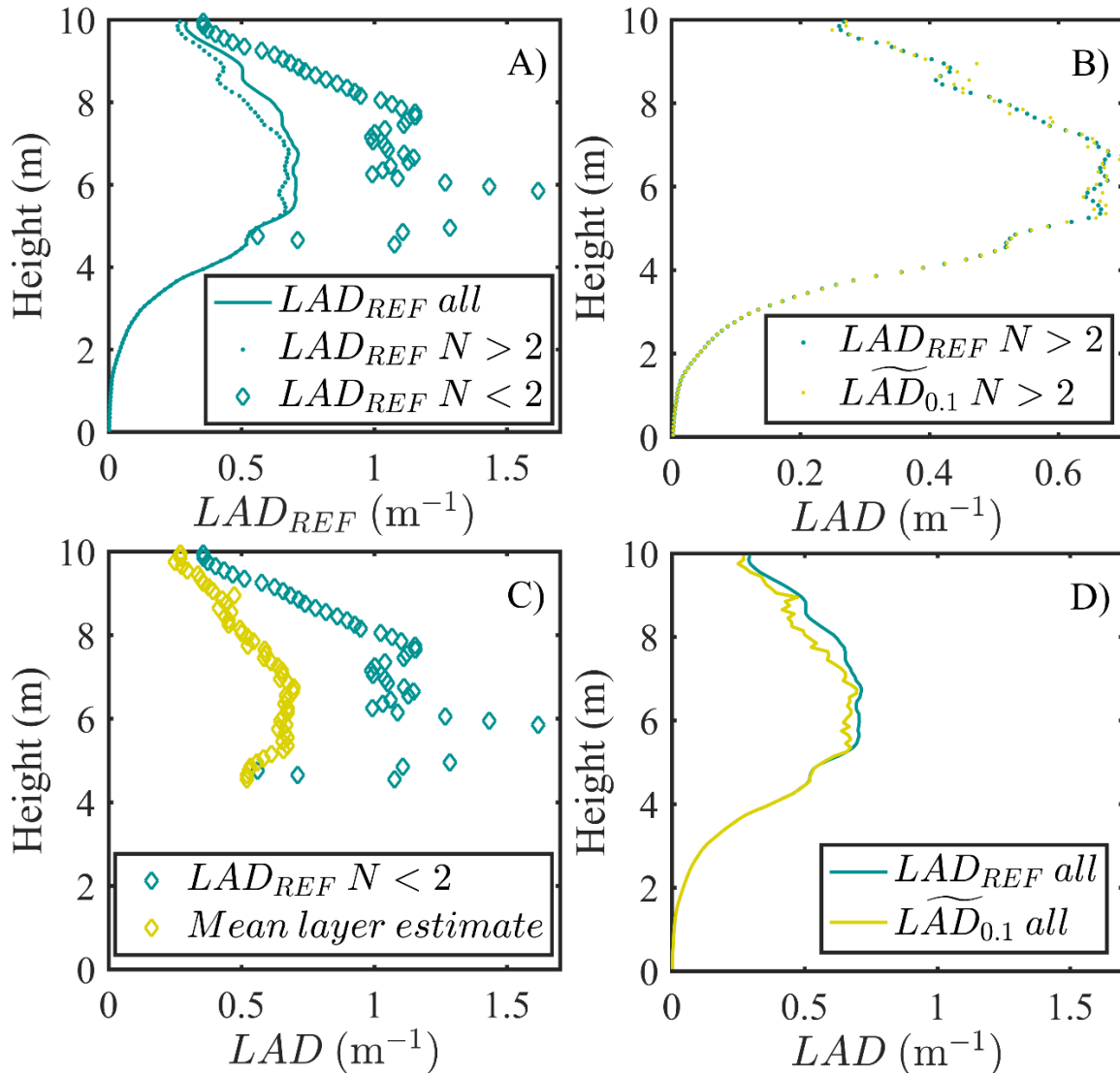
394 in Fig. 1 (LAD_{REF}). Here, the traversal algorithm was applied to the five scans at a 0.10 m voxel size.
395 Blank pixels are unexplored voxels, which reveal occluded locations in the upper part of the canopy.

396

397 Figure 4 shows horizontal distributions of LAD estimations at four heights above ground
398 level, providing insights on the 3D distribution of $\widehat{LAD}_{0.1}$ at plot scale. Comparison with the same
399 plots shown for LAD_{REF} in Figure 2 allowed to identify main discrepancies with the reference field. At
400 low height ($H=3$ m), estimates were very close to the reference for all voxels. From 5 m height and
401 above, the general structure and variability in the LAD distribution was captured, but a few
402 discrepancies can be identified. In particular, horizontal variations in $\widehat{LAD}_{0.1}$ were less smooth than the
403 reference, with inaccuracy resulting from poor sampling. Typically, local differences were identified,
404 mostly visible in dense clumps and a few voxels that were not sampled (in white). At 6 m height, some
405 large blocks of unexplored voxels occurred, most of them being located in dense locations (LAD
406 higher than 2 m^{-1}). In upper canopy ($H=9$ m), large errors occurred (higher than 100%) and unexplored
407 voxels were very frequent (~30% of voxels). These unexplored locations were typically structured in
408 large blocks (until 4-m width), covering both dense and light vegetation density.

409 The following sections provide quantitative analyses of estimations under various configurations.

410



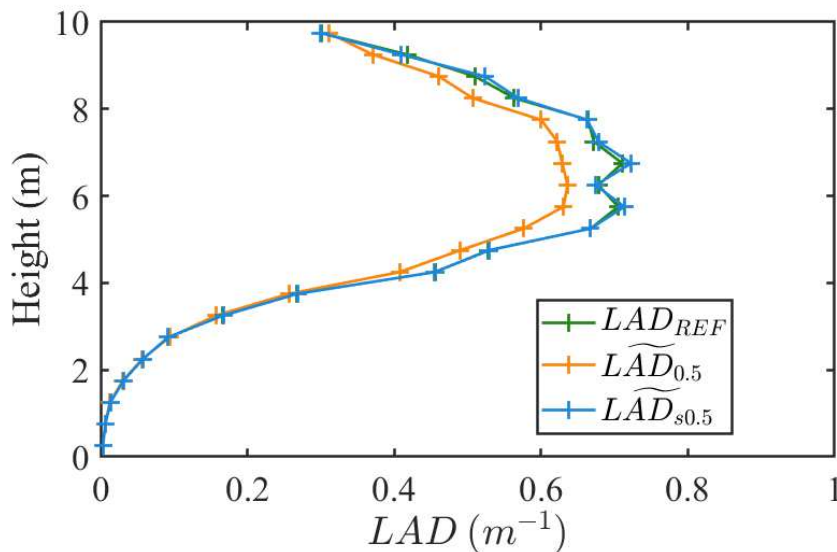
412
 413 **Figure 5.** Comparison of LAD profiles at 0.1 m voxel size with 5 scans. A) Comparison of LAD_{REF}
 414 profiles for all voxels (blue curve), sampled voxels ($N > 2$, blue dots) and unsampled voxels ($N < 2$ (blue
 415 diamonds). B) Comparison of LAD_{REF} and $\widetilde{LAD}_{0.1}$ from sampled voxels ($N > 2$, blue and yellow dots,
 416 respectively). C) Comparison of LAD_{REF} and $\widetilde{LAD}_{0.1}$ from unsampled ($N < 2$, blue and yellow dots,
 417 respectively). D) Comparison of LAD_{REF} and $\widetilde{LAD}_{0.1}$ profiles with all voxels (blue and red yellow,
 418 respectively).

419
 420 In this section, we consider LAD estimations of LAD_{REF} at 0.1 m voxel size. Therefore, we must not
 421 expect any effect of subgrid vegetation heterogeneity on estimations, as this is exactly the resolution

422 of the reference field. Figure 5 compares profiles obtained for references and estimated voxels, which
 423 were respectively explored by less or more than 5 beams. The analysis of reference profiles (**Fig5A**)
 424 allowed to pinpoint sampling quality differences depending on reference LAD values. As expected,
 425 the unsampled voxels at a given height exhibited systematically the highest reference LAD values,
 426 well above the mean LAD_{REF} (**Fig5A**), and were systematically underestimated, because the “mean
 427 layer” value (that was attributed to unsampled voxels) was strictly lower than actual references in
 428 unsampled voxels (**Fig5C**). For explored voxels, the agreement between the reference and the
 429 estimation was very good (**Fig5B**), which confirms the unbiasedness of point estimates used in the
 430 present study (initially proposed by Pimont et al. 2018). This demonstrates that the overall
 431 underestimation (**Fig5D**) was caused by unsampled voxels (even if the “mean layer” value was
 432 assessed to these voxels). This confirms the hypothesis presented in introduction, where we suggested
 433 that the heterogeneity of sampling was likely to induce biases in mean estimates. Moreover, these
 434 differences in LAD_{REF} between sampled and unsampled voxels decreased in the upper part of the
 435 canopy (**Fig5A**). Hence voxels had similar averages near canopy top. This suggests that sampling
 436 restriction was a random process where occlusion was very severe (in the upper part), contrary to
 437 where occlusion began (mid canopy), where occlusion concerned more frequently the denser voxels.

438

439 *3.3. Influence of subgrid vegetation heterogeneity*



440

441 **Figure 6.** Comparison of five scans design profiles of $\overline{LAD}_{0.5}$ and $\overline{LAD}_{s0.5}$ (generated from the
442 smoothed reference $LAD_{REF\ s0.5}$) with the reference LAD_{REF} . Profiles were averaged in 0.5-m layer to
443 ease comparisons, so that a single reference could be used for both estimates (labelled LAD_{REF} for
444 simplicity), as $LAD_{REF\ s0.5}$ and LAD_{REF} were identical at the 0.5 m resolution by definition.

445
446 At coarser voxel sizes, convexity effects associated with subgrid vegetation heterogeneity are expected
447 to affect LAD estimations, also leading to underestimations when, for example, comparing $\overline{LAD}_{0.5}$ to
448 references (Figure 6). However, a part of this negative bias could have been caused by the presence of
449 unsampled voxels. In order to formerly attribute this negative bias to subgrid vegetation heterogeneity,
450 we used the estimates ($\overline{LAD}_{s0.5}$) derived from smoothed reference fields $LAD_{REF\ s0.5}$ (Figure 6),
451 corresponding to a scenario in which no subgrid vegetation heterogeneity can occur. In this case,
452 estimates ($\overline{LAD}_{s0.5}$) were very close to reference profile, demonstrating that the role of the sampling
453 bias was negligible in this scenario. The negative bias observed with $\overline{LAD}_{0.5}$ was hence attributed to
454 subgrid vegetation heterogeneity.

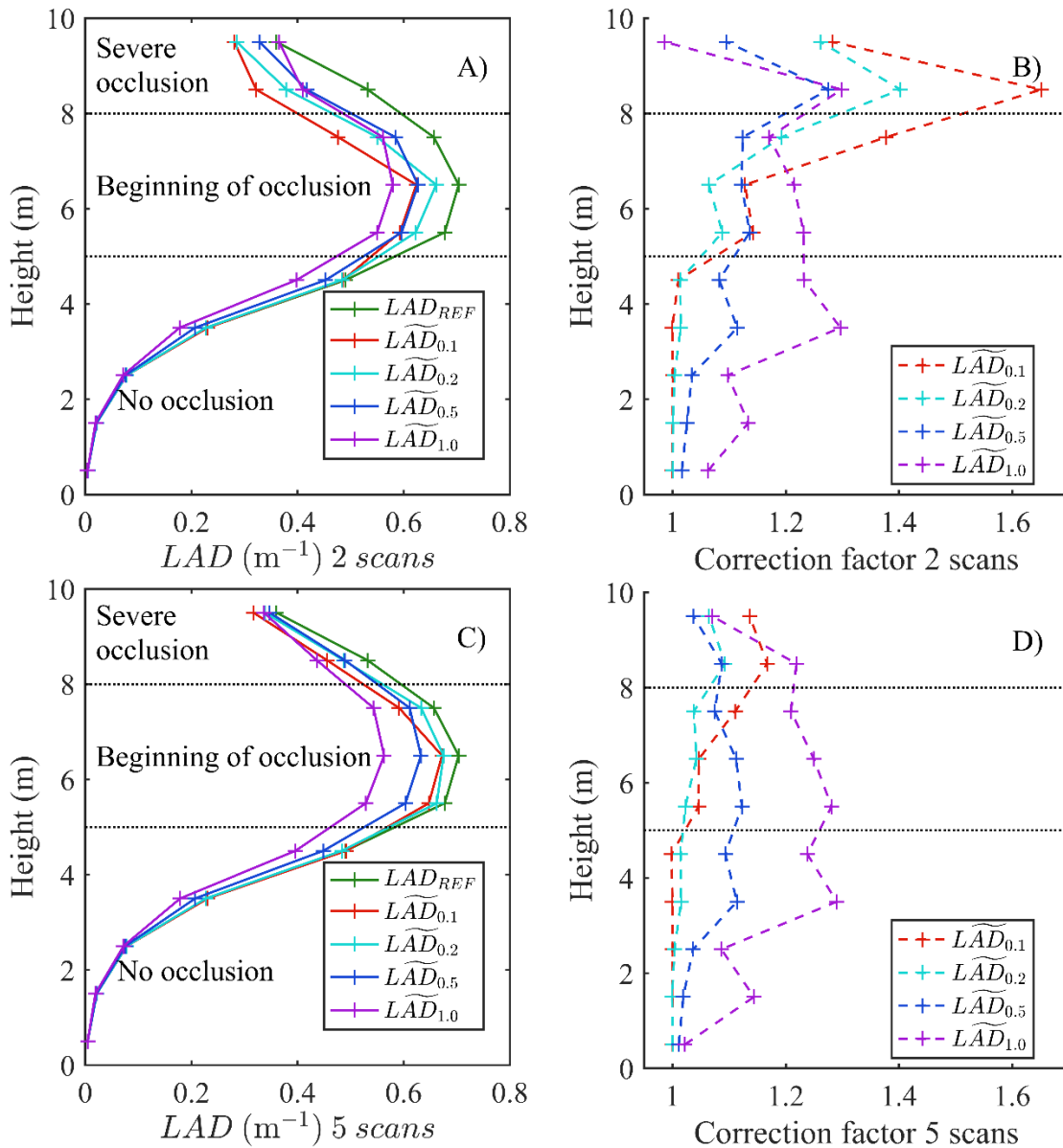
455

456

457

458 3.4. Overall sensitivity of estimation to voxel sizes

459 Because subgrid vegetation heterogeneity and sampling depends on voxel size, the amplitude of
460 negative biases is expected to vary with voxel size, which is investigated in next section.



461

462 **Figure 7.** A) Vertical profiles retrieved with \widetilde{LAD} estimates with various voxel sizes for the two scans

463 design. B) Vertical profiles of the multiplicative correction coefficient required to retrieve reference

464 value ($\frac{LAD_{REF}}{LAD}$) using the two scan design. C) and D) Same as A and B, but for the five scans design. In

465 these plots, LAD were averaged in 1-m layer when voxel size was smaller than 1 m to ease

466 comparison between estimates across voxel sizes. Dotted lines roughly pinpoint height boundaries

467 between layers with no occlusion, layers where occlusion starts to occur (typically within vegetation

468 clumps) and layers with widespread occlusion patterns (both within and outside vegetation clumps).

469

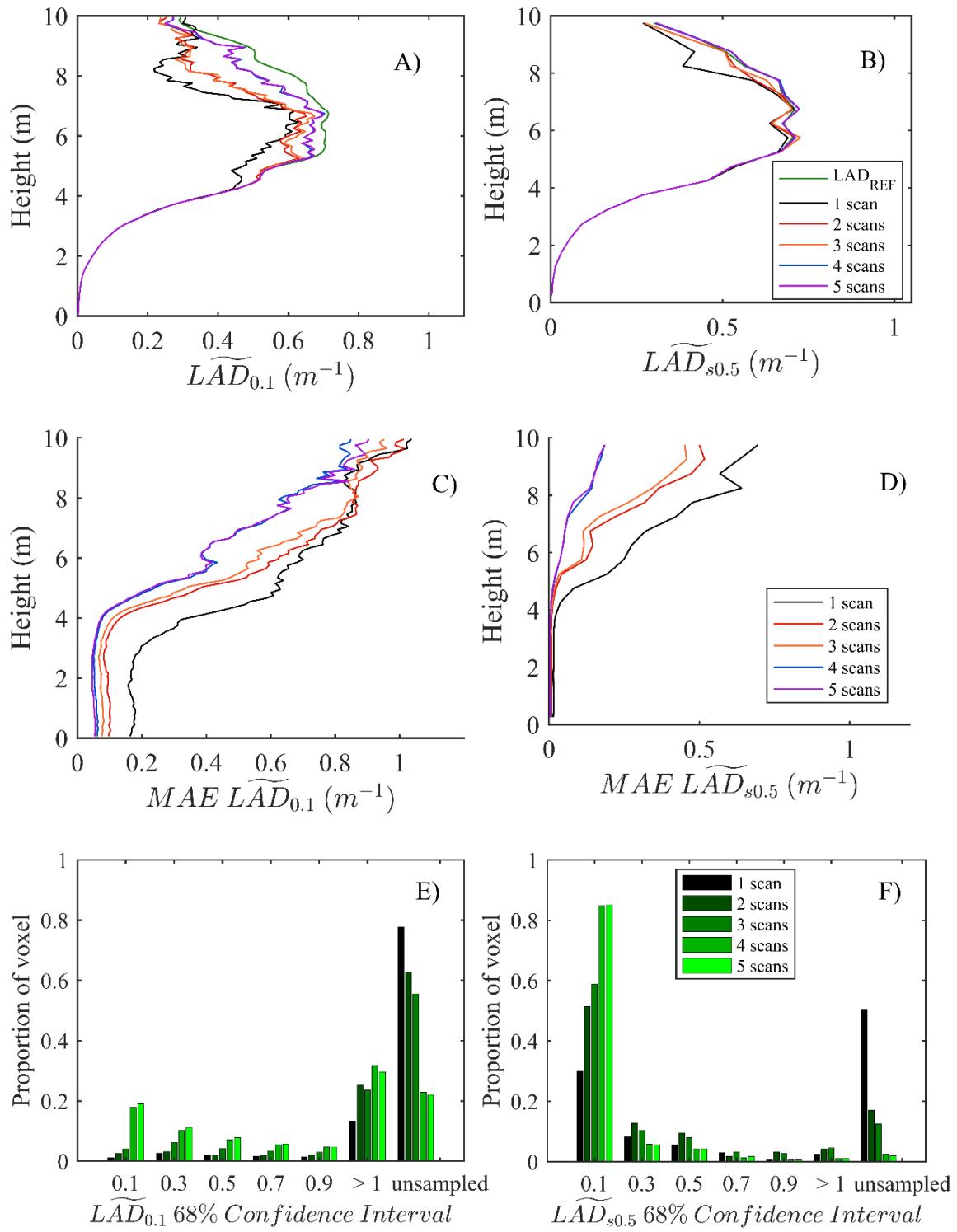
470 In Figure 7A, we show the overall sensitivity of estimates of \overline{LAD} profiles to voxel size,
471 including both sampling and subgrid vegetation heterogeneity biases, for two scans and five scans
472 designs. All voxel sizes adequately captured the shape of the reference profile. However, various
473 amplitudes of underestimations were observed depending on voxel sizes, scan design and height in the
474 canopy. The significance of these effects was confirmed by a statistical analysis (Appendix D). In
475 well explored layers, i.e. below 5 m, coarser voxel sizes led to strong underestimations, requiring
476 highest correction factors for both scan designs. However, as occlusion occurred and increased in
477 highest layers, the two smallest voxel sizes yielded strongest and variable underestimations, of greater
478 magnitude than with the 0.5 m and 1 m voxel sizes (7B). On the contrary, even in mid and top layers,
479 the biases observed with 0.5 m and 1 m voxel sizes were similar to those below 5 m, resulting in a
480 relatively homogenous bias in the whole canopy.

481 According to the results of previous sections 3.2 and 3.3, underestimations resulted from two
482 different mechanisms, namely, sampling bias for small voxels (e.g. 0.1 m, **Fig 5**) and convexity bias
483 associated with subgrid vegetation heterogeneity for coarse voxels (e.g. 0.5 m, **Fig 7**). Our analysis
484 and vegetation manipulation allowed to disentangle and quantify these biases. We found that their
485 magnitude and their spatial distribution was variable, but that bias arising from sampling limitations
486 was very unstable and of stronger magnitude.

487

488 3.5. *Sensitivity of estimates to variations in sampling design (without subgrid heterogeneity*
489 *effect)*

490



491

492 **Figure 8.** A. Vertical profiles of $\widetilde{LAD}_{0.1}$ for 1 to 5 scan design. B. Vertical profiles of $\widetilde{LAD}_{s0.5}$ for 1

493 to 5 scan designs. Profiles of Mean Absolute Error (MAE, in fraction) are presented accordingly for

494 each voxel size for $\widetilde{LAD}_{0.1}$ and $\widetilde{LAD}_{s0.5}$ in subplots C) and D), respectively. MAE were normalized

495 by the mean reference values within each horizontal layer. Distributions of 68% confidence intervals

496 with 1 to 5 scan designs (colored bars from black to green, respectively) for $\overline{LAD}_{0.1}$ and $\overline{LAD}_{0.5}$ in
497 subplots C) and D), respectively. Bins have a 0.2 m^{-1} extent centered on x-axis labels.

498

499 Figure 8. presents \overline{LAD} profiles and corresponding Mean Absolute Error (MAE), for various
500 number of scans (1 to 5 scans design) and two voxel sizes (0.1 m and 0.5 m). Normalized MAE
501 inform on the accuracy of the estimation at voxel scale. For the coarsest resolution, we use the
502 smoothed reference field to eliminate the subgrid vegetation heterogeneity effect.

503 The shape of the reference profile was roughly captured for all scan designs and voxel sizes (**Fig.**
504 **8AB**). Below 8 m, estimations were closer to reference (**Fig. 8AB**) and errors decreased as the voxel
505 size and/or the number of scans increased (**Fig. 8CD**). The negative bias was less pronounced in the
506 upper part of the canopy (above 8 m), as already noticed with large voxels (**Fig 7**), with more
507 randomly distributed occlusion voxels (with regards to voxel density). For example, the bias was the
508 same for 5 and 1 scans for the 0.1m voxel size in the upper part of the canopy (**Fig 8A**). Both 4 and 5
509 scans design yielded relatively close estimates to the reference, but their MAE were very poor in upper
510 canopy layers (~100% error), remaining similar to the single scan design MAE (**Fig 8C**). At 0.5 m
511 voxel size (**Fig 8B**), with the 4 and 5 scan design, the estimated profile (blue and violet lines) closely
512 matched the reference. In this context, coarser voxels exhibit an increased sampling rate with respect
513 to fine voxels (as for multiplying scan positions), leading to significantly better estimated profiles and
514 smaller MAE values (thanks to the suppression of the subgrid vegetation heterogeneity effect).
515 However, the single scan design, and in a lesser extent 2 and 3 scans designs, remained affected by
516 systemic negative bias and large occlusion (drop down of estimations) above 8 m, even with 0.5 m
517 voxels (**Fig 8B**). Conversely to 1-2-3 viewpoints, MAE were much lower than with 0.1 m voxel size
518 and were limited to 0.7 at top height, which showed that the estimated local voxel value was more
519 reliable. Remarkably, the 4-5 scans design had MAE limited to 0.2 at top height, which is consistent
520 with a 20% error on a given voxel in average.

521 Figure **8EF** shows the impact of number of scan and voxel size on the distributions of 68%
522 confidence intervals (CI) in sampled voxels, and on the occurrence of unsampled voxels. Increasing
523 the voxel size and number of scans decreased occurrences of unexplored voxels, as suggested above.

524 In particular, 80% of voxels were unexplored for a single scan at 0.10 m voxel size (i.e. most critical
525 case), implying that estimations relied on only 20% of the vegetation scene. More generally, more
526 than 50% of the voxels are unexplored when three scans or less are used, which highlights that such a
527 0.1 m voxel size is not really practical given TLS resolution and number of scans.

528 For explored voxels, increase in number of scans and voxel sizes tends to decrease CI width,
529 leading to higher frequencies of small errors. In particular, whatever the scan design, increasing the
530 voxel size shifted the occurrence of largest CI (>1) and unexplored voxels (i.e. infinite error) towards
531 the lowest CI values. Typically, the single scan design was almost unable to produce estimations with
532 CI lower than 1 m^{-1} at 0.10 m voxel size, whereas 30% of estimations reached the smallest confidence
533 level with 0.5 m voxel size. Similarly, with the 4-5 scan designs, 20% of the estimates had the smallest
534 confidence level at 0.10 m voxel size, while 85% of voxels are estimated with CI lower than 0.2 m^{-1} at
535 0.5 m voxel size. As a result, at least 50% of the voxels were unexplored with the 0.1 m voxel size,
536 even for 5 scan designs. These points highlight the fact that the sampling used in the present numerical
537 experiment, which was representative of field conditions (a few scans per forestry plot, with 0.036°
538 resolution) was too limited to provide reliable results at the 0.1 m voxel size, contrary to coarser
539 voxels, which on the other hand, raise the problem of subgrid vegetation heterogeneity (**Fig 6, Fig 7**).

540

541 **4. Discussion**

542 The simulation framework used in this study allowed to evaluate the process of estimations of LAD
543 with TLS thanks to exact references. In particular, we analysed the sensitivity to vegetation and
544 sampling heterogeneities. We disentangled combined effects associated with variations in voxel sizes
545 and scan designs with respect to true numerical references, and biases were formerly attributed to two
546 main mechanisms, thanks to our numerical design and the used of unbiased point estimators. Our
547 study confirmed the effect associated with subgrid vegetation heterogeneity, which leads to
548 underestimation when voxels are too coarse to explicitly account for it. This effect is discussed in
549 section 4.1. Our study also demonstrates that an overall negative bias can arise when relying on mean
550 layer estimates for correcting for the presence of unsampled voxels. The bias was not homogeneously

551 distributed within the vegetation scene and typically occurs when occlusion begins to be significant (at
552 mid-canopy height in the present example), as it more frequently affects clumps of dense voxels, but
553 can diminish in highest vegetation layer, despite strong occlusion. This effect is discussed in section
554 4.2. Beyond overall biases in vertical profiles, we also studied the sensitivity of voxel-scale
555 predictions to voxel size and scan design, as the objective of voxelized TLS method is to measure the
556 tri-dimensional distribution of the LAD. The distribution of confidence intervals showed that the voxel
557 sizes required for a reasonable accuracy (here typically 0.5m), could be larger than the voxel size
558 required to mitigate subgrid vegetation heterogeneity. The discussion ends with recommendations and
559 future work.

560

561 *4.1. Subgrid vegetation heterogeneity should be accounted for to provide reliable* 562 *estimations*

563

564 In the present study, underestimations occurred when voxel size increased (i.e. 0.5 m, **Fig. 6**),
565 and the negative bias of the estimated LAD profile (i.e. for a horizontal layer of voxels) could reach
566 30% for 1 m voxel size (**Fig. 7**). We found that such underestimation did not occur when point clouds
567 were generated with vegetation that was smoothed at voxel size ($\widetilde{\text{LAD}}_{s0.5}$, **Fig. 6**), which
568 demonstrated that this effect was caused by subgrid vegetation heterogeneity. This confirms prior
569 observations on natural branches of three different species scanned with two different instruments
570 (Soma et al., 2018), but also on trees (Béland et al., 2014a), which already attributed underestimations
571 to vegetation heterogeneity. One can notice that the effect of voxel size observed in Soma et al. (2018)
572 was more important, requiring a 1.7 correction factor with 0.7 m voxels compared to 1.3 for 0.5 m
573 voxels in the present simulations. Such differences could arise from more heterogeneous vegetation
574 structure in natural vegetation than in our example reference field, in which heterogeneity was
575 resolved at 0.1 m scale. However, these findings demonstrate the efficiency of numerical experiments
576 for testing hypothesis and evidence the associated mechanism (here heterogeneity effect on LAD

577 estimation), since it was possible to numerically manipulate vegetation (here smoothing), for a formal
578 attribution of bias source.

579 Similarly to approaches developed for passive optical measurements (e.g. hemispherical
580 photography), identifying canopy gaps and angular segmentation the scene is efficient to partially
581 correct for clumping (Yan et al., 2019). Yet, the subgrid vegetation heterogeneity in each segment of
582 2D images remains along the optical depth, i.e. in the third dimension. Our results illustrate the
583 potential of an appropriate segmentation level to limit the clumping effect, making TLS voxel-based
584 approaches more efficient compared to 2D gap fraction methods for the separation of gaps and crowns
585 and particularly fitted for application of LAI retrieval approaches, including path length distribution
586 models whether at crown (Hu et al., 2018) or voxel scales (Pimont et al., 2019; Soma et al., 2018).

587

588 4.2. *Sampling limitations generate heterogeneous and variable underestimations*

589

590 When occlusion was limited (typically 5 scans and 0.5 m voxel size), the clumping effect was
591 the main source of underestimation in LAD. However, when sampling limitation became significant,
592 typically above 5 m and with 0.1 m voxels ($\widetilde{\text{LAD}}_{0.1}$, **Fig 7AB** – red curve), underestimations of
593 heterogeneous magnitudes occurred without any subgrid heterogeneity, demonstrating the presence of
594 a second mechanism of underestimation. This negative bias arose from the spatial correlations
595 between sampling limitations and LAD densities, as clumps of dense volumes were more frequently
596 occluded. Indeed, scan positions located in the open ease the sampling of empty or light voxels and
597 might amplify the oversampling of light voxels described above, while dense voxel are rarely explored
598 and less accounted for in mean layer LAD calculation. This spatial structure of occlusion can result in
599 up to 25% underestimations (**Fig. 8A**) of the average LAD profile despite 5 scans are performed. In
600 these non-sampled volumes, we assigned the mean layer estimated values, which results from sampled
601 voxels in dense layers, i.e. mostly empty and low density voxels, generally equal to 0.4 m^{-1} in average,
602 i.e. much less than dense voxel LAD values. As less and less beams coming from below canopy
603 remained available in the canopy middle, clumps of dense occluded voxels are more frequent, and

604 such underestimation mean bias was worse until 8 m height. Above 9 m, this bias was dampened (**Fig.**
605 **8AB**), because the correlation between dense voxels and occluded area is much less pronounced. In
606 other words, large spots of unexplored voxels in top-canopy affected almost randomly both empty and
607 dense above volumes (**Fig. 4**). As a result, such strong exposure to occlusion led to underestimations
608 of significant but very heterogeneous magnitudes along vertical profiles (Fig 7B). It suggests that such
609 underestimations depend on both the canopy structure and the degree of occlusion, which are both
610 hardly predictable. On the contrary, coarser voxel sizes were much less affected by occlusion, and
611 were rather affected by a significant but spatially homogenous negative bias attributed to clumping
612 effect. In these conditions, it seems a better strategy to avoid erratic negative biases resulting from the
613 strong occlusion inherent to small voxel size by using coarser voxel sizes corrected for a more
614 constant and measurable bias and less prone to large occlusion spots.

615 However, independently of the voxel size, such a negative bias is expected to systematically
616 occur in layers where occlusion starts, as soon as the dense voxels are clumped. Thus the “mean layer”
617 method for dealing with unsampled voxels is not satisfying. Additionally, it raises sharp errors locally.

618 Strong underestimations at layer and plot scales are in accordance with previous observations
619 from ground perspectives (Schneider et al., 2019). Increasing the number of scans and the voxel size
620 limited occurrence of unexplored voxels, so that estimations were generally improved. However,
621 highest number of scans with smallest voxel size or single scan with 0.5 m voxel size were still
622 affected. Hence, the appropriate combination of scan design and voxel size directly depends of the
623 vegetation density and of the level of detail required to match user’s aim. Typically, retrieving LAD
624 profiles shapes and values with 20% error can be achieved either by 5 scans with 0.1 m voxels or by a
625 single scan with 0.5 m voxels in sparse vegetation structures (**Fig 8AB**). However, in dense vegetation
626 and/or when estimates at voxel scale are valuable (i.e. for 3D radiative transfer modelling), it seems
627 mandatory to rely on both multiple viewpoints and coarse voxels to retrieve unbiased estimates and
628 reliable voxel LAD value (**Fig. 8CD**). Similarly to (Schneider et al., 2019), until 80% of voxels can be
629 occluded in critical cases (**Fig. 8EF**).

630 By calculating confidence intervals at voxel scale, our study gives insights on the drop in local
631 accuracy within voxel resulting from poor sampling conditions. This differs from the usual approach

632 for quantifying occlusion consisting in counting voxel with at least one beam passing through, thus
633 ignoring errors generated in poorly sampled voxels. Confidence intervals indicated that estimations in
634 poorly sampled voxel can be affected by 1 m^{-1} errors, which is considerable regarding the range of
635 reference LAD values. These confidence intervals can be considered as a statistical alternative to the
636 use of mean nearest neighbour point (Wilkes et al., 2017) in order to assess the sampling densities of
637 given areas, in particular when used approaches are sensitive to non-uniform distribution of points (i.e.
638 for geometrical modelling of trees). Small voxels were frequently affected by large errors, until 35%
639 of the scene. With the 4-5 scans design with 0.1 m voxels, $\text{CI} > 1 \text{ m}^{-1}$ are even more frequent than with
640 a single scan (**Fig 8E**). This higher frequency of large CI was due to the poor exploration of voxels
641 which would have remained unexplored with a single scan, producing unreliable estimates. Locally,
642 such errors were worse than replacing estimates by the mean layer value, suggesting that the reduction
643 of occluded voxel percentage should not be the single criteria for sampling accuracy, and raised the
644 question of the appropriate criterion for occlusion. Determining whether to replace voxel by a fixed
645 value or to keep the estimation given by few beams is particularly difficult in clumped situations given
646 the strong correlation observed between occlusion areas and high LAD values areas. This advocates
647 for the development of correction techniques that are not based on numerical thresholds (Soma et al.
648 2020).

649

650 *4.3. Recommendations and research needs*

651

652 All identified estimation biases were negative. The vegetation heterogeneity effect could yield
653 up to 20% underestimations (**Fig. 6, Fig. 7BD**), while unsampled voxels led to profiles that could be
654 locally less than half of reference (**Fig. 8A**), and even worst locally. In the field, these negative biases
655 are entangled and add up, which can potentially induce a global stronger bias than observed in these
656 simulations. Hence, these biases should not be overlooked and we strongly recommend to limit their
657 magnitude and correct them when possible. This study confirms observations of Béland et al., 2014a
658 at the scale of individual trees, i.e. increasing voxel size is an efficient way to level up sampling at
659 voxel scale. Such coarser discretization level should be considered with caution, as it leads to the

660 aggregation of gaps and clumps, as well as the aggregation of occluded and explored areas in same
661 voxels. However, negative bias resulting from occurrence of occlusion are unstable and of higher
662 magnitude and should be avoided in first place, suggesting coarse voxel sizes remain an advantageous
663 option. To determine the appropriate voxel size, the ability to correct for subgrid vegetation
664 heterogeneity (Soma et al., 2018) and to assess accuracy with confidence intervals are key steps.

665 In light of our results, we recommend to use unbiased estimators and to rely on multiple
666 viewpoints. Tested vegetation scene had a 10 m extent in each dimension, with no or few vegetation
667 below 3 m height. The density of scans was relatively high (2.5 m between scans) compared to other
668 studies (Schneider et al., 2019; Wilkes et al., 2017). However, the simulated vegetation features,
669 which were representative of Mediterranean canopies (Simioni et al., 2016), displayed much denser
670 clumps of LAD values (0.6 m^{-1} in average) compared to retrieved LAD in other vegetation types
671 ($\sim 0.3 \text{ m}^{-1}$). In such dense context, occlusion occurred from 5 m height, while accurate sampling has
672 been observed until 35-m height with multi-echoes instrument and in other field conditions (Schneider
673 et al., 2019). This proves that *a priori* recommendations concerning visibility analysis can hardly be
674 applied to a range of vegetation structures and different forest types. Providing exhaustive guidelines
675 based on easily measurable stand parameters, i.e. number of large stems (Abegg et al., 2017), presence
676 of understorey or mean crown width could help to plan appropriate scan design across sites.
677 Additionally, the terrestrial LiDAR simulated in this study was a single echo scanner, typically
678 representative of the FARO Focus 3D, which is used in operational context for French forest
679 monitoring. Other scanners, such as RIEGL VZ400, a multi-echoes LiDAR, have better exploration
680 skills and are less sensitive to occlusion. Then, as suggested in (Wilkes et al., 2017), recommendations
681 should also be adapted to the type and resolution of instruments. Yet, beam density also decreases at a
682 given distance of these multi-echoes instruments, so that such scanners are theoretically exposed to the
683 same bias arising from occlusion, as confirmed by observations in tropical canopies (Schneider et al.,
684 2019). The bias arising from subgrid vegetation heterogeneity is independent of sampling accuracy
685 and, as a result, concerns all instruments. The presence of vegetation in the lower part of the canopy
686 (understorey) might generate more occlusion, requiring to increase the number of scans. Hence, a
687 particular attention should be paid to set instrument in order to obtain viewpoints that enable to

688 efficient plot sampling, for example in open locations (Wilkes et al., 2017). Since the number of scans
689 can generally not be increased in field conditions, a trade off on voxel size must be found. Therefore,
690 we recommend to use medium voxel size, typically close to 0.5 m, with an empirical correction to
691 account for vegetation heterogeneity in the subgrid volume. Scanning from the above perspective (i.e.
692 UAV LiDAR) can provide significant reduction of occluded areas close to tops of canopies, but
693 suffered from a limited penetration of beams in the mid-crown (Morsdorf et al., 2018; Schneider et al.,
694 2019), so that similar negative biases associated with unsampled voxels and subgrid vegetation
695 heterogeneity as identified in the present study should exist.

696 In the present study, we used the mean layer assumption (replacing unsampled voxels by the
697 mean of LAD estimates obtained at corresponding layer) to extrapolate missing data. Such method is
698 theoretically similar to ignoring these missing values in spatial averages, and suffer from strong local
699 inaccuracy in dense or empty areas differing from mean LAD value. Additionally, poorly sampled
700 voxels had frequent unreliable estimates, which might be locally more erroneous than replacing by the
701 biased mean layer value. First, it shows the benefit of quantifying uncertainties of estimations in
702 locations of interests (i.e. dense voxels), which should be included in further analysis when estimating
703 LAD. Second, it raises the question of the definition of an occluded voxel.

704 In our numerical experiment, the leaf orientation factor, the leaf fraction, and the distance
705 effect were known and corrected so that our estimators of the LAD were unbiased in each voxel to
706 limit confounding errors in the numerical experiment. However, such parameters are challenging to
707 retrieve in field conditions and their estimations are expected to induce errors (Yan et al., 2019), which
708 might affect the outcome of the present study. The leaf orientation factor can be for example retrieved
709 by leaves triangulations if LiDAR resolution is fine enough (Bailey and Mahaffee, 2017). Several
710 classification methods have been proposed to classify leaves *versus* wood hits (Wang et al., 2020).
711 Then, the leaf fraction factor can be retrieved as the ratio between leaves and total number of points in
712 a given voxel. Eventually, the effect of the distance to scanner can be either estimated from laboratory
713 experiments (Soma et al. 2018) or physical modelling of beam divergence (Béland et al., 2011). The
714 occlusion bias highlighted in the present work should also be explored in other vegetation scenes with
715 absolute references, e.g. with point clouds arising from the sampling of 3D tree models. We showed

716 that underestimations constantly affected dense voxels where occlusion begin, but that voxels can be
717 affected independently of their vegetation density in background and/or top layers, including empty
718 areas. The consequences of this sampling bias in those partially explored layers directly depend on the
719 representativeness of remaining sampled voxels, which can either results in top profile
720 overestimations, underestimations or even in a limited bias. Then, we recommend to analyse the
721 magnitude of this bias for a variety of vegetation structures in light of optical depth and correlation
722 length in vegetation (Appendix B). This variability in spatial bias distribution, vegetation structure and
723 the difficulty to define an occluded volume demonstrate the need for spatially adaptive methods to
724 deal both with poorly sampled and fully occluded voxels. Given the clumped structure of vegetation,
725 using the information of surrounding voxels is a promising approach. Such method relying on kriging
726 has been developed and tested both on a virtual scene (e.g. in Appendix B), and then applied to field
727 data (Soma et al., 2020). Corrections yielded encouraging results with a buffering of errors due to
728 sampling limitations and a more adaptive guess of unexplored voxels values, holding the potential to
729 retrieve better voxel and plot LAD estimations.

730 **5. Conclusion**

731

732 The present work disentangles some of biases related to LAD estimations with TLS through a
733 numerical study. This simulation framework was mandatory to fully understand and quantify the
734 magnitude of these major biases, because they mainly occur in dense and clumped canopy context, for
735 which field references are often not available. Moreover, such a design allows vegetation
736 manipulation, such as smoothing to ease analysis. First, simulations allowed to characterize a negative
737 bias for LAD estimation due to subgrid vegetation heterogeneity in voxels, arising from dense and
738 light density areas being encompassed within same volumes. Second, we assessed sampling accuracy
739 with confidence intervals and identified a spatial correlation between occlusion and dense LAD in
740 layers where occlusion of the laser beams starts to occur. This spatial correlation leads to
741 underestimation of LAD profile, even when the mean layer value was attributed to unsampled voxels.
742 Several methods to limit occurrence of unsampled voxels and increase sampling rate were efficient to

743 damper underestimations. We confirm that combining the information from several scan positions was
744 critical to improve the reliability of estimates. Moreover, the negative bias arising from unsampled
745 voxels was prone to large and unstable variations of magnitude depending on the canopy structure.
746 Then, we recommend to solve in first place the occurrence of unsampled areas, typically by using
747 estimations from coarse voxels complemented by an adequate treatment of subgrid vegetation
748 heterogeneity effect. In the present study, we found that at least 0.5 m voxel size was appropriate, with
749 a mandatory correction for vegetation heterogeneity. The development of methods to extrapolate or
750 correct LAD values in occluded volumes (e.g. kriging) remains a challenging but critical need. The
751 simulation framework developed in our study cannot serve as sole reference for quantifying the
752 magnitude of similar biases occurring with the diversity of natural vegetation and acquisition
753 parameters, as many uncertainties that are inherent to the field conditions are neglected in numerical
754 experiments. Yet, this study paves the way for additional research and field measures needed to
755 characterize in operational context the biases identified in these simulations.

756

757 **Author Contributions**

758 M.S. conceived and designed the experiments; M.S. performed the experiments, as well as the
759 analyses, with useful guidelines of F.P. and J-L.D. . M.S. wrote the paper, with critical contributions
760 of the other co-authors.

761

762 **Funding**

763 This research was funded by Institut National de la Recherche Agronomique (INRA) and by Conseil
764 Régional Provence-Alpes-Côte d'Azur (LiDARForFuel, grant n° APR-EX 2014_04163 and PhD grant
765 n° 2015_07468).

766

767 **6. References**

768 Abegg, M., Kükenbrink, D., Zell, J., Schaepman, M.E., Morsdorf, F., 2017. Terrestrial laser
769 scanning for forest inventories-tree diameter distribution and scanner location impact on

770 occlusion. *Forests* 8. <https://doi.org/10.3390/f8060184>

771 Bailey, B.N., Mahaffee, W.F., 2017. Rapid measurement of the three-dimensional distribution
772 of leaf orientation and the leaf angle probability density function using terrestrial LiDAR
773 scanning. *Remote Sens. Environ.* 194, 63–76. <https://doi.org/10.1016/j.rse.2017.03.011>

774 Béland, M., Baldocchi, D.D., Widlowski, J.L., Fournier, R.A., Verstraete, M.M., 2014a. On
775 seeing the wood from the leaves and the role of voxel size in determining leaf area
776 distribution of forests with terrestrial LiDAR. *Agric. For. Meteorol.* 184, 82–97.
777 <https://doi.org/10.1016/j.agrformet.2013.09.005>

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

781 Béland, M., Widlowski, J.L., Fournier, R.A., Cote, J.F., Verstraete, M.M., 2011. Estimating
782 leaf area distribution in savanna trees from terrestrial LiDAR measurements. *Agric. For.*
783 *Meteorol.* 151, 1252–1266. <https://doi.org/10.1016/j.agrformet.2011.05.004>

784 Cifuentes, R., Van der Zande, D., Farifteh, J., Salas, C., Coppin, P., 2014. Effects of voxel
785 size and sampling setup on the estimation of forest canopy gap fraction from terrestrial
786 laser scanning data. *Agric. For. Meteorol.* 194, 230–240.
787 <https://doi.org/10.1016/j.agrformet.2014.04.013>

788 Côté, J.F., Fournier, R.A., Egli, R., 2011. An architectural model of trees to estimate forest
789 structural attributes using terrestrial LiDAR. *Environ. Model. Softw.* 26, 761–777.
790 <https://doi.org/10.1016/j.envsoft.2010.12.008>

791 Durrieu, S., Allouis, T., Fournier, R., Véga, C., Albrech, L., 2008. Spatial quantification of
792 vegetation density from terrestrial laser scanner data for characterization of 3D forest
793 structure at plot level, in: *SilviLaser 2008*. pp. 325–334.

794 Gollob, C., Ritter, T., Wassermann, C., Nothdurft, A., 2019. Influence of scanner position and

795 plot size on the accuracy of tree detection and diameter estimation using terrestrial laser
796 scanning on forest inventory plots. *Remote Sens.* 11, 1–30.
797 <https://doi.org/10.3390/rs11131602>

798 Grau, E., Durrieu, S., Fournier, R., Gastellu-etchegorry, J., Yin, T., 2017. Estimation of 3D
799 vegetation density with Terrestrial Laser Scanning data using voxels . A sensitivity
800 analysis of influencing parameters. *Remote Sens. Environ.* 191, 373–388.
801 <https://doi.org/10.1016/j.rse.2017.01.032>

802 Hosoi, F., Nakai, Y., Omasa, K., 2010. Estimation and error analysis of woody canopy leaf
803 area density profiles using 3-d airborne and ground-based scanning lidar remote-sensing
804 techniques. *IEEE Trans. Geosci. Remote Sens.* 48, 2215–2223.
805 <https://doi.org/10.1109/TGRS.2009.2038372>

806 Hosoi, F., Omasa, K., 2006. Voxel-based 3-D modeling of individual trees for estimating leaf
807 area density using high-resolution portable scanning lidar. *IEEE Trans. Geosci. Remote*
808 *Sens.* 44, 3610–3618. <https://doi.org/10.1109/TGRS.2006.881743>

809 Hu, R., Bournez, E., Cheng, S., Jiang, H., Nerry, F., Landes, T., Saudreau, M., Kastendeuch,
810 P., Najjar, G., Colin, J., Yan, G., 2018. Estimating the leaf area of an individual tree in
811 urban areas using terrestrial laser scanner and path length distribution model. *ISPRS J.*
812 *Photogramm. Remote Sens.* 144, 357–368.
813 <https://doi.org/10.1016/j.isprsjprs.2018.07.015>

814 Huang, P., Pretzsch, H., 2010. Using terrestrial laser scanner for estimating leaf areas of
815 individual trees in a conifer forest. *Trees - Struct. Funct.* 24, 609–619.
816 <https://doi.org/10.1007/s00468-010-0431-z>

817 Kamoske, A.G., Dahlin, K.M., Stark, S.C., Serbin, S.P., 2019. Leaf area density from airborne
818 LiDAR: Comparing sensors and resolutions in a temperate broadleaf forest ecosystem.
819 *For. Ecol. Manage.* 433, 364–375.

820 <https://doi.org/https://doi.org/10.1016/j.foreco.2018.11.017>

821 Li, Y., Guo, Q., Tao, S., Zheng, G., Zhao, K., Xue, B., Su, Y., 2016. Derivation, Validation,
822 and Sensitivity Analysis of Terrestrial Laser Scanning-Based Leaf Area Index. *Can. J.*
823 *Remote Sens.* 42, 719–729. <https://doi.org/10.1080/07038992.2016.1220829>

824 Morsdorf, F., Kükenbrink, D., Schneider, F.D., Abegg, M., Schaepman, M.E., 2018. Close-
825 range laser scanning in forests: Towards physically based semantics across scales.
826 *Interface Focus* 8. <https://doi.org/10.1098/rsfs.2017.0046>

827 Norman, J.M., Campbell, G.S., 1989. Canopy structure, in: Percy, R.W., Ehleringer, J.R.,
828 Mooney, H.A., Rundel, P.W. (Eds.), *Plant Physiological Ecology: Field Methods and*
829 *Instrumentation*. Springer Netherlands, Dordrecht, pp. 301–325.
830 https://doi.org/10.1007/978-94-009-2221-1_14

831 Pimont, F., Allard, D., Soma, M., Dupuy, J.-L.J.-L., 2018. Estimators and confidence
832 intervals for plant area density at voxel scale with T-LiDAR. *Remote Sens. Environ.*
833 215, 343–370. <https://doi.org/10.1016/j.rse.2018.06.024>

834 Pimont, F., Dupuy, J.L., Caraglio, Y., Morvan, D., 2009. Effect of vegetation heterogeneity
835 on radiative transfer in forest fires. *Int. J. Wildl. Fire* 18, 536–553.
836 <https://doi.org/10.1071/WF07115>

837 Pimont, F., Soma, M., Dupuy, J.-L., 2019. Accounting for Wood, Foliage Properties, and
838 Laser Effective Footprint in Estimations of Leaf Area Density from Multiview-LiDAR
839 Data. *Remote Sens.* 11, 1580. <https://doi.org/10.3390/rs11131580>

840 Pimont, F., Soma, M., Dupuy, J.-L., Rigolot, E., Jean, F., 2016. Estimating canopy load and
841 bulk density distribution using calibrated T-LiDAR indices, in: *Proceedings for the 5th*
842 *International Fire Behavior and Fuels Conference*. International Association of Wildland
843 *Fire*, Missoula, Montana, USA, Portland, Oregon, USA, pp. 3–8.

844 Ruel, J.J., Ayres, M.P., 1999. Jensen ’ s inequality predicts effects of environmental variation.

845 Tree 5347, 361–366. [https://doi.org/10.1016/S0169-5347\(99\)01664-X](https://doi.org/10.1016/S0169-5347(99)01664-X)

846 Schlather, M., Malinowski, A., Menck, P.J., Oesting, M., Storkorb, K., 2015. Analysis,
847 Simulation and Prediction of Multivariate Random Fields with Package RandomFields.
848 J. Stat. Softw. 63. <https://doi.org/10.18637/jss.v063.i08>

849 Schneider, F.D., Kükenbrink, D., Schaepman, M.E., Schimel, D.S., Morsdorf, F., 2019.
850 Quantifying 3D structure and occlusion in dense tropical and temperate forests using
851 close-range LiDAR. Agric. For. Meteorol. 268, 249–257.
852 <https://doi.org/10.1016/j.agrformet.2019.01.033>

853 Simioni, G., Marie, G., Huc, R., 2016. Influence of vegetation spatial structure on growth and
854 water fluxes of a mixed forest: Results from the NOTG 3D model. Ecol. Modell. 328,
855 119–135. <https://doi.org/10.1016/j.ecolmodel.2016.02.004>

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

859 Soma, M., Pimont, F., Durrieu, S., Dupuy, J.-L.J.-L., 2018. Enhanced Measurements of Leaf
860 Area Density with T-LiDAR: Evaluating and Calibrating the Effects of Vegetation
861 Heterogeneity and Scanner Properties. Remote Sens. 10, 1580.
862 <https://doi.org/10.3390/rs10101580>

863 Stark, S.C., Leitold, V., Wu, J.L., Hunter, M.O., de Castilho, C. V, Costa, F.R.C., McMahon,
864 S.M., Parker, G.G., Shimabukuro, M.T., Lefsky, M.A., Keller, M., Alves, L.F., Schiatti,
865 J., Shimabukuro, Y.E., Brandão, D.O., Woodcock, T.K., Higuchi, N., de Camargo, P.B.,
866 de Oliveira, R.C., Saleska, S.R., 2012. Amazon forest carbon dynamics predicted by
867 profiles of canopy leaf area and light environment. Ecol. Lett. 15, 1406–1414.
868 <https://doi.org/10.1111/j.1461-0248.2012.01864.x>

869 Wang, D., Momo Takoudjou, S., Casella, E., 2020. LeWoS: A universal leaf-wood

870 classification method to facilitate the 3D modelling of large tropical trees using
871 terrestrial LiDAR. *Methods Ecol. Evol.* 11, 376–389. [https://doi.org/10.1111/2041-](https://doi.org/10.1111/2041-210X.13342)
872 [210X.13342](https://doi.org/10.1111/2041-210X.13342)

873 Weiss, M., Baret, F., Smith, G.J., Jonckheere, I., Coppin, P., 2004. Review of methods for in
874 situ leaf area index (LAI) determination Part II. Estimation of LAI, errors and sampling.
875 *Agric. For. Meteorol.* 121, 37–53. <https://doi.org/10.1016/j.agrformet.2003.08.001>

876 Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., Gonzalez de Tanago, J., Bartholomeus,
877 H., Brede, B., Herold, M., 2017. Data acquisition considerations for Terrestrial Laser
878 Scanning of forest plots. *Remote Sens. Environ.* 196, 140–153.
879 <https://doi.org/10.1016/j.rse.2017.04.030>

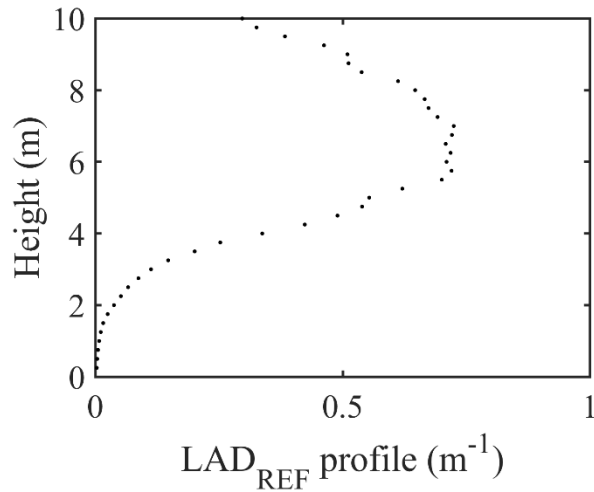
880 Yan, G., Hu, R., Luo, J., Weiss, M., Jiang, H., Mu, X., Xie, D., Zhang, W., 2019. Review of
881 indirect optical measurements of leaf area index: Recent advances, challenges, and
882 perspectives. *Agric. For. Meteorol.* 265, 390–411.
883 <https://doi.org/10.1016/j.agrformet.2018.11.033>

884 Yun, T., Cao, L., An, F., Chen, B., Xue, L., Li, W., Pincebourde, S., Smith, M.J., Eichhorn,
885 M.P., 2019. Simulation of multi-platform LiDAR for assessing total leaf area in tree
886 crowns. *Agric. For. Meteorol.* 276–277, 107610.
887 <https://doi.org/10.1016/j.agrformet.2019.06.009>

888 Zhao, K., García, M., Liu, S., Guo, Q., Chen, G., Zhang, X., Zhou, Y., Meng, X., 2015.
889 Terrestrial lidar remote sensing of forests: Maximum likelihood estimates of canopy
890 profile, leaf area index, and leaf angle distribution. *Agric. For. Meteorol.* 209–210, 100–
891 113. <https://doi.org/10.1016/j.agrformet.2015.03.008>

892

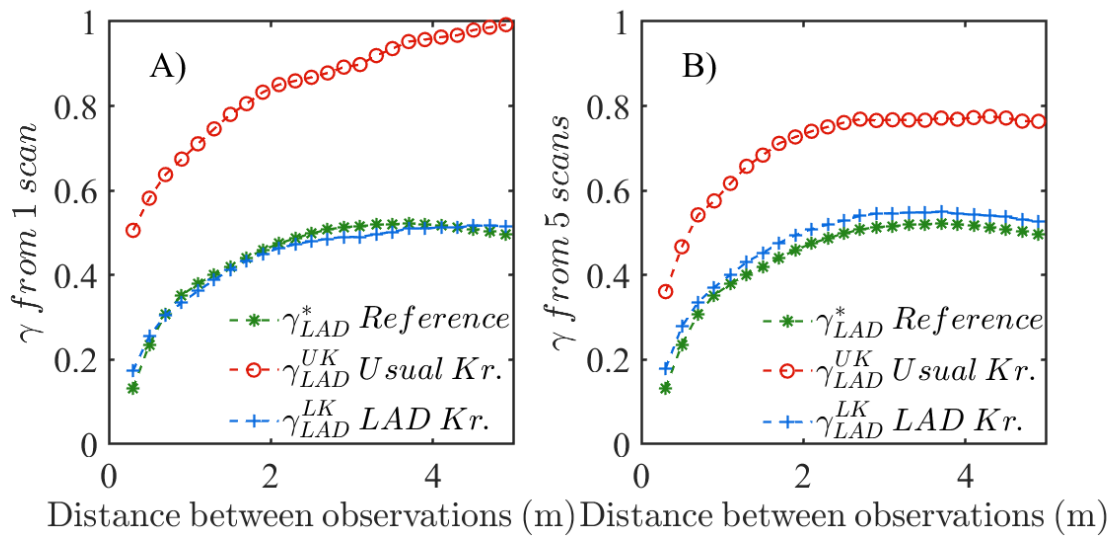
893 **Appendix A**



894

895 A1. Vertical profile of simulated reference Leaf Area Density 3D field.

896 **Appendix B**



897

898 B1. Variogram of the reference vegetation scene (green crosses). Estimated variograms from

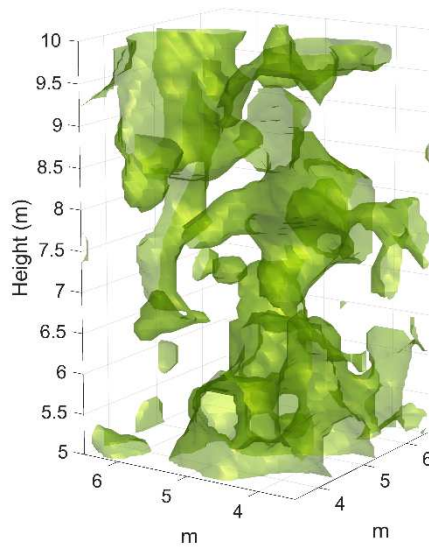
899 \widehat{LAD} computed for usual kriging method (red circle) and with the kriging method specific to LAD

900 (blue cross) for the scan design corresponding to A) Single viewpoint; B) Five viewpoints. For more

901 details, please refers to Soma et al. 2020.

902

903 **Appendix C**



904

905 C1. Isosurface of the simulated vegetation scene from RandomFields package. To ease visualization,
 906 only a subsample of the scene is displayed (core of the plot).

907

908 **Appendix D**

909 We assessed the significance of voxel size and number on scans effects on estimations though
 910 a statistical analysis. Voxels resulting from the fine and coarse discretization levels of the
 911 vegetation scene cannot be considered as independent observations when comparing voxel
 912 sizes, as voxels share the same vegetation in occupied volumes, with respect to their size (e.g.
 913 several 0.1 m voxels are contained in the 0.5 m voxel at the same location). To account for
 914 such specificity, all estimations were smoothed to the largest voxel size voxels, i.e. 1 m, and
 915 tests were performed at this scale. Smoothing was done by averaging 0.1 m – 0.2 m and 0.5 m
 916 voxels with respects to the 1 m voxel size grid. As voxels arise from the same vegetation
 917 scene, observations (voxels) cannot be considered independent between treatments (voxel size
 918 * number of scans), but should be considered as a repeated measure on the same individual.
 919 Hence, we performed a two-way ANOVA with repeated measures using R software (R Core
 920 Team (2016). R: A language and environment for statistical computing. R Foundation for
 921 Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>). The analysis was

922 carried out with estimation errors ($LAD_{REF,i} - \widehat{LAD}_i$) as predictor, for the 0.1 m, 0.2 m, 0.5 m
 923 and 1 m voxel sizes and for the two scans and five scans designs :

924
$$(LAD_{REF,i} - \widehat{LAD}_i) \sim IDvox / (Voxel\ size + Number\ of\ scans)$$

925 with $IDvox$ the voxel ID used for the within-subject model computation.

Effects	D.F.	F-value	P-Value
IDvox	999	-	-
IDvox : Voxel size	3	27.1	2e-16 ***
IDvox : Number of scans	1	34.31	6.4e-9 ***

926
 927 In accordance with graphical observations, both effects of voxel size and of number of scans
 928 are significant on the error in estimation of LAD value.