Landscape Ecol (2011) 26:59–68
DOI 10.1007/s10980-010-9533-7

RESEARCH ARTICLE

Spatial pattern corrections and sample sizes for forest density estimates of historical tree surveys

Brice B. Hanberry · Shawn Fraver · Hong S. He · Jian Yang · Dan C. Dey · Brian J. Palik

Received: 23 April 2010 / Accepted: 9 September 2010 / Published online: 22 September 2010

Abstract The U.S. General Land Office land surveys document trees present during European settlement. However, use of these surveys for calculating historical forest density and other derived metrics is limited by uncertainty about the performance of plotless density estimators under a range of conditions. Therefore, we tested two plotless density estimators, developed by Morisita and Pollard, for two, three, and four trees per survey point under simulated ranges of tree densities, non-uniform densities, and different tree spatial distributions. Based on these results, we developed estimator corrections and determined number of survey points needed for reliable density estimates. The Morisita estimator was accurate for densities ranging from 5 to 1,000 trees per unit area, non-uniform densities, random and regular spatial distribution, and outperformed the Pollard estimator. Estimators using points with two or three trees did need a simple correction to account for overestimation. Likewise, for clustered distributions, depending on the number of trees per survey point and the amount of clustering, there should be adjustment for a range of under and overestimation. Sample sizes for survey points with three or four trees should be at least 200 survey points, and 1,000 survey points will have density estimates within ±10% tolerance range of actual density. For survey points with two trees, the minimum sample size should be 600 survey points, and 2,000 survey points should be the target value. These results provide guidelines for researchers to improve density estimates of historical forests.

Keywords General Land Office · Morisita · Point-centered quarter · Pollard · Presettlement forests · Public Land Survey

Electronic supplementary material The online version of this article (doi:10.1007/s10980-010-9533-7) contains supplementary material, which is available to authorized users.

B. B. Hanberry (✉) · H. S. He
Department of Forestry, University of Missouri, 203 Natural Resources Building, Columbia MO 65211, USA
e-mail: hanberryb@missouri.edu

S. Fraver · B. J. Palik
USDA Forest Service, Northern Research Station, Grand Rapids MN 55744, USA

J. Yang
Institute of Applied Ecology, Chinese Academy of Sciences, 110016 Shenyang, China

D. C. Dey
USDA Forest Service, Northern Research Station, University of Missouri, 202 Natural Resources Building, Columbia MO 65211, USA

Introduction

Ecological management plans and forestry prescriptions for restoration of native ecosystems increasingly need historical information on forest community structure and composition, including assessments of the natural range of variability. Early land surveys conducted under the auspices of the General Land Office (GLO) provide a record of historical vegetation across much of the United States at the time of European settlement. Although the intent of the GLO was that of property division, the records included information about 'bearing' trees (species, diameter, distance, and bearing from surveying boundaries). These data lend themselves to numerous applications related to understanding historical reference conditions, however, uncertainty exists in the use of GLO records for these purposes. Problems due to assumptions and surveyor bias have been well-documented (e.g., Bourdo 1956; Almendinger 1996; Bouldin 2008). Bias or imprecision in initial density estimates will propagate through derived metrics, such as basal area, biomass, and carbon storage. Without correction of these problems, researchers may produce questionable results that are unsuitable for guiding ecological management and restoration.

Early surveys of the United States roughly follow an angle-order design, such as the point-centered quarter, which involves dividing the area around each survey point into four equal quarters, and measuring the distance, species, and diameter at breast height of the nearest tree in two to four quadrants. Therefore, any density estimation (i.e., number of trees per unit area) for historical forests must be derived from a plotless density estimator. Existing estimators include formulas proposed by Cottam and Curtis (1956), Morisita (1957), and Pollard (1971). Indeed, most researchers used density estimates derived from the original or modified Cottam and Curtis estimator (e.g., Ward 1956; Nelson 1997; He et al. 2000; Bolliger and Mladenoff 2005). However, Pollard (1971) showed that the Cottam and Curtis (1956) estimator was biased mathematically and Bouldin (2008) illustrated poor performance of the Cottam and Curtis estimator for areas with non-uniform density.

There are fundamental challenges in tree density estimation from GLO data beyond the choice of estimator: the assumption of randomness for tree spatial location, spatial unevenness of density, wide density ranges, adjustments for fewer than four trees per point, and minimum number of points needed to generate a robust result. Estimators assume complete spatial randomness in tree locations, which does not always hold true for forests. Morisita's estimator in particular addresses this assumption, because it reduces the effects of non-random spatial distribution by estimating density at each survey point and then averaging the densities. However, this method has not been widely adopted in most historical forest reconstructions using GLO data. In addition, tree density can vary greatly both in magnitude and in spatial intensity (i.e., number of points per unit area). Such density variations are dictated by the underlying physical environments and past disturbance history (Schulte et al. 2005). Research has not explored whether existing estimators are suitable over a range of tree densities and with variation in tree density. Although in theory estimators are flexible for number of quadrants with trees, how well the estimators perform when applied to fewer than four trees per survey point has not been studied. Lastly, researchers need to know what sample size (i.e., number of survey points) is required for robust tree density estimations from GLO data (see Bouldin 2008).

Our objective in this study was to further improve forest density estimates derived from GLO data. Specifically, we examined two seemingly unbiased estimators, the Morisita (1957) and Pollard (1971) estimators, for two, three, and four trees per survey point (1) with ranges of tree densities (from 5 to 1,000 trees/area unit), (2) with variability in tree density (non-uniform density), and (3) with different tree spatial distributions (random, regular, and clustered). Thus, our study covered the full range of forest conditions likely encountered by early surveyors. Based on these results, we developed adjustments that produce more accurate density estimates. Finally, we determined the appropriate number of survey points needed for reliable density estimates.

Methods

GLO data

Detailed descriptions of the GLO survey methods can be found in White (1983); we provide here a brief overview. Unsettled territories were divided into townships measuring 6 miles (9.6 km) on a side. Townships were further divided into one-square-mile (1.6×1.6 km) sections. Surveyors typically set wooden posts at the intersection of section lines, as well as the midpoints between section corners. At each survey point, two to four bearing trees were typically marked for the purpose of relocating corner stakes. Though surveyor instructions varied somewhat, they usually required that bearing trees be selected in different quadrants (NE, SE, SW, NW) relative to the survey point. For each of these trees, the species, diameter, distance and bearing to survey point was recorded, thus providing the basic data set addressed here.

Tree density estimators

Our objective was to assess the performance of the Morisita (1957) and Pollard (1971) tree density estimators when applied to GLO survey data. The two estimators are as follows:Pollard density estimator:

$$\lambda = \frac{q(qn-1)}{\pi \sum_{i=1}^{n} \sum_{j=1}^{q} r_{ij}^2} \tag{1}$$

Morisita density estimator:

$$\lambda = \frac{(q-1)}{\pi n} \sum_{i=1}^{n} \frac{q}{\sum_{j=1}^{q} r_{ij}^2} \tag{2}$$

where λ (density) is the number of trees/unit area, q is the number of quadrants with surveyed trees (2, 3, or 4), n is the number of points, and r is the survey point-to-tree distance.

Design of simulation experiments

To assess the accuracy of these density estimators, we repeatedly simulated forest structure using a range of reasonable tree densities and tree spatial patterns (random, regular, and clustered; Table 1). Instead of simulating a large forested landscape (see Bouldin 2008), we simply centered a 'survey point' within a square plane (of 100 × 100 area units), populated the plane with simulated trees, and determined density using the Morisita and Pollard estimators. Such a design allowed us to produce unit-independent results that are scalable over various areas. For each simulation, tree locations were generated randomly using Python (http://docs.python.org; see code in Appendix A in Electronic supplementary material). The number of simulated trees followed a Poisson distribution, with the constraint that at least four trees be located per simulation. If two trees were within 0.25 units of each other, the tree generated last was eliminated. Because GLO surveyors typically recorded either two or four trees per survey point, we likewise ran our simulations based on two or four trees. We also evaluated the special case of three trees per survey point, as these are occasionally encountered in survey data, although we caution researchers that these points with three trees at times result from data errors, such as overlapping surveys. This process was repeated under various predetermined density–spatial pattern combinations, allowing the comparisons of estimated and simulated (i.e., actual known) densities. To test a range of densities, we ran 60 trials of 600 simulated survey points, calculating a density estimate for each trial of 600 points, for each density and number of trees per survey point. We then determined the necessary number of survey points to achieve specific tolerances.

We first assessed the estimators at varying densities with random spatial distribution. We generated densities ranging from 5 to 25 trees/unit area in steps of 5 (i.e., 5, 10, 15, 20, and 25); 30–100 trees/unit area in steps of 10; and 200–1,000 trees/unit area in steps of 200, for a total of 18 densities tested. Although these target densities were set, the number of simulated trees was drawn from a Poisson distribution, which introduced variability in the simulated densities (mean maximum trees per survey point was two to three times greater than the mean minimum). Further, at the lowest densities, quadrants occasionally contained no trees. In these cases, we assigned trees to the vacant quadrant using a random bearing and a distance of 70 m. We chose this distance because it represents the tail of the distribution of actual distances surveyors recorded in a dataset of Missouri GLO trees (i.e., 2% of recorded distances were 70 m or greater), and because simulated density estimates stabilized more quickly than other values when using this distance. Filling these vacant quadrants additionally increased variation in simulated densities among points.

Recognizing, however, that survey points were unlikely to have uniform tree densities throughout a landscape (as simulated above, albeit with low variability around a set density), we also varied the density within each simulation trial, using two levels of variability, and again with random spatial distribution. For moderately variable densities, we

Table 1 Design overview for testing Morisita and Pollard density estimators with two to four trees per survey point, simulating a density range for random, regular, and clustered distributions, and non-uniform density for random distribution

	Distribution		
	Random	Regular	Clustered
Uniform density range (trees/unit area)	5–1,000[a]	5–1,000	20–1,000
Non-uniform density low, high (trees/unit area)	5, 10	10–100[b]	20–100[b]
	5, 15		
	5, 20		
	5, 25		
	5, 50		
	5, 100		
	5, 250		
	20, 40		
	20, 60		
	20, 80		
	20, 100		
	20, 200		
	20, 500		
	20, 1,000		

Sample size determination occurred for the density range and random distribution, and sample size testing was for non-uniform density and regular or clustered distributions

[a] Sample size determination

[b] Sample size test for Morisita estimator only

simulated densities such that each survey point had an equal probability of containing 5 trees/unit area and either 2, 3, 4, or 5 times greater and then 20 trees/unit area and 2, 3, 4, or 5 times greater. This resulted in a ratio of mean maximum to mean minimum trees per survey point of 5–16. For highly variable densities, we simulated densities such that each survey point had an equal probability of 5 trees/unit area and either 10, 25, or 50 times greater, and 20 trees/unit area and either 10, 25, or 50 times greater. This resulted in a ratio of mean maximum to mean minimum trees per survey point of 30–137. Thus, we simulated random spatial patterns under three levels of density variability: low (the above random distribution), moderate with eight levels, and high variability with six levels.

We further tested the estimators under both regular and clustered tree distributions (Fig. 1). The regular distribution was formed by randomly generating integer points only, effectively forming a grid, and randomly adding or subtracting a random offset value between 0 and 1 continuous (i.e., decimal) units, allowing the gridlines to move in relation to the center point. The random presence of points along the gridlines along with offset of the gridlines allowed us to characterize a forest that had uniform spacing and also was shaped by natural processes. We used the same density levels as for random spatial distribution (5–25 trees/unit area in steps of 5; 30–100 trees/unit area in steps of 10; and 200–1,000 trees/unit area in steps of 200, for a total of 18 densities). We evaluated the clustered distribution by varying both the number of clusters (5, 10, 25, 50, and 100 clusters/unit area) and the mean number of trees per cluster (4, 6, 8, or 10), yielding densities of 20–1,000 trees/unit area (e.g., 5 clusters $\times$ 4 trees per cluster $=$ 20 trees/unit area) and 20 density levels. The number of clusters and trees per cluster were drawn from a Poisson distribution. The cluster radius was varied by controlling the maximum distance a "child" tree location could be placed relative to the original ("parent") tree. Child trees were added by random bearing, with maximum radii set at 10 units, and then 15 units.

To evaluate the density estimators under various trials, we used the ratio of the density estimate mean and the actual simulation density mean as an intuitive gauge of performance. We also determined the bias and precision of the density estimators by comparing estimated and simulated (i.e., actual) densities. Bias was considered to be the mean difference between the estimated and simulated densities, and precision was taken to be standard deviation of the differences.

Finally, we used the first simulation, for varying densities with random spatial distribution, to determine the minimum number of survey points needed to achieve a specified level of confidence in density estimates. We report the minimum number as the threshold below which ca. 10 or more trials fell outside the tolerance limit. We evaluated tolerances at ±5%, 10%, and 20% of the simulated means. We also examined the effects on sample size of 1) high density variability (any density from 10 to 100 trees/unit area) in combination with regular spatial distribution and 2) high density variability (any density from 20 to 100 trees/unit area) in combination with clustered spatial distribution. For clustered distribution, we set the mean number of clusters at 10 and the

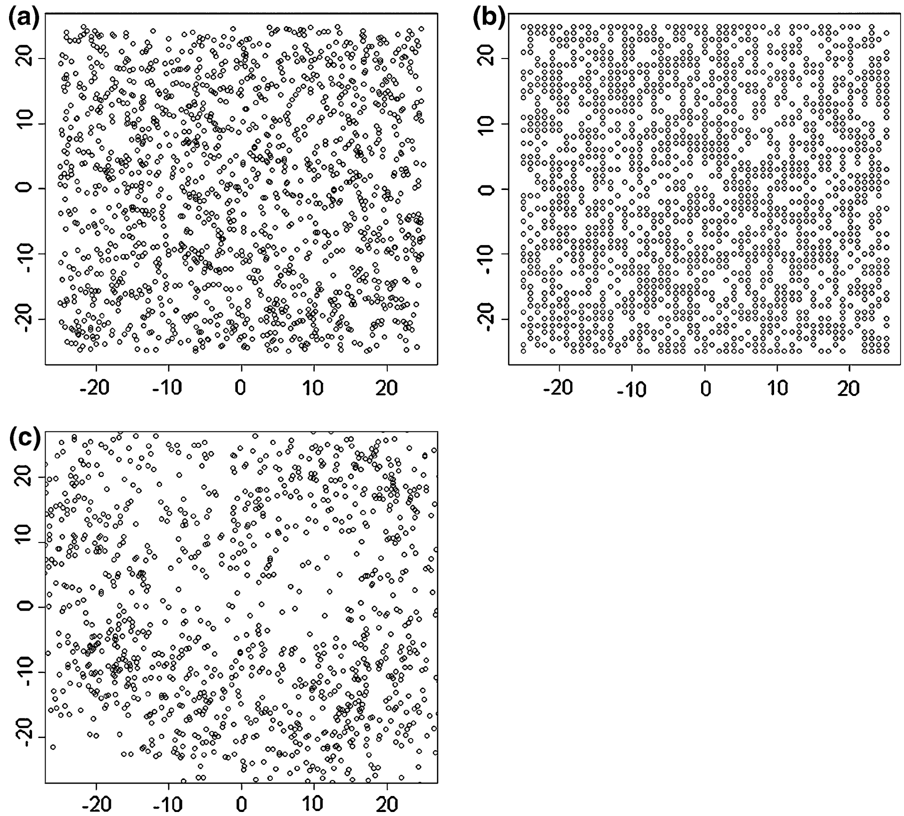

Fig. 1 Examples of simulated (**a**) random, (**b**) regular, and (**c**) clustered spatial distribution for about 1,500 points for a window of 50 × 50. The density is greater and window is smaller than in simulations to emphasize the point patterns

maximum distance at 15 units for tree points from the center of the cluster.

Results

Simulations based on random tree distribution with low density variability showed that both the Morisita and Pollard estimators generally had low bias and high precision when predicting tree density from points with four bearing trees and densities of 10 trees/unit area and greater (Table 2). We excluded from analyses densities of 5 trees/unit area, because the Pollard estimator overestimated density by a ratio of 1.1, likely owing to vacant quadrants and the resulting great density variation. When predicting density based on two or three bearing trees, both estimators produced overestimates. The Morisita estimator had positive ratios of 1.22 (two bearing trees) and 1.18 (three bearing trees), while the Pollard estimator had ratios of 1.21 (two bearing trees) and 1.18 (three bearing trees), after excluding the lowest density (5 trees/area unit). Based on these results, estimates can be corrected by simply dividing by 1.22 (Morisita, two bearing trees), 1.21 (Pollard, two bearing trees), or 1.18 (both estimators, three bearing trees). All the simulations included this correction.

Simulations based on random tree distribution with moderate to high density variability established superior performance of the Morisita estimator compared to the Pollard estimator (Table 3, values for combined two, three, and four trees). The Morisita estimator produced densities ranging from 0.97 to 1.01 of the actual density. The Pollard estimator produced densities as low as 0.55 of the simulation mean with moderate density variability, and as low as 0.10 with high density variability.

Simulations based on regular tree distribution showed that both the Morisita and Pollard estimators had relatively low bias and high precision across a range of simulated tree densities (Table 4). The Pollard estimator produced overestimates at a density of 5 trees/unit area (1.03–1.09) and underestimates (0.95–0.98) until reaching a density of 20 trees/unit area. Here, the Morisita estimator produced slight underestimates, down to an overall ratio of 0.97 with

Table 2 Estimator summary statistics for random spatial distribution before and after adjustment (*Adj*) and all simulated densities (10–25 trees/unit area in steps of 5, 30–100 in steps of 10, and 200–1,000 in steps of 200), 60 trials, and 600 survey points

Estimator	Ratio[a]	Adj Ratio	Bias[b]	Adj Bias	SD[c]	Adj SD
Morisita two trees	1.22	1.00	43.91	0.19	73.38	29.44
Morisita three trees	1.18	1.00	37.85	1.79	59.12	16.72
Morisita four trees	1.00		0.97		10.65	
Pollard two trees	1.21	1.00	41.35	–0.29	61.98	10.58
Pollard three trees	1.18	1.00	36.08	0.29	54.05	8.73
Pollard four trees	1.00		1.12		7.70	

[a] Mean density estimate/mean simulation density

[b] Mean of (density estimate – simulation density)

[c] Standard deviation of bias

two bearing trees. The ratios also increased slightly (0.96–1.02) with increasing tree density, except for the Morisita estimator for two bearing trees, which decreased (0.99–0.94) over the density range.

Simulations based on clustered tree distribution clearly demonstrated that the Morisita surpassed the Pollard estimator across a range of cluster configurations (Table 5). Overall, the Morisita estimator produced density estimates of 1.01 (four trees), 1.08 (three trees), and 1.10 (two trees) of the actual density value. Comparable figures for the Pollard estimator were 0.67, 0.70, and 0.69, respectively. The density estimates were most strongly affected by changes in the number of clusters. Specifically, the estimator performance increased with increasing number of clusters, with the Morisita estimator based on four bearing trees increasing from 0.70 (5 clusters/unit area) to 1.06 (100 clusters/unit area) of the simulation mean, and the Pollard estimator based on four trees increasing from 0.29 (5 clusters/unit area) to 0.81 (100 clusters/unit area). The cluster radii did have an effect on the Morisita estimator for points with four trees and the Pollard estimator (Appendix B in Electronic supplementary material). The Morisita estimator with four trees increased neutrally (around 1.00) from 0.96 of the simulation value with a maximum distance of 10 units to 1.05 with a maximum distance of 15 units, whereas for example, the Pollard estimator with four trees improved from 0.57 to 0.76 of the simulation value as the maximum radii increased. The number of trees per cluster did not affect the ratio. Estimates can be corrected by dividing by overall ratios of 1.10 (Morisita, two bearing trees) or 1.08 (Morisita, three bearing trees). Additionally, a low and high range of density estimate can incorporate effects of cluster number, for example, dividing by 1.21 at the low range and by 1.08 at the high range for the Morisita estimate with two bearing trees.

Lastly, we determined the minimum number of survey points needed to achieve the tolerance limits of ±5%, 10%, and 20% of the simulated density means based on random spatial distribution (Table 6). Not surprisingly, estimates based on four bearing trees required fewer survey points than those based on two or three bearing trees. Changing from random to regular spatial patterns had no effect on the number of survey points needed to achieve a given tolerance (Table 7). In contrast, clustered patterns required larger numbers of points (Table 8). For example, for the ±5% tolerance level, the minimum number increased from 600 (Morisita, four bearing trees) to 4,000.

Discussion

The GLO data represent an invaluable source of information regarding forest conditions at the time of European settlement for much of the U.S. However, bias or imprecision in metrics derived from these data, such as tree density, may produce faulty interpretations of conditions, which in turn may lead to misguided management strategies. Our evaluation of the Morisita and Pollard tree density estimators for GLO data contribute to the recent interest in assessing bias and improving interpretation of GLO data (e.g., Kronenfeld and Wang 2007; Bouldin 2008).

Table 3 Estimator (combined for two, three, and four trees) summary statistics for random spatial distribution (after adjustment) and non-uniform densities, 60 trials, and 600 survey points

Estimator	Ratio[a]	Bias[b]	SD[c]	Density low, high
Morisita two, three, four trees	0.99	−0.12	0.46	5, 10
Morisita two, three, four trees	1.00	0.02	0.54	5, 15
Morisita two, three, four trees	0.98	−0.26	0.62	5, 20
Morisita two, three, four trees	0.99	−0.10	0.84	5, 25
Morisita two, three, four trees	0.99	−0.19	2.11	5, 50
Morisita two, three, four trees	0.98	−0.82	3.02	5, 100
Morisita two, three, four trees	0.99	−1.33	9.17	5, 250
Morisita two, three, four trees	0.99	−0.33	1.45	20, 40
Morisita two, three, four trees	0.99	−0.26	2.33	20, 60
Morisita two, three, four trees	1.00	0.08	2.74	20, 80
Morisita two, three, four trees	0.99	−0.48	3.13	20, 100
Morisita two, three, four trees	1.01	0.70	7.01	20, 200
Morisita two, three, four trees	0.99	−1.30	15.52	20, 500
Morisita two, three, four trees	1.00	−0.90	31.45	20, 1,000
Pollard two, three, four trees	0.95	−0.39	0.17	5, 10
Pollard two, three, four trees	0.82	−1.80	0.21	5, 15
Pollard two, three, four trees	0.71	−3.63	0.27	5, 20
Pollard two, three, four trees	0.63	−5.68	0.29	5, 25
Pollard two, three, four trees	0.38	−17.24	0.78	5, 50
Pollard two, three, four trees	0.21	−41.54	1.48	5, 100
Pollard two, three, four trees	0.09	−115.51	4.27	5, 250
Pollard two, three, four trees	0.89	−3.39	0.70	20, 40
Pollard two, three, four trees	0.75	−10.18	0.84	20, 60
Pollard two, three, four trees	0.64	−18.06	1.15	20, 80
Pollard two, three, four trees	0.56	−26.58	1.26	20, 100
Pollard two, three, four trees	0.33	−72.82	2.22	20, 200
Pollard two, three, four trees	0.15	−218.43	7.86	20, 500
Pollard two, three, four trees	0.08	−466.21	17.10	20, 1,000

[a] Mean density estimate/mean simulation density

[b] Mean of (density estimate – simulation density)

[c] Standard deviation of bias

The Morisita estimator strongly outperformed the Pollard estimator in situations of non-uniform density and clustered distribution, that is, under typical forest conditions. The Morisita estimator can handle greater point-to-survey point variation because density estimation occurs for each point, rather than averaged over all survey points as in the Pollard estimator. The Pollard estimator was unable to compensate for non-uniform density, and we were unable to develop adjustments to systematically correct the estimates produced under non-uniform densities, as there was no obvious pattern. Kronenfeld and Wang (2007) also documented density underestimation by the Pollard estimator for non-uniform density and clustered spatial distribution, and furthermore, density overestimation with dispersal (i.e., regular spatial distribution).

Nevertheless, the Morista estimator produced density overestimates for clustered spatial distribution. Although the actual spatial distribution of the forest in question is unknown, there is likely to be a range of spatial patterns across the landscape. Researchers could present both a density estimate for random and regular forests and a density estimate for clustered forests, representing potential endpoints of a landscape density gradient. Again, estimates can be adjusted by simply dividing by the ratios provided above. Furthermore, density estimates varied considerably depending on the number of clusters, and therefore researchers also might consider incorporating a low and a high density estimate to account for the potential effect of clustering. We are uncertain why there was so much variability in Morisita density estimates when survey points had four trees (ratios of

Table 4 Estimator summary statistics for regular spatial distribution (after adjustment) and all simulated densities (5–25 trees/unit area in steps of 5, 30–100 in steps of 10, and 200–1,000 in steps of 200), 60 trials, and 600 survey points

Estimator	Ratio[a]	Bias[b]	SD[c]
Morisita two trees	0.97	−9.17	21.47
Morisita three trees	0.98	−1.45	11.66
Morisita four trees	0.99	−0.20	9.75
Pollard two trees	1.00	1.52	11.61
Pollard three trees	1.00	2.31	10.70
Pollard four trees	0.99	1.03	8.53

[a] Mean density estimate/mean simulation density

[b] Mean of (density estimate – simulation density)

[c] Standard deviation of bias

0.70–1.06 of the actual density). In general we do not encourage loss of information, but in this case we suggest eliminating the most distant tree, thus converting the survey points to three trees, thereby reducing the range of density estimates. This differs from excluding all but the nearest tree, in the expectation that removing non-random trees will reduce the effect of surveyor bias (Leitner et al. 1991).

In addition to adjustments for clustering, correction is both necessary and easily accounts for overestimation at survey points with fewer than four trees. Most GLO survey points have fewer than four trees, as it was not until 1833 that surveying protocol required four trees, and then only at the intersection of section corners within a township (White 1983). Without correction, error from density overestimates will accumulate with subsequent extrapolations, such as for dominance (Leitner et al. 1991) or biomass and carbon estimations (Rhemtulla et al. 2009).

Sample sizes (i.e., number of survey points used in estimates) need to be large enough for accuracy but small enough to allow as much heterogeneity in density as possible. Density estimators produce only one estimate, and as sample size increases, so does the areal extent of the density estimate. Using the Morisita estimator for survey points with four trees, and accounting for the possibility of clustering, the number of survey points used for estimating density probably should be at least 200, whereas 1,000 survey points likely would fall within the ±10% tolerance range of actual density. For survey points with two trees, minimum sample size probably should be 600 survey points, with 2,000 survey points achieving ±10% tolerance. Density estimates using only a few survey points (see Bouldin 2008) are unlikely to be correct. Given that each township includes 85 survey points, excluding the outer boundaries, and 48 points along the boundaries, an area including ca. 10–20 townships would be sufficient to produce reliable density estimates, depending on the number of trees per point. Rather than following a strict number of survey points for density estimation, we recommend using units from a land type classification system for density estimation.

Table 5 Estimator summary statistics for clustered spatial distribution after adjustment for different number of clusters per survey point, pooled across a range of trees/cluster (4, 6, 8, and 10) and two maximum cluster radii (10 and 15 units), for 60 trials and 600 survey points

Estimator	All clusters (5, 10, 25, 50, and 100 clusters/point)			5 clusters/point			25 clusters/point			100 clusters/point		
	Ratio[a]	Bias[b]	SD[c]	Ratio[a]	Bias[b]	SD[c]	Ratio[a]	Bias[b]	SD[c]	Ratio[a]	Bias[b]	SD[c]
Morisita two trees	1.10	20.20	13.18	1.21	6.93	7.28	1.14	19.91	22.24	1.08	38.32	48.96
Morisita three trees	1.08	16.09	9.71	1.03	1.02	5.17	1.09	13.05	13.02	1.09	40.25	31.62
Morisita four trees	1.01	1.69	11.20	0.70	−10.19	5.18	0.96	−5.28	14.01	1.06	28.04	28.02
Pollard two trees	0.69	−62.83	34.64	0.36	−22.62	9.68	0.53	−72.25	34.67	0.82	−88.04	67.21
Pollard three trees	0.70	−61.93	34.22	0.33	−23.80	9.82	0.53	−71.79	34.70	0.83	−84.08	65.16
Pollard four trees	0.67	−68.09	36.09	0.29	−24.93	10.15	0.49	−77.39	35.87	0.81	−95.90	70.03

[a] Mean density estimate/mean simulation density

[b] Mean of (density estimate – simulation density)

[c] Standard deviation of bias

Table 6 Minimum number of survey points needed to achieve specified tolerance (deviation from simulated mean) limits based on 60 trials (for each density from 5 to 25 trees/unit area in steps of 5 and mean for all densities from 30 to 100 in steps of 10 and 200–1,000 in steps of 200) and random spatial distribution

Estimator	Tolerance (%)	Number of points	Trials exceeding tolerance					
			Each density					All densities[a]
			5	10	15	20	25	30–100 and 200–1,000
Morisita two trees	±5	3,000	0	1	0	4	1	10.85
Morisita three trees	±5	1,000	15	7	6	4	5	4.92
Morisita four trees	±5	600	1	7	4	3	3	4.69
Pollard two trees	±5	600	58	8	3	2	2	5.46
Pollard three trees	±5	600	16	3	1	1	0	1.69
Pollard four trees	±5	600	60	11	11	3	0	0.92
Morisita two trees	±10	1,000	7	7	6	4	4	6.31
Morisita three trees	±10	200	6	8	10	9	8	7.77
Morisita four trees	±10	200	1	3	2	1	3	2.54
Pollard two trees	±10	200	26	1	4	1	2	2.31
Pollard three trees	±10	200	3	0	3	2	1	0.69
Pollard four trees	±10	200	21	1	2	0	0	0.38
Morisita two trees	±20	200	6	6	6	4	4	4.77
Morisita three trees	±20	50	7	10	8	6	9	7.31
Morisita four trees	±20	50	2	2	1	3	1	2.54
Pollard two trees	±20	50	8	3	2	0	3	3.38
Pollard three trees	±20	50	1	1	1	0	2	1.08
Pollard four trees	±20	50	1	0	0	0	0	0.23

[a] Mean trials

Table 7 Minimum number of survey points needed to achieve specified tolerance (deviation from simulated mean) limits based on 60 trials, for non-uniform densities of 10–100 trees/unit area and regular spatial distribution

Estimator	Tolerance					
	±5%		±10%		±20%	
	Number of points	Mean trials	Number of points	Mean trials	Number of points	Mean trials
Morisita two trees	3,000	8	1,000	4	200	4
Morisita three trees	1,000	9	200	7	50	8
Morisita four trees	600	10	200	1	50	6

Table 8 Minimum number of survey points needed to achieve specified tolerance (deviation from simulated mean) limits based on 60 trials, for non-uniform densities of 20–100 trees/unit area and clustered[a] spatial distribution

Estimator	Tolerance					
	±5%		±10%		±20%	
	Number of points	Mean trials	Number of points	Mean trials	Number of points	Mean trials
Morisita two trees	10,000	10	2,000	10	600	2
Morisita three trees	4,000	5	1,000	7	200	4
Morisita four trees	4,000	9	1,000	4	200	7

[a] Clusters of 10, 2–10 trees per cluster, maximum radius of 15 from center

Our results demonstrate the range of errors possible given the unqualified application of tree density estimators of GLO data. This information, along with provided corrections, will increase the reliability of future studies that rely on density estimates from GLO records. Researchers traditionally have used faulty estimators with sample sizes that either were too small or excessively large. They have not corrected for survey points with two trees nor adjusted for possible clustered spatial distribution, or conversely, avoided the potential pitfalls of density estimation and thus limited the usefulness of the GLO surveys. Although previous work may present general trends, research in the future can improve density estimates by incorporating these guidelines.

Acknowledgments We thank William Dijak for programming assistance. Support was provided by the National Fire Plan and the USDA Forest Service, Northern Research Station.

References

Almendinger JC (1996) Minnesota's bearing tree database. Minnesota Department of Natural Resources, St. Paul, Minnesota

Bolliger J, Mladenoff DJ (2005) Quantifying spatial classification uncertainties of the historical Wisconsin landscape (USA). Ecography 28:141–156

Bouldin J (2008) Some problems and solutions in density estimation from bearing tree data: a review and synthesis. J Biogeogr 35:2000–2011

Bourdo EA (1956) A review of the General Land Office survey and of its use in quantitative studies of former forests. Ecology 37:754–768

Cottam G, Curtis JT (1956) The use of distance measures in phytosociological sampling. Ecology 37:451–460

He HS, Mladenoff DJ, Sickley T, Guntenspergen GG (2000) GIS interpolations of witness tree records (1839–1866) for northern Wisconsin at multiple scales. J Biogeogr 27:1031–1042

Kronenfeld BJ, Wang Y-C (2007) Accounting for surveyor inconsistency and bias in estimation of tree density from presettlement land survey records. Can J Forest Res 37:2365–2379

Leitner LA, Dunn CP, Guntenspergen GR, Stearns F, Sharpe DM (1991) Effects of site, landscape features, and fire regime on vegetation patterns in pre-settlement southern Wisconsin. Landscape Ecol 5:203–217

Morisita M (1957) A new method for the estimation of density by the spacing method, applicable to non-randomly distributed populations. Seiri Seitai 7:134–144 (in Japanese)

Nelson JC (1997) Presettlement vegetation patterns along the 5th principal meridian, Missouri territory, 1815. Am Midl Nat 137:79–94

Pollard JH (1971) On distance estimators of density in randomly distributed forests. Biometrics 27:991–1002

Rhemtulla JM, Mladenoff DJ, Clayton MK (2009) Historical forest baselines reveal potential for continued carbon sequestration. PNAS 106:6082–6087

Schulte LA, Mladenoff DJ, Burrows SN, Sickley TA, Nordheim EV (2005) Spatial controls of pre-Euro-American wind and fire disturbance in northern Wisconsin (USA) forest landscapes. Ecosystems 8:73–94

Ward RT (1956) The beech forests of Wisconsin: changes in forest composition and the nature of the beech border. Ecology 37:407–419

White CA (1983) A history of the rectangular survey system. Bureau of Land Management, Government Printing Office, Washington DC, USA