Estimators and confidence intervals for plant area density at voxel scale with T-LiDAR

François Pimont, Denis Allard, Maxime Soma, Jean-Luc Dupuy

1   Estimators and confidence intervals for plant area density at voxel scale with T-LiDAR

2

3   François Pimont[1,*], Denis Allard[2], Maxime Soma[1], Jean-Luc Dupuy[1]

4

5   [1] URFM, INRA, 84914, Avignon, France

6   [2] BioSP, INRA, 84914, Avignon, France

7   [*] Corresponding author. francois.pimont@inra.fr

8

9   **Abstract:**

10   Terrestrial LiDAR becomes more and more popular to estimate leaf and plant area density.

11   Voxel-based approaches account for this vegetation heterogeneity and significant work has

12   been done in this recent research field, but no general theoretical analysis is available.

13   Although estimators have been proposed and several causes of biases have been identified,

14   their consistency and efficiency have not been evaluated. Also, confidence intervals are

15   almost never provided.

16   In the present paper, we solve the transmittance equation and use the Maximum Likelihood

17   Estimation (MLE), to derive unbiased estimators and confidence intervals for the attenuation

18   coefficient, which is proportional to leaf area density. The new estimators and confidence

19   intervals are defined at voxel scale, and account for the number of beams crossing the voxel,

20   the inequality of path lengths in voxel, the size of vegetation elements, as well as for the

21   variability of element positions between vegetation samples. They are completed by

22   numerous numerical simulations for the evaluation of estimator consistency and efficiency, as

23   well as the assessment of the coverage probabilities of confidence intervals.

24   • Although commonly used when the beam number is low, the usual estimators are strongly

25   biased and the 95% confidence intervals can be $\approx \pm 100\%$ of the estimate.

26   • Our unbiased estimators are consistent in a wider range of validity than the usual ones,

27   especially for the unbiased MLE, which is consistent when the beam number is as low as 5.

28   The unbiased MLE is efficient, meaning it reaches the lowest residual errors that can be

29   expected (for an unbiased estimator). Also the unbiased MLE does not require any bias

30   correction when path lengths are unequal.

31   • When elements are small (or voxel is large), 10^3 beams entering the voxel leads to some

32   confidence intervals $\approx \pm 10\%$, but when elements are larger (or voxel smaller), it can remain

33   wider than $\pm 50\%$, even for a large beam number. This is explained by the variability of

34   element positions between vegetation samples. Such a result shows that a significant part of

35   residual error can be explained by random effects.

36   • Confidence intervals are much smaller (± 5 to 10%) when LAD estimates are averaged over

37   several small voxels, typically within a horizontal layer or in the crown of individual plants.

38   In this context, our unbiased estimators show a reduction of 50% of the radius of confidence

39   intervals, in comparison to usual estimators.

40   Our study provides some new ready-to-use estimators and confidence intervals for attenuation

41   coefficients, which are consistent and efficient within a fairly large range of parameter values.

42   The consistency is achieved for a low beam number, which is promising for application to

43   airborne LiDAR data. They entail to raise the level of understanding and confidence on LAD

44   estimation. Among other applications, their usage should help determine the most suitable

45   voxel size, for given vegetation types and scanning density, whereas existing guidelines are

46   highly variable among studies, probably because of differences in vegetation, scanning design

47   and estimators.

48

49

50   *Keyword*: terrestrial LiDAR; TLS; LAI; LAD; element size; bias; consistency; efficiency

51

52   *Highlights:*

53   • Voxel-based estimations of LAD/PAD may lack of consistency and efficiency

54   • We propose new estimators based on theoretical derivation and numerical simulations

55   • Estimators for confidence intervals are also provided

56   • New estimators should help determine the most appropriate voxel resolution

57

58   1. Introduction

59   The amount and spatial distribution of foliage in a tree canopy have a fundamental function in

60   ecosystems by affecting energy and mass fluxes through photosynthesis and transpiration.

61   Moreover, canopy structure may reveal plant adaptation strategies to their physical or biotic

62   environment (Norman and Campbell, 1989). Canopy foliage has other important ecological

63   functions since it constitutes the crown fuels involved in high intensity forest fires (Keane,

64   1995) and its spatial structure may determine the habitat quality for animal species (Vierling

65   et al., 2008). Terrestrial LiDAR (Light Detection And Ranging), referred to hereinafter as

66   TLS (Terrestrial LiDAR System) recently emerged as a promising tool to estimate leaf/plant

67   area density (LAD/PAD) distribution for individual plants and forest plots. Although similar

68   traversal algorithms have recently been used with high resolution airborne data, acquisitions

69   still suffer from substantial occlusion. This occlusion could be reduced with large flight strip

70   overlaps (Kükenbrik et al., 2017), which would lead to a promising application of methods

71   initially developed for TLS. Two classes of methods are commonly applied to derive LAD

72   distributions with TLS. First, the leaf area density profile can be measured through a gap

73   fraction approach (Jupp et al., 2009; Zhao et al., 2011). Rigorous statistical analysis using

74   maximum likelihood estimator (MLE) has been applied to the gap fraction equation inversion,

75   leading to robust estimates of LAD and leaf angulation profiles (Zhao et al., 2015).

76   Unfortunately, the gap fraction approach does not explicitly account for spatial correlation in

77   vegetation distribution (Zhao et al., 2015), whereas spatial correlation in heterogeneous media

78   are known to modify transmission laws and free path distribution (Davis and Marshak, 2004;

79   Pimont et al., 2009; Larsen and Clark, 2014). A clumping factor is thus required (Chen and

80   Cihlar, 1995; Zhao et al., 2011). Stochastic geometry entails to explicitly account for such

81   clumping, but only to determine the leaf area index, LAI (Allard et al., 2013), which is the

82   integral of the LAD over the vertical. The second class of methods is voxel-based and

83   explicitly account for clumping at scales larger than voxel size. They entail to assess not only

84   the vegetation vertical profile, but the full 3D distribution of area or mass density. Several

85   approaches have been developed: the voxel-based profiling (Hosoi and Omasa, 2006 & 2007;

86   Bailey and Mahafee, 2017a), the relative density index (Durrieu et al., 2008; Pimont et al.,

87   2015), the modified contact frequency (Béland et al., 2011) and the Beer-Lambert approach

88   (Béland et al., 2014b; Grau et al., 2017; Bailey and Mahafee, 2017a). These theoretical

89   indices can be readily applied or combined with field measurements through a calibration

90   phase (e.g., in Pimont et al., 2015).

91   The application of physical principles such as turbid media and contact frequency to

92   voxelized-TLS data raises several problems that folds in two categories. The first one deals

93   with departure from ideal measurements due to TLS "flaws". An idealized TLS would send

94   an infinite number of infinitely thin beams on any voxel. The actual diameter of the beam (on

95   the order of a few mm) is responsible for partial hits (Hebert and Koktov, 1992; Béland et al.,

96   2011; Grau et al., 2017). There is also uncertainty regarding beam intensity, due to the noise

97   instrument gain, that affects the detection (Grau et al., 2017). Béland et al., (2011) proposed

98   an approach that accounts for partial hits and intensity through a calibration of intensity and

99   view factors. Another aspect that has received little attention until now is the number of

100   beams entering the voxel (sampling size). The beam number depends on the distance to

101   scanner, the direction and scanner resolution, as well as the interaction with vegetation which

102   limits the number of beams reaching a given background voxel (occlusion). A basic rule of

103   the thumb is to only consider measurements with beam number larger than 10 (Béland et al.,

104   2014a) or to compute indices in large voxels, which leads to a fairly large beam number in

105   most cases (Bailey and Mahafee, 2017a).

106   The second cause of departure from ideal measurement is vegetation "flaws". An idealized

107   vegetation would be made of leaf only, assumed to be infinitely small elements with random

108   distribution of position and orientation. The actual orientation and size of elements alter

109   transmission laws (Larsen and Clark, 2014) and can be accounted for as in Béland et al.

110   (2011), where leaf orientation is separately measured and the interaction between a single leaf

111   and a beam is modelled. Element and branch orientations have been reported to be of

112   secondary importance in comparison to other sources of errors (Grau et al., 2017; Seielstadt et

113   al., 2011; Pimont et al., 2015). However, the assumption of spherical leaf inclination is not

114   valid in many cases (Pisek et al., 2013), which suggests that the assessment of the proper

115   angle distribution is likely to reduce errors. A recent method based on triangulation entails to

116   estimate the orientation factor with a TLS, provided that leaves are large enough to be

117   individually sampled by several beams (Bailey and Mahafee, 2017b). Regarding element size,

118   Grau et al. (2017) reports little effects when elements are much smaller than grid size. This

119   effect has been demonstrated to vary with voxel size (Béland et al., 2014b). Finally, several

120   methods based on return intensity have been proposed to separate leaf from wood returns and

121   to account for it (Béland et al., 2011, 2014a), even though such a method can not be applied to

122   all TLS (Pimont et al., 2015). Despite these known issues, a detailed analysis of the

123   consistency is still missing.

124   The determination of confidence intervals on LAD estimates has received little attention until

125   now. If such estimators are known in the context of gap fraction approaches (Zhao et al.,

126   2015), confidence interval for voxel-based approaches are seldom proposed (with the

127   exception of Pimont et al., 2015). Most error evaluations are based on simple comparison to

128   experimental data, in which various sources of bias and dispersion may interact. This might

129   explain why there is no consensus about the selection of voxel size among studies (e.g.,

130   Béland et al., 2011, 2014a; Bailey and Mahafee, 2017a; Grau et al., 2017). More generally, a

131   rigorous statistical analysis of estimators such as in Zhao et al. (2015) is still missing for

132   voxel-based approaches.

133   In the present study, we focus on some of the vegetation "flaws". We set our approach in the

134   framework of random set theory, stochastic geometry and stereology (Stoyan et al., 1987;

135   Schneider and Weil, 2008). In stochastic geometry, random distributions of geometrical

136   objects such as points, segments and disks are analyzed and analytical expressions are derived

137   for geometrical characteristics such as mean volume and area, specific surface etc. Stereology

138   is concerned with the estimation of those quantities with limited probing, in particular in

139   lower dimensions, such as beams probing a canopy voxel. We develop generalized estimators

140   towards two different approaches: i) the resolution of the transmittance equation (also called

141   Beer-Lambert law), ii) the maximum likelihood. Our developments are theoretical and

142   validated through numerical simulations. They include bias corrections for the beam number,

143   the element size, as well as for the variability of element positions between vegetation

144   samples. Estimators for variance are also provided and can be used to compute confidence

145   intervals. In order to facilitate the reading of the manuscript, most of the mathematical

146   development are detailed in supplementary materials for reference and only the main

147   equations are presented in the manuscript. Numerical simulations are used to compare the

148   new estimators to usual ones (Beer-Lambert, Modified Contact Frequency), through the

149   analysis of their consistency (i.e. bias size) and of their efficiency at the scale of a single voxel

150   or a group of voxels (i.e. 95% error). The application of the new estimators and their

151   confidence interval are then discussed, especially in the context of the determination of the

152   most appropriate voxel size.

153

154   **2. Background regarding the estimation of PAD/LAD through the attenuation**

155   **coefficient**

156   This section summarizes the existing knowledge regarding the estimation of attenuation

157   coefficient in voxel-based approaches and defines a few notations.

158

159   *2.1. Beer-Lambert law formulation for TLS*

160   The transmittance T in small and randomly distributed vegetation elements with no scattering

161   follows an exponential attenuation along a path of length δ, known as the Beer-Lambert law

162   (Nilson, 1971; Ross, 1981):

$$T = e^{-\int_0^\delta \lambda(z)dz} \tag{1}$$

163   with λ the attenuation coefficient of the medium (m^{-1})

164   The plant area density (PAD, m^{-1}) is related to the attenuation coefficient (λ, m^{-1}):

$$PAD = \lambda/G \tag{2}$$

165   where G is the plant projection function, which is frequently assumed to be equal to 0.5.

166

167   The complementary to one of the transmittance is the absorbance, A. TLS can be used to

168   estimate the absorbance of a vegetation sample with the relative density index I (also denoted

169   RDI in text), defined for a volume of vegetation, further referred to as the voxel *V* with

170   volume denoted $|V|$. The RDI is the ratio between the number of hits within the voxel (N_i), to

171   the number of beams that reaches the voxel (N):

$$1 - T(\lambda) = A(\lambda) \approx I = \frac{N_i}{N} \tag{3}$$

172   When beams are aligned with one cell face for cubic voxels or when the geometry of voxels is

173   spherical (Durrieu et al., 2008), the lengths of the different paths are equal. This is generally

174   not true and a first order approximation of Eq. 1 can be obtained using (Béland et al., 2014b;

175   Grau et al., 2017):

$$T \approx e^{-\lambda \bar{\delta}} \qquad (4)$$

176   where $\bar{\delta}$ is the mean path length within the voxel.

177   Taking the logarithm of the transmittance equation (Eq. 4) and combining with Eq. 3 leads to

178   the usual estimator of the attenuation coefficient:

$$\hat{\lambda} = -\frac{\log{(1 - I)}}{\bar{\delta}} \qquad (5)$$

179   This estimator (later referred to as the usual Beer-Lambert estimator) assumes that the

180   attenuation coefficient is constant in the voxel, that the vegetation elements are infinitely

181   small and that path length variations are negligible. Unequal path lengths involve the variance

182   of path length within voxels (Grau et al., 2017). An empirical correction that depends on

183   voxel orientation is described in Béland et al. (2014b) for cubic voxels. Another approach is

184   to use the secant method to solve the exponentially-weighted transmittance equation, since

185   such an equation does not have an explicit solution (Bailey and Mahafee, 2017a).

186    $\hat{\lambda}$ is not defined when I = 1, i.e. when no beam travels beyond the voxel. This occurs with

187   probability $\left(1 - e^{-\lambda \delta}\right)^{N}$ which is, for example, equal to 0.01 when $\lambda \delta = 1$ and N = 10.

188   Although this probability is very low when N is high, such an event may happen quite often in

189   any large voxelized scene, especially when the vegetation is dense. These cases can simply be

190   ignored, as proposed in Béland et al. (2014a), considering these cases as "occluded", but it

191   will be shown later that it leads to biases and loss of efficiency.

192

193   *2.2. Modified Contact frequency formulation for TLS*

194   The contact frequency of vegetation elements CF is the number of contacts per unit length of

195   point quadrat (probe) (Warren Wilson, 1960):

$$CF = \frac{\sum_{j=1}^{N} C_j^l}{N\delta} \tag{6}$$

196   where C_j^l is the number of leaf contacts for the jth probing and δ the probe length, and

197   assuming that the probing number N is large. It is related to plant area density in a similar

198   manner as the attenuation coefficient:

$$PAD \approx CF/G \tag{7}$$

199   With a TLS, the laser represents a virtual probe that is intercepted by vegetation. However,

200   the contact number cannot exceed one and only a fraction of the volume is explored by the

201   beam. A direct application of this method with TLS thus leads to an underestimation of the

202   attenuation coefficient (Bailey and Mahafee, 2017a). This method is adapted in Béland et al.,

203   (2011) for TLS data, thank to the volume fraction concept and leads to:

$$\tilde{\lambda} = \frac{\sum_{j=1}^{N} \mathbf{1}_{z_j < \delta_j}}{\sum_{j=1}^{N} z_j} = \frac{N_i}{N\bar{z}} = \frac{I}{\bar{z}} \tag{8}$$

204   where z_j is the length of the path actually explored (free path) by the jth beam and $\mathbf{1}_B$ is the

205   indicator function of event B ($\mathbf{1}_{z_j < \delta_j} = 1$ if the jth beam hits vegetation inside the voxel and 0

206   otherwise).

207   This formulation assumes that the explored volume is statistically representative of the

208   unexplored volumes. However, it does not assume the equality of path lengths (contrary to

209   Beer-Lambert estimator).

210

211   *2.3. Accounting for the size of vegetation elements*

212   The finite size of elements (ie. the size of elements is larger than zero in the real world)

213   induces a bias in the above estimators when the element size is not negligible when compared

214   to voxel size (Béland et al., 2014a). Assuming that the beams are parallel, let S_1 and S be

215   respectively the cross sections of the element size and voxel volume, which will be assumed

216   constant for simplicity. The probability that a given beam crosses the voxel containing p

217   elements is $\left(1 - \frac{S_1}{S}\right)^p$, as shown in Campbell and Norman (1998, chapter 15).

218   The volume cross section is given by $S = \frac{|V|}{\delta}$, where δ is the path length. The contribution of

219   a single leaf to the attenuation coefficient of the voxel V is:

$$\lambda_1 = \frac{S_1}{|V|} = \frac{S_1}{S}\frac{1}{\delta} \tag{9}$$

220   Since $\lambda = p\lambda_1$, the transmittance of the voxel is:

$$T \approx \left(1 - \frac{S_1}{S}\right)^p = (1 - \lambda_1\delta)^{\frac{\lambda}{\lambda_1}} \tag{10}$$

221   When the element size is not neglected is, the Beer-Lambert estimator is modified as follow:

$$\widehat{\lambda_p} = \frac{\lambda_1 \log(1 - I)}{\log(1 - \lambda_1\delta)} \tag{11}$$

222   As show with slightly different notations in Béland et al. (2014a), $\widehat{\lambda_p} = \frac{\hat{\lambda}}{-\frac{1}{\lambda_1\delta}\log(1-\lambda_1\delta)}$, which

223   converges to $\hat{\lambda}$ when $R = \frac{S}{S_1} = \frac{1}{\lambda_1\delta}$ is large, or equivalently, when $\lambda_1\delta$ tends to 0.

224

225   In the discussion section of Béland et al. (2014a), the same correction factor is suggested to

226   apply to the modified contact frequency, so that the modified contact frequency for finite size

227   element would be:

$$\widetilde{\lambda_p} = -\frac{\lambda_1\bar{\delta}}{\log\left(1 - \lambda_1\bar{\delta}\right)}\frac{I}{\bar{z}} \tag{12}$$

228

229   Depending whether elements are small needles or broad leaves, $\lambda_1\delta$ typically range between

230   $2\,10^{-5}\delta^{-2}$ and $5\,10^{-3}\delta^{-2}$ (See supplementary S1). This means that $\lambda_1\delta$ ranges between

231   0.002 and 0.5 when the voxel size is about 10 cm, and between 0.0002 and 0.05 when the

232   voxel size is 30 cm. $\lambda_1\delta$ is smaller than 0.005 when the voxel size is on the order of 1 m.

233

234 **3. Mathematical formulation**

235

236 In this section, we develop the mathematical framework leading to unbiased estimators for

237 point, variance and confidence intervals of the attenuation coefficient. The proofs are given in

238 Supplementaries S2 and S3.

239

240 *3.1. Set-up and notations*

241 We assume a finite number of elements and we rely on the notations defined in 2.3. For a

242 given quantity A, such as an estimator, a variance or a confidence interval radius, $\hat{A}$ denotes

243 the quantity as derived from the Beer-Lambert law, whereas $\tilde{A}$ is derived from the Maximum

244 Likelihood Estimator approach. As shown below, the MLE generalizes the Modified Contact

245 Frequency introduced in Béland et al. (2011), so that these symbols are consistent with

246 section 2 (Eq. 8). Furthermore, the use of upper case letters, such as Λ, refers to our new

247 estimators presented below, whereas lower case letters, such as λ, refers to the usual ones

248 presented in the background section.

249

250 We briefly present some stochastic geometry material and refer to the literature for a more

251 detailed exposition (Stoyan et al., 1987; Schneider and Weil, 2008). p elementary objects

252 identical in shape and size are located at random within the voxel *V*. If one representative of

253 these objects is denoted *B*, the vegetation elements, denoted *X*, corresponds to the union of all

254 objects:

$$X = \bigcup_{k=1}^{p} B(x_k) \tag{13}$$

255 where $B(x_k)$ denotes the element *B* shifted to the random location $x_k \in V$.

256

257   One special case of interest is the Boolean model, for which the number of objects *p* is

258   distributed according to a Poisson distribution (Stoyan et al., 1987). A remarkable feature of

259   Boolean models is that the intersection of a Boolean model by a random line is also a Boolean

260   model with intensity $\lambda = S_1 \frac{p}{|V|}$, where S_1 is the cross-section of *B* perpendicular to the line. As

261   a consequence, the lengths *Y* of segments with no intersection with *X*, called free paths in the

262   present context, are distributed according to an exponential random variable with parameter λ:

$$P(Y > y) = e^{-\lambda y}, y > 0 \qquad (14)$$

263   When the object number is not assumed to be large, the distribution becomes:

$$P(Y > y) = (1 - \lambda_1 y)^{\lambda/\lambda_1} \qquad (15)$$

264   with $\lambda_1 = \frac{S_1}{|V|}$.

265   Let *M* be the segment corresponding to the intersection between a beam and the voxel *V*.

266   Depending on voxel shape and size, its length δ follows a distribution Δ. The distance actually

267   traveled in voxel Z by a beam corresponds to:

$$Z = \min(Y, \Delta) \qquad (16)$$

268   The probability distribution of Z is derived from Eq. 15:

$$f_Z(z; \delta) = \begin{cases} \lambda(1 - \lambda_1 z)^{\lambda/\lambda_1 - 1} & \text{when } z < \delta \\ (1 - \lambda_1 \delta)^{\lambda/\lambda_1} & \text{when } z = \delta \end{cases} \qquad (17)$$

269   In the Z distribution, the density of Y for Y> δ is cumulated at z = δ due to the "min" operator

270   in Eq. 16. Note that there are <u>two components of randomness</u> in this set-up: a random

271   realization of a Boolean model *X,* on the one hand (i.e. element positions for a vegetation

272   sample) and a random beam *M* over the cross-section of the voxel *S*, on the other hand (i.e.

273   instrument sampling).

274

275   Let us consider a given realization of the Boolean model X and N beams $\{m_j\}_{j \leq N}$ distributed

276   over the voxel. We can define the RDI (Eq. 3) as the fraction of beams hitting the canopy

277   elements as:

$$I(X, \{m_j\}_{j \leq N}) = \frac{N_i(X)}{N} = \frac{\sum_{j=1}^{N} \mathbf{1}_{z_j(X, m_j) < \delta_j(m_j)}}{N} \tag{18}$$

278   Let us now denote $I_\infty(X)$ the asymptotic RDI, which is the expectation of $I(X)$ with respect to

279   the instrument sampling ($i.e.\ N \to \infty$):

$$I_\infty(X) = E_M[\mathbf{1}_{Y(X) < \Delta}] \tag{19}$$

280

281   *3.2. Point, Variance and Confidence Interval of the relative density index (RDI)*

282   The RDI defined in Eq. (18) for N beams has the same expectation as the asymptotic RDI

283   over all configurations X, since beams are drawn randomly across *S*.

284   This expectation is, according to Eq. (15):

$$E_X[I_\infty(X)] = E_{M,X}[\mathbf{1}_{Y(X) < \Delta}] = \frac{1}{S} \int_S P(Y < \delta(s)) ds = 1 - \frac{1}{S} \int_S (1 - \lambda_1 \delta(s))^{\lambda/\lambda_1} ds \tag{20}$$

285   As pointed out above, the variance of I has two components. The conditional variance

286   formula provides:

$$Var(I(X)) = E_X[Var_M(I(X))] + Var_X(E_M[I(X)])$$
$$= E_X[Var_M(I(X)|I_\infty(X))] + Var_X(I_\infty(X)) \tag{21}$$

287   Now, assuming that the N beams are independent and identically distributed, one gets

288   $Var_M(I)(X) = \frac{I_\infty(X)(1 - I_\infty(X))}{N}$, since I is simply a proportion estimated on a sample of size N.

289   The variance $Var_X(I_\infty(X))$ is due to the variability of element positions in a vegetation sample.

290   This variance becomes negligible when the vegetation sample is made of a large number of

291   small area elements, but it cannot be evaluated in closed form for actual configurations of *X*.

292   It is instead approximated by numerical simulations, as described in section 4 (for $\lambda_1 \bar{\delta} < 0.3$):

$$\mathrm{Var}_X(\mathrm{I}_\infty(X)) \approx \sigma^2_{\mathrm{I}_\infty}\big(\mathrm{I}_\infty(X)\big) = 0.230\lambda_1\overline{\delta}\,\mathrm{I}_\infty(X)^{1.903-2.30\lambda_1\overline{\delta}}\big(1 - \mathrm{I}_\infty(X)\big) \tag{22}$$

293   From now on, we drop the dependence to *X* and *M* for the ease of notations. Putting these

294   results together leads to the following estimator for the variance of I:

$$\sigma_I^2 = \frac{\mathrm{I}(1-\mathrm{I})}{\mathrm{N}} + \sigma^2_{\mathrm{I}_\infty}\big(\mathrm{I}, \lambda_1\overline{\delta}\big) \tag{23}$$

295   Hence to the following Wald confidence interval for the expectation of the asymptotic RDI,

296   I_∞:

$$\mathrm{I} \pm z_{\alpha/2}\sqrt{\frac{\mathrm{I}(1-\mathrm{I})}{\mathrm{N}} + \sigma^2_{\mathrm{I}_\infty}(\mathrm{I}, \lambda_1\overline{\delta})} \tag{24}$$

297   where $z_{\alpha/2}$ is the usual $1 - \frac{\alpha}{2}$ quantile of the standard Gaussian distribution.

298

299   The Wald interval is known to have a lower-than-expected coverage probability when the

300   empirical proportion (here, the RDI) is close to 0 or 1. It means that the actual value of λ is

301   less frequently within the estimated interval than expected (Brown et al., 2001). As an

302   example, it is obviously the case when I=0 (or 1), since the true value is supposed to be 0 (or

303   1) at 100 %. This is problematic since both cases are quite frequent in TLS data. Among

304   others, the Agresti-Coull interval is a simple alternative to the Wald interval recommended in

305   Brown et al. (2001). Its formulation is similar to that of the Wald interval:

$$\mathrm{I}_c \pm z_{\alpha/2}\sqrt{\frac{\mathrm{I}_c(1-\mathrm{I}_c)}{\mathrm{N}_c} + \sigma^2_{\mathrm{I}_\infty}(\mathrm{I}_c, \lambda_1\overline{\delta})} \tag{25}$$

306   with corrected values of "I" and "N" defined as follow:

$$\begin{cases} \mathrm{I}_c = \dfrac{\mathrm{I} + \frac{z_{\alpha/2}^2}{2\mathrm{N}}}{1 + \frac{z_{\alpha/2}^2}{\mathrm{N}}} \\ \mathrm{N}_c = \mathrm{N} + z_{\alpha/2}^2 \end{cases} \tag{26}$$

307   This correction leads to confidence intervals that are not centered on I and wider than the

308   Wald interval. Agresti and Coull interval is known to have a higher-than-expected coverage

309   probability when N is small, which is not fully satisfactory, but safer than the Wald interval

310   (Brown et al., 2001).

311

312   *3.3. Point and Variance Estimation of the unbiased Beer-Lambert estimator*

313   The Beer-Lambert estimator derives from solving the transmittance of the voxel medium

314   (section 2.1) and thus rely on the empirical relative density index. In supplementary S2, we

315   extend this approach to define the unbiased Beer-Lambert estimator $\hat{\Lambda}$, valid for close-to-

316   equal path lengths:

$$
\hat{\Lambda} = \begin{cases} -\dfrac{1}{\overline{\delta_e}}\left(log(1-I) + \dfrac{I}{2N(1-I)}\right) & \text{when } I < 1 \\[2ex] \dfrac{log(2N+2)}{\overline{\delta_e}} & \text{when } I = 1 \end{cases} \tag{27}
$$

317   with the mean "effective" path length:

$$
\overline{\delta_e} = mean\left(-\frac{log(1-\lambda_1\delta_j)}{\lambda_1}\right) \tag{28}
$$

318   When $\lambda_1 \ll \frac{1}{\overline{\delta}}$, it simplifies to $\overline{\delta_e} \approx \overline{\delta}$ (mean path length).

319

320   The first term in Eq. (27) when I<1 accounts for the size of elements (Eq. (11)). The second

321   term is a bias correction for the instrument sampling, that compensates a systematic bias

322   caused by the convexity of the log function (See Supplementary S2 for details). Such a bias

323   has never been reported before, to the best of our knowledge.

324   When I=1, the above formulation is derived from the application of the Beer-Lambert law to

325   the center of the Agresti-Coull interval, which is more robust than I.

326

327   For unequal path lengths, the transmittance equation can be approximated as a second order

328   polynomial in λ, which leads to the following unbiased estimator:

$$\hat{\Lambda}_2 = \frac{1}{a_e}\left(1 - \sqrt{1 - 2a_e\hat{\Lambda}}\right) \tag{29}$$

329   with $\hat{\Lambda}$ in Eq. (27) and a_e the ratio between empirical variance to mean of the effective path

330   length:

$$a_e = var\left(-\frac{\log(1 - \lambda_1\delta_j)}{\lambda_1}\right)\Big/ mean\left(-\frac{\log(1 - \lambda_1\delta_j)}{\lambda_1}\right) \tag{30}$$

331   Notice that when $\lambda_1 \ll \frac{1}{\delta}$, $a_e \approx \frac{\sigma_\delta^2}{\bar{\delta}}$, with σ_δ^2 the empirical variance of path lengths.

332

333   The variances of both unbiased estimators $\hat{\Lambda}$ and $\hat{\Lambda}_2$ can be derived from the variance of the

334   RDI (Eq. 23), as shown in Supplementary S2:

$$\sigma_{\hat{\Lambda}}^2 = \begin{cases} \frac{1}{\bar{\delta}_e^2}\left(\frac{I}{N(1-I)} + \frac{\sigma_{I_\infty}^2(I, \lambda_1\bar{\delta})}{(1-I)^2}\right)\left(1 - \frac{1}{2N(1-I)}\right)^2 & \text{when } I < 1 \\ \frac{1}{\bar{\delta}_e^2}\left(2 + \frac{1}{N} + (2N+2)^2\sigma_{I_\infty}^2\left(\frac{2N+1}{2N+2}, \lambda_1\bar{\delta}\right)\right) & \text{when } I = 1 \end{cases} \tag{31}$$

335   and

$$\sigma_{\hat{\Lambda}_2}^2 = \sigma_{\hat{\Lambda}}^2\left(1 + 2a_e\hat{\Lambda} + 4\left(a_e\hat{\Lambda}\right)^2\right) \tag{32}$$

336   These estimators account for the instrument sampling (with the 1/N term when I<1), the

337   asymptotic variability of element positions between vegetation samples (terms with $\sigma_{I_\infty}^2$) and

338   the convexity of the log function (third factor when I<1). As above, the case corresponding to

339   I=1 is based on the center of the Agresti-Coull interval.

340

341   *3.4. Point and Variance Estimation from Maximum Likelihood Estimation*

342   The following estimator is derived from Maximum Likelihood (Kay, 1993, Chapter 7), that

343   uses the full information provided by the TLS, namely the actual distribution of free paths in

344   the voxel $\{z_j\}_{j=1,N}$ the N free paths. In supplementary S3, we compute the analytical MLE

345   from the expected free path distribution (Eq. (17)). These derivations show that the modified

346   contact frequency proposed in Béland et al. (2011) is indeed the MLE, which demonstrates its

347   asymptotic consistency when N is large. This formulation is extended to the case of elements

348   of finite size, thanks to the "effective" mean free path:

$$\overline{z_e} = \text{mean}\left(-\frac{\log(1 - \lambda_1 z_j)}{\lambda_1}\right) \tag{33}$$

349   The MLE is asymptotically-normal, meaning that its residuals become normal when N is

350   large (Kai, 1993). However, the MLE is biased when the number of beams is finite. In

351   Supplementary S3, we account for this bias, which leads to the following unbiased MLE:

$$\widetilde{\Lambda} = \frac{I}{\overline{z_e}} - \frac{\overline{\mathbf{1}_{z<\delta}z_e}}{N\overline{z_e}^2} \tag{34}$$

352   with

$$\overline{\mathbf{1}_{z<\delta}z_e} = \text{mean}\left(-\frac{\mathbf{1}_{z_j<\delta_j}\log(1 - \lambda_1 z_j)}{\lambda_1}\right) \tag{35}$$

353   Compared to the Beer-Lambert estimator, this approach does not require any correction for

354   unequal path lengths.

355

356   In supplementary S3, we rigorously compute the variance of $\widetilde{\Lambda}$ with bias correction for

357   instrument sampling and the variability of element positions between vegetation samples:

$$\sigma_{\widetilde{\Lambda}}^2 = \frac{I}{N\overline{z_e}^2}\left(1 - \frac{\overline{\mathbf{1}_{z<\delta}z_e}}{NI\overline{z_e}}\right)^2 + \frac{\sigma_{I_\infty}^2(I_b, \lambda_1\overline{\delta})}{\overline{\delta}_e^2(1 - I_b)^2} \tag{36}$$

358   With

$$I_b = \min\left(I, 1 - \frac{1}{2N + 2}\right) \tag{37}$$

359   The factor involving $\overline{\mathbf{1}_{z<\delta}z_e}$ expresses the bias correction for the instrument sampling and the

360   term with $\sigma_{I_\infty}^2$ derives from the variability of vegetation samples, as for the variance of the

361   unbiased Beer-Lambert estimator.

362

363　　*3.5. Cramer-Rao bound for variance*

364　　The Cramer-Rao bound is the theoretical lower bound of the variance of unbiased estimators

365　　(Kay, 1993, Chapter 3), meaning that an unbiased estimator with variance as small as the

366　　Cramer-Rao bound is optimal. In Supplementary S3, we show that when the vegetation

367　　samples are not fixed (i.e. when TLS beams are shot on variable vegetation samples), the

368　　Cramer-Rao bound is:

$$\text{CRB}_\lambda = \frac{\lambda^2}{\text{NI}_\infty(\lambda)} \tag{38}$$

369　　This analytical formulation is helpful, since the value of the Cramer-Rao bound can be

370　　analytically computed (integrating Eq. (20) to compute $I_\infty(\lambda)$), when both attenuation

371　　coefficient and voxel geometry are known. This is the case in the numerical simulations

372　　developed in sections 4 and 5, which thus provides a way to evaluate the efficiency of the

373　　unbiased estimators, that are expected to have empirical variances as close as possible to

374　　CRB_λ. A perfect match corresponds to the most efficient estimator.

375　　It is important to notice however, that this theoretical bound can never be achieved when the

376　　variability of vegetation samples has a significant contribution to the variance of the RDI,

377　　since this variability is not accounted for in this theoretical bound (see Supplementary S3 for

378　　more details).

379

380　　*3.6. Estimating confidence intervals for a voxel or a group of voxel*

381　　From unbiased estimators $\widetilde{\Lambda}$ and $\sigma_{\widetilde{\Lambda}}^2$, the confidence interval at a risk level α can naturally be

382　　estimated as:

$$\widetilde{\Lambda} \pm z_{\alpha/2}\sqrt{\sigma_{\widetilde{\Lambda}}^2} \tag{39}$$

383　　However, such a formulation is expected to have the same limitations as the Wald interval for

384　　the RDI, when the probabilities of interception (RDI) are low or high. This interval is thus

385   expected to lead to lower-than-expected coverage probabilities in voxels with low and high

386   density. As for the Agresti-Coull interval, an alternative is to replace I by I_c and N by N_c in

387   estimation of $\widetilde{\Lambda}$ and $\sigma_{\widetilde{\Lambda}}^2$, leading to:

$$\widetilde{\Lambda}_c \pm z_{\alpha/2}\sqrt{\sigma_{\widetilde{\Lambda}_c}^2} \tag{40}$$

388   If estimations at voxel scale is a key outcome of TLS, the scale of interest is often larger,

389   typically the individual plants, the horizontal slice of vegetation, or the forestry plot. In this

390   case, the variable of interest is not the single voxel estimation, but the average attenuation

391   coefficient in a group of voxels. For a group of n_v voxels and assuming independence between

392   voxels, the confidence interval on the mean of attenuation coefficient estimators is:

$$\frac{1}{n_v}\sum_{n_v}\widetilde{\Lambda} \pm \frac{z_{\alpha/2}}{n_v}\sqrt{\sum_{n_v}\sigma_{\widehat{\Lambda}}^2} \tag{41}$$

393   The 95% errors, defined as the radius of the confidence interval at 95%, for a single voxel or a

394   group of voxel are thus:

$$E95_{\widetilde{\Lambda}} = 1.96\sigma_{\widetilde{\Lambda}} \tag{42}$$

395   And

$$E95_{\widetilde{\Lambda}}^{n_v} = \frac{1.96}{n_v}\sqrt{\sum_{n_v}\sigma_{\widetilde{\Lambda}}^2} \tag{43}$$

396   Similar quantities can be defined for the unbiased Beer-Lambert estimators $\widehat{\Lambda}$ and $\widehat{\Lambda}_2$, as well

397   as the bound of the 95% error, based on the Cramer-Rao bound for variance, which is the

398   lower bound of 95% error for an unbiased estimator (i.e. no unbiased estimator can lead to

399   smaller errors).

400

401 **4. Design of the numerical experiments**

402

403 *4.1. Overview*

404 The aim of our numerical simulations is to compare the estimates of attenuation coefficients

405 to their true values. Simple configurations are generated to simulate replicates of virtual TLS

406 point clouds in voxels filled with idealized vegetation of known properties. The numerous

407 replicates enable to compute various statistics through a MonteCarlo approach, to evaluate the

408 consistency and efficiency of each estimator, as well as the consistency of variance and

409 confidence interval estimators. For each estimator, we compute its expectation, variance and

410 95 % errors. The 95 % errors are estimated as the 95th percentile of the absolute residuals and

411 are evaluated for a single voxel and a group of voxels (here $n_v = 100$). We also compute the

412 expectation of variance estimators, as well as the coverage probability of the estimated

413 confident intervals, which is the empirical frequency at the true value of the attenuation

414 coefficient belongs to the estimated confidence interval.

415 These statistics are computed for various values of attenuation coefficients, element sizes,

416 voxel sizes, and beam numbers. Simulations are run for two different configurations described

417 below. The first configuration, described in details in subsection 4.2, assumes finite element

418 size (meaning that $\lambda_1 > 0$) and equal path lengths. Equal path lengths imply that the second

419 order correction for the Beer-Lambert estimators (i.e. $\widehat{\Lambda}^2$) is not required. The computation of

420 RDI and distance travelled is done over the actual distributions of vegetation elements

421 (generated with random positions), each corresponding to a fixed vegetation sample *X*.

422 Simulations are run for a large number of vegetation samples, which entails to compute the

423 asymptotic variance of the RDI, $\sigma^2_{I_\infty}$, (when the beam number is infinite) for each value of

424 attenuation coefficient and element size. The second set of simulations (subsection 4.3) is

425 specifically designed to evaluate the correction for unequal path lengths involved in unbiased

426   Beer-Lambert estimator $\widehat{\Lambda}_2$. For simplicity, we assume that vegetation elements are infinitely

427   small ($\lambda_1 = 0$) and that the voxel is a sphere, so that the free-path distribution and the

428   Cramer-Rao bound can be analytically solved (See Supplementary S4 for details). We

429   separate both sets of simulations to facilitate the presentation of results.

430   In order to simplify the numerical experiment design and the presentation of the results, we

431   build dimensionless quantities, namely the beam depth ($L_j = \lambda \delta_j$) and the voxel depth ($L =$

432    $\lambda \bar{\delta}$). Likewise, $y_j = \lambda z_j$ is the free depth for beam j. We can notice that I can be computed

433   from the distribution of depth $\{y_j\}$ since $\mathbf{1}(z_j < \delta_j) = \mathbf{1}(y_j < L_j)$. The element depth is $L_1 =$

434    $\lambda_1 \bar{\delta}$. More generally, dimensionless quantities can be derived for all quantities of interest

435   developed above, as shown in Supplementary S5. The practical interest of this substitution is

436   that computations can be done for a series of voxel depth values, and easily extrapolated to λ

437   by simply dividing results by $\bar{\delta}$, instead of running simulations for series of $(\lambda, \bar{\delta})$ values.

438

439   *4.2. Numerical simulations for finite-size elements and equal path lengths*

440   Simulated vegetation samples correspond to flat square elements that are randomly distributed

441   in a voxel, parallel to one face of the voxel. The virtual beams are sent perpendicular to

442   elements, so that the path lengths are equal to cube size (Figure 1). The voxel and element

443   depths are $L = \lambda \delta$ and $L_1 = \lambda_1 \delta$, respectively.

444
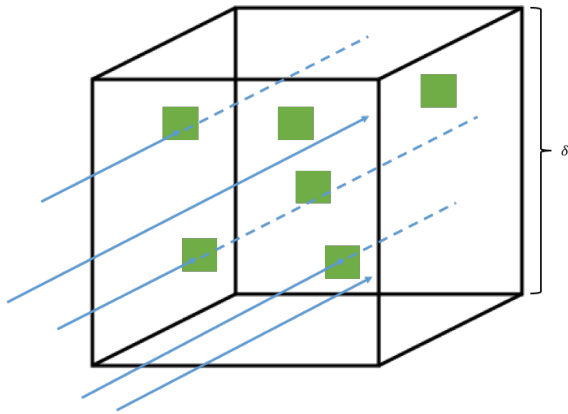
445

446   Fig. 1. Illustration of a numerical simulation of TLS beams over finite-size elements and equal path

447   length δ. Each square element has a size equal to $\sqrt{S_1} = \sqrt{\lambda_1 \delta^3} = \sqrt{L_1}\delta$.

448

449   For each value of L and L_1 tested (Table 1), we simulate 10000 vegetation samples on which

450   we shoot M batches of N virtual beams, with N between 5 and 10000. Batches serve as

451   replicates of TLS shooting, to compute the different statistics (estimator expectation, variance,

452   95% error, variance and confidence interval estimator, confidence interval coverage

453   probability). The batch numbers are selected so that the total number of beams MN is

454   constant, equal to $= 10^8$. This number is large enough for the convergence of the different

455   statistics, despite the replicate number M decreases with N, since the variance of the estimates

456   sharply decays with N.

457

458   **Table 1.** Parameter values in numerical simulations

Parameter	Values
Voxel depth (L)	0.05, 0.1, 0.5, 1,1.5, 2, 2.5, 3
Beam number (N)	3,5,7,10,15,20,30,40,50,75,100,150,200,300,400, 500,750 1000,5000,10000
Element depth (L_1)	$\begin{cases} Subsection\ 4.2: 0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.5 \\ Subsection\ 4.3: 0 \end{cases}$

459   Nb: in subsection 4.3, elements are assumed infinitely small so that L_1=0.

460

461   We use simulations with the largest beam number ($N = 10000 \approx \infty$) to estimate the

462   asymptotic variance $\sigma^2_{I_\infty}$, for various values of L_1 and L. We remind that this variance is due

463   to the variability of element position between vegetation samples *X*. When $L_1 < 0.3$, which is

464   the case for most vegetation when voxels are greater than 10 cm (See Supplementary S1),

465    $\sigma^2_{I_\infty}$ can be estimated from a simple empirical function of L_1 and RDI (Fig. 2 and Eq. (22)).
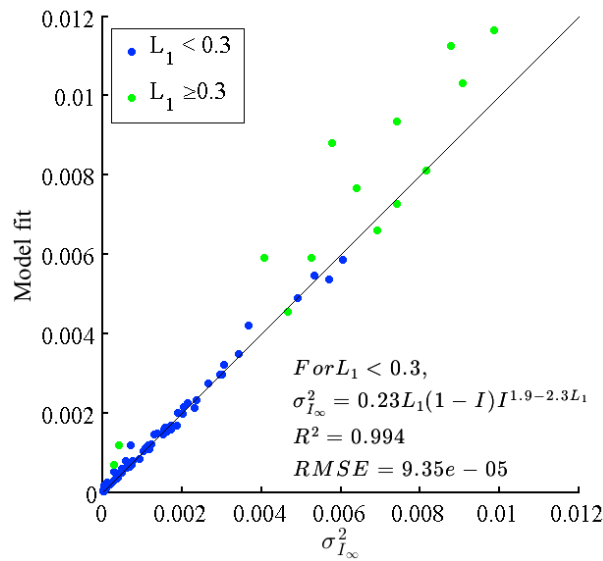
466

467   **Fig. 2.** Empirical model for the variance of the asymptotic relative density index I_∞. $\sigma^2_{I_\infty}$ is caused by

468   the variability of element positions between vegetation samples.

469

470   *4.3. Numerical simulations for unequal path lengths*

471   In a sphere with radius R, the voxel depth is:

$$L = \lambda\overline{\delta} = \lambda\frac{V}{S} = \lambda\frac{\frac{4}{3}\pi R^3}{\pi R^2} = \frac{4}{3}\lambda R \qquad (44)$$

472   The distribution of dimensionless optical depth L_j is, with u between 0 and 1 (Supplementary

473   S4):

$$\text{PDF}\left(L_j = \frac{3}{2}L\sqrt{1 - u^2}\right) = 2u \qquad (45)$$

474   For each value of L (Table 1), we simulate a total of 10^8 virtual beams with lengths $\{L_j\}$ and

475   the corresponding free path lengths $\{y_j\}$, that respectively follow Eq. (45) and the exponential

476   law. As in subsection 4.2, virtual beams are grouped in batches of N beams to simulate M

477   replicates of TLS shooting and to compute the same statistics as above.

478

479   **5. Numerical simulation results**

480   In this section, we show the statistics described in section 4. Subsection 5.1. corresponds to

481   finite size element simulations (described in subsection 4.2), whereas subsection 5.2.

482   corresponds to unequal path simulation (described in subsection 4.3). Expectation, variance

483   and 95% error enable to compare the consistency and efficiency of the usual estimators ($\hat{\lambda}$ and

484     $\tilde{\lambda}$) and the new ones ($\widehat{\Lambda}$, $\widehat{\Lambda}_2$ and $\widetilde{\Lambda}$). We also show the expectation of the variance estimators

485   ($\sigma^2_{\widehat{\Lambda}}$, $\sigma^2_{\widehat{\Lambda}_2}$ and $\sigma^2_{\widetilde{\Lambda}}$), the confidence interval radiuses ($E95_{\widehat{\Lambda}}$, $E95_{\widehat{\Lambda}_2}$ and $E95_{\widetilde{\Lambda}}$), and the coverage

486   probabilities of estimated confidence intervals.

487

488   *5.1. Estimator performance for finite size elements*

489

490   *5.1.1 Estimator consistency*

491   Figure 3 shows the expectation of estimator derived from the MLE, as a function of the beam

492   number. Blue dots corresponds to the modified contact frequency estimator $\tilde{\lambda}$ (Eq. 8), which is

493   the biased MLE for infinitely small elements. Green dots corresponds to the unbiased MLE $\widetilde{\Lambda}$

494   (Eq. 34), that accounts for element size and beam number. Since expectations are normalized

495   by the true value of λ, estimators are consistent when expectations equal one, and deviations

496   from 1 quantifies the bias. Subplots A, B and C correspond to small elements (L_1=0.001) for

497   three voxel depths L, whereas subplots D, E and F correspond to "larger" elements compared

498   to voxel size (L_1=0.1) for the same values of voxel depths.

499　　Even when elements are small, the modified contact frequency is positively biased when N is

500　　small (subplots A, B, C). Vertical blue lines show the thresholds of N for which the bias of $\tilde{\lambda}$

501　　is larger than 1%. These thresholds range between N=30 and N=75, depending on L. When

502　　elements are larger (subplots D, E, F), the positive biases remain for large values of N, so that

503　　the 1% threshold is not reached. On the contrary, the unbiased MLE $\tilde{\Lambda}$ (green dots) shows

504　　biases always lower than 1% (green lines) when elements are small, even when N is as small

505　　as 3 for L_1=0.001 (subplots A, B, C) and 5 when L_1=0.1 (subplots D, E, F).
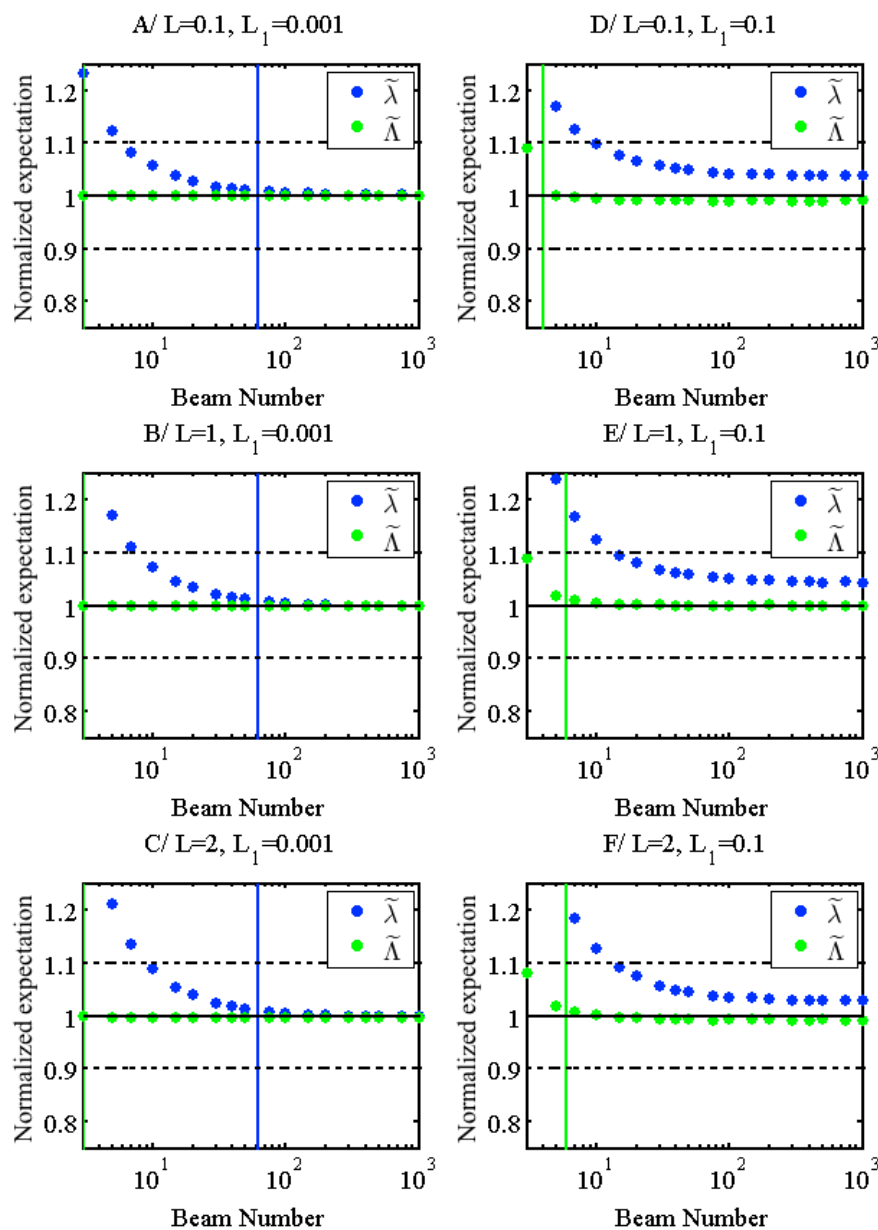

506

507   Fig. 3. Expectations of the attenuation coefficient estimators derived from the MLE, as a function of

508   the beam number. Blue dots corresponds to the modified contact frequency estimator $\tilde{\lambda}$ (Eq. 8), which

509   is the biased MLE for infinitely small elements. Green dots corresponds to the unbiased MLE $\widetilde{\Lambda}$ (Eq.

510   34), that accounts for element size and beam number. Estimators are normalized by their true value λ,

511   so that they are consistent when the expectation equals to one. The vertical lines correspond to the

512   lowest values of N leading to a bias smaller than 1% in blue and green for respectively the biased and

513   unbiased estimators.

514

515   Figure 4 is similar to figure 3, but for the usual ($\hat{\lambda}$) and unbiased ($\widehat{\Lambda}$) Beer-Lambert estimators.

516   Trends are similar with two main differences. First, the biases are slightly larger with Beer-

517   Lambert estimators than with the MLE when the vegetation density is low to moderate (L<=1,

518   subplot A, B, D and E), whether bias are corrected or not. Second, the bias of the usual Beer-

519   Lambert estimator decreases until becoming negative for small values of N when the

520   vegetation density is high (subplots C and F). Such a decay is attributed to the occurrence of

521   cases in which all beams are intercepted inside the voxel (I=1), referred to as "occluded" in

522   Béland et al. (2014a). In this cases, the usual Beer-Lambert estimator is $+\infty$, but is ignored in

523   expectation computation to avoid divergence. Attenuation coefficient estimates are thus

524   bounded by $\frac{\log(N)}{\delta}$, leading to this negative bias of increasing magnitude when N is small.

525   This trend is also visible for $\widehat{\Lambda}$, but it is far less pronounced and it occurs for lower values of

526   N. This demonstrates the benefit of the definition of $\widehat{\Lambda}$, which is extended when I=1 with

527   Agresti-Coull interval centers. The range of consistency of $\widehat{\Lambda}$, however, is clearly narrower

528   than the one of the unbiased MLE $\widetilde{\Lambda}$.

529

530

531
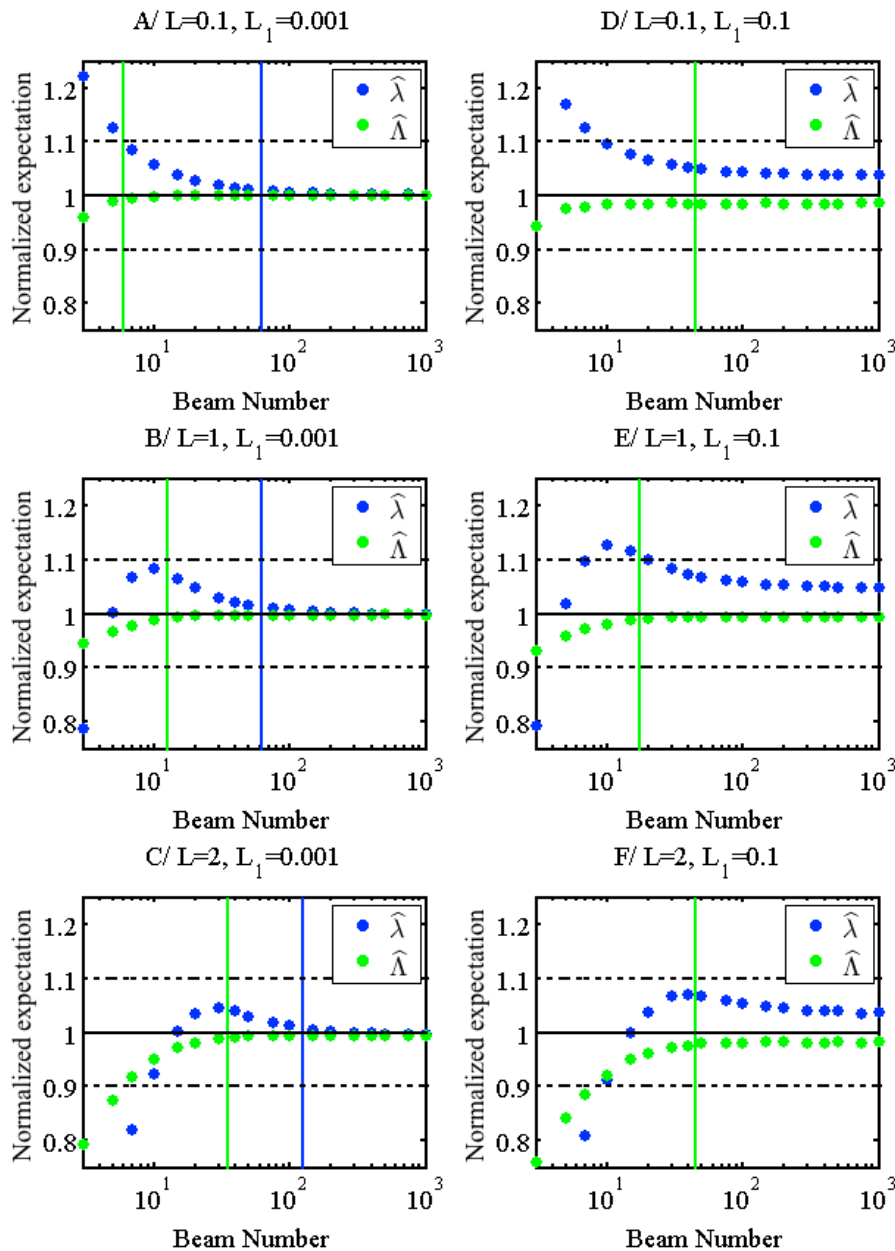
Fig. 4. Same as Figure 3 for the usual and unbiased Beer-Lambert estimators.

Expectations of the Beer-Lambert-attenuation-coefficient estimators, as a function of the beam number. Blue dots corresponds the usual Beer-Lambert estimator ($\hat{\lambda}$). Green dots corresponds to the unbiased Beer-Lambert estimator ($\hat{\Lambda}$), that accounts for element size and beam number and extended definition when I=1. Estimators are normalized by their true value λ, so that they are consistent when the expectation equals to one. The vertical lines correspond to the lowest values of N leading to a bias smaller than 1% in blue and green, for respectively the biased and unbiased estimators.

539

540   The computations of biases are done for all values of parameters in Table 1 and lead to the

541   range of validity for three levels of consistency (1%, 5%, 10%) for the four estimators

542   summarized in Table 2. As expected from Figures 3 and 4, the usual Beer-Lambert and

543   modified contact frequency are biased in a much wider range than the corrected indices

544   introduced in section 3, and generally requires smaller elements and a larger beam number to

545   be consistent. Also, the biases of the Beer-Lambert estimators increase with density, which

546   leads to less straightforward formulations of range of consistency, since their definition vary

547   with both L_1 and L. MLE thus has wider range of validity than Beer-Lambert estimators.

548

549   **Table 2.** Range of consistency of the four estimators of attenuation coefficient for three consistency

550   thresholds (biases smaller than 1, 5 and 10 %). NB: According to numerical simulations, the bias

551   thresholds are never reached for values of L_1 and L that are out of the ranges provided below (even

552   when N =10000).

| Index	Consistency (1%)	Consistency (5%)	Consistency (10%)
$\tilde{\lambda}$	$L_1 \leq 0.01$ and N $\geq$ 100	$\begin{cases} L_1 \leq 0.01 \text{ and N} \geq 20 \\ L_1 \leq 0.05 \text{ and N} \geq 30 \end{cases}$	$\begin{cases} L_1 \leq 0.01 \text{ and N} \geq 10 \\ L_1 \leq 0.1 \text{ and N} \geq 20 \end{cases}$
$\tilde{\Lambda}$	$\begin{cases} L_1 \leq 0.01 \text{ and N} \geq 3 \\ L_1 \leq 0.1 \text{ and N} \geq 5 \\ L_1 \leq 0.2 \text{ and N} \geq 15 \\ L_1 \leq 0.3 \text{ and N} \geq 30 \end{cases}$	$\begin{cases} L_1 \leq 0.05 \text{ and N} \geq 3 \\ L_1 \leq 0.1 \text{ and N} \geq 5 \\ L_1 \leq 0.3 \text{ and N} \geq 10 \end{cases}$	$\begin{cases} L_1 \leq 0.1 \text{ and N} \geq 3 \\ L_1 \leq 0.2 \text{ and N} \geq 5 \\ L_1 \leq 0.3 \text{ and N} \geq 7 \\ L_1 \leq 0.5 \text{ and N} \geq 10 \end{cases}$
$\hat{\lambda}$	$L \leq 2$ and $L_1 \leq 0.01$ and N $\geq$ 100	$L_1 \leq 0.05$ and N $\geq$ 40	$\begin{cases} L \leq 2 \text{ and } L_1 \leq 0.01 \text{ and N} \geq 10 \\ L \leq 2.5 \text{ and } L_1 \leq 0.1 \text{ and N} \geq 20 \\ L \leq 3 \text{ and } L_1 \leq 0.1 \text{ and N} \geq 30 \end{cases}$
$\hat{\Lambda}$	$\begin{cases} L \leq 0.5 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 7 \\ L \leq 1 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 10 \\ L \leq 1.5 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 15 \\ L \leq 2 \text{ and } L_1 \leq 0.05 \text{ and N} \geq 40 \\ L \leq 2.5 \text{ and } L_1 \leq 0.005 \text{ and N} \geq 75 \\ L \leq 3 \text{ and } L_1 \leq 0.001 \text{ and N} \geq 75 \end{cases}$	$\begin{cases} L \leq 1 \text{ and } L_1 \leq 0.1 \text{ and N} \geq 5 \\ L \leq 1.5 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 10 \\ L \leq 2 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 15 \\ L \leq 2.5 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 40 \\ L \leq 3 \text{ and } L_1 \leq 0.1 \text{ and N} \geq 75 \end{cases}$	$\begin{cases} L \leq 1.5 \text{ and } L_1 \leq 0.2 \text{ and N} \geq 5 \\ L \leq 2 \text{ and } L_1 \leq 0.3 \text{ and N} \geq 10 \\ L \leq 2.5 \text{ and } L_1 \leq 0.3 \text{ and N} \geq 20 \\ L \leq 3 \text{ and } L_1 \leq 0.3 \text{ and N} \geq 40 \end{cases}$

553

554   *5.1.2 Estimator efficiency*

555   Figure 5 shows the empirical variances (multiplied by δ^2 for the generality of results) of

556   estimator derived from the MLE, similarly to Figure 3. As expected, the variances decay with

557   the beam number. When the elements are large and the density is moderate to high (subplots

558   E and F), variances remain significantly larger than zero, even for large beam numbers,

559   because of the variability between vegetation samples. The variances of the biased and

560   unbiased estimators are similar in magnitude, the former being slightly larger when the beam

561   number is small. Both variances are very close to the theoretical Cramer-Rao bound (in

562   black), when L_1 is small (subplots A, B and C). Since $\widetilde{\Lambda}$ is unbiased when L_1 is small, it can

563   thus be considered as efficient. When the elements are large and the vegetation is dense

564   (subplots E and F), the variance of $\widetilde{\Lambda}$ is much larger than the Cramer-Rao bound, even when N

565   is large. This is because the Cramer-Rao bound does not account for asymptotic variability

566   due to the variability of vegetation samples.

567   The green dashed lines corresponds to the expectation of the estimator of the variance of $\widetilde{\Lambda}$,

568   namely $\sigma_{\widetilde{\Lambda}}^2$ (Eq. 36). Its expectation is very close to the empirical expectation of the variance

569   of $\widetilde{\Lambda}$ (green dots), demonstrating the consistency of our variance estimator when the beam

570   number is larger than 5.

571

572
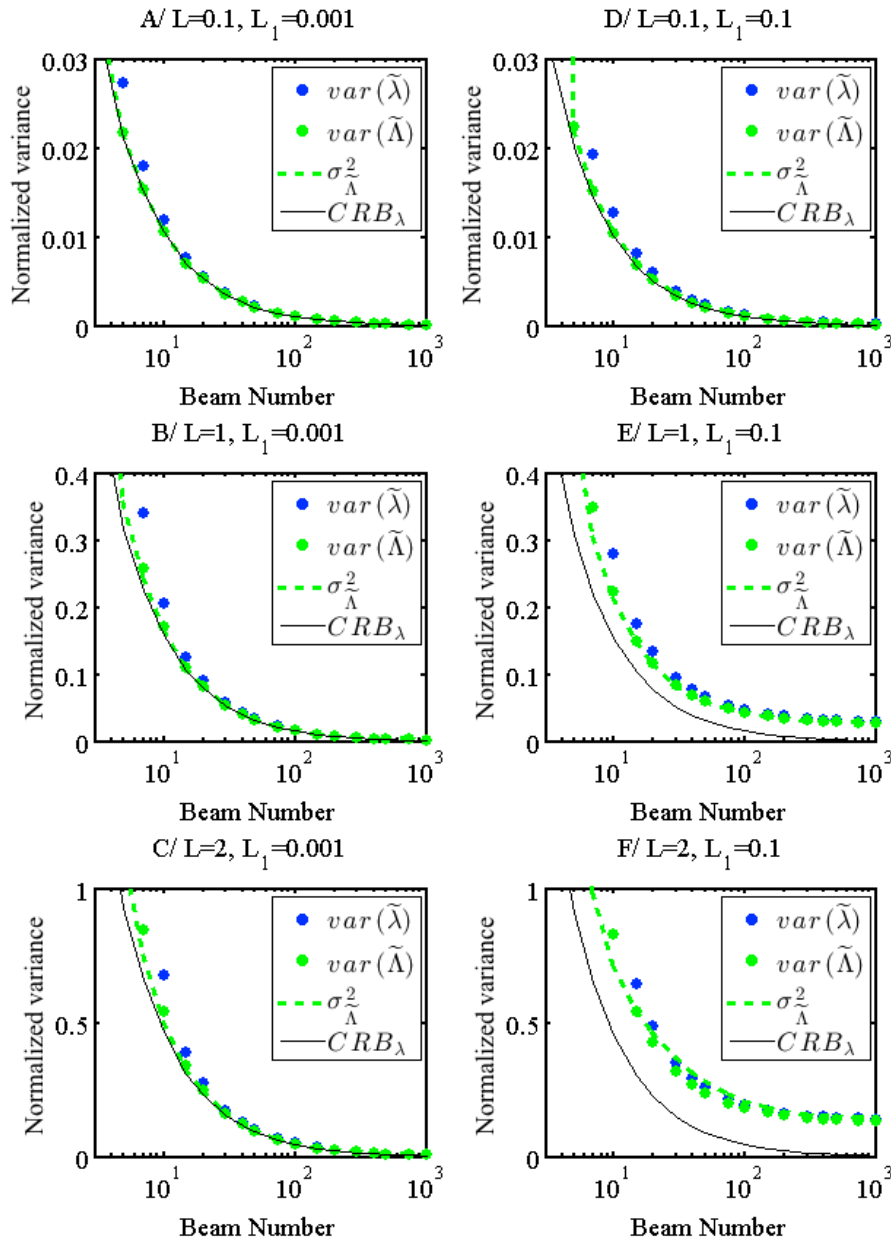
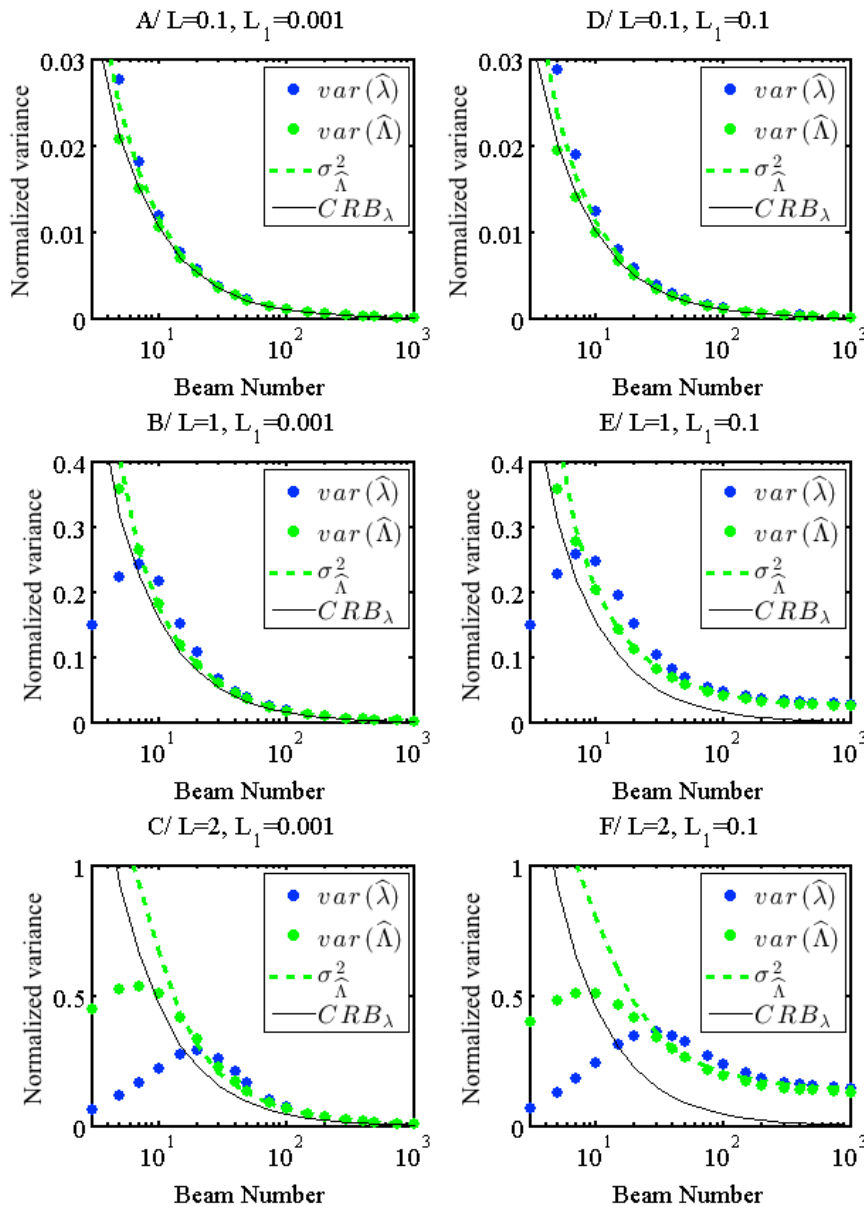Fig. 5. Normalized variances of attenuation coefficient estimators derived from the MLE, as a function

of the beam number. Normalized variances correspond to variances multiplied by δ^2. Blue dots

corresponds to the variance of the modified contact frequency estimator $\tilde{\lambda}$. Green dots corresponds to

the variance of the unbiased MLE $\widetilde{\Lambda}$. Green dashed lines correspond to the dimensionless expectation

of the variance estimator $\sigma^2_{\widetilde{\Lambda}}$. The black line corresponds to Cramer-Rao bound for the variance of

unbiased estimator.

579

580   Figure 6 is similar to Figure 5 for the variances of Beer-Lambert estimators. Although trends

581   are similar, it is worth noting that the variance of $\widehat{\Lambda}$ is greater than the Cramer-Rao bound for

582   small elements when vegetation is dense (Fig. 6C), showing that $\widehat{\Lambda}$ is suboptimal and less

583   efficient than the MLE. Also, the expectation of the variance estimator $\sigma_{\widehat{\Lambda}}^2$ can significantly

584   overestimate the empirical variance, showing a lack of consistency for this estimator. This

585   situation mostly occurs in range of data where $\widehat{\Lambda}$ itself is biased (dense vegetation, low

586   number of beams). The variance of the basic Beer-Lambert law can often be lower than the

587   Cramer-Rao bound. This is simply another evidence that this estimator is strongly biased (due

588   to the I=1 cases), since it would otherwise be greater than the Cramer-Rao bound.

589

590

591   Fig. 6. Same as Figure 5 for Beer-Lambert estimators. Normalized variances of Beer-Lambert

592   attenuation coefficient estimators, as a function of the beam number. Normalized variances correspond

593   to variances multiplied by δ^2. Blue dots corresponds to the variance of the usual Beer-Lambert

594   estimator $\hat{\lambda}$. Green dots corresponds to the variance of the unbiased Beer-Lambert estimator $\hat{\Lambda}$. Green

595   dashed lines correspond to the dimensionless expectation of the variance estimator $\sigma^2_{\hat{\Lambda}}$. The black line

596   corresponds to Cramer-Rao bound for the variance of unbiased estimator.

597

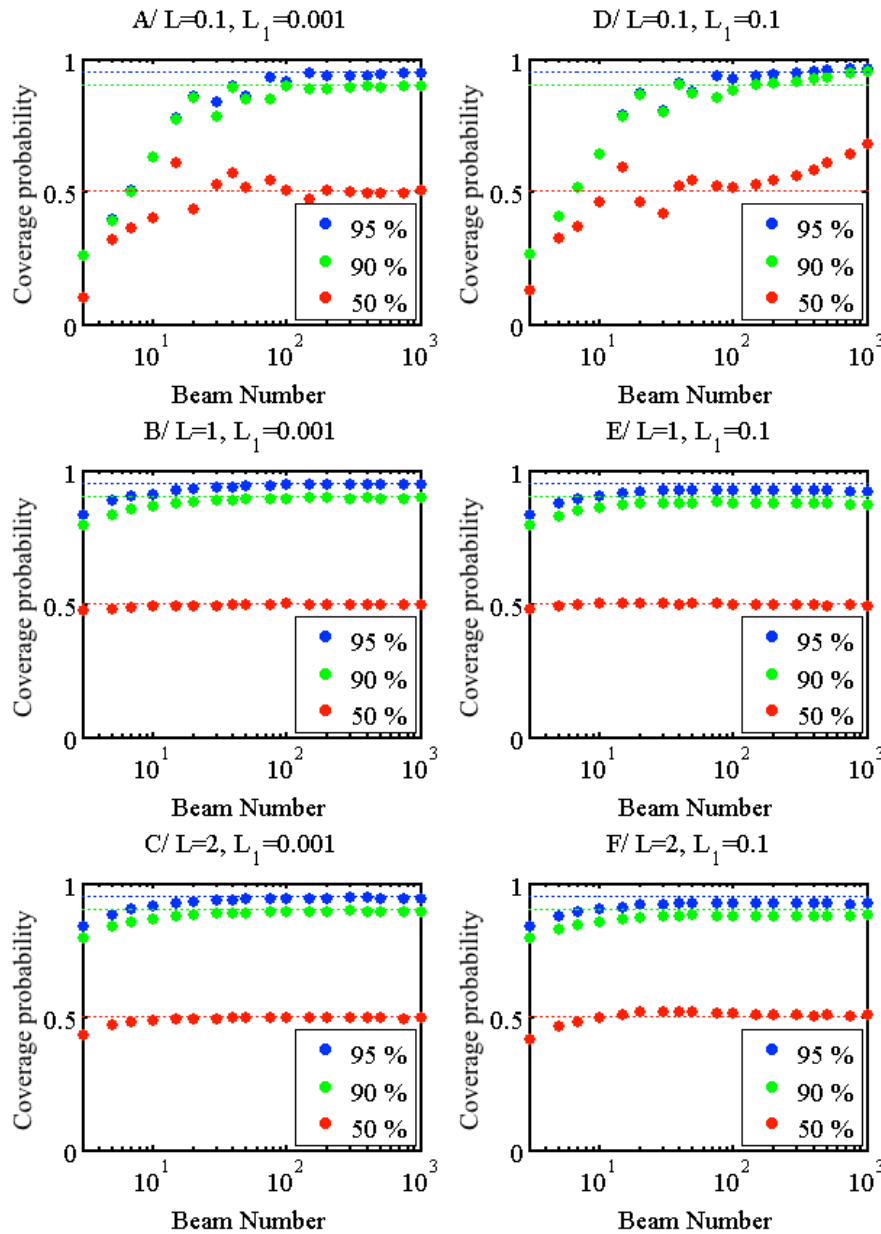598   *5.1.3 Coverage probabilities of the estimated confidence intervals*

599   Figure 7 shows the coverage probabilities of the estimated confidence interval based on

600   unbiased MLE, $\widetilde{\Lambda} \pm z_{\alpha/2}\sqrt{\sigma_{\widetilde{\Lambda}}^2}$ for three confidence levels (50, 90 and 95%). When the

601   confidence intervals are correctly estimated, empirical coverage probabilities should match

602   the confidence level (dashed lines). Estimated confidence intervals are satisfactory in most

603   cases, with the exception of low density when the beam number is low (subplots A and B), for

604   which the true value is less frequently in the confidence interval than expected.

605

606

607

608

609

610   Fig. 7. Coverage probabilities of the estimated confidence interval $\widetilde{\Lambda} \pm z_{\alpha/2}\sqrt{\sigma_{\widetilde{\Lambda}}^2}$ (computed with I),

611   function of the beam number, for 3 levels of confidence (50, 90 and 95%).
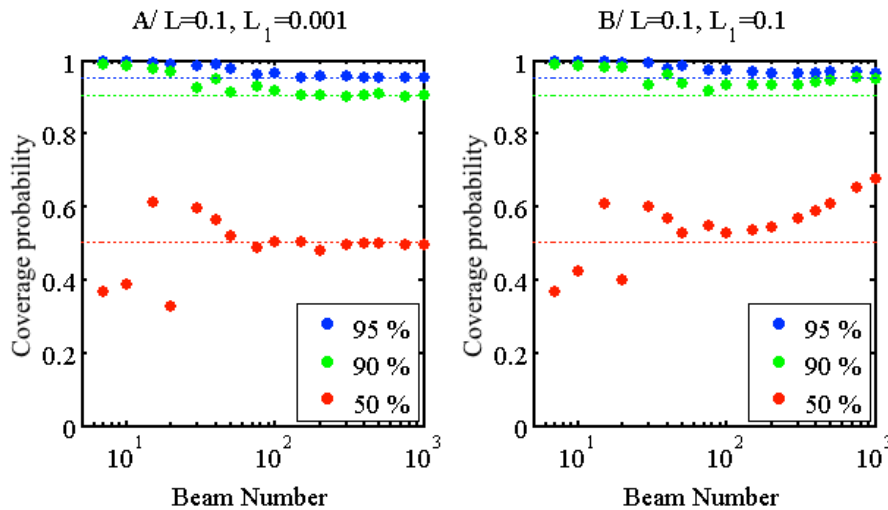
612

613   As explained in section 3.5, the alternative interval estimation based on Agresti-Coull

614   correction (Eq. (40)) leads to higher-than-expected coverage rates, as shown in Figure 8,

615   which is safer when density is low. Very similar intervals can be obtained for the unbiased

616   Beer-Lambert $\widehat{\Lambda}$ (not shown).

34

Fig. 8. Coverage probabilities of the estimated confidence interval $\widetilde{\Lambda}_c \pm z_{\alpha/2}\sqrt{\sigma^2_{\widetilde{\Lambda}_c}}$ (based on the Agresti-Coull values I_c and N_c), function of the beam number, for 3 confidence levels (50, 90 and 95%).

Coverage probabilities are computed for all simulated cases. For a given confidence level (90%, 95%), we can determine the range of parameter values (beam number, element and voxel depths) for which the coverage probabilities match the expected value (0.9, 0.95), within 5% for both the usual formulation of confidence interval (Eq. (39)) and for the "Agresti-Coull" interval (Eq. (40)). We find that using the "Agresti-Coull" interval increases the range of validity, when L is estimated lower than 0.5, but that the usual formulation performs better for higher densities. We thus adopt the following partical rule

$$
\begin{cases}
\widetilde{\Lambda}_c \pm z_{\frac{\alpha}{2}}\sqrt{\sigma^2_{\widetilde{\Lambda}_c}}, & \text{when } \tilde{L} \leq 0.5 \\
\widetilde{\Lambda} \pm z_{\frac{\alpha}{2}}\sqrt{\sigma^2_{\widetilde{\Lambda}}}, & \text{when } \tilde{L} > 0.5
\end{cases}
\tag{46}
$$

We summarize the ranges of validity of confidence intervals defined as in Eq. 46 in Table 3. Confidence intervals are consistent in a fairly large range of parameters. As in the results presented above, the range of validity of the unbiased "Beer-Lambert" confidence interval is

35

633   not as wide as the one of the unbiased MLE, especially when the voxel depth is larger than 2,

634   for which more than 100 beams are required.

635

636   **Table 3.** Range of validity of the confidence intervals at rate $\alpha = 0.90$ and 0.95. We consider that the

637   confidence interval is consistent, when the empirical probability reaches the expected level within 5%.

638   When L is estimated lower than 0.5, we use the "Agresti-Coull" confidence interval, whereas the usual

639   formulation is used otherwise (Eq. 46).

| Index	$\alpha = 0.90$ and Coverage probability within $0.9 \pm 5\%$	$\alpha = 0.95$ and coverage probability within $0.95 \pm 5\%$
$\widetilde{\Lambda}$	$\begin{cases} L \geq 0.1 \text{ and } L_1 \leq 0.1 \text{ and } N \geq 10 \\ L_1 \leq 0.1 \text{ and } N \geq 100 \end{cases}$	$\begin{cases} L \geq 0.1 \text{ and } L_1 \leq 0.1 \text{ and } N \geq 10 \\ L_1 \leq 0.1 \text{ and } N \geq 20 \end{cases}$
$\widehat{\Lambda}$	$\begin{cases} L \in [0.5; 2] \text{ and } L_1 \leq 0.05 \text{ and } N \geq 40 \\ L_1 \leq 0.01 \text{ and } N \geq 100 \\ L_1 \leq 0.05 \text{ and } N \geq 200 \end{cases}$	$\begin{cases} L \leq 2 \text{ and } L_1 \leq 0.05 \text{ and } N \geq 30 \\ L \leq 2.5 \text{ and } L_1 \leq 0.01 \text{ and } N \geq 100 \\ L_1 \leq 0.05 \text{ and } N \geq 150 \end{cases}$

640

641

642   *5.1.4. 95% errors for a single voxel and a group of voxels*

643   Figure 9 shows the expectation of the 95% errors for the MLE estimators in the same setting

644   as before. When the beam number is small and the density is low, this percentage can largely

645   exceed 100%. In these cases, the estimates remain very uncertain, although close to optimal

646   (Cramer-Rao-95%-error bound in black). The accuracy increases with vegetation density and

647   beam number. However, the 95% errors remain very high even for large beam number, when

648   elements are large because of the variability of vegetation samples (subplots E and F). At the

649   scale of a single voxel, using $\tilde{\lambda}$ or $\widetilde{\Lambda}$ leads to similar errors, which may be disappointing. This

650   is explained by the fact that the bias corrections accounted for in $\widetilde{\Lambda}$ are significant in a range

651   of parameter values for which variances are fairly large. Results are very different, however,

652   when errors are computed after averaging over several voxels (here, 100 voxels), which leads

653   to much smaller errors (Figure 10): using the unbiased estimator ($\widetilde{\Lambda}$) rather than the usual

654　　modified contact frequency ($\tilde{\lambda}$) leads to a reduction of the error on the order of 50%, typically

655　　in cases with low beam number or large elements, demonstrating the interest of bias

656　　corrections. When elements are large, 95%-errors are below 10% with $\tilde{\Lambda}$, when the beam

657　　number is greater than 100.

658　The expectation of the radiuses of the confidence interval $E95_{\tilde{\Lambda}}$ and $E95_{\tilde{\Lambda}}^{n_v}$ (green dashed line

659　　in Figure 9 and 10) are very close to the expectation of the 95% error, showing again that the

660　　estimated confidence intervals are consistent.
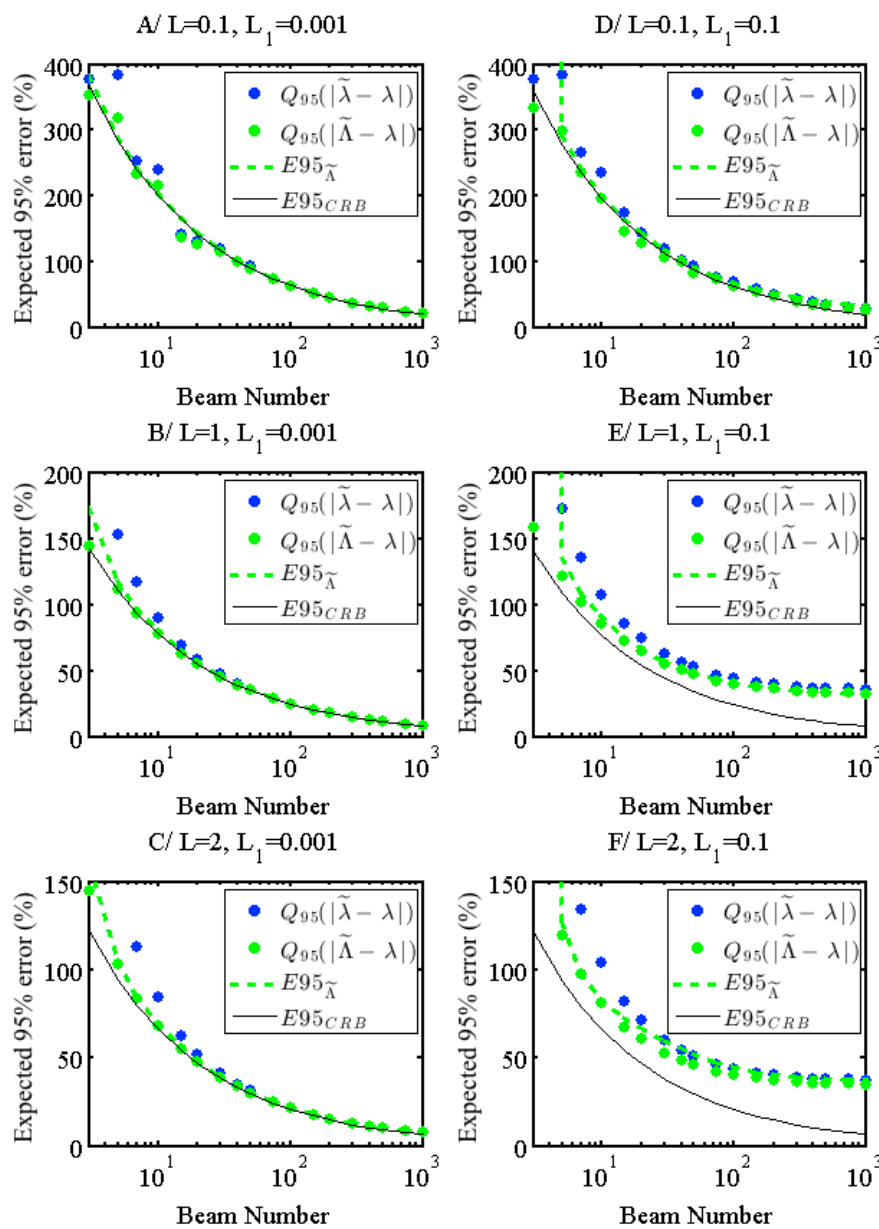
661

662   Fig. 9. Expectations of the 95% error, expressed in percentage of λ, as the 95th percentile of the

663   absolute residual to λ for $\tilde{\lambda}$ (blue dots) and $\widetilde{\Lambda}$ (green dots). We show for comparison the estimated

664   radius of the confidence interval $E95_{\widetilde{\Lambda}}$ (green dashed line) and the radius bound derived from the

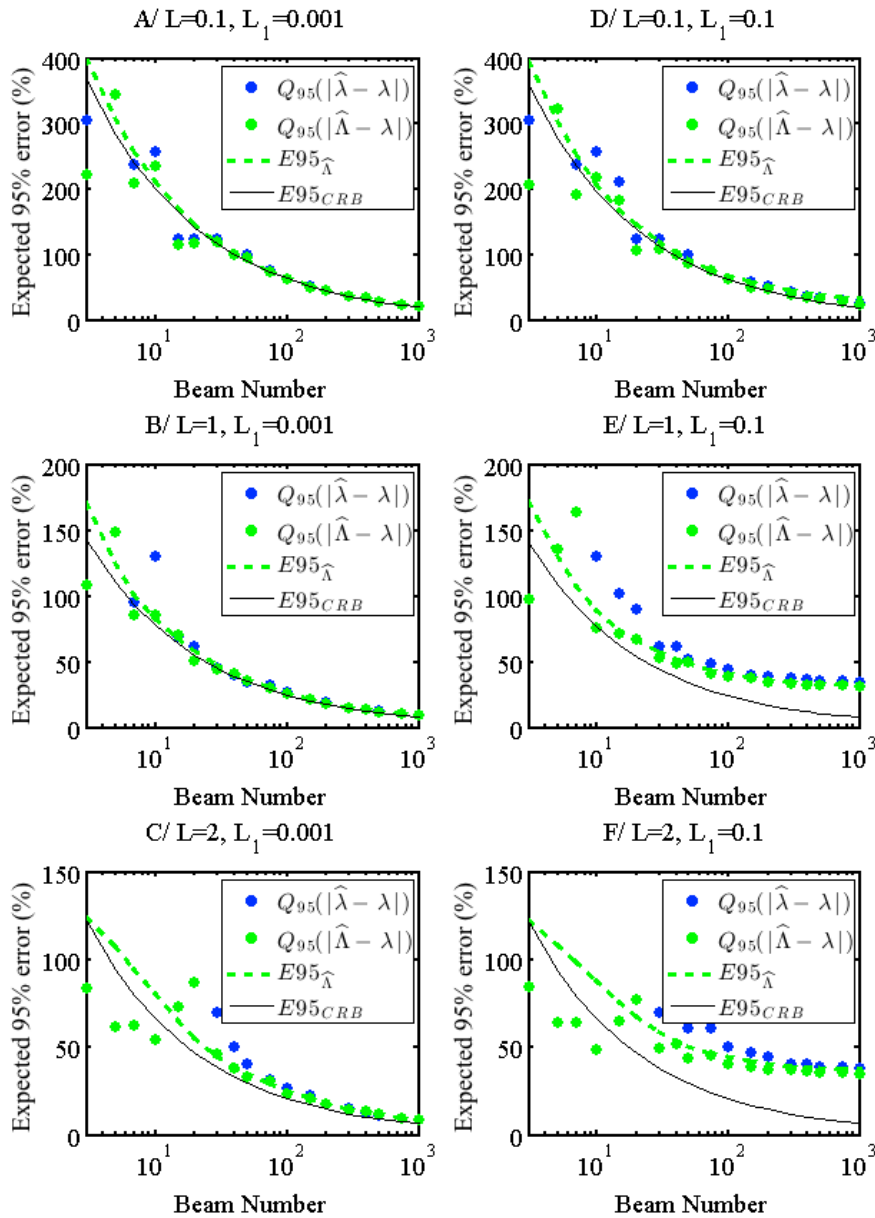665   Cramer Rao bound $E95_{\mathrm{CRB}}$ (black line).

666
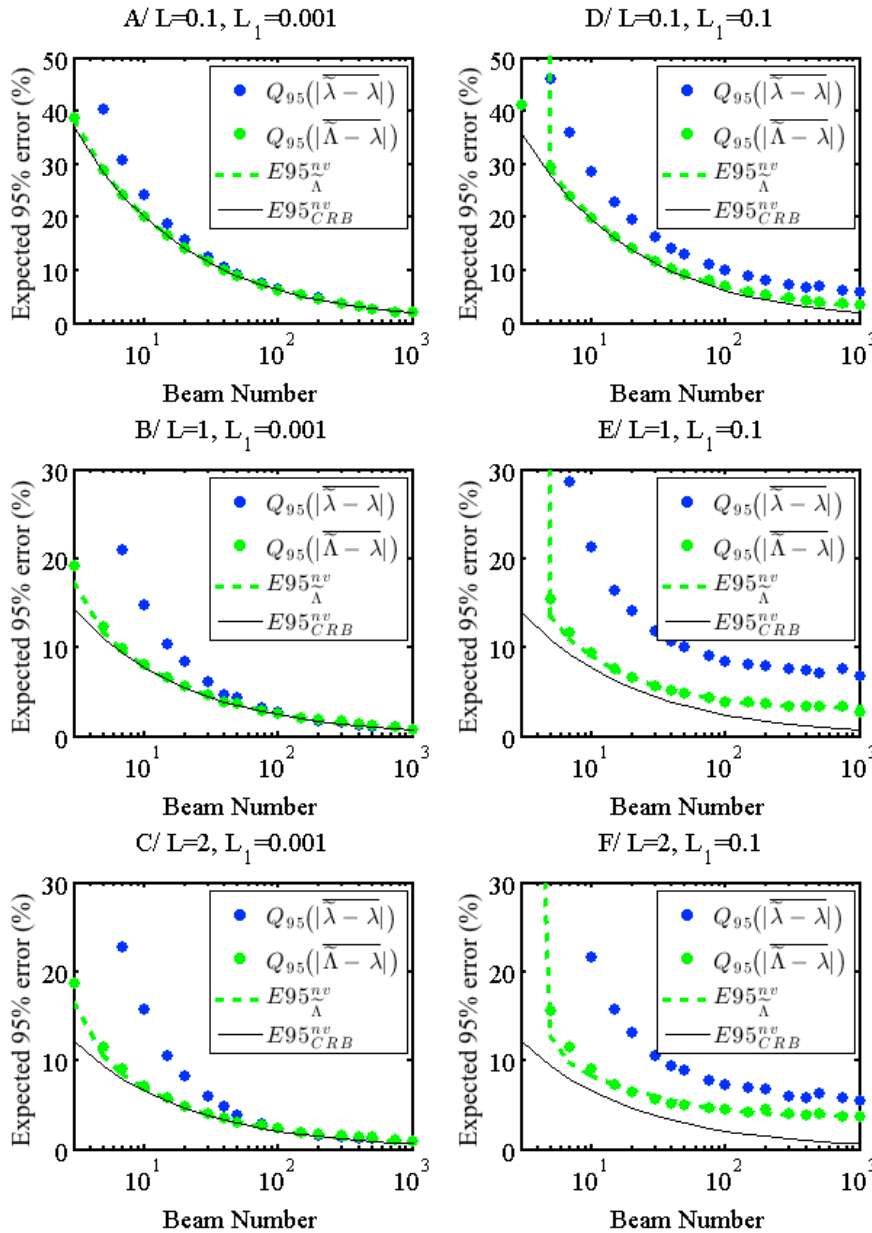
667

668   Fig. 10. Same as Fig. 9 for an average over 100 voxels. Expectations of the 95% error over nv=100

669   voxels, expressed in percentage of λ, as the 95th percentile of the absolute averaged residual to λ for $\tilde{\lambda}$

670   (blue dots) and $\widetilde{\Lambda}$ (green dots). We show for comparison the estimated radius of the confidence

671   interval $E95_{\widehat{\Lambda}}^{nv}$ (green dashed line) and the radius bound derived from the Cramer Rao bound $E95_{CRB}^{nv}$

672   (black line).

673

674   Figures 11 and 12 show similar trends to Figures 9 and 10, for both biased and unbiased Beer-

675   Lambert estimators. When blue dots are missing (usual Beer-Lambert), they correspond to

676   cases where I=1 in more than 5% of the voxel, so that the 95% error is in this case infinite. As

677   expected from previous results, the estimator of the radius of the 95% confident interval

678   (green dashed lines) is not consistent for the average over 100 voxels, in ranges where when $\widehat{\Lambda}$

679   is biased (typically L and L_1 large, Fig. 12F).

680

681   Fig. 11. Same as Fig. 9 for the Beer-Lambert estimators. Expectations of the 95% error, expressed in

682   percentage of λ, as the 95th percentile of the absolute residual to λ for $\hat{\lambda}$ (blue dots) and $\widehat{\Lambda}$ (green dots).

683   We show for comparison the estimated radius of the confidence interval $E95_{\widehat{\Lambda}}$ (green dashed line) and

684   the radius bound derived from the Cramer Rao bound $E95_{CRB}$ (black line).

685

686
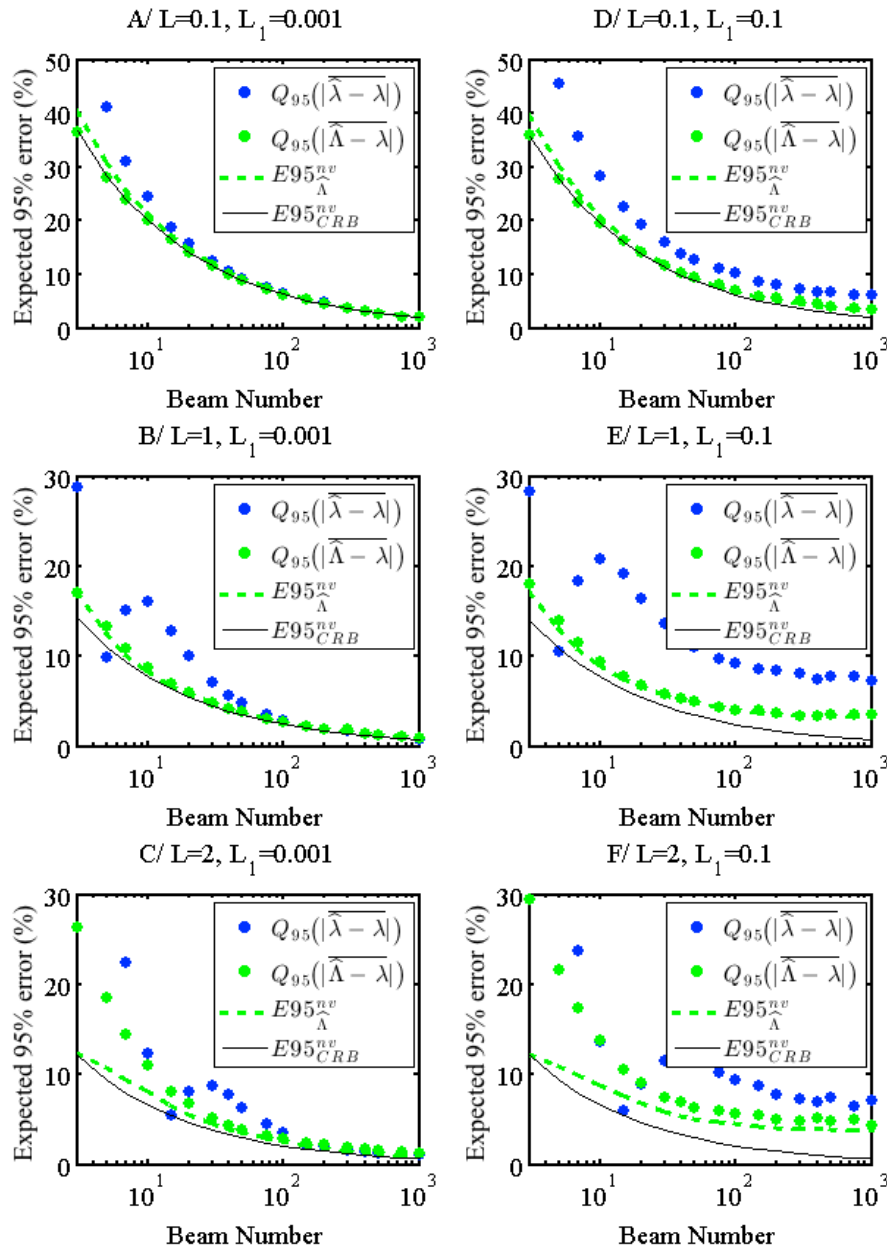
687

688   Fig. 12. Same as Fig. 10 for the Beer-Lambert estimators. Expectations of the 95% error over nv=100

689   voxels, expressed in percentage of λ, as the 95th percentile of the absolute averaged residual to λ for $\hat{\lambda}$

690   (blue dots) and $\hat{\Lambda}$ (green dots). We show for comparison the estimated radius of the confidence

691   interval $E95_{\hat{\Lambda}}^{nv}$ (green dashed line) and the radius bound derived from the Cramer Rao bound $E95_{\text{CRB}}^{nv}$

692   (black line).

693

694

695

696   *5.2. Estimator performance for unequal path lengths*

697

698   In this subsection, we show the statistics of estimators computed with simulations described

699   in subsection 4.3, in the context of a spherical voxel (unequal path lengths). Since MLE

700   performance is similar to the results shown in the previous section, it is not shown again.

701   Here, we focus on the comparison between Beer-Lambert estimators $\hat{\lambda}$ and $\widehat{\Lambda}_2$ (similar to $\widehat{\Lambda}$,

702   but which includes the correction for unequal path lengths). This is of major importance, since

703   Beer-Lambert law is mostly applied to cubic voxels, for which path lengths are generally not

704   equal. In this subsection, we assume that elements are infinitely small for simplicity (L_1=0).

705

706   *5.2.1 Estimator consistency*

707   Figure 13 shows the expectations of $\hat{\lambda}$ and $\widehat{\Lambda}_2$ similarly to Fig. 4. For low density (Fig. 13A),

708   the expectation of $\hat{\lambda}$ and $\widehat{\Lambda}_2$ are similar to those obtained with equal path lengths (Fig. 4A).

709   When density is higher (Fig. 13 B and C), the basic Beer-Lambert estimator $\hat{\lambda}$ is negatively

710   biased, and the bias does not tend to zero when the beam number is large.
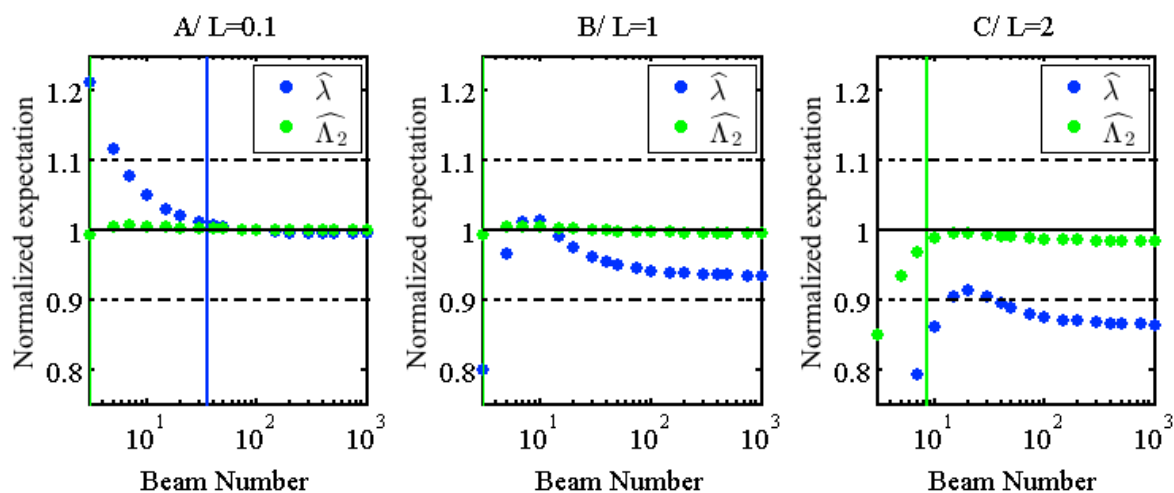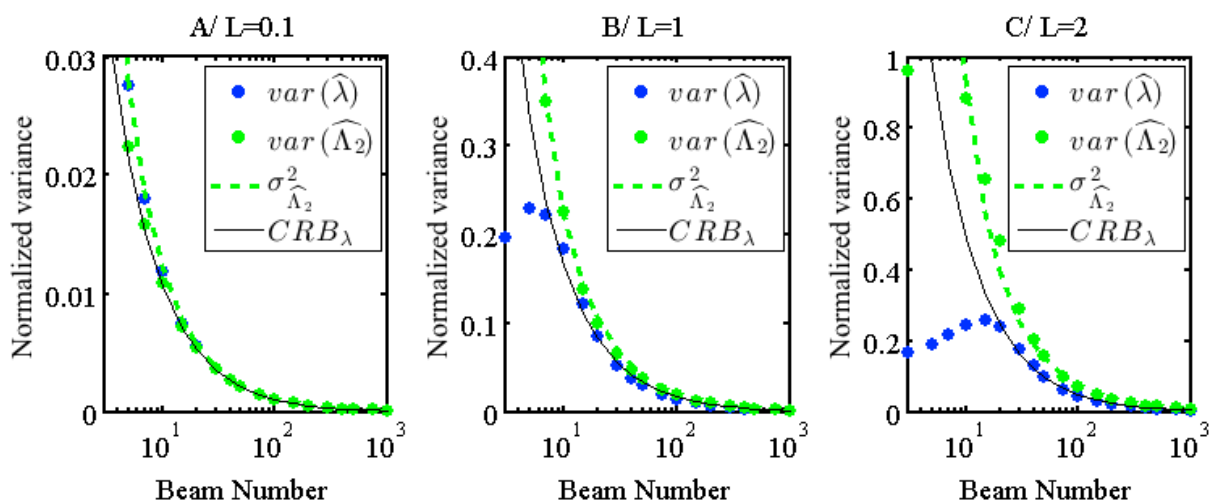

711

712   Fig. 13. Same as Figure 4 for the unequal path lengths. Expectations of the Beer-Lambert-attenuation-

713   coefficient estimators, as a function of the beam number. Blue dots corresponds the usual Beer-

714   Lambert estimator ($\hat{\lambda}$). Green dots corresponds to the unbiased Beer-Lambert estimator ($\widehat{\Lambda}_2$), that

715   accounts for unequal path length, element size and beam number and extended definition when I=1.

716   Estimators are normalized by their true value λ, so that they are consistent when the expectation equals

717   to one. The vertical lines correspond to the lowest values of N leading to a bias smaller than 1% in

718   blue and green, for respectively the biased and unbiased estimators.

719

720   *5.2.2 Estimator efficiency*

721   Figure 14 show the variances of $\hat{\lambda}$ and $\widehat{\Lambda}_2$ similarly to Fig. 6. For low density (Fig. 14A), the

722   variances of $\hat{\lambda}$ and $\widehat{\Lambda}_2$ are similar to those obtained with equal path lengths (Fig. 6A). When

723   density is higher (Fig. 14 B and C), the variance of $\widehat{\Lambda}_2$ (Green dots) is much larger, which is

724   mostly explained by the variability of the empirical path lengths. $\sigma^2_{\widehat{\Lambda}_2}$ (green dashed line)

725   provides consistent estimators for the variance of $\widehat{\Lambda}_2$, at least when the beam number is larger

726   than 10. This variance, however, is significantly larger than the Cramer-Rao bound (black

727   line), showing that $\widehat{\Lambda}_2$ is suboptimal when N is smaller than 100, contrary to the MLE, that

728   reaches the Cramer-Rao bound (not shown). Again, $\hat{\lambda}$ being biased, its variance cannot be

729   assessed against the Cramer-Rao bound to evaluate its efficiency.

730

731   Fig. 14. Same as Figure 6 for unequal path lengths. Normalized variances of Beer-Lambert attenuation

732   coefficient estimators, as a function of the beam number. Normalized variances correspond to

733   variances multiplied by δ^2. Blue dots corresponds to the variance of the usual Beer-Lambert estimator

734     $\hat{\lambda}$. Green dots corresponds to the variance of the unbiased Beer-Lambert estimator $\widehat{\Lambda}_2$. Green dashed

735   lines correspond to the dimensionless expectation of the variance estimator $\sigma^2_{\widehat{\Lambda}_2}$. The black line

736   corresponds to Cramer-Rao bound for the variance of unbiased estimator.

737

738   *5.2.3. 95% error of estimators for a single and a group of voxels*

739   The coverage probabilities of the estimated confidence interval $\widehat{\Lambda}_2 \pm z_{\alpha/2}\sqrt{\sigma^2_{\widehat{\Lambda}_2}}$ are similar to

740   those shown for $\widehat{\Lambda} \pm z_{\alpha/2}\sqrt{\sigma^2_{\widehat{\Lambda}}}$ in Figure 8 (and thus not shown). More interestingly, Figures 15

741   and 16 show the expectation of the 95% error, as in Figure 11 and 12. The 95% errors are

742   significantly reduced at the scale of a single voxel when the density is high and N is large

743   (Fig. 15C). As for the other bias correction, the error reduction is limited in other cases since

744   estimators are too uncertain. When averaged at the scale of several voxels, the benefit of the

745   correction for unequal path lengths is clearly visible when the beam number is moderate and

746   large (Fig. 16B and C). In these cases, 95% errors are always greater than 7 and 12%, even

747   when the beam number is large. Bias correction leads to an important reduction of the error,

748   that becomes close to the Cramer-Rao bound. However, contrary to unbiased MLE for which

749   the Cramer Rao bound is reached with unequal path length (not shown here, but logical since

750   the formulation is not affected by the path length), the Beer-Lambert estimator is not optimal

751   even after bias corrections. This is especially true when the beam number is small, because of

752   the variability induced by the empirical correction factor a_e. This demonstrates that the

753   unbiased MLE is clearly more efficient than the (unbiased) Beer-Lambert estimator, since

754   those estimators are mostly computed in cubic voxels.

755
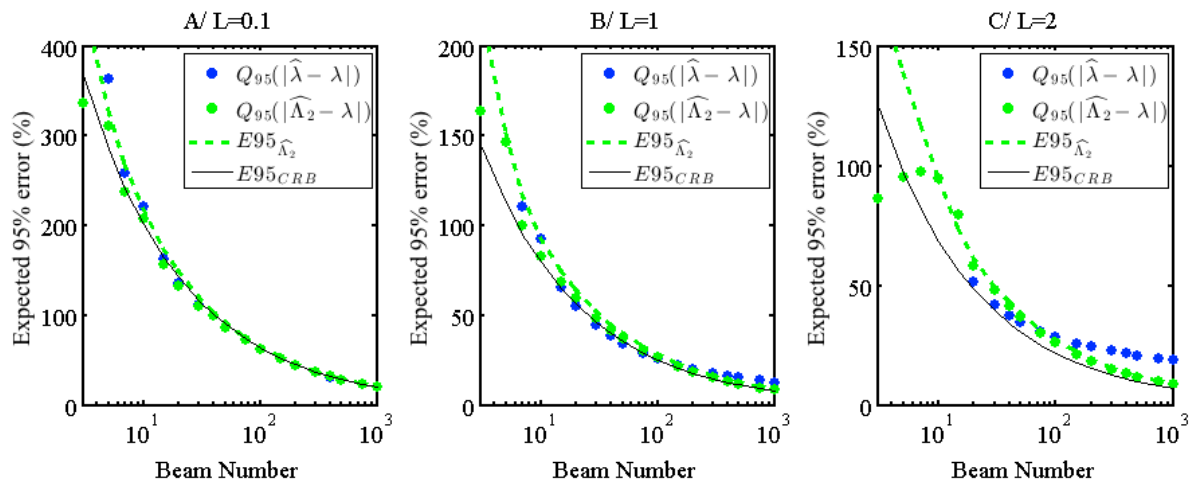
Fig. 15. Same as Fig. 11 for the unequal path lengths. Expectations of the 95% error, expressed in percentage of λ, as the 95th percentile of the absolute residual to λ for $\hat{\lambda}$ (blue dots) and $\widehat{\Lambda}_2$ (green dots). We show for comparison the estimated radius of the confidence interval $E95_{\widehat{\Lambda}_2}$ (green dashed line) and the radius bound derived from the Cramer Rao bound $E95_{\mathrm{CRB}}$ (black line).
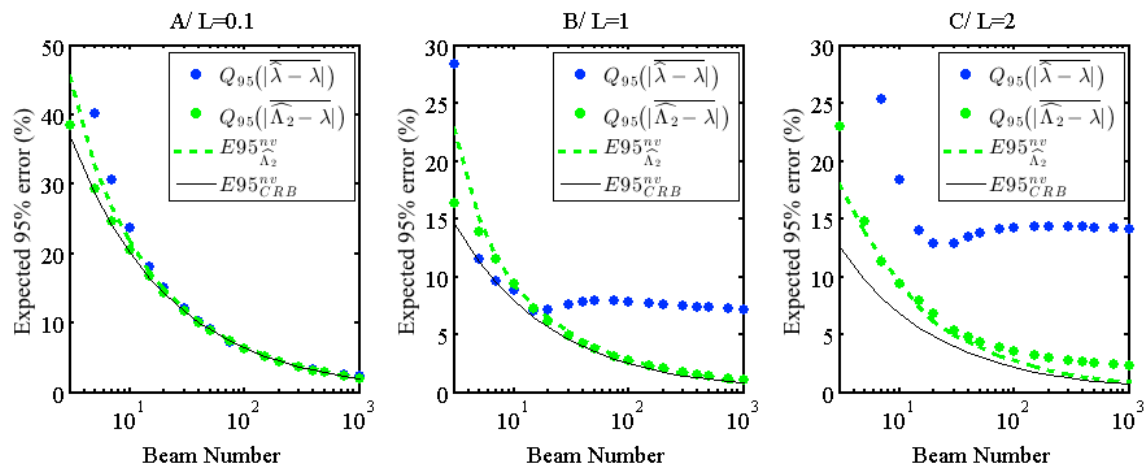
Fig. 16. Same as Fig. 12 for unequal path length. Expectations of the 95% error over nv=100 voxels, expressed in percentage of λ, as the 95th percentile of the absolute averaged residual to λ for $\hat{\lambda}$ (blue dots) and $\widehat{\Lambda}_2$ (green dots). We show for comparison the estimated radius of the confidence interval $E95_{\widehat{\Lambda}_2}^{nv}$ (green dashed line) and the radius bound derived from the Cramer Rao bound $E95_{\mathrm{CRB}}^{nv}$ (black line).

45

770

771

772

773   ***5. Discussion***

774

775   The method proposed here is based on a mathematical formulation of the estimation problem

776   in a formal statistical framework. Technical derivations are detailed in several supplementary

777   materials for reference, to facilitate the reading of the manuscript. This theoretical part is

778   completed with numerical simulations, for validation and determination of the range of

779   validity of our new estimators. As every modeling approach, both mathematical formulation

780   and simulations are based on assumptions that are not necessarily valid in the field. Here, we

781   assume that the spatial distribution of beams is random, whereas the actual one has a periodic

782   pattern, potentially altered by occlusion. Also, we assume simple square leaves for our

783   vegetation elements. More realistic patterns for beam shooting and vegetation elements can be

784   simulated (Grau et al., 2017). However, a drawback of this later approach is that it limits both

785   theoretical derivations and simulation number. The benefits of our simplifying assumptions

786   are that the mathematical framework can be deeply explored and that the cost of numerical

787   simulations is very limited, so that a full statistical analysis of estimator performance can be

788   done, over a wide range of parameter values (here element size, voxel size and beam number).

789

790   Our derivations entails to propose some new ready-to-use, analytical estimators for the

791   attenuation coefficient, which is proportional to PAD/LAD. These estimators generalize the

792   ones proposed in several pioneering studies. In this sense, our unbiased Beer-Lambert

793   estimator combines the effects of finite-element size and unequal path lengths that are already

794   identified respectively in Béland et al. (2014a), and Béland et al. (2014b) and Grau et al.

795   (2017). Regarding the effect of the element size, we choose to explicitly correct our estimator

796   for the associated bias, rather than to restrict its range of validity to largest voxels, as in

797   Béland et al. (2014a). For unequal path lengths, our formulation is more general than the

47

798   empirical correction proposed in Béland et al. (2014b), since it does not assume a particular

799   shape for the voxel. Contrary to the secant method, used for example in Bailey and Mahafee

800   (2017b), our formulation is analytical and easy to implement. Also, our Beer-Lambert

801   estimators are defined even when the RDI is equal to one, whereas earlier formulation

802   considered this special case as "occluded" (Béland et al., 2014a), leading to a negative bias as

803   shown above. Our approach also demonstrates that the modified contact frequency,

804   introduced in Béland et al. (2011), is indeed the Maximum Likelihood Estimator of the

805   mathematical problem. It extends the modified contact frequency initially developed for

806   infinitely small elements to the case of finite-size elements in a theoretically-based

807   formulation. This formulation slightly differs from the one proposed in the discussion in

808   Béland et al. (2014a), in which the correction term that accounts for finite-element size is the

809   same as for the Beer-Lambert law (see Eq. (12)). Such a proposition is not supported by the

810   Maximum Likelihood, but numerical consequences are probably limited. More importantly,

811   our formulation includes some bias corrections that depend on the beam number, for both

812   approaches (Beer-Lambert and MLE), whereas usual estimators are shown to be positively

813   biased to more than 20 % when the beam number is small. To the best of our knowledge, such

814   an effect has never been reported before.

815

816   Numerical simulations show that the new estimators are consistent for a much wider range of

817   parameter values (element size, attenuation coefficient, beam number), than the usual ones.

818   The range of consistency of the unbiased MLE $\widetilde{\Lambda}$ is wider than the one of the unbiased Beer-

819   Lambert estimator. Interestingly, the beam number required to reach consistency of the

820   unbiased MLE depends on the element depth L_1 only. Contrary to the unbiased Beer-Lambert

821   estimator, for which a larger number of beams is required when the vegetation density

822   increases, the unbiased MLE is not affected by the actual value of the attenuation coefficient.

823   This is practically convenient, since the attenuation coefficient is unknown prior to the

824   computation of the estimates when dealing with field data. When $L_1 \leq 0.1$, which is the case

825   for most vegetation when the voxel size is larger than 5 cm, $\widetilde{\Lambda}$ is consistent (bias smaller than

826   1%) as soon as the beam number is larger than 5. This is important when computing the

827   attenuation in small voxels, because the beam number in often low (e.g. Béland et al., 2009),

828   especially when vegetation is dense because of occlusion. Another potential application is

829   airborne LiDAR, for which the point density is much lower than TLS. For comparison, more

830   than 100 beams are required in vegetation with $L_1 \leq 0.01$ and $L \leq 2$ to reach the same

831   consistency with the usual Beer-Lambert estimator as with the unbiased MLE. Another

832   benefit of the MLE when compared to the Beer-Lambert estimator is that it does not require

833   any bias correction when path lengths are unequal. This is all the more important, that the bias

834   correction for unequal path lengths tends to reduce the efficiency of this estimator (in

835   comparison with Cramer-Rao bound). We also demonstrate that the unbiased MLE is

836   efficient, at least when the element size is small, since it reaches the Cramer-Rao bound. This

837   means that no unbiased estimator can have a smaller error than this estimator, so that the

838   unbiased MLE can be considered as optimal. This result is of major importance, since it

839   shows that there is no need for further correction, as long as the assumptions leading to these

840   results are valid.

841

842   Our mathematical derivations allow us to derive estimators of the variance and hence,

843   confidence interval for the unbiased estimators. Analyzing their coverage properties shows

844   that they are generally consistent, especially when using the formulation based on both the

845   "Agresti-Coull" interval and the basic interval, depending whether the voxel density is low or

846   high (Eq. 46). Providing such confidence intervals fill a gap for voxel-based approach, as

847   done earlier for gap fraction methods in Zhao et al. (2015). The first outcome of our

848   confidence interval estimates is that the prediction at voxel scale is uncertain, especially when

849   the voxels are small with errors larger than 100%, since the beam number is low and the

850   variability of the vegetation sample is high. When elements are large, uncertainty remains

851   high, even when scanning density is very high, because of the variability of element positions

852   within vegetation samples. The confidence intervals of the mean attenuation coefficient in a

853   larger volume (eventually discretized in small voxels) is much narrower, since the total beam

854   number is larger and the impact of the variability of vegetation samples is dampened. Our

855   numerical results, however, might be affected by our assumptions (random position, square

856   flat leaves, random beams), so that it would be worthwhile in the future to evaluate the

857   asymptotic residual variability in the context of realistic vegetation, for example following the

858   approach detailed in Grau et al. (2017).

859

860   Until now, most of the evaluation of the performance of voxel-based estimators was based on

861   the analysis of residual error between estimations and field data. The different sources of bias

862   and dispersion were thus merged. We believe that the applications of the new estimators and

863   their confidence interval should help to choose the appropriate resolution. Small voxels lead

864   to a smaller probability to get larger gaps that invalidate the assumption of random

865   distribution and results in an underestimation of LAD (Béland et al., 2014a). In other words,

866   transmission laws are wider-than-exponential in presence of spatial correlations (Davis and

867   Marshak, 2004; Pimont et al., 2009; Larsen and Clark, 2014). The question of resolution is

868   critical, since the recommendations vary among studies from some millimeters to 2 m (e.g.

869   Hosoi and Asama, 2006; Grau et al. 2017; Béland et al., 2011; Béland et al., 2014a; Pimont et

870   al., 2015; Bailey and Mahafee, 2017b). Among others, these studies deal with various

871   vegetation (various element size from needles to large leaves, various spatial distribution,

872   single tree vs forestry plot, etc.) as well as various scanning density (from single scan to high

873   density scanning). Also, the formulations vary among studies, some biases being corrected in

874   some studies, while not corrected in others. Most of them are affected by the positive biases

875   caused by the beam number and the element size (with the exception of Béland et al., 2014a

876   in which element size sensitivity is evaluated). Such biases are stronger at high resolution

877   since the beam number is lower and the element path is larger. Some of them are affected by

878   the negative biases of the usual Beer-Lambert estimator when path lengths are unequal, or

879   when the RDI is equal to 1. Again, such biases both vary with voxel resolution and

880   orientation. A general use of the unbiased MLE $\widetilde{\Lambda}$, for instance, should cancel these biases and

881   thus gives the opportunity to focus on the remaining sources of bias and dispersion pointed

882   out in the introduction. Among them, we can cite the TLS "flaws" (partial hit and detection

883   threshold) and the remaining vegetation "flaws" (element orientation and clumping, leaf and

884   wood separation). Vegetation heterogeneity is especially concerned by the issue of the voxel

885   size. At the end, the computation of confidence intervals could also help determine the

886   resolution that minimizes errors, since the resolution that minimizes the confidence interval

887   radius of the average attenuation coefficient within a given volume could be selected.

888

889   **6. Conclusion**

890

891   The present work provides an innovative approach of TLS point clouds, based on both

892   analytical derivations and numerical simulations to propose some new efficient estimators of

893   the attenuation coefficient, which is proportional to the LAD/PAD. These estimators are

894   designed for TLS point clouds of high density, so that they mostly concern TLS, although

895   their consistency with low beam number is also promising for their application to airborne

896   scanner. Among them, the unbiased MLE is consistent and efficient in a wider range of

897   parameter values than the usual estimators. It accounts for statistical biases associated with

898   beam number and element size. Although the biases caused by partial hit and clumping at

899   larger scale are not included, this new estimator should improve the choice of voxel

900   resolution, since it corrects several biases that depends on resolution and that might have been

901   mixed up in some earlier studies. Also, this work provides some estimators for the confidence

902   intervals of the attenuation coefficient within a volume containing one or several voxels,

903   increasing our knowledge of PAD/LAD regarding measurement accuracy by TLS, which is

904   probably lower than expected when voxels are small, and again contributes to the

905   determination of the best resolution.

906

907

909

910   ***References***

911   Allard, D., Lopez-Lozano, R., Baret, F., 2013 Modeling forest canopies with a hierarchical

912   multi-ring Boolean model for estimating Leaf Area Index. Spat. Stat-Neth 5, 42-56

913

914   Bailey, B.N., Mahafee, W.F., 2017a. Rapid, high-resolution measurement of leaf area and leaf

915   orientation using terrestrial LiDAR scanning data. Meas. Sci. Technol. 28(6).

916

917   Bailey, B.N., Mahafee, W.F., 2017b. Rapid measurement of the three-dimensional

918   distribution of leaf orientation and the leaf angle probability density function using terrestrial

919   LiDAR scanning. Remote Sens. Environ. 194, 63-76.

920

921   Béland, M., Widlowski, J.-L., Fournier, R.A., Côté, J.-F., Verstraete, M., 2011. Estimating

922   leaf area distribution in savanna trees from terrestrial LiDAR measurements. Agric. For.

923   Meteorol. 151, 1252–1266.

924

925   Béland, M., Baldocchi, D.D., Widlowski, J.-L., Fournier, R.A., Verstraete, M.M., 2014a. On

926   seing the wood from the leaves and the role of voxel size in determining leaf area distribution

927   of forests with terrestrial LiDAR. Agric. For. Meteorol. 184, 82–97.

928

929   Béland, M., Widlowski, J.-L., Fournier, R.A., 2014b. A model for deriving voxel-level tree

930   leaf area density estimates from ground-based LiDAR. Environ. Model. Softw. 51, 184–189.

931

932   Brown, L.D., Cai T.T., DasGupta, A., 2001. Interval Estimation for a Binomial Proportion.

933   Stat. Sci. 16(2), 101-133.

934

935　　Chen, J.M., Cihlar, J., 1995. Quantifying the effect of canopy architecture on optical

936　　measurements of leaf area index using two gap size analysis methods. Geosci. Remote Sens.

937　　IEEE Trans. 33, 777–787.

938

939　　Campbell　G.S. & Norman　J.N. (1998) An　Introduction　to　Environmental　Biophysics 2nd

940　　ed. Springer, New York.

941

942　　Davis, A.B., Marshak, A., 2004. Photon propagation in heterogeneous optical media with

943　　spatial　correlations:　enhanced　mean-free-paths　and　wider-than-exponential　free-path

944　　distributions. J. Quant. Spectrosc. Ra. 84, 3-34.

945

946　　Durrieu, S., Allouis, T., Fournier, R.A., Véga, C., Albrech, L., 2008. Spatial quantification of

947　　vegetation density from terrestrial laser scanner data for characterization of 3D Forest

948　　structure at plot level. In Proceedings of the SilviLaser, Edinburgh, UK, 17–19 September

949　　2008.

950

951　　Grau., E., Durrieu, S., Fournier, R., Gastellu-Etchegorry, J.P., Yin, T., 2017. Estimation of 3D

952　　vegetation density with Terrestrial Laser Scanning data using voxels. A sensitivity analysis of

953　　influencing parameters. Remote Sens. Environ. 191, 373 - 388.

954

955　　Hebert, M., Kroktov, E., 1992. 3-D Measurements from imaging laser radars: How good are

956　　they? Int. J. Image Vis. Comput. 10, 170–178.

957

958   Hosoi, F., Omasa, K., 2006. Voxel based 3D modeling of individual trees for estimating leaf

959   area density using high-resolution portable scanning LiDAR. IEEE Trans. Geosci. Remote

960   Sens. 44, 3610–3618.

961

962   Hosoi, F., Omasa, K., 2007. Factors contributing to accuracy in the estimation of the woody

963   canopy leaf area density profile using 3D portable LiDAR imaging. J. Exp. Bot. 58, 3463–

964   3473.

965

966   Jupp, D.L.B., Culvenor, D.S., Lovell, J.L., Newnham, G.J., Strahler, H., and Woodcock, C.E.,

967   2009. Estimating forest LAI profiles and structural parameters using a ground- based laser

968   called "Echidna". Tree Physiol. 29, 171–181.

969

970   Kay, S.M., 1993. *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice

971   Hall, 595 p.

972

973   Keane, R.E., 2015. Wildland fuel fundamentals and application. Springer International

974   Publishing, Switzerland, 191 p.

975

976   Kukenbrink, D., Schneider, F.D., Leiterer, R., Schaepman M.E., Morsdorf F., 2017.

977   Quantification of hidden canopy volume of airborne laser scanning data using a voxel

978   traversal algorithm. Remote Sens. Environ. 194, 424-436.

979

980   Nilson, T., 1971. A theoretical analysis of the frequency of gaps in plant stands. Agric.

981   Meteorol. 8 (C), 25–38.

982

983   Norman, J.M., Campbell, G.S., 1989. Canopy structure. In: Pearcy, R.W., Ehleringer,

984   J.,Mooney, H.A., Rundel, P.W. (Eds.), Plant Physiological Ecology – Field Methodsand

985   Instrumentation. Chapman and Hall, London, pp. 301–325.

986

987   Larsen, M.L., Clark, A.S., 2014. On the link between particle size and deviations from the

988   Beer-Lambert-Bouguer law for direct transmission. J. Quant. Spectrosc. Ra. 133, 646-651.

989

990   Pimont. F., Dupuy, J.-L., Caraglio, Y., Morvan, D., 2009. Effect of vegetation heterogeneity

991   on radiative transfer in forest fires. Int. J. Wildland Fire 18, 536–553.

992

993   Pimont. F., Dupuy, J.-L., Rigolot, R., Prat, V., Piboule, A., 2015. Estimating Leaf Bulk

994   Density Distribution in a Tree Canopy Using Terrestrial LiDAR and a Straightforward

995   Calibration Procedure. Remote Sens. *7*(6), 7995-8018.

996

997   Pisek, J., Sonnentag, O., Richardson, A.D., ottus, M.M., 2013. Is the spherical leaf inclination

998   angle distribution a valid assumption for temperate and boreal broadleaf tree species? Agric.

999   For. Meteorol. 169, 186–194.

1000

1001   Ross, J., 1981. The Radiation Regime and Architecture of Plant Stands. Task for Vegetation

1002   Sciences 3; Springer: The Hague, The Netherlands, p. 391.

1003

1004   Ruel, J.J., Ayres, M.P., 1999. Jensen's inequality predicts effects of environmental variation.

1005   Trends Ecol. Evol., 14, 361–366.

1006

1007   Schneider, R., Weil, W., 2008. Stochastic and integral geometry. Springer Science &

1008   Business Media.

1009

1010   Seielstad, C., Stonesifer, C., Rowell, E., Queen, L., 2011. Deriving Fuel Mass by Size Class

1011   in Douglas-fir (Pseudotsuga menziesii) Using Terrestrial Laser Scanning. Remote Sens. 3,

1012   1691-1709.

1013

1014   Stoyan, D., W. Kendall, J. Mecke. 1987. Stochastic Geometry and Its Applications. J. Wiley

1015   & Sons/Akademie-Verlag, Berlin 1987, 345 pp.

1016

1017   Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., Clawges, R.M., 2008. Lidar:

1018   shedding new light on habitat characterization and modeling. Frontiers Ecol. Environ. 6, 90–

1019   98.

1020

1021   Warren Wilson J., 1960. Inclined point quadrats. New Phytol. 59(1), 1-8.

1022

1023   Zhao, F., Yang, X., Schull, M.A., Román-Colón, M.O., Yao, T., Wang, Z., Zhang, Q., Jupp,

1024   D.L., Lovell, J.L., Culvenor, D.S., 2011. Measuring effective leaf area index, foliage profile:

1025   and stand height in New England forest stands using a full-waveform ground-based lidar.

1026   Remote Sens. Environ. 115, 2954–2964.

1027

1028   Zhao, K., García, M., Liu, S., Guo, Q., Chen, G., Zhang, X., Zhou, Y., Meng, X., 2015.

1029   Terrestrial lidar remote sensing of forests: Maximum likelihood estimates of canopy profile,

1030   leaf area index, and leaf angle distribution. Agric. For. Meteorol. 209-201, 100-113.

1031

1032   ***Supplementary Material***

1033   ***Supplementary S1. Expected range of the optical depth of an element of vegetation in a***

1034   ***voxel***

1035   Following the definition of λ_1 in Eq. (9), the optical depth of an element in a cubic voxel of size δ is

$$L_1 = \lambda_1 \delta = \frac{S_1}{S} \approx \frac{S_1}{\delta^2} \qquad \text{(S1-1)}$$

1036   For a needle of radius r and length l, this leads to:

$$\lambda_1 \delta \approx \frac{2\pi r l}{4\delta^2} \qquad \text{(S1-2)}$$

1037   For a (small) needle of diameter 2r = 0.5 mm and length l = 5 cm, we have:

$$\lambda_1 \delta_{min} \approx 2 \; 10^{-5} \delta^{-2} \qquad \text{(S1-3)}$$

1038   For a flat leaf of radius r, this leads to:

$$\lambda_1 \delta \approx \frac{2\pi r^2}{4\delta^2} \qquad \text{(S1-4)}$$

1039   For a (large) leaf of diameter 2r = 10 cm, we have:

$$\lambda_1 \delta_{max} \approx 5 \; 10^{-3} \delta^{-2} \qquad \text{(S1-5)}$$

1040

1041

1042

1043   ***Supplementary S2. Point estimators and their variance based on Beer-Lambert law***

1044   The usual Beer-Lambert estimator is based on the RDI. It assumes that (i) a mean path length $\overline{\delta}$ (~V/S) is

1045   suitable to account for unequal path length in the voxel and (ii) $-\log(1 - I)$ is a good estimator $\hat{\kappa}$ of $\kappa =$

1046 $-log(1 - E(I))$ (Note that κ is the optical depth of the voxel). However, in both cases, the non-linearity of the

1047   *log* function limits the validity of such assumptions and makes the standard estimator $\hat{\lambda}$ a biased estimator of λ.

1048

1049   Taking the log(1-x) of Eq. (20), we have:

$$\log\big(1 - E(I)\big) = \log\left(\frac{1}{S}\iint_{s\in S}(1 - \lambda_1\delta(s))^{\lambda/\lambda_1}dS\right) \tag{S2-1}$$

1050   The Lemma proved below enables to approximate the logarithm in (S2-1) and gives a second order

1051   approximation of $\log\big(1 - E(I)\big)$, as a function of the actual attenuation coefficient λ. Combining (S2-1) with the

1052   Lemma leads to:

$$\log\big(1 - E(I)\big) \approx -\overline{\delta_e}\lambda + \frac{1}{2}\sigma_{\delta_e}^2\lambda^2 \tag{S2-2}$$

1053   where the effective mean path length $\overline{\delta_e}$ and its variance $\sigma_{\delta_e}^2$ are defined as the mean and variance of the

1054   effective path lengths $\delta_{e,j} = -\frac{log(1-\lambda_1\delta_j)}{\lambda_1}$.

1055   We then consider the issue of the bias associated with $\log(1 - I)$. A bias correction can be computed applying

1056   the approximation S6-5 to the function $g(x) = \log(1 - x)$, which depends on $g''(I)$ and the variance of I (given

1057   by Eq. 23). An unbiased estimator of $\hat{\kappa}$ is thus

$$\hat{\kappa} = -\widehat{\log(1 - E(I))} = -\log(1 - I) - \frac{1}{2}\left(\frac{I(1 - I)}{N} + \sigma_{I_\infty}^2\right)(1 - I)^{-2} \tag{S2-3}$$

1058   Combining (S2-2) and (S2-3) leads to a second order polynomial in λ that can be solved to derive the corrected

1059   estimator $\widehat{\Lambda}_2$ accounting for unequal path lengths (S2-7).

$$\log(1 - I) - \frac{1}{2}\left(\frac{I(1 - I)}{N} + \sigma_{I_\infty}^2\right)(1 - I)^{-2} = -\overline{\delta_e}\lambda + \frac{1}{2}\sigma_{\delta_e}^2\lambda^2 \tag{S2-4}$$

1060   We first derive the estimator for nearly equal path length $\widehat{\Lambda}$. When path lengths are nearly constant ($\sigma_{\delta_e}^2 \approx 0$), the

1061   equation S2-4 in λ leads to:

$$\widehat{\Lambda} = -\frac{1}{\overline{\delta_e}}\left(\log(1 - I) + \frac{I}{2N(1 - I)} + \frac{\sigma_{I_\infty}^2}{2(1 - I)^2}\right) \approx -\frac{1}{\overline{\delta_e}}\left(\log(1 - I) + \frac{I}{2N(1 - I)}\right) \tag{S2-5}$$

1062   Numerical simulations show that $\frac{\sigma_{I_\infty}^2}{2(1-I)^2}$ is always very small compared to log(1-I) in the range of interest. It is

1063   thus neglected in the rest of the study.

1064

1065   For unequal path length, the second order polynomial in λ of (S2-4) can be rewritten (since by definition of $\widehat{\Lambda}$,

1066 $\log(1 - E(I)) = -\overline{\delta_e}\widehat{\Lambda}$):

$$\frac{1}{2}\sigma_{\delta_e}^2\lambda^2 - \overline{\delta_e}\lambda + \overline{\delta_e}\widehat{\Lambda} = 0 \qquad\qquad \text{(S2-6)}$$

1067   Assuming that $2\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}\widehat{\Lambda}$, is smaller than 1, we can solve the polynomial and keep the smallest root. This leads to

1068   the $\widehat{\Lambda}_2$ estimator that accounts for unequal path lengths:

$$\widehat{\Lambda}_2 = \frac{\overline{\delta_e}}{\sigma_{\delta_e}^2}\left(1 - \sqrt{1 - 2\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}\widehat{\Lambda}}\right) \qquad\qquad \text{(S2-7)}$$

1069   The above indices are not defined when I=1, since the Beer-Lambert approach does not provide any insight

1070   regarding the attenuation coefficient rather than "probably high". As explained in section 3.2, the center of the

1071   confidence interval can be estimated as a function of N, by the Agresti-Coull interval. With $z_{\alpha/2}^2 = 1$, it is:

$$I_c = \frac{1 + \frac{1}{2N}}{1 + \frac{1}{N}} = 1 - \frac{\frac{1}{2N}}{1 + \frac{1}{N}} = 1 - \frac{1}{2N+2} \qquad\qquad \text{(S2-8)}$$

1072   Since I_c is at the center of the confidence interval, $-\frac{\log(1-I_c)}{\overline{\delta_e}}$ is a more robust estimator for λ in this context:

$$\widehat{\Lambda} = \frac{\log(2N+2)}{\overline{\delta_e}} \qquad\qquad \text{(S2-9)}$$

1073   The estimator of the variance of $\widehat{\Lambda}$ is derived from (S6-2). Let $g(x) = -\left(\log(1-x) + \frac{x}{2N(1-x)}\right)$

$$g'(x) = \frac{1}{1-x} - \frac{(1-x) - x(-1)}{2N(1-x)^2} = \frac{1}{1-x}\left(1 - \frac{1}{2N(1-x)}\right) \qquad\qquad \text{(S2-10)}$$

1074   We can thus define the estimator of variance of $\widehat{\Lambda}$ as:

$$\sigma_{\widehat{\Lambda}}^2 = \frac{\sigma_I^2}{\overline{\delta_e}(1-I)^2}\left(1 - \frac{1}{2N(1-E(I))}\right)^2 \qquad\qquad \text{(S2-11)}$$

1075   Estimating the variance of $\widehat{\Lambda}_2$ as defined in (S2-7) can be done using (S6-2):

$$\sigma_{\widehat{\Lambda}_2}^2 = \sigma_{\widehat{\Lambda}}^2\left(\frac{\overline{\delta_e}}{\sigma_{\delta_e}^2}\frac{\frac{1}{2}2\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}}{\sqrt{1 - 2\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}\widehat{\Lambda}}}\right)^2 = \frac{\sigma_{\widehat{\Lambda}_2}^2}{1 - 2\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}\widehat{\Lambda}} \approx \sigma_{\widehat{\Lambda}_2}^2\left(1 + 2\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}\widehat{\Lambda} + 4\left(\frac{\sigma_{\delta_e}^2}{\overline{\delta_e}}\widehat{\Lambda}\right)^2\right) \qquad \text{(S2-12)}$$

1076

1077

1078   ***Proof of lemma (Eq. S2-2)***

1079   With $g(x) = \log(x)$, the integral formulation of (S6-1) leads to:

$$\frac{1}{S} \iint\limits_{s \in S} \log\big((1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}\big) dS \approx \log\left(\frac{1}{S} \iint\limits_{s \in S} (1 - \lambda_1 \delta(s))^{\frac{\lambda}{\lambda_1}} dS\right) - \frac{1}{2} \frac{var\big((1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}\big)}{\overline{(1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}}^2} \quad \text{(S2-13)}$$

1080   The left member is:

$$\frac{1}{S} \iint\limits_{s \in S} \log\big((1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}\big) dS = \frac{\lambda}{\lambda_1} \overline{\log(1 - \lambda_1 \delta)} = -\overline{\delta_e} \lambda \qquad \text{(S2-14)}$$

1081   With $g(x) = \exp(\lambda x)$, $var\big((1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}\big) = var\left(g\left(\frac{log(1 - \lambda_1 \delta(s))}{\lambda_1}\right)\right)$

1082   According to (S6-2),

$$var\big((1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}\big) \approx g'\left(\overline{\frac{log\big(1 - \lambda_1 \delta(s)\big)}{\lambda_1}}\right)^2 var\left(\frac{log\big(1 - \lambda_1 \delta(s)\big)}{\lambda_1}\right)$$

$$= \lambda^2 exp\left(\lambda \overline{\frac{log\big(1 - \lambda_1 \delta(s)\big)}{\lambda_1}}\right)^2 var(\delta_e) \qquad \text{(S2-15)}$$

1083   Since at the first order, $g(\overline{x}) = \overline{g(x)}$,

1084   we can write, $exp\left(\lambda \overline{\frac{log(1 - \lambda_1 \delta(s))}{\lambda_1}}\right) \approx \overline{exp\left(\lambda \frac{log(1 - \lambda_1 \delta(s))}{\lambda_1}\right)} = \overline{(1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}}$, and then:

$$\frac{1}{2} \frac{var\big((1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}\big)}{\overline{(1 - \lambda_1 \delta(s))^{\lambda/\lambda_1}}^2} \approx \frac{1}{2} \lambda^2 var(\delta_e) \qquad \text{(S2-16)}$$

1085   Combining this with the above results leads to:

$$\log\left(\frac{1}{S} \iint\limits_{s \in S} (1 - \lambda_1 \delta(s))^{\frac{\lambda}{\lambda_1}} dS\right) \approx -\overline{\delta_e} \lambda + \frac{1}{2} \sigma_{\delta_e}^2 \lambda^2 \qquad \text{(S2-17)}$$

1086

1087 ### *Supplementary S3. Point and variance estimators based on MLE*

1088 *S3.1. Log likelihood and MLE of the attenuation coefficient*

1089 Let use denote $\{z_j\}_{j=1,N}$ the N free paths with respective path lengths $\{\delta_j\}_{j=1,N}$.

1090 From Eq. (17), the likelihood of Z is:

$$\mathcal{L}(\lambda; z_1, z, \dots, z_N) = \prod_{j=1}^{N} f_Z(z_j; \delta_j) = \prod_{z_j < \delta_j} \lambda\left(1 - \lambda_1 z_j\right)^{\lambda/\lambda_1 - 1} \prod_{z_j = \delta_j} \left(1 - \lambda_1 z_j\right)^{\lambda/\lambda_1} \qquad \text{(S3-1)}$$

1091 The ML estimator is the value $\tilde{\lambda}$ that cancels the first derivative of $\mathcal{L}$ (Kay, 1993, chapter 7). Deriving the

1092 logarithm of the likelihood and equating to zero provides

$$\frac{d\log\mathcal{L}}{d\lambda} = \frac{N_i}{\lambda} + \sum_{j=1}^{N} \frac{\log\left(1 - \lambda_1 z_j\right)}{\lambda_1} = 0 \qquad \text{(S3-2)}$$

1093 Hence, with $z_{ej} = -\frac{\log(1 - \lambda_1 z_j)}{\lambda_1}$

$$MLE_\lambda = \frac{\lambda_1 N_i}{\sum_{j=1}^{N} \log\left(1 - \lambda_1 z_j\right)} = \frac{I}{\overline{z_e}} \qquad \text{(S3-3)}$$

1094

1095 *S3.2. Bias correction terms for the MLE*

1096 The bias correction for the MLE is derived from (S6-6) with $f(x,y) = \frac{x}{y}$, since the MLE is $\frac{I}{\overline{z_e}}$. The three terms

1097 corresponding to bias corrections are, assuming that beams are independent:

$$-\frac{1}{2}\sigma_I^2 \frac{\partial^2 f}{\partial x^2}(I, \overline{z_e}) = -\frac{1}{2}\sigma_I^2 \times 0 = 0 \qquad \text{(S3-4)}$$

1098

$$-\frac{1}{2}\sigma_{\overline{z_e}}^2 \frac{\partial^2 f}{\partial y^2}(I, \overline{z_e}) = -\frac{1}{2}\sigma_{\overline{z_e}}^2 \frac{2I}{\overline{z_e}^3} = -\frac{I}{N\overline{z_e}^3} var(z_e) \qquad \text{(S3-5)}$$

1099

$$-\sigma_{I,\overline{z_e}} \frac{\partial^2 f}{\partial x \partial y}(I, \overline{z_e}) = \sigma_{I,\overline{z_e}} \frac{1}{\overline{z_e}^2} = \frac{1}{N\overline{z_e}^2} covar\left(\mathbf{1}_{z_j < \delta_j}, z_e\right) \qquad \text{(S3-6)}$$

1100 Combining S3-4, S3-5, and S3-6 leads to:

$$\tilde{\Lambda} = \frac{I}{\overline{z_e}} - \frac{I}{N\overline{z_e}^3} var(z_e) + \frac{1}{N\overline{z_e}^2} covar\left(\mathbf{1}_{z_j < \delta_j}, z_e\right) \qquad \text{(S3-7)}$$

1101 Although practically computable for a distribution of $\{z_j\}$ and $\{\delta_j\}$, the variance of the estimator can not be

1102 analytically derived, so that it is not possible to estimate the variance and thus the confidence interval for the

1103   estimation. The next subsection is dedicated to the development of analytical estimator for $var(z_e)$ and

1104 $covar\left(\mathbf{1}_{z_j<\delta_j}, z_e\right).$

1105

1106   *S3.3. Estimates for $var(z_e)$ and $covar\left(\mathbf{1}_{z_j<\delta_j}, z_e\right)$*

1107   From variance formulae:

$$var(z_e) = E[z_e{}^2] - E[z_e]^2 \tag{S3-8}$$

1108   From the probability distribution (Eq. (17)) and definition of z_e, the expectation of z_e is:

$$E[z_e] = \frac{1}{S}\int_S\left(\int_0^{\delta(s)} -\frac{\log(1-\lambda_1 z)}{\lambda_1}\lambda(1-\lambda_1 z)^{\lambda/\lambda_1-1}dz + -\frac{\log(1-\lambda_1\delta(s))}{\lambda_1}(1-\lambda_1\delta(s))^{\lambda/\lambda_1}\right)ds \tag{S3-9}$$

1109   Integrating by parts the integral leads to:

$$E[z_e] = \frac{1}{S}\int_S\left(\left[\frac{\log(1-\lambda_1 z)}{\lambda_1}(1-\lambda_1 z)^{\frac{\lambda}{\lambda_1}}\right]_0^{\delta(s)}\right.$$
$$\left. -\int_0^{\delta(s)} -\frac{1}{1-\lambda_1 z}(1-\lambda_1 z)^{\frac{\lambda}{\lambda_1}}dz - \frac{\log(1-\lambda_1\delta(s))}{\lambda_1}(1-\lambda_1\delta(s))^{\frac{\lambda}{\lambda_1}}\right)ds \tag{S3-10}$$
$$= \frac{1}{S}\int_S\left(\int_0^{\delta(s)}(1-\lambda_1 z)^{\frac{\lambda}{\lambda_1}-1}dz\right)ds = \frac{P(Z<\delta)}{\lambda} = \frac{E\left[\mathbf{1}_{z_j<\delta_j}\right]}{\lambda}$$

1110   This demonstrates that:

$$\lambda = \frac{E\left[\mathbf{1}_{z_j<\delta_j}\right]}{E[z_e]} = \frac{I_\infty}{\overline{z_e}^\infty} \tag{S3-11}$$

1111   Similarly,

$$E[z_e{}^2] = \frac{1}{S}\int_S\left(\int_0^{\delta(s)}\left(\frac{\log(1-\lambda_1 z)}{\lambda_1}\right)^2\lambda(1-\lambda_1 z)^{\lambda/\lambda_1-1}dz + \left(\frac{\log(1-\lambda_1\delta(s))}{\lambda_1}\right)^2(1-\lambda_1\delta(s))^{\lambda/\lambda_1}\right)ds \tag{S3-12}$$

1112   Integrating twice by parts the integral leads to:

$$E[z_e{}^2] = \frac{1}{S} \int_S \left(\left[-\left(\frac{\log(1-\lambda_1 z)}{\lambda_1}\right)^2 (1-\lambda_1 z)^{\frac{\lambda}{\lambda_1}} \right]_0^{\delta(s)} \right.$$

$$-\int_0^{\delta(s)} \frac{2}{1-\lambda_1 z}\left(\frac{\log(1-\lambda_1 z)}{\lambda_1}\right)(1-\lambda_1 z)^{\frac{\lambda}{\lambda_1}} dz + \left(\frac{\log(1-\lambda_1\delta(s))}{\lambda_1}\right)^2 (1 \tag{S3-13}$$

$$\left. -\lambda_1\delta(s)^{\frac{\lambda}{\lambda_1}}\right)ds = \frac{2}{\lambda}\frac{1}{S}\int_S \left(\int_0^{\delta(s)} -\frac{\log(1-\lambda_1 z)}{\lambda_1}\lambda(1-\lambda_1 z)^{\frac{\lambda}{\lambda_1}-1}dz \right)ds$$

$$= \frac{2}{\lambda}E\left[\mathbf{1}_{z_j<\delta_j}z_e\right]$$

Thus, $$var(z_e) = \frac{2}{\lambda}E\left[\mathbf{1}_{z_j<\delta_j}z_e\right] - E[z_e]^2 \tag{S3-14}$$

And $$covar\left(\mathbf{1}_{z_j<\delta_j}, z_e\right) = E\left[\mathbf{1}_{z_j<\delta_j}z_e\right] - E\left[\mathbf{1}_{z_j<\delta_j}\right]E[z_e] \tag{S3-15}$$

1113

1114   *S3.4. Point estimator*

1115   Plugging S3-14 and S3-15 in, S3-7 leads to:

$$\widetilde{\Lambda} = \frac{I}{\overline{z_e}} - \frac{1}{N}\frac{I}{\overline{z_e}^3}\left(\frac{2}{\lambda}E\left[\mathbf{1}_{z_j<\delta_j}z_e\right] - E[z_e]^2\right) + \frac{1}{N\overline{z_e}^2}\left(E\left[\mathbf{1}_{z_j<\delta_j}z_e\right] - E\left[\mathbf{1}_{z_j<\delta_j}\right]E[z_e]\right) \tag{S3-16}$$

1116   And since $E\left[\mathbf{1}_{z_j<\delta_j}\right] \approx I$, $E[z_e] \approx \overline{z_e}$, $E\left[\mathbf{1}_{z_j<\delta_j}z_e\right] \approx \overline{\mathbf{1}_{z_j<\delta_j}z_e}$ and $\lambda \approx \frac{I}{\overline{z_e}}$:

$$\widetilde{\Lambda} = \frac{I}{\overline{z_e}} - \frac{\overline{\mathbf{1}_{z_j<\delta_j}z_e}}{N\overline{z_e}^2} \tag{S3-17}$$

1117

1118   *S3.5. Variance estimator*

1119   The variance for the MLE is derived from (S6-4) with $f(x,y) = \frac{x}{y}$, since the MLE is $\frac{I}{\overline{z_e}}$. The three terms

1120   corresponding to bias corrections are:

$$\sigma_I^2\left(\frac{\partial f}{\partial x}\right)^2(I,\overline{z_e}) = \frac{\sigma_I^2}{\overline{z_e}^2} = \frac{I(1-I)}{N\overline{z_e}^2} = \frac{I}{N\overline{z_e}^2} - \frac{I^2}{N\overline{z_e}^2} \tag{S3-18}$$

1121

$$\sigma_{\overline{z_e}}^2\left(\frac{\partial f}{\partial y}\right)^2(I,\overline{z_e}) = \frac{var(z_e)}{N}\frac{I^2}{\overline{z_e}^4} = \frac{1}{N}\left(\frac{2}{\lambda}\overline{\mathbf{1}_{z_j<\delta_j}z_e} - \overline{z_e}^2\right)\frac{I^2}{\overline{z_e}^4} = \frac{2}{\lambda}\frac{I^2}{N\overline{z_e}^4}\overline{\mathbf{1}_{z_j<\delta_j}z_e} - \frac{I^2}{N\overline{z_e}^2} \tag{S3-19}$$

1122

$$2\sigma_{I,\bar{z}_e}\left(\frac{\partial f}{\partial x}\right)(I,\bar{z}_e)\left(\frac{\partial f}{\partial y}\right)(I,\bar{z}_e) = 2\frac{\left(\overline{\mathbf{1}_{z_j<\delta_j}z_e} - I\bar{z}_e\right)}{N}\left(\frac{1}{\bar{z}_e}\right)\left(-\frac{I}{\bar{z}_e^2}\right)$$

$$= -\overline{\mathbf{1}_{z_j<\delta_j}z_e}\frac{2I}{N\bar{z}_e^3} + \frac{2I^2}{N\bar{z}_e^2}$$

(S3-20)

1123   Since $\lambda \approx \frac{I}{\bar{z}_e}$, summing S3-18 to S3-20 leads to:

$$var\left(\frac{I}{\bar{z}_e}\right) \approx \frac{I}{N\bar{z}_e^2}$$

(S3-21)

1124   (S3-14) can be rewritten:

$$\widetilde{\Lambda} = \frac{I}{\bar{z}_e}\left(1 - \frac{\overline{\mathbf{1}_{z_j<\delta_j}z_e}}{NI\bar{z}_e}\right) = \text{MLE}_\lambda\left(1 - \frac{\overline{\mathbf{1}_{z_j<\delta_j}z_e}}{NI\bar{z}_e}\right)$$

(S3-22)

1125 MLE_λ is corrected by a factor of which the variance is assumed to be small when compared to the variance of the

1126   standard MLE, so that, we can write (when I>0):

$$\sigma_{\widetilde{\Lambda}}^2 = \frac{I}{N\bar{z}_e^2}\left(1 - \frac{\overline{\mathbf{1}_{z_j<\delta_j}z_e}}{NI\bar{z}_e}\right)^2$$

(S3-23)

1127   This formulation does not account for the asymptotic variability of attenuation coefficient estimators, caused by

1128   the variability of element positions in vegetation samples. However, this asymptotic variability can be estimated

1129   from Eq. (31):

$$\lim_{N\to\infty}\sigma_{\widetilde{\Lambda}}^2 = \frac{\sigma_{I_\infty}^2\left(I,\lambda_1\bar{\delta}\right)}{\bar{\delta}_e^{-2}(1-I)^2}$$

(S3-24)

1130   And thus:

$$\sigma_{\widetilde{\Lambda}}^2 = \frac{I}{N\bar{z}_e^2}\left(1 - \frac{\overline{\mathbf{1}_{z_j<\delta_j}z_e}}{NI\bar{z}_e}\right)^2 + \frac{\sigma_{I_\infty}^2\left(I_b,\lambda_1\bar{\delta}\right)}{\bar{\delta}_e^2(1-I_b)^2}$$

(S3-25)

1131   With I_b defined as deal with case I=1 as in Supplementary S2:

$$I_b = \min\left(I, 1 - \frac{1}{2N+2}\right)$$

(S3-26)

1132   *S3.6. Cramer Rao bound*

1133   The Fisher information (Kay, 1993, Chapter 3) measures the amount of information that is carried about the

1134   attenuation coefficient, by the set of distances travelled by beams within a voxel $\{x_i\}_{i\leq N}$. It is defined as:

$$I_F(\lambda) = E\left[\left(\frac{d\log\mathcal{L}}{d\lambda}\right)^2\right] = -E\left[\frac{d^2\log\mathcal{L}}{d\lambda^2}\right]$$

(S3-27)

1135   Since the log likelihood function of Z (S3-1) is twice differentiable, the Fisher information can be expressed as:

$$I_F(\lambda) = \frac{E[N_i]}{\lambda^2} = \frac{NI_\infty(\lambda)}{\lambda^2}$$

(S3-28)

1136   The Cramer-Rao bound is the inverse of the Fisher Information:

$$\mathrm{CRB}_\lambda = \frac{\lambda^2}{\mathrm{NI}_\infty(\lambda)} \hspace{4cm} \text{(S3-28)}$$

1137   In the case of a spherical voxel, the Cramer-Rao bound can be analytically computed (see section S4.2).

1138   ***Supplementary S4. Technical derivation in a spherical voxel***

1139   *S4.1. Distribution of optical depths in a spherical voxel*
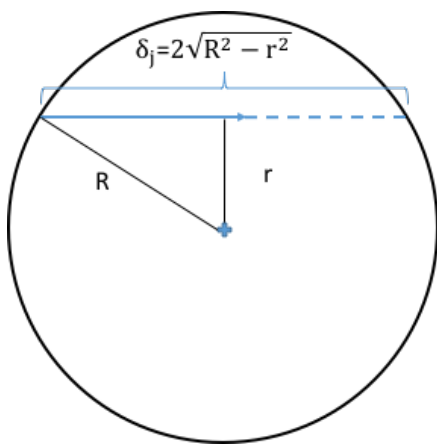
1140   In this subsection, we present the derivation leading to the PDF of optical depths $\{L_j\}$ in a spherical voxel,

1141   required for the numerical simulation described in subsection 4.3.

1142

1143   If r is the distance between the beam and the sphere center, the path length of the beam within the sphere is:

$$L_j = \lambda \delta_j = \lambda 2\sqrt{R^2 - r^2} \qquad (S4\text{-}1)$$

1144

1145

1146   Fig. S4-1. Illustration of the numerical simulation of a TLS beam crossing a spherical voxel with radius R and

1147   path length δ_i.

1148

1149   Assuming a constant beam density within the sphere, the beam density with length L_j is:

$$P\left(\lambda \delta_j(r) \le L_j < \lambda \delta_j(r + dr)\right) = \frac{2\pi r dr}{\pi R^2} = \frac{2r dr}{R^2} \qquad (S4\text{-}2)$$

1150   With $u = \frac{r}{R}$ between 0 and 1 and $h(u) = \sqrt{1 - u^2}$, it becomes:

$$P\left(\lambda 2Rh(u) \le L_j < \lambda 2Rh(u + du)\right) = 2udu \qquad (S4\text{-}3)$$

1151   Or equivalently, using (44):

$$P\left(L_j(u) \le L_j < L_j(u + du)\right) = 2udu \qquad (S4\text{-}4)$$

1152   with $L_j(u) = \frac{3}{2}L\sqrt{1 - u^2}$

1153

1154   *S4.2. Cramer-Rao bound for dimensionless spherical voxels*

1155   As defined in Supplementary S5.2, the Cramer-Rao bound for a spherical voxel is given by:

$$CRB_L = \frac{L^2}{NE(I)} \tag{S4-5}$$

1156   For the Beer-Lambert law, the expectation of the relative density index E(I) can be expressed as a function of the

1157   optical depth Li:

$$E[I] = 1 - \int_0^1 e^{-L_i(u)} 2u du \tag{S4-6}$$

1158   With $y = \frac{3}{2}L\sqrt{1-u^2}$, $dy = \frac{3}{2}L\frac{1}{\sqrt{1-u^2}}\frac{1}{2}(-2u du)$ so that $u du = \frac{4}{9L^2}y dy$ and:

$$\int_0^1 e^{-L_i(r)} 2u du = -\frac{8}{9L^2}\int_{\frac{3}{2}L}^0 e^{-y}y dy = \frac{8}{9L^2}\int_0^{\frac{3}{2}L} e^{-y}y dy \tag{S4-7}$$

1159   Integrating by parts:

$$\int_0^1 e^{-L_i(r)} 2u du = \frac{8}{9L^2}\left([-ye^{-y}]_0^{\frac{3}{2}L} - \int_0^{\frac{3}{2}L} -e^{-y}dy\right) = \frac{8}{9L^2}\left(1 - e^{-\frac{3}{2}L} - \frac{3}{2}Le^{-\frac{3}{2}L}\right) \tag{S4-8}$$

1160   Which leads to:

$$E[I] = 1 - \frac{8}{9L^2}\left(1 - e^{-\frac{3}{2}L} - \frac{3}{2}Le^{-\frac{3}{2}L}\right) \tag{S4-9}$$

1161   Thus

$$CRB_L = \frac{L^2}{N\left(1 - \frac{8}{9L^2}\left(1 - e^{-\frac{3}{2}L} - \frac{3}{2}Le^{-\frac{3}{2}L}\right)\right)} \tag{S4-10}$$

1162

1163

1164　　***Supplementary S5: Dimensionless quantities used in numerical simulations***

1165　　*S5.1. Finite element simulations (Described in section 4.2):*

1166　　The dimensionless quantities of interest for these numerical simulations are:

$$\mathbf{1}_e = \frac{\delta_e}{\delta} = -\frac{\log(1 - \lambda_1\delta)}{\lambda_1\delta} = -\frac{\log(1 - L_1)}{L_1} \tag{S5-1}$$

$$\hat{I} = \hat{\lambda}\delta = \begin{cases} -\log(1 - I) & \text{when } I < 1 \\ +\infty & \text{when } I = 1 \end{cases} \tag{S5-2}$$

$$\hat{L} = \hat{\Lambda}\delta = \begin{cases} -\frac{1}{1_e}\left(\log(1 - I) + \frac{I}{2N(1 - I)}\right) & \text{when } I < 1 \\ \frac{1}{1_e}\log(2N + 2) & \text{when } I = 1 \end{cases} \tag{S5-3}$$

$$\sigma_{\hat{L}}^2 = \sigma_{\hat{\Lambda}}^2\delta^2 = \begin{cases} \frac{I}{1_e^2(1 - I)}\left(\frac{1}{N} + h_\infty(I, L_1)\right)\left(1 - \frac{1}{2N(1 - I)}\right)^2 & \text{when } I < 1 \\ \frac{2N - 1}{1_e^2}\left(\frac{1}{N} + h_\infty\left(\frac{1}{2N + 2}, L_1\right)\right) & \text{when } I = 1 \end{cases} \tag{S5-4}$$

$$\overline{y_e} = \lambda\overline{z_e} = -\lambda\overline{\frac{\log(1 - \lambda_1 z_J)}{\lambda_1}} = -\frac{L}{L_1}\overline{\log\left(1 - \frac{L_1}{L}y_J\right)} \tag{S5-5}$$

$$\overline{\mathbf{1}_{y<L}y_e} = \lambda\overline{\mathbf{1}_{z<\delta}y_e} = -\lambda\overline{\frac{\mathbf{1}_{z<\delta}\log(1 - \lambda_1 z_J)}{\lambda_1}} = -\frac{L}{L_1}\overline{\mathbf{1}_{y<L}\log\left(1 - \frac{L_1}{L}y_J\right)} \tag{S5-6}$$

$$\tilde{I} = \tilde{\lambda}\delta = \frac{L}{\overline{y}}I \tag{S5-7}$$

$$\tilde{L} = \tilde{\Lambda}\delta = \frac{L}{\overline{y_e}}\left(I - \frac{\overline{\mathbf{1}_{y<L}y_e}}{N\overline{y_e}}\right) \tag{S5-8}$$

$$\sigma_{\tilde{L}}^2 = \sigma_{\tilde{\Lambda}}^2\delta^2 = \frac{L^2 I}{N\overline{y_e}^2}\left(1 - \frac{\overline{\mathbf{1}_{y<L}y_e}}{NI\overline{y_e}}\right)^2 + \frac{\sigma_{I_\infty}^2(I_b, L_1)}{1_e^2(1 - I_b)^2} \tag{S5-9}$$

$$CRB_L = \frac{L^2}{NE(I)} = \frac{L^2}{N\left(1 - (1 - L_1)^{L/L_1}\right)} \tag{S5-10}$$

1167

1168　　*S5.2. Unequal path length simulations (Described in section 4.3)*

1169　　The ratio of the volume to cross section of the voxel $\frac{V}{S}$ is $\overline{\delta}^\infty$, the limit of $\overline{\delta}$ when N tends to infinity, since a

1170　　constant surface density of beams is assumed. Thus, the asymptotic optical depth is:

$$L = \lambda\frac{V}{S} = \lambda\overline{\delta}^\infty \tag{S5-11}$$

1171　　With

$$\overline{L_e} = \lambda\overline{\delta_e} = \lambda\overline{\delta} \tag{S5-12}$$

1172   and

$$\sigma_L^2 = \sigma_\delta^2 \lambda^2 \tag{S5-13}$$

1173   The dimensionless quantities of interest are:

$$1_e = \frac{\overline{L_e}}{L} \tag{S5-14}$$

$$\hat{I} = \hat{\lambda}\delta = \begin{cases} -\dfrac{1}{1_e}\log(1 - I) & \text{when } I < 1 \\ +\infty & \text{when } I = 1 \end{cases} \tag{S5-15}$$

$$\hat{L} = \hat{\Lambda}\delta = \begin{cases} -\dfrac{1}{1_e}\left(\log(1 - I) + \dfrac{I}{2N(1 - I)}\right) & \text{when } I < 1 \\ \dfrac{1}{1_e}\log(2N + 2) & \text{when } I = 1 \end{cases} \tag{S5-16}$$

$$\hat{L}_2 = \hat{\Lambda}_2 \frac{V}{S} = \frac{V}{S}\frac{\overline{\delta}}{\sigma_\delta^2}\left(1 - \sqrt{1 - 2\frac{\sigma_\delta^2}{\overline{\delta}}\hat{\Lambda}}\right) = \frac{L\overline{L_e}}{\sigma_L^2}\left(1 - \sqrt{1 - 2\frac{\sigma_L^2}{L\overline{L_e}}\hat{L}}\right) \tag{S5-17}$$

$$\sigma_{\hat{L}}^2 = \sigma_{\hat{\Lambda}}^2\left(\frac{V}{S}\right)^2 = \begin{cases} \dfrac{I}{1_e{}^2 N(1 - I)}\left(1 - \dfrac{1}{2N(1 - I)}\right)^2 & \text{when } I < 1 \\ \dfrac{2 - \dfrac{1}{N}}{1_e{}^2} & \text{when } I = 1 \end{cases} \tag{S5-18}$$

$$\sigma_{\hat{L}_2}^2 = \sigma_{\hat{\Lambda}_2}^2\left(\frac{V}{S}\right)^2 = \sigma_{\hat{L}}^2\left(1 + \frac{\sigma_L^2}{L\overline{L_e}}\hat{L}\right)^2 \tag{S5-19}$$

$$CRB_L = \frac{L^2}{N\left(1 - \dfrac{8}{9L^2}\left(1 - e^{-\frac{3}{2}L} - \dfrac{3}{2}Le^{-\frac{3}{2}L}\right)\right)} \tag{S5-20}$$

1174

1175   ***Supplementary S6. Empirical expectation and variance of the function f of random***

1176   ***variable X or two random variables X and Y and method for bias correction.***

1177   Here, f is assumed continue and twice differentiable with continuous second derivative. Then the following

1178   second order approximations hold:

1179

$$\overline{f(X)} \approx f(\overline{X}) + \frac{1}{2}\sigma_X^2 f''(\overline{X}) \qquad \text{(S6-1)}$$

$$\text{Var}\big(f(X)\big) \approx \big(f'(\overline{X})\big)^2 \sigma_X^2 \qquad \text{(S6-2)}$$

$$\overline{f(X,Y)} \approx f(\overline{X},\overline{Y}) + \frac{1}{2}\sigma_X^2 \frac{\partial^2 f}{\partial x^2}(\overline{X},\overline{Y}) + \frac{1}{2}\sigma_Y^2 \frac{\partial^2 f}{\partial y^2}(\overline{X},\overline{Y}) + \sigma_{X,Y}\frac{\partial^2 f}{\partial x \partial y}(\overline{X},\overline{Y}) \qquad \text{(S6-3)}$$

$$\text{Var}\big(f(X,Y)\big) \approx \left(\frac{\partial f}{\partial x}(\overline{X},\overline{Y})\right)^2 \sigma_X^2 + \left(\frac{\partial f}{\partial y}(\overline{X},\overline{Y})\right)^2 \sigma_Y^2 + 2\frac{\partial f}{\partial x}(\overline{X},\overline{Y})\frac{\partial f}{\partial y}(\overline{X},\overline{Y})\sigma_{X,Y} \qquad \text{(S6-4)}$$

1180

1181   (S6-1) and (S6-3) are used to compute the bias correction for the estimators of $f(\overline{X})$ and $f(\overline{X},\overline{Y})$, since

$$\overline{f(X) - \frac{1}{2}\sigma_X^2 f''(X)} \approx f(\overline{X}) \qquad \text{(S6-5)}$$

1182   And

$$\overline{f(X,Y) - \frac{1}{2}\sigma_X^2 \frac{\partial^2 f}{\partial x^2}(X,Y) - \frac{1}{2}\sigma_Y^2 \frac{\partial^2 f}{\partial y^2}(X,Y) - \sigma_{X,Y}\frac{\partial^2 f}{\partial x \partial y}(X,Y)} \approx f(\overline{X},\overline{Y}) \qquad \text{(S6-6)}$$

1183

1184   **<u>Proof:</u>**

1185

1186   The proof of (S6-1) is the following. For a set of N value x_j of the random variable X:

$$f(x_j) = f\left(\overline{x}_j + (x_j - \overline{x}_j)\right) \approx f(\overline{x}_j) + (x_j - \overline{x}_j)f'(\overline{x}_j) + \frac{1}{2}(x_j - \overline{x}_j)^2 f''(\overline{x}_j)$$

1187   Summing over N and dividing by N leads to:

$$\frac{1}{N}\sum_{j=1}^{N} f(x_j) \approx f(\overline{x}_j) + \frac{f'(\overline{x}_j)}{N}\sum_{j=1}^{N}(x_j - \overline{x}_j) + \frac{1}{2}\frac{f''(\overline{x}_j)}{N}\sum_{j=1}^{N}(x_j - \overline{x}_j)^2$$

1188   The second term is by definition equal to 0. Thus, with σ_X^2 the usual unbiased estimates for variance:

$$\overline{f(x_j)} \approx f(\overline{x}_j) + \frac{1}{2}\frac{N-1}{N}\sigma_X^2 f''(\overline{x}_j) \approx f(\overline{x}_j) + \frac{1}{2}\sigma_X^2 f''(\overline{x}_j)$$

1189

1190   The proof of (S6-2) is similar:

$$\left(f(x_j) - f(\bar{x}_j)\right)^2 \approx \left((x_j - \bar{x}_j)f'(\bar{x}_j)\right)^2$$

1191   Thus,

$$\frac{1}{N-1}\sum_{j=1}^{N}\left(f(x_j) - f(\bar{x}_j)\right)^2 \approx \frac{f'(\bar{x}_j)^2}{N-1}\sum_{j=1}^{N}(x_j - \bar{x}_j)^2$$

1192   So that,

$$\frac{1}{N-1}\sum_{j=1}^{N}\left(f(x_j) - f(\bar{x}_j)\right)^2 \approx f'(\bar{x}_j)^2\sigma_X^2$$

1193

1194   Similar derivations for a function of two variables lead to (S6-3) and (S6-4). Then (S6-5) and (S6-6) follow.