Eur J Forest Res (2012) 131:1513–1527
DOI 10.1007/s10342-012-0619-4

A new composite *k*-tree estimator of stem density

Steen Magnussen

Received: 7 October 2011 / Revised: 3 January 2012 / Accepted: 17 February 2012 / Published online: 9 March 2012

Abstract This study presents a generally applicable and robust *k*-tree composite estimator of density. We propose to estimate stem density by a weighted average $\left(\hat{\lambda}_{\text{aic}}\right)$ of 16 individual density estimators. The weights given to individual estimators are inversely proportional to the relative fit (Akaike's corrected information criterion) of each estimator to the assumed distribution of observed *k*-tree distances. The performance of the proposed estimator is evaluated in simulated simple random sampling with $k = 3$ and 6 in 58 forest stands (54 actual and 4 simulated) and 600 replications. Sample sizes were 15 and 30 locations per stand. Eleven estimators were novel, including three designed for regular spatial patterns. Absolute stand-level bias with $k = 6$ varied from 0.1 to 8.1% (mean 1.8%), and a bias larger than 6% was limited to 3 stands with either pronounced density gradients or a strong clustering of stem locations. Root mean squared errors were approximately 16% ($k = 6$ and $n = 15$) versus 12% for sampling with comparable fixed-area plots. Coverage of computed 95% confidence intervals ranged from 0.72 to 0.99 (median = 0.98 with $n = 15$ and 0.95 with $n = 30$), with 98% of all intervals achieving a coverage of 0.85 or better. In seven stands used in an assessment of a novel spatial point pattern reconstruction *k*-tree density estimator (RDE) by Nothdurft et al. (Can J For Res 40:953–967, 2010), the average absolute bias of $\hat{\lambda}_{\text{aic}}$ with $k = 6$ was 1.5 versus 0.7% for $\hat{\lambda}_{\text{RDE}}$.

Communicated by A. Weiskittel.

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

Keywords Plotless sampling · Fixed-count sampling · Point-to-tree distance · Distance sampling · Akaike's information criterion

Introduction

k-tree sampling (viz. plotless, fixed-count, and point-to-object sampling) remains an attractive option for both ecological surveys (Augustin et al. 2009; Kint et al. 2004; Sandrine et al. 2003; Sheil et al. 2003) and forest inventories (Kleinn and Vilčko 2006a; Nothdurft et al. 2010; Staupendahl 2008) because they offer advantages in terms of cost, flexibility and control over sample yields (Jonsson et al. 1992; Lynch and Wittwer 2003). In forestry, *k*-tree sampling with a sample size of *n* typically involves the measurement of a number of desired tree attributes (size, height, etc.) of the *k* trees located nearest (sensu Euclidean distance) to each of the randomly or systematically selected sample locations.

A *k*-tree sampling estimator of per tree attribute, if formulated as a ratio estimator, can be approximately unbiased (Lynch and Wittwer 2003; Sutherland 2006, pp. 319–320) as long as sample locations are chosen at random, and the covariance between the tree attribute and the square of the distance to the *k*th tree is weak (Lessard et al. 2002; Lessard et al. 1994; Saborowski and Šmelko 1998). In contrast, *k*-tree estimators of density and, by extension, associated ratio estimators of attribute value per unit area are not, because sample inclusion probabilities of selected trees remain unknown. Unbiased inclusion probabilities of sample trees can only be computed through a *k*-order Voronoi tessellation of the area surrounding a sample location (Fehrmann et al. 2011; Kleinn et al. 2009); in practice, an inefficient task that offsets all practical advantages of a constant sample yield.

A novel approach toward mitigation of the bias problem is inspired by algorithms for reconstruction of spatial point patterns from cumulative spherical contact and nearest-neighbor distance distribution functions (Tscheschel and Stoyan 2006). Nothdurft et al. (2010) devised an iterative simulation procedure (RDE, reconstruction-based density estimation) that generates a stem map of predicted and observed ($n \times k$) stem locations. To work, the coordinates of all k trees must be known to compute $n \times k \times (k-1)$ inter-tree distances, as well as the $n \times 10 \times k$ nearest neighbor distances, from 10 test point locations virtually arranged in a regular grid centered on each sample plot. In applications with seven actual data sets with diverse patterns of stem locations, the estimates of bias for $k = 6$ and approximately 4 samples/ha ranged from −0.9 to 1.8%. In simulations with complete spatial randomness (csr) of tree locations, and more samples per hectare, the stand-level bias was <1%. The requirements of RDE to measure a large number of additional distances limit logistical advantages of distance sampling vis-à-vis sampling with fixed-area plots, unless field measurements are automated.

Kronenfeld (2009) suggested a nonlinear, model-based estimator inspired by observed rates of bias reduction with increasing k in hexagonal and maximally clustered patterns of stem locations. Test results with 13 simulated test patterns were encouraging.

The bias problem in density estimators for distance sampling remains a detractor despite recent successes and many other partially successful attempts to mitigate bias through the design of new estimators or modifications to existing ones (for examples, Barabesi and Marcheselli 2004; Byth 1982; Cox et al. 1997; Engeman et al. 1994; Kendall and Moran 1963; Kleinn and Vilčko 2006b; Lewis 1975; Magnussen et al. 2008; McIntire and Fajardo 2009; Parker 1979; Patil et al. 1982; Popescu et al. 2003; Prodan 1968). Today we have approximately unbiased density estimators for distance sampling applicable to a few well-defined spatial distributions of tree locations (Delince 1986; Eberhardt 1967; Kleinn and Vilčko 2006b; Magnussen et al. 2008; Persson 1964; Picard et al. 2005). Unfortunately, outside their domain of optimal performance bias can be serious (Magnussen et al. 2008). In many practical situations, the analyst would lack the information needed to select the most appropriate estimator.

To capitalize on the advantages of distance sampling, a robust and approximately un-biased density estimator is needed; robust in the sense that the probability of an important bias is low. A practically important bias could be, for example, a bias that in absolute terms amounts to more than 30% of the precision in a comparative fixed-area plot sampling experiment (Cochran 1977, p. 15). A recently proposed adaptive composite density estimator, $\hat{\lambda}_{\text{com}}$, for distance sampling (Magnussen et al. 2011) was designed with robustness and general applicability in mind. The estimator is a weighted linear combination of three top-ranked k-tree density estimators (Magnussen et al. 2008) with negatively correlated bias and weights determined from the sample data. Although the average and range of bias in applications with nine actual stands was superior to individual component results, the risk of a relative bias larger than 10% was a concern.

This study proposes a new, composite k-tree density estimator with component weights determined by the relative fit of observed distances to the assumed distribution of these distances. Hence, the estimator does not rely on a calibration model (Magnussen et al. 2011). Specifically, the proposed estimator is a weighted sum of 16 individual k-tree density estimators. The performance of the proposed estimator is evaluated in simulated random sampling from 54 actual stem-mapped forest stands and 4 simulated stem maps.

Materials and methods

Composite estimator of stem density

We propose a composite density estimator, $\hat{\lambda}_{\text{aic}}$, for the actual but unknown stem density λ. The proposed estimator is a weighted average of 16 individual k-tree density estimators. Thirteen of the component estimators take the form of a parametric density-dependent probability distribution function (pdf) for $r_{k,i}$ (the distance to the k-nearest tree from the ith randomly selected sample location, $i = 1,\ldots, n$). Details of the 13 pdfs and their associated estimators of λ are given in "Appendix 1: maximum likelihood estimators of density (λ)". Eleven of the 13 estimators are novel, and three of the novel estimators are designed for regular point patterns. For these 13 estimators, the density estimates were obtained by a method of maximum likelihood (MLE). The remaining three components were (1) $\hat{\lambda}_{kv}$, by Kleinn and Vilčko (2006b), (2) $\hat{\lambda}_{\text{pers}}$, by Persson (1964), and (3) $\hat{\lambda}_{kf}$, by Kronenfeld (2009). Specifically,

$$\begin{aligned}\hat{\lambda}_{kv} &= (k-1) \times \pi^{-1} \sqrt[-0.5]{\bar{r}_k \times \bar{r}_{k-1}} \\ \hat{\lambda}_{\text{pers}} &= k \times \pi^{-1} \text{median}\left(r_{k,i}\right), \quad i = 1,\ldots,n \\ \hat{\lambda}_{kf} &= f\left(\hat{c}, \hat{\lambda}_{\text{mor}}\right)\end{aligned} \tag{1}$$

where $\hat{\lambda}_{\text{mor}}$ is Morisita's k-tree method-of-moments estimator of λ under csr, and $\hat{\lambda}_{kf}$ is the estimator that minimizes the following sum of squared deviations: $\sum_{i=1}^{k}\left(2i(2i-c-1)^{-1}\lambda_{kf} - \hat{\lambda}_{\text{mor}}\right)^2$ where c and λ_{kf} are

constants to be estimated by nonlinear least squares methods (Ratkowsky 1983). The three estimators in (1) are not likelihood-based estimators, but for the purpose of assigning weights, it was decided to assign each with a log-likelihood of the data under the assumption of csr and conditional on the estimated value of λ. The number of parameters in the three estimators in (1) is 2, 1, and 2, respectively.

In $\hat{\lambda}_{\text{aic}}$, component weights are proportional to the inverse of Akaike's corrected information criterion (AICc, Burnham and Anderson 2004), which is minus twice the log-likelihood of the n k-tree distances under the assumed model and estimated model parameters plus a small sample-size correction for the number of model parameters. Specifically, we let AICc_j be the AICc for the jth estimator ($j = 1,\ldots,16$) and let $\text{AIC}_{\min}$ denote the minimum of AICc_j, so then

$$\hat{\lambda}_{\text{aic}} = \frac{\sum_1^{16} \exp\left(0.5\left(\text{AIC}_{\min} - \text{AICc}_j\right)\right)\hat{\lambda}_j}{\sum_{j=1}^{16} \exp\left(0.5\left(\text{AIC}_{\min} - \text{AICc}_j\right)\right)} = \sum_{j=1}^{16} \hat{w}_j \hat{\lambda}_j \quad (2)$$

According to Burnham and Anderson (2004), the term $\exp\left(0.5\left(\text{AIC}_{\min} - \text{AICc}_j\right)\right)$ is an estimate of the probability that model j provides a better fit to data than the model associated with $\text{AIC}_{\min}$. As a variance estimator for $\hat{\lambda}_{\text{aic}}$, we used the following approximation:

$$\hat{\text{var}}\left(\hat{\lambda}_{\text{aic}}\right) = \sum_{j=1}^{16} \hat{w}_j \hat{\text{var}}\left(\hat{\lambda}_j\right) \quad (3)$$

where $\hat{var}\left(\hat{\lambda}_j\right)$ is the estimated variance of $\hat{\lambda}_j$ obtained via estimates of the sampling variance of the parameters in the jth estimator, and we used an application of the delta method (Oehlert 1992) to the known relationship between the parameters in the jth pdf and λ (see "Appendix 1: maximum likelihood estimators of density (λ)") and MLEs of the variance of the estimated parameters. For the three estimators in (1), a variance estimator is not provided. Instead, the variance was equated to the variance under csr and the given estimate of density (Pollard 1971). A 95% student's t quantile confidence interval for the true but unknown stem density was computed for each sample.

Comparison estimators of stem density

The adaptive composite k-tree estimator of density, $\hat{\lambda}_{\text{com}}$, by Magnussen et al. (2011) served as a comparison estimator. It has demonstrated a low to moderate bias across a wide range of stem densities and spatial patterns of tree locations. $\hat{\lambda}_{\text{com}}$ is a weighted sum of three k-tree density estimators: $\hat{\lambda}_{\text{gam·poi}}$, the Gamma–Poisson estimator by Magnussen et al. (2008); $\hat{\lambda}_{kv}$, the estimator by Kleinn and Vilčko (2006b); and $\hat{\lambda}_{\text{per}}$, the median density estimator by Persson (1964). Specifically,

$$\hat{\lambda}_{\text{com}}(k) = \sum_{c=\{\text{gam·poi},kv,\text{per}\}} \hat{w}_c(\mathbf{X}) \times \hat{\lambda}_c, \quad 0 \le \hat{w}_c(\mathbf{X}) \le 1, \quad \sum_{c=\{\text{gam·poi},kv,\text{per}\}} \hat{w}_c(\mathbf{X}) = 1 \quad (4)$$

where the weights $\hat{w}_c(\mathbf{X})$ are functions of three ancillary statistics obtained from the sample data. The following set of multinomial logistic functions (Greene 2003, pp. 857–862) were used to compute estimator weights

$$\begin{aligned}
\hat{w}_{gp} &= \left(1 + \hat{\omega}_{kv} + \hat{\omega}_{\text{per}}\right)^{-1} \\
\hat{w}_{kv} &= \omega_{kv}\left(1 + \hat{\omega}_{kv} + \hat{\omega}_{\text{per}}\right)^{-1} \\
\hat{w}_{\text{per}} &= \hat{\omega}_{\text{per}}\left(1 + \hat{\omega}_{kv} + \hat{\omega}_{\text{per}}\right)^{-1} \\
\hat{\omega}_{kv} &= \exp\left(-0.70 - 0.55k + 2.31\widehat{MV} - 0.24n - 1.03\widehat{VR}\right) \\
\hat{\omega}_{\text{per}} &= \exp\left(-0.50 - 0.47k + 1.81\widehat{MV} - 0.15n + 0.22\widehat{VR}\right)
\end{aligned} \quad (5)$$

where $\widehat{MV}$ is the ratio of the mean to variance of the n distances to the k-nearest tree from the n random sample locations, and $\widehat{VR}$ is the ratio of the observed to the expected variance in a Poisson forest of the distances to the k-nearest tree. An expression for the expected variance can be found in Pollard (1971, equation 18); it requires an estimate of λ for which we used $\hat{\lambda}_{\text{mle}} = (n \times k - 1)\pi^{-1} \times \bar{r}_k^{-2}$ (Pollard 1971).

Based on a good root mean squared error performance in comparative study of k-tree density estimators, we also used $\hat{\lambda}_{kv}$ by Kleinn and Vilčko (2006b) as a comparison estimator (Magnussen et al. 2008). Finally, we also compared $\hat{\lambda}_{\text{aic}}$ to $\hat{\lambda}_{kf}$ (Kronenfeld 2009).

The RDE proposed by Nothdurft et al. (2010) would have been a logical fourth k-tree comparison estimator. However, for the study design adopted here (see next), it would place an excessive demand on computing time. Hence, it was not included; but, where possible, a comparison will be made with published results for $\hat{\lambda}_{\text{RDE}}$.

Evaluation of estimators

The performance of $\hat{\lambda}_{\text{aic}}$, $\hat{\lambda}_{\text{com}}$, $\hat{\lambda}_{kv}$, and $\hat{\lambda}_{kf}$ was evaluated in replicated Monte–Carlo (MC) simulations of simple random sampling from 58 forest stands (54 actual and 4 simulated; see "Appendix 2" for details). Performance was assessed with respect to bias (estimated density minus actual density), precision (viz. MC estimate of sampling error), root mean squared error (RMSE = square root of the sum of bias squared and the MC estimate of variance), and coverage of computed 95% confidence intervals for the true density (coverage is the proportion of intervals that

included the true density). Bias, precision, and RMSE are expressed in percent of the actual stem density.

The performance evaluation took place with k values of 6 and 3, sample sizes of 15 and 30, and 600 replications of each design setting and stand. A k value of 6 was chosen because, in forestry, it is considered a favorable compromise between field effort and the anticipated bias and precision (Lynch and Rusydi 1999; Prodan 1968; Staupendahl 2008). A k value of 3 was included to demonstrate that the performance of $\hat{\lambda}_{\text{aic}}$ deteriorates at a much slower rate with decreasing k than what can otherwise be expected from k-tree density estimators (Kleinn and Vilčko 2006b; Kronenfeld 2009; Magnussen et al. 2008; Steinke and Hennenberg 2006). With 600 replications, the average relative MC error (Koehler et al. 2009) of $\hat{\lambda}_{\text{fix}}$ with $k = 6$ and $n = 30$ was 0.35% and the error of an estimate of coverage was 0.8%.

Simple random sampling of sample locations was done with replacement in stands with an area <3 ha. In larger stands, sampling was conducted without replacement from a grid with 1 m^2 cells (to avoid duplication of sampled distances). To minimize edge effects, the sampling was restricted to locations with a minimum distance to the nearest stand boundary equal to the radius of a fixed-area plot holding an average of k trees (Magnussen et al. 2011). The average density estimated with co-located fixed-area plot sampling $\left(\hat{\lambda}_{\text{fix}}\right)$ within this buffered area served as the actual stand density when estimating bias and RMSEs of $\hat{\lambda}_{\text{aic}}$, $\hat{\lambda}_{\text{com}}$, $\hat{\lambda}_{kv}$, and $\hat{\lambda}_{kf}$. Thus, it is assumed that the density in the buffered exclusion zone is equal to the density in the stand interior. The use of identical sample locations for all estimators and avoidance of sampling at stand edges was done to reduce, to the largest extent possible, any differential edge effects on the estimators.

The design-unbiased fixed-area density estimator (De Vries 1986) is

$$\hat{\lambda}_{\text{fix}}(k) = \frac{1}{n \times A_k} \sum_{i=1}^{n} n_i \tag{6}$$

where A_k is the area of a circle containing, on average, k trees in simple random sampling with replacement, and n_i is the number of trees sampled at location i. The applicable circle radius for each study stand was determined through trial-and-error MC simulations with 2,000 sampling locations on a regular grid.

Results

Stand-level distribution of r_k

The distribution of r_k plays a central role for the performance of 13 of the proposed density estimator. The stand-level mean and median of r_k in the 58 stands was, in 46 cases (80%), within 5% of the expected value under csr of tree locations. The top row in Fig. 1 shows the distribution of the ratio of the expected to the observed mean of r_k. The distribution is similar for $k = 3$ and 6 with a mean (median) of 0.97 (0.99), and 95% of the stands with ratios between 0.82 and 1.04. Extreme ratios were observed in stands 6 (0.58) and 10 (1.06).

The distributions of the ratio of the expected to the observed variance of r_k are shown in the middle row of Fig. 1. Again, the effect of k is subtle. Approximately half the stands exhibited a variance of r_k that was less than expected under csr. There was a large spread in the ratio, with 95% falling in the interval from 0.21 to 3.71. An increase in k from 3 to 6 lowered the ratio slightly. Extreme variance ratios were observed in stands 6 (0.06) and 57 (4.44).

Skewness ratios are summarized in the third row of Fig. 1. Here, an effect of k is discernable: the mean of the distribution with $k = 3$ have shifted approximately 0.4 to the right relative to the distribution for $k = 6$. With $k = 3$ the mean (median) skewness ratio was 1.01 (0.70) and with $k = 6$ it was 0.62 (0.55). Approximately 70% of the stands exhibit a stronger skewness in the distribution of r_k than under csr. An interval from 0.09 to 2.62 captured approximately 95% of the ratios, but extremes ranged from −6.0 in stand 12 to 8.5 in stand 58. Increasing k from 3 to 6 invoked a statistically significant change from a right- to a left-skewed distribution of r_k in stands 12 and 58, both with extreme skewness ratios.

For a known stem density, the cumulative distribution function (cdf) of r_k under csr is known (Thompson 1956). For each stand, the cdf of r_k under csr was compared to its empirical counterpart (ecdf) estimated from 1,000 random sampling locations. With $k = 3$ the ecdfs deviated significantly from the distributions under csr ($P < 0.05$, permutation test (Good 1993) with a Kolmogorov–Smirnov [KS] $D_{\max}$ test statistic (Fattorini et al. 2001) and 400 replications). With $k = 6$ the number of rejections was 24. Stands with a significant departure from csr are marked in Tables 2, 3, 4, and 5 in Appendix B.

In practical applications with relatively small stand-level sample sizes, the statistical power to reject csr is low. For example, with $n = 30$, we would expect only 10 rejections (stands 2, 4, **5**, **6**, **8**, 9, **30**, 44, **56**, **57**), and with $n = 15$ only 6 (stands set in bold in previous list of 10). Hence, in most of the 58 stands, the distribution of sampled distances will appear consistent with the distribution under csr. Figure 2 shows four examples of observed and expected cdfs for $k = 3$ and 6. The examples were selected at random among stands with a rejection of csr.

The best fit (sensu minimum $D_{\max}$) to the stand-wide empirical distributions of r_k was obtained with seven of the

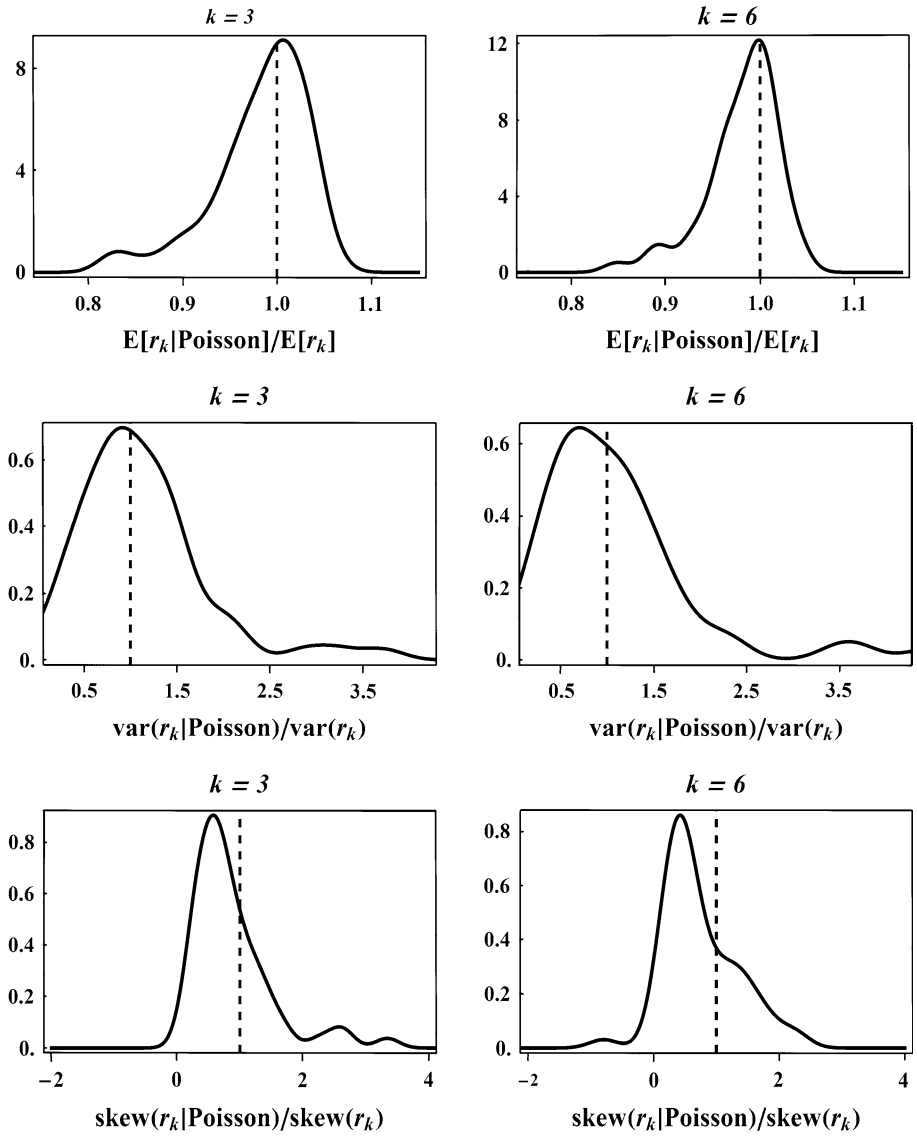

Fig. 1 Among-stand distribution of the ratio of expected (Poisson forest) to observed mean (*top*), variance (*middle*), and skewness (*bottom*) of r_k

proposed models (Table 1). In accordance with the above results, f_{poi} achieved the best fit in 34 stands, followed by $f_{\text{poi·tri}}$ ($f_{\text{poi·sq}}$) with a best fit in 9 stands. Distributions $f_{\text{ig·poi}}$ and $f_{\text{uni·poi}}$ ($f_{\text{poi·poi}}$) were best in 5 (4) stands. Finally, $f_{\text{birn·poi}}$ was best in just one stand (6). When the best stand-specific model was different from f_{poi}, the $D_{\max}$ fit-statistic for the best fit model was (on average) about 1/3 of the $D_{\max}$ for f_{poi} and the associated P-value was approximately 12 times larger.

In the Monte Carlo simulations with $n = 15$ and 30, each of the 13 proposed distribution models for r_k achieved the minimum AICc in 249 cases or more (out of a total of 58 [stands] × 2 [k] × 2 [n] × 600 [replicates] = 132 000). Again, f_{poi} dominated with a best fit in 48% of all cases. Three distributions (poi·hx, erl·poi, and gp·poi) achieved the best fit at a rate between 0.2 and 0.6%. The remaining nine models were best [sensu min (AICc)] at a rate between 1.1 and 11.7%.

Bias

Taken over the 58 stands, the average estimate of absolute bias in $\hat{\lambda}_{\text{aic}}$ with $k = 3$ was 2.6% (median: 2.1%, max: 9.7%); with $k = 6$ it was 1.8% (median: 1.3%, range: 0.1% to 8.1%). In comparison, the average absolute bias in $\hat{\lambda}_{\text{com}}$ was 6.5% with $k = 3$ (max 18.2%) and 2.5% with $k = 6$ (max 7.7%). A larger average absolute bias was encountered in $\hat{\lambda}_{kv}$ (12.0% with $k = 3$ and 8.7% with $k = 6$) and $\hat{\lambda}_{kf}$ (7.5% with $k = 3$ and 6.8% with $k = 6$). Histograms of the absolute bias in $\hat{\lambda}_{\text{aic}}$ and the three comparison estimators $\left(\hat{\lambda}_{kv},\ \hat{\lambda}_{kf}, \text{and } \hat{\lambda}_{\text{com}}\right)$ are

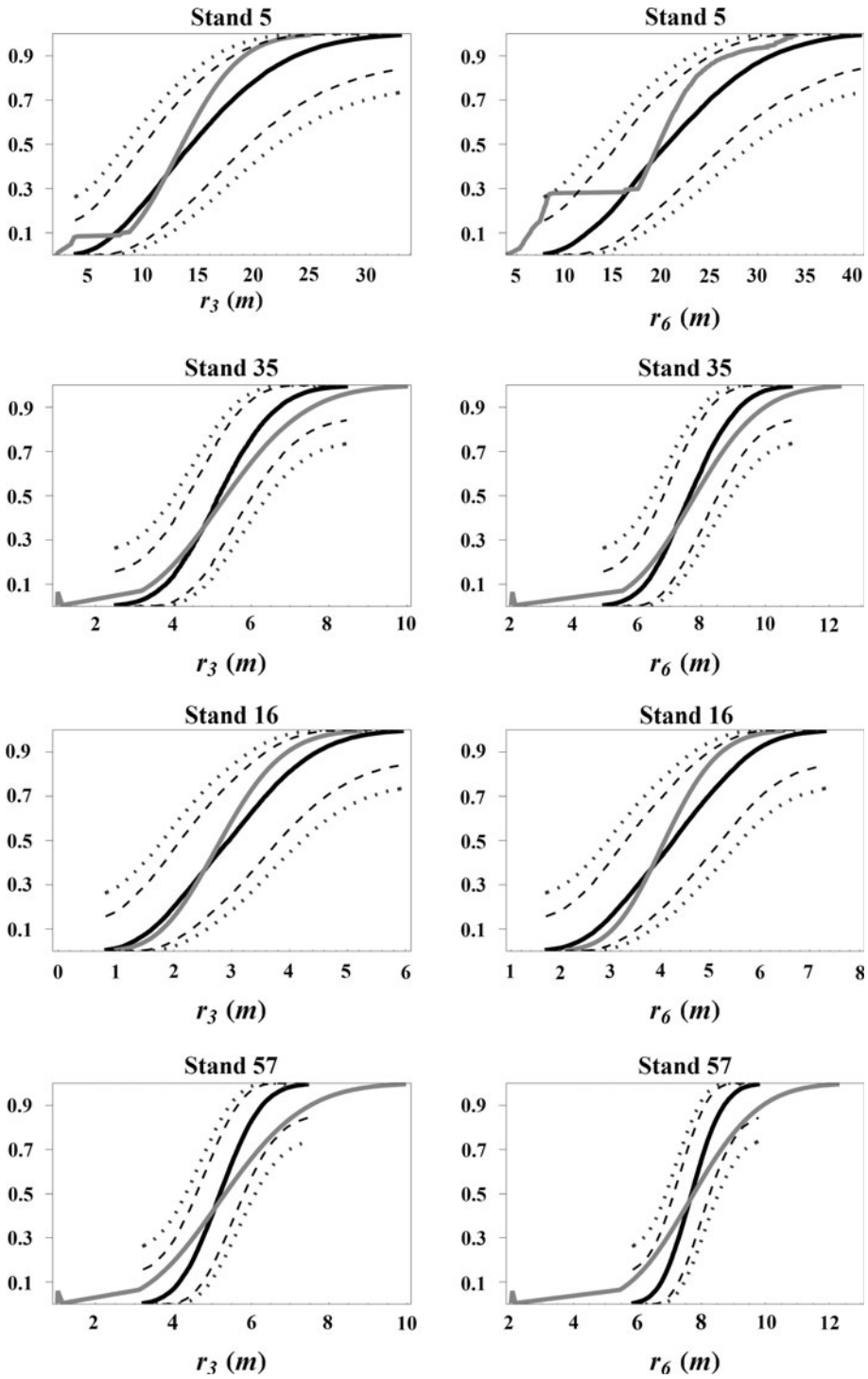

Fig. 2 Randomly selected examples of observed (*black*) and expected (*gray*) distribution functions of r_k in four stands with a rejection of the null hypothesis of csr of stem locations. Estimated 95% confidence bands (Owen 1995) for sample sizes of 15 (*dotted*) and 30 (*dashed*) are shown

in Figs. 3 ($k = 3$) and 4 ($k = 6$). It is apparent that the absolute bias in $\hat{\lambda}_{\text{aic}}$ is <3% in most stands and over 6% in just 3. With $k = 3$ the distribution of absolute bias in $\hat{\lambda}_{kv}$, $\hat{\lambda}_{kf}$, and $\hat{\lambda}_{\text{com}}$ was clearly inferior to that of $\hat{\lambda}_{\text{aic}}$. With $k = 6$ the distributions of absolute bias in $\hat{\lambda}_{\text{com}}$ and $\hat{\lambda}_{\text{aic}}$ are not too dissimilar, while those of $\hat{\lambda}_{kv}$, and $\hat{\lambda}_{kf}$ remain inferior to that of $\hat{\lambda}_{\text{aic}}$. On a stand-by-stand basis, the absolute bias in $\hat{\lambda}_{\text{aic}}$ was less than the bias of the least biased comparison estimator in 40 cases with $k = 3$ and in 37 cases with $k = 6$. With $\hat{\lambda}_{\text{aic}}$ and $k = 3$ an absolute bias larger than 6% was limited to stands 3, 6, and 30 (as per Tables 2, 3, 4, and 5 in Appendix 2); with $k = 6$ it was limited to stands 9, 40, and

Table 1 Distribution model(s) for r_k with best fit (in 600 replications) to observed distances in samples of size 15 and 30 (KS-test)

Best model(s)	Stand
Poi	1, 3, 7, 14, 15, 17, 18, 20, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 37, 38, 39, 42, 43, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55
Poi·sq, poi·tr	10,12, 26, 28, 34, 35, 44, 50, 57
Ig·poi	2, 4, 5, 8, 56
Poi·poi, uni·poi	1, 9, 16, 30
Birn·poi	6

Results are pooled across values of k and n. Multiple models in a row indicate models with similar D_{max} and AICc values. See "Appendix 1: maximum likelihood estimators of density (λ)" for model details

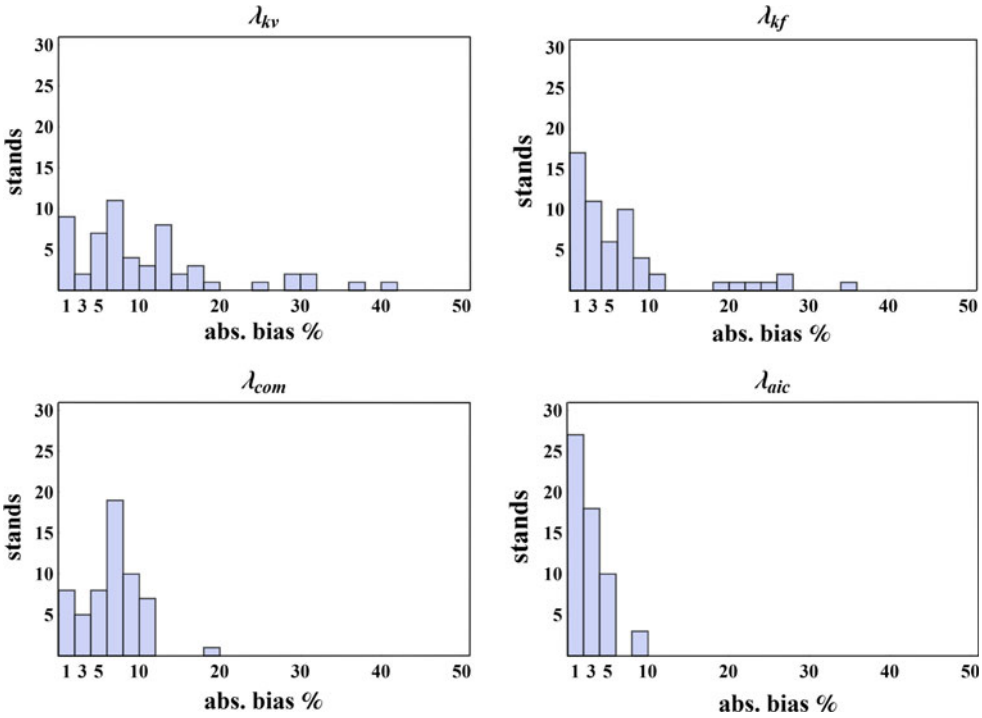

Fig. 3 *Histograms* of absolute bias with $k = 3$ in 58 stands for $\hat{\lambda}_{aic}$ and three comparison estimators $\left(\hat{\lambda}_{kv}, \hat{\lambda}_{kf}, \text{and } \hat{\lambda}_{com}\right)$

56. Note the lack of overlap between $k = 3$ and 6 in stands with an extreme bias. The spatial pattern of tree locations in the stands with an absolute bias >6% can be characterized as clustering at several (overlapping) scales, a positive skewness (>0.7) in the distribution of r_k, and a quadratic trend in Ripley's L_4 function (not shown). In these stands, four distribution models for r_k [poi·poi, uni·poi, ig·poi, and birn·poi cf. "Appendix 1: maximum likelihood estimators of density (λ)"] were given the largest weight in $\hat{\lambda}_{aic}$.

In six stands with a semi-regular pattern of stem locations (stands 10, 28, 34, 44, 57, and 58), the mean absolute bias of $\hat{\lambda}_{kv}$, $\hat{\lambda}_{kf}$, $\hat{\lambda}_{com}$, and $\hat{\lambda}_{aic}$ with $k = 3$ was 2.1, 1.7, 8.0, and 2.3%, respectively. With $k = 6$ the corresponding figures were 1.0, 1.4, 2.2, and 1.4%. Hence, in semi-regular stands where $k = 6$ the four estimators appear to incur an acceptable level of bias; with $k = 3$ only $\hat{\lambda}_{com}$ failed. In the six stands, the mixture models poi·tr, poi·hx, and poi·sq for r_k achieved the lowest AICc values in 74% of the simulated samples, with the remainder split between models poi and par·poi.

Seven stands (4, 25, 34, 37, 39, 46, and 50 as per Tables 2, 3, 4, and 5 in Appendix 2) were also used for stem density estimation with the RDE k-tree distance method (Nothdurft et al. 2010). A direct comparison of bias results is possible for $k = 6$. Accordingly, the average absolute bias was 1.5% for $\hat{\lambda}_{aic}$, 2.6% for $\hat{\lambda}_{com}$, 5.3% for $\hat{\lambda}_{kf}$ versus 0.7% for $\hat{\lambda}_{RDE}$. In five of the seven stands, the absolute bias of $\hat{\lambda}_{RDE}$ was the lowest overall.

To assess the potential for improvements to $\hat{\lambda}_{aic}$, for each setting of stand, k, n, and replication, we picked the density estimate (out of 16) closest to the actual (known) stand density; call it $\hat{\lambda}_{opt}$. With $k = 3$ the average (across stands) minimum absolute bias of $\hat{\lambda}_{opt}$ was 1.4%; with

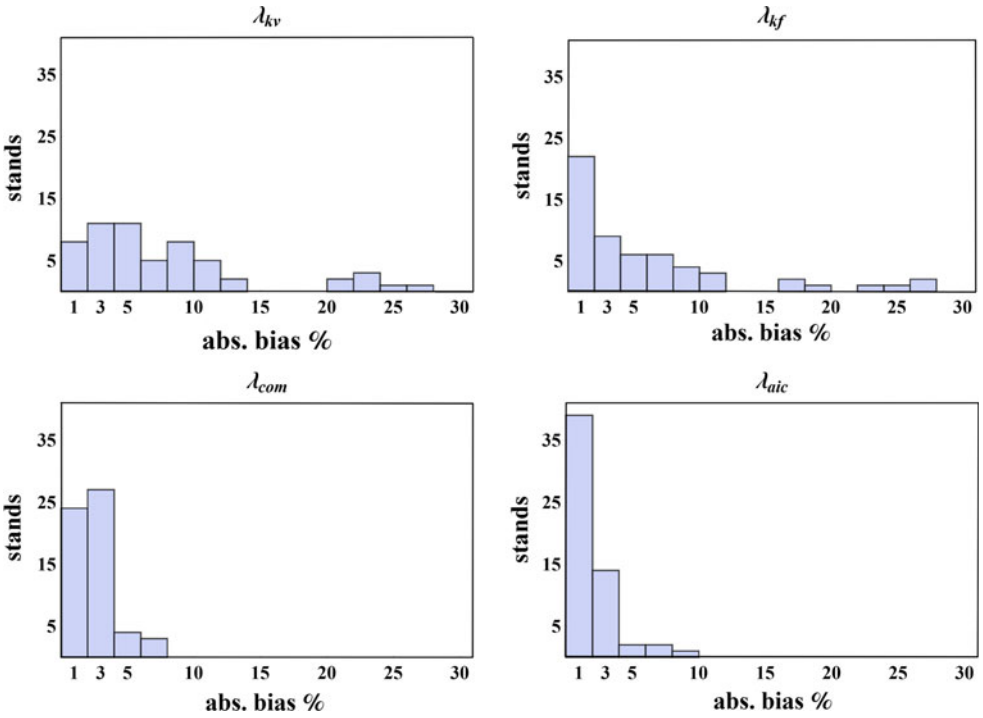

Fig. 4 *Histograms* of absolute bias with $k = 6$ in 58 stands for $\hat{\lambda}_{\text{aic}}$ and three comparison estimators $\left(\hat{\lambda}_{kv}, \hat{\lambda}_{kf}, \text{and } \hat{\lambda}_{\text{com}}\right)$

$k = 6$ it was 1.2%. In other words, the optimal bias is approximately two-thirds of the bias in $\hat{\lambda}_{\text{aic}}$. The among-stand standard deviations of $\left|\hat{\lambda}_{\text{aic}} - \hat{\lambda}_{\text{fix}}\right|$ and $\left|\hat{\lambda}_{\text{opt}} - \hat{\lambda}_{\text{fix}}\right|$ were similar (differences <5%). To improve $\hat{\lambda}_{\text{aic}}$, the range of component estimates of λ_{stand} needs to be widened. In the MC simulations, the chance of including λ_{stand} in the range spanned by the 16 estimates was, for the 'average' stand, 62% (range: 14–99%).

Root mean squared error (RMSE)

Averaged over stands, the relative RMSE of $\hat{\lambda}_{\text{aic}}$ was 25% ($k = 3$, $n = 15$), 21% ($k = 3$, $n = 30$), 16% ($k = 6$, $n = 15$), and 14% ($k = 6$, $n = 30$). The worst stand-specific RMSE was approximately 2.5 times larger than the average, while the best was approximately 2/3 of the average. Across the 58 (stands) × 2 (k) × 2 (n) settings, the RMSEs for $\hat{\lambda}_{\text{aic}}$ were about 40% larger than for $\hat{\lambda}_{\text{fix}}$ (correlation ≈ 0.9). RMSEs for $\hat{\lambda}_{\text{aic}}$ were also larger than those for $\hat{\lambda}_{\text{com}}$ (13% points) and $\hat{\lambda}_{kv}$ (10% points), but at par with RMSEs for $\hat{\lambda}_{kf}$. Estimator RMSE ranks across stands followed, by and large, the ranking of estimator bias.

Relative RMSEs reported by Nothdurft et al. (2010) for the seven stands (4, 25, 34, 37, 39, 46, and 50) where $k = 6$ are not directly comparable to results from this study due to a difference in sampling intensity of approximately 4 sample locations/ha. However, a rescaling to our sample sizes of 15 and 30 suggested that RMSE_{aic} were larger across the seven sites (approximately 32%). RMSE_{com} practically matched those obtained with $\hat{\lambda}_{\text{RDE}}$.

Precision

MC estimates of precision for sampling with fixed-area plots holding an average of k trees could be summarized by a linear function of k, n^{-1} and λ_{stand} as in:

$$\hat{\sigma}_{\text{fix}} \approx 0.0072 - 0.00087k + 0.0022n^{-0.5} + 0.1431\lambda_{\text{stand}}, \quad \hat{R}^2_{\text{adj}} = 0.91 \tag{7}$$

confirming that the within-stand standard deviation of density increases with stand density (Taylor 1971). MC estimates of precision for λ_{stand} with $k = 3$ were, on average, 30% larger than for $\hat{\lambda}_{\text{fix}}$ (correlation = 0.96). With $k = 6$ the difference was 19%. Precision of $\hat{\lambda}_{kv}$ was, as a rule, better than for $\hat{\lambda}_{\text{fix}}$ (approximately 9%) while precision values of $\hat{\lambda}_{\text{com}}$ were larger (33% for $k = 3$ and 20% for $k = 6$). Precision of $\hat{\lambda}_{kf}$ was better than with $\hat{\lambda}_{\text{aic}}$ (approximately 5%).

Coverage of nominal 95% confidence intervals

Nominal 95% Student's t confidence intervals computed from estimates of $\hat{\lambda}_{\text{aic}}$ and $\hat{\sigma}\left(\hat{\lambda}_{\text{aic}}\right)$ achieved an average coverage of 0.96 ($k = 3$, $n = 15$), 0.94 ($k = 3$, $n = 30$), 0.97 ($k = 6$, $n = 15$), and 0.94 ($k = 6$, $n = 30$). Inter-

Table 2 Summary of stands 1–13

Stand	Description [KS-test statistics]	Area (ha)	λ_0	λ_B	References
1	Barro Colorado Island Gatun Lake, Central Panama. Center for Tropical Forest Science. *Leguminosae*- and *Bombacaceae*-dominated wet tropical forest. Data from the 1990 census	50.0	7,361	7,369	Condit et al. (1998)
2	Forest DYNAMICS PLOT in a sub-montane evergreen broadleaf forest in Northeastern Taiwan	25.0	4,474	4,486	Sun et al. (2007)
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				
3	Miombo Woodlands savannah, n. Zambia	13.3	400	409	Kleinn and Vilčko (2006b)
4	Old-growth *Pinus Palustris* Wade Tract pine savanna in Thomas County, GA, USA. Center part of the 39.4-ha surveyed plot	4.0	146	161	Platt et al. (1988)
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				
5	As in 4, but limited to trees with a DBH $\geq$ 66 cm	27.0	46	49	
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				
6	As in 4, but limited to trees with a DBH < 66 cm	27.0	83	81	
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				
7	A natural stand of maple and hickory in East Lansing, MI, USA	1.0	1,217	1,219	Hatch et al. (1975)
8	As in 7, hickory only		703	652	
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				
9	As in 7, maple only		514	618	
	$[P(F(r_k)\|\mathrm{H}_0) < 0.05$ for $k = 3$ and 6]				
10	Norway spruce, Tharandter Wald (Ger.)	0.2	630	652	Illian et al. (2008, p. 404)
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				
11	Oak and beech, Manderscheid-198, Rheinland-Pfalz (Ger.)	0.6	381	359	Pommerening (2002)
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 6]$				
12	As in 11, oak only		125	117	
	$[P(F(r_k)\|\mathrm{H}_0) < 0.05$ for $k = 3$ and 6]				
13	As in 11, beech only		256	250	
	$[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]				

Stands with a rejection of the null hypothesis H_0: the empirical distribution function (F) of r_k is consistent with the distribution in a Poisson forest (csr of stem locations) are listed with details of the KS-test results. See text for details of test

Table 3 Summary of stands 14–24

Stand	Description [KS-test statistics]	Area (ha)	λ_0	λ_B	References
14	Douglas-fir, immature (North), Victoria Watershed (BC, Canada)	0.8	1,889	1,888	Trofymow et al. (1998)
15	Douglas-fir, immature (South) – –	0.5	883	945	
16	Douglas-fir, mature (North) – –	1.1	1,081	1,102	
	$[P(F(\mathrm{r_k})\|\mathrm{H}_0) < 0.05$ for $k = 3$ and 6]				
17	Douglas-fir, mature (South) – –	1.2	639	627	
18	Douglas-fir, old (North) – –	0.9	673	679	
19	Douglas-fir, old (South) – –	0.8	481	466	
	$[P(F(\mathrm{r_k})\|\mathrm{H}_0) < 0.05$ for $k = 6]$				
20	Tropical moist forest, Serra do Mare State Park, São Paulo St. (BRA). Plot A, elev. 0–50 m a.s.l.	1.0	1,630	1,664	Alves et al. (2010)
21	Tropical moist forest, Plot D, elev. 50–100 m a.s.l.	1.0	1,338	1,342	
22	Tropical moist forest, Plot E, elev. 50–100 m a.s.l.	1.0	1,244	1,234	
23	Tropical moist forest, Plot J, elev. 100–500 m a.s.l.	1.0	1,833	1,834	
24	Tropical moist forest, Plot K, elev. 500–1,200 m a.s.l.	1.0	1,866	1,837	

Table 4 Summary of STIPSI stands (Schöpfer 1967) (http://www.fva-bw.de/forschung/bui/stipsi_en.html)

Stand	Species [KS-test statistics]	Abt	Area (ha)	λ_0	λ_B
25	F/S/B	7	4.9	432	435
26	S	12	3.8	596	588
	$[P(F(\mathrm{r_k})\mid \mathrm{H_0}) < 0.05$ for $k = 3$ and 6]				
27	F/S/sH	13	5.2	318	325
28	B	15	5.0	171	167
29	B/O	16	4.1	121	126
	$[P(F(r_k)\mid H_0) < 0.05$ for $k = 6]$				
30	S	17	6.4	370	374
	$[P(F(r_k)\mid H_0) < 0.05$ for $k = 3$ and 6]				
31	mH	18	16.0	365	361
32	mH	19	10.5	337	338
33	mH	20	5.6	411	413
34	B	21	4.1	242	247
	$[P(F(\mathrm{r_k})\mid H_0) < 0.05$ for $k = 3$ and 6]				
35	P/B/O	22	5.1	297	306
	$[P(F(\mathrm{r_k})\mid H_0) < 0.05$ for $k = 3$ and 6]				
36	M/O/sH	27	6.6	226	228
	$[P(F(r_k)\mid H_0) < 0.05$ for $k = 3$ and 6]				
37	S	28	3.5	424	446
38	A/iH	29	16.1	342	338
39	F/S/P	30	7.8	839	846
40	B/O/iH	32	4.1	709	747
	$[P(F(r_k)\mid H_0) < 0.05$ for $k = 3$ and 6]				
41	O	33	4.2	685	705
	$[P(F(r_k)\mid H_0) < 0.05$ for $k = 3$ and 6]				
42	P	34	4.1	616	616
43	O/sH	40	13.0	499	508
44	P	50	4.0	379	377
	$[P(F(r_k)\mid H_0) < 0.05$ for $k = 3$ and 6]				
45	B/F	54	4.8	335	325
46	B/O/sH	61	4.1	483	492
47	O	64	5.2	945	952
48	O	65	4.2	551	558
49	S/F/sH	66	6.2	552	568
50	O/sH	67	6.5	882	889
51	O/sH	69	6.5	658	664
52	B/sH	84	3.1	441	413
53	S/F/sH	88	5.2	322	325
54	B/O/iH	90	6.2	357	360

All duplicate stem locations were dropped. *B Fagus sylvestris*, *F Abies alba*, *O Quercus petraea*, *M Acer platanus*, *P Pinus sylvestris*, *S Picea abies*, *mH* mixed hardwoods, *iH* intolerant hardwoods, *sH* shade tolerant hardwoods, *abt* STIPSI stand reference number. Minimum tree size of mapped trees: 7 cm at DBH

quartile ranges in coverage had a length of approximately 0.05 regardless of sampling design. Coverage below 0.85 was seen in just four combinations of stands × sampling design. The minimum coverage was 0.52.

Coverage of confidence intervals computed from estimates of $\hat{\lambda}_{\mathrm{fix}}$ and $\hat{\sigma}\left(\hat{\lambda}_{\mathrm{fix}}\right)$ was, as expected, close to their nominal value (mean 0.94, minimum = 0.56). But approximately 4% had coverage below 0.85. Results for $\hat{\lambda}_{\mathrm{com}}$ and $\hat{\lambda}_{kf}$ were similar (mean of 0.94, minimum of 0.31), and 11% of intervals had coverage of 0.85 or less. Results for $\hat{\lambda}_{kv}$ were generally poor (mean = 0.77).

Discussion

The appeal of composite estimators, like $\hat{\lambda}_{\mathrm{aic}}$ and $\hat{\lambda}_{\mathrm{com}}$, is a more constant performance across a broad application

Table 5 Summary of simulated stands

Stand	Name [KS-test statistics]	Area (ha)	λ_0	λ_B	Comment
55	Poisson	9.0	800	798	Csr
56	Matérn $[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]	9.0	1,581	1,584	$\lambda_{\text{parents}} = 0.02$, offspring (cluster) radius ~uniform (2, 8 m), no. of offspring per cluster ~betabinomial (5, 51, 90). Mean cluster size is ~8 (Picard et al. 2005)
57	Uniform $[P(F(r_k)\|H_0) < 0.05$ for $k = 3$ and 6]	9.0	300	296	Hard-core process with a minimum tree-to-tree distance is 4.0 m (Bondesson and Fahlten 2003)
58	Strauss $[P(F(r_k)\|H_0) < 0.05$ for $k = 6]$	1.0	100	103	Soft-core strauss process. Disk radius 3.0 m, permeability of disk: 0.4 (Strauss 1975)

spectrum than what is possible with single-component estimators (White et al. 2008). As demonstrated here and elsewhere, many *k*-tree density estimators may perform well in a narrow range of spatial patterns of stem locations (Clayton and Cox 1986; Delince 1986; Diggle 1977). For example, the $\hat{\lambda}_{kv}$ estimator by Kleinn and Vilčko (2006b) performs well in quasi-regular stands, but performance worsens (bias increases) dramatically as the pattern of stem locations becomes more and more irregular or clustered. In stands with a csr of stem location, the estimator by Morisita (1954) appears to be the best choice.

The proposed composite AIC-weighted *k*-tree stem density estimator appears to have mitigated a recalcitrant bias problem in *k*-tree density estimation to a level that may be considered acceptable in stand-level applications. A bias >6% appears to be restricted to sampling in: (1) strongly clustered subpopulations in multi-species or multi-layered stands or (2) stands with significant density gradients. With typical stand-level sample sizes (say 4 sample locations/ha), the precision of a stem density estimate will not be great, whether obtained with small, fixed-area plots or *k*-tree sampling with $k = 6$. Thus a typical bias in the 1–3% range will not seriously skew a statistical inference regarding the point estimate of stem density (Cochran 1977, p. 15). Overall, $\hat{\lambda}_{\text{aic}}$ presents a tangible improvement, in terms of bias, over otherwise relatively low-biased estimators like $\hat{\lambda}_{\text{com}}$ (Magnussen et al. 2011) and $\hat{\lambda}_{kf}$ (Kronenfeld 2009). The improvement was especially noteworthy for $k = 3$ because the bias in $\hat{\lambda}_{\text{aic}}$ did not display the expected strong increase with decreasing *k* (Kleinn and Vilčko 2006b; Kronenfeld 2009). In that sense $\hat{\lambda}_{\text{aic}}$ is robust. Currently, and limited to the results from seven stands, it appears that the stem-pattern reconstruction *k*-tree estimator by Nothdurft et al. (2010) remains the only nearly unbiased *k*-tree density estimator. Implementation issues and computation time (rate of convergence) will decide on the practical uptake of this estimator.

In other aspects (RMSE and coverage of 95% confidence intervals), $\hat{\lambda}_{\text{aic}}$ does not offer any advantage, at least not when compared to $\hat{\lambda}_{\text{com}}, \lambda_{kv}, \hat{\lambda}_{\text{kf}}$, and $\hat{\lambda}_{\text{RDE}}$. From a modeling perspective, the proposed composite estimator, $\hat{\lambda}_{aic}$, has more appeal than the model-based, adaptive composite estimator $\hat{\lambda}_{\text{com}}$ (Magnussen et al. 2011) because weights given to each component estimator are decided by an objective information criterion (self-weighting), thus eliminating the need for a model-based calibration.

The 11 novel density-dependent distribution models ("Appendix 1: maximum likelihood estimators of density (λ)") for the distance to the *k*-nearest tree from a random sample location expand the choices of *k*-tree density estimators available to an analyst. The three novel maximum likelihood *k*-tree density estimators for quasi-regular stands (Miyagawa 2009) should be attractive for applications in plantation forests and stands, subject to density management (Kleinn et al. 2009; Lynch and Rusydi 1999; Lynch and Wittwer 2003).

The success of any composite estimator depends on: (1) whether the true density lies within the range of individual component estimates or not, and (2) weighting each component proportional to a statistic that measures 'goodness-of-fit'. To (1) an increase in the chance that the range of component estimates includes the actual density from its current level of 0.62–1.00 requires adding density-dependent distribution models for r_k with less variance than that under csr, and possibly also left-skewed distribution models. Ten of the distribution models used here had a variance greater than that under csr and were all right-skewed with 'extra-Poisson' variance. The desired distribution may be constructed by a convolution of a pdf of r_k applicable to a regular configuration of tree locations with a pdf of a measurement error process (Carroll et al. 2009; 2011). Further research is needed to test whether this approach is computationally feasible. To (2) the low reliability of any small-sample 'goodness-of-fit' statistics remains a challenge in stand-level applications where sample sizes rarely, if ever, exceed 30 (Meeker et al. 1992; Vokoun et al. 2001). This challenge is compounded by the fact that optimizing the fit of data to an assumed distribution for r_k is not

equivalent to minimizing bias of the ensuing estimate of stem density. The inverse relationship between r_k and local stem density is the core of this enigma.

This study was limited to applications with a stand-level estimation of stem density. Forest enterprises may also conduct stratified sampling for estimation of stem density in a suite of strata. In this situation, the variance of distances r_k from a random sample location to the k-nearest tree will reflect a sum of within- and between-stand variation. Hence, the variance of stratum-specific distributions of r_k will be larger than the average within-stand variance. For this reason, the proposed estimator, $\hat{\lambda}_{\text{aic}}$, can be expected to perform even better in stratified random sampling.

Although the proposed new composite estimator did mitigate the bias problem in k-tree density estimates, it must be admitted that the bias reduction comes at the cost of a complex estimator and associated implementation issues. Overall, it seems that we are still in a position in which we do not know a priori which estimator to choose to minimize bias, since we usually do not have prior knowledge of the type of stem spatial distribution. And for forest inventory purposes at least, the bias that might be encountered in both density estimates and attributes per unit area estimates could be important.

Conclusions

In an extensive testing with simulated sampling in 58 stands (54 real, 4 simulated), the proposed composite k-tree density estimator—constructed as an AIC-weighted sum of 16 estimators of density—performed well in terms of bias. For a wide range of small-sample applications in forestry, the bias of the proposed estimator will be unimportant relative to the precision of a density estimated from a small sample. Eleven new k-tree density estimators, suited for stands with either a regular or an irregular pattern of stem locations, enriched the options of estimation.

Acknowledgments Drs. F. Lutz and A. Nothdurft provided many valuable comments on an earlier draft of this manuscript. Stem location data from Brazil were taken from the project "Functional Gradient (FAPESP 03/12595-7)". Permit: COTEC/IF 41.065/2005 and IBAMA/CGEN 093/2005 supported by the State of São Paulo Research Foundation (FAPESP) and kindly made available by Dr. L. F. Alves. Data from stands in Table B3 were kindly made available through the assistance of Dr. A. Nothdurft. A thank you also goes to Drs C. Kleinn for the Miombo data, I. F. Sun for the Fushan data, and W. J. Platt for the Tall pines data. Data from the Smithsonian Barro Colorado Island research plot (Condit and Hubbell 1998; Hubbell et al. 1999) were downloaded from the Center for Tropical Forest Research site http://ctfs.si.edu (last accessed May 22nd, 2009). Two anonymous journal reviewers provided helpful comments and critique to an earlier version of this manuscript.

Appendix 1: maximum likelihood estimators of density (λ)

The probability density function (pdf) for the Euclidean distance from a randomly selected sample location to the k-nearest tree in an area (stand) of interest is assumed to follow a parametric model $f_{\text{model}}(r_k|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of model-dependent parameters to be estimated from n observed distances by conventional maximum likelihood techniques (Gould et al. 2010, p. 13). A known function (g) defines the relationship between density λ and $\boldsymbol{\theta}$. The desired maximum likelihood estimator (MLE) of density $\hat{\lambda}_{\text{model}}$ is obtained by applying g to the MLE of $\boldsymbol{\theta}$. Estimates of precision were computed by application of the delta-technique (Oehlert 1992) to $g\left(\hat{\theta}\right)$ and the estimated covariance matrix of the parameters.

The basic distribution model for r_k applies to a csr (Poisson) of tree locations (Thompson 1956). It is

$$f_{\text{poi}}(r_k|\lambda) = 2(\pi\lambda)^k r_k^{2k-1} \exp\left(-\pi\lambda r_k^2\right)[(k-1)!]^{-1} \tag{8}$$

with the following bias-corrected estimator of density (Pollard 1971).

$$\hat{\lambda}_{\text{poi}} = \frac{nk-1}{nk}\hat{\lambda}_{\text{MLE}}, \quad \hat{\lambda}_{\text{MLE}} = nk\left(\pi \sum_{i=1}^{n} r_{k.i}^2\right) \tag{9}$$

All other distribution models in this study are derived by: (1) assuming that r_k is distributed as per $f_{\text{poi}}(r_k|\lambda)$ but with λ varying at random across space according to a parameter-mixing distribution $h(\lambda|\phi)$, or (2) assuming a mixture distribution of two density-dependent distributions of r_k, of which the Poisson distribution model in (8) is one.

Parameter-mixture distributions

We examined 46 potential parameter-mixing distributions $h(\lambda|\phi)$ with support on the domain of positive real numbers. Eight had a suitable closed-form solution to the convolution integral

$$f_{\text{model}\cdot i}(r_k|\phi) = \int_0^\infty f_{\text{poi}}(r_k|\lambda)h(\lambda|\phi) \tag{10}$$

Others had either no closed-form solution or were computationally too demanding and time-consuming to be considered for this study.

By assuming that r_k is distributed as under a Poisson model and λ is distributed uniformly on the closed interval $[\lambda_{\min}, \lambda_{\max}]$, the following parameter-mixture *pdf* applies to r_k:

$$f_{\text{uni·poi}} = \frac{2r_k^{2k-3}r_k^{-2k}}{\pi(\lambda_{\max}-\lambda_{\min})\Gamma(k)} \left(\Gamma\left(1+k,\pi r_k^2\lambda_{\min}\right)-\Gamma\left(1+k,\pi r_k^2\lambda_{\max}\right)\right) \quad (11)$$

where $\Gamma(x)$ is the gamma function with argument x. The ensuing bias-corrected estimator of density is $\hat{\lambda}_{\text{uni·poi}} = 0.5(nk-1)n^{-1}k^{-1}\left(\hat{\lambda}_{\min}+\hat{\lambda}_{\max}\right)$.

With λ assumed distributed as a random gamma variate with parameters α and β, the resulting parameter-mixture distribution of r_k is (Magnussen et al. 2008)

$$f_{\text{gam·poi}}(r_k|\alpha,\beta) = \frac{2\pi^k\Gamma(k+\alpha)(\beta r_k)^k(\pi\beta r_k^2+1)^{-\alpha-k}}{(k-1)!\Gamma(\alpha)} \quad (12)$$

which leads to the following density estimator: $\hat{\lambda}_{\text{gam·poi}} = \hat{\alpha}\hat{\beta}$.

With λ assumed distributed as a random Erlang-variate with parameters γ and η, the parameter-mixture of r_k is

$$f_{\text{poi·erl}}(r_k|\gamma,\eta) = \frac{2\eta^\gamma\pi^k r_k^{2k-1}(\eta+\pi r_k^2)^{-a-ki}\Gamma(\gamma+k)}{(k-1)!\Gamma(\gamma)} \quad (13)$$

and the estimator of density is $\lambda_{\text{erl·poi}} = \hat{\gamma}\hat{\eta}$.

A multiplicative convolution of λ with a random variable δ having a triangular distribution on the closed interval $[\delta_{\min},\delta_{\max}]$ and expectation 1 yields the following parameter-mixture pdf for r_k

$$f_{\text{tri·poi}}(r_k|\lambda,\delta_{\min},\delta_{\max}) = \frac{4(P1+P2)}{\pi^2\lambda^2 r_k^5\Gamma(k)(\delta_{\max}-1)},$$
$$P1 = \frac{\pi\lambda r_k^2\delta_{\max}\left(\Gamma\left(k+1,\pi\lambda r_k^2\delta_{\min}\right)-\Gamma\left(k+1,\pi\lambda r_k^2\delta_{\max}\right)\right)+\Gamma\left(k+2,\pi\lambda r_k^2\delta_{\max}\right)-\Gamma\left(k+2,\pi\lambda r_k^2\delta_{\min}\right)}{\delta_{\max}-\delta_{\min}},$$
$$P2 = \frac{\pi\lambda r_k^2\left(\left(\Gamma(k+1,\pi\lambda r_k^2\delta_{\min})-\Gamma\left(k+1,\pi\lambda r_k^2\right)\right)-\Gamma\left(k+2,\pi\lambda r_k^2\delta_{\min}\right)+\Gamma\left(k+2,\pi\lambda r_k^2\right)\right.}{\delta_{\min}-1} \quad (14)$$

with estimator $\hat{\lambda}_{\text{tri·poi}} = \hat{\lambda}$.

Taking λ to be distributed as a random inverse Gaussian variate with parameters μ and σ, the parameter-mixture of r_k is

$$f_{\text{ig·poi}}(r_k|\mu,\sigma) = 2\sqrt{2}\pi^{k-\frac{1}{2}}\sqrt{\sigma}\mu^{k-\frac{1}{2}}r_k^{2k-1}\exp\left(\sigma\mu^{-1}\right)\left(\frac{\sigma}{2\pi\mu^2 r_k^2+\sigma}\right)^{\frac{k}{2}-\frac{1}{4}} \times K_{0.5-k}\left(\left(\mu^{-1}\sqrt{\left(2\pi r_k^2\mu^2+\sigma\right)}\right)\right)[(k-1)!]^{-1} \quad (15)$$

where K is the Bessel function of the second kind. We have $\hat{\lambda}_{\text{ig·poi}} = \hat{\mu}$.

A Laplace parameter-mixture distribution gave rise to

$$f_{\text{lapl·poi}}(r_k|\tau,\upsilon) = \pi^k\exp\left(-\tau\upsilon^{-1}\right)r_k^{2k-1}\left(\pi\upsilon r_k^2-1\right)^{-k-1} \times\left(\upsilon^{k+1}\left(\Gamma(k+1)-\Gamma\left(k+1,\tau\left(\upsilon\pi r_k^2-1\right)\upsilon^{-1}\right)\right)\Gamma^{-1}(k)\tau\right. + \exp\left(2\tau\upsilon^{-1}\right)\left(\tau\left(\pi\upsilon r_k^2-1\right)\right)^{k+1}E_{-k}\left[\tau\left(\pi r_k^2+\upsilon^{-1}\right)\right] \times\upsilon^{-1}\Gamma^{-1}(k) \quad (16)$$

where E_{-k} [] is the exponential integral function and $\hat{\lambda}_{\text{lapl·poi}} = \hat{\tau}$.

A Pareto parameter-mixture distribution leads to

$$f_{\text{par·poi}}(r_k|\psi,\zeta) = 2\pi^\zeta\zeta\psi^\zeta r_k^{2\zeta-1}\Gamma\left(k-\zeta,\pi\psi r_k^2\right)\Gamma^{-1}(k) \quad (17)$$

with the following density estimator $\hat{\lambda}_{\text{par·poi}} = \hat{\psi}\hat{\zeta}\left(\hat{\zeta}-1\right)^{-1}$. During applications, it became clear that computing times would become excessive despite the apparent simplicity of the expression. All estimates of ζ were unstable and numerically very large. To fix matters, we decided to keep ζ constant at a value of 200. Thus, this distribution has only one free parameter.

Finally, with a Birnbaum–Saunders parameter-mixture distribution, we got

$$f_{\text{birn·poi}}(r_k|\omega,\xi) = \sqrt{2}\pi^{k-\frac{1}{2}}\tilde{a}^{\frac{1}{\omega^2}}r_k^{2k-1}\left(\xi\left(2\pi\omega^2 r_k^2+\xi\right)\right)^{\frac{1}{4}(-2k-1)} \times\left(\sqrt{\xi\left(2\pi\omega^2 r_k^2+\xi\right)}K_{k-0.5}\left(\omega^{-2}\sqrt{2\pi\omega^2 r_k^2\xi^{-1}+1}\right)\right. + \xi K_{k+0.5}\left(\omega^{-2}\sqrt{2\pi\omega^2 r_k^2\xi^{-1}+1}\right)\Big) \times\sqrt[-0.5]{\xi}\omega^{-1}\Gamma^{-1}(k) \quad (18)$$

and $\hat{\lambda}_{\text{birn·poi}} = (2+\hat{\omega}^2)2^{-1}\hat{\xi}^{-1}$.

Appendix 2

A brief description and reference is provided for the 54 actual tree stem location patterns (Tables 2, 3, 4, and 5) and the 4 simulated patterns (Table 5). The density estimator λ_0 is the total number of recorded stems divided by the total area of a stand (unit: stems $\times$ ha^{-1}). The density estimator λ_B refers to the density of stems located at a distance from the stand boundary equal to or greater than the radius in a fixed-area plot holding, on average, $k = 6$ trees.

References

Alves LF, Vieira SA, Scaranello MA, Camargo PB, Santos FAM, Joly CA, Martinelli LA (2010) Forest structure and live aboveground biomass variation along an elevational gradient of tropical Atlantic moist forest (Brazil). For Ecol Manag 260:679–691

Augustin NH, Musio M, von Wilpert K, Kublin E, Wood SN, Schumacher M (2009) Modeling spatiotemporal forest health monitoring data. J Am Stat Assoc 104:899–911

Barabesi L, Marcheselli M (2004) Design-based ranked set sampling using auxiliary variables. Environ Ecol Stat 11:415–430

Bondesson L, Fahlten J (2003) Mean and variance of vacancy for hard-core disc processes and applications. Scand J Stat 30:797–816

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. Sociol Methods Res 33:261–304

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

Carroll RJ, Delaigle A, Hall P (2009) Nonparametric prediction in measurement error models. J Am Stat Assoc 104:993–1003

Carroll RJ, Delaigle A, Hall P (2011) Testing and estimating shape-constrained nonparametric density and regression in the presence of measurement error. J Am Stat Assoc 106:191–202. doi: 10.1198/jasa.2011.tm10355

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 (1998) Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a comparison with other plots. Springer, Berlin

Condit R, Foster RB, Hubbell SP, Sukumar R, Leigh EG, Manokaran N, de Lao LS, Dallmeier F, Comiskey J (1998) Assessing forest diversity from small plots: calibration using species-individual curves from 50 ha plots. In: Dallmeier F, Comiskey JA (eds) Forest biodiversity research, monitoring and modeling: conceptual background and old world case studies. Parthenon, Paris, pp 247–268

Cox DD, Cox LH, Ensor KB (1997) Spatial sampling and the environment: some issues and directions. Environ Ecol Stat 4:219–233

De Vries PG (1986) Sampling theory for forest inventory. A teach yourself course. Springer, Berlin

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

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

Fattorini L, Greco L, Naddeo S (2001) The use of χ^2 and Kolmogorov–Smirnov statistics in permutation-based pairwise comparison. Metron Int J Stat 59:11–20

Fehrmann L, Gregoire T, Kleinn C (2011) Triangulation based inclusion probabilities: a design-unbiased sampling approach. Environ Ecol Stat. doi:10.1007/s10651-011-0177-9

Good P (1993) Permutation tests. Springer, New York

Gould W, Pitblado J, Poi B (2010) Maximum likelihood estimation with Stata, 4th edn. Stata Press, College Station

Greene WH (2003) Econometric analysis, 5th edn. Prentice Hall, Upper Saddle River

Hatch CR, Gerrard DJ, Tappeiner JC II (1975) Exposed crown surface area: a mathematical index of individual tree growth potential. Can J For Res 5:224–228

Hubbell SP, Foster RB, O'Brien ST, Harms KE, Condit R, Wechsler B, Wright SJ, Loo de Lao S (1999) Light gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science 283:554–557

Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns. Wiley, Chichester

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 PP (1963) Geometrical probability. Charles Griffin, London

Kint V, De Wulf R, Noël 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) Design-unbiased estimation for point-to-tree distance sampling. Can J For Res 36:1407–1414

Kleinn C, Vilčko F (2006b) A new empirical approach for estimation in *k*-tree sampling. For Ecol Manag 237:522–533

Kleinn C, Vilčko F, Fehrmann L, Hradetzky J (2009) About the evaluation of *k*-tree samples. Zur Auswertung der *k*-baum-probe. Allgemeine Forst- und Jagdzeitung 180:228–237

Koehler E, Brown E, Haneuse J-PA (2009) On the assessment of Monte Carlo error in simulation-based statistical analyses. Am Stat 63:155–162

Kronenfeld BJ (2009) A plotless density estimator based on the asymptotic limit of ordered distance estimation values. For Sci 55:283–292

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

Lynch TB, Rusydi R (1999) Distance sampling for forest inventory in Indonesian teak plantations. For Ecol Manag 113:215–221

Lynch TB, Wittwer RF (2003) *n*-Tree distance sampling for per-tree estimates with application to unequal-sized cluster sampling of increment core data. Can J For Res 33:1189–1195

Magnussen S, Kleinn C, Picard N (2008) Two new density estimators for distance sampling. Eur J For Res 127:213–224

Magnussen S, Fehrman L, Platt W (2011) An adaptive composite density estimator for distance sampling. Eur J For Res 131:307–320

McIntire EJB, Fajardo A (2009) Beyond description: the active and effective way to infer processes from spatial patterns. Ecology 90:46–56

Meeker WQJ, Escobar LA, Hill DA (1992) Sample sizes for estimating the Weibull hazard function from censored samples. IEEE Trans Reliab 41:133–138

Miyagawa M (2009) Order distance in regular point patterns. Geog Anal 41:252–262
Morisita M (1954) Estimation of population density by spacing method. In: Contributions from the Department of Biology, Faculty of Science, vol 69, pp 187–197
Nothdurft A, Saborowski J, Nuske RS, Stoyan D (2010) Density estimation based on k-tree sampling and point pattern reconstruction. Can J For Res 40:953–967
Oehlert GW (1992) A note on the delta method. Am Stat 46:27–29
Owen AB (1995) Nonparametric likelihood confidence bands for a distribution function. J Am Stat Assoc 90:516–521
Parker KR (1979) Density estimation by variable area transect. J Wildl Manag 43:484–492
Patil SA, Kovner JL, Burnham KP (1982) Optimum nonparametric estimation of population density based on ordered distances. Biometrics 35:597–604
Persson O (1964) Distance methods: the use of distance measurements in the estimation of seedling density and open space frequency. Studia Forestalia Suecica 15:1–68
Picard N, Kouyaté AM, Dessard H (2005) Tree density estimations using a distance method in Mali savanna. For Sci 51:7–18
Platt WJ, Evans GW, Rathbun SL (1988) The population dynamics of a long-lived conifer (*Pinus palustris*). Am Nat 131:491–525
Pollard JH (1971) On distance estimators of density in randomly distributed forests. Biometrics 27:991–1002
Pommerening A (2002) Approaches to quantifying forest structures. Forestry 75:305–324
Popescu SC, Wynne RH, Nelson RF (2003) Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Can J Rem Sens 29:564–577
Prodan M (1968) Punktstichprobe für die Forsteinrichtung (in German). Forst- und Holzwirt 23:225–226
Ratkowsky DA (1983) Nonlinear regression modeling. Marcel Dekker, New York
Saborowski J, Šmelko S (1998) Zur Auswertung von Stichprobeninventuren mit variablen Probeflächengrößen. Allgemeine Forst- und Jagdzeitung 169:71–75
Sandrine R, Philippe S, Michel A (2003) Applying distance-sampling methods to spotlight counts of red foxes. J Appl Ecol 40:32–43
Schöpfer W (1967) Ein Stichprobensimulator für Forschung und Lehre. Allg F Jagdztg 138:267–273
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
Staupendahl K (2008) The modified six-tree-sample—a suitable method for forest stand assessment. Allg F Jagdztg 179:21–33
Steinke I, Hennenberg KJ (2006) On the power of plotless density estimators for statistical comparisons of plant populations. Can J Bot 84:421–433
Strauss DJ (1975) A model for clustering. Biometrika 62:467–475
Sun IF, Chen YY, Hubbell SP, Wright SJ, Noor N (2007) Seed predation during general flowering events of varying magnitude in a Malaysian rain forest. J Ecol 95:818–827
Sutherland WJ (2006) Ecological census techniques—a handbook. Cambridge University Press, Cambridge, pp 141–142
Taylor ER (1971) Aggregation, variation, and the mean. Nature 189:2732–2735
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. Information Report BC-X-374. Canadian Forest Service, Victoria BC, Canada, 47 pp
Tscheschel A, Stoyan D (2006) Statistical reconstruction of random point patterns. Comput Stat Data Anal 51:859–871
Vokoun JC, Rabeni CF, Stanovick JS (2001) Sample-size requirements for evaluating population size structure. N Am J Fish Manag 21:660–665
White NA, Engeman RM, Sugihara RT, Krupa HW (2008) A comparison of plotless density estimators using Monte Carlo simulation on totally enumerated field data sets. BMC Ecol 8:6. doi:10.1186/1472-6785-8-6