Horvitz-Thompson–like estimation with distance-based detection probabilities for circular plot sampling of forests

Kasper Kansanen1 | Petteri Packalen2 | Matti Maltamo2 | Lauri Mehtätalo1

1 School of Computing, University of Eastern Finland, Joensuu, Finland
2 School of Forest Sciences, University of Eastern Finland, Joensuu, Finland

Correspondence
Kasper Kansanen, School of Computing, University of Eastern Finland, Joensuu, Finland.
Email: kasper.kansanen@uef.fi

Funding information
Academy of Finland, Grant/Award Number: 310073

Abstract
In circular plot sampling, trees within a given distance from the sample plot location constitute a sample, which is used to infer characteristics of interest for the forest area. If the sample is collected using a technical device located at the sampling point, e.g., a terrestrial laser scanner, all trees of the sample plot cannot be observed because they hide behind each other. We propose a Horvitz-Thompson–like estimator with distance-based detection probabilities derived from stochastic geometry for estimation of population totals such as stem density and basal area in such situations. We show that our estimator is unbiased for Poisson forests and give estimates of variance and approximate confidence intervals for the estimator, unlike any previous methods. We compare the estimator to two previously published benchmark methods. The comparison is done through a simulation study where several plots are simulated either from field measured data or different marked point processes. The simulations show that the estimator produces lower or comparable error values than the other methods. In the sample plots based on the field measured data, the bias is relatively small—relative mean of errors for stem density, for example, varying from 0.3% to 2.2%, depending on the detection condition. The empirical coverage probabilities of the approximate confidence intervals are either similar to the nominal levels or conservative in these sample plots.

Keywords
Circular plot sampling, forest remote sensing, stochastic geometry, terrestrial laser scanning

1 | INTRODUCTION

Circular plot sampling is a commonly used method in forest inventory (Gregoire and Valentine, 2007). In this form of sampling, a location is selected as a center point of a plot and the surrounding area within a given distance is observed from that point. The objective is to gather information on forest characteristics of interest, such as stem density \( N \) (the number of tree stems per unit area) and basal area \( G \) (the sum of the cross-sectional areas of the stems at breast height per unit area). We focus on \( N \) and \( G \) in this study. If the sampling is performed visually by a person, they can move slightly from the sampling point to observe trees that are not visible from the sampling point because of some sort of an obstacle. However, if the sampling is performed with a technical device, such as a laser...
scanner, the device cannot usually be moved from its original position. Hence, these sampling methods are susceptible to errors caused by incomplete detection. Lately, terrestrial laser scanning (TLS) has gained popularity as a tool for circular plot sampling and it is therefore topical to consider the detection problems in these sampling situations.

TLS is a method of close-range remote sensing where a LiDAR scanner is operated on ground level to produce a three-dimensional point cloud of the surroundings. Typically, the scanner remains stationary on a tripod and rotates 360 degrees in the horizontal plane, while the laser scans in the vertical direction in some device-dependent opening angle. In forestry, the strength of TLS is that accurate measurements of the forest structure below the canopy can be made. Methods for detecting individual tree objects from the point cloud have been developed (see, eg, Strahler et al., 2008; Lovell et al., 2011; Liang et al., 2012; Raumonen et al., 2015; Ravaglia et al., 2017), meaning that tree-level information for forest inventory purposes can be gathered using TLS. In this study, we focus strictly on tree stems (or more specifically, their cross-sections at a certain reference height) and ignore other parts of trees.

TLS data collection can be performed either as a single- or a multiscan setup. The difference between these two setups is that with single-scan one forest area is scanned only once, whereas with multiscan, the same area is scanned from several different locations with overlapping scanning areas to produce a more accurate point cloud, having echoes from different sides of trees and possibly from all of the trees in the area. However, scanning from several locations is more time-consuming and there is the added difficulty of combining the data from different scans. From an estimation point of view, several scans in nearby locations are suboptimal compared to single scans from distant locations. To give an example, let us consider a situation where an inventory of a large forest area is needed. For multiscan, you would place a scanner to five locations in such a way that four of the scan positions are on the boundary of the plot and one is in the middle of the plot. This produces measurements from one field plot, the central one—at least if the idea is to derive more precise measurements from a field plot. The scanning areas not overlapping with the central one would still be in essence single-scan plots. If movement in the inventory area does not significantly add to costs of the data collection mission, then by placing the five plots apart from each other would give you data from five measurement points, not one, with the same costs, covering more of the inventory area and taking into account possible spatial autocorrelations and trends in the forest.

We focus on the single-scan case where individual tree stems have already been detected from the point cloud, meaning that at least locations and diameters at breast height (DBH) of these detected trees are available. Naturally, the locations and DBH will have estimation errors (see, eg, Lovell et al., 2011). The magnitude of these estimation errors depends on the data collection parameters, the underlying forest structure, and the method used to derive the location and DBH. We shall not consider this error source, not because it is not significant, but to focus on solving problems in a different part of the forest inventory pipeline. The problem with single-scan TLS that we are focusing on is that not all trees can be detected, which can lead to underestimation of forest characteristics of interest, such as stem density or basal area. There are different possible reasons for not detecting all of the trees. For example, the tree stems closer to the scanner produce nonvisible areas behind them so that trees further away from the scanner located in the nonvisible areas cannot be seen (Figure 1). Undergrowth, low branches, or other objects can block the submitted laser pulse. Small trees far away from the scanner can produce very few laser returns. We consider here only the effect of tree stems producing nonvisible areas behind them; generalization to other obstacles is trivial.

Correcting the problem of nondetection in single-scan TLS has garnered some attention. The use of gap probabilities of a Poisson forest has been proposed as a means for correction (eg, Lovell et al., 2011). Duncanson et al. (2014) and Astrup et al. (2014) proposed models for detection probability based on traditional distance sampling. Seidel and Ammer (2014) proposed a correction factor based on the nonvisible area of the TLS plot. Olofsson
and Olsson (2018) used only the area visible to the scanner as a sampling window. Kuronen et al. (2019) improved on the work of Olofsson and Olsson (2018) by modifying the visible area based on what we call a detection condition, producing a weight for every tree that depends on its DBH and the detection condition: is the tree detected only if it is fully outside of the nonvisible area, or if the center point is visible, or if any small visible part of stem is enough for detection, or something in between, a partial visibility? The premise for the work of Kuronen et al. (2019) was that the estimator of Olofsson and Olsson (2018) produced large under- and overestimation in Poisson forests when the detection condition was either full visibility or any visibility, respectively, and deduced that this was because the area from which trees could be detected in these cases was not the same as the area visible from the scanner. However, Kuronen et al. (2019) also found that their estimator has a positive bias in the Poisson process case.

Kansanen et al. (2016) proposed estimators for stem density that correct a nondetection problem in the case of individual tree detection from airborne laser scanning (ALS) data. In ALS, unlike TLS, the scanner is operated from an aircraft and the laser measurements are attained from above, mainly from the forest canopy. They assumed that a tree could be left undetected if it was covered by the larger tree crowns. The most promising of the estimators was a Horvitz-Thompson–like estimator where the detection probability for each tree was calculated based on the planar set formed as a union of the projections of the larger tree crowns on ground that was morphologically transformed based on the detection condition. This method requires the maximum radii of detected tree crowns and the locations of their centers.

Here, we take a similar approach for calculating detection probabilities in the TLS case, and more generally, any circular plot sampling case with similar problems with incomplete observation. Instead of the size of the tree crowns, we assume that there is a certain hierarchy on trees being left undetected based on their distances from the scanner. The detection condition is taken into account by a tuning parameter that controls a morphological transformation of the area visible, or nonvisible, from the scanning location. The detection probabilities are used in a Horvitz-Thompson–like estimator to produce estimates of a population total of interest, such as stem density or basal area. The main difference between our estimator and the estimator of Kuronen et al. (2019) is that our detection probabilities are distance-based, whereas their construction is area-based. Hence, it could be easier to combine our method with other, possibly distance sampling based, methods that take into account some other effects on tree detection from TLS data that depend on the distance from the scanner. The performance of the estimator is compared with those of Olofsson and Olsson (2018) and Kuronen et al. (2019) in circular sampling plots that are either fully simulated from point process models or based on field measurements of diameters and locations of all trees of rectangular fixed-area plots.

2 | METHODOLOGY

We consider the forest as a realization of a marked point process (see, eg, Illian et al., 2008) \( M = \{ (x_i, d_i, m_i) \} \), where \( x_i \) are the locations of the stem centers, generated by some spatial point process, \( d_i \) are the DBH of the trees, governed by some distribution, and \( m_i \) is a mark of interest related to tree \( i \). If we are interested in a stem density estimate \( \hat{N} \), then \( m_i = 1 \). If we are interested in total basal area \( \hat{G} \), then \( m_i \) is the basal area of tree \( i \). Other marks, such as volume or biomass, are also possible. We observe \( n \) trees from a total of \( N \) in some window of interest \( W \) (ie, fixed-area sample plot), which, without loss of generality, is centered at the origin of the plane. “To observe” means here that the locations, DBHs, and marks of interest of the trees are known to us, at least approximately.

We model the tree stems as discs \( B(x_i, d_i/2) \) centered at \( x_i \) with radius \( d_i/2 \). We assume that the trees can be ordered based on their distances from the origin \( r_i - d_i/2 \), where \( r_i \) is the distance of the stem center from the origin. In other words, the trees are ordered based on the shortest distance to their outer bark. Henceforth we assume that the index \( i \) is ordered, \( i = 1 \) being the tree closest to the origin, \( i = 2 \) being the second closest, and so forth. We also assume that no tree disc \( B(x_i, d_i/2) \) covers the origin.

We assume that the ability to detect tree \( i \) is related to the trees closer to the origin than tree \( i \). A tree forms a subset of plane \( A_i \) that cannot be seen from the origin as a union of the stem disc \( B(x_i, d_i/2) \) and the area between the two tangent lines of the disc running through the origin (see Figure 1).

We suggest the following for the detection probability \( p \) of tree \( i \). If \( i = 1 \), the tree is detected for sure, \( p(r_1) = 1 \). Otherwise,

\[
\begin{align*}
    p_c(r_i, d_i) &= \begin{cases} 
    1 & \text{if } \frac{l[\partial B(o, r), \bigcup_{j=1}^{i-1} A_j] B(o, d_i/2)}{2 \pi r_i}, \quad \alpha < 0 \\
    1 & \text{if } \frac{l[\partial B(o, r), \bigcup_{j=1}^{i-1} A_j]}{2 \pi r_i}, \quad \alpha = 0 \\
    1 & \text{if } \frac{l[\partial B(o, r), \bigcup_{j=1}^{i-1} A_j] B(o, d_i/2)}{2 \pi r_i}, \quad \alpha > 0 
    \end{cases}
\end{align*}
\]

(1)

where \( l[\partial B(o, r), X] \) is the total length of the arcs of a circle with radius \( r \) that is inside the set \( X \) and \( \alpha \) is a tun-
ing parameter that controls the morphological transformations $\Theta$ (erosion) and $\odot$ (dilation). The value of $\alpha$ has to be fixed before estimation.

The reasoning behind the detection probabilities in Equation (1) is as follows. If we assume that a tree center point is uniformly distributed on a circle of radius $r_i$, then the probability that it is located in some certain arc of length $l$ is $l/(2\pi r_i)$, the proportional length of the arc. Hence, the probability for a tree to be hidden can be calculated based on the lengths of arcs generated by the nonvisible areas $A_j$, $j < i$. This is achieved by taking a union of $A_j$, which forms the total nonvisible area that could affect the detection of tree $i$, performing a morphological transformation related to a detection condition, and calculating the proportional length of the circle arcs inside that transformed set. The parameter $\alpha$ controls the detection condition. If $\alpha = 0$, the center point of the tree must be visible for detection (center point visibility detection condition, abbreviated as center in the tables and figures). If $\alpha = -1$, the tree is hidden only if its disc is fully inside the nonvisible area. In other words, any visible part produces detection (any visibility detection condition, abbreviated as any). If $\alpha = 1$, the tree must be fully visible, or fully outside the union of $A_j$, for detection (full visibility detection condition, abbreviated as full). Examples of the nondetectable areas related to these detection conditions are presented in Figure 2. After the probability for a tree to be hidden has been calculated, the probability of being detected can be calculated as a probability of the complement event.

With these detection probabilities, a Horvitz-Thompson-like estimator for a population total of interest $\tau = \sum_{i=1}^{N} m_i$ can be formed:

$$\hat{\tau} = \sum_{i=1}^{N} \frac{m_i I_i}{p_\alpha(r_i, d_i)}.$$  \hspace{1cm} (2)

$I_i$ is an indicator variable that equals 1 if tree $i$ is detected and 0 if it is not detected. Hence, in practice, only the detection probabilities and marks of detected trees are needed.

FIGURE 2  Example of detection probability calculation for tree $i$ (the black disc) in the plot of Figure 1. Trees $j$, $j < i$ (the black circles) that are closer to the plot center (the red cross) than tree $i$ produce areas of nondetection (colored with gray) related to nonvisible areas behind those trees ($A_j$) and a geometrical transformation dependent on the size of tree $i$ and the detection condition. The red circle arcs, coming from a circle with radius corresponding to the distance of tree $i$ center point from the center of the plot, are the locations where the tree center point could be located and left undetected. The proportional length of the arcs is the probability to be hidden, and the detection probability can be calculated as a probability of the complement event. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.
However, we take into account the invisible areas produced by undetected trees when we calculate the detection probabilities. That is, although a proper detection has not occurred, for example, DBH estimation is not possible, these trees increase the nonvisible area and produce a nonzero term \(A_j\) to \(\bigcup A_j\) in Equation (1). Note that this is not a concern in the case of any visibility detection condition, as the nonvisible areas created by trees that are not detected are fully contained in the nonvisible areas created by the detected trees.

It should be noted that there is no mathematical reason to constrain \(\alpha\) from \(-1\) to \(1\). It could be almost any real number. Values \(\alpha > 1\) would mean that a tree has to avoid contact with the nonvisible area created by other trees by some margin related to \(\alpha\) and its DBH to be detected. On the other hand, \(\alpha < -1\) would mean that trees could be detected even if they were fully inside the nonvisible area, again by some margin. When \(\alpha \rightarrow -\infty\), the detection probabilities approach zero, meaning that no trees apart from the one closest to the plot center should be observed. Depending on how division by zero is interpreted, the stem number estimate is either one or infinite number of trees. Hence, it would be natural to bound \(\alpha\) from above in such a way that the detection probabilities are nonzero.

### 2.1 Unbiasedness of the estimator for homogeneous Poisson process

As a shorthand, let us write \(p_i\) for the probability of detection for tree \(i\) and \(S_i\) for the parts of the origin-centered circle with radius \(r_i\) that belong to the set from which tree \(i\) could not be detected, generated by the trees closer to the origin. Let us write \(S_i^c\) for the parts of the circle that do not belong to the aforementioned set, and hence, the locations from which tree \(i\) could be detected. Let us also write \(|S_i|\) for the proportional length of \(S_i\), so that \(|S_i| + |S_i^c| = 1\). Now

\[
E[\hat{\tau}] = \sum_{i=1}^{N} \frac{m_i E[I_i]}{p_i}.
\]

Here, \(I_i\) is an indicator function of tree \(i\) belonging to \(S_i^c\). When the homogeneous Poisson process is considered in polar coordinates, the angles of the points are uniformly distributed, and hence, the distribution of a point at a distance \(r_i\) is uniform on a circle of radius \(r_i\). From this uniformity, it follows that the probability of a point hitting a certain subset of the circle is proportional to the length of that subset. Hence, \(E[I_i] = 1 \times |S_i^c| + 0 \times |S_i| = |S_i^c|\). But now our construction for the probabilities of detection is also \(p_i = |S_i^c|\), and it follows that

\[
E[\hat{\tau}] = \sum_{i=1}^{N} m_i = \tau,
\]

hence the estimator is unbiased.

### 2.2 Estimated variance of the estimator

When the assumptions related to our Horvitz-Thompson–like estimator are met—the sequential nature of detection holds, the detection condition is true, and the data are generated by a process where the points are distributed uniformly around the origin, for example, the homogeneous Poisson process—the detection probabilities we approximate in Equation (1) are the true detection probabilities, and hence, our estimator coincides with the Horvitz-Thompson estimator. An unbiased estimator for the variance of the Horvitz-Thompson estimator is known (see, eg, Thompson, 2012, p. 70):

\[
\hat{\vartheta}(\hat{\tau}) = \sum_{i=1}^{n} \left( \frac{1}{p_i} - 1 \right) m_i^2 + 2 \sum_{i=1}^{n} \sum_{j>i} \left( \frac{1}{p_i p_j} - \frac{1}{p_{ij}} \right) m_i m_j,
\]

where \(p_{ij}\) is the probability to include both \(i\) and \(j\) in the sample, or in our case, to detect both trees \(i\) and \(j\). The indexing goes over the detected trees, not the whole population. Through conditional probability, we can write \(p_{ij} = p_{ij|l} p_l\), where \(p_{ij|l}\) is the probability of detecting \(j\) when \(i\) is detected. Because of our sequential construction, the probability of detecting \(j\) always takes into account the fact that \(i\) has been detected for \(j > i\). Hence, \(p_{ij|l} = p_j\) and \(p_{ij} = p_i p_j\), and furthermore,

\[
\hat{\vartheta}(\hat{\tau}) = \sum_{i=1}^{n} \left( \frac{1}{p_i} - 1 \right) m_i^2.
\]

The estimated variance makes it possible to produce approximate confidence intervals (Thompson, 2012, p. 70)

\[
\hat{\tau} \pm t_{\alpha/2,n-1} \sqrt{\hat{\vartheta}(\hat{\tau})},
\]

where \(t_{\alpha/2,n-1}\) is the critical value corresponding to the selected nominal confidence level from the \(t\)-distribution with \(n - 1\) degrees of freedom. The \(t\)-distribution can be replaced with a standard normal distribution if the number of detected trees is large. We follow the pragmatic rule given by Thompson (2012) and use the \(t\)-distribution if the
number of detected trees is less than 50. We note that a large portion of the simulated circular plots we use here for testing have less than 50 detected trees.

2.3 Comparison to the estimator of Kuronen et al.

The estimator of Kuronen et al. (2019) is of the same general form as ours (Equation (2)). However, instead of weighting the detections with detection probabilities \( p_\alpha(r_i, d_i) \) which we have formulated in Equation (1), the detections are weighted with weights \( w_i \) that correspond to surface areas of the visible (or nonvisible) planar set that has been morphologically transformed according to the detection condition. Using the notation that we have presented the weights can be written as

\[
\begin{align*}
    w_i &= \begin{cases} 
        1 - \frac{|(\cup_{j=1}^N A_j) \cap B(o, d_i/2)|}{|W|}, & \alpha = -1 \\
        1 - \frac{|(\cup_{j=1}^N A_j)|}{|W|}, & \alpha = 0 \\
        1 - \frac{|(\cup_{j=1}^N A_j) \cup B(o, d_i/2)|}{|W|}, & \alpha = 1
    \end{cases} 
\end{align*}
\]

(4)

The weights of Kuronen et al. (2019) did not include the parameter \( \alpha \) to their construction, but considered the detection conditions as totally separate cases. They did, however, consider a different kind of a weighting for the partial visibility detection condition, related to the visible proportion of the boundary of the stem disc, whereas in our construction, partial visibility cases are handled by setting \( \alpha \) to a value that differs from \(-1, 0, \) and 1 and are related to the proportion of visible stem disc radius.

Comparison of our weights in Equation (1) and the weights of Kuronen et al. (2019) in Equation (4) shows that the main difference between the two constructions is the sequential nature of our detection probabilities: in our construction, only the trees that are before tree \( i \) in the ordering affect the weight of tree \( i \), whereas in the construction of Kuronen et al. (2019), all of the trees, including tree \( i \), have an effect. The premises of the constructions are also different. While we approach the situation from a probabilistic perspective, Kuronen et al. (2019) are following the idea that a tree that is detected should belong to such a sampling window from which it can be detected, and hence, the size of the sampling window should change for every tree. Kuronen et al. (2019) was not able to show unbiasedness of their estimator—in fact, they showed that the estimator has positive bias even in the case of Poisson process data—and they did not present an estimator for the variance either.

2.4 Characterizing deviations from the Poisson assumption

As our estimator can be shown to be unbiased when the data are generated by a Poisson process, it is of interest to see how the bias of the estimator behaves when the spatial structure of the data deviates from this Poisson assumption to either clustering or regularity, and if the magnitude of this deviation has an effect on the magnitude of the bias. We use the \( L \)-function, derived from the \( K \)-function as

\[
L(r) = \sqrt{\frac{K(r)}{\pi}}, \quad r \geq 0
\]

to characterize the spatial structure of a point pattern in a plot (see, e.g., Illian et al., 2008, Chapter 4.3). The \( K \)-function is related to the expected number of points of a point process that are closer than distance \( r \) to a point of the process. The \( L \)-function of the Poisson process is known: \( L(r) = r \). For a clustered process, \( L(r) > r \), and for a regular process, \( L(r) < r \).

We estimated the \( L \)-functions for all the simulated plots and point patterns by using the R package spatstat (Baddeley et al., 2015) with isotropic edge corrections (see, e.g., Illian et al., 2008, Section 4.2.2) with \( r \) ranging from 0 to 5 m. If the point pattern contained only one point, a case where \( L \)-function cannot be estimated, we took the convention of characterizing it as coming from a Poisson process and set \( \hat{L}(r) = r \). To produce measures of deviation from the Poisson process, we used signed maximum deviation between \( \hat{L} \) and the identity line. First find the distance at which maximum absolute deviation occurs,

\[
r^* = \arg \max_{r \in [0,5]} |r - \hat{L}(r)|.
\]

Then, the \( L \)-based measure of deviation is given by

\[
r^* - \hat{L}(r^*),
\]

producing positive values if the pattern shows stronger signs of regularity than clustering, and negative values if the pattern shows stronger signs of clustering than regularity.

We also used this deviation measure to characterize whole data sets by first calculating the mean \( L \)-function over all \( n \) plots of the data as
\[ \bar{L}(r) = \frac{1}{n} \sum_{i=1}^{n} \hat{L}_i(r), \]

and then calculating the deviation measure for the mean \( L \)-function. The mean \( L \)-function can be seen as an estimator of the expected \( L \)-function of the process that has generated the data, and as \( L(r) = r \) for Poisson process is also an expectation, this is quite natural way to characterize the deviation of a process from the Poisson process. The \( L \)-functions were estimated in R (R Core Team, 2019) with the function \texttt{Lest} in the R package \texttt{spatstat} (Baddeley et al., 2015).

### Materials

#### 3.1 Field data

We use field data measured in 2017 in Eastern Finland. It consists of 111 mapped square 30 × 30 m plots placed to the area of about 43 000 ha. The plots were sampled from a systematic network using a priori information of stand maturity and dominant tree species. Tree species were determined and tree height and DBH were measured for all trees with a DBH ≥ 5 cm. Tree locations were determined using a variant of the triangulation method proposed by Korpela et al. (2007). Field measurements are documented in detail in Räty et al. (2019).

The central 10 × 10 m² of every plot was covered with a triangular grid where the distance between points was at least 1 m. Every grid point was used as a center for a simulated circular sampling plot with radius 10 m. If a tree stem covered the center point, it was removed. This produced 111 × 126 = 13 986 simulated plots using true field data. We use the term “simulated plot” here to make it clear that the data have not been gathered with remote sensing systems such as TLS from the plot center. Instead, the field measurements of tree locations and DBH are used as realizations of a point process and used as if circular plot sampling via a remote sensing system and the subsequent identification of tree objects was performed. Statistics related to \( N, G \), and the \( L \)-function-based deviation measures in these plots are presented in Web Table 1.

#### 3.2 Process-simulated data

Circular field plots of radius 10 m were simulated from six different types of marked point processes: Poisson process, nonoverlapping discs process, Gibbs hard core process, Log-Gaussian Cox process, and a plantation-like process (generation of point patterns from these processes is more thoroughly depicted in Section 3.2.2). Two different variants of the nonoverlapping process, three variants of the Gibbs hard core process and five variants of the Log-Gaussian Cox process were simulated. The theoretical intensity \( N \) of the plantation process was 833 stems per hectare in all simulations, due to the construction of the process (described in Section 3.2.2). In the other cases, process intensities 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 stems per hectare were used. These intensities were chosen to cover the variation in stem densities in natural forests and have been previously used by, eg, Olofsson and Olsson (2018). The number of plots generated was 10 000 for the Poisson process, 1000 with every intensity, and 2000 per other point processes, 200 with every intensity. Special attention was given for the Poisson process because it is theoretically in line with the assumptions of our estimator and hence demonstrating the performance of the estimator in Poisson process data is of great importance. In all cases, the point process was simulated in a larger simulation window to take into account possible effects at the plot boundary. The trees that did not come into contact with the plot boundary at 10 m, and hence could not have any effect on the estimation, were removed. The number of points in the simulation window was always first generated from a Poisson distribution with expected value \( N | W | \), where \( |W| \) is the area of the simulation window, an exception being again the plantation process. Only point patterns where the central point of the plot was not covered by a stem disc were accepted. All simulations were done with R (R Core Team, 2019). Statistics related to the simulated \( N, G \), and the \( L \)-function-based deviation measures are presented in Web Table 2 and example plots are presented in Web Figure 1.

#### 3.2.1 Generating tree diameters

An adjusted version of the parameter recovery method of Siipilehto and Mehtätalo (2013) was followed to produce Weibull distributions for the DBH. Mean DBH \( D \) and basal area \( G \) pairs of 6 cm and 3 m²/ha, 12 cm and 12 m²/ha, 15 cm and 20 m²/ha, and 21 cm and 35 m²/ha were taken as a starting point. Then, for shape parameter \( \gamma \) and scale parameter \( \beta \), the sum of differences between the means and quadratic means

\[ |\beta \Gamma(1 + 1/\gamma) - D| + |\beta^2 \Gamma(1 + 2/\gamma) - G/(N \pi /40000)| \]

was minimized. This produced 40 different DBH distributions with expected basal areas almost exactly agreeing with the four given mean values of \( G \) and expected DBH differing from the given mean values of DBH to keep \( N \) and \( G \) compatible. In other words, for every intensity, four dif-
different DBH distributions corresponding to different mean DBH and basal area were used to simulate the DBH marks for the trees. It should be noted that these intensities and distributions are parameters of the random processes, and hence, the simulated stem densities, mean DBH, and basal areas will be different from these numbers and vary according to the nature of the process.

3.2.2 Generating marked point patterns

Poisson process (abbreviated as Poisson in the tables and figures) forests with complete spatial randomness of points and independent and identically distributed DBH were generated in a circular window of radius 11 m, with the given intensities and DBH distributions. Nonoverlapping discs process (nonoverlapping), producing point patterns where the stem discs cannot overlap, was used to simulate forests with a more realistic hard core property. The first point location was generated uniformly in a circular window of radius 11 m and mark was generated from the DBH distribution. For points 2,…,M, a suggested point with a mark from the corresponding DBH distribution was generated uniformly in the same window. If the point was located in such a way that its stem disc did not overlap the previous discs, it was accepted; otherwise, it was rejected and the marked point was simulated again. At every step 2,…,M, 10 000 attempts to simulate were allowed. If the algorithm failed to simulate an acceptable point during these attempts, the simulation was regarded as finished and the resulting point pattern accepted. A second version of nonoverlapping discs process (nonoverlapping 10), where the nonoverlapping areas were discs with diameter 10×DBH of the corresponding stem, was also simulated similarly.

Another type of hard core process where the points are at least 1 m apart was simulated as a Gibbs hard core process (Gibbs 1). The points were first simulated in a 40×40 m square window. The locations of points were simulated from a Gibbs process with known number of points (see, eg, Illian et al., 2008). Then an independent sample was taken from the DBH distribution to assign marks to the points. A similar scheme was used to simulate a Gibbs 1.5 m hard core process (Gibbs 1.5) and a Gibbs 2 m hard core process (Gibbs 2).

Log-Gaussian Cox process (Cluster) with Matern covariance function was used to produce plots with spatial clustering. The smoothness and variance parameters of the covariance function were fixed to 2 and 1, whereas the range was varied between 2, 4, 6, 8, and 10 m, producing five different variants of a cluster process. The points were simulated in a 40×40 m square window. The Gaussian field was simulated with the R package RandomFields (Schlather et al., 2019) to produce the density field over the window. The points were then generated according to that density with the spatstat function rpoint (Baddeley et al., 2015).

As a special case, plantation-type data were also simulated. Trees were located in a 3×4 m grid in 40×40 m² window, leading to stand density of 833 stems per hectare. DBHs were simulated randomly from distributions determined as in Section 3.2.1 using the target values of D and G given there. The plot center point was located randomly among the trees in such a way that the whole plot was inside the square window.

4 EVALUATION OF RESULTS

Relative root-mean-squared errors

\[ \text{RMSE}\% = \frac{100}{\bar{y}} \times \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \]

and relative means of errors (sometimes called relative bias, even though it is only an estimate of it)

\[ \text{ME}\% = \frac{100}{\bar{y}} \times \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)}{n} \],

where \( \hat{y}_i \) is the estimated value, \( y_i \) is the true value, \( \bar{y} \) is the mean of the true values, and \( n \) is the number of plots (\( n = 13\,986 \) for the field data, 10,000 for Poisson process, and 2000 for the other processes), were used to evaluate and compare the estimation errors between different estimators. The estimators that we benchmark our method against are the estimators of Olofsson and Olsson (2018) and Kuronen et al. (2019). These estimators are the same when the detection condition is center point visibility. We also report the errors of estimation based on the detected trees with no corrections. It should be noted that both Olofsson and Olsson (2018) and Kuronen et al. (2019) assumed that the stem discs belong to the visible area, whereas we assume that they belong to the nonvisible area. In other words, in Equations (1) and (4), the nonvisible area \( A^*_j \) produced by tree \( j \) should be replaced by \( A^*_j = A_j \setminus B(x_j, d_j/2) \) to conform with Olofsson and Olsson (2018) and Kuronen et al. (2019). To compare only the differences of distance-based and area-based construction on the estimation, we wanted to keep all other effects the same and decided to modify these estimators to include the stem discs to the nonvisible area. This should not affect the results greatly; the important part is to have the detection conditions matching the nonvisible areas, in other words that the simulated trees have been correctly classified as...
TABLE 1  Estimation errors of stem density (N) and basal area (G) over the simulated plots based on the field data for detection conditions full visibility, center point visibility, and any visibility, and for estimators by us (HT), Olofsson and Olsson (2018) (O & O), and Kuronen et al. (2019). Lowest (absolute) values for each case are bolded. The estimators O & O and Kuronen et al. coincide in the center point case.

| Condition | Estimator | N  | RMSE% | ME% | G  | RMSE% | ME% |
|-----------|-----------|----|-------|-----|----|-------|-----|
| Full      | HT        | 8.0| 2.2   | 11.0| 4.1|       |     |
|           | O & O     | 9.0| −3.7  | 9.6 | −2.7|       |     |
|           | Kuronen et al. | 8.5| 3.3   | 12.0| 5.8|       |     |
|           | detected  | 22.5| −15.4| 18.3| −13.5|       |     |
| Center    | HT        | 6.2| 1.2   | 7.6 | 2.1|       |     |
|           | Kuronen et al. | 6.3| 1.6   | 7.8 | 2.8|       |     |
|           | detected  | 16.4| −10.8| 12.6| −8.7|       |     |
| Any       | HT        | 4.6| 0.3   | 4.8 | 0.4|       |     |
|           | O & O     | 10.2| 6.5  | 10.4| 7.4|       |     |
|           | Kuronen et al. | 4.6| 0.4   | 4.8 | 0.6|       |     |
|           | detected  | 10.7| −6.7 | 7.7 | −4.7|       |     |

either detected or not detected. In the result tables, we refer to our estimator as “HT,” the estimator of Olofsson and Olsson (2018) as “O & O,” the estimator of Kuronen et al. (2019) as “Kuronen et al.,” and finally, the estimator based on just the detected trees as “detected.” We evaluate the performance of the estimators with detection conditions full, center, and any. The estimators of Olofsson and Olsson (2018) and Kuronen et al. (2019) are the same when the detection condition is center point visibility.

5 RESULTS

In the plots based on the field measured data, the HT-like estimator produces better or very similar error values than the other methods in most cases (Table 1). The exception is the estimation of G when the detection condition is full visibility, where the estimator of Olofsson and Olsson (2018) produces lowest error values.

The results in the plots simulated from marked point processes are presented in Table 2 for the Poisson process data and Web Tables 3-5 for the others. When comparing the HT-like estimator and the estimator of Kuronen et al. (2019), the HT-like estimator produces better or very similar error values in data simulated from the Poisson process, Gibbs processes, and the nonoverlapping discs models. Especially in the Poisson process case, the ME% values of stem density estimation are very close to zero, which is consistent with the unbiasedness of the estimator for homogeneous Poisson process. In the clustered data, the estimator of Kuronen et al. (2019) produces lower error values than the HT-like estimator. Both of the estimators produce larger estimation errors in the clustered data than in the other types of data. Both of the estimators show overestimation in the regular data and underestimation in the clustered data. The estimator of Olofsson and Olsson (2018) does not show this behavior, but shows overestimation with full visibility detection condition and underestimation with any visibility detection condition in all of the simulated data sets. For the Gibbs hard core data sets and full visibility detection condition, it produces lower error values than the other two estimators.

Figure 3 represents the relationship between bias and deviation of a data set from the Poisson process assumption. The ME% of the different data sets are plotted against the L-function-based deviation measures calculated from the mean L-functions of the data sets. The Poisson process data, having deviation measure closest to zero, have ME% closest to zero, indicating unbiasedness. When the deviation measure grows larger, the errors start to also grow, showing that larger deviations into regularity produce larger overestimations, up to a point—in the very regular Gibbs 2 m hard core data and plantation data, the ME% starts to drop off. Due to the large hard core distances in these processes, the plots are not able to achieve the same stem densities as the less regular processes, and as Figure 4 shows for Poisson data, the stem density has a clear effect on estimation errors. On the other hand, deviations to clustering produce underestimation, and the magnitude of deviation affects the magnitude of underestimation. The deviation and detection condition also have a connection, the any visibility detection condition results being least

TABLE 2  Estimation errors of stem density (N) and basal area (G) over the Poisson process simulated plots for detection conditions full visibility, center point visibility, and any visibility, and for estimators by us (HT), Olofsson and Olsson (2018) (O & O), and Kuronen et al. (2019). Lowest (absolute) values for each case are bolded. The estimators O & O and Kuronen et al. coincide in the center point case.

| Condition | Estimator | N  | RMSE% | ME% | G  | RMSE% | ME% |
|-----------|-----------|----|-------|-----|----|-------|-----|
| Full      | HT        | 6.1| 0.0   | 13.6| 0.3|       |     |
|           | O & O     | 11.1| −7.5 | 16.1| −8.3|       |     |
|           | Kuronen et al. | 6.1| 0.7   | 16.5| 1.7|       |     |
|           | detected  | 28.3| −21.5| 34.0| −23.3|       |     |
| Center    | HT        | 4.8| 0.0   | 7.8 | 0.1|       |     |
|           | Kuronen et al. | 4.8| 0.2   | 8.2 | 0.6|       |     |
|           | detected  | 20.3| −15.1| 23.8| −16.0|       |     |
| Any       | HT        | 3.4| 0.0   | 5.0 | 0.0|       |     |
|           | O & O     | 12.0| 8.5  | 15.2| 9.6|       |     |
|           | Kuronen et al. | 3.4| 0.1   | 5.1 | 0.2|       |     |
|           | detected  | 11.9| −8.4 | 14.1| −8.8|       |     |
The relationships of the estimated stem densities and their estimated standard errors to the simulated stem densities in the Poisson process data are presented in Figure 4. The estimated stand densities are well centered around the identity line, with more variance in the estimates as the simulated stem density increases. The estimation results get better when going from full visibility to center point visibility, and further to any visibility detection condition, represented by the more concentrated point cloud. This effect is also shown by the incrementally smaller RMSE% values (Table 2). The estimated standard errors increase as the simulated stand density increases, as does the variation in the estimated standard errors. The errors decrease affected by the magnitude of clustering or regularity, and full visibility detection condition being affected the most. The field-data-based data set is somewhat regular, having $L$-function-based deviation of roughly 0.5. It should be noted that although the data sets have “on average” a certain type of behavior, e.g., heavily clustered or slightly regular, the plots in these data sets can have widely varying deviation measures (see Web Tables 1 and 2). However, the mean behavior seems to have the largest effect on the bias of the estimator.

To examine the estimator variance and the approximate confidence intervals, we calculated the 90%, 95%, and 99% intervals for stem density and basal area in the simulated data sets with different visibility conditions and then calculated how many of the simulated stem densities and basal areas belonged to their corresponding confidence intervals. The results about stem density are presented in Table 3 and Web Table 6. In the Poisson and nonoverlapping discs data sets, the observed coverage probabilities of the confidence intervals are very close to the nominal confidence levels with all of the detection conditions tested. In the slightly regular field data set, the percentages are mostly in line with the 99% confidence level, but the 90% and 95% intervals are slightly conservative. In the Gibbs hard core process data sets with strong regularity, the visibility condition has a larger effect on the goodness of the interval estimates. This can be especially seen in the 90% interval case, where under the full visibility condition, only 87.1% of the simulated stem densities belong to their corresponding intervals, whereas under the any visibility condition, the value is 93.2%, going from anticonservative to conservative intervals. However, the 99% intervals for Gibbs 1 m hard core process have very similar nominal and observed coverage probabilities. For Gibbs 1.5 and 2 m hard core process, most of the intervals are anticonservative, whereas for the plantation data set, they again are mostly conservative. In the clustered data sets, all of the intervals are anticonservative, indicating that the variances have been underestimated in addition to the underestimation present in the estimated stem densities. The results relating to basal area estimation, presented in Web Table 7, show similar effects than the stem density estimation results.
FIGURE 4  Horvitz-Thompson–like estimates of stem densities and their estimated standard errors as functions of the simulated stem density in the Poisson process data for detection conditions full visibility, center point visibility, and any visibility. The color corresponds to the number of sample plots in the corresponding square area of the figure. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

TABLE 3  The percentages of simulated stem densities belonging to their approximate confidence intervals for detection conditions full visibility, center point visibility, and any visibility.

| Data | Condition | Interval | 90%  | 95%  | 99%  |
|------|-----------|----------|------|------|------|
| Field | Full      | 94.0     | 97.5 | 99.5 |      |
|       | Center    | 94.1     | 97.4 | 99.3 |      |
|       | Any       | 93.4     | 96.5 | 98.8 |      |
| Poisson | Full     | 90.0     | 94.9 | 98.7 |      |
|       | Center    | 89.9     | 94.5 | 98.7 |      |
|       | Any       | 90.5     | 94.9 | 98.4 |      |

when moving from the full visibility to any visibility detection condition.

6 | DISCUSSION

Our estimator has a clear connection to distance sampling (see, eg, Buckland et al., 2004), which is commonly used in estimating for example sizes of wild animal populations from a sample of observations at different distances from the observer. Most commonly, a parametric model describing the probability to detect an animal at a certain distance is fitted to the sample data, and then a Horvitz-Thompson–like estimator is used, either with the fitted probabilities for each observation depending on its distance, or the mean probability for every observation. Fitting of a parametric detection probability model is also the approach of Duncanson et al. (2014) and Astrup et al. (2014) to estimating population totals from TLS data in forestry. Our construction could be seen as a nonparametric distance sampling estimator—provided that the value of the tuning parameter is known—where the detection probabilities are calculated directly from the data for each observation individually. It could also be seen as distance sampling with empirical detection probabilities under an assumed Poisson model.

The strength of our construction, when compared to Kuronen et al. (2019) and other previously published methods, is the ability to prove the unbiasedness of the estimator when the data are generated by a Poisson process—Kuronen et al. (2019) showed that there is a positive
bias in their estimator even in this case—and a way to calculate approximate confidence intervals. In both cases, the sequential nature of the detection probabilities, i.e., only the trees closer to the plot center affecting the detection of trees further away, is of great use. In addition to the unbiasedness in Poisson process data, Figure 3 shows that there is a natural relation between deviations from Poisson process and the bias of the estimator: regularity leads to overestimation, clustering to underestimation.

The HT-like estimator with distance-based detection probabilities produces in most simulation cases estimation errors that are lower than or comparable to the errors produced by the estimators of Olofsson and Olsson (2018) and Kuronen et al. (2019). The estimator works very well in the Poisson process data, as is to be expected. The estimator is slightly biased in the slightly regular field data—the L-based measure of the field data set is between the measure of the Poisson process data set and the more regular hard core processes. However, the bias is to a “safe” direction because the confidence intervals are conservative in this case. The results indicate that the estimator should not be used “as is” for strongly clustered or regular data. In these cases, it might be possible to tune the estimator by adjusting the parameter $\alpha$. An example of unintentional tuning is, for example, the estimator of Olofsson and Olsson (2018) producing lower ME% of $G$ in the field data with full visibility detection condition. This is because the correction is “wrong”: when compared to the area from which trees can be detected based on the detection condition, a larger area is used for weighting, pushing the estimates down. Intentional tuning in our case would entail calculating estimates with several values of parameter $\alpha$ in a training data set and choosing a value that minimizes some error value as the “right” detection condition for that type of data.

Figure 4 further demonstrates the properties of our estimator when the data are Poisson distributed. The unbiasedness is shown by the concentration of the stem density estimates along the identity lines. The increasing simulated stem density produces increasing estimates of standard error. As the number of stems increases, so does the number of detected stems, which introduces more positive terms to the formula of the variance estimator (Equation (3)), which produces a larger variance. The increase caused by a single tree is connected to its detection probability; small probability leads to a large increase.

As it has been mentioned, the value of parameter $\alpha$ has to be fixed before estimation. The aforementioned tuning in training data is one option for finding this value. Another one is to have some prior information on the data collection device, such as TLS scanner, and the methodology used for locating trees and estimating the DBH, namely, what detection condition does the pairing follow. For example, it might be clear that because of the point cloud density provided by the TLS and the accuracy of the detection method, the DBH and location of a tree can be estimated only when the tree center point is visible.

An interesting question for further studies is if aggregation of the detection probabilities could benefit the estimation by reducing variance and making it more robust. In distance sampling, it is quite common to integrate over distance to achieve just one mean detection probability that is used for every detected object (see, e.g., Buckland et al., 2004). Here, it would be similarly possible to calculate the mean detection probability in a plot simply through the individual detection probabilities. However, this would also aggregate over the detected DBH, and the effect on, e.g., basal area estimation is unclear. Also, the number of detected trees per plot can be quite low, possibly leading to a mean detection probability that does not well portray the detections on that plot. The detection probabilities could also be aggregated over plots in a forest inventory, by fitting a detection function depending on distance and DBH to the probabilities. This would produce “corrected” detection probabilities for every tree, and if the spatial structure of the forest inventory area was sufficiently homogeneous, this might provide improvements to estimation over the whole inventory.

The construction could be generalized in several ways. For example, other objects than tree stems producing non-visible areas behind them could be added in when calculating the detection probabilities, as long as the object is detected from the data in such a way that the geometry of the non-visible area it produces is known. The visibility condition of a tree could depend on its size and distance from the scanner. Other effects connecting distance and detection probability, such as the number of laser returns per unit area that diminishes as a function of distance, could be added to the calculation of detection probabilities, either through the parameter $\alpha$ or as additional weights.

7 | CONCLUSIONS

We have presented a Horvitz-Thompson–like estimator with distance-based detection probabilities for single scan TLS and evaluated its performance in a simulation study consisting of plots based on field measured data and simulated from several different point processes. The estimator produces error values comparable to or lower than the two benchmark methods in most cases. A notable exception is the clustered data. The estimator is unbiased when the data come from a Poisson process. The variance of the estimator can be estimated and approximate confidence intervals can be formed. The detection probability construction is easily generalizable to allow objects other than tree stems
to affect detection and addition of different distance-based effects on detection. The estimator has a tuning parameter that could be used to find an appropriate detection condition for a scanner in training data.

ACKNOWLEDGMENTS
This work was funded by the Academy of Finland (decision number 310073). We would like to thank Juha Heikkinen, Mikko Kuronen, and Mari Myllymäki for useful discussions on the topic.

DATA AVAILABILITY STATEMENT
The process simulated data, described in Section 3.2, that support the findings in this paper are openly available in OSF, at http://doi.org/10.17605/OSF.IO/P875Q (Kansanen, 2020). The field data and the sample plots derived from it are not shared.

ORCID
Kasper Kansanen https://orcid.org/0000-0002-3645-3402
Petteri Packalen https://orcid.org/0000-0003-1804-0011
Matti Maltamo https://orcid.org/0000-0002-9904-3371
Lauri Mehtätalo https://orcid.org/0000-0002-8128-0598

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, 666–676.
Baddeley, A., Rubak, E. and Turner, R. (2015) Spatial Point Patterns: Methodology and Applications with R. London: Chapman and Hall/CRC Press.
Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. (2004) Advanced Distance Sampling. Oxford: Oxford University Press.
Duncanson, L., Cook, B., Hurtt, G. and Dubayah, R. (2014) An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. Remote Sensing of Environment, 154, 378–386.
Gregoire, T. and Valentine, H. (2007) Sampling Strategies for Natural Resources and the Environment. New York: Chapman and Hall/CRC.
Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008) Statistical Analysis and Modelling of Spatial Point Patterns. Chichester: John Wiley & Sons, Ltd.
Kansanen, K. (2020) Simulation data used in Horvitz–Thompson-like estimation with distance-based detection probabilities for circular plot sampling of forests by Kansanen, Packalen, Maltamo and Mehtätalo. OSF: 10.17605/OSF.IO/P875Q.
Kansanen, K., Vauhkonen, J., Lähiivaara, T. and Mehtätalo, L. (2016) Stand density estimators based on individual tree detection and stochastic geometry. Canadian Journal of Forest Research, 46, 1359–1366.
Korpela, I., Tuomola, T. and Välimäki, E. (2007) Mapping forest plots: An efficient method combining photogrammetry and field triangulation. Silva Fennica, 41, 457–469.
Kuronen, M., Henttonen, H.M. and Myllymäki, M. (2019) Correcting for nondetection in estimating forest characteristics from single-scan terrestrial laser measurements. Canadian Journal of Forest Research, 49, 96–103.
Liang, X., Litkey, P., Hyypia, J., Kaartinen, H., Vastaranta, M. and Holopainen, M. (2012) Automatic stem mapping using single-scan terrestrial laser scanning. IEEE Transactions on Geoscience and Remote Sensing, 50, 661–670.
Lovell, J., Jupp, D., Newham, G. and Culvenor, D. (2011) Measuring tree stem diameters using intensity profiles from ground-based scanning lidar from a fixed viewpoint. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 46–55.
Mehtätalo, L. (2019) Imfor: Functions for Forest Biometrics. URL: https://CRAN.R-project.org/package=Imfor.
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, 365–377.
R Core Team. (2019) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
Raumonen, P., Casella, E., Calders, K., Murphy, S., Åkerblom, M. and Kaasalainen, M. (2015) Massive-scale tree modelling from TLS data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II 3, 189–196.
Ravaglia, J., Bac, A. and Fournier, R.A. (2017) Extraction of tubular shapes from dense point clouds and application to tree reconstruction from laser scanned data. Computers & Graphics, 66, 23–33. Shape Modeling International 2017.
Räty, J., Packalen, P. and Maltamo, M. (2019) Nearest neighbor imputation of logwood volumes using bi-temporal ALS, multispectral ALS and aerial images. Scandinavian Journal of Forest Research, 34, 469–483.
Schlather, M., Malinowski, A., Oesting, M., Boecker, D., Strokorb, K., Engelke, S. et al. (2019) RandomFields: Simulation and Analysis of Random Fields. R package version 3.3.6.
Seidel, D. and Ammer, C. (2014) Efficient measurements of basal area in short rotation forests based on terrestrial laser scanning under special consideration of shadowing. iForest - Biogeosciences and Forestry, 7, 227–232.
Siipilehto, J. and Mehtätalo, L. (2013) Parameter recovery vs. parameter prediction for the Weibull distribution validated for Scots pine stands in Finland. Silva Fennica, 47(4), 1–22.
Strahler, A.H., Jupp, D.L.B., Woodcock, C.E., Schaaf, C.B., Yao, T., Zhao, F. et al. (2008) Retrieval of forest structural parameters using a ground-based lidar instrument (echidna®). Canadian Journal of Remote Sensing, 34, S426–S440.
Thompson, S.K. (2012) Sampling. 3rd edition, Hoboken, NJ: Wiley.

SUPPORTING INFORMATION
Web Figures and Tables, referenced in Sections 3 and 5, are available with this paper at the Biometrics website on
Example codes and data are also available as Supporting Information. The functions related to our estimator are available in the R package `lmfor`, available on CRAN at [https://CRAN.R-project.org/package=lmfor](https://CRAN.R-project.org/package=lmfor) (Mehtätalo, 2019). See the documentation for function `HTest_cps` for further information.

**How to cite this article:** Kansanen K, Packalen P, Maltamo M, Mehtätalo L. Horvitz-Thompson–like estimation with distance-based detection probabilities for circular plot sampling of forests. *Biometrics.* 2021;77:715–728. [https://doi.org/10.1111/biom.13312](https://doi.org/10.1111/biom.13312)