

A Plotless Density Estimator based on the Asymptotic Limit of Ordered Distance Estimation Values

Abstract: Estimation of tree density from point-tree distances is an attractive option for quick inventory of new sites, but estimators that are unbiased in clustered and dispersed situations have not been found. Noting that bias of an estimator derived from distances to the k^{th} nearest neighbor from a random point tends to decrease with increasing k , a method is proposed for estimating the limit of an asymptotic function through a set of ordered distance estimators. A standard asymptotic model is derived from the limiting case of a clustered distribution. The proposed estimator is evaluated against 13 types of simulated generating processes, including random, clustered, dispersed and mixed. Performance is compared with ordered distance estimation of the same rank, and with fixed-area sampling with the same number of trees tallied. The proposed estimator consistently performs better than ordered distance estimation, and nearly as well as fixed-area sampling in all but the most clustered situations. The estimator also provides information regarding the degree of clustering or dispersion.

Keywords: Plotless density estimation, distance sampling, asymptotic estimation, forest inventory, point patterns

Introduction

Estimating the density of a population of trees is a primary goal of forest inventories. Sampling within fixed-area plots is the most common, conceptually simple and statistically robust inventory method. However, the time-intensive nature of inventory work has led many authors to investigate the possibility of using distance or angle measurements rather than counts to estimate density (Cottam and Curtis 1956). Formulas or algorithms to estimate density from distance or angle measurements are often referred to as plotless density estimators (PDEs) because they do not involve any predefined sampling region.

Two advantages of PDEs have been suggested in the literature. First, it has variously been claimed that plotless sampling requires less time or effort than fixed-area sampling (Kenning 2005, Engeman et al. 1994). In quantitative evaluations, however, fixed-area sampling has resulted in lower standard error values than plotless sampling for a given number of trees tallied (Eberhardt 1967) or area searched (Steinke and Hennenberg 2006, Picard et al. 2005).

A second advantage of PDEs is the ability to facilitate inventory planning for a target accuracy level, especially when the nature of the tree pattern is initially unknown (Kleinn and Vilčko 2006). In fixed area sampling, the combination of plot size and plot count required to achieve a given standard error will depend upon actual population density. The latter, however, cannot be known prior to inventory work. Because the number of individuals sampled at each plot is predetermined, PDEs facilitate planning towards a desired accuracy level absent prior knowledge of population density.

Design-unbiased estimation of density from point-to-tree distances is possible if the probabilities of each tree being included in the sample space are known, but this requires detailed knowledge of the spatial arrangement of neighboring trees which is

generally impractical (Kleinn and Vilčko 2006). Instead, most PDEs assume a particular type of generating process, most commonly a Poisson process resulting in complete spatial randomness (CSR). Unfortunately, CSR is rarely encountered in real-world situations. Violation of CSR does not bias fixed-area sampling, although estimation variance is affected. However, violation of CSR does bias the results of PDEs by as much as 100% (Engeman et al. 1994), making their application risky in real-world settings.

One particular type of PDE that has attracted much interest is the ordered distance estimator of Morisita (1957, in Engeman et al. 1994) and examined more fully by Pollard (1971). The estimator of distance rank k is computed from measured distances between each randomly placed point and the k^{th} nearest tree. One attractive feature of this estimator is the simplicity both of the sampling scheme and of computation. Additionally, the choice of distance rank may be adjusted to achieve a desirable balance between estimation accuracy and sampling effort (Nielson et al. 2004).

Clustering and dispersal are scale-dependent phenomena, and so their effects will vary as distance to trees beyond the nearest (i.e. 2nd nearest, 3rd nearest, etc.) are measured. Let λ_k (the “base estimator”) denote any ordered distance estimator of distance rank k . As in all PDEs, λ_k will be biased for clustered and dispersed distributions. However, the causal factors resulting in clustering or dispersion are likely to be limited to a certain range of scales. Therefore, one might expect bias to decrease with increasing distance rank k , converging asymptotically to the true density value. Based on this concept, an asymptotic ordered distance estimator λ_A is proposed, defined as the asymptotic limit of ordered distance estimators as k increases to infinity:

$$\lambda_A = \lim_{k \rightarrow \infty} (\lambda_k) \quad (1)$$

It is hypothesized that this limit will equal the true density value for any base estimator that is unbiased under CSR. No proof is given; instead, the soundness of the hypothesis will be considered through simulation.

To derive the proposed estimator, it is convenient to express λ_k as a function of λ_A . To automate the process, a least-squares solution is applied to an asymptotic model of the form:

$$\lambda_k = f(k; D, A) + \varepsilon \quad (2)$$

where D is the true density, A is a set of one or more additional parameters defining the shape of the curve, and ε indicates error in the model.

The characteristics of the asymptotic estimation procedure will depend greatly on the base estimator λ_k as well as the functional form of (2). For the present study, only the ordered distance estimator of Morisita (1957, in Picard et al. 2005) was considered. For the functional form, a variety of candidate functions were tested, including logarithmic, exponential and power functions on both raw and transformed data. Many functions were found to produce a good model fit, but were unreliable in predicting density from small data sets. Ultimately, a simple function derived from the limiting case of a clustered distribution was found to produce better results than any other tested under a variety of spatial distributions. In its simplest form, the function is:

$$\lambda_k = \frac{k}{k + b} D + \varepsilon \quad (3)$$

The estimators λ_A and β of D and b , respectively, are determined by seeking values that result in the least-squared-error fit of (3) to λ_k computed from observed distance measurements.

The next section examines the properties of bias when ordered distance estimation is applied to dispersed and clustered situations, and derives the asymptotic model function given in (3).

Ordered Distance Estimation Bias in Dispersed and Clustered Situations

A large number of PDEs have been developed over the past six decades; the reader is referred to Picard et al. (2005) for a comprehensive list. For simplicity, the present paper uses as a base estimator only the ordered distance estimator developed by Morisita (1957, in Picard et al. 2005) and examined more fully by Pollard (1971). Let $\overline{r_k^2}$ denote the average squared distance from each random point to the k^{th} nearest tree. Further let λ_k denote the corresponding ordered distance estimator of rank k . Pollard (1971) showed that under CSR,

$$\lambda_k = \frac{nk}{nk - 1} \cdot \frac{k}{\pi \overline{r_k^2}} \quad (4)$$

is an unbiased estimator of density whose variance is

$$\sigma^2(\lambda_k) = \frac{\lambda_k^2}{nk - 2} \quad (5)$$

For convenience, we follow Engeman et al. (1994) in referring to (4) as the ordered distance estimator (ODE) of rank k , although the reader should be aware that other estimators can be derived from ordered distance measurements.

The variance (5) of the ODE can be compared to that associated with a fixed area sample large enough to contain nk trees on average. Again under CSR, the actual number of trees contained in such an area will follow a Poisson distribution with mean and variance of nk . The variance of the density estimate is therefore approximately D^2/nk , and the ratio of variances of density estimates for ODE and fixed area sampling is approximately $nk/(nk-2)$, which is nearly unity for large sample sizes.

Three previous simulation studies (Steinke and Hennenberg 2006, Nielson et al. 2004, Engeman et al. 1994) evaluated the bias characteristics of (4) for various distance ranks as applied to clustered and dispersed point processes. When applied to a moderately dispersed point process, Nielson et al. (2004) reported relative bias (*RBIAS*) monotonically decreasing from ~40% for λ_1 to less than 5% for λ_{10} . In a maximally dispersed distribution (hexagonal lattice), however, Engeman et al. (1994) reported a non-monotonically decreasing trend, with *RBIAS* values of 90%, 16% and 29% for λ_1 , λ_2 and λ_3 respectively. When applied to clustered point processes, maximum negative *RBIAS* values of 75%, 61%, and 48% for λ_1 , λ_2 , and λ_3 , and 31% and 20% for λ_3 and λ_6 were reported by Engeman et al. (1994) and Steinke and Hennenberg (2006) respectively. Nielson et al. (2004) reported a slight departure from monotonicity, however, with negative *RBIAS* generally approaching approaching zero from λ_1 to λ_{10} but dipping slightly from λ_5 to λ_6 for one point process and from λ_9 to λ_{10} for another. Thus, previous studies confirm the generally asymptotic nature of λ_k with increasing k for both clustered

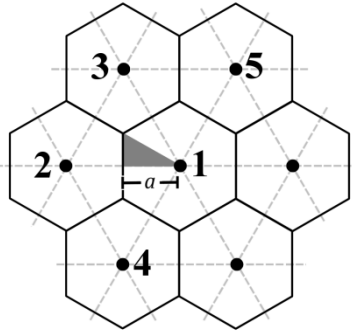


Figure 1:

Maximally dispersed distribution (hexagonal lattice) of trees. Numbers indicate distance ranks k of each tree from any sample location within the shaded triangle.

and dispersed point processes, but slight departures from monotonicity are observed in some cases.

Further insight can be gained from examination of certain extreme point processes which are analytically tractable. Maximum dispersion occurs when trees are located on a hexagonal lattice (Figure 1). On such a lattice, geometric relationships of random sample locations to trees are repeated over triangular regions (shaded triangle in Fig. 1) representing $1/12$ of the hexagonal Thiessen polygon surrounding each tree. Let a represent the length of a perpendicular segment from a tree to the edge of its enclosing hexagon, and A the area of the shaded triangle, such that:

$$a = \frac{1}{\sqrt{2} \cdot \sqrt{3} \cdot \sqrt{D}} \quad , \quad A = \frac{1}{12D}$$

The expectation of λ_1 is the integral of (4) over the shaded triangle, divided by the area of the triangle:

$$E(\lambda_1) = \frac{nk}{nk-1} \cdot \frac{1}{A} \int_0^a \int_0^{\frac{x}{\sqrt{3}}} \frac{1}{\pi(x^2 + y^2)} dy dx \quad (6)$$

Unfortunately, integration at $x=0$ leads to an expectation of infinity, as is the case under CSR as well (Pollard 1971). Therefore the expectation of $1/\lambda_k$ is considered instead. This is the mean area (μ_k) of Cottam and Curtis (1956), representing the average area of a circle centered on the original point and extending to the k^{th} nearest tree. $E(\mu_1)$ is again defined by the double integral, this time omitting the constant $nk/(nk-1)$ for simplicity:

$$E(\mu_1) \cong \frac{1}{A} \int_0^a \int_0^{\frac{x}{\sqrt{3}}} \pi(x^2 + y^2) dy dx \quad (7)$$

Integration yields an expected value of approximately $0.5038/D$, corresponding to $1/E(\mu_1) \cong 1.98D$.

Within each triangle, the nearest five trees are uniquely defined as shown in Figure 1, and the expectations of μ_2 to μ_5 can be calculated in a similar manner. The ratios of $k/E(\mu_k)$ to the true density D are shown in Figure 2 up to $k=5$. These calculations provide further evidence that bias does not decrease monotonically for the extreme case. Similar computation is possible for higher values of k , and for alternative configurations such as a

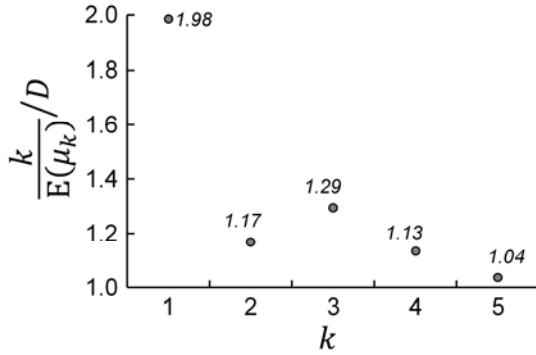


Figure 2:
 Expected value of ranks 1-5 ordered distance estimators when applied to maximally dispersed point process (hexagonal lattice).

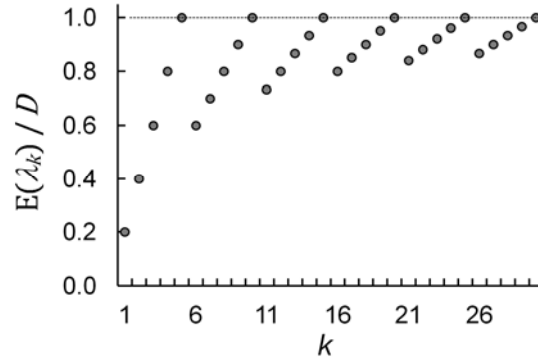


Figure 3:
 Expected value of ranks 1-30 ordered distance estimators when applied to Pielou's (1959) maximally clustered point pattern.

square lattice. However, computation quickly becomes tedious as multiple trees must be considered. Therefore, further exploration of the properties of λ_k in dispersed situations was deferred to the simulation study.

Although clustering may take many forms, Pielou (1959) presented a simple distribution that may be considered maximally clustered. The distribution is a special case of the Matérn process, and is created in two steps. First, cluster centers are generated according to a homogeneous point process with intensity D_p . Second, exactly c trees are located coincidentally at each cluster center location. The density D of the resulting tree pattern is cD_p . However, it is clear that if only the distance to the nearest tree is recorded, the pattern is indistinguishable from a simple homogeneous point process of intensity D_p . Therefore:

$$E(\lambda_1) = D_p = \frac{1}{c} D \quad (8)$$

More generally, let m_k denote the distance rank of the cluster containing the k^{th} tree, such that m_1 to m_c are equal to 1, m_{c+1} to m_{2c} are equal to 2, etc. In other words, the clustering process entails that:

$$m_k = \left\lfloor \frac{k-1}{c} \right\rfloor + 1 \quad (9)$$

where $\lfloor x \rfloor$ is the greatest integer function. Since cluster centers are located according to a Poisson process, the expected average squared distance to the k^{th} tree is the same as that to the m_k^{th} tree for a random point pattern with density $\frac{1}{c} D$. It follows that the expected value of λ_k is:

$$E(\lambda_k) = \frac{k}{m_k} \cdot \frac{1}{c} D \quad (10)$$

Figure 3 illustrates ratios of $E(\lambda_k)$ to D for $c=5$. Two interesting results emerge. First, for all $k \leq c$, $E(\lambda_k)/D$ is a linear function of k :

$$E(\lambda_k) = \frac{D}{c} k \quad (11)$$

Since there exist an infinite number of pairs $[D', c']$ such that $D'/c' = D/c$, it is impossible in this case to determine the true values of c and D from λ_k . In other words, in the

extreme case distance-based density estimation is theoretically impossible unless sampling extends beyond the number of trees contained in a single cluster. This logic would seem to apply to any plotless density estimator.

Fortunately, most clustered ecological distributions depart from the extreme case in that (i) coincident points are rare, and (ii) the number of points per cluster is rarely uniform. These vagaries will tend to smooth out the functional form of $E(\lambda_k)$. Therefore, we sought a monotonically increasing function with asymptote D that approximates (10). If only the i^{th} term in each group of c is considered (i.e. k is restricted to $k = c \left\lfloor \frac{x-i}{c} \right\rfloor + i$), then the points in Figure 3 are connected by the function:

$$E(\lambda_k) = \frac{k}{k + (1 - \alpha)(c - 1)} D \quad (12)$$

where:

$$\alpha = \frac{i - 1}{c - 1}$$

(12) can be reduced to the simpler form in (3) by defining the parameter $b=(1-\alpha)(c-1)$. However, it may be useful to recover the value of c which indicates the number of trees per cluster. Unfortunately, c and α are codependent and cannot be determined independently. As an approximation, we consider the mid-point of each series defined by $\alpha=0.5$, so that $c=2b+1$ and the asymptotic function can be expressed as:

$$E(\lambda_k) = \frac{2k}{2k + c - 1} D \quad (13)$$

In this manner, the parameter c can be considered to be a measure of the average cluster size.

Although (13) is derived from a clustered distribution, it was found to work well in dispersed situations as well. In this case, the parameter c will take on values less than one but greater than zero. No theoretical basis for this relationship is known at the present time.

Methods

To fit (13) to a given set of observed λ_k , we seek λ_A and \hat{c} that minimize:

$$\varepsilon = \sum_{k \in K} \left(\frac{2k}{2k + \hat{c} - 1} \lambda_A - \lambda_k \right)^2 \quad (14)$$

where K represents a set of distance ranks. Because (14) contains only two unknown parameters, a solution can be determined from as few as two distance ranks. If distance ranks k_1 and k_2 are used, an exact solution is given by:

$$\lambda_A = \frac{k_1 - k_2}{\left(\frac{k_1}{\lambda_{k_1}} - \frac{k_2}{\lambda_{k_2}} \right)} \quad (15)$$

$$\hat{c} = 2 \frac{\lambda_{k_1} - \lambda_{k_2}}{\left(\frac{\lambda_{k_2}}{k_2} - \frac{\lambda_{k_1}}{k_1} \right)} + 1 \quad (16)$$

If greater than two distance ranks are used, ε is minimized using the least squares method. Because (13) is non-linear, the Gauss-Newton algorithm was used to search for

Table 1: Summary of pattern simulation processes.

Type	Description	Relevant Parameter Values
A	random	-
B	slightly dispersed	$R=1.25$
C	dispersed	$R=1.5$
D	very dispersed	$R=1.75$
E	square lattice	-
F	hexagonal lattice (maximally dispersed)	-
G	3-tree clustered	$c=3, \alpha = 1$
H	5-tree clustered	$c=5, \alpha = 1$
I	10-tree clustered	$c=10, \alpha = 1$
J	5-tree loosely clustered	$c=5, \alpha = 0.5$
K	5-tree tightly clustered	$c=5, \alpha = 1.5$
L	dispersed clusters	$R=1.5 \mid c=5, \alpha = 1$
M	dispersed with gaps	$c=10, \alpha = 0.6 \mid R=1.25$

an optimal solution. Numerous references on the least-squares method and Gauss-Newman search algorithm exist; we followed Wolberg (2006).

Any search algorithm requires initialization to a set of parameter values, and these initial values may affect the final solution. Solution of (13) using the Gauss-Newman algorithm was found to be very robust when the initial value of D was correctly oriented (i.e. above or below the observed λ_k). However, in some cases the correct orientation was not obvious from the observed λ_k . Therefore, two runs of the algorithm were performed for each set of λ_k , with initial parameters of:

$$\begin{aligned}\lambda_{A_1} &= 2 \times \min(\lambda_k) - \max(\lambda_k) \\ \lambda_{A_2} &= 2 \times \max(\lambda_k) - \min(\lambda_k) \\ \hat{c}_1 &= \hat{c}_2 = 5\end{aligned}$$

The solution that yielded the lowest sum of squared errors was then selected.

To evaluate the relative bias and variance characteristics of the proposed estimator, 13 generating processes representing a wide range of tree patterns were developed, including one random, five dispersed, five clustered and two mixed processes. Table 1 provides a summary of these generating processes.

Type A was a random Poisson process. The actual number of trees in a given plot was selected at random from a Poisson distribution with mean D , and individual trees were placed by choosing random x - and y - coordinates.

Types B-F were dispersed processes. Types B-D were created via a repulsion process developed by the author. First, trees were initialized using a simple Poisson process as in Type A. Next, nearest-neighbor pairs were determined, and the closest pairs were incrementally pushed apart until the overall value of the nearest neighbor statistic (R) of Clark and Evans (1954) was within 0.02 of a target value. To avoid edge effects, nearest neighbors were determined on a torus; if a tree was pushed off the edge of the unit square during the repulsion process, it was placed back onto the other side. R was computed from non-toroidal distances using the correction factor of Donnelly (1978). Benchmark values of this statistic are $R=0$ (maximally clustered), $R=1$ (Poisson random), and $R=2.13$ (maximally dispersed). Target values of $R=1.25$, $R=1.5$ and $R=1.75$ were used to generate pattern types B, C and D respectively. An index of nearest neighbors was continuously updated to facilitate simulation efficiency; nevertheless,

$R=1.75$ was the maximum degree of dispersion that could be consistently achieved. To simulate extremely dispersed situations, pattern types E and F were defined as a square lattice ($R=2.00$) and a hexagonal lattice ($R=2.15$).

Types G-K were clustered patterns created via a Poisson clustering (or Matérn) process (Picard et al. 2005, Bailey and Gatrell 1995). Because density D was separately defined, we chose to model this process using two parameters instead of the usual three: the average number of trees per cluster (c) and the density-independent spread of trees around each cluster center (σ). Parent locations were generated via a simple Poisson process at a density of D/c . A number of trees selected randomly from a Poisson distribution with mean c were then placed around each parent location using the bivariate normal distribution with standard deviation $\sigma\sqrt{1/\pi D}$. Pattern types were organized into two groups to separate the effects of cluster size and spread of trees within a cluster. To observe the effect of cluster size, pattern types G, H and I shared the parameter value $\sigma=1$, with c taking values of 3, 5 and 10 respectively. To observe the effect of relative distance between trees in a cluster, pattern types J, H and K shared the parameter value $c=5$, with σ taking values of 0.5, 1 and 1.5 respectively.

By combining the repulsion and clustering processes outlined above, it is possible to generate additional types of patterns that are realistic but not often represented in simulation studies. Two such pattern types were defined. Type L simulated dispersed clusters, characterized by clustering at short distances and dispersion at longer distances. The repulsion process with target $R=1.5$ was used to generate parent locations at a density of D/c . Trees in a cluster were then generated using the Poisson clustering process defined above at $c=5$ and $\sigma=1$. Type M simulated a dispersed pattern with gaps, characterized by dispersion at short distances and clustering at longer distances. First, a seed pattern was generated via the clustering process at $c=10$ and $\sigma=0.6$ to create a patchy landscape with clumps of trees. Trees were then dispersed via the repulsion process using a target of $R=1.25$, allowing the trees expand to take up much but not all of the simulation space.

Each of the above processes was used to generate 120,000 simulated plots. This number was chosen to balance the need for a large sample size with the computation time required to produce simulations. For all processes, trees were placed on a unit grid at a mean density of $D=200$. Distances to the nearest 50 trees from the grid center were recorded in a master database (available on request). Distances from all 120,000 simulated plots were used to calculate typical values of λ_1 to λ_{50} for each generating process. The best fit of (13) was determined, and the parameters λ_A (predicted density) and \hat{c} (predicted trees per cluster) were recorded. To provide an indication of the validity of the functional form of the asymptotic model, the error in the model fit was measured as:

$$\varepsilon = \sum_{k=1}^{50} \sqrt{\left(\lambda_k - \frac{2k}{2k + \hat{c} - 1} \lambda_A\right)^2} / 50 \quad (17)$$

and was computed separately for each generating process.

To test effectiveness under normal sampling scenarios, the bias and variance of the proposed estimator were calculated for 35 sampling schemes which varied in the number of plots ($n=10, 20, 30, 40, 50, 60, 100$) and number of trees sampled per plot ($\kappa=2, 3, 4, 5, 6$). A given trial group consisted of 120,000/ n trials of n plots each. For

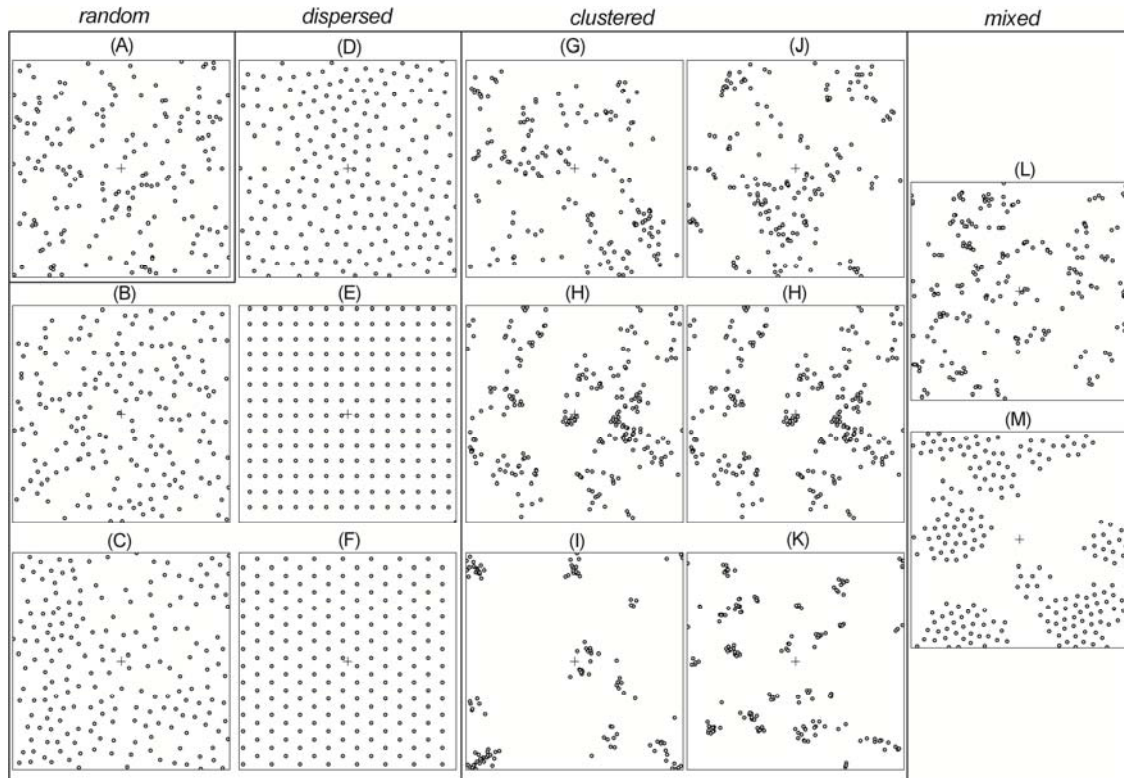


Figure 4:
 Samples of tree patterns simulated by each generating process.

each trial t of a given pattern type, the base estimators λ_1 thru λ_k were calculated from (4), and the Gauss-Newton algorithm was used to determine the parameters λ_A and \hat{c} . The relative bias, standard deviation and root mean squared error for each trial group were calculated as:

$$RBIAS = \frac{|\bar{\lambda}_A - 200|}{200} \quad (18)$$

$$RSD = \frac{\sqrt{\sum(\lambda_A - \bar{\lambda}_A)^2}}{200} \quad (19)$$

$$RRMSE = \sqrt{RBIAS^2 + RSD^2} \quad (20)$$

The mean and standard deviation of \hat{c} were also recorded.

Although typical bias and variance characteristics of other estimators can be acquired from the literature, comparison is difficult because generating processes and sampling schemes vary from author to author. Therefore, we compared the properties of the proposed asymptotic estimator with (i) the ordered distance estimator of the same rank, and (ii) an equivalent fixed-area sampling scheme. The ODE was chosen because it was one of three recommended by Engeman et al. (1994), and because it uses a similar sampling scheme to the proposed estimator. Fixed-area sampling was chosen because no PDE has been shown to result in a lower RRMSE for a given search effort. For a given sampling scheme, the same number of plots was used for the asymptotic estimator, ODE and fixed area estimate. Plot size for fixed-area sampling was determined by calculating

Table 2:

Average value of the nearest neighbor statistic (R) for each simulated point process.

Process	Mean (R)
A	1.00
B	1.26
C	1.51
D	1.74
E	1.94
F	2.09
G	0.71
H	0.61
I	0.48
J	0.75
K	0.36
L	0.61
M	1.25

Table 3:

Predicted density (λ_A), number of trees per cluster (\hat{c}) and error of the model fit (ε) for asymptotic estimation applied to 50 rank distances for 120,000 plots.

	Process	λ_A	\hat{c}	ε	
random	A	200.1	1.0	0.00	
	dispersed	B	198.9	0.7	0.00
		C	197.7	0.4	0.01
		D	197.7	0.2	0.01
		E	199.9	0.0	0.02
		F	199.8	0.0	0.03
clustered	G	197.5	3.2	0.01	
	H	197.5	5.0	0.00	
	I	199.0	10.1	0.01	
	J	195.1	4.1	0.01	
	K	200.4	5.8	0.01	
	mixed	L	196.1	2.9	0.01
M		199.3	5.1	0.01	

the area required to include, on average, a number of trees equal to the distance rank of a given sampling scheme. The master database described above was used to determine density based on distances (for the ODE) or counts (for fixed area sampling). RBIAS, RSD and RRMSE were computed for each sampling scheme using (18), (19) and (20), and these were compared with the corresponding values of the asymptotic estimator.

Results

A typical example of tree patterns resulting from the 13 generating processes is shown in Figure 4 (note that pattern type H is shown twice in the figure). Due to random chance, a total of 5 plots (4 of type I, 1 of type A) contained less than 50 trees. This occurred most frequently for type I due to the fact that the number of parent clusters was lowest for this generating process. Every simulated plot contained at least 30 trees. The average computed value of the nearest neighbor statistic R (Clark and Evans 1954) is provided in Table 2 for reference. Note that the computed values of R for the lattice patterns do not equate with the theoretical values due to application of the edge correction factor of Donnelly (1978).

Trends in λ_k calculated using the full set of 120,000 simulated plots are shown in Figure 5 for each generating process. In all processes the ODE values conformed generally to an asymptotic model with convergence toward the true density value of 200. However, the hexagonal and square lattices exhibited significant non-monotonicity. Only values up to $k=20$ are shown in Figure 5, but this non-monotonicity continued up to λ_{50} , the maximum distance rank measured. The dispersed processes converged more quickly to the true density value, with $\lambda_{50}=201.3$ for the maximally dispersed process (F) as compared to $\lambda_{50}=182.6$ for the 10-tree clustered process (I).

The results of fitting (13) to the base estimators calculated from all plots for ranks 1-50 are shown in Table 3. Predicted density (λ_A) ranged from 195.1 (J) to 200.4 (K), showing a tendency towards underestimation rather than overestimation for both clustered and dispersed processes. The predicted number of trees per cluster (\hat{c}) closely matched the number of trees per cluster for the 3-, 5- and 10-tree cluster generating

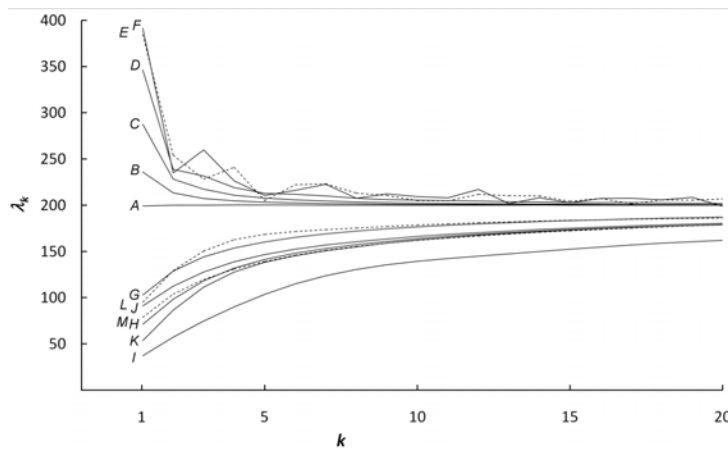


Figure 5: Trend in ordered distance estimates (λ_k) with distance rank (k) using all simulated plot data for thirteen generating processes, up to $k=20$. Dashed lines for processes E , L and M are used for visual separation.

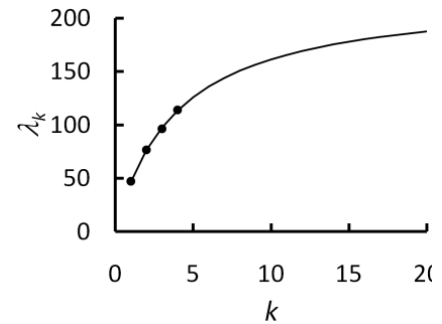


Figure 6: Sample fit of asymptotic curve to ordered distances estimators of ranks 1-4. ODEs are taken from first 50 simulated plots of process type I. The asymptotic limit of 224.0 is 12% higher than the true density.

processes $G-I$. However, this value was also affected by the spread of trees around the cluster center, varying from 4.1 for the 5-tree loosely clustered process (J) to 5.8 for the 5-tree tightly clustered process (K). The values of \hat{c} for the mixed processes were somewhat lower than the number of trees per cluster used in the component cluster-generating process, as might be expected. Among dispersed processes, \hat{c} decreased from 1.0 for the random process (A) to 0.0 for the maximally dispersed process (F). Error in the model fit (ε) was generally low, and was highest for the lattice processes (E and F) due to non-monotonicity.

Figure 6 shows the best-fit curve for the first 50-plot, 4-tree trial of process type I (10 trees per cluster) to illustrate the curve-fitting procedure. Ordered distance estimates for this trial were $\lambda_1=45.9$, $\lambda_2=76.5$, $\lambda_3=96.3$ and $\lambda_4=114.2$. The best-fit asymptotic function had parameters of $\lambda_A=224.0$ and $\hat{c}=8.8$. Mean and standard deviation of parameter values for the trial group were $\lambda_A = 201.6 (\pm 27.5)$ and $\hat{c}=10.9 (\pm 2.3)$.

Across all trial groups, RSD tended to decrease with both the number of plots per trial and the number of trees per plot, whereas $RBIAS$ decreased with the number of trees per plot but remained fairly consistent with respect to the number of plots per trial (data not shown). For simplicity and to allow comparison with the results of Picard et al. (2005), we focus on the case of 50 plots with four trees sampled per plot. Figure 7 shows $RBIAS$ and RSD of the asymptotic estimator for the 13 point processes investigated. $RRMSE$ can be interpreted from the diagram as the Euclidean distance from the origin to each point. Overall, the estimator is characterized by very low bias ($RBIAS < 3\%$) and low variability ($RSD < 10\%$) for random and dispersed distributions. For clustered distributions, bias and variability were somewhat higher, but usually under 15%. The estimator performed worst on the tight 5-tree clustering process (K) ($RBIAS=27\%$, $RSD=25\%$).

When compared to the equivalent ordered distance estimator, the asymptotic estimator resulted in lower $RRMSE$ for all point processes except the random (A), slightly dispersed (B) and 5-tree tightly clustered (K) processes (Figure 8). $RRMSE$ for

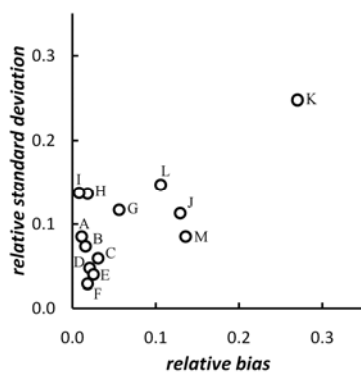


Figure 7: Relative bias and standard deviation of asymptotic estimator for the 4-tree, 50-plot sampling scheme.

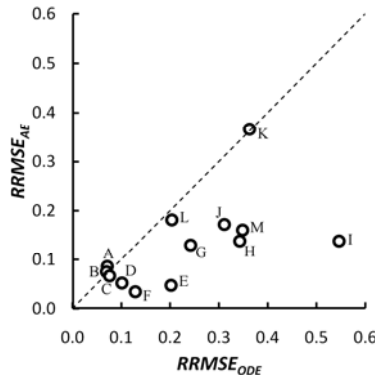


Figure 8: Comparison of relative root mean square errors of the asymptotic estimator vs. ordered distance estimator for 4-tree, 50-plot sampling scheme.

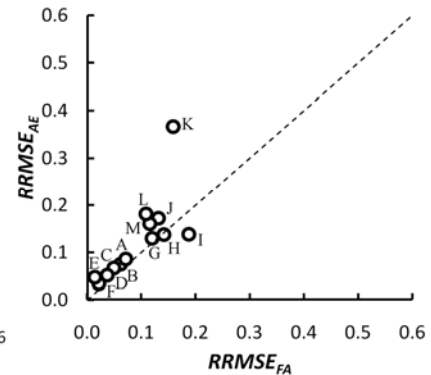


Figure 9: Comparison of relative root mean square errors of the asymptotic estimator vs. fixed-area sampling for the 4-tree, 50-plot sampling scheme.

the asymptotic estimator was less than half that of the ODE for five of the 13 generating processes. The improvement over ordered distance estimation was primarily due to reduced bias; variability of the two estimators was approximately the same (data not shown).

In comparison to fixed-area sampling (Figure 9), the asymptotic estimator resulted in slightly lower RRMSE for the 5-tree and 10-tree clustered processes (*H* and *I*). Fixed-area sampling resulted in lower RRMSE for the remaining processes. Overall, however, the differences between the asymptotic and fixed area estimators were not as substantial as the differences between the asymptotic estimator and the ODE. RRMSE for the fixed area estimator was less than half that of the asymptotic estimator for the square lattice (*E*) and 5-tree tightly clustered processes (*K*).

The average number of trees per cluster predicted in trials (Table 4) was similar to that predicted from the full 120,000 simulated plots. However, there was greater variability for the 10-tree clustered process (*I*) and the 5-tree tightly clustered process (*K*), contributing to the variance in estimated density for these processes.

The above results summarize the 50-plot, 4-tree sampling scheme. Performance of the asymptotic estimator in comparison to fixed-area sampling generally improved as the number of trees per plot increased, but not as the number of plots increased. Figure 10 shows the ratio of RRMSE for the asymptotic estimator vs. fixed-area sampling averaged across all 13 generating processes for all sampling schemes tested. For a given number of trees per plot, the asymptotic estimator performed well relative to fixed area sampling when 20-40 plots were sampled, with the best relative performance for the 6-tree, 40-plot sampling scheme.

RRMSE of the asymptotic estimator was actually slightly lower than that of fixed-area sampling for 52 of the 455 combinations of generating process \times sampling scheme, including the square (*E*) and hexagonal (*F*) lattices for distance ranks 5-6 for all numbers of plots, and the dispersed-with-holes process (*M*) for distance ranks 4-6 when the number of plots was 30 or less (results not shown). At the other extreme, the 2-tree, 10-plot sampling scheme resulted in a single trial with an asymptotic density estimate of

Table 4:

Predicted number of trees per cluster under the 50-plot, 4-tree sampling scheme for each generating process.

Process Type	Number of Clusters (c)	
	mean	st. dev.
A	1.1	0.34
B	0.7	0.28
C	0.4	0.21
D	0.2	0.14
E	0.1	0.12
F	0.1	0.09
G	2.8	0.78
H	4.9	1.28
I	10.9	2.27
J	3.1	0.85
K	8.7	2.48
L	3.8	1.05
M	3.6	0.90

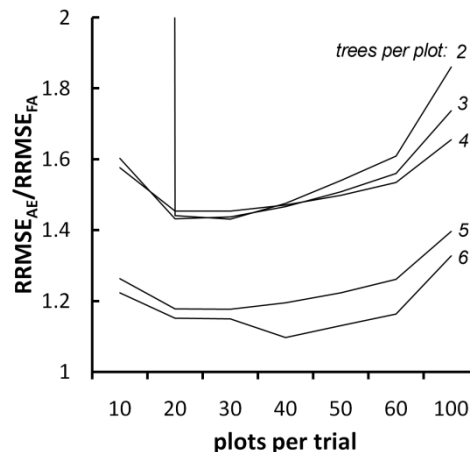


Figure 10:
 Average RRMSE ratio of asymptotic estimator vs. fixed-area sampling for various sampling schemes.

nearly one billion for both process types *I* and *K*. This occurred when the first and second trees were virtually the same distance from grid center in all 10 plots.

Discussion

The asymptotic ordered distance estimator performed well in simulated trials when applied to tree patterns created from dispersed, clustered and mixed point-generating processes, and was consistently as good as or better than ordered distance estimation of the same rank. However, it should be taken under consideration that the ODE requires only one distance measurement per random point, whereas the proposed estimator requires a number of measurements equal to the distance rank.

Although other estimators were not evaluated, we note that the ODE of rank 3 was one of three recommended by Engeman et al. (1994). Also, visual comparison of the results in Figure 7 to those of Picard et al (2005, Figure 3) suggests that the asymptotic estimator may perform approximately on par with the two estimators which performed best in that study (the Kendall-Moran estimator and Picard et al.'s own estimator) in clustered situations. In dispersed situations, however, the asymptotic estimator also performs well, while the Kendall-Moran estimator has been shown to perform poorly (Engeman et al. 1994) and Picard et al's (2005) estimator does not seem to be applicable. These observations are based on visual observation only and should be treated with caution. No attempt was made in this study to provide a more rigorous quantitative comparison of λ_A with other plotless density estimators except the ODE.

As with all PDEs evaluated in previous studies, the bias and variance characteristics of the proposed asymptotic estimator are generally worse than fixed-area sampling when the number of trees tallied is held constant. However, the differences are minor for all but the most clustered distributions tested, and indeed the proposed estimator performed better than fixed-area sampling in some situations.

Given this performance, asymptotic estimation may be considered a reasonable option when little is known about the tree distribution prior to sampling. A good strategy appears to be to sample five or more trees at each of between 20 to 40 plots (Figure 10). If more effort is available, this will be best spent sampling more trees per plot, rather than sampling more plots. This result is similar to what Nielson et al. (2004) found with regards to ordered distance estimation. If sampling of more plots is desirable, then fixed-area sampling is probably warranted. The advantage of decreased variance with large sample size does not mitigate bias, and so a perfectly unbiased estimator will always perform comparatively better as sample size increases. Furthermore, the area covered by larger inventories is unlikely to be uniform with respect to tree pattern, which will have unknown effects on the bias and variance of the proposed estimator.

The parameter \hat{c} determined during the curve-fitting procedure also provides additional information on tree pattern. Although this parameter was initially defined as representing the number of trees per cluster, it is best interpreted more generally as indicating the degree of clustering or dispersion. Indeed, the number of trees per cluster had less effect on the accuracy of the estimator than did the tightness of clustering. This is due to the fact that the number of trees per cluster affected the steepness more than the shape of the curve through the ordered distance estimates (process types *G, H, I* in Figure 5), whereas the tightness of clustering affected the shape of the curve more than the steepness (process types *J, H, K* in Figure 5).

Since the functional form of the asymptotic model was derived from an extreme case of a clustered distribution, the fact that it performed well in dispersed situations compared to fixed-area sampling is somewhat surprising. It is possible that some unifying process may underlie both dispersed and clustered distributions. The relative success of the estimator seems to hinge not only on the quality of fit of the asymptotic model function (13), but also on the limitations of that function. In searching for candidate functions, this author's first instinct was to identify functions with several parameters that would provide maximum shape flexibility. Because the curve-fitting procedure requires extrapolation of the asymptotic limit from a small number of sample points, however, it quickly became obvious that such shape flexibility would allow wildly errant density estimates in certain situations. The proposed asymptotic model function (13) has only two parameters and enforces severe constraints on the shape of the fitted curve.

It is interesting to consider similarities between the estimator presented here and that of Picard et al. (2005). The latter differed from previous PDEs in that it utilized the entire distribution of distance values at a given rank, rather than a single summary statistic. The asymptotic estimator uses only a single summary statistic at each rank, but examines the distribution of these statistics across multiple distance ranks. The relatively good performance of both estimators is promising, and demonstrates the benefit of analyzing full distributions rather than summary statistics.

Observations made during an inventory can improve understanding of tree pattern and inform estimation accuracy. For example, in field situations it may be noted whether or not distinct clusters are discernable, and if so, whether the number of trees per cluster exceeds the sampled distance rank. The quality of fit of the asymptotic function to the ordered distance estimates may also be used to evaluate the accuracy of the estimator. Large errors were usually the result of over-extrapolation in extremely clustered

situations. To account for this, it would be prudent to consider the ratio of \hat{c} to the highest distance rank measured: a ratio greater than one is an indication of uncertainty. In the extreme case, when distance measurements to all observed ranks are essentially identical at every plot, both \hat{c} and λ_A can take on unreasonably high values. This happened twice in simulation for the smallest sampling scheme applied to the most extreme clustering processes (I and K). Such situations should be easy to spot if the data are examined carefully.

The asymptotic estimation procedure is based on the premise that patterns of clustering or dispersal are the result of processes that operate on a limited scale. In reality, every point distribution is the result of a variety of processes operating at multiple scales. The proposed estimator does not account for broad-scale differences in pattern or non-uniformity in the underlying generating process.

Conclusion

The proposed asymptotic estimator presents a viable alternative to fixed-area sampling. This and other PDEs should be avoided when the inventory region is large and/or the underlying pattern is thought to be heterogeneous. The estimator is appropriate, however, for quick preliminary inventory over a homogenous region, especially when the appropriate plot size for fixed-area sampling is difficult to know in advance. The asymptotic estimation procedure is easy to understand and implement, and automatically provides additional information on the structure of the sampled tree pattern through estimation of the clustering parameter c . Although Gauss-Newton optimization was used here to automate curve-fitting over numerous simulated trials, the best-fit curve can be found for a single inventory through trial and error using any spreadsheet program.

References:

Bailey, TC and Gatrell, AC. 1995. Interactive Spatial Data Analysis. London: Longman Group Limited.

Clark, PJ and Evans, FC. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* **35**: 445-453.

Cottam, G and Curtis, JT. 1956. The use of distance measures in phytosociological sampling. *Ecology* **37**:451-460.

Donnelly, K. 1978. Simulations to determine the variance and edge-effect of total nearest neighbor distance. *In Simulation Methods in Archaeology. Edited by I. Hodder.* Cambridge University Press, London. pp. 91-95.

Eberhardt, LL. 1967. Some developments in 'distance sampling'. *Biometrics*, **23(2)**:207-216.

Engeman, RM, Sugihara, RT, Pank, LF and Dusenberry, WE. 1994. A comparison of plotless density estimators using Monte Carlo simulation. *Ecology* **75**: 1769-1779.

Kenning, RS, Ducey, MJ, Brissette, JC and Gove, JH. 2005. Field efficiency and bias of snag inventory methods. *Canadian Journal of Forest Research* **35**: 2900-2910.

Kleinn, C and Vilčko, F. 2006. Design-unbiased estimation for point-to-tree distance sampling. *Canadian Journal of Forest Research*. **36**:1407-1414.

Nielson, RM, Sugihara, RT, Boardman, TJ and Engeman, RM. 2004. Optimization of ordered distance sampling. *Environmetrics* **15**: 119-128.

Picard, N, Kouyaté, AM and Dessard, H. 2005. Tree density estimations using a distance method in Mali savanna. *Forest Science*, **51(1)**: 7-18.

Pielou, EC. 1959. The use of point to plant distances in the study of the pattern of plant populations. *Journal of Ecology*, **47**: 607-613.

Pollard, JH. 1971. On distance estimators of density in randomly distributed forests. *Biometrics*, **27**: 991-1002.

Steinke, I, and Hennenberg, KJ. 2006. On the power of plotless density estimators for statistical comparisons of plant populations. *Canadian Journal of Botany*, **84**: 421-432.

Wolberg, J. 2006. Data analysis using the method of least squares: Extracting the most information from experiments. Berlin: Springer.