Eur J Forest Res (2008) 127:213–224
DOI 10.1007/s10342-007-0197-z

ORIGINAL PAPER

Two new density estimators for distance sampling

S. Magnussen · C. Kleinn · N. Picard

Received: 13 December 2006 / Accepted: 30 October 2007 / Published online: 5 December 2007

Abstract Two new density estimators for *k*-tree distance sampling are proposed and their performance is assessed in simulated distance sampling from 22 stem maps representing a wide range of natural to semi-natural forest tree stands with random to irregular (clustered) spatial distribution of trees. The new estimators are model-based. The first (Orbit) computes density as the inverse of the average of the areas associated with each of the *k*-trees nearest to a sample location. The area of the *k*-th tree is obtained as a prediction from a linear regression model while the area of the first is obtained via a Poisson probability integral. The second (GamPoi) is based on the expected distribution of distance to the *k* nearest tree in a forest where the local distribution of trees is random but the stem density varies from sample location to sample location as a gamma distribution. In a comprehensive assessment with 17 promising reference estimators, a subset composed of Morisita's, Persson's, Byth's, Kleinn's, Orbit, and GamPoi was significantly better, in terms of relative root mean square error (RRMSE), than average. GamPoi emerged as the better estimator for sample sizes larger than or equal to 30. For smaller sample sizes, both Kleinn's and Morisita's appear attractive.

Communicated by M. Köhl.

S. Magnussen (✉)
Natural Resources Canada, Canadian Forest Service, 506 West Burnside Road, Victoria, BC, Canada V8Z 1M5
e-mail: steen.magnussen@nrcan.gc.ca

C. Kleinn
Institute of Forest Management, Georg-August-Universität Göttingen, Büsgenweg 5, 37077 Gottingen, Germany

N. Picard
CIRAD, BP 4035, Libreville, Gabon

Keywords Gamma × Poisson distribution · Poisson distribution · Spatial point pattern · Robust estimation · Root mean square error · Bias · Precision · Sample size · *k*-tree sampling · Fixed count sampling

Introduction

Distance sampling, also known as plotless sampling, has a long history in forest inventory (see Kleinn and Vilčko 2006a for a historic overview). In forest inventory applications the distance and attributes of interest are measured on the *k*-trees closest to a sample location (Prodan 1968; Persson 1964). Distance sampling has been favoured when plot (quadrat) sampling would be difficult or too costly (Picard et al. 2005; Sheil et al. 2003; Lessard et al. 2002; Hall 1991). Distance sampling is often faster than fixed-area plot sampling with an average of *k*-trees per plot since the number of measurements to take at each plot location is fixed in advance: measurement efforts are independent of local tree density. Expediency in stand-wise forest-management inventories has been another motivating factor (Lynch and Rusydi 1999; Lessard et al. 1994; Jonsson et al. 1992; Prodan 1968). In sampling for indices of forest structure, distance sampling can be very efficient (Kint et al. 2004; Pretzsch 1997). Continued technological advances in range measurements and field computing add to the attraction of distance sampling because measurements can be completed ever faster. Also, in natural populations characterized by an irregular, possibly clustered, distribution of trees, the precision of a density estimate from distance sampling can be better than the precision obtained with fixed area plot sampling (Lessard et al. 2002).

Bias of estimates obtained from distance-sampling, and a sensitivity of estimators to the spatial distribution of trees are recognized impediments to wider use (Schreuder 2004; Engeman et al. 1994; Byth 1982; Persson 1971). Bias stems from the problem of obtaining an unbiased inclusion area for the selected trees. Although one can define a design-unbiased inclusion area for k-tree distance sampling (Kleinn and Vilčko 2006b), it would require the identification of, and the measuring of distances to, a large number of additional trees, possibly as many as $2k^2$ or even $2k^3$. Such a burden eliminates any practical advantage over sampling with fixed area plots.

A large number of improved density estimators have been proposed to combat the recognized problems of bias and lack of robustness (Engeman et al. 1994; Patil et al. 1982; Byth 1982; Diggle 1977; Morisita 1957). Adjusting the effective number of trees enclosed in a circle with radius equal to the distance to the k (or $k + 1$) nearest tree or the radius of this circle is an intuitive device for bias-reduction (Kleinn et al. 2006a; Delince 1986; Prodan 1968; Eberhardt 1967; Morisita 1957). Robust estimators often originate in the observation that estimates derived from point-to-tree distances and those from tree-to-tree distances are biased in opposite directions (Clayton and Cox 1986; Byth 1982). By measuring the two types of distances at a single sample location the expectation is a more robust and, ideally, also a less biased estimator. Alternatively, an estimator may be tailored to a specific point pattern such as, those generated by a Matérn process (Picard et al. 2005) or a regular tree spacing (Persson 1964).

Notwithstanding tangible progress and clear successes of this or that estimator in a particular assessment, it is nevertheless clear that we do not yet have a uniformly best estimator suitable for practical application. In this study we propose two new model-based density estimators for distance sampling. One is tailored towards application in natural- or semi natural tree populations with an irregular distribution of trees and pronounced tendency towards clustering. The second is designed for populations with a Poisson distribution of tree locations (complete spatial randomness) and with robust properties in mind—robust in the sense that it is not sensitive to departures from complete spatial randomness.

The performance of the two new estimators is assessed in simulated distance sampling from 12 actual and 10 simulated forest tree stem maps representing 22 actual sites. Seventeen estimators, including a density estimator for fixed area plot sampling, act as a reference set for ranking and recommendations. Reference estimators are a mixture of parametric, nonparametric, and so-called robust estimators. Published records led us to expect that these estimators would perform well in natural and semi-natural populations. Relative root mean square error (RRMSE) of a density estimate is the primary criterion of performance since the above problem of bias and robustness originates in the estimate of density. A bias in density will also propagate to other tree attributes when expressed on a per unit area basis.

Materials and methods

Variables and symbols are defined in Appendix A. Random variables are set in upper case and observed values of the random variable is set in lower case.

Two new density estimators

The first new density estimator attempts status as a robust estimator by using all k distances and not just the distance to the k nearest tree r_k. This idea stems from the observation that a single location density estimated by $k \times \pi^{-1} \times r_k^{-2}$ often fluctuates greatly with k. If all k distances were combined to produce a single estimate, then one might expect a dampening effect, and, hopefully, a more robust estimator.

We propose to allocate to each tree $(1,\ldots,k)$ an area A_j, $j = 1,\ldots,k$ and then calculate the density as k divided by the sum of these k areas. Assigned areas are defined to give an approximately unbiased estimator under an assumed random (Poisson) point pattern. The area we allocate to the j nearest tree is that of an annulus with an inner radius of $\sqrt{r_{j-1} \times r_j}$ and an outer radius of $\sqrt{r_j \times r_{j+1}}$ i.e. an area equal to $\pi \times r_j\,(r_{j+1} - r_{j-1})$. An example is in Fig. 1 for $k = 6$ where the six annuli are numbered and identified by varying grey tones. Let A_j denote the (random) area of the annulus containing the j-th tree. For a single sample location the estimator is

$$\Lambda_{\mathrm{ORBIT}} = k \times \left(\sum_{j=1}^{k} A_j\right)^{-1} = \frac{k}{\pi(R_k R_{k+1} - R_0 R_1)} \qquad (1)$$

The name "orbit" is coined as a figurative mnemonic inspired by viewing the k-trees in Fig. 1 as occupying k concentric orbits. There is no observation of R_0 and R_{k+1}; they must be predicted. The inner radius R_0 is taken as the radius of a disk cantered at the sample location and for which the probability is 0.5 that it does not contain a randomly located tree in a co-located disk with radius R_1. This definition gives $R_0 \equiv \sqrt{\log(2)}R_1$ and thus $\hat{A}_1 = \pi \times R_1\left(R_2 - \sqrt{\log(2)}R_1\right)$. A prediction of A_k and hence R_{k+1} is obtained from a linear regression:

$$\hat{A}_k = k \times \hat{\beta}, \hat{\beta} = \frac{\sum_{j=1}^{k-1} j \times A_j}{\sum_{j=1}^{k-1} j^2} \qquad (2)$$

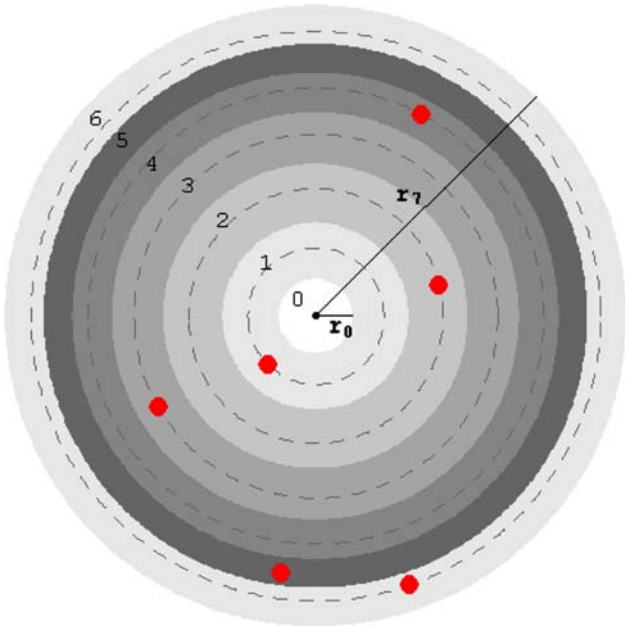

Fig. 1 Geometric principle of annuli areas assigned each of k-trees nearest to a sample location (center black dot). Six nearest trees are indicated by red dot. Annuli are grey-shaded and numbered 1,…, 6. Circles with radii = r_k, k = 1,…, 6 are indicated by dashed lines

where $\hat{\beta}$ is the estimate of the trend in the linear regression of A_j on j through origo.

We expect Λ_{ORBIT} to have robust properties because all k distances are used in the estimator. In order to see this, we express Λ_{ORBIT} as a function of the k distances:

$$\Lambda_{\text{ORBIT}} = \frac{28}{\pi\left(R_3(13R_4 - 2R_2) - 2R_1R_2 - 9\sqrt{\log(2)}R_1^2\right)}, \quad k = 4 \tag{3}$$

and

$$\Lambda_{\text{ORBIT}} = \frac{330}{\pi\left(85R_5R_6 - 6(R_4R_5 + R_3R_4 + R_2R_3 + R_1R_2) - 61\sqrt{\log(2)}R_1^2\right)}, \quad k = 6 \tag{4}$$

The expected value of Λ_{ORBIT} given a Poisson distribution of trees with density λ can be computed from the expected area of annuli 2,…, $k-1$, and R_1^2. The latter is $\pi \times \lambda^{-1}$ while the expected area of an annulus is

$$E\left(A_j|\text{Poisson}(\lambda)\right) = \frac{\Gamma(j+0.5)(\Gamma(j-1)\Gamma(j+1.5) - j!\Gamma(j-0.5))}{\lambda \times \Gamma(j-1)\Gamma(j)\Gamma(j+1)} \tag{5}$$

Approximate values of the first five annuli (j = 2,…, 6) are $1.03084\lambda^{-1}$, $1.01243\lambda^{-1}$, $1.00667\lambda^{-1}$, $1.00416\lambda^{-1}$, and $1.00284\lambda^{-1}$. From these expectations and the assumption that the expected value of an inverse is approximately equal to the inverse of the expected value we obtain $E\,(\Lambda_{\text{ORBIT}}|k = 4, \text{Poisson}\ (\lambda)) \approx 1.0141\lambda$ and E $\Lambda_{\text{ORBIT}}|k = 6, \text{Poisson}\ (\lambda)) \approx 1.0033\lambda$. That is, Λ_{ORBIT} is asymptotically $(n \to \infty)$ approximately unbiased for $k = 4$ and $k = 6$ when trees are located completely at random and the assumptions are correct. Λ_{ORBIT} is a random variable of density at a single sample location and will often fluctuate erratically across sample locations. With n sample locations (i = 1,…, n), a ratio of means estimator (Schreuder et al. 1993, p. 85) will have less bias and a lower sampling variance than $E(\Lambda_{\text{ORBIT}})$. Hence we opt for the following sample-based estimate of density

$$\hat{\lambda}_{\text{ORBIT}} = \frac{n \times k}{\pi \sum_{i=1}^{n}\left(r_{ik}\hat{r}_{ik+1} - \sqrt{\log(2)}r_{i1}^2\right)} \tag{6}$$

A variance estimate for $\hat{\lambda}_{\text{ORBIT}}$ is obtained by application of the delta technique (Kendall and Stuart 1969).

Our second estimator is also model-based. In natural or semi-natural populations the spatial distribution of trees is often irregular. Sampling with fixed-area plots in these populations often gives rise to a frequency distribution of trees per plot that resembles a negative binomial distribution (Wolf 2005; Druckenbrod et al. 2005; Turechek and Madden 1999; White and Bennetts 1996; Upton and Fingleton 1985; Diggle 1983; Eberhardt 1967). A negative binomial distribution of counts implies that the local density λ_i of trees varies from one sample location to another as in a gamma distribution with a mean of λ (Johnson et al. 1992, p. 408). Stated otherwise, the spatial point pattern arises from a gamma-mixed Poisson process. The mixing generates an over-dispersion (extra Poisson variance) of tree counts in fixed-area plot sampling. Under a gamma × Poisson model the probability density function (pdf) of the distance r_k from a (random) sample location to the k nearest tree is

$$f(r_k|\alpha,\beta) = 2\pi^k r_k^{2k-1} \Gamma(k+\alpha)\Gamma^{-1}(\alpha)\Gamma^{-1}(k)\beta^{-\alpha}\left(\beta^{-1} + \pi r_k^2\right)^{-(k+\alpha)} \tag{7}$$

where α and β are the two gamma parameters with $\lambda = \alpha \times \beta$. Equation 7 is obtained from a convolution of the pdf of r_k under a Poisson model (Thompson 1956) with a gamma distribution. Estimates of α and β and hence $\hat{\lambda}_{\text{GAMPOI}} = \hat{\alpha} \times \hat{\beta}$ are ideally obtained by maximizing the

log-likelihood of sample data r_{ik}, $i = 1,\ldots, n$ or, in case of unstable estimates, by the method of moments whereby the observed mean and variance of r_{ik} is equated to their expectations under a gamma × Poisson model. Estimation by the method of moments can be an efficient alternative for relatively small sample sizes ($n \leq 30$) when the log-likelihood surface is flat in the direction of α. When this happens, the estimate of α is unstable and lacks precision. The expected mean and variance of the distance to the k nearest tree are

$$E[R_k|\alpha > 0.5, \beta] = \frac{\Gamma(\alpha - \frac{1}{2})\Gamma(k + \frac{1}{2})}{\sqrt{\pi \times \beta}\Gamma(\alpha)\Gamma(k)} \tag{8}$$

and

$$\begin{aligned} &\mathrm{var}(R_k|\alpha > 1, \beta) = \pi^{-1}\beta^{-1} \\ &\left[k(\alpha - 1)^{-1} - \Gamma^2\left(k + \frac{1}{2}\right)\Gamma^2\left(\alpha - \frac{1}{2}\right)\Gamma^{-2}(k)\Gamma^{-1}(\alpha)\right] \end{aligned} \tag{9}$$

Notice, the restrictions on α are unimportant in practical applications since $\alpha > 1$when data come from a Poisson or a gamma × Poisson process, but the restrictions must be acknowledged during estimation.

Distance sampling may be constrained to a maximum distance $r_{\max}$ for the k nearest tree. For example, sampling may be excluded from a buffer zone around the perimeter of a forest stand or limited to avoid time-consuming displacements and measurements (see Engeman et al. 1994 for a list of Batcheler-Bell estimators). A constraint on sampling distances must be reflected in the pdf of r_k in (6) by dividing the right-hand side by the probability

$$\begin{aligned} P(R_k < r_{\max}|\alpha, \beta) &= (-1)^{\alpha}\Gamma^{-1}(\alpha)\Gamma^{-1}(k)\Gamma(\alpha + k) \\ &\times \mathrm{Beta}\left(-\beta^{-1}\pi^{-1}r_{\max}^{-2}, \alpha, 1 - \alpha - k\right) \end{aligned} \tag{10}$$

Reference set of density estimators

A set of 17 density estimators listed in Table 1 is used as a reference set in the assessment of λ_{ORBIT} and $\lambda_{\mathrm{GAMPOI}}$. The set includes a majority of published and practically important estimators. We think it is unlikely that an estimator that could potentially outperform the best in the reference set has been missed. The fixed-area plot estimator is a benchmark estimator since it is design-unbiased (Avery and Burkhart (1983). Estimation procedures for the KM and PICARD estimators were non-standard. For KM one needs to compute a Boolean integral of the union of the area of k overlapping circles. Computing the Boolean integral with the Mathematica® software (version 5.2 2005; Wolfram 1999) to four significant digits capped the estimation time of one KM density estimate to 2 min on a 1.8 GHz Pentium PC (i.e. about 3 weeks to finish the 66,000 estimates). In order to obtain a PICARD estimate one has to optimize a pseudo likelihood of the data. The optimization is usually quite difficult due to a highly irregular, discontinuous and spiked likelihood surface. Finding a maximum of a single likelihood surface with a working precision of eight significant digits often took more than an hour. Even so, the maximum could change considerably if the optimization procedure was changed, from the default Nelder–Mead algorithm (Nelder and Mead 1965), to a simulated annealing, a genetic algorithm, or Newton–Raphton's. It would not be feasible to evaluate this estimator over a large number of settings and replicates with standard estimation procedures. Since the estimator has shown promise (Picard et al. 2005), an attempt to resolve the estimation problem was undertaken. A linear Log-likelihood interpolation function in ω, μ, R, and r_k with 2.7 million support points on a regular grid of values was estimated first. For a given observed distance r_k the interpolation function is linear in the unknowns ω, μ, and R, and could be optimized for the entire sample by standard methods. Optimization for a single sample took about a minute. Forty tests comparing results from the new estimation procedure to those obtained by the Nelder–Mead algorithm showed only minor differences (<4%) in estimates of density without indications of a systematic trend.

Simulated distance sampling

Distance sampling with $k = 4,6$ and nominal sample sizes of $n = 10, 20,\ldots, 50$ locations was simulated for 22 sites with stem mapped locations of trees (see below). Each sample was replicated 300 times. At a sample location the distances specified for the two new and 17 reference estimators were measured and recorded. After completion of a sample, an estimate of stem density was obtained for each estimator as outlined above and in Table 1. Sample locations were excluded from a buffer zone around the perimeter of each stem map but otherwise placed at random. The width of the buffer zone was large enough to ensure that no distance from a sample location to the six nearest trees exceeded the width of the buffer. Consequently there was no effective censoring of sampled distances. Limiting sample positions to an interior portion of a stand does change the reference population accordingly. For the sake of transparency we maintained the target density as the total number of trees divided by the total area.

Effective sample size

Some of the reference estimators (Table 1) require either fewer or more than k distances to be measured at a sample

Table 1 Reference set of density estimators (λ trees m^{-2})

Estimator	$\hat{\lambda} =$	Reference	Distances measured	Remarks
FIXED-AREA	$a_{\text{plot}}^{-1} \times \bar{n}_{\text{trees per plot}}$	Avery and Burkhart (1983)	n.a.	Benchmark estimator
MORISITA	$\pi^{-1}\bar{r}_k^{-2}\Gamma^2(k+1)\Gamma^{-2}(k)$	Morisita (1957)	k	Moment estimator under a Poisson model
PERSSON	$k \times \pi^{-1}\tilde{r}_k^{-2}$	Persson (1964)	k	Median distance estimator
BYTH	$k \times 2^{-1.5}\bar{r}_k \times \bar{z}_k$	Byth (1982)	$2k$	Assumed robust properties
KLEINN	$\pi^{-1}k \times \bar{r}_k \times \bar{r}_{k+1}$	Kleinn and Vilčko (2006a)	$k + 1$	Slightly modified to improve performance
THOMPSON	$(n \times k - 1) \times n^{-1}\pi^{-1} \times E^{-1}(r_k^2)$	Thompson (1956)	k	Bias adjusted maximum likelihood estimator for Poisson point pattern
TRANSECT	$(n \times k - 1) \times n^{-1} \times E^{-1}(L_k) \times 10^{-1}$	Parker (1979)	k	Ten meter wide variable length transects. See also Sheil et al. (2003)
CLAYTON$_7$	$(0.45 + 0.08E(\delta_{t_1 > 1.5r_1}))\bar{r}_1^{-1}\bar{t}_1^{-1}$	Clayton and Cox (1986)	2	Conditional (robust) density estimator. See also Cox (1976)
PICARD	$\hat{\omega} \times \hat{\mu}$	Picard et al. (2005)	k	ω is the parent density and μ is the density of off-springs in a Matérn process
DIGGLE	$(n \times k - 1) \times \pi^{-1}E^{0.5}(r_k^2) \times E^{0.5}(z_k^2)$	Diggle (1977)	$2k$	Assumed robust properties
KM	$(n \times k - 1) \times \text{Area}^{-1}\left(\bigcup_{j=1}^{k} C_i\right)$	Kendall and Moran (1963)	k	Density of search area. Search area is the union of a daisy chain of k nearest neighbour circles
CLAYTON$_5$	$(0.35 + 0.01E(\delta_{z_1 > 2r_1}))\bar{r}_1^{-1}\bar{z}_1^{-1}$	Clayton and Cox (1986)	2	Conditioned (robust) density estimator. See also Cox (1976)
LEWIS	$0.767E(r_2^2) - 0.065E(r_1^2)$	Lewis (1975)	2	Assumed robust properties
KEULS	$(n \times k - 1) \times n^{-1}\pi^{-1} \times E(r_k^{-2})$	Keuls et al. (1963)	k	Recommended by Payandeh and Ek (1986)
ANGLE	$k(k \times n - 1)\pi^{-1}\sum_{j=1}^{k} r_1(\theta \mid \theta \in [(2\pi(j-1), 2\pi j))$	Pollard (1971)	k	Distance to nearest tree in each of k equal intervals of polar direction θ
CLAYTON$_6$	$(0.44 + 0.08E(\delta_{t_1 > 2r_1}))\bar{r}_1^{-1}\bar{t}_1^{-1}$	Clayton and Cox (1986)	2	Conditioned (robust) density estimator. See also Cox (1976)
PATIL	$(n^{2/3} - 1)n^{-1}\pi^{-1}\hat{r}_{1([n^{2/3}])}^{-2}$	Patil et al. (1982)	1	$\lfloor \cdot \rfloor$ is the floor function and subscript $(\cdot)$ denotes an interpolated order-statistic

Estimators are listed in order of overall RMSE ranking across all simulated sampling

location. Furthermore, prescribed field procedures also vary by the type of distances to measure. Thus, the time it will take to complete distance measurements for a nominal number of sample locations varies from estimator to estimator. In order to make a fair comparison, we require that they all use the same number of time units to complete sampling. In consequence, a faster procedure will sample more locations than specified by the nominal sample size n, and vice versa.

We assume that field teams are equipped with modern laser range finders, GPS receiver stations, and other electronic equipments to expedite and facilitate both the identification of the measurements to be made and the measurements themselves. It is also assumed that this equipment has all but eliminated any time difference between estimators requiring the same number of distance measurements but with different field procedures. Accordingly, the time to complete a moderate to small number of distance measurements ($k \leq 6$) will be roughly proportional to the number of distances measured (Lynch and Rusydi 1999). Without modern equipment the time is likely to increase with the square of the number of distances (Lessard et al. 1994) due to an increasingly complex task. Leaning on data in Lessard et al. (1994), Lindsey et al. (1958), and Lynch and Rusydi (1999), we fixed the time factor for one distance measure at 1 unit, and the time factor for finding a sample location and walking to the next at 22 units. Travel time to the first sample location is also 22 units. For example, an estimator that requires k distance measures at n sample locations requires $n(22 + k)$ time units to complete.

Time allotted to each estimator is $n(22 + k)$. Accordingly (cf. Table 1), BYTH and DIGGLE achieve an effective sample size n^* of $0.87n$ for $k = 4$, and $0.87n$ for $k = 6$. Corresponding numbers for KLEINN are $n^* = 0.96n$ and $n^* = 0.97n$, respectively. $\text{CLAYTON}_{5,6,7}$ and LEWIS require only 2 measurements per sample location, hence their effective sample sizes are $1.08n$ ($k = 4$) and $1.17n$ ($k = 6$). Finally, for PATIL—with just one measurement per location—the effective sample sizes are $1.13n$ ($k = 4$) and $1.22n$ ($k = 6$). For all others $n^* \equiv n$. In all simulated sampling we use n^* rounded to the nearest integer as the actual (effective) sample size. All results are based on the effective sample size.

Sites and stem maps

Simulated distance sampling was done by random placement of sample locations within a stem mapped site with known coordinates of all trees of interest. Twelve actual stem maps and ten simulated stem maps were used (Table 2). All stem maps are for natural or semi-natural tree stands. They range in size from 0.45 to 50.0 ha and in density from 146 to 3,960 trees per ha (Table 2). Six of the simulated stem maps (Kokani, Sokouna, Korokoro, Dndouantien, Woro, and Amba) were generated according to parameters of a Matérn process with parameters estimated from actual stem maps in Mali Savanna (Picard et al. 2005). The 22 spatial patterns of tree locations cover a range from purely random (Poisson) to clumped (Sokouna). As such they are viewed as a representative sample of irregular natural to semi-natural spatial point patterns.

Performance assessment

Relative root mean square error of an estimate of density serves as the primary statistic for assessing the performance of an estimator. The RRMSE is widely used for comparing the performance of estimators (Picard et al. 2005; Engeman et al. 1994; Clayton and Cox 1986). The RRMSE combines scaled estimates of error and bias into a single statistic

$$\begin{aligned} \text{RRMSE}\left(\hat{\Lambda}_M\right) &= \Lambda^{-1}\sqrt{\text{var}\left(\hat{\Lambda}_M\right) + \left(\hat{\Lambda}_M - \Lambda\right)^2} \\ &= \Lambda^{-1}\text{RMSE}\left(\hat{\Lambda}_M\right) \end{aligned} \tag{11}$$

where Λ is the true tree density as per Table 2, $\hat{\Lambda}_M$is an estimate of density obtained from a sample using estimators M (i.e. one of the 19 detailed above and in Table 1), and $\text{var}\left(\hat{\Lambda}_M\right)$ is the observed variance of $\hat{\Lambda}_M$ across the 300 replications of a sampling design and serves as an empirical approximation to the true error variance. The scaling by Λ^{-1} allows us to compare the RRMSE across sites (maps) with contrasting stem densities.

Users of density estimators for distance sampling prefer an estimator that performs well across a range of spatial point patterns. They need to be assured—by rigorous testing—that their favourite estimator performs to acceptable standards. A preferred estimator must also perform well across a range of sample sizes. It is therefore the average of the RRMSE across a wide range of sites (maps) and replicates that is our primary yardstick of performance. The RRMSE should also be stable across sites and not be overly sensitive to a particular point pattern. The standard deviation of site-specific estimates of the RRMSE gauges this attribute.

In order to simplify reporting of results, we recognize that interest will focus on the best performing estimators. Hence results are only detailed for a subset of seven estimators deemed best in terms of the RRMSE. Readers interested in an estimator not found in this subset can

Table 2 Stem maps used in simulated distance sampling

Name	Location	Area (ha)	Trees per ha (λ)	Source	Remarks
Immature-N	Victoria Watershed District (Vancouver Island, BC, Canada)	0.77	3421	Trofymow et al. (1998)	Douglas-fir dominated stands. See also Getzin et al. (2006) for statistics on spatial distribution
Immature-S	–	0.45	3960	–	–
Mature-N	–	1.07	2913	–	–
Mature-S	–	1.20	2529	–	–
Old-N	–	0.92	2272	–	–
Old-S	–	0.71	2052	–	–
Barro Colorado Island	Gatun Lake, Central Panama	50.00	380	Center for Tropical Forest Science (http://www.ctfs.si.eduUT)	*Leguminosae* and *Bombacaceae* dominated wet tropical forest. Data from 1990 census. See also Condit et al. (1996) and He and Hubbell (2003) for more details
Lansing	East Lansing, MI, USA	7.90	1217	Gerrard (1969)	A natural stand of maple-hickory
Longleaf pine	Wade Tract Preserve near Thomasville Georgia	4.00	146	http://www.talltimbers.org/wadetract.html	See also Rathbun and Cressie (1994) and Cressie (1991 Chap. 8.2) for detailed analyses of the stem map
Miombo	Miombo Woodlands, N. Zambia	13.44	370	Kleinn and Vilčko (2006a)	Trees with DBH > 7 cm
Sherman Block 1		4.00	875	Center for Tropical Forest Science (http://www.ctfs.si.edu)	Tropical moist forest, trees with DBH > 1 m. 1997 census data. Further details are in Condit et al. (2004)
Sherman Block 2		1.96	969	–	–
Poisson	n.a.		500		Simulated
Kokani (Mali)	(8°28′ W;12°24′ N)	0.50	600	Picard et al. (2004)	Tree savanna
Sokouana (Mali)	(7°18′ W;12°52′ N)	0.50	790	–	–
Korokoro (Mali)	(7°24′ W;12°45′ N)	0.50	1348	–	–
Ndouantien (Mali)	(7°17′ W;12°46′ N)	0.50	1486	–	–
Woro (Mali)	(7°29′ W;13°41′ N)	0.50	364	–	–
Amba (Mali)	(3°30′ W;15°15 Ń)	0.50	198	–	–
Matérn 1	n.a.	1.00	500	–	$R = 5$, $\omega = 100$, $\mu = 5$
Matérn 2	n.a.	1.00	500	–	$R = 2$, $\omega = 100$, $\mu = 5$
Matérn 3	n.a.	1.00	500	–	$R = 1$, $\omega = 100$, $\mu = 1$

obtain specific results upon request to the corresponding author.

For each of the 10 combinations of nominal sample size (n) and k, we first identified a set of best estimators, and then selected the seven estimators that appeared most frequently in these selected sets. We began the selection process by ranking the RRMSE of an estimator on a given site (map) from 1(lowest = best) to 19 (highest = worst) and then computed its average rank across the 22 sites and the standard error of its rank average. A t-test then identified estimators with an average rank significantly ($P < 0.05$) better than average (i.e. 19/2 = 9.5) for the particular combination of n and k. Between five and seven estimators fell into this category. Finally, the seven estimators that appeared most often in these 10 selected sets were chosen for reporting. Estimates of bias (average sampling error) and standard error (standard deviation of estimates in replicate sampling) are detailed for the selected sub-set of estimators.

Results

Seven top-ranked estimators

The seven estimators designated as significantly better than average in terms of the RRMSE were (in ascending order of global average): FIXED, MORISITA, ORBIT, GAMPOI, PERSSON, BYTH, and KLEINN. None of the remaining 12 estimators achieved an average n and k specific rank better than 5. Runner ups were THOMPSON, TRANSECT and KM. Apart from a few promising site-specific results, none of the 12 estimators dropped from the reporting appear to be attractive for forest inventory distance sampling in natural or semi-natural stands of trees with a random or irregular spatial distribution.

RRMSE

Average RRMSE of FIXED across 22 sites and replications were, as expected, the lowest across all settings of nominal sample size (n) and k (Table 3). However, for a small sample size ($n = 10$) the advantage of FIXED over the best distance sampling estimator (MORISITA, KLEINN) seems to be unimportant (≤ 0.01). Larger sample sizes convey a RRMSE advantage to FIXED in the range of 0.02–0.03.

The RRMSE closest to that of FIXED was obtained with MORISITA for $n \leq 20$ and GAMPOI for $n \geq 30$ (Table 3). Notice that the differences between the top three RRMSE values are rather modest (≤ 0.01). The composition of the best three estimators changed with sample size and, to a lesser degree, with k. For example, for $n = 10$ and $k = 4$ we have MORISITA $\approx$ KLEINN $\approx$ BYTH, while at $n = 50$ and $k = 6$ we have GAMPOI $\approx$ PERSSON $\approx$ ORBIT. The reason behind these changes is a marked difference in the rate at which RMSE declines with increasing sample size. While the decline is most rapid for GAMPOI and PERSSON, it is about 1.5–3.0 times slower for MORISITA and KLEINN. Consequently, the former two gain in performance with n while the latter two decline. Increasing k from 4 to 6 lowered RRMSE, more at lower n values than at higher values. PERSSON, ORBIT, and GAMPOI appear to benefit more from this increase at $n = 10$ than any of the other four estimators.

The estimator with the most consistent RRMSE performance across sites (lowest standard deviation of the RRMSE) was, as expected, FIXED, but GAMPOI was a close runner-up with PERSSON in third place. The robustness of ORBIT was not markedly different from that of MORISITA, BYTH, or KLEINN.

Relative bias

Of the plot-less estimators the lowest relative bias over all sites (maps) was achieved by KLEINN where the estimated

Table 3 The RRMSE of seven top-ranked density estimators

	$k = 4$					$k = 6$				
$n =$	10	20	30	40	50	10	20	30	40	50
FIXED	0.24 (0.07)	0.17 (0.05)	0.14 (0.05)	0.12 (0.05)	0.11 (0.04)	0.20 (0.06)	0.15 (0.05)	0.12 (0.05)	0.11 (0.04)	0.10 (0.04)
MORISITA	*0.24 (0.07)*	*0.19 (0.08)*	0.18 (0.08)	0.17 (0.09)	0.16 (0.09)	*0.21 (0.07)*	*0.17 (0.10)*	0.15 (0.08)	0.14 (0.08)	0.14 (0.09)
PERSSON	0.29 (0.07)	0.21 (0.06)	0.18 (0.06)	0.17 (0.06)	0.15 (0.06)	0.24 (0.07)	0.18 (0.06)	0.16 (0.06)	0.14 (0.06)	0.13 (0.06)
BYTH	0.25 (0.11)	0.22 (0.09)	0.18 (0.08)	0.17 (0.08)	0.15 (0.08)	0.23 (0.09)	0.21 (0.07)	0.18 (0.06)	0.16 (0.06)	0.14 (0.06)
KLEINN	0.24 (0.11)	0.22 (0.08)	0.20 (0.09)	0.19 (0.09)	0.19 (0.09)	0.21 (0.07)	0.18 (0.08)	0.17 (0.08)	0.16 (0.09)	0.16 (0.09)
ORBIT	0.26 (0.07)	0.20 (0.07)	0.18 (0.07)	0.17 (0.08)	0.16 (0.08)	0.22 (0.07)	0.17 (0.07)	0.15 (0.08)	0.14 (0.08)	0.14 (0.08)
GAMPOI	0.29 (0.06)	0.21 (0.05)	*0.17 (0.04)*	*0.15 (0.04)*	*0.14 (0.04)*	0.25 (0.07)	0.18 (0.05)	*0.15 (0.05)*	*0.14 (0.05)*	*0.12 (0.05)*

Entries are averages of site specific estimates of sRMSE. Among site standard deviations of the RRMSE are in parentheses. The RRMSE closest to FIXED is italised

Table 4 Relative error (bias) of density estimates

	$k = 4$					$k = 6$				
n =	10	20	30	40	50	10	20	30	40	50
FIXED	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)	0.01 (0.05)
MORISITA	−0.12 (0.10)	−0.12 (0.11)	−0.13 (0.11)	−0.13 (0.11)	−0.13 (0.11)	−0.08 (0.09)	−0.10 (0.10)	−0.10 (0.10)	−0.10 (0.10)	−0.11 (0.10)
PERSSON	−0.06 (0.08)	−0.06 (0.08)	−0.07 (0.08)	−0.08 (0.08)	−0.09 (0.08)	−0.05 (0.07)	−0.06 (0.07)	−0.07 (0.08)	−0.07 (0.08)	−0.07 (0.08)
BYTH	−0.15 (0.09)	−0.16 (0.10)	−0.16 (0.10)	−0.17 (0.10)	−0.17 (0.10)	−0.12 (0.09)	−0.13 (0.09)	−0.13 (0.09)	−0.13 (0.09)	0.00 (0.10)
KLEINN	*0.03 (0.13)*	*0.02 (0.13)*	*0.01 (0.13)*	*0.00 (0.13)*	*0.00 (0.13)*	*0.03 (0.11)*	*0.01 (0.10)*	*0.01 (0.11)*	*0.00 (0.11)*	*−0.00 (0.11)*
ORBIT	−0.10 (0.10)	−0.10 (0.10)	−0.11 (0.10)	−0.12 (0.10)	−0.12 (0.10)	−0.08 (0.10)	−0.09 (0.10)	−0.09 (0.10)	−0.10 (0.10)	−0.10 (0.10)
GAMPOI	0.06 (0.09)	0.03 (0.08)	0.02 (0.08)	0.02 (0.08)	0.02 (0.08)	0.06 (0.08)	0.04 (0.08)	0.03 (0.08)	0.03 (0.08)	0.03 (0.08)

Entries are the average of site specific estimates. Among site standard deviations are in parentheses. The error closest to FIXED is italised

bias declined from 0.03 for $n = 10$ to less than 0.01 for $n \geq 40$ (Table 4) with no n and k specific estimate significantly different from 0 (t-test, $\alpha = 0.05$, $P > 0.26$). However, this estimator also has the largest among-site variation in bias, so the bias on any particular site could be substantial. The expected unbiasedness of FIXED was confirmed; the average of about 0.01 across all sample designs does not differ significantly from zero ($P > 0.37$).

ORBIT and GAMPOI are biased but in opposite directions. For ORBIT the relative bias varied between −0.08 and −0.12 with an among-site standard deviation of 0.10. Results are at par with those of MORISITA. GAMPOI was with a relative bias in the interval 0.02–0.06 the runner-up to KLEINN but considerably more stable across sites than the latter. The GAMPOI among-site standard deviation of estimated bias was about 0.08 or approximately 40% lower than for KLEINN.

Relative standard error

Relative standard errors were, on average, lowest for BYTH and highest for PERSSON (Table 5). As sample size increase the estimator differences in relative errors decreased. At $n \geq 40$ they appear unimportant for practical applications. The two new estimators were among the least precise of the seven. Plot sizes in FIXED were designed to hold an average of k-trees in repeat sampling. They were therefore smaller than is typical in forest inventories—a fact that is reflected in a precision slightly (4–7%) below that of the best plot-less estimator.

Discussion and conclusions

Despite a recognized problem with bias, point-to-tree distance sampling for density estimation remains a popular option, especially in highly dense forest, and in difficult terrain (Kleinn and Vilčko 2006a; Picard et al. 2005; Sheil et al. 2003; Lessard et al. 2002; Engeman et al. 1994). Since the pioneering work of Morisita (1957) and later Persson (1964) a multitude of density estimators have been proposed with various claims of robustness, less bias, or improved precision (Picard et al. 2004; Engeman et al. 1994, for example). Although many comparisons of estimators have been done, one is left with an impression that there is no uniformly best estimator. This may be, in part, explained by different sub-sets of estimators appearing in each assessment, different k-values, and a comparison of results across differences in effort (time) (Kleinn and Vilčko 2006a; Picard et al. 2005; Engeman et al. 1994).

Differences in details of applied computational procedures for an estimator can also contribute to some of the

Table 5 Relative standard error of density estimates

	$k = 4$					$k = 6$				
$n =$	10	20	30	40	50	10	20	30	40	50
FIXED	0.23 (0.05)	0.16 (0.04)	0.13 (0.03)	0.12 (0.03)	0.10 (0.02)	0.20 (0.05)	0.14 (0.04)	0.11 (0.03)	0.10 (0.03)	0.09 (0.02)
MORISITA	0.23 (0.05)	0.16 (0.03)	0.13 (0.03)	0.11 (0.03)	0.10 (0.02)	0.20 (0.05)	*0.14 (0.03)*	0.11 (0.03)	0.10 (0.02)	0.09 (0.02)
PERSSON	0.29 (0.07)	0.20 (0.05)	0.17 (0.04)	0.14 (0.03)	0.13 (0.03)	0.24 (0.07)	0.17 (0.05)	0.14 (0.04)	0.12 (0.04)	0.11 (0.03)
BYTH	*0.22 (0.05)*	*0.16 (0.04)*	*0.12 (0.03)*	*0.11 (0.02)*	*0.10 (0.02)*	*0.19 (0.05)*	0.14 (0.04)	*0.11 (0.03)*	*0.09 (0.02)*	*0.08 (0.02)*
KLEINN	0.22 (0.05)	0.17 (0.04)	0.13 (0.03)	0.11 (0.03)	0.10 (0.02)	0.20 (0.06)	0.15 (0.04)	0.12 (0.03)	0.10 (0.03)	0.09 (0.02)
ORBIT	0.25 (0.05)	0.17 (0.04)	0.14 (0.03)	0.12 (0.02)	0.11 (0.02)	0.21 (0.05)	0.14 (0.03)	0.12 (0.03)	0.10 (0.02)	0.09 (0.02)
GAMPOI	0.26 (0.06)	0.18 (0.04)	0.15 (0.03)	0.13 (0.03)	0.11 (0.03)	0.22 (0.06)	0.16 (0.04)	0.13 (0.03)	0.11 (0.03)	0.10 (0.03)

Entries are means of site specific estimates. Standard deviations of site estimates are in parentheses. The lowest error is italised

apparent differences in the performance of an estimator. Many density estimators allow a choice between a density estimate obtained as the mean of n location-specific estimates and one obtained for the entire sample. The former yields a mean of ratios estimator and the latter a ratio of means estimator. Where possible, we opted for the ratio of means as it is the more stable (Cochran 1977). Finally, a lack of a consistent naming convention for estimators means that the same estimator, or minor variants of it, has appeared under different names in different studies (including our own).

The two new estimators were compared to a comprehensive sub-set of estimators that were selected based on published results leading us to expect good performance. By adjusting nominal sample sizes to the effective sample size we recognize that prescribed field procedures take unequal amounts of time. A fair comparison must address this issue. Our approach was pragmatic and simple. Local conditions and logistics may dictate different adjustments than pursued here. Setting k to four and six was motivated by the observation that the performance of all estimators deteriorated rather significantly when k was lowered from four to three. The extra effort required for a few extra measurements is rewarded by a sharp drop in the RRMSE of about 0.10 for $n \leq 20$. No estimator based on one or two distance measurements ranked in the top seven, and this is but another illustration that time-adjusted performance increases rapidly with an increase in k. In forest inventories, $k = 6$ appears to cap the bias to acceptable levels (Kleinn and Vilčko 2006a; Lynch and Rusydi 1999; Prodan 1968).

The 22 sites (maps) used in this study span a wide spectrum of spatial point patterns representative of natural or semi-natural forest tree stands. Consequently, our study has been a comprehensive assessment of the two new estimators and those in the reference set. Their performance on sites with a more or less regular distribution of trees remains to be studied. One might reasonably expect that GAMPOI will not perform particularly well in such patterns since it is based on a model of extra-Poisson variation in density. ORBIT, on the other hand, may, as many other robust estimators, perform almost as well on a regular as on an irregular point pattern (Kleinn and Vilčko 2006a; Engeman et al. 1994; Byth 1982; Diggle 1977; Persson 1971).

Both ORBIT and GAMPOI reached the top six tier of distance-sampling density estimators in terms of the RRMSE performance due to a very good performance when sample sizes are relatively large. This is encouraging since it demonstrates opportunities for even better estimators (Kleinn and Vilčko 2006b). For $n \geq 30$ GAMPOI emerged as the best choice, with ORBIT as a contender. Both are model-based, and these results confirmed that a precise estimate of a distribution of distances or a regression slope requires more than a small sample ($n \geq 20$). The success of the GAMPOI estimator suggests that mixing distributions other than a gamma distribution might be worth investigating. In preliminary studies (not shown) we entertained both an inverse Gaussian and a triangular distribution for the mixing process but found that numerical challenges in maximum likelihood estimation of model parameters were too onerous for practical applications. They were not unlike the problems we encountered during estimation of the parameters of the Matérn process (Picard et al. 2004).

Based on the RRMSE alone MORISITA followed by KLEINN would be our recommendation for $n \leq 20$ but the convincing performance of KLEINN in terms of bias does tip the balance in its favour.

Acknowledgements We are grateful to Drs. F. He (University of Alberta, Edmonton) and T. Trofymow (Canadian Forest Service, Victoria) for making available the Chronosequence stem maps, to the Center of Tropical Forest Science for the Barro Colorado Island and Sherman stem maps, to the Tall Timbers research group for the Wade County Longleaf pine data, to CIRAD for data from the six Mali Savannah sites, and Dr. Gerrard (US Forest Service) for the Lansing Maple-Hickory stem map. The Miombo data are from a research project funded by the United Kingdom Department for International Development (DFID) in its Forest Research Programme (Project R8305).

Appendix A

Notations

Symbol	Definition	Remarks
λ	Density	Unit: trees m^{-2}
Λ	Random variable of density	Unit: trees m^{-2}
R_k	Random variable of distance from sample location to k nearest tree	
r_k	Distance from sample location to k nearest tree	Unit: m
$\bar{r}_k$	The sample mean of distance to k nearest tree	The mean is taken over n viz. n^* sample locations
$\tilde{r}_k$	Median distance to k nearest tree	
A_j	Area of random annulus centred at sample location and with inner radius $\sqrt{R_{j-1}R_j}$ and outer radius $\sqrt{R_jR_{j+1}}$	
a_{plot}	Area of circular sample plot	Unit: m^2
n	Sample size	No. of sample locations
n^*	Effective sample size	
$\Gamma(z)$	$\int_0^\infty e^{-t}t^{z-1}dt$	Euler's gamma function
$E(X)$	Expected value of a random variable X	In a sampling context: the mean of observed values
L_k	Minimum length of a rectangle, with width W and anchored at a sample location, that holds exactly k-trees	Unit: m
z_k	Distance to k nearest neighbour from k nearest tree to a sample location	
t_k	Distance to k nearest neighbour at location T from k nearest tree at location K to sample location (S) provided that $\overrightarrow{KS} \cdot \overrightarrow{TK} \leq 0$	T is excluded from the half-plane defined by S and the line through (K) and perpendicular to $\overrightarrow{KS}$
C_j	Search circle j centred at sample location if $j = 1$, and at nearest neighbour tree to nearest neighbour tree for $j > 1$	A tree can only be selected once as a nearest neighbour to a sample location viz. another tree. Radii in search circles are the distance from the centre to the nearest neighbour tree
$\delta_{\text{condition}}$	An indicator variable taking the value 1 if condition is satisfied and 0 otherwise	

References

Avery TE, Burkhart HE (1983) Forest measurements, 3rd edn. McGraw-Hill, New York

Byth K (1982) On robust distance-based intensity estimators. Biometrics 38:127–135

Clayton G, Cox TF (1986) Some robust density estimators for spatial point processes. Biometrics 42:753–767

Cochran WG (1977) Sampling techniques. Wiley, New York

Condit R, Hubbell SP, Lafrankie JV, Sukumar R, Manokaran N, Foster RB, Ashton PS (1996) Species-area and species-individual relationships for tropical trees: a comparison of three 50-ha plots. J Ecol 84:549–562

Condit R, Aguilar A, Hernández A, Pérez R, Lao S, Angehr G, Hubbell SP, Foster RB (2004) Tropical forest dynamics across a rainfall gradient and the impact of an El Niño dry season. Ecology 20:51–72

Cox TF (1976) The robust estimation of the density of a forest stand using a new conditioned distance method. Biometrika 63:493–499

Cressie NAC (1991) Statistics for spatial data. Wiley, New York

Delince J (1986) Robust density estimation through distance measurements. Ecology 67:1576–1581

Diggle PJ (1977) A note on robust density estimation for spatial point patterns. Biometrika 64:91–95

Diggle PJ (1983) Statistical analysis of spatial point patterns. Academic Press, New York

Druckenbrod DL, Shugart HH, Davies I (2005) Spatial pattern and process in forest stands within the Virginia piedmont. J Veg Sci 16:37–48

Eberhardt LL (1967) Some developments in 'distance sampling'. Biometrics 27:207–216

Engeman RM, Sugihara RT, Dusenberry WE (1994) A comparison of plotless density estimators using Monte Carlo simulation. Ecology 75:1769–1779

Gerrard DJ (1969) Competition quotient: a new measure of the competition affecting individual forest trees. Mich. St. Univ. Agric. Exp. Stn. Report 20

Getzin S, Dean C, He F, Trofymow JA, Wiegand K, Wiegand T (2006) Spatial patterns and competition of tree species in a Douglas-fir chronosequence on Vancouver Island. Ecography 29:671–682

Hall JB (1991) Multiple-nearest-tree sampling in an ecological survey of Afromontane catchments forests. For Ecol Manage 42:245–266

He F, Hubbell SP (2003) Percolation theory for the distribution and abundance of species. Phys Rev Letters 91:198103.1–198103.4

Johnson NL, Kotz S, Kemp AW (1992) Univariate discrete distributions. 2nd edn. Wiley, New York

Jonsson B, Holm S, Kallur H (1992) A forest inventory method based on density-adapted circular plot size. Scand J For Res 7:405–421
Kendall MG, Moran PAP (1963) Geometric probability. Charles Griffin, London
Kendall MG, Stuart A (1969) The advanced theory of statistics. 3rd edn. Griffin, London
Keuls M, Over HJ, De Wit CT (1963) The distance method for estimating densities. Statist Neerl 17:71–91
Kint V, De Wulf R, Noel L (2004) Evaluation of sampling methods for the estimation of structural indices in forest stands. Ecol Model 180:461–476
Kleinn C, Vilčko F (2006a) A new empirical approach for estimation in k-tree sampling. For Ecol Manage 237:522–533
Kleinn C, Vilčko F (2006b) Design-unbiased estimation for point-to-tree distance sampling. Can J For Res 36:1407–1414
Lessard VC, Reed D, Monkevich N (1994) Comparing n-tree distance sampling with point and plot sampling in northern Michigan forest types. N J Appl For 11:12–16
Lessard VC, Drummer TD, Reed DDA (2002) Precision of density estimates from fixed-radius plots compared to n-tree distance sampling. For Sci 48:1–6
Lewis SM (1975) Robust estimation of density for a two-dimensional point-process. Biometrika 62:519–521
Lindsey AA, Barton JD, Miles SR (1958) Field efficiencies of forest sampling methods. Ecology 39:428–444
Lynch TB, Rusydi R (1999) Distance sampling for forest inventory in Indonesian teak plantations. For Ecol Manage 113:215–221
Morisita M (1957) A new method for the estimation of density by the spacing method applicable to non-randomly distributed populations. Physiol Ecol 7:134–144
Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7:308–313
Parker KR (1979) Density estimation by variable area transect. J Wildl Manage 43:484–492
Patil SA, Kovner JL, Burnham KP (1982) Optimum nonparametric estimation of population density based on ordered distances. Biometrics 35:597–604
Payandeh B, Ek AR (1986) Distance methods and density estimators. Can J For Res 16:918–924
Persson O (1964) Distance methods: The use of distance measurements in the estimation of seedling density and open space frequency. Skogshögskolan 15:68 pp
Persson O (1971) The robustness of estimating density by distance measurements. Sampling and modelling biological populations and population dynamics Penn State Statistics Series. Statistical Ecology, Pennsylvania State University Press, University Park, Pennsylvania
Picard N, Nouvellet Y, Sylla ML (2004) Relationship between plot size and the variance of the density estimator in West African savannas. Can J For Res 34:2018–2026
Picard N, Kouyaté AM, Dessard H (2005) Tree density estimations using a distance method in Mali Savanna. For Sci 31:7–18
Pollard JH (1971) On distance estimators of density in randomly distributed forests. Biometrics 27:991–1002
Pretzsch H (1997) Analysis and modeling of spatial stand structures. Methodological considerations based on mixed beech-larch stands in Lower Saxony. For Ecol Manage 97:237–253
Prodan M (1968) Punktstichprobe für die Forsteinrichtung (in German). Forst- und Holzwirt 23:225–226
Rathbun SL, Cressie N (1994) A space-time survival point process for a longleaf pine forest in southern Georgia. J Am Stat Assoc 89:1164–1174
Schreuder HT (2004) Sampling using a fixed number of trees per plot. USDA Forest Service, Research Note. RMRS, 3 pp
Schreuder HT, Gregoire TG, Wood GB (1993) Sampling methods for multiresource forest inventory. Wiley, New York
Sheil D, Ducey MJ, Sidiyasa K, Samsoedin I (2003) A new type of sample unit for the efficient assessment of diverse tree communities in complex forest landscapes. J Trop For Sci 15:117–135
Thompson HR (1956) Distribution of distance to the n-th neighbour in a population of randomly distributed individuals. Ecology 37:394
Trofymow JA, Porter G, Blackwell B, Arksey R, Marshall V, Pollard D (1998) Chronosequences for research into the effects of converting coastal British Columbia old-growth forests to managed forests: an establishment report. Can For Serv Inf Rep BC-X-374
Turechek WW, Madden LV (1999) Spatial pattern analysis of strawberry leaf blight in perennial production systems. Phytopathology 89:421–433
Upton GJG, Fingleton B (1985) Spatial data analysis by example. Volume I. Point patterns and Quantitative data. Wiley, Chichester
White GC, Bennetts RE (1996) Analysis of frequency count data using the negative binomial distribution. Ecology 77:2549–2557
Wolf A (2005) Fifty year record of change in tree spatial patterns within a mixed deciduous forest. For Ecol Manage 215:212–223
Wolfram S (1999) The mathematica book. 4th edn. Wolfram Media/Cambridge University Press, Champaign