A monte carlo appraisal of tree abundance and stand basal area estimation in forest inventories based on terrestrial laser scanning

Canadian Science Publishing | Canadian Journal of Forest Research

A Monte Carlo appraisal of tree abundance and stand basal area estimation in forest inventories based on terrestrial laser scanning

Journal:	*Canadian Journal of Forest Research*
Manuscript ID	cjfr-2017-0462.R2
Manuscript Type:	Article
Date Submitted by the Author:	24-Sep-2018
Complete List of Authors:	Corona, Piermaria; Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria Di Biase, Rosa; Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria; Università della Tuscia, DIBAF - Department for innovation in Biological, Agro-food and Forest systems Via San Camillo de Lellis s.n.c Fattorini, Lorenzo; Università di Siena, Dipartimento di Economia Politica e Statistica. Piazza San Francesco 8 D'Amati, Michela; Università di Siena, Dipartimento di Economia Politica e Statistica. Piazza San Francesco 8
Keyword:	plot sampling, TLS-based detection, distance sampling, hybrid inference, simulation study
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

A Monte Carlo appraisal of tree abundance and stand basal area estimation in forest inventories based on terrestrial laser scanning

P. Corona[(1)], R.M. Di Biase[(2)*], L. Fattorini[(3)], M. D'Amati[(4)]

[(1)] piermaria.corona@crea.gov.it Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA), Research Centre for Forestry and Wood, Italy.

[(2)] rmdibiase@unitus.it Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA), Research Centre for Forestry and Wood, Italy. PhD student at University of Tuscia, Viterbo, Italy.

[(3)] lorenzo.fattorini@unisi.it University of Siena, Italy.

[(4)] damati@student.unisi.it University of Siena, Italy

* Corresponding author: Rosa Maria Di Biase, University of Tuscia, Via San Camillo de Lellis s.n.c.- 01100 Viterbo, Italy. Telephone number: 3274948887. Email address: rmdibiase@unitus.it

Abstract

Non-detection of trees is an important issue when using single-scan TLS in forest inventories. A hybrid inference approach is adopted. Quoting from distance sampling, a detection function is assumed, so that the inclusion probability of each tree included within each plot can be determined. A simulation study is performed to compare the TLS-based estimators corrected and uncorrected for non-detection with the Horvitz-Thompson estimator based on conventional plot sampling, in which all the trees within plots are recorded. Results show that single-scan TLS provides more efficient estimators with respect to those provided by the conventional plot sampling in the case of low density forests when no distance sampling correction is performed. In low density forests, uncorrected estimators lead to a small bias (1-6%), increasing with plot size. Therefore, care must be taken in enlarging the plot radius too much. The bias increases in forests with clustered spatial structures and in dense forests, where the bias levels (30-50%) deteriorate the performance of uncorrected estimators. Even if the bias-corrected estimators prove to be effective in reducing the bias (below 15%), these reductions are not sufficient to outperform conventional plot sampling. Therefore, there is no convenience in using TLS based estimation in high density forests.

Keywords: plot sampling; TLS-based detection; distance sampling; hybrid inference; simulation study.

1. Introduction

Terrestrial Laser Scanning (TLS) demonstrated to be a promising tool for plot-level field inventories (e.g. Liang et al. 2011, Moskal and Zeng 2012, Kankare et al. 2015, Liang et al. 2016). The main advantages of using TLS lie in its capability to document the forest details automatically and at very fine (millimeter) spatial scales (Fardusi et al. 2017). In particular, the main advantage is the possibility to record the exact three-dimensional forest stand structure at any inventory occasion. This fact enables subsequent measurements on structural attributes not even considered before and very accurate time-series analyses. Two features particularly relevant for permanent survey programs like e.g. the National Forest Inventories.

In theory, several scans and high scanning resolution can provide complete information about the structure of the forest inventory plot. The statistical treatment of multi-scan data has been recently considered by Saarela et al. (2017). On the other hand, in single-scan operations, the amount of data is small compared to multi-scan and the registration of different scans is not needed. Therefore, both measurement and processing operations are faster and fully automatic. For these reasons, the single-scan approach is the most appealing in forest inventories, especially in those performed at large-scale (Liang et al. 2016).

TLS single location scans capture only a portion of the total tree volume, due to occlusion effects. However, the non-detection of trees by single-scan TLS can be accounted for under the statistical estimation framework of forest inventories by a hybrid inference approach (Corona et al. 2014). Distinctively, quoting from distance sampling, a detection function can be assumed to give the probability of detecting a tree as function of the distance from the point where TLS is located.

Ducey and Astrup (2013) and Astrup et al. (2014) suggested the use of distance sampling (Buckland et al. 2001) to identify detection probabilities and developed adjusted estimates for tree abundance and stand basal area. Ducey and Astrup (2013) compared conventional and TLS-based plot sampling in 9 forest stands that contained between 8 and 22 fixed radius plot, while Astrup et al. (2014) compared the two strategies in 166 plots in 12 mature stands. In both cases, conclusions were that TLS-based estimates are similar to those achieved by conventional plot sampling. However, because the comparisons were performed on sample data, unaware of the true number of trees and stand basal area, conclusions about the relative precision of the two estimation strategies are hard to reach.

The purpose of this paper is to perform an artificial comparison of conventional plot sampling, in which all the trees within field plots are accurately recorded by forest crews, vs TLS-based sampling, that, being based on an automatic detection of trees, allows for the use of larger plots. The comparison is performed by means of a Monte Carlo study based on simulated forests with low and high densities and random, trended and aggregated spatial patterns. Considering the time saved when trees are automatically recorded by a TLS device, the plot radii here tested in the TLS-based sampling are two-three times greater than those adopted in the standard estimation, in order to determine if: i) the information loss due to undetected trees can be compensated by the enlargement of the sampled area; ii) the non-detection adjustments performed by means of distance sampling technique provide improvements with respect to the non-adjusted estimates.

For avoiding ambiguities, it is worth noting that this study is completely design-based, i.e. populations are fixed and the uncertainty of estimators only stems from the random selection of plots onto the study area.

2. Preliminaries on conventional plot sampling estimation

Denote by U a collection of trees of size greater than a pre-fixed threshold within a study region $\mathcal{A}$ of size A. Suppose that the abundance N, i.e. the number of stems in the study region, and the total basal area T are the attributes to be estimated. Denoting the basal area of a single tree y_j, $j \in \mathsf{U}$, the total basal area is obviously given by

$$T = \sum_{j \in \mathsf{U}} y_j$$

Moreover, it is worth noting that even the abundance N may be viewed as the total of a dummy variable such that $y_j = 1$ for each tree $j \in \mathsf{U}$. Being both totals, T and N are usually estimated in forest surveys by using the plot sampling scheme joined with the well-known Horvitz-Thompson (HT) estimation criterion (Gregoire and Valentine 2008, chapter 7).

Plot sampling is usually performed by randomly locating a point onto the study region $\mathcal{A}$ enlarged by a buffer of width equal or greater to the pre-fixed plot radius r. Then, a plot of radius r is centered at the random point and the sample $\mathsf{S} \subset \mathsf{U}$ is the set of trees lying within the plot. The enlargement of the study region eliminates any edge effects (Gregoire and Valentine 2008, section 7.5), because each tree within the study area has a fixed inclusion zone given by the field plot centered at the tree location, which is completely included within the enlarged area. Therefore, the first-order inclusion probability of each tree is a / A^*, where $a = \pi r^2$ is the plot size and A^* is the size of the enlarged region.

From these inclusion probabilities, the HT estimator of abundance is

$$\hat{N}_{HT} = A^* \frac{n}{a} \tag{1}$$

where n is the number of trees observed within the plot, and the HT estimator of the total basal area is

$$\hat{T}_{HT} = \frac{A^*}{a} \sum_{j \in \mathsf{S}} y_j \tag{2}$$

From the general theory of the HT estimator (Gregoire and Valentine 2008, section 7.4.2), the two estimators are design-unbiased. The design-based variance of (1) is

$$V_{\hat{N}}^2 = \frac{A^*}{a} N + 2\frac{A^*}{a^2} \sum_{h>j \in \mathrm{U}} a_{jh} - N^2 \tag{3}$$

and the design-based variance of (2) is

$$V_{\hat{T}}^2 = \frac{A^*}{a} \sum_{j \in \mathrm{U}} y_j^2 + 2\frac{A^*}{a^2} \sum_{h>j \in \mathrm{U}} a_{jh} y_j y_h - T^2 \tag{4}$$

where a_{jh} is the size of the intersection of the two plots of radius r centered at the trees j and h. Obviously, any estimator based on a sole plot, even if unbiased, is destined to be highly imprecise. Usually, R plots are randomly and independently located within the enlarged study region in accordance with the sampling protocol usually referred to as uniform random sampling (URS). Under URS, the R plots give rise to R independent samples $\mathrm{S}_1,...,\mathrm{S}_R$ that in turn give rise to R identically and independent abundance estimates $\hat{N}_1,...,\hat{N}_R$ each of them with expectation N and variance $V_{\hat{N}}^2$ as well as R identically and independent basal area estimates $\hat{T}_1,...,\hat{T}_R$ each of them with expectation T and variance $V_{\hat{T}}^2$. Accordingly, the arithmetic mean of the R estimates,

$$\hat{\bar{N}}_R = \frac{1}{R}\sum_{i=1}^{R} \hat{N}_i \tag{5}$$

and

$$\hat{\bar{T}}_R = \frac{1}{R}\sum_{i=1}^{R} \hat{T}_i \tag{6}$$

are unbiased and consistent estimators of N and T, with variance $V_{\hat{N}}^2 / R$ and $V_{\hat{T}}^2 / R$, respectively. Unbiased and consistent estimators of the variances of (5) and (6) are given by $S_{\hat{N}}^2 / R$ and $S_{\hat{T}}^2 / R$, respectively, where

$$S_{\hat{N}}^{\ 2} = \frac{1}{R-1}\sum_{i=1}^{R}(\hat{N}_i - \bar{\hat{N}}_R)^2$$

and

$$S_{\hat{T}}^{\ 2} = \frac{1}{R-1}\sum_{i=1}^{R}(\hat{T}_i - \bar{\hat{T}}_R)^2$$

Moreover, owing to the Central Limit Theorem, $\bar{\hat{N}}_R$ and $\bar{\hat{T}}_R$ approach normality as R increases. Therefore, for a sufficiently large R, the confidence intervals $\bar{\hat{N}}_R \pm 2S_{\hat{N}} / \sqrt{R}$ and $\bar{\hat{T}}_R \pm 2S_{\hat{T}} / \sqrt{R}$ have an approximate coverage of 0.95.

3. TLS-based plot sampling estimation

Plot sampling surveys are performed by forest crews travelling the selected plots, counting the trees within plots and measuring their basal areas as well as any other attribute of interest. An alternative way to perform plot-based surveys could be the placement of a TLS device in the plot centers to automatically count the trees within the plot and measure their basal area. However, owing to the shadow provided by trees located near the device, some trees, especially those far from the center, remain undetected. Therefore, the actual sample is constituted by the set $\mathsf{D} \subset \mathsf{S}$ of the $m \le n$ detected trees. Accordingly, the HT estimators of abundance and basal area would be

$$\tilde{N}_{HT} = \sum_{j \in \mathsf{D}} \frac{1}{\pi_j} \tag{7}$$

and

$$\tilde{T}_{HT} = \sum_{j \in \mathsf{D}} \frac{y_j}{\pi_j} \tag{8}$$

where π_j would represent the probability that tree j is detected by a TLS device randomly located over the enlarged area at a distance smaller than the plot radius r. Unfortunately, the occlusion provided by the trees located near tree j renders very cumbersome the determination of this probability. Indeed, the inclusion zone of tree j, i.e. the region onto which the TLS device should be located to give rise to detection, is the circle of radius r centered at tree j (as in the conventional plot sampling scheme) minus the area shielded by the neighboring trees from which tree j cannot be detected by the device (see Figure 1). Therefore, the determination of the size of the inclusion zones is prohibitive, especially in the case of highly dense forest stands. This fact precludes the determination of the first order inclusion probability and the subsequent use of the HT estimators (7) and (8).

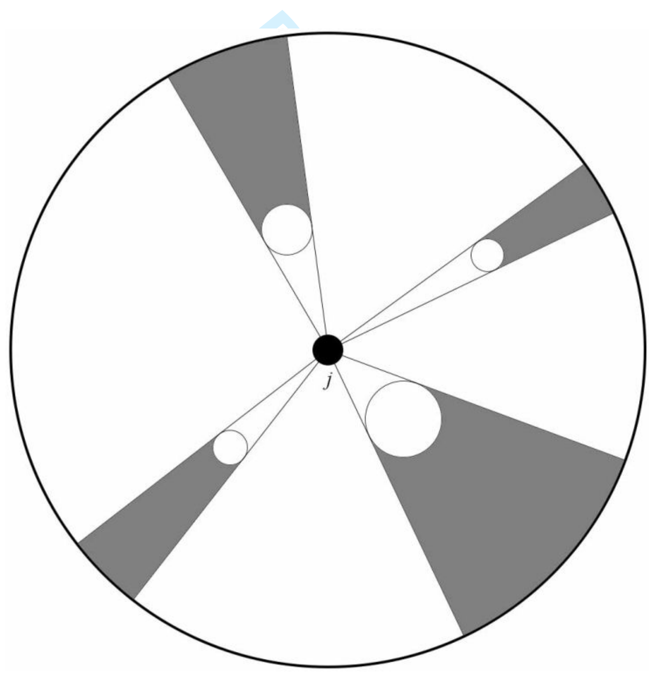

Figure 1. Example of inclusion region for tree j when four neighboring trees are within the circle of radius r centered at tree j. The inclusion region is the whole circle minus the shaded areas.

When detection probabilities cannot be quantified, e.g. when sampling populations of elusive animals, a pure-design based approach cannot be pursued. In these cases, distance sampling may be a suitable solution (Ducey and Astrup 2013, Astrup et al. 2014). When detections occur at points, as in the TLS case, distance sampling is referred to as point transect sampling (PTS). When detections occur along transects, distance sampling is referred to as line transect sampling. As pointed out by Thomas et al. (2010), distance sampling lies in the framework of the so-called composite or hybrid inference (see also Corona et al. 2014) in which the inclusion probabilities π_j are partially determined by the design, i.e., in this case, by the random location of the TLS device (that is real), and partially determined by some assumptions about the detection process.

There are several versions of distance sampling. Probably, the most simple and familiar version is the so-called Conventional Distance Sampling (CDS), i.e. the approach performed by the CDS engine in the Distance 6 software (see Thomas et al. 2010). This approach simply assumes that detection only depends on distance. Even if the approach cannot be considered entirely design-based, Buckland et al. (2004, Section 10.3.3) point out that modeling "is kept to a minimum" by this approach. On the other hand, an extension of CDS allows inclusion of covariates (e.g. tree diameter) other than distance, thus involving a more complex modelling of the detection mechanism. The approach is performed by the so-called MCDS engine in the Distance 6 software (Thomas et al. 2010) and has been pursued by Ducey and Astrup (2013) and Astrup et al. (2014). However, stated the design-based nature of this study, the CDS approach seems to be the more appropriate, avoiding complex modelling of detection.

Under CDS, the key assumption (Buckland et al. 2001, Chapter 2) can be rephrased in the TLS framework as: *a*) the detection probability of any tree only depends on its distance from the TLS device, i.e. there exists a detection function

$$g(x) = \Pr(j \in \mathsf{D} \mid X_j = x) \quad 0 \le x \le r \ , \ j \in \mathsf{U}$$

that gives the probability of detecting any tree $j \in \mathsf{U}$ when its distance X_j from the TLS device is equal to *x*; *b*) $g(0) = 1$, i.e. if the tree location coincides with the TLS location, detection is sure.

From assumptions *a*) and *b*) it follows that: *i*) the detection probability is equal for any tree, i.e. $\pi_j = \pi_0$ for any $j \in \mathsf{U}$; and *ii*) the probability density function of the areas $V_j = \pi X_j^2$ of the circles having as radii the distances of the observed trees from TLS is equal for any observed tree $j \in \mathsf{D}$ and is given by

$$h(v) = \left\{A^* \pi_0\right\}^{-1} g\left(\sqrt{v/\pi}\right), \quad 0 \le v \le a$$

in such a way that $\pi_0 = \left\{A^* h(0)\right\}^{-1}$ (e.g. Buckland et al. 2001, section 3.1). Accordingly, $h(0)$ can be estimated from the v_j s, where $v_j = \pi x_j^2$ and $\left\{x_j, j \in \mathsf{D}\right\}$ is the set of the *m* observed distances. The estimation can be performed in a parametric way, assuming an analytic form for $g(x) = g(x;\boldsymbol{\theta})$ depending on a vector $\boldsymbol{\theta}$ of unknown parameters that will be estimated from the v_j s, or in a nonparametric way, without specifying the analytical form of $g(x)$.

Once an estimate $\hat{h}(0)$ of $h(0)$ has been achieved, the HT estimators (7) and (8) reduce to the distance sampling estimators

$$\tilde{N}_{HT} = A^* \hat{h}(0) m \tag{9}$$

and

$$\tilde{T}_{HT} = A^* \hat{h}(0) \sum_{j \in \mathsf{D}} y_j \tag{10}$$

respectively. It is worth noting that distance sampling estimators differ from genuine HT estimators. They are biased with cumbersome design-based variances, say $V_{\tilde{N}}^2$ and $V_{\tilde{T}}^2$, accomplishing the uncertainty induced by the design plus the uncertainty induced by the estimation of $h(0)$.

Under URS, *R* estimates of type (9) and (10) are achieved by estimating $h(0)$ from the sole trees detected within the corresponding plot. Obviously, a more precise estimation of $h(0)$ would occur if it was based on the pooled set $D_{Rpool} = \bigcup_{i=1}^{R} D_i$ of all the distances recorded within the R plots. Therefore, denoting by $\hat{h}_{Rpool}(0)$ the estimate of $h(0)$ achieved from the pooled set of distances, any plot gives rise to the following estimate of abundance

$$\widetilde{N}_i = A^* \hat{h}_{Rpool}(0) m_i \quad , i = 1,\ldots,R$$

and basal area

$$\widetilde{T}_i = A^* \hat{h}_{Rpool}(0) \sum_{j \in D_i} y_j \ , i = 1,\ldots,R$$

where m_i is the number of detected trees within the i-th plot. Therefore, the definitive distance sampling estimators of N and T based on the R plots are given by

$$\widetilde{\widetilde{N}}_R = \frac{1}{R}\sum_{i=1}^{R} \widetilde{N}_i = A^* \hat{h}_{Rpool}(0) \overline{m}_R \tag{11}$$

and

$$\widetilde{\widetilde{T}}_R = \frac{1}{R}\sum_{i=1}^{R} \widetilde{T}_i = A^* \hat{h}_{Rpool}(0) \overline{y}_R \tag{12}$$

where $\overline{m}_R$ is the average number of trees detected within the R plots and

$$\overline{y}_R = \frac{1}{R}\sum_{i=1}^{R} \sum_{j \in D_i} y_j$$

is the average amount of basal area recorded with the R plots.

Because any distance sampling estimator necessitates a model to assume how detections occur and because any model is invariably wrong in a design-based approach, where models are not allowed, any distance sampling estimator is invariably biased, i.e. the difference between the true total and the expectation of the estimator made with respect to all the possible samples generally differs from zero.

The bias is one of the main concern regarding distance sampling (see e.g. the recent paper by Prieto Gonzales et al. 2017). Indeed, as stated by Särndal and Lundström (2005, p. 98) "if an estimator is greatly biased, it is poor consolation that its variance is low". Because it is impossible to know in advance if a biased estimator is heavily or slightly biased, mean squared error (MSE) rather than variance should be estimated in these cases.

Fortunately, URS ensures R independent samples of trees detected within the plots in such a way that a bootstrap estimator of the MSE can be attempted. From the set of the R plots, B bootstrap sets of R plots are selected with replacement and estimators (11) and (12) are computed, achieving B bootstrap estimates $\left\{\tilde{\tilde{N}}^*_{R,b}, \tilde{\tilde{B}}^*_{R,b}\,;\, b=1,\ldots,B\right\}$. Then, in accordance with Shao and Tu (1995, chapter 3), the bootstrap MSE estimators of (11) and (12) are given by

$$M\hat{S}E_{\tilde{N}} = \frac{1}{B}\sum_{b=1}^{B}(\tilde{\tilde{N}}^*_{R,b} - \tilde{\tilde{N}}_R)^2 \qquad (13)$$

and

$$M\hat{S}E_{\tilde{T}} = \frac{1}{B}\sum_{b=1}^{B}(\tilde{\tilde{T}}^*_{R,b} - \tilde{\tilde{T}}_R)^2 \qquad (14)$$

respectively. Confidence intervals can also be derived from the appropriate quantiles of the bootstrap distributions.

Alternatively, a less cumbersome solution is obtained by neglecting tree occlusion within the plots, assigning to each detected tree the "pseudo" inclusion probability a/A^*, as if detection within plots was perfect. In this case the estimation of abundance and total basal area proceeds as in the conventional plot sampling described in section 2, with the difference that m is used in (1) instead of n and the summand in (2) is extended to the set D of detected trees instead to the complete sample S. The estimators arising in this way from R plots will be denoted by $\breve{\bar{N}}_R$ and $\breve{\bar{T}}_R$ to avoid confusion with the conventional plot sampling estimators $\hat{\bar{N}}_R$ and $\hat{\bar{T}}_R$. Obviously, the uncorrected estimators are likely to be highly biased, but probably with variances smaller than those of the distance sampling estimators that are inflated by the estimation of $h(0)$. Therefore, MSE rather than variance should be estimated also in this case. That can be done, *mutatis mutandis*, by the same bootstrap procedure adopted for the distance-based estimators.

4. Simulation study

In order to compare the estimators arising from conventional plot sampling with those arising from TLS-based surveys, corrected and uncorrected for non-detection, a simulation study was performed on a set of artificial populations depicting some realistic situations. Because the man-made recording and mensuration of trees within a plot may be time very consuming, especially in dense stands, usually in conventional plot sampling, plot radius is of about 10 m. On the other hand, when trees are automatically recorded by a TLS device, the plot radius can be increased relevantly, in order to compensate for the undetected trees near the edge of the plot obscured by those near the TLS device. Therefore, the purpose of the simulation was to assess if the enlargement of the plots allowed by the use of TLS devices may compensate the non-detection of some trees, especially those far from the device, and if the application of the distance-based procedure reduced the downward bias affecting the uncorrected estimators.

4.1 Artificial populations

A quadrat of size 100 ha was taken as the study area. Within the area, low density forests of 20 000 trees (density of 200 trees per ha) and high density forests of 500 000 trees (density of 5 000 trees per ha) were generated. In order to consider several spatial patterns, the tree locations were distributed over the area: (I) completely at random; (II) in accordance with a spatially-trended process in which the coordinates of tree locations were independent random variables of type $100\times(1-u^2)$ with u uniformly distributed on $(0,1)$; (III) in accordance with a clustered process in which 10 cluster centers were randomly distributed over the area and in each equal-sized cluster, tree locations were generated from a bivariate normal distribution centered at the cluster center and having independent marginal distributions with standard deviation 80 m in the case of low density forests and 100 m in the case of high density forests.

Once a tree location was generated, the corresponding tree was determined by the circle centered at the location, with radius independently generated from a log-normal distribution with mean 15 cm and coefficient of variation of 55% in the case of low density forests and mean 4 cm and coefficient of variation of 75% in the case of high density forests. Circles/trees overlapping previously generated circles/trees were removed and the process was repeated.

Figure 2 shows the resulting six forests for the two densities (low and high) and the three spatial patterns (random, trended, clustered). Trees with radius smaller than 2.25 cm were discarded from the target populations. Therefore, abundance and basal area vary throughout populations, with about 10 trees discarded in the low density forests (0.05%) and about 150 000 in the high density forests (30%). It is worth noting that the trees with radius smaller than 2.25 cm were still considered in the simulations because they contribute to obscure the detection of those trees in the target populations.

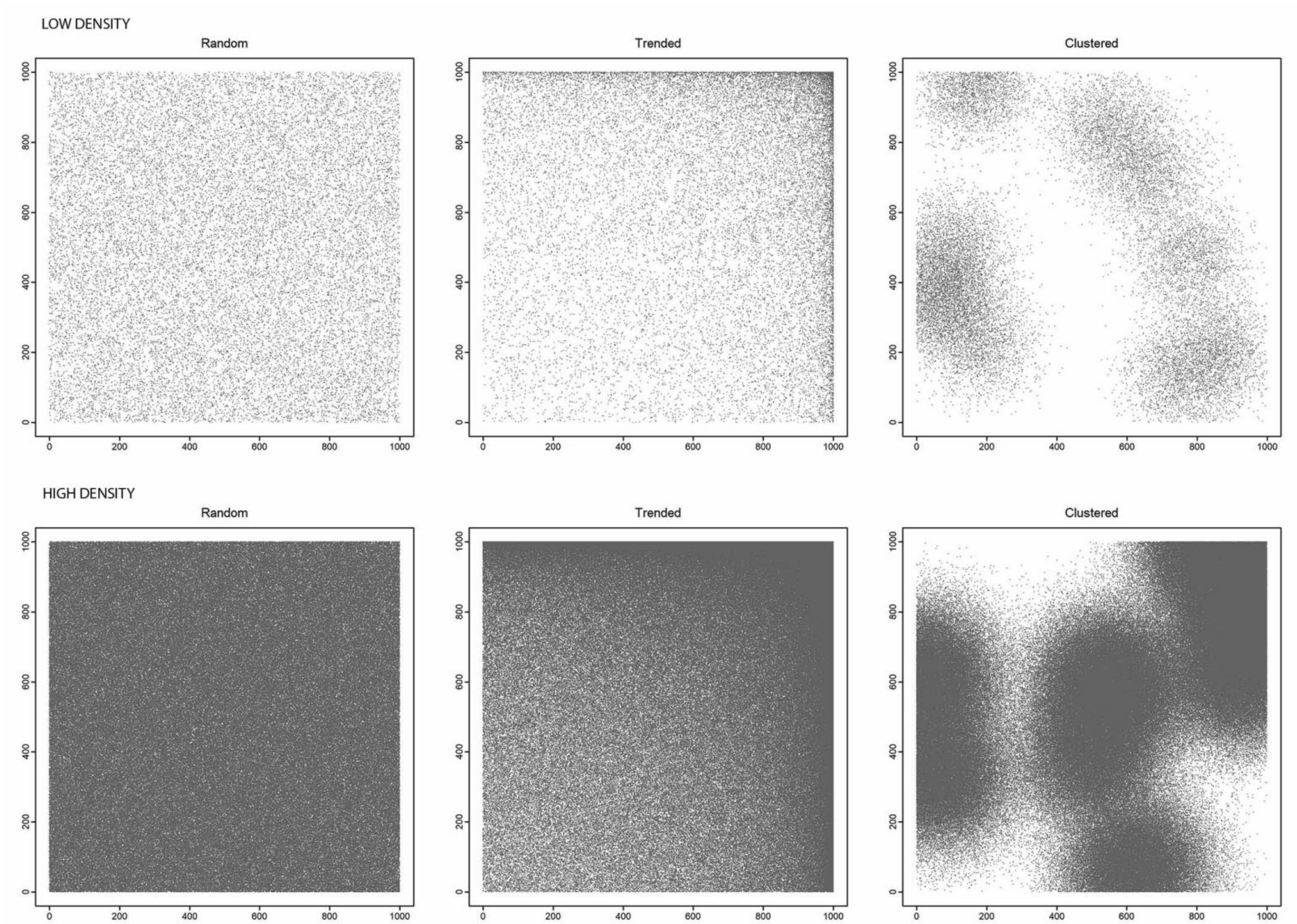

Figure 2. Graphic representations of the six forests adopted in the simulation study.

4.2 TLS-detection within circles

A first simulation step was performed to have insights on the shape of the detection function that may be supposed to rule the detection of threes within plots, that in turn determined the shape of $h(v)$. For each combination of forest densities and spatial patterns, $M = 10\ 000$ plots of radius $r = 10, 15, 20, 30$ m were replicated within the study area enlarged by a buffer equal to *r*. Within each plot $i = 1,...,M$, a naive detection mechanism was supposed in order to simplify simulation, i.e. the trees detected by the TLS device were supposed to be those subtending cones (to the center of the plot) not completely filled by the others trees nearer to the device (see Figure 3). Notwithstanding the simplicity of the supposed detection model, the computations adopted to analytically determine the detected trees were quite cumbersome and are reported in section S1 of the Supplementary material file.

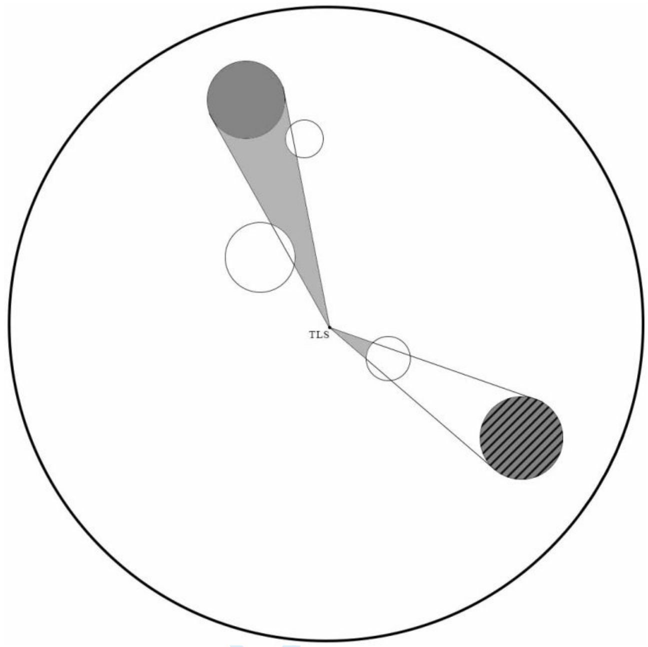

Figure 3. Example of TLS detection within a plot. The patterned tree is not detected because it subtends a cone completely occluded by a tree nearer to the TLS device; the plain gray one is detected because its cone is partially occluded by trees nearer to the TLS device.

For any simulated plot *i*, the set D_i of the m_i trees detected out of the n_i trees contained within the plot was determined and the detection frequency $p_i = m_i / n_i$ was computed, together with the m_i detected distances $\{x_j,\ j \in \mathsf{D}_i\}$, that, in turn, determined the areas of the corresponding circles $\{v_j,\ j \in \mathsf{D}_i\}$, with $v_j = \pi x_j^2$. The area range $(0, a)$ was divided into 20 intervals of equal width and the frequencies f_{ih} $(h = 1, \ldots, 20)$ of the intervals were computed. Then, the average number and the average fraction of trees detected within plots of radius *r* was empirically evaluated by means of

$$E(m) = \frac{1}{M}\sum_{i=1}^{M} m_i \quad , \quad E(p) = \frac{1}{M}\sum_{i=1}^{M} p_i$$

respectively, while the probability density function $h(v)$ was empirically approximated by the average interval frequencies

$$f_l = \frac{1}{M} \sum_{i=1}^{M} f_{ih} \text{ , } l = 1,\dots,20$$

For each combination of forest densities, spatial patterns, and plot radii, Table 1 reports the average fraction of detected trees, while Figure 4 reports the graphs of the average interval frequencies approximating $h(v)$.

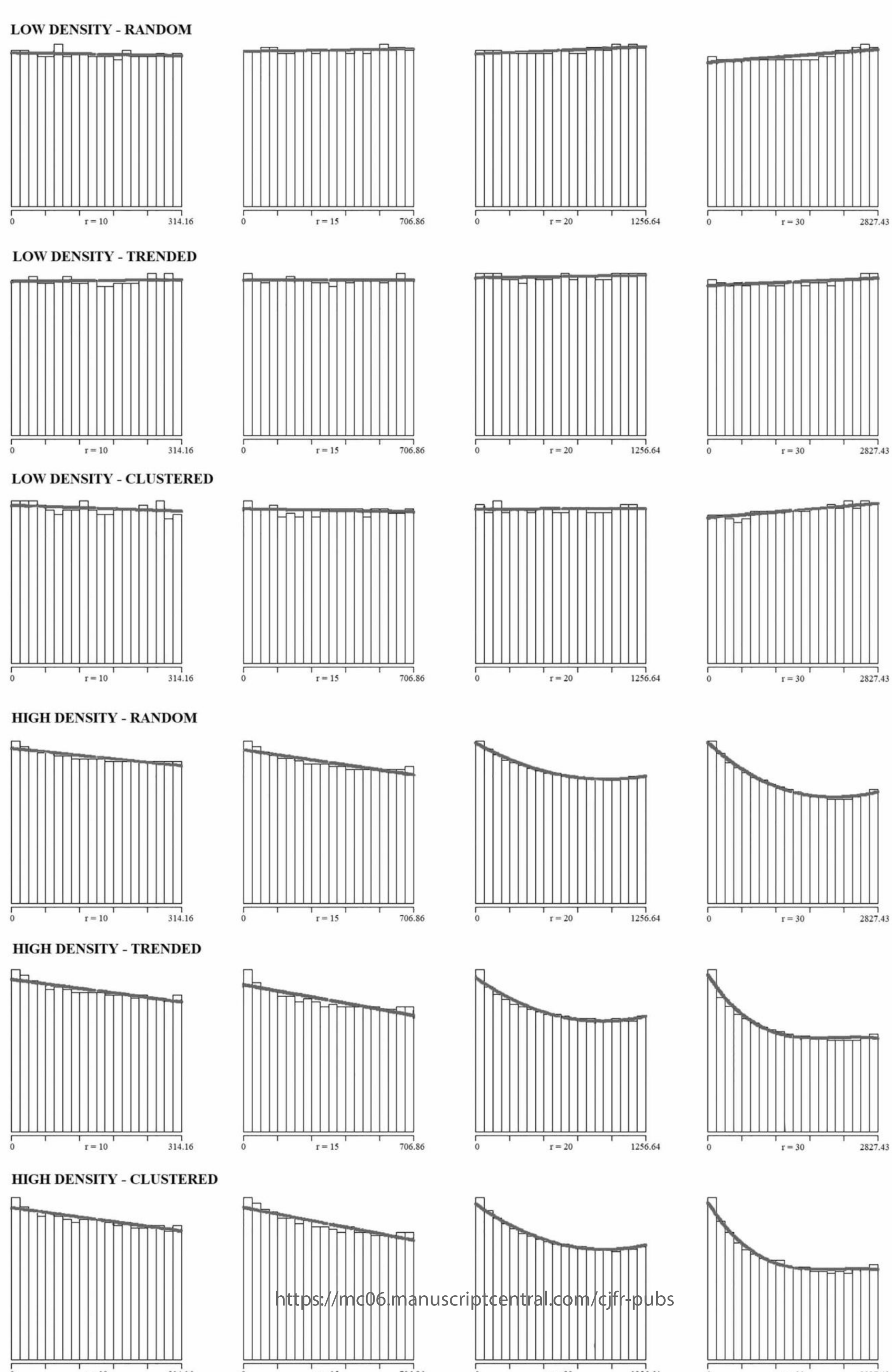
LOW DENSITY - RANDOM
LOW DENSITY - TRENDED
LOW DENSITY - CLUSTERED
HIGH DENSITY - RANDOM
HIGH DENSITY - TRENDED
HIGH DENSITY - CLUSTERED
0
r = 10
314.16
r = 15
706.86
r = 20
1256.64
r = 30
2827.43

Figure 4. Monte Carlo probability density functions of the areas of detected trees within plots of radius $r = 10,15,20,30$ m in low and high density forests with random, trended and clustered spatial patterns.

The interval frequencies in low density forests with random and trended spatial pattern resulted very similar to the uniform distribution in $(0, a)$. That revealed a quite perfect detection within plots with fraction of detected trees near to one (see Section S2 of the Supplementary Material File). In the case of the low density forest with clustered pattern, detection was less good with fractions of detected trees of about 70-80%. On the other hand, in high density forests the frequencies of first intervals decreased with areas/distances and the decrease was more marked as the plot radius increased, especially under trended and clustered patterns.

Based on this preliminary analysis of the detection frequencies, we opted for a nonparametric fitting performed by means of an orthogonal series approximation (Barabesi and Fattorini 1996), in order to provide a method likely to perform well under a wide range of situations. More precisely, $h(v)$ was approximated by the first *K* terms of a Legendre polynomial series, i.e.

$$h(v) \approx \frac{1}{a}\left\{1 + \sum_{k=1}^{K} \theta_k \phi_k\left(\frac{v}{a}\right)\right\} \ , \ 0 \le v \le a$$

where

$$\theta_k = E\left\{\phi_k\left(\frac{V}{a}\right)\right\} \approx \sum_{l=1}^{20} \phi_k\left(\frac{\bar{v}_l}{a}\right) f_l \ , \ k = 1,\ldots,K$$

$\bar{v}_l$ was the central value of the *l*-th interval, ϕ_k was the *k*-th shifted Legendre polynomial while *K* was the first integer such that $\theta_{KK} > \{1 + E(m)\}\theta_K^2$ and

$$\theta_k^2 = E\left\{\phi_k^2\left(\frac{V}{a}\right)\right\} \approx \sum_{L=1}^{20} \phi_k^2\left(\frac{\bar{v}_l}{a}\right) f_l \ , \ k = 1,\ldots,K$$

(see section S3 of the Supplementary Material file for further details). The fitted distributions are reported in Figure 4 and are graphed by bold lines. The closeness of the interval frequencies with the bold lines evidenced the effectiveness of the orthogonal polynomial fitting. In most cases, $K=1$ term was sufficient to provide the best fitting, with the exceptions of plots of radius $r=20,30$ for the high density forests. In these cases, $K=2$ terms were necessary with 20 m-plot radius and $K=3$ terms were necessary for 30 m-plot radius.

These results suggested the use of Legendre polynomial series to achieve the $h(0)$ estimates in equations (11) and (12) as suitable alternative to more widely applied methodologies, such as the flexible, semi-parametric procedure performed by the CDS engine in the Distance 6 software, with a parametric key function chosen by the Akaike model selection criterion among uniform, half-normal, hazard rate and exponential detection functions, paired with zero or more series adjustment terms chosen among cosine, Hermite or simple polynomial series (Thomas et al. 2010). As opposite to CDS engine that adopts the maximum likelihood criterion to estimate the parameters of the key function and the coefficients of the series terms, giving rise to convergence problems in some cases (e.g. Ducey and Astrup 2013), the use of Legendre polynomial series only involves very simple moment estimates of the first polynomial coefficients. Therefore, it can be easily implemented in simulation saving much of the computational time that would be involved to automatically insert the use of Distance 6 in the simulation codes.

4.3 Sampling and estimation

To simulate conventional and TLS-based plot sampling, $M=10\ 000$ simulation runs were performed for each combination of the six artificial forests, plot radius $r=10,15,20,30$ m, and number of replicated plots $R=9,16,25,36$ for a total of 96 cases. At each simulation run, R plots were randomly and independently located within the quadrat of size 100 ha enlarged by a buffer of width equal to the plot radius (URS). In the case of plots of 10 m radius, the trees located within the plots were completely enumerated and estimators (5) and (6) were computed in accordance with the conventional plot sampling protocol. Moreover, for any of the four plot radii, the set of detected trees was determined within each plot by the steps delineated in section S2 of the Supplementary Material file and distance sampling estimators (11) and (12) were computed together with the uncorrected estimators performed as if detection was perfect. As already stated, Lagrange polynomials were adopted to estimate $h(0)$ in equations (11) and (12), i.e.

$$\hat{h}_{Rpool}(0) = \frac{1}{a}\left\{1 + \sum_{k=1}^{K} \hat{\theta}_k \phi_k(0)\right\}$$

where

$$\hat{\theta}_k = \frac{1}{m_{Rpool}} \sum_{j \in D_{Rpool}} \phi_k\left(\frac{v_j}{a}\right),\ k = 1, \ldots, K$$

m_{Rpool} was the total number of detected trees in the pooled sample, while K was the first integer such that $\hat{\theta}_{KK} > (1 + m_{Rpool})\hat{\theta}_K^2$ where

$$\hat{\theta}_{kk} = \frac{1}{m_{Rpool}} \sum_{j \in D_{Rpool}} \phi_k^2\left(\frac{v_j}{a}\right),\ k = 1, \ldots, K$$

(see section S3 of the Supplementary Material file for details).

4.4 Performance indicators

At the end of the simulation, ten thousand estimates of type (11) and (12) together with the corresponding uncorrected estimates were achieved for each of the 96 cases together, while ten thousand estimates of type (5) and (6) were achieved for each combination of forest types and R values when $r = 10$. These collections determined the Monte Carlo distributions from which the performance of the estimators were empirically evaluated by means of the relative bias (RB=expectation minus true parameter value divided by the true parameter value) and the relative root mean squared error (RRMSE=square root of the mean squared error divided by the true parameter value). Because the conventional plot sampling gave rise to unbiased estimators, only the RRMSEs were considered in these cases.

It is worth noting that the design-based nature of the study is apparent from the simulation structure in which populations were kept fixed, thus univocally determining detection, which was established from the position of a tree with respect to its neighboring trees. In this way, uncertainty and estimator properties only stemmed from sampling, i.e. from the locations of sample plots that were newly generated at each run.

5. Results and discussion

Tables 2, 3 and 4 report the percent values of RBs and RRMSEs for the three spatial patterns and any combination of density, number of plots (R) and plot radius (r).

In low density forests, the use of uncorrected TLS-based estimators generated a downward bias that rapidly increased with plot radius and was not reduced by increasing the number of replicated plots. For forests with a random spatial pattern, the bias of abundance estimators increased from -1% with plot of 10 m radius to -7% with plots of 30 m radius. In the case of trended and clustered spatial patterns, the bias increased from -6% to -14% and from -5% to 16%. For basal area estimation, the results were very similar, even if bias levels were slightly smaller than that achieved for abundance. On the other hand, the use of corrected TLS-based estimators was effective in reducing the level of bias. In this case, bias increased from $\pm 1\%$ to -2%, from -1% to -4% and from -1% to -5% for random, trended and clustered patterns, respectively. Even better results were achieved for basal area estimation, where bias was almost completely eliminated. Unfortunately, the estimation of the detection function involved in distance sampling generated a further uncertainty that deteriorated the precision of the corrected estimator, providing RRMSEs that were invariably greater than those achieved with the uncorrected counterparts. Therefore, in low density forests, when the bias of the uncorrected estimators is small, there is little reason to adopt corrected estimators. Moreover, the TLS-based uncorrected estimators with enlarged plot radius compared favorably with respect to the (unbiased) plot sampling estimator with 10 m radius plots. RRMSE reductions were sometimes relevant, being in some cases greater than ten percentage points, and were more marked for basal area estimation. However, care must be taken in enlarging the plot radius too much in order to avoid unsuitable levels of bias over 5%. Practically speaking, in the case of low density forests, the use of uncorrected estimator with plot of radius two time greater than that adopted in the conventional plot sampling is the best option. Finally, a less than obvious result must be discussed. When conventional and TLS-based sampling were both performed at 10 m radius plots, the conventional plot sampling estimators performed invariably worse than the uncorrected TLS-based estimators, even if all the trees were detected in the conventional plot sampling while some were lost in the TLS-based uncorrected case. It should be considered that when trees in the plot were few, few of them were undetected and both the estimators provided similar, low estimates of abundance and basal area. On the other hand, when trees in the plot were many, many of them, especially those at the edge of the plot, were undetected. In these cases, the conventional plot sampling provided high estimates of abundance and basal area, while the estimates provided by the uncorrected estimator were much smaller and nearer to those achieved when trees are few. This reduced the variance of the uncorrected estimator with respect to that provided by the conventional plot sampling. This reduction in variability was not compensated by the presence of bias that, in the case of low density forest, was not excessive, therefore providing values of RRMSE smaller than those achieved by conventional plot sampling.

Regarding high density forests, the density of 500 000 trees per ha was very high, corresponding e.g. to young planted stand, where trees would often be so small that they would not be measured. Indeed, about the 30% of the generated trees were discarded from the target population. In these forests, the use of uncorrected TLS-based estimators generated inacceptable amounts of downward bias that rapidly increased with plot radius and did not reduce with the number of replicated plots. For abundance estimation, the bias increased from -13% with plot of 10 m radius to -37% with plots of 30 m radius for random spatial patterns, from -27% to -51% for trended patterns and from -27% to -59% for clustered patterns. The results for basal area estimation were similar even if the bias levels were slightly smaller. Also in this case, the use of corrected TLS-based estimators was effective in reducing the level of bias that in this case ranged from -3% to -6%, from -7% to -16% and from -5% to -12% under random, trended and clustered pattern, respectively. Better results were again achieved for basal area estimation, where bias ranged from 2% to -8% in the worst situations. Obviously, the presence of massive downward bias deteriorated the precision of the uncorrected TLS-based estimators, that resulted invariably worse that the corrected counterparts. Especially for plots of 20/30 m radius, the RRMSEs were sometimes two-three times greater than those achieved after the distance sampling corrections. However, also in this case, the estimation of the detection function involved in distance sampling generated a further uncertainty that makes the performance of corrected estimators comparable, but not generally better, than those arising from the use of conventional plot sampling with plots of smaller radius. Practically speaking, in the case of high density forests, there is no reason to adopt TLS based estimation.

The whole simulation study was repeated locating plots in accordance with the sampling scheme usually referred to as tessellation stratified sampling (TSS). To this purpose, the study area enlarged by a buffer of width equal to the plot radius was partitioned into $R = 9, 16, 25, 36$ quadrats of equal size and, at each simulation run, a plot was randomly located within each quadrat. Then, estimation was performed repeating the same computations adopted in the URS case. The simulation results are reported in section S4 of the Supplementary Material file. The superiority of TSS with respect to URS was proven in several studies (Barabesi and Franceschi 2011, Barabesi et al. 2012, Barabesi and Fattorini 2013). Also in this case, TSS provided improvement in precision for both conventional and TLS-based plot sampling, but in relative terms the results were similar to the URS case: TLS-based plot sampling provided improvement with respect to conventional plot sampling only in the cases of low density forests, when no distance sampling correction was performed.

It is worth noting that the estimation of the sampling variances was neglected in this study. While we recognized the importance of estimating variance, we avoided this step owing to the time consuming computations involved in performing the bootstrap estimation of variance in the case of TLS-based plot sampling (equations 13 and 14). These computations would have elongate the simulation time much beyond the two months that were necessary to achieve the present results.

6. Conclusions

Forest inventories are rapidly evolving as novel approaches arise and new techniques and tools become available. However, implementation within operative processes should be evidence-based, i.e. based on objective, reliable assessment (Corona 2016, 2018). The results of the simulation study demonstrate that the use of single-scan TLS for the automatic detection of trees provides gains in estimation precision – with respect to the conventional plot sampling performed within plot of smaller size - only in the case of low density forests when no distance sampling correction is performed. Care must be taken in enlarging the plot radius too much in order to avoid unsuitable increases of bias over 5%. On the other hand, there is no convenience in using TLS-based estimation in high density forests.

It should be noticed that these conclusions must be viewed only as a first attempt for evaluating TLS-based sampling. They are indeed obtained on the basis of an artificial simulation study in which the process of tree detection has been necessarily simplified. First, it is assumed that a tree is detected even if just a small sliver of it is visible from the scan point: that represents an optimistic assumption given the current state of algorithms for this task. Second, it is assumed that only tree stems cause occlusion, that may be realistic only in stands with a relatively open understory and where the trees have no crowns in the horizontal plane where searching is performed. These simplifications constitute a significant limitation of the simulation study that may give rise to optimistic evaluation of TLS detection and hence of the performance of TLS-based estimators. However, on this issue it should be noticed that very similar simplifications are also adopted in two recent simulation studies on single-scan TLS-based sampling (Olofsson and Olsson 2018, Kuronen et al. 2018) owing to the difficulties in simulating more realistic situations.

In order to effectively exploit the advantage of TLS under permanent forest inventory applications, due to the possibility of a suitable monumentalization of the sample plots, advancement of multi-scan or mobile TLS approaches must be sought, in terms of both measurement and processing cost-effectiveness. Distinctively, mobile TLS coupled by Simultaneous Localization and Mapping methods currently seems to be the most promising option (e.g. Ryding et al. 2015), whose potential needs to be properly investigated under a large-scale forest inventory framework.

References

Astrup, R., Ducey, M.J., Granhus, A., Ritter, T., and von Lüpke, N. 2014. Approaches for estimating stand-level volume using terrestrial laser scanning in a single-scan mode. Canadian Journal of Forest Research **44**(6): 666-676. doi:10.1139/cjfr-2013-0535.

Barabesi, L., and Fattorini, L. 1996. A nonparametric orthogonal series estimator of plant density with truncated samples. Italian Journal of Applied Statistics **8**(3): 563-576.

Barabesi, L., and Fattorini, L. 2013. Random versus stratified location of transects or points in distance sampling: theoretical results and practical considerations. Environmental and Ecological Statistics **20**(2): 215–236. doi:10.1007/s10651-012-0216-1.

Barabesi, L., and Franceschi, S. 2011. Sampling properties of spatial total estimators under tessellation stratified designs. Environmetrics **22**(3): 271–278. doi:10.1002/env.1046.

Barabesi, L., Franceschi, S., and Marcheselli, M. 2012. Properties of design-based estimation under stratified spatial sampling. Annals of Applied Statistics **6**(1): 210–228. doi:10.1214/11-AOAS509.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., and Thomas, L. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford.

Corona, P., Fattorini, L., Franceschi, S., Scrinzi, G., and Torresan, C. 2014. Estimation of standing wood volume in forest compartments by exploiting airborne laser scanning information: model-based, design-based, and hybrid perspectives. Canadian Journal of Forest Research **44**(11): 1303-1311. doi:10.1139/cjfr-2014-0203.

Corona, P. 2016. Consolidating new paradigms in large-scale monitoring and assessment of forest ecosystems. Environmental Research **144**(pt B): 8-14. doi:10.1016/j.envres.2015.10.017.

Corona, P. 2018. Communicating facts, findings and thinking to support evidence-based strategies and decisions. Annals of Silvicultural Research **42**: 1-2. doi: 10.12899/ASR-1617.

Ducey, M.J., and Astrup, R. 2013. Adjusting for nondetection in forest inventories derived from terrestrial laser scanning. Canadian Journal of Remote Sensing **39**(5): 410-425. doi: 10.5589/m13-048.

Fardusi, M.J., Chianucci, F., and Barbati, A. 2017. Concept to Practices of Geospatial Information Tools to Assist Forest Management and Planning under Precision Forestry Framework: a review. Annals of Silvicultural Research **41**(1): 3-14. doi:10.12899/asr-1354.

Gregoire, T.G., and Valentine, H.T. 2008. Sampling strategies for natural resources and the environment. Chapman & Hall, Boca Raton (FL), USA.

Kankare, V., Liang, X., Vastaranta, M., Yu, X., Holopainen, M., and Hyyppä, J. 2015. Diameter distribution estimation with laser scanning based multisource single tree inventory. ISPRS Journal of Photogrammetry and Remote Sensing **108**: 161-171. doi:10.1016/j.isprsjprs.2015.07.007.

Kuronen, M., Henttonen, H.M., and Myllymäki, M. 2018. Correcting for nondetection in estimating forest characteristics from single-scan terrestrial laser measurements. Canadian Journal of Forest Research, online first. doi: 10.1139/cjfr-2018-0072.

Liang, X., Litkey, P., Hyyppä, J., Kaartinen, H., Kukko, A., and Holopainen, M. 2011. Automatic plot-wise tree location mapping using single-scan terrestrial laser scanning. The photogrammetric Journal of Finland **22**(2): 37-48.

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M., and Vastaranta, M. 2016. Terrestrial laser scanning in forest inventories. ISPRS Journal of Photogrammetry and Remote Sensing **115**:63-77. doi:10.1016/j.isprsjprs.2016.01.006.

Moskal, L.M., and Zheng, G. 2012. Retrieving Forest Inventory Variables with Terrestrial Laser Scanning (TLS) in Urban Heterogeneous Forest. Remote Sensing **4**(1): 1-20. doi:10.3390/rs4010001.

Olofsson, K., and Olsson, H. 2018. Estimating tree stem density and diameter distribution in single-scan terrestrial laser measurements of field plots: a simulation study. Scandinavian Journal of Forest Research **33**(4): 365-377.doi: 10.1080/02827581.2017.1368698.

Prieto Gonzalez, R., Thomas, L., and Marques, T.A. 2017. Estimation bias under model selection for distance sampling detection functions. Environmental and Ecological Statistics **24**(3): 399-414. doi:10.1007/s10651-017-0376-0.

Ryding, J., Williams, E., Smith, M.J., and Eichhorn, M.P. 2015. Assessing handheld mobile laser scanners for forest surveys. Remote Sensing **7**(1): 1095-1111. doi:10.3390/rs70101095.

Saarela, S., Breidenbach, J., Raumonen, P., Grafström, A., Ståhl, G., Ducey, M.J., and Astrup, R. 2017. Kriging prediction of stand level forest information using mobile laser scanning data adjusted for non-detection. Canadian Journal of Forest Research, 47(9):1257-1265. doi: 10.1139/cjfr-2017-0019.

Särndal, C. E., and Lundström, S. 2005. Estimation in Survey with Nonresponse. Wiley, New York.

Shao, J., and Tu, D. 1995. The Jackknife and Bootstrap. Springer-Verlag, New York.

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A., and Burnham, K.P. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology **47**(1): 5-14. doi:10.1111/j.1365-2664.2009.01737.x.

Table 1. Monte Carlo values of the average fraction of detected trees in the six forests.

Plot radius	Low density			High density		
	Random	Trended	Clustered	Random	Trended	Clustered
10	97%	92%	68%	61%	62%	57%
15	96%	96%	73%	56%	58%	55%
20	95%	95%	77%	53%	54%	53%
30	93%	93%	80%	46%	49%	48%

Table 2. Percent values of RB and RRMSE of TLS-based corrected (TLS C) and uncorrected (TLS U) estimators compared with those of the conventional plot sampling estimator (PLOT) in artificial forests with trees located in accordance with a random spatial pattern.

No. of plots	Plot radius	Estimator	Low density				High density			
			Abundance		Basal area		Abundance		Basal area	
			RB	RRMSE	RB	RRMSE	RB	RRMSE	RB	RRMSE
9	10	TLS C	1	36	1	41	-3	13	1	15
		TLS U	-1	15	0	24	-13	14	-10	13
		PLOT		15		24		7		9
	15	TLS C	-1	25	0	28	-4	12	1	13
		TLS U	-3	11	-2	16	-20	21	-15	17
	20	TLS C	-2	20	0	23	-5	12	2	13
		TLS U	-4	11	-3	14	-26	27	-21	22
	30	TLS C	-2	17	0	19	-6	14	2	14
		TLS U	-6	12	-5	13	-37	38	-31	32
16	10	TLS C	-1	26	0	30	-3	10	1	11
		TLS U	-2	11	-1	17	-13	14	-10	12
		PLOT		11		18		5		7
	15	TLS C	-1	19	0	21	-4	10	2	10
		TLS U	-3	9	-2	13	-20	20	-15	16
	20	TLS C	-1	16	0	18	-5	10	2	10
		TLS U	-4	9	-3	11	-26	26	-21	21
	30	TLS C	-2	13	0	14	-6	11	3	11
		TLS U	-7	10	-5	10	-37	37	-31	32
25	10	TLS C	0	22	0	25	-3	8	1	9
		TLS U	-2	9	-2	14	-13	14	-10	11
		PLOT		9		14		4		5
	15	TLS C	-1	15	0	17	-4	8	2	8
		TLS U	-3	7	-2	10	-20	20	-15	16
	20	TLS C	-1	13	0	14	-4	9	2	9
		TLS U	-4	7	-3	9	-26	26	-21	21
	30	TLS C	-2	11	0	12	-6	10	3	9
		TLS U	-7	9	-5	9	-37	37	-31	32
36	10	TLS C	0	17	0	20	-3	7	1	8
		TLS U	-2	7	-1	12	-13	13	-10	11
		PLOT		7		12		3		4
	15	TLS C	-1	13	0	14	-4	7	2	7
		TLS U	-3	6	-2	9	-20	20	-15	15
	20	TLS C	-1	10	0	12	-4	8	3	7
		TLS U	-4	7	-3	7	-26	26	-20	21
	30	TLS C	-2	9	0	10	-5	8	3	8
		TLS U	-6	8	-5	8	-37	37	-31	31

Table 3. Percent values of RB and RRMSE of TLS-based corrected (TLS C) and uncorrected (TLS U) estimators compared with those of the conventional plot sampling estimator (PLOT) in artificial forests with trees located in accordance with a trended spatial pattern.

No. of plots	Plot radius	Estimator	Low density				High density			
			Abundance		Basal area		Abundance		Basal area	
			RB	RRMSE	RB	RRMSE	RB	RRMSE	RB	RRMSE
9	10	TLS C	-1	65	1	69	-7	42	-1	44
		TLS U	-6	38	-4	42	-27	32	-23	29
		PLOT		46		47		43		39
	15	TLS C	-3	54	-1	56	-9	37	-2	39
		TLS U	-9	32	-7	34	-35	37	-30	33
	20	TLS C	-2	52	0	54	-12	35	-4	36
		TLS U	-11	30	-8	31	-41	42	-36	38
	30	TLS C	-3	48	0	50	-16	33	-8	33
		TLS U	-14	28	-11	29	-51	51	-46	47
16	10	TLS C	-1	49	1	52	-7	31	-1	33
		TLS U	-6	28	-4	31	-27	30	-23	27
		PLOT		34		35		32		29
	15	TLS C	-3	41	0	44	-9	29	-3	30
		TLS U	-9	25	-7	26	-35	36	-30	32
	20	TLS C	-3	39	0	40	-12	27	-4	28
		TLS U	-11	24	-8	24	-41	42	-36	37
	30	TLS C	-4	36	0	38	-16	27	-8	25
		TLS U	-14	23	-11	23	-51	51	-46	46
25	10	TLS C	-2	38	0	40	-7	25	-1	27
		TLS U	-6	23	-5	25	-27	29	-23	25
		PLOT		28		29		25		23
	15	TLS C	-2	33	0	35	-9	24	-3	24
		TLS U	-9	21	-7	22	-35	36	-30	31
	20	TLS C	-3	31	0	33	-11	23	-4	22
		TLS U	-11	20	-8	20	-41	42	-36	37
	30	TLS C	-4	29	-1	30	-15	23	-7	21
		TLS U	-14	20	-11	19	-51	51	-46	46
36	10	TLS C	-1	32	1	34	-7	22	-1	22
		TLS U	-6	20	-4	21	-27	28	-23	25
		PLOT		24		24		21		19
	15	TLS C	-2	28	0	29	-9	21	-2	20
		TLS U	-9	18	-6	18	-35	36	-30	31
	20	TLS C	-3	26	0	27	-11	20	-4	19
		TLS U	-11	18	-8	17	-41	42	-36	37
	30	TLS C	-4	24	-1	25	-15	21	-7	18
		TLS U	-14	19	-11	17	-51	51	-46	46

Table 4. Percent values of RB and RRMSE of TLS-based corrected (TLS C) and uncorrected (TLS U) estimators compared with those of the conventional plot sampling estimator (PLOT) in artificial forests with trees located in accordance with a clustered spatial pattern.

No. of plots	Plot radius	Estimator	Low density				High density			
			Abundance		Basal area		Abundance		Basal area	
			RB	RRMSE	RB	RRMSE	RB	RRMSE	RB	RRMSE
	10	TLS C	-2	56	0	61	-6	38	1	41
		TLS U	-6	41	-5	46	-27	36	-23	34
		PLOT		44		48		38		38
	15	TLS C	-2	49	0	52	-7	38	1	41
9		TLS U	-8	39	-6	42	-38	42	-33	39
	20	TLS C	-4	46	-1	48	-8	37	-1	40
		TLS U	-11	37	-9	39	-46	49	-42	45
	30	TLS C	-3	45	0	48	-12	37	-4	38
		TLS U	-15	37	-12	37	-59	60	-55	56
	10	TLS C	-1	42	0	45	-5	29	1	32
		TLS U	-5	31	-4	34	-27	32	-23	30
		PLOT		33		36		29		29
	15	TLS C	-3	36	-1	39	-6	29	1	31
16		TLS U	-8	30	-6	31	-38	40	-33	36
	20	TLS C	-3	35	0	37	-8	29	0	30
		TLS U	-11	29	-8	30	-46	48	-42	44
	30	TLS C	-5	34	-1	36	-12	29	-4	29
		TLS U	-16	30	-13	29	-59	59	-55	56
	10	TLS C	-1	33	0	36	-5	24	1	25
		TLS U	-5	25	-4	28	-27	31	-23	27
		PLOT		27		29		23		23
	15	TLS C	-2	29	0	31	-6	23	1	25
25		TLS U	-8	24	-6	25	-38	40	-33	35
	20	TLS C	-3	28	0	29	-8	23	0	24
		TLS U	-10	24	-8	24	-47	47	-42	43
	30	TLS C	-4	28	0	30	-11	24	-4	23
		TLS U	-16	25	-12	25	-59	59	-55	56
	10	TLS C	-1	28	0	30	-5	20	1	21
		TLS U	-5	21	-4	23	-28	30	-23	26
36		PLOT		22		24		19		19
	15	TLS C	-3	25	-1	26	-6	20	1	21
		TLS U	-8	21	-6	21	-38	39	-33	35

20	TLS C	-3	24	0	25	-8	20	0	20
	TLS U	-11	21	-8	21	-47	47	-42	43
30	TLS C	-4	23	0	25	-11	21	-4	19
	TLS U	-16	23	-12	22	-59	59	-55	56

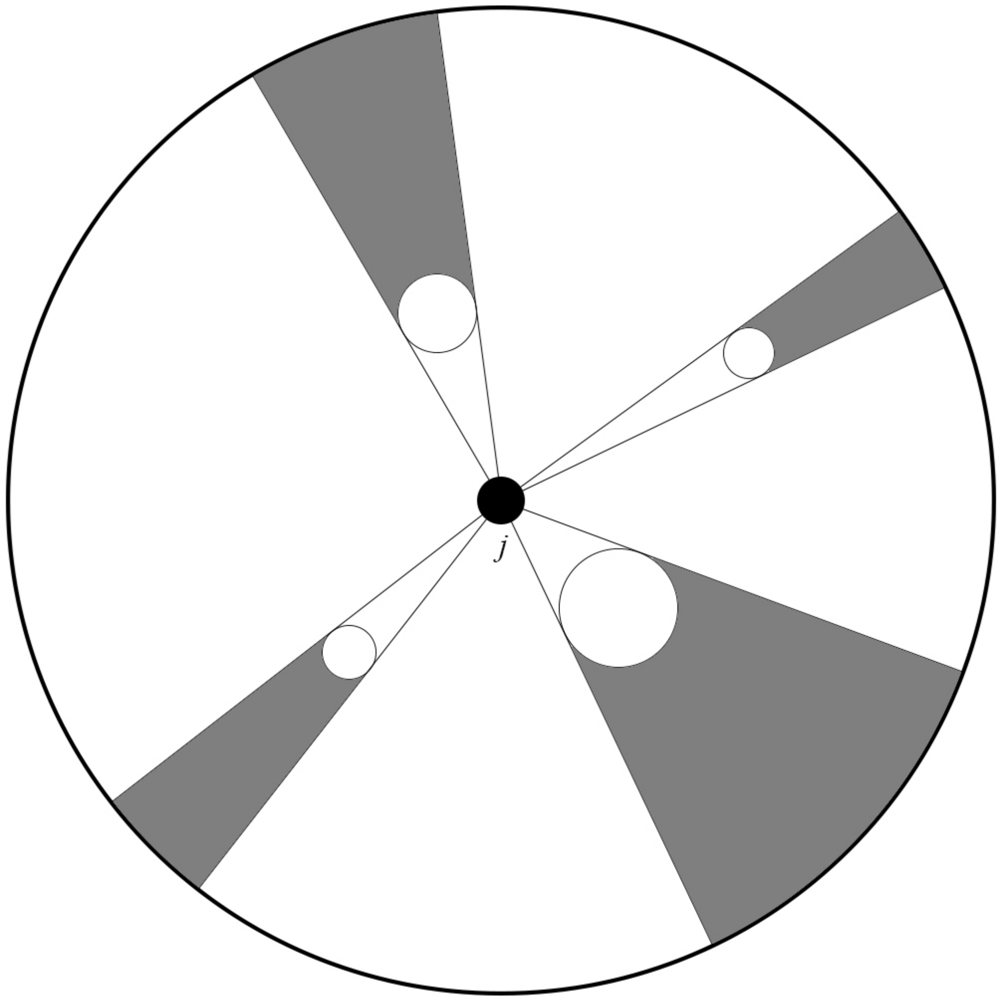

Figure 1. Example of inclusion region for tree j when four neighboring trees are within the circle of radius r centered at tree j. The inclusion region is the whole circle minus the shaded areas.

85x85mm (300 x 300 DPI)

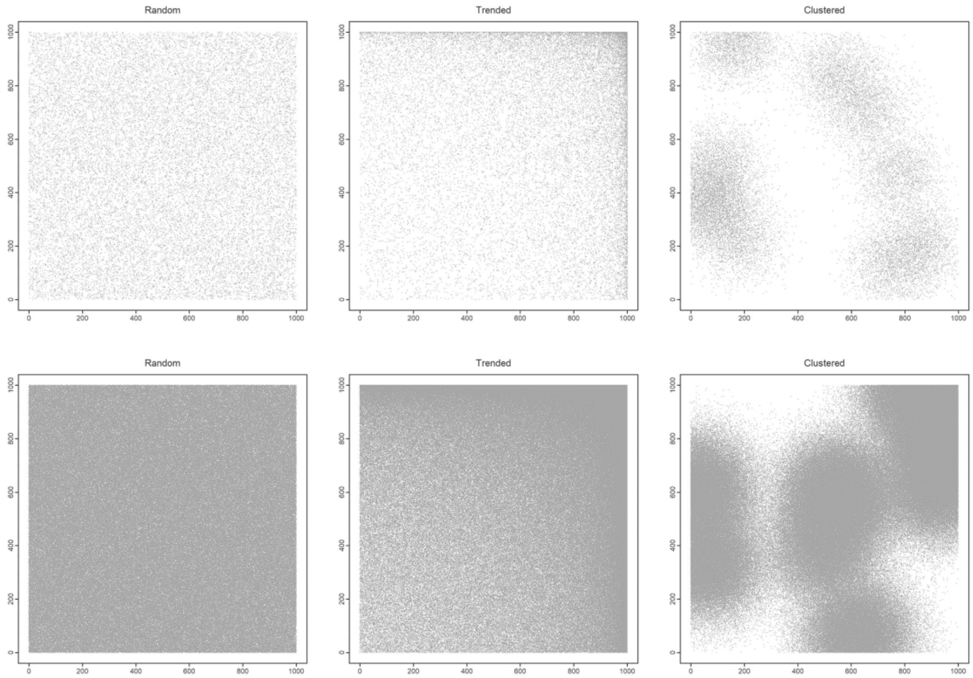

Figure 2. Graphic representations of the six forests adopted in the simulation study.

181x128mm (300 x 300 DPI)

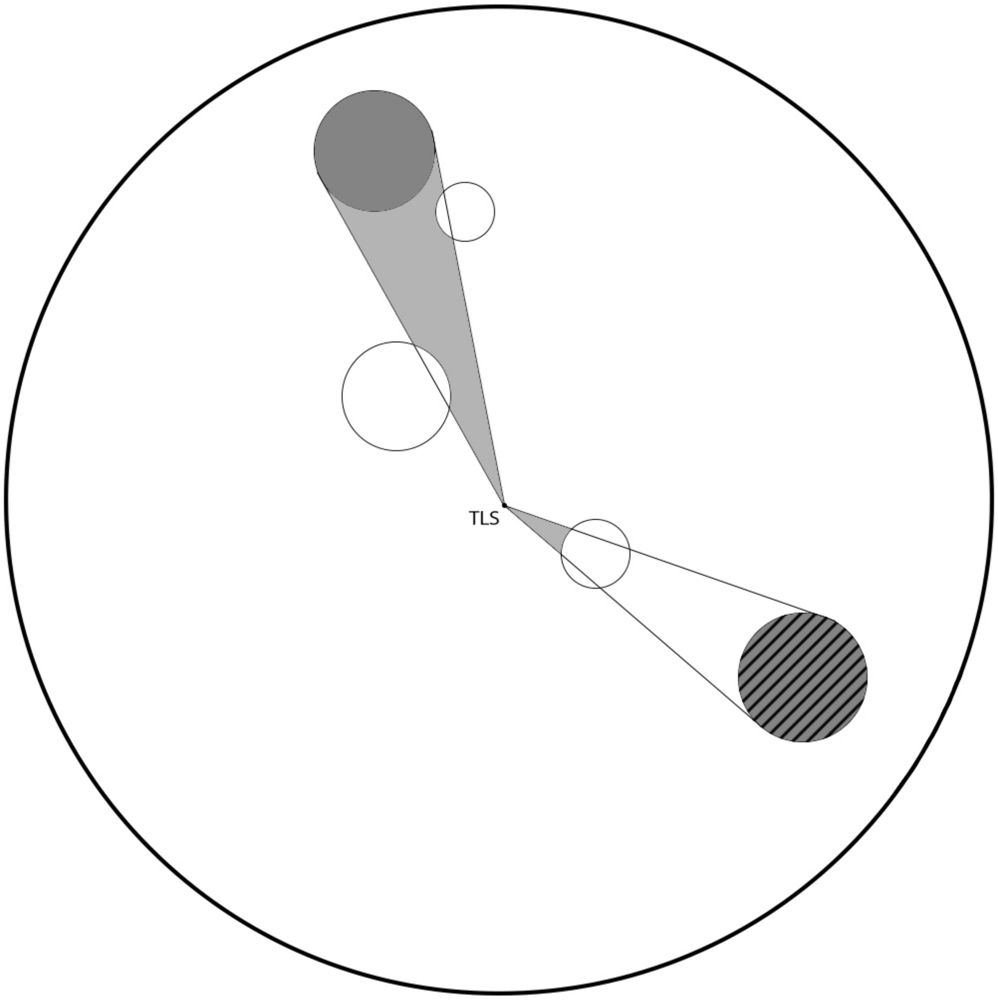

Figure 3. Example of TLS detection within a plot. The patterned tree is not detected because it subtends a cone completely occluded by a tree nearer to the TLS device; the plain gray one is detected because its cone is partially occluded by trees nearer to the TLS device.

85x85mm (300 x 300 DPI)

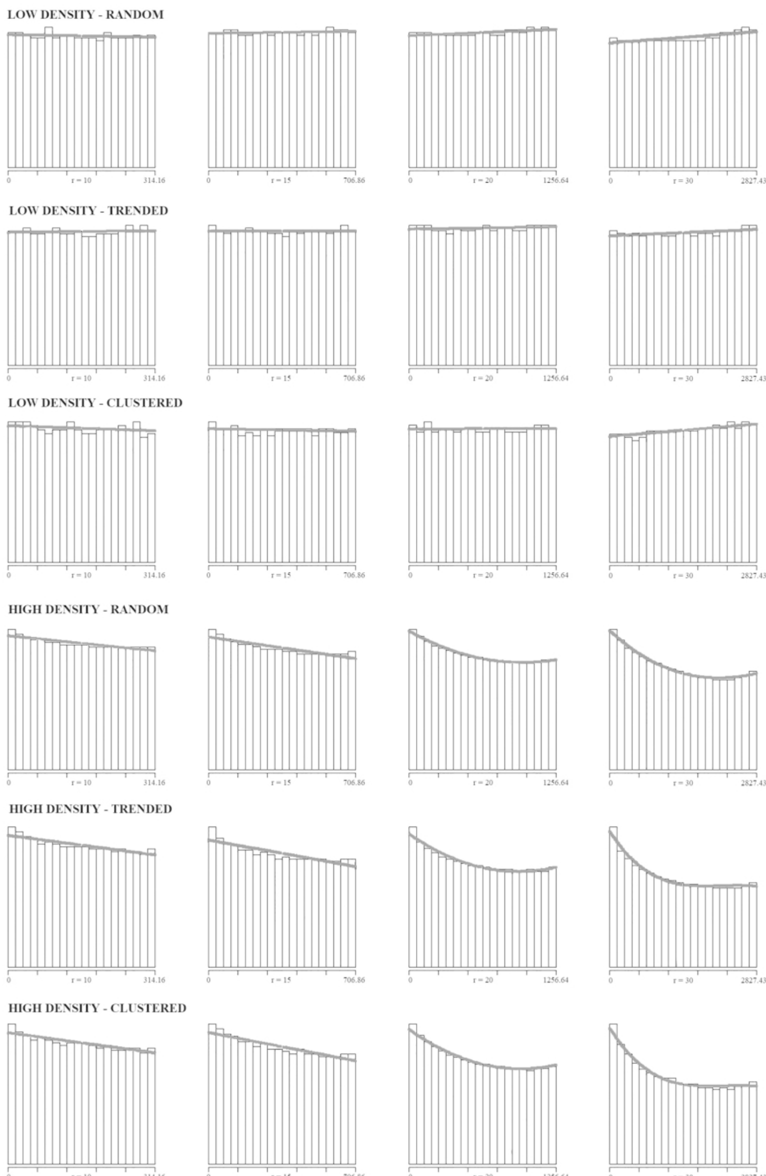

Figure 4. Monte Carlo probability density functions of the areas of detected trees within plots of radius r=10,15,20,30 m in low and high density forests with random, trended and clustered spatial patterns.

153x235mm (300 x 300 DPI)