Received: 5 June 2023 | Accepted: 9 June 2023

DOI: 10.1002/ecs2.4647

ECOSPHERE
AN ESA OPEN ACCESS JOURNAL

SYNTHESIS & INTEGRATION

Methods, Tools, and Technologies

# Surveyor and analyst biases in forest density estimation from United States Public Land Surveys

**Charles V. Cogbill**

Harvard Forest, Harvard University, Petersham, Massachusetts, USA

**Correspondence**
Charles V. Cogbill
Email: cvcogbill@gmail.com

**Funding information**
National Science Foundation, Grant/Award Numbers: DEB-1065656, DEB-1065702, DEB-1241868

**Handling Editor:** Debra P. C. Peters

**Abstract**

Accurate forest density estimates based on United States Public Land Surveys have long been questioned because of doubts about randomness of both the surveyors' selection of witness trees and the underlying tree dispersion. This study analyzes the surveyor sampling of witness trees in six Midwestern states in the mid-1800s. It develops universal methods for identification, quantification, and correction of bias, and then calculation of unbiased density. Applying these techniques produces unbiased site-specific densities before Euro-American settlement, which are the essential baseline for determining historic changes in forest structure. Previous analysts used untested assumptions, inaccurate estimators, unknown or unrealistic sampling designs, and omitted or poorly corrected surveyor bias, resulting in hundreds of unreliable density estimates. The surveyors' recording of the empirical distance, bearing, and diameter of witness trees documented the exact sampling design. The intended design and deviation from it are investigated with a combination of descriptive statistics, probability theory, computer simulations, analogue geometric models, and modern stand conditions. Herein, analyst bias is eliminated using the robust Morisita II density estimator matching the predominant sampling design of two trees in opposing semicircles. Six widespread surveyor biases deviating from the nearest tree to the corner are evaluated. Quadrant bias and diameter bias for medium-sized trees are subsumed under newly framed design and small tree biases. Two novel surveyor positional biases (pair angle and near-post) are introduced here. Previously recognized azimuthal and species biases are analyzed with new techniques. Widely postulated surveyor bias for certain species was found to be minimal. Bias correction and density estimation are applied in detail over 68 townships in northern Wisconsin. The estimated historical forest density in northern Wisconsin, corrected for bias and small tree truncation, averaged 323 trees/ha ≥20 cm. Over 80 Midwestern subregions, surveyors bypassed an estimated 17% of the nearest trees due to their position, resulting in an average bias correction of +24% over the base density. If censored trees below a 20-cm "veil-line" are considered, the surveyors bypassed 48% of the nearest trees >12.7 cm in diameter. This study resolves a 70-year-old conundrum of surveyor and analyst biases in historical density estimation.

**KEYWORDS**
analyst error, historical ecology, Morisita, near-post bias, northern Wisconsin, pair angle analysis, plotless density estimator (PDE), public land survey (PLS), randomness, surveyor bias, veil-line truncation, witness tree

# INTRODUCTION

Unbiased estimates of forest density before Euro-American perturbations are a touchstone for all applications of historical ecology throughout the United States. Although density estimates are derived through a seemingly esoteric exploration of historic surveying methodology and the statistics of point pattern analysis, accurate density estimates have potentially profound influence and broad applications. The use of historic land surveys has been invaluable to current research ranging from the influence of Indigenous populations and the prevalence of open ecosystems (e.g., Abrams et al., 2022; Hanberry, Bragg, et al., 2020; Whitney & Steiger, 1985), to the role of fire and the restoration and management of forests (e.g., Baker et al., 2023; Knight et al., 2020; Meunier & Shea, 2020), to the determination of aboveground biomass of undisturbed forests (e.g., Hanberry & He, 2015; Rhemtulla et al., 2009), or to the parameterization and validation of theoretical vegetation models (e.g., Blankenship et al., 2021; Raiho et al., 2022). Unfortunately most previous density estimates lack any evaluation of bias or methodology for assuring the accuracy of historic survey data. Beyond the obvious interesting descriptions of long-lost ecosystems, the density of past forests is the ideal baseline for comparison with present or future forest structure. Elimination of surveyor bias and correction of previous forest density misestimates are thus fundamental to the accurate historical biogeography of the United States.

The United States Public Land Survey (PLS) is the only comprehensive quantitative sample of the composition and structure of the forests of the United States before Euro-American settlement. These surveys divided the federal land, from Ohio westward, into six-by-six-mile (9.66 × 9.66 km) townships and subsequently subdivided each township into 36 one-mile-square (1.61 × 1.61 km) sections (Price, 1995). At each section (mile mark) and quarter-section (half-mile mark) corner along township exterior and interior subdivision section lines, posts were set in the ground and surveyors were instructed (ca. 1805) to mark "two or more adjacent trees in opposite direction as nearly as may be" (White, 1983). Surveyors recorded the common name, diameter, and distance and bearing from the corner post to typically two or four "bearing trees," or after 1834 interchangeably called "witness trees" as used herein (Bourdo, 1956; Schulte & Mladenoff, 2001; Stewart, 1935; White, 1983). Although this Public Land Survey System (PLSS) was designed for the pragmatic purposes of surveying and legal description, it also yields a systematic, plotless ecological sample of the forests of that time. For over a century, myriad studies have reconstructed early species composition from these records. Recently measurements of the distance to witness trees have been used to infer historical forest structure (e.g., Baker et al., 2023; Bourdo, 1956; Cottam, 1949; Goring et al., 2016; Knight et al., 2020; Kronenfeld & Wang, 2007; Manies et al., 2001; Paciorek et al., 2021; Rhemtulla et al., 2009).

Most researchers reanalyzing PLS data have either not addressed the problem of surveyor bias (58% of the 956 historical PLS papers in Cogbill et al., 2018) or given it only cursory, qualitative consideration (23%), frequently simply acknowledging Bourdo's (1956) concerns about nonrandom sampling of trees. Of the 18% of the papers that investigated bias in more detail, Eric Grimm's work (1981, 1984) has been by far the most influential in casting doubt on the assumptions of randomness. Grimm mustered evidence from PLS surveys in Minnesota and a critical review of Bourdo's analysis to propose that "… the most specious uses of land surveys have been the purported statistical analysis of surveyor bias and computation of absolute tree densities" (Grimm, 1981). In particular, Grimm summarily concludes that the "assumption of random dispersion is unrealistic (and) … the surveyors were inevitably biased for or against certain sizes and species" (Grimm, 1984). This absolute dismissal of both spatial and tree sample randomness leads to the proposition that PLS samples are not amenable to accurate density determination. Continuing doubts about these two aspects of randomness have heretofore stymied progress in plotless density estimation.

An accurate forest density estimate is a straightforward mathematical function of the measured distances, if the underlying dispersion of the sampled trees follows complete spatial randomness (CSR), if the surveyors are unbiased in the choice of bearing trees, and if a distance-based plotless density estimator—PDE—fits the sampling design (Cogbill et al., 2018). The correctness of each of these three assumptions is unknown a priori; therefore, the density determination from PLS data is problematic. Historical surveys present additional impediments to density estimation as the exact sampling methods are often ill-defined, sampling cannot be replicated, and simulations are based on unknown true density or tree spatial pattern. While randomness of both underlying spatial dispersion of trees at different scales (CSR) and surveyor's choice of witness trees (nearest tree to the corner) is a sufficient condition for unbiased density determinations (Cogbill et al., 2018), this paper will prove that randomness of pattern and sampling are not necessary conditions for valid results.

The primary purpose in correction of the accumulated surveyor bias in PLS surveys is to produce an explicit and accurate determination of forest density at specific sites before land clearance or forest management. This is only possible when the modern analyst uses an appropriate density estimation methodology, minimizes errors in interpretation of survey design, and rectifies any surveyor bias imbedded in these data as outlined below. The resultant forest composition and structure provide a baseline reference for modern forest changes and enable an overall understanding of historical forest dynamics.

## Surveyor bias

Proposed reasons for surveyor bias are varied: subjective choice of trees (Sears, 1921, 1925); tree size (Kenoyer, 1930); species (Lutz, 1930; Shanks, 1938); economic value (Lutz, 1930); longevity (Cottam & Curtis, 1949); quadrant position (Bourdo, 1956); ease of marking (Bourdo, 1956); spatial pattern of the trees (Bourdo, 1956; Siccama, 1971); age (Grimm, 1981); conspicuousness (Grimm, 1981); tree health (Almendinger, 1996); bearing (Manies, 1997); or ease of locating (Schulte & Mladenoff, 2001). Eric Bourdo, uniquely qualified as both a surveyor and an ecologist, presented the definitive exploration of PLS surveying techniques in Michigan and the potential for bias in surveyor selection of witness trees (Bourdo, 1954, 1956). Integrating historical instructions, empirical evidence from PLS data, current surveying practice, and statistical analyses of measured distances, Bourdo concluded that surveyor choice of trees was not random, specifically favoring medium-sized trees or particular species. Bourdo (1956) recommends that any researcher should investigate the characteristics of PLS sampling at the specific site to determine the possibility and degree of any bias.

Surveyor bias in early land surveys has been addressed in over 190 scientific publications over 100 years (e.g., Sears, 1921; see Cogbill et al., 2018). Four sources of surveyor bias are commonly considered for density estimations in PLS publications: (1) the sampling design deviates from one witness tree per quadrant or from random among quadrants; (2) certain species are favored or avoided as witness trees; (3) medium-sized trunks are preferred; and (4) trees are undersampled near the cardinal azimuths (Table 1; e.g., Hanberry, Yang, et al., 2012). The latter three of these biases result in the choice of a substitute tree at a greater distance from the post than the nearest tree. Despite acknowledgment of the qualitative importance of errors in density estimation, there has been little consistency in the methods for quantitative determination of bias, and estimates of bias vary widely in magnitude (Table 1). Therefore, the purpose of this paper is to develop and apply uniform methods for identifying, quantifying, and correcting surveyor biases in PLS surveys, and to reach a clearer overall understanding of their magnitude and effects on resulting reconstructions of historical forests.

# MATERIALS AND METHODS

## PLS surveys

Analyses in this paper are based on data extracted from PLS surveys overseen by the General Land Office (GLO) in the Midwest United States. These surveys cover six current states (Minnesota, Wisconsin, Michigan, Indiana, Illinois, Ohio) documenting 642,354 corners, of which 526,852 included 1,138,614 witness trees sampled from 1786 to 1907 (Table 2; PalEON database: http://www.paleonproject.org; see *Data Availability Statement*). Detailed analyses for 80 representative subregions include 505,080 witness trees at 238,096 corners, herein referred to as "in the Midwest." In Minnesota, Wisconsin, and the Upper Peninsula and the northern half of the Lower Peninsula of Michigan, subregions were chosen to represent stratified geography, vegetation types, and date of survey. In this northern domain (i.e., Goring et al., 2016), subregions consist of 44 nominal countywide units (average 2796 $km^2$/unit, 5861 witness trees/unit), proportionally sampling approximately one third of available data. In Illinois, Indiana, the southern half of the Lower Peninsula of Michigan, and the southeastern quarter of Ohio, digitization of survey data is ongoing, but still herein has complete geographic coverage. The available data, comprising approximately 40% of the townships in this southern domain (i.e., Paciorek et al., 2021), are aggregated into 36 contiguous subregions

**TABLE 1** Literature identification and estimation of various surveyor bias corrections in density determinations from reanalysis of public land survey (PLS) data.

| Study | Estimator | Quadrant[a] | Ward[b] | Azimuth | Species | Near[c] | Size[d] |
|---|---|---|---|---|---|---|---|
| Ward (1954) | Distance and RP | * | 1.15 | | | | |
| Bourdo (1956) | Distance | ✓ | | | ✓ | | * |
| Anderson and Anderson (1975) | Shanks | ✓ | 1.11 | | | | |
| Delcourt and Delcourt (1974, 1977) | Distance | | | | ± | | |
| Delcourt (1976) | Shanks | | 1.16 | | * | | |
| Kline and Cottam (1979) | Cottam | | | | ✓ | | |
| Grimm (1981) | Distance | * | | | * | | * |
| Dorney and Dorney (1989) | Distance | | | | ✓ | | ± |
| Leitner et al. (1991) | Shanks and distance | * | | | ± | | ± |
| Bowles et al. (1998) | Shanks | ✓ | | | | | ✓ |
| Edgin and Ebinger (1997) | Shanks | * | | | | | * |
| Clark (2000) | Cottam | * | | | | | ± |
| Manies et al. (2001)[e] | Shanks | ✓ | 1.15 | * | ± | | ± |
| Fralish et al. (2002) | RP, Cottam, N1 | * | | | | | |
| Anderson et al. (2006) | Cottam and RP | * | 1.15 | ✓ | | | |
| Kronenfeld and Wang (2007)[f] | Pollard | 0.94 | | 1.23 | 1.03 | 1.19 | 0.81 |
| Bouldin (2007)[g] | Morisita | 2.16 | 1.15 | 1.73 | * | 1.37 | |
| Bouldin (2009)[h] | Morisita | * | | 1.13 | | 1.51 | ~1.40 |
| Fralish and McArdle (2009) | Cottam | | 1.19 | | | | ± |
| Liu et al. (2011) | Distance | * | | | * | | * |
| Williams and Baker (2011) | Voronoi | | | ± | ✓ | | 1.06 |
| Hanberry, Palik, et al. (2012)[i] | Morisita | 1.24 | 1.21 | 1.18 | 1.60 | 2.33 | * |
| Kronenfeld (2014)[j] | Pollard | | | | 1.05 | | * |
| Tulowiecki (2014)[j] | Pollard | | | | 1.02 | | 1.07 |
| Goring et al. (2016)[k] | Morisita | 1.10 | 1.04 | 1.18 | | 1.30 | 0.85 |
| Paciorek et al. (2021)[l] | Morisita | 1.04 | 1.00 | 1.06 | | 1.13 | 0.86 |

*Note*: ✓, considered but not significant; ±, significant minor in a few situations; *, significant but not quantified.

Abbreviations: N1, absolutely nearest tree; RP, random pair.

[a]Surveyor choice of quadrants reflects pairs sector geometry.

[b]Inverse of correction factor invoked for Ward assumption (two nearest quadrants), corrects inappropriate 2nQ assumption.

[c]Joint spatial correction for all biases against nearest tree.

[d]Factor to adjust density to above a standard diameter particular to each study.

[e]Includes Manies (1997) thesis.

[f]Weighted average for township and quarter corners in Holland Land Company Land, NY, technically not a PLS, but a prototype for later surveying methodology.

[g]Bouldin (2007), public communication, Ecological Society of America Annual Meeting, abstract: "New methods for the analysis of an important historical data set," abstract: https//eco.confex.com/eco/2007/techprogram/P4973.HTM; poster: "Estimating American presettlement forest parameters from General Land Office data: problems and some solutions," https//slideplayer.com/slide/5056273/; corrections for the same MN database as Hanberry, Palik, et al. (2012).

[h]Bouldin (2009), public communication, "One hundred fifty years of tree density decline in the Huron Mountains of Michigan, USA." Annual Report, Huron Mt. Wildlife Foundation. www.hmwf.org/archives/reports/; includes methods from Bouldin (2008) for Huron Mts., MI, uncertain since density is calculated from a diameter frequency fit; bias approximated.

[i]Quantification ill-defined; bias-based approach based on mean empirical deviation of quadrant, azimuth, and lumped species and diameter values from random quadrant; 2:1 adjacent:opposite quadrant expectation; equal 30° azimuth sectors; and similarity of size and diameter to line trees in Missouri. Identical methods for corrections used in other publications for Minnesota, Wisconsin, and Michigan.

[j]Based on relative distance test with sum of significant bias equaling density bias.

[k]Average of interior-section corners over Wisconsin, Michigan, and Minnesota.

[l]Average of all corners over Indiana and Illinois, methodology of this study excludes $\nu$.

**TABLE 2** Witness trees at section and quarter-section public land survey corners in the PalEON database sampled by state and frequency by number of trees at each corner.

| State | Dates | Open non-tree (%) | No. treed corners | Trees per corner 1 (%) | 2 (%) | 3 (%) | 4 (%) | No. witness trees |
|---|---|---|---|---|---|---|---|---|
| Ohio | 1786–1808 | 1.5 | 3753 | 15 | 84 | 0.03 | 0.2 | 6856 |
| Indiana | 1799–1846 | 12 | 60,687 | 4 | 96 | 0.2 | 0.04 | 119,131 |
| Illinois | 1804–1849 | 50 | 35,616 | 13 | 84 | 1 | 2 | 68,369 |
| Michigan | 1815–1855 | 3 | 128,193 | 0.3 | 99 | 0.1 | 0.4 | 256,999 |
| Wisconsin | 1832–1866 | 6 | 155,005 | 4 | 87 | 1 | 8 | 330,327 |
| Minnesota | 1835–1907 | 43 | 143,598 | 5 | 67 | 3 | 25 | 356,932 |
| Midwest | 1786–1907 | 18.0 | 526,852 | 3.9 | 85.5 | 1.3 | 9.3 | 1,138,614 |

(average 9523 km$^2$/unit, 9234 trees/unit). The surveys from counties or divisions of the states are classified into 375 local corner types categorized by combinations of county or subregion; township exterior or interior lines (done at different times by different surveyors); section or quarter-section corners; and number of witness trees per corner.

The methods of identification, quantification, and correction of surveyor bias are applicable to any survey where the species, diameter, distance, and bearing of witness trees were recorded (e.g., California: Knight et al., 2020). This paper will present detailed derivations for each bias in the same area of northern Wisconsin (see *Data Availability Statement*). This case study uses 14,787 witness trees at 6609 PLS corners in 68 scattered townships over nine counties (Ashland County east to Florence County and south to Rush County), herein referred to as "in northern Wisconsin." These are the same townships used by Manies et al. (2001) from the generally mesic mixed conifer–hardwood ecosystems originally surveyed in 1847–1865 by 20 different surveyors.

## Surveyor choice of witness trees

Surveyors' choice of witness trees is key to determining surveyor biases. The default witness tree choices in Midwest PLS surveys are the closest trees to the survey corner. This is consistent with the explicit methodology used earlier by surveyors in New York and Pennsylvania who in the early 1800s marked the "nearest tree(s)" to each corner. PLS surveyors ostensibly followed directions from the Surveyor General for each Survey District and the 1833 instructions for Ohio, Indiana, and the Territory of Michigan explicitly required marking the "nearest adjoining tree… as near as may be to the corner" (White, 1983). This "nearest tree conjecture" is identical to the assumption that the surveyor had no bias that would cause him to bypass the nearest tree and replace it with a more remote tree (higher distance rank). Surveyor bias regardless of reason can be framed as the surveyor simply not using the nearest tree. For example, modern reconstructions of late-1800s PLS surveys done in four Western conifer forests indicate that 1.6%–3.9% (mean 3.0%) of the witness trees were not originally the ones nearest to the post (White, 1976; Williams & Baker, 2010). This replacement of the nearest tree, however, would be constrained by limited availability of a tree more expedient than the nearest. Importantly, the bypassing of trees creates ambiguity because the ignored trees potentially conceal the corner location or confuse the corner relocation. Although doubt may persist about any long-past data collection, the absolutely nearest tree assumption defines an unbiased choice of random witness trees and thus a straightforward estimation of tree density from its distance to the corner (Bollinger et al., 2004; Bouldin, 2008; Bourdo, 1956; Grimm, 1984; Kronenfeld, 2014; Kronenfeld & Wang, 2007; Liu et al., 2011; Nelson, 1997; Whitney, 1994; Williams & Baker, 2010).

Unbiased density estimates are best restricted to "bearing" trees (distance and bearing recorded) located at the regularly spaced, pre-located section or quarter-section corners with a post (Liu et al., 2011). Firstly, the location of the section line and the location of the corners (e.g., Bourdo, 1956; Tulowiecki et al., 2015; Wang, 2004) are unrelated to trees because the post is at a predetermined, precisely surveyed, grid location. Secondly, surveyor selection of the nearest tree at that corner is necessarily a random sample from all trees. Thus, surveyor biases (nonrandom samples) are limited to deviation from the intended spatial geometry of witness trees at the corner and the designation of non-nearest witness trees within that design.

## Survey designs

The Surveyor General's instructions define the putative sampling design used by PLS surveyors (White, 1983).

Proof of the actual sampling designs used in the surveys is preserved in the number of witness trees at each corner and in their distances and bearings as actually measured by the surveyors. In the Midwest PLS, two witness trees were recorded for the vast majority of corners (85.5%, Table 2). The direction-of-travel for the survey line further defined two sides of the line or two sections at quarter-section corners (e.g., Kronenfeld & Wang, 2007; Manies et al., 2001). Around 2-tree corners, the quadrant orientation of pairs of witness trees can be "same-sided" (in adjacent quadrants on the same side of a surveyed section line), "adjacent-across" (in adjacent quadrants across the section line), or "diagonal" (in diagonal quadrants across the section line) (Figure 1). The same-sided: adjacent-across:diagonal (SAD) percentage ratios for a pool of corners indicate the actual pair designs used: 33:33:33 for two nearest quadrants (2nQ); 100:0:0 for same-sided (2sH); 0:50:50 for separate halves by sections (2S); *x*:50–*x*:50 for opposing semicircles misaligned with the section line (2oH); or 0:0:100 for diagonal quadrants (2oQ). These SAD ratios form a template where the quadrant distribution is the first indication of a particular sampling design (cf. Anderson & Anderson, 1975; Bourdo, 1956; Fralish et al., 2002; Grimm, 1984).

At the 2-tree corners in Midwest PLS surveys, the aggregated SAD quadrant ratios straightforwardly correspond to two standard empirical sampling designs (Table 2). Starting in 1786, exterior corners in Ohio, Michigan, southern Wisconsin before 1845, and northern Indiana after 1825 were predominantly sampled with two trees on the same side inside the township (2sH). In northern Wisconsin after 1846 and throughout Minnesota, this 2sH design was applied at exterior-section corners. The same-sided design was virtually restricted to township lines, but in scattered counties in Minnesota, the surveyors also used a one-sided (i.e., same-sided sections) design at interior-section corners. At all the remaining exterior and interior corners in all Midwestern states, the sampling design generally followed a majority diagonal configuration variably mixed with same-sided and adjacent-across pairs (2oH or 2S design). Many of the later interior-section corners had only two bearing trees recorded (ostensibly 2-tree corner), but they also could have two blazed and unrecorded witness trees (actually a 4-tree corner) as allowed in the instructions of 1846 (Grimm, 1984; White, 1983).

The presence of ≥50% of witness tree pairs in the diagonal orientation is indicative of an equal halves design, not a quadrant-based design of random placement of witness trees among quadrants. The strictly "opposite half" random design (either a section across the line-of-travel—2S—or alternatively a semicircle centered in the opposite quarter—2oH), would also include ≤50% of the trees in adjacent orientations (same-sided, adjacent-across) in the opposing half. In the Midwest, 58.7% of 2-tree interior corners were in a diagonal orientation, indicating a sample more restricted within the opposite half than following a strict 1:1 same-sided plus adjacent-across:diagonal ratio (Table 2). This nonexclusive, opposite 2-tree pattern, first documented by Bourdo (1956), is the surveyors' empirical execution of the "in opposite direction as nearly as may be" instruction. Significantly, two trees were in different sections, but not restricted to random quadrants (i.e., 2:1 ratio). The "opposite" orientation of witness trees also minimizes any ambiguity of the corner location as it lies sandwiched between marked trees.

Four-tree corners (9.3% of all Midwest) were prevalent only from northern Illinois after 1837 (16% of all state's corners), and after 1846 from northern Wisconsin (14%) and all of Minnesota (24%) (Table 2). Virtually all (98.8%) of these 4-tree samples were at section corners with 61% on the township's interior lines. Some 95% of the 4-tree corners of the Midwest had one witness tree in each quadrant/section. The exterior 4-tree section corners in half the townships in northern Wisconsin, however, are effectively 2-tree corners with all witness trees in duplicate pairs inside the township. Additionally, most of the apparent 4-tree exterior corners in Michigan are the joining of two township surveys, with each using two trees on the inside of their respective townships. These surveys were done by different surveyors at different times, and due to surveying inaccuracy and the convergence of meridians, they are not necessarily tied to the exact same corner. Thus, these apparent

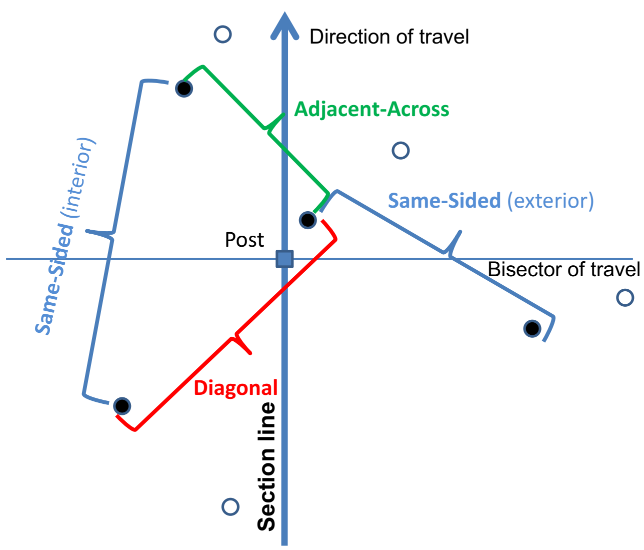


**FIGURE 1** Survey designs for witness tree pairs at corners on a typical public land survey eastern exterior section line. Solid circles are the nearest trees in the four quadrants (4-tree corner) and open circles are other nearby trees. Brackets connect chosen trees according to four possible 2-tree designs (same-sided interior or exterior; adjacent-across; diagonal).

4-tree corners are actually dual 2-tree same-sided (2sH) corners and there are few actual 4-tree corners in Michigan.

There were relatively few 1-tree (3.9%) or 3-tree (1.3%) corners in the six states (Table 2; Liu et al., 2011) as they represent various irregular designs, such as "corner tree" (replacing post), or missing second or fourth tree. One-tree corners were most common in the prairie and savannas at the western edge of the region (e.g., Illinois 12% of all state corners; Wisconsin 9%; Minnesota 5%) where many second trees were "not found" or "too far to measure." In northern Wisconsin, 1-tree corners comprise 1.62% of all corners and are predominantly (88%) single witness tree with "no other tree around." Most of the single tree corners were interior-quarter corners (79%) and the witness tree is at a great (64% >10 m, mean = 33.2 m) distance. Three-tree corners were most common in Minnesota (3%), Illinois (1%), and Wisconsin (1%; Liu et al., 2011) where many are incomplete 4-tree designs (skipped section, bearing and distance not recorded, three nearest trees). In northern Wisconsin, 3-tree corners comprised 0.84% of all corners and represent four different alternative designs: duplicate trees in the same quarter (32%); corner tree plus two witness trees (23%); "no other tree near" (14%); or undefined missing or extra tree (30%).

## Analyst-induced error

### Determination of sampling design and choice of density estimator

Virtually all theoretical models, simulations, and modern field samples using PDEs do not question the nature of sampling design because it is perfectly known and implemented. In contrast, a researcher using historical data collected by a PLS surveyor must reconstruct the sampling design and account for any subjective deviations. Furthermore, the analyst must match the PDE with the actual sampling design. These analyses are constrained by the distances of two or four nearby witness trees at each of widely spaced corners in possibly different forest types. Despite myriad PDEs, the most appropriate and flexible equation known to fit these conditions is the Morisita estimator (Bouldin, 2008; Cogbill et al., 2018; Goring et al., 2016; Hanberry et al., 2011; Levine et al., 2017; Morisita, 1957; Shen et al., 2020):

Morisita Plotless Density Estimator:

$$\lambda_M = \left[\frac{(g \times k - 1)}{\pi \times N}\right] \times \left[\sum_{i=1}^{N} \frac{k}{\left(\sum_{j=1}^{k} r_{ij}^2\right)}\right], \tag{1}$$

where $\lambda$ is density, $g$ is distance rank order, $k$ is the number of equiangular sectors, $N$ is the number of corners, $r_{ij}$ is the distance from post to the $g$th nearest tree in the $j$th sector at the $i$th corner.

The Morisita PDE has the advantage of being evaluated at a single point (corner) and thus relaxes the random assumption and bias for any broadscale or inhomogeneous (nonstationary) spatial pattern. Various simulations of sampling (R Core Team, 2019; R code in *Data Availability Statement*) of regionally inhomogeneous and locally nonrandom dispersions at widely different densities and patterns demonstrate that the Morisita II ($g = 1$, $k = 2$) estimate is broadly robust and averages $0.975 \pm 0.016$ SD of the true value (Cogbill et al., 2018). In contrast, the commonly used Cottam PDE is strongly biased ($0.655 \pm 0.211$ SD) for the same nonstationary and nonrandom dispersion of trees (Cogbill et al., 2018; Cottam & Curtis, 1956; Morisita, 1954):

Cottam Plotless Density Estimator:

$$\lambda_C = \frac{g \times k}{4 \times \left[\sum_{i=1}^{N}\sum_{j=1}^{k} r_{ij}/(k \times N)\right]^2}. \tag{2}$$

For example, at 2-tree interior quarter-section corners in northern Wisconsin (see *Data Availability Statement*: Northern Wisconsin Case Study), the recorded witness trees have a mean 5.19-m distance ($r$) from the post. Insertion of this mean distance into the Cottam PDE above yields an estimated density of 185.5 trees/ha. In contrast, using the Morisita II PDE, the measured distances yield a 325.8 trees/ha density. Unfortunately, the widely used Cottam estimator is highly biased and unsuitable for PLS density estimation.

The forest spatial pattern in the Midwest PLS is dominated by regional inhomogeneity (Cogbill et al., 2018). Significantly, if the trees were dispersed with CSR, the simulation of mean distance for 325.8 trees/ha is 3.92 m from the post. The northern Wisconsin data indicate that the spatial pattern was strongly inhomogeneous with a Pielou's index of randomness ($\overline{r^2}/\lambda$) of $\alpha = 11.2$ ($\alpha \gg 1.27$ {random}, $p < 0.001$; Pielou, 1959). Indeed, the empirical densities were highly heterogeneous (Morisita II individual corner estimates from 0.025 to 4663 trees/ha with a mean $326 \pm 403$ SD trees/ha) and hyper-skewed with a long tail of far distances. The inverse square of these distances results in a strong negative skewness of Morisita density (median 201 trees/ha $\ll$ mean).

Simulation and sampling (R code in *Data Availability Statement*) of an inhomogeneous pattern (exponential gradient of 20–2363 trees/ha with a mean density of 325 trees/ha) yield a mean distance of 5.21 m to the post and a Morisita median density of 189 trees/ha, which is very similar to the northern Wisconsin sample (5.19 m, 201 trees/ha). Additionally, the Morisita II ($g = 1$, $k = 2$) estimator contains minor bias due to local regular spatial pattern found in Midwestern forests (e.g., −1.6% for 1.5 m Matern II inhibition, spatstat package in R; Cogbill et al., 2018). Thus, the appropriate simulation analogue for PDE density determination is a dual-scale inhibited-inhomogeneous ($InH^2$) model. For northern Wisconsin PLS surveys, a realistic nonrandom $InH^2$ dispersion model (exponential gradient from 30 to 1030 with a mean of 402 trees/ha and a Matern II inhibition of 1.5 m) is used for expectations and verifications.

### Diameter limit bias

For meaningful estimates of density, a minimum diameter must be determined and consistently applied. Without an explicit lower diameter limit, any cited density is indefinite because including smaller trees will always increase that density. In various PLS studies, a diameter limit has been assumed based on surveyor instructions (e.g., 12.7 cm [5 inches]; 1833 instructions—Bourdo, 1956; Hanberry, 2021; Hanberry, Yang, et al., 2012; Manies & Mladenoff, 2000; Rhemtulla et al., 2009; White, 1983) or reported diameters (e.g., 5–15 cm [2–6 inches]; Bourdo, 1956; Cottam & Curtis, 1949; Grimm, 1984; Liu et al., 2011; Williams & Baker, 2011). Unfortunately, these diameter limits are absolute minimums and do not account for selective bias against trees larger than that limit. Regardless of the inconsistent minimum diameter in PLS surveys, or the lack of evidence of consistent adherence to that limit, the traditional method of correcting diameter bias is to assign a reference diameter, not recalculate the estimate (e.g., Cox & Hart, 2015; Dyer, 2001; Kronenfeld & Wang, 2007; Rhemtulla & Mladenoff, 2010; Zhang et al., 2000).

### Distance to center of the witness tree correction (ρ)

The correct distance metric for all density analyses is the distance from the corner to the center of the tree—its true point position. If the analyst does not add half the diameter of the tree to the surveyor's post-to-tree distance, the distance to the center of the tree is confounded. Of the 194 PLS density publications reviewed for this paper, only six researchers (i.e., Anderson et al., 2006; Bouldin, 2008; Bourdo, 1954, 1956; Fralish et al., 2002; Paciorek et al., 2021; Williams & Baker, 2011) explicitly indicated that the analyst added the tree radius to the measured distance parameter. When the squared distances are averaged and inverted, there will be a decrease in the density estimate if the radii are added (Ashby, 1972; Bourdo, 1956). The error for radius omission will depend on the nearness to the post, tree diameter, tree density, and the estimator used. For example, in Midwest subregions, the average correction for radius bias using the Morisita estimator is $\rho = 0.864$ ($n = 12$, range $\rho = 0.916$–$0.776$) for an average 16% overestimate (PalEON data). Adding the empirical radius to the distance to tree before any other calculations will produce density estimates without radius bias. In northern Wisconsin interior-quarter corners, Morisita calculations already include radius ($\rho = 1.000$), so the potential radius correction ($\rho = 0.891$) is unnecessary.

## DETERMINATION OF SURVEYOR BIASES

This paper presents a comprehensive framework for density estimation for historical Midwestern surveys. The primary density calculation is based on the post-to-tree-center distances measured by the surveyors. This raw density is corrected for various biased surveyor choices of witness trees due to their position, size, or species. For each bias type, the theoretical context within PLS surveys, its previous application, method of identification, generic quantitative determination, and then correction are addressed in the separate subsections that follow below. Analytical derivation of estimated biases is verified by comparing with simulations of sampling various tree spatial patterns under different densities, designs, and bias assumptions (*Data Availability Statement*, PlotlessPatternSimulationver-4.R). Then the specific calculation of the bias is demonstrated for the same area of northern Wisconsin serving as a guide for practitioners. Finally, this formulation is extended to estimate the surveyor biases and yield density estimates across the Midwest.

### Design bias

To avoid inappropriate calculation, the actual survey design must be confirmed before imposition of any correction. Many studies have adjusted density estimates for the so-called quadrant bias where witness trees deviate from one tree per quadrant with a random distribution among all quadrants (e.g., Bouldin, 2008, 2010; Kronenfeld & Wang, 2007; Manies, 1997; Warde & Petranka, 1981). The most widely applied adjustment is for the Ward assumption—two nearest quadrants (2nQ) design (Anderson & Anderson, 1975; Bouldin, 2008; Cogbill et al., 2018; Ward, 1954). Alternative restrictions are based on sampling designs using the absolutely nearest trees regardless of quadrant (Kronenfeld & Wang, 2007) or assuming a fixed proportion of witness trees in combined adjacent (same-sided plus adjacent-across) quadrants (Anderson et al., 2006; Anderson & Anderson, 1975; Bouldin, 2008; Edgin & Ebinger, 1997; Hanberry, Yang, et al., 2012).

### Design bias correction (κ)

Corrections for sampling designs (κ) accommodate deviation from the sampling model fixed in the assumptions of the PDE the analyst utilizes. If the sampling design is two trees in opposing halves (2oH) or four trees in quarters (point-centered quarter, PCQ), the Morisita PDF model is congruent with the design and the base density is statistically unbiased ($\kappa = 1$). When the design deviates by including more than one tree in an equiangular sector, the Morisita equation overestimates density and $\kappa < 1$. The two absolute nearest trees (N2) sampling contains extra trees in the same half (dup plus same-sided orientations) with a design correction of $\kappa = 0.723$ (theoretically; Thompson, 1956) to $\kappa \approx 0.749$ (R simulation sampling of a CSR population in *Data Availability Statement*). If the surveyor used a two nearest quadrant design (2nQ, Ward assumption), there are excess trees (~33%) in the same-sided orientation yielding an overestimate of density with $\kappa \approx 0.824$ (R simulation of sampling of a CSR population with the Morisita PDE; Cogbill et al., 2018). The restriction of witness trees to one predetermined side of the section line (2sH) completely truncates one half of the sampling area and a $\kappa = 2.000$ (derived from theory confirmed in R simulation) corrects the resulting underestimate of density.

Some 94.8% of all treed corners in the Midwest conform to either 2- or 4-tree corners and the remainder are odd designs represented in 1- or 3-tree corners. Exceptions to the standard sampling designs occur when there are multiple witness trees (dups) in the same section/quadrant (2.5% of all Midwest treed corners); trees are too far (1.3%); trees have omitted distances (3.9%); or include a corner tree (0.27%) (Tables 2 and 3; PalEON data). In northern Wisconsin interior quarter-section corners, 94.0% had two trees in predominantly opposite orientations (SAD: 13:26:60; consistent with mixed opposite halves—2S, 2oH). Thus, the dominant survey design was two halves ($\kappa \approx 1$) and harmonious with Morisita PDEs with negligible design or quadrant biases (*Data Availability Statement*: Northern Wisconsin Case Study).

## Pair angle bias

The empirical evidence of the sampling geometry at 2-tree corners is the aggregated distribution of the difference (pair angle) between the two bearings of witness tree pairs. At 2-tree corners, quadrants occupied by witness trees (SAD ratio) represent a crude approximation of sample design, but the empirical bearings of witness trees around the corner are precise evidence for the position of the trees sampled by surveyors. In contrast to corners with four trees in separate sections, surveyors had a flexible choice of two trees at 2-tree corners (Figure 2). Considering the nearest tree (N1) as one of a pair, the second tree of the pair can be in seven different nearest positions depending on the sampling geometry: (1) the absolute second nearest (N2, "dup" if in same quadrant); (2) the second nearest quadrant (2nQ); (3) the second nearest on the same side of the direction-of-travel (2sH); (4) the nearest on the opposite side of the section line (2S); (5) the nearest in the opposite semicircle (2oH); (6) the nearest in the opposite quarter (2oQ); or (7) the nearest in a restricted opposite acute angle (2O, "nearly opposite"). Any other choice would bypass nearest trees and violate the "nearest tree conjecture" reflecting surveyor bias.

The "pair angle" parameter gives the spatial geometry between 0° and 180° for each pair and in aggregate indicates the sampling design independent of quadrants (Figure 3). The pair angle is reciprocal and does not differentiate between a controlling tree and its second. Moreover, the absolute nearest tree may be excluded because of its size, relation to the survey lines, or other overriding bias. Regardless of distance rank in the original quadrant orientation, each sampled tree is the nearest within its actual, perhaps restricted, angular sector. Rather than a crude restriction to fixed quadrant (SAD) classes, the pair angle distribution is continuous and flexible without reference to absolute azimuth, direction-of-travel, or section boundaries.

The seven 2-tree scenarios are templates of possible sampling designs (Figure 2). Expectations are derived from modeling of geometric probabilities. In practice, there is a continuous variation of designs reflected in the pooled frequency from random within equal half (180°, semicircular) sectors to moderately restricted to a nearly opposite acute (<90°) sector. In graphs of aggregated angles, the former has an frequency of a ramped line (long-dashed line in Figure 3, $y = \{2 \times (N/c) \times x\}/180°$, where $N$ is the total number of pairs, $x$ is the midpoint of the angle class, and $c$ is the number of classes) from none at 0° to maximum at 180° and the latter has a steeply sloping line (solid line $y = \{4 \times (N/c) \times (180° - x)\}/180°$) starting at 0 near 90° and rising without a plateau to a maximum at 180°.

**TABLE 3** Empirical 2-tree sample designs in quadrants from exterior and interior corners in public land surveys in the Midwest.

| | | Exterior 2-tree | | | | Interior 2-tree | | | |
|---|---|---|---|---|---|---|---|---|---|
| Region | Date | Dup. (%) | Same (%) | Adj. Acr. (%) | Diag. (%) | Dup. (%) | Same (%) | Adj. Acr. (%) | Diag. (%) |
| Ohio | 1786–1808 | 4 | 24 | 22 | 50 | 4 | 18 | 19 | 59 |
| S Indiana | 1799–1825 | 3 | 19 | 20 | 58 | 3 | 15 | 22 | 60 |
| N Indiana | 1826–1846 | 10 | 74 | 6 | 9 | 6 | 15 | 25 | 53 |
| Illinois | 1804–1849 | 3 | 20 | 22 | 55 | 3 | 21 | 22 | 54 |
| S Michigan | 1815–1835 | 2 | 80 | 2 | 17 | 3 | 3 | 42 | 52 |
| NLP Michigan | 1836–1857 | 1 | 88 | 3 | 8 | 1 | 11 | 16 | 73 |
| UP Michigan | 1840–1855 | 2 | 64 | 9 | 25 | 1 | 7 | 10 | 82 |
| S Wisconsin | 1832–1845 | 4 | 79 | 12 | 5 | 3 | 19 | 26 | 53 |
| N Wisconsin sec | 1846–1866 | 2 | 71 | 21 | 6 | 0 | 14 | 20 | 66 |
| N Wisconsin qtr | 1846–1866 | 1 | 8 | 34 | 57 | 1 | 13 | 27 | 59 |
| Minnesota sec | 1847–1907 | 5 | 68 | 20 | 7 | 1 | 42 | 25 | 32 |
| Minnesota qtr | 1847–1907 | 1 | 14 | 34 | 51 | 1 | 14 | 30 | 56 |

*Note*: S, south; N, north; LP and UP, Lower and Upper Peninsula; sec, sections; qtr, quarter-sections; dup., duplicate pair in the same quadrant; same, pair in adjacent quadrants on the same side of survey line; adj.acr., pair in adjacent quadrants across the survey line; diag, diagonal pair in opposite quadrants on different sides and not adjacent. Pair orientation relative to direction-of-travel on the survey line and based on bearing parameters from 80 representative county-scale subsets (244,720 corners) of the PalEON database.

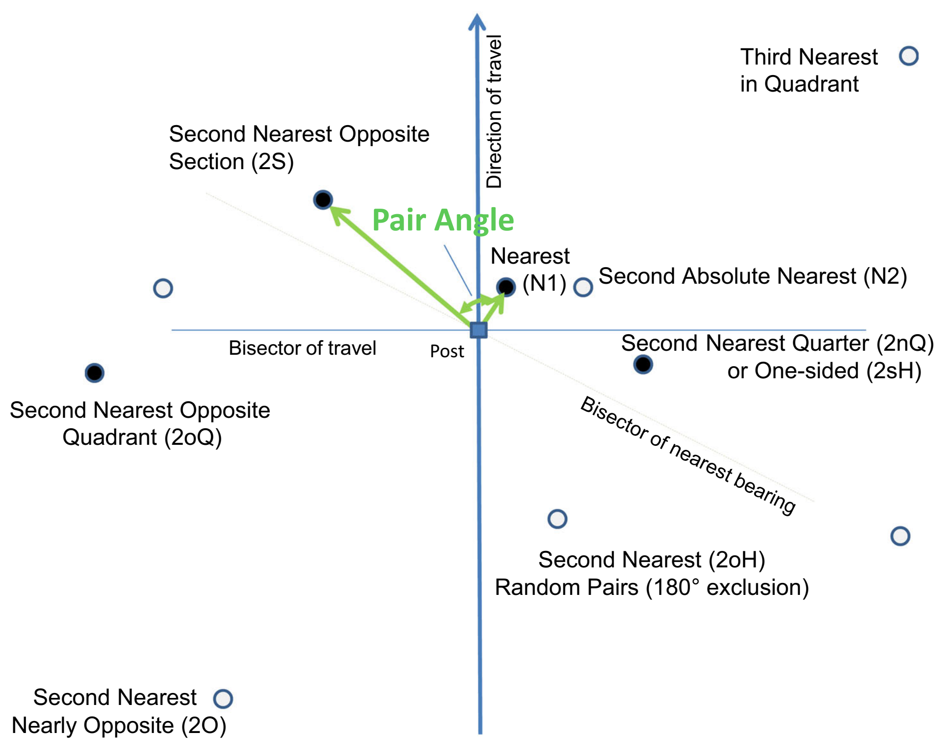


**FIGURE 2** Geometry of various potential nearest witness trees around survey corner. Solid circles are the nearest trees in the four quadrants and open circles are additional witness trees including various second nearest possibilities under different sampling designs. The pair angle illustrates a 2-tree pair: the absolute nearest (N1) and the second nearest on the opposite side of the line-of-travel (2S design).

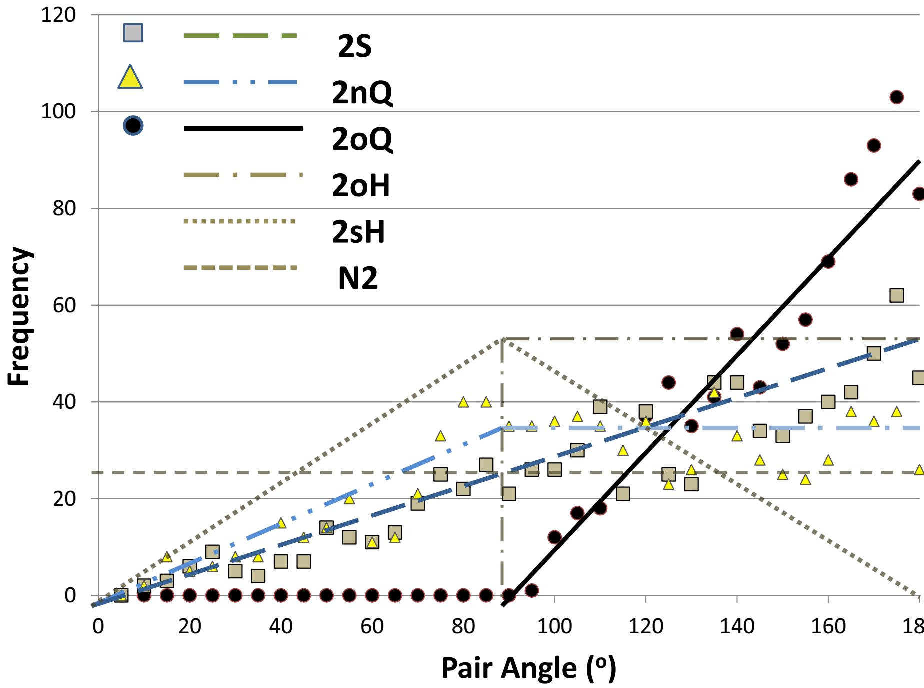


**FIGURE 3** Pooled pair angle observations under different sampling design scenarios. Lines are the theoretical distribution derived from analogue geometric probability models of pairs in an isotropic population. Symbols are results of simulated sampling of random stands with 20 replications at 50 corners (Cogbill et al., 2018; R Core Team, 2019). Design scenarios include section across direction-of-travel (2S, squares, ramp); two nearest quadrants (2nQ, triangles, rise to plateau); opposite quarter (2oQ, circles, steep ramp >90°); opposite half regardless of the section lines (2oH, step function at 90°); one-sided (2sH, circumflex); absolutely two nearest (N2, flat line).

Pair angle analysis of PLS samples refines the sampling design used in that survey. For curves representing deviation from the discrete opposite half ($\kappa = 1$) model, the fit rotates from an opposite half to a highly restricted opposite quarter model as bias increases (Figure 4). Thus, the theoretical frequency line becomes steeper and the *x*-intercept increases (Figure 3). Interestingly, the degree of shifting, that is bypassing of nearest trees due to surveyor bias, results in a family of curves where the equal halves expectation and the empirical fit lines initially cross at 120° and approach crossing at 135° with restriction to the opposite quarter. Graphically, trees with small pair angles, below the left dashed line (light shaded area in Figure 5), are at least partially bypassed and are replaced by trees with a large pair angle, above the right dashed line (dark shaded area). The degree of bias is simply the geometrically determined proportion of either of these opposing triangles (bypassed trees equal the number of replacement trees) or half the total shaded areas. Any bypassed nearest trees are censored. Because only one tree of each pair must be bypassed to change the pair angle, the percentage of bypassed trees is half the percentage of switched pairs in the model. Referring to the original azimuths, this can be envisioned as one tree in half of the sample circle and the second transposed to an uncensored sector in the second half. The proportion of switched pairs is equivalent to the amount of spatially censored (unsampled) angles ($\varepsilon = 180° \times$ % bypassed trees/100). Because each sampled tree is the nearest in its own sector, the degree of bias equals the proportion of the total uncensored sectors within the entire circle ($\Sigma\alpha = 360° - \varepsilon$). As the Morisita II density is based on the distance to the sampled trees within an assumed semicircular sector, the absolute density is underestimated due to the replacement of the nearest tree with a farther tree in a more restricted sector.

## Pair angle bias correction ($\theta$)

The pair angle biased density calculation can be corrected to the actual density per full unit area by a bias factor ($\theta = 360°/\{360° - \varepsilon\}$) (Kronenfeld & Wang, 2007; Morisita, 1954). If the proportion shifted ($\varepsilon$) is 0, the frequency of the pair angles follows the dashed line (Figure 4) and $\theta = 1.00$ because the design is in perfectly opposite halves (2S or 2oH). On the other hand, if 50% of the pairs contain a shift (e.g., same-sided or adjacent-across quarter trees are replaced by a diagonal opposite-quarter tree), the frequency follows a line from 0 at 90° to twice the halves maximum at 180° (solid line in Figure 4). Here the design inferred is a 90° sector (not

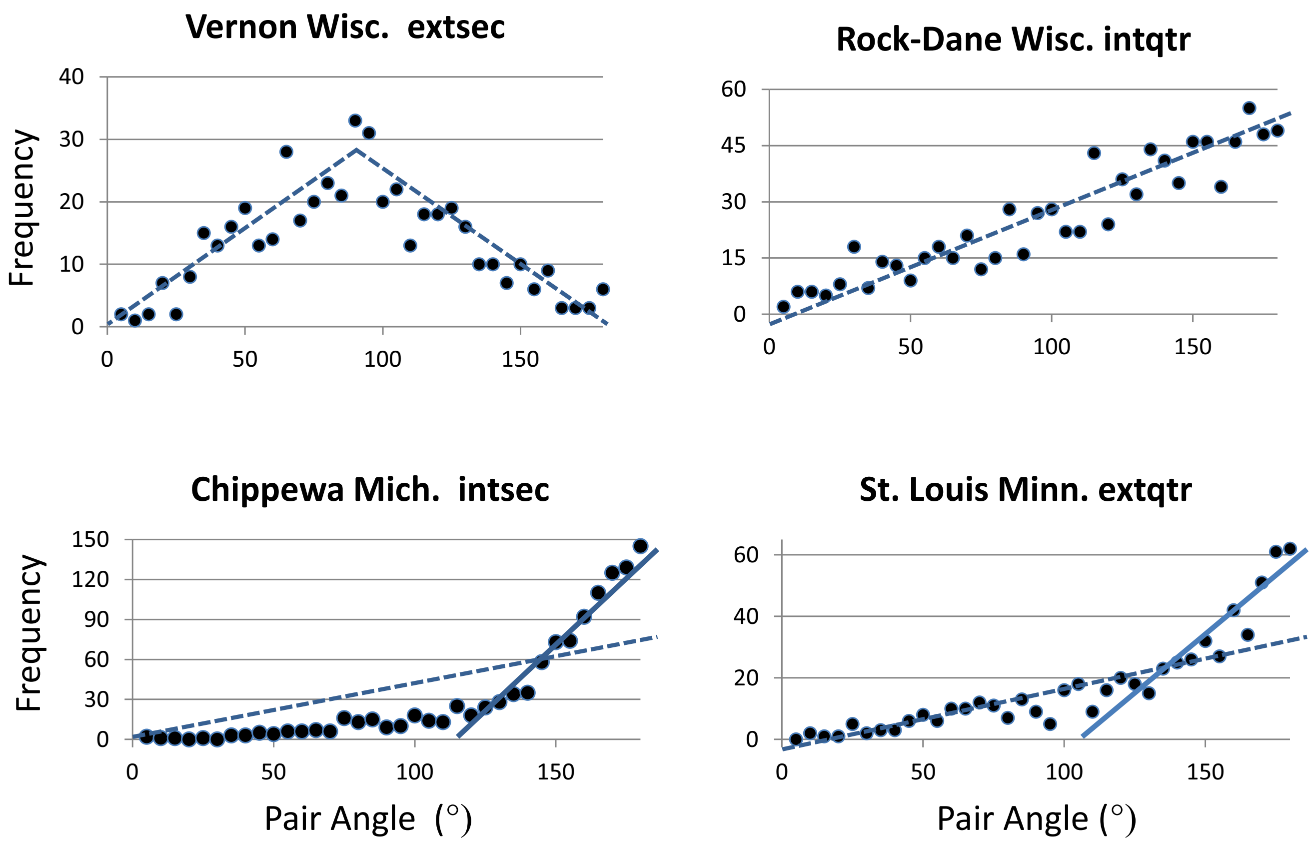


**FIGURE 4** Selected pooled pair angle frequencies from empirical public land survey regional subsamples. Short dashed line is the theoretical fit for separate halves (2S) design and the solid line is the empirical regression fit for opposite quarter (2oQ) design. Sample designs inferred for Vernon, Rock-Dane, Chippewa, and St. Louis are 2sH, 2S, approaching 2oQ, and hybrid 2S-2oQ, respectively.

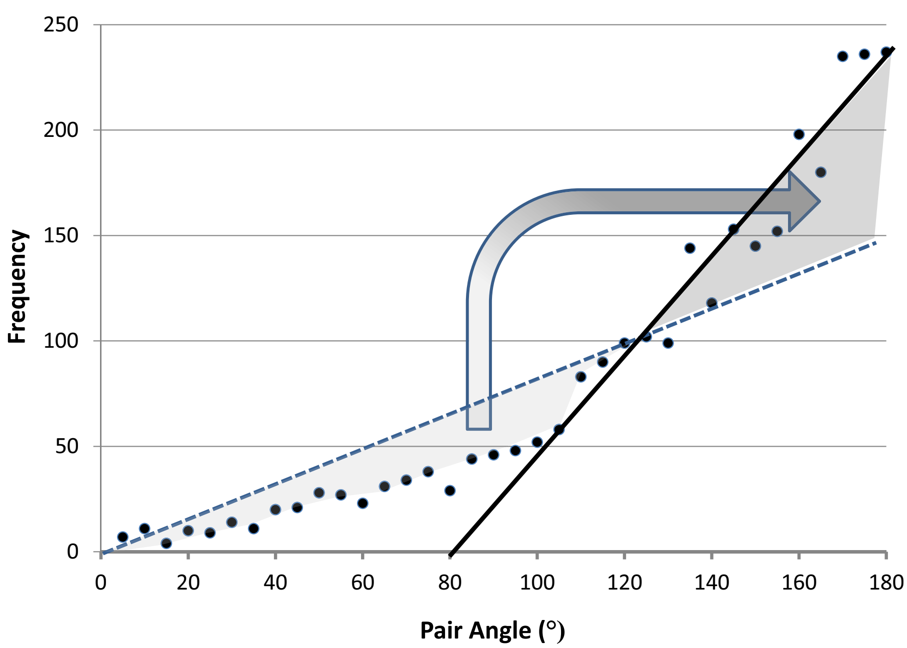


**FIGURE 5** Graphical analysis of surveyor bias in sampling design from 3579 interior quarter-section corners in public land survey from northern Wisconsin. The circles are the empirical frequency while the short dashed line is the theoretically expected frequency for two opposite sections (2S) and the solid line is the expected frequency for the opposite halves (2oH) design. The light shaded area indicates bypassed trees, and the dark shaded area is that to which the trees were shifted (shaded arrow).

necessarily a quadrant) and the expected θ correction is 1.33 (coupled with $\kappa = 1$). The bypassing of trees is a continuous variable from partial switching leaving some low-angle pairs ($1 \le \theta \le 1.33$) to extremes with highly restricted nearly opposite pairs ($1.33 < \theta < \infty$).

When a sampling design includes more trees in the same half (2nQ, N2, 2sH) than present in the strict 2-tree two-halves model, the transposition of trees is reversed with some large pair angles graphically moved to the area of small angles (Figure 3). Surveyor bias then forms a continuum from a few (2nQ) pairs, through half (N2) pairs, to all (2sH, "circumflex") pairs shifted to the same side. The corrections for these designs are initially reflected in the κ factor, but a fine-tuning of the design involves a modification of the κ factor by the θ factor. With exterior corners, the pair angle pool is occasionally prorated between purely same-sided corners inside the township (2sH, $\kappa = 2$) and the noncompliant corners outside the township (2S, $\kappa = 1$). Then the 2-tree design correction ($\theta = 1 - \{[\%\text{ adjacent-across plus diagonal}]/100\}$) reflects the proportion between the one-sided and opposite side corners. The specific design correction for composite exterior lines with the right limb of the circumflex elevated is the design κ (=2) times the pair angle (θ) bias.

In practice, the determination of the percentage of bypassed trees due to pair angle bias is calculated as the discrete summation of the difference between the theoretical equal halves (2oH) line and the pooled pair angle values (Figure 3; Table 2). In northern Wisconsin, half of the sum of the two shaded areas indicates that 15.4% of the pairs had a bypassed tree. This implies that 7.7% (0.154/2) of all the trees were bypassed. This bias can be corrected by a sampling design factor $\theta(1/[1 - \%\ \text{shifted}/100]) = 1.083$. The replacement trees are restricted to an equivalent of a 166° sector (supplement of $0.077 \times 180°$). The 13.9° sector of bypasses for the second tree is taken away from an equal half, and thus underestimates the actual density if calculated with an equal half equation (2oH). Nearly the same rotation from equal halves sampling is also found in the mean pair angle of 134.2°, while an average pair angle of 120° is found in an R simulation of sampling of unbiased, equal halves witness trees in a CSR population. The average pair angle deviation from 120° can then provide a preliminary index of the pair angle bias.

The corresponding quadrant SAD ratio of 13:26:60 in northern Wisconsin also implies a sample design with witness trees switched from a strictly opposite section (2S) to a more directly opposite position. The 10.5% excess witness trees in the diagonal quadrants infer a crude surrogate of pair angle correction equal to $1 + 0.105 = 1.105$. Together with 26% of the 2-tree corners with adjacent-across orientation, this indicates a strong nonexclusive preference for witness tree pairs across the section line. The pair angle analysis reiterates the general sample design of two trees in opposite halves (2oH, $\kappa = 1$), but it refines the correction to a +8.3% deviation ($\theta = 1.083$) from that strict design.

## Azimuthal bias

The empirical evidence for azimuthal bias in historical PLS samples is the aggregated distribution of the witness trees' bearing directions (Figure 6). The underlying spatial pattern of the trees is assumed to be independent of direction from the corner (isotropic), so even if not CSR, witness trees are expected to have random bearings. Early surveys from New York to Minnesota have a deficit of witness trees near the cardinal bearing directions and a concomitant surplus of witness trees in the ordinal (center of quadrants) bearing directions (Anderson et al., 2006; Bouldin, 2008; Goring et al., 2016; Hanberry, Yang, et al., 2012; Kronenfeld & Wang, 2007). Obviously the surveyors at least partially avoided using witness trees along the surveyed lines. This bias is a logical consequence of clearing the line-of-travel for sight lines and an unobstructed path for running the chain by the chainmen. Additionally, the line-of-travel was blazed by "line" trees up to 3 m on either side of the section line (1833 instructions; White, 1983). Keeping witness trees separated from the section line prevented confusing the witness trees with blazed line trees or the section within which they lay. Past analyses, however, have only sporadically considered the line-of-travel (i.e., Clark, 2000; Kronenfeld & Wang, 2007; Manies, 1997) or corner types (i.e., Anderson et al., 2006; Bouldin, 2008; Bourdo, 1956; Liu et al., 2011). It is initially unclear whether azimuthal bias is found at all cardinal compass directions, just the line-of-travel, all section lines, the perpendicular initially unsurveyed quarter-section lines, and/or equally at particular corner types. More importantly, the angular bias (width) and its degree (proportion affected) of each potentially bypassed sector are undetermined.

Conceptually, azimuthal bias assumes that bypassed witness trees are censored in wedge-shaped sectors (ω) and transposed to complementary uncensored (α) sectors (Figure 7). The maximum bias is expected at small angles near the section lines. In addition, the bias width is seldom fixed, and decreasing censoring will tend to blend into the inherent variability of the unbiased sector. The amount of bias is a trade-off between a strictly limited sector minimizing the censored area and an expanded uncensored area. This limits the biased trees remaining (undetected) in either the censored or uncensored areas.

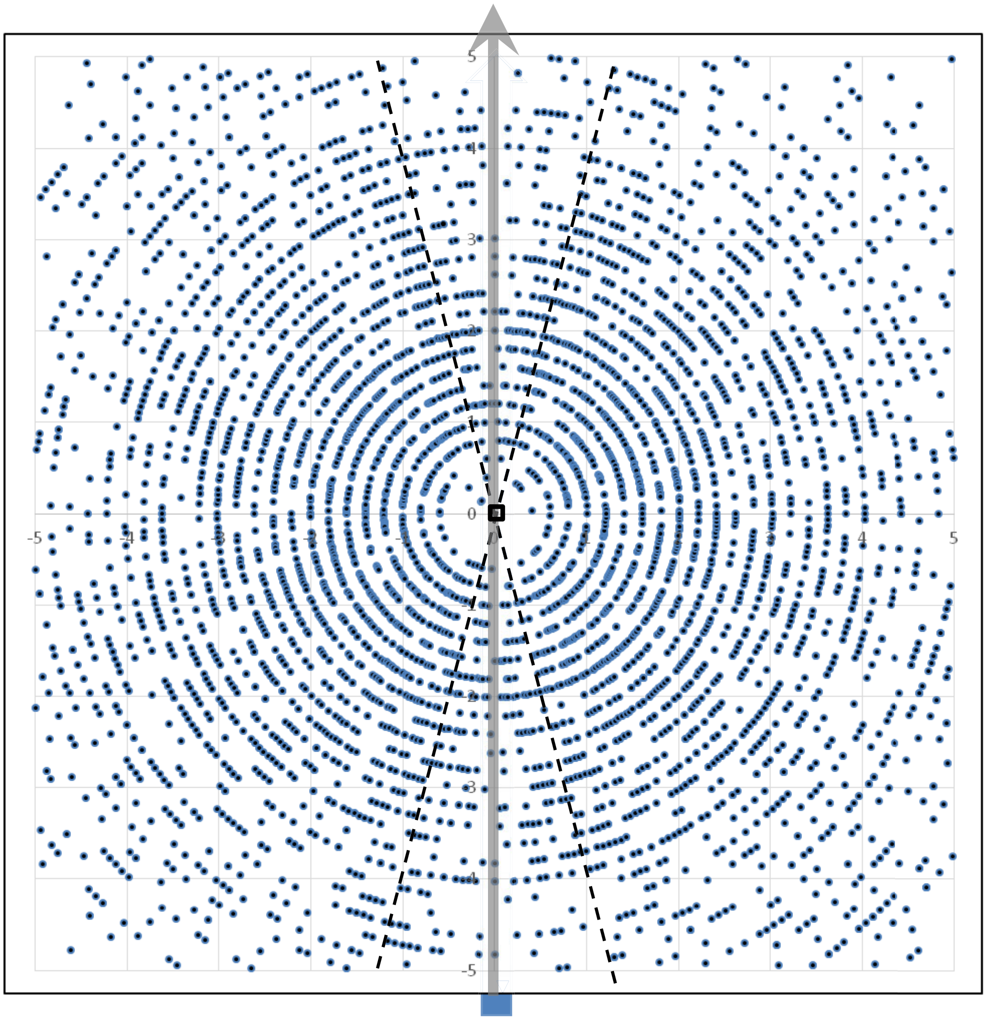

**FIGURE 6** A 10 × 10 m detail of the composite geometry of witness trees (circles) around corner posts at 3579 2-tree interior quarter-section public land survey corners in northern Wisconsin. Gray arrow is the line-of-travel of the survey along the section line between two sections. The dark square is the post and the dotted lines are 15° from the line-of-travel ($\omega = 60°$, "hourglass"). The mean distance to all witness trees is 5.19 m, the median distance is 3.71 m, and 95% are within 10.8 m of the post.

If the actual bias width is less than the set width, any error will tend to be minimized by trees transposed to a wide uncensored sector. If the width of bias is greater than the standard, the weaker bias at greater angles will likewise minimize error. This is illustrated by 2-tree interior-quarter corners from northern Wisconsin (Figure 6) where bias (shifts detected) for the 30° wedge along the line-of-travel (4.9%) is maximized compared with either larger (i.e., 4.3% at 40°) or smaller angles (i.e., 3.8% at 20°). This assumption implies an unbiased sector ($\alpha$) of 30° per quarter-section ($\Sigma\omega = 120°$) or "Maltese Cross" bias at section corners (Figures 7a and 8). For most quarter-section corners, the perpendicular unsurveyed quarter-section line (90° and 270° to line-of-travel) has minimal azimuthal bias and the bias is only within 15° per quarter ($\Sigma\omega = 60°$) or "hourglass" bias along the line-of-travel (Figures 7b and 8).

The generic method of quantifying surveyor bias determines the difference between proportion of witness trees expected to occur in a sector and those observed in that sector (Kronenfeld & Wang, 2007). The missing trees are framed as trees bypassed by the surveyor and biased sampling of their replacement. With azimuthal bias, the angular proportion is a convenient proxy for expected unbiased frequency (% expected = 100 × angle°/360°, confirmed in R simulation) in the wedge. If the nearest potential witness tree within a censored ($\omega$) sector near a section line is bypassed because of its position, it is replaced by a farther tree in the uncensored ($\alpha$) sectors (Figure 9). Specifically, azimuthal bias identifies the proportion of trees remaining in set ($\omega = 90° - \alpha$) sectors with a significant depletion of witness tree frequency from that expected along the cardinal directions (Kronenfeld & Wang, 2007).

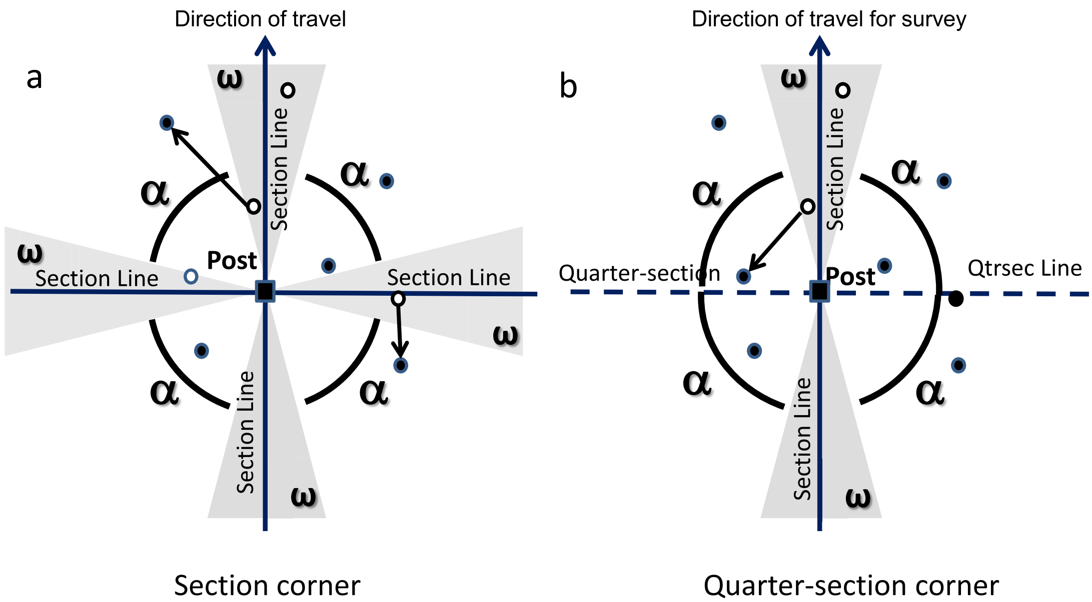


**FIGURE 7** Azimuthal geometry and potential bias around cardinal directions at public land survey corners. The left (a) is a typical section corner ("Maltese Cross" bias) and the right (b) is a typical quarter-section corner ("hourglass" bias). The α sectors are unbiased while their shaded complement ($\omega = 90° - \alpha$ per quarter) are potentially censored. Solid circles are available witness trees while open circles are bypassed trees and replaced by tree at arrow.

## Azimuthal bias correction (ζ)

The initial density calculated from a restricted sector is an underestimate because only a subset of the full area is sampled (cf. Hanberry, Yang, et al., 2012; Kronenfeld & Wang, 2007). Nevertheless, all sampled trees whether remaining in the biased sector, transposed, or already in the uncensored sector are the nearest trees in their own particular sector. Thus, no further correction of density for the distance rank is necessary. If completely restricted to the unbiased sectors, the correction for density due to azimuthal bias will be the reciprocal of the uncensored area $\zeta = 1/\Sigma\alpha$ (Morisita, 1954). If partially bypassed (some witness trees remain in the censored sector), the correction effectively prorates the complement of trees in the censored sector (% expected − % remaining) and the overall azimuthal bias is the ratio of the proportion of trees in the unbiased sector divided by its width ([% unbiased sectors/360°]/[% bypassed/100]; Kronenfeld & Wang, 2007). The azimuthal correction is then the reciprocal of bias ($\zeta = 1/\text{bias} = [1 - \%\ \text{remaining}/100]/[\Sigma\alpha/360°]$).

Calculation of azimuthal bias depends on the width of the censored sector and the proportion of trees left remaining in that sector. Azimuthal bias is neither a step function at a set width (e.g., 30° wedge) nor regularly distributed within a censored sector. The bypassing is partial and diminishes as the azimuthal width (angle around the section line) increases. A fixed width is a compromise between incorporating the majority of the bias and not including unbiased sectors. For example, at interior-section corners in northern Wisconsin, a bias wedge is obvious in all four cardinal directions, but only in the line-of-travel at quarter-section corners. Conceptually in graphs of azimuth frequencies, the surveyor bypassed trees in the lightly shaded areas near the section lines and they were switched (arrows) to the dark shaded areas (Figure 9). The two 30° wedges ($\Sigma\omega = 60°$) at quarter-section corners have 12.34% witness trees remaining, while the four biased sectors at section corners ($\Sigma\omega = 120°$) total 17.8% nonbypassed trees. The resulting azimuthal bias correction factor for density calculated from the bypassing of 4.33% of the nearest trees at quarter corners is $\zeta = 1.052$. For 4-tree section corners, the correction for bypassing trees is $\zeta = 1.231$. Furthermore, corrections for St. Louis Minnesota, Northeast Indiana, Chippewa Michigan, and Menominee-Shaw Wisconsin in Figure 8 are $\zeta = 1.23$, 1.00, 1.23, and 1.07, respectively. The corrections at section corners are substantially greater than those at quarter-section corners, in part due to avoidance of a doubled area. Azimuthal bias at section corners throughout the Midwest is symmetric in the four cardinal directions ($\Sigma\omega = 120°$, "Maltese Cross"). Quarter-section corners in northern Michigan have the same four-sector ($\Sigma\omega = 120°$) bias, while all other Midwest quarter-section corners have azimuthal bias, sometimes minimal, just along the line-of-travel ($\Sigma\omega = 60°$, "hourglass").

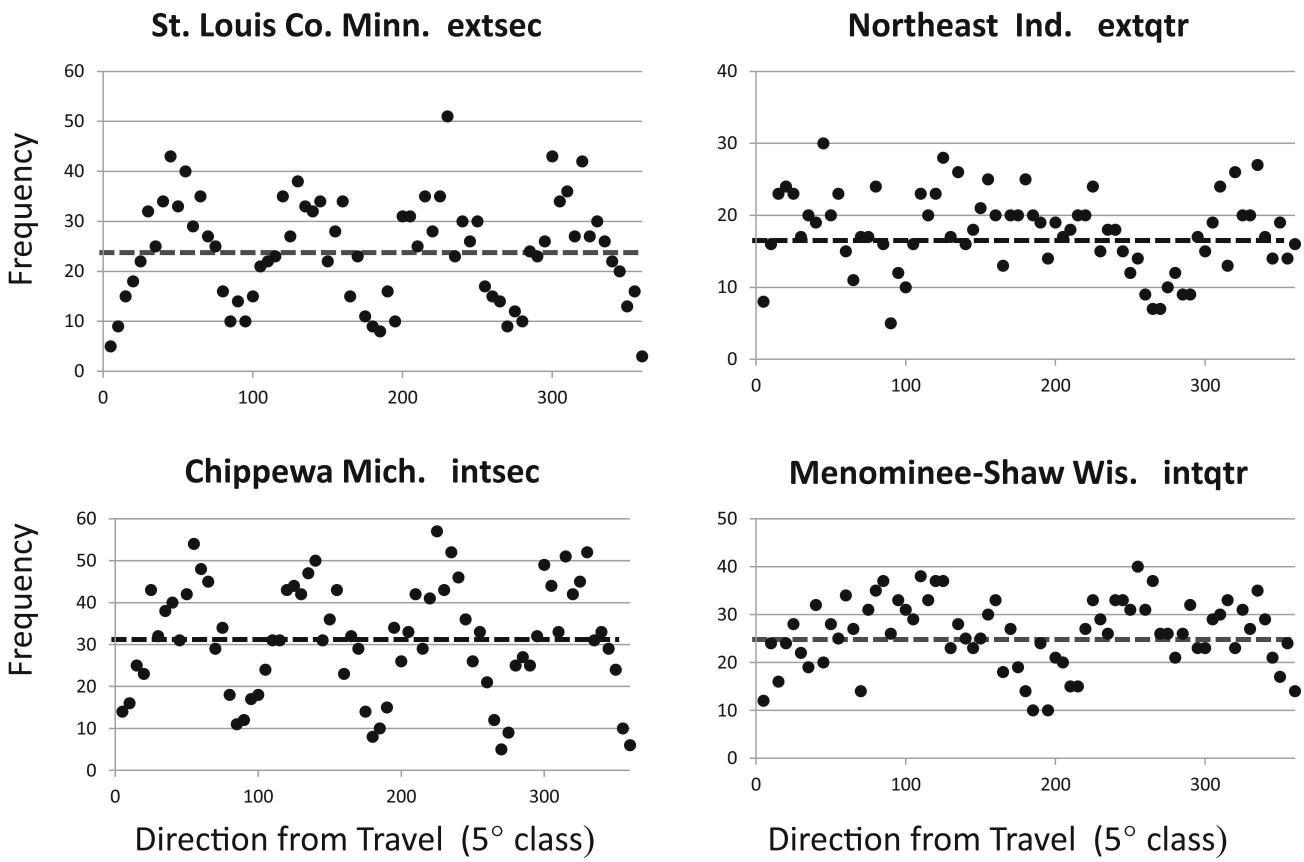


**FIGURE 8** Azimuthal frequencies from selected public land survey regional subregions. Solid circles are frequency of trees in 5° bearing classes relative to the line-of-travel (360° forward or 180° backward). Dashed line is theoretical unbiased expectation with a 16% CV (determined in random R simulation). Patterns displayed are "double M" of all cardinal directions (St. Louis County, MN: 433 exterior-section 4-tree corners and Chippewa County, MI: 1123 interior-section 2-tree corners), "single M" of travel direction only (Menominee-Shaw, WI: 941 interior-quarter 2-tree corners), and relatively flat with low bias (Northeast Indiana: 634 exterior-quarter 2-tree corners).

Analysis of interior-quarter corners from northern Wisconsin indicates that the bias due to azimuth angle increases as the bias angle increases, reaching an empirical plateau ($\zeta \approx 1.055$) at 15° on either side of the section line ($\Sigma\omega = 60°$, Figure 7) with a peak ($\zeta = 1.058$) for a 17° angle from the line ($\Sigma\omega = 68°$). The 15° bias angle on either side of the survey line (6 − 5° classes ≡ 30° wedge) is the minimum value, which captures nearly maximum net bias around the line-of-travel (Figure 8). Although the degree of bias varies by site, this critical bias angle is shared throughout the Midwest.

## Azimuth bias alternative: Transect section-line correction ($\xi$)

Azimuthal bias is just a convenient approximation of the association of bypassed trees with the section line. The more rigorous model is the distance from the line-of-travel, not the angle in which breadth increases with distance from the post. Thus, a rectangular transect spanning either side of the section line represents an alternative to an angle for estimating section-line bias. As with azimuthal bias, the difference between the expected number of witness trees in the transect and those actually observed is considered bypassing of trees, which are then replaced by another tree outside of the biased swath. With azimuthal bias, the expected number of witness trees in a pie-shaped wedge is directly proportional to the angle. Within transects, the expected number is not a simple function of the area of the swath, but varies with the underlying density and spatial pattern of the trees. The expected number of witness trees in the rectangular transect is determined by R simulated sampling (R Core Team, 2019; in *Data Availability Statement*). In northern Wisconsin 2-tree interior-quarter corners, 18.86% of all witness trees are found within a 2-m-wide transect along the surveyed section line, while the R simulation of the representative InH$^2$ population yields an expected 27.87% of the trees in that swath (Appendix S1: Table S1; Cogbill et al., 2018). The estimated 9.02% net bypassed trees indicate a "transect" correction factor of

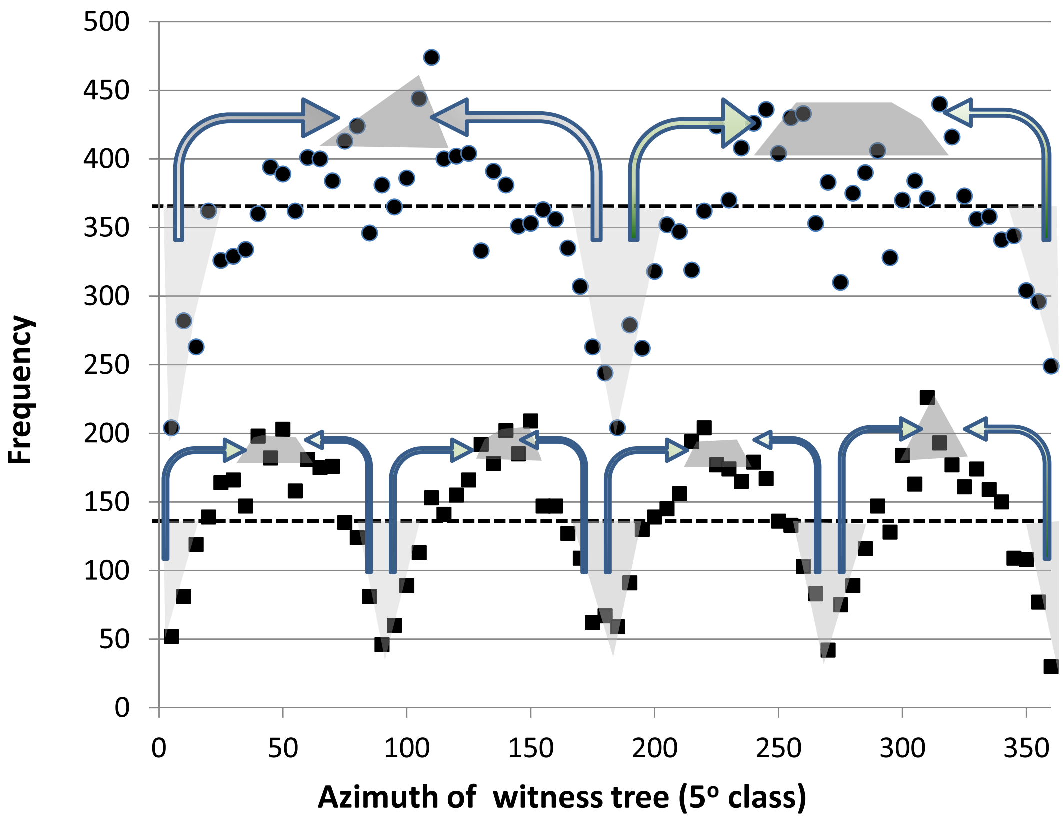


**FIGURE 9** Graphical analysis of azimuthal bias. Solid circles in the upper panel are 25,827 interior quarter-section trees and squares in lower panel are 9972 interior-section trees from public land surveys in Wisconsin. Azimuth is relative to line-of-travel (360° forward). The light shaded area indicates bypassed trees (% expected − % remaining), and the dark shaded area is the area receiving the shifted (shaded arrow) trees.

$\xi = (1 - 0.189)/(1 - 0.279) = 1.125$ for bias over the total transect. This proportion increases with transect width reaching a plateau at 0.9–1.7 m from the section line (1.8–3.4 m wide transect). The maximum correction ($\xi = 1.134$) is for 1.3 m from the line-of-travel, but the general standard of 1 m from the section line is a conservative estimate of the width of the transect bias. In contrast, the total hourglass azimuth wedge indicates 4.36% bypassed trees and a correction of $\zeta = 1.052$.

Most of the 4.66% difference in bypassed trees between the azimuth and transect section-line estimators lies within 1 m of the post (Figure 10, red circle) that is completely within the transect swath (4.40% bypasses) and only minimally (0.64% bypasses) overlaps within the azimuth wedge (purple region). The censored areas sampled by the transect and azimuth methods are different (Figure 10) and the corrections are conditioned on a different geometry of switching. The estimate for section-line correction of the transect method beyond 1 m from the post ($\xi = 1.066$) is +1.4% more than the total azimuthal bias ($\zeta = 1.052$).

The transect along the section line has 5.2% bypassed trees beyond 1 m from the post, four times more trees bypassed than in the perpendicular future quarter-section lines (Appendix S1: Table S1). The azimuthal method also shows an exaggerated difference between the section line and the quarter-section lines that actually have negative bypass (a net movement of trees into the quarter-section sector presumably from the section line or other biases). In addition, the transect bias is not strictly symmetrical as 27% of the trees behind the corner are estimated to be bypassed, while 21% of the trees ahead of the corner are bypassed (Figures 6 and 10; Appendix S1: Table S1). Apparently, at interior-quarter-section corners in northern Wisconsin, the surveyors preferentially bypassed (perhaps removed) trees along the previously surveyed line, while they utilized a slightly greater proportion of witness trees in the not-yet-surveyed part of the section line. These patterns of section-line bias at 2-tree corners confirm a particular selective bias against trees in the 2-m-wide line-of-travel with much lower to nil bias on the perpendicular quarter-section lines.

The calculation of transect bias is dependent on an a priori estimate of density and requires simulation of sampling for expected proportions. The transect and azimuthal section-line methods demonstrate different bypass patterns but agree with a section-line bias of 22%–24% bypassed trees within the individual censored

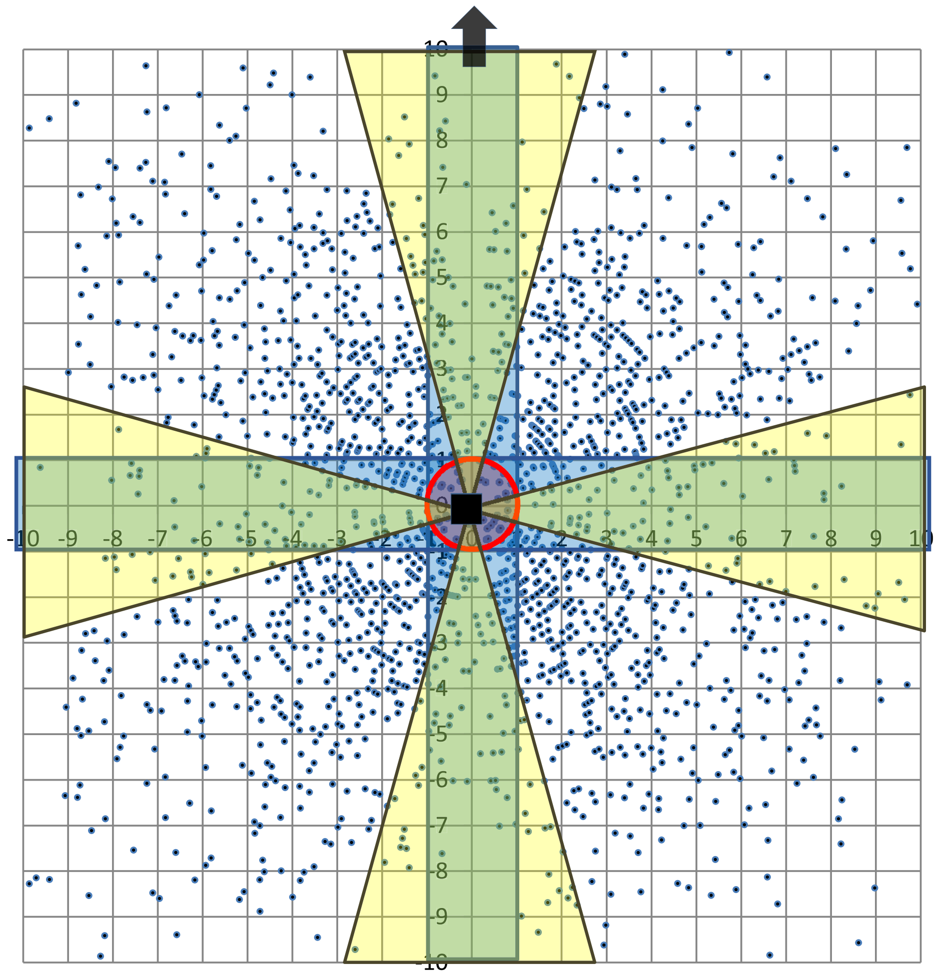


**FIGURE 10** A 20 × 20 m detail of the composite geometry of witness trees (dots) around corner posts at 599 4-tree exterior-section public land survey corners in northern Wisconsin. The mean distance for the 4-tree corners is 5.91 m, the median distance is 4.94 m, the mean distance to the nearest tree of the four is 3.09 m, and 95% of the witness trees are less than 11.9 m from the post. The small square is the post (scaled to 30 cm square) at the corner of four sections and the black arrow is the direction-of-travel of the survey. The colored areas are the three alternative sampling strategies to determine bias along the intersection of two section lines. The four yellow-plus-green triangles are 15° from the section lines and the angle azimuth ("Maltese Cross") bias area. The red-plus-purple and orange circle is the 2-m-diameter near-post bias area, which, when combined with the yellow, is the union ("Celtic Cross") bias sector. The two 2-m-wide blue-plus-green rectangles are the alternative transect section line ("Greek Cross") bias sector.

sectors (local %) beyond the near-post circle in northern Wisconsin (Appendix S1: Table S1). Although less discriminating, the azimuthal bias calculation does not require any assumptions of density, a spatial dispersion model, nor simulation of expectations, and thus is simpler and more easily applied than the transect method.

## Four-tree section-line bias

The 4-tree corner sampling design became established after 1837. At 813, 4-tree exterior- and interior-section corners in northern Wisconsin, the azimuth bias within the four-way 30° wedges around the section lines ($\omega = 120°$, 15.5% bypassed, 17.8% remaining) yields a correction of $\zeta = 1.232$ (Figure 10; Appendix S1: Table S3). The bias within the four wedges increases with widening bias angle reaching a maximum at 19° ($\omega = 152°$, 23.4% bypassed, $\zeta = 1.297$) and adds +6.6% to the azimuthal bias. This azimuthal expansion for 4-tree corners is not only potentially doubling the number of section lines over the 2-tree corners but also has proportionally more bypassing and an expanded wedge angle.

The qualitative estimate for transect section-line bias at 4-tree corners is a central circle with rectangular arms (Celtic Cross bias). Using two crossed 2-m-wide transects incorporating the line-of-travel and the subsequent perpendicular secondary section lines, 20.3% of trees remain in the transects beyond the central circle (Figure 10). This implies a $\xi = 1.231$ correction for the transect method, which is virtually the same as for the full azimuthal method. At 4-tree corners in northern Wisconsin, the two section-line methods demonstrate a roughly consistent 40%–45% local bypassing of trees in both azimuths and transects beyond the central circle (Appendix S1: Table S3). At 4-tree corners, there is more intense bypassing near the section lines beyond 1 m from the post (≈47% local bypassing in four arms) than at 2-tree corners (≈24% local bypassing in two arms; Appendix S1: Tables S1 and S3).

## Near-post bias

As found in transect bias, the area within 1 m of the post contains fewer witness trees in that small circle than other areas of the swath around the section line. It is not immediately evident, however, whether few trees is simply a function of density or the surveyor's choice. At interior quarter-section corners, the nearest witness tree (of two) has an observed mean distance of 3.38 m from the post. In contrast, an R simulation of a CSR pattern at 326–402 trees/ha density yields an expected mean for the nearest tree of 2.77–2.49 m, respectively. Much of the difference from the empirical observations is due to the non-CSR spatial pattern and the true density, but even sampling of a simulated InH$^2$ dispersion with a realistic gradient of densities and a mean of 402 trees/ha indicates an expected distance of 3.04 m for the absolutely nearest tree. The residual increased distance for chosen trees is consistent with surveyors regularly bypassing the tree nearest the post and sampling a farther tree.

The empirical evidence for this novel "near-post" bias is the distinct lack of distances less than 5 links (1.01 m) in the Midwest (Figure 6; PalEON data; Grimm, 1984). For example, distances from 7158 witness trees at interior-quarter corners in northern Wisconsin displayed virtually no 0–3 link (0–0.6 m) distances, few trees at 4–5 links (0.8–1.0 m), and increasing relative frequency to a peak near 10 links (2.0 m) (Figure 11). Employing simulations dramatically highlights the depletion of distances less than 1 m from the post regardless of density or spatial pattern (Figure 11; Appendix S2). The transposition of these trees then results in a peak frequency (at 1.5–2.5 m) above that expected in the simulations. Because short distances themselves signify very dense populations

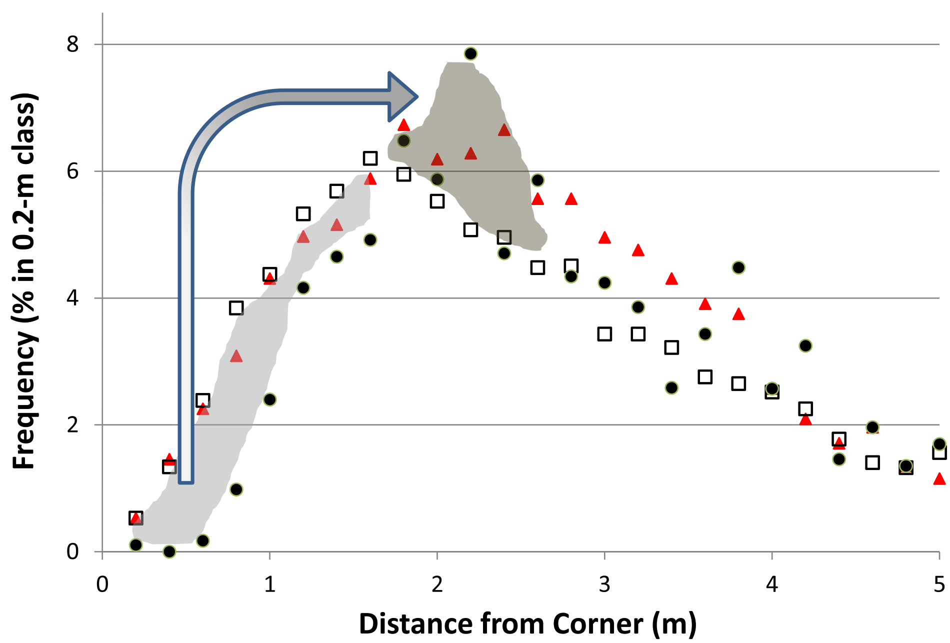


**FIGURE 11** Geometric analysis of frequency of witness tree distances from post and near-post bias in public land surveys (PLS) of interior quarter-section corners in northern Wisconsin (Figure 5). Solid circles are empirical PLS distances of witness trees (mean 3.38 m, median 2.70 m, 95% <7.05 m), red triangles are simulated sampling (R code in *Data Availability Statement*) of a random population (homogenous Poisson pattern of 402 trees/ha, distance: mean 2.75 m, median 2.50 m, 95% <5.18 m) and open squares simulated sampling of a dual inhomogeneous-inhibited (InH$^2$) distribution (gradient and 1.5 m between tree inhibition, with average 402 trees/ha, distance: mean 3.03 m, median 2.38 m, 95% <7.35 m). Graphically the gray area represents the shortfall of short distances (near-post bias) and the arrow indicates the replacement of censored trees at farther distances in dark area.

(<1 m distance implies >3000 trees/ha density), bypassing a nearby tree has the strong effect of decreasing (underestimating) PDE-derived density (λ). The shortfall of expected unbiased trees near the post validates that "near-post" bias exists and needs correction separate from azimuthal bias.

The survey corners were sites of intense activity by the entire survey team, and establishment of the corner post and section lines had priority over witness tree locations. The preparation of a well-marked corner left few witness trees in the immediate vicinity of the post. The lack of witness trees near the post is the logical result of several factors: corner trees replacing the post; trees cleared because they interfere with proper corner marking such as placement of the posts or incorporation of stone piles; or the tree's position ambiguous in relation to the corner or section line (overlaps *Azimuthal bias* above).

## Near-post bias correction ($\nu$)

As farther trees replace any "corner" or near-post trees, a near-post bias correction (υ) will offset the underestimate based on the nearest replacement. For example, at interior-quarter corners in northern Wisconsin, only 1.940% of the witness trees are found within 1 m of the post (Figure 6) while the simulations of an $InH^2$ pattern at 402 trees/ha indicate that 6.306% of the witness trees are expected to be within this distance (Appendix S2: Equation S3). This implies that 69.2% of trees ≤1 m from the post (4.365% of all trees) were bypassed. There are even proportionally more bypassed trees at shorter distances (94.5% of trees within 50 cm are bypassed). Since the ultimate near-post correction is disproportionate to the bypass total in the 2-m-diameter circle, the squaring and inverse of short distances have an inordinate effect on density.

The difference between the expected and observed frequency of witness trees quantifies near-post bias. Some 98% of the witness trees within 1 m of the post are the nearest in the pair and if all are switched to the sector's next (second) nearest tree in Morisita II PDE simulations, a $Q_{<1} = 0.801$ of true density results (Appendix S2: Equation S5). The bias with no alteration (no bypasses, $Q_0 = 0.980$) is found in R simulations of typical $InH^2$ dispersions, analogous to spatial patterns in northern Wisconsin (Appendix S2: Equation S4). The interpolation of the actual bias between none displaced from the 1-m circle around the post and full elimination yields an underestimate of density (near-post bias $= 1 - 0.692 \times [0.980 - 0.801] = 0.875$, Appendix S2: Equation S1). The correction ($\upsilon = 1/0.875 = 1.142$) accounts for surveyor near-post bias in northern Wisconsin under assumed $InH^2$ dispersion (Figure 11; Cogbill et al., 2018).

The total net capture of bypassed trees increases as the distance from the post increases and the near-post correction reaches a peak ($InH^2$ $\nu = 1.148$) at 1.4 m from the post. Any addition to near-post bias beyond 1 m, however, is less influential and partially redundant because the 1.4-m-radius circle incorporates many bypasses that are jointly up to 1.0 m from the line-of-travel (see *Azimuth bias alternative* above). The 1-m near-post range is arbitrary and probably underestimates the total range of influence, but given the crossover between empirical and expected values above 1 m, this value is herein adopted as a conservative standard for near-post bias in Midwest PLS.

## Corner tree bias

A corner tree is the ultimate form of near-post surveyor bias. As explicitly permitted by survey instructions after 1815, a tree can replace a corner post (White, 1983). In this case, the witness tree itself is "squared off" and scribed as a "corner tree" with neither post nor a distance and bearing. A corner tree is a logical consequence of the radius of the tree or root swell being greater than the distance to the corner, so that neither post nor standard witness trees are necessary. Designated corner trees or corners with witness trees that have one at a 0 (not missing) distance comprise 0.268% of all witness trees in Midwest PLS surveys (PalEON database). Because the number and distances of witness trees are confounded by a corner tree, these corners are not considered 2-tree corners.

Corners across the Midwest have a mean basal area of 22.7 $m^2$/ha (Paciorek et al., 2021), or alternatively interpreted as 0.227% of all points (possible corners) in the region falling within the area occupied by the cross section of tree stems. This implies that the corner trees come from only 18% greater area (0.268% observed vs. 0.227% expected) than corners actually falling within the area of a witness tree trunk. This defines an average 21.9-cm-radius circle (1.09 links) around the corner position containing corner trees that have a quadratic mean radius of 17.8 cm. Corner trees are virtually the same size as other witness trees (17.7-cm radius) and were regularly selected from an area slightly beyond where their actual diameter covered the corner point. The replacement by corner trees is consistent with the undersampling of other near-post distances, and the logical explanation is that all witness trees within their own radius (i.e., 5–20 cm, 0.25–1 links) of the corner were "squared off" and converted to posts and the majority of those within 100 cm (5 links) of the post were selectively cleared (or bypassed).

### Near-post bias alternative: Ratio of $W2$:$W1$ (MoR bias)

Additional empirical evidence for spatial bias in historical PLS samples is the ratio of the nearest to second nearest witness tree distances (Cogbill et al., 2018). If the nearest trees are bypassed and replaced by farther trees, the distance to the nearest will increase and the ratio of the nearest witness tree distance ($W1$) to the second tree (unchanged) distance ($W2$) in a pair will typically decrease. The MoR index is sensitive to the forest density, spatial pattern, and close trees being bypassed. The mean of the second tree/nearest tree distance ratios ($\text{MoR} = \overline{W2/W1}$) of a random forest (CSR) sampled without bias is 2.571 in theory (Cogbill et al., 2018) or $\text{MoR} \approx 2.519$ for no bias in R simulations of an $\text{InH}^2$ model at 402 trees/ha (Appendix S2: Equation S4). Similar R modeling indicates a $\text{MoR} \approx 1.909$ if all trees <1 m from the post switched to the next nearest (Appendix S2: Equation S3). In northern Wisconsin, the actual MoR is 2.036, indicating significant bias due to the bypassing of nearest witness trees. Interpolation of the empirical value between no bypasses and all trees bypassed implies the equivalent of 79.2% of the nearest trees (4.99% of all trees) bypassed. This compares to near-post bias (4.37% bypassed) confirming the parallel between MoR and bypassing of trees close to the post. When inserted into the analogous density correction equation as near-post bias, the MoR 4.99% of bypassed trees yields a density correction $\upsilon = 1.169$ for MoR bias (Appendix S2: Equation S2). The additional +0.63% bypassing and +2.33% density correction of the MoR over that of near-post is consistent with MoRs indiscriminate scope including trees beyond 1 m from the post regardless of the type of bias.

The weighted average of the MoRs of distances at all 2-tree $\kappa = 1$ (opposite half) corners in the Midwest is $\text{MoR} = 2.29$. This value independently confirms that the surveyors were biased across the region and were replacing the nearest trees with ones farther away. This index of surveyor bias converts into an average density correction of $\upsilon = 1.075$ over the Midwest (Appendix S2: Equation S2). The analogous near-post bias over the same sample is $\nu = 1.079$. The paired corrections ($\nu$, $\upsilon$) in 233 county/subregions are not statistically different ($t$ test: $t = 0.861$, df = 222, ns; Sokal & Rohlf, 1981) and the residuals indicate that the two indices are equivalent over a broadscale. Therefore, the MoR is a simple check of near-post bias and an effective statistical proxy of surveyor bias near the post without the more complex direct calculation of near-post bias.

## Diameter bias

### Diameter determination

Determination of witness tree diameters contains several subjective uncertainties. PLS surveyors did not measure the diameters of every tree, but presumably estimated the diameters visually (e.g., Almendinger, 1996; Bourdo, 1956; Grimm, 1984; Stewart, 1935; White, 1983; PalEON data). Thus, diameter estimates were regularly rounded to even units of inches, multiples of 10 inches, or feet (e.g., 14, 30, 48 inches) (Dyer, 2001; PalEON data). Reconstruction of PLS surveys, from the late 1800s in conifer forests in the West, even indicate that reported diameters correspond better with basal diameter rather than the dbh (Habeck, 1994; White, 1976; Williams & Baker, 2010). In Midwest PLS surveys, the diameter height is never reported, but without any contrary evidence, the estimation by eye, blazing at arm height, increased visibility, minimal taper, and traditional forestry practice, all argue for a dbh (Bourdo, 1956; Stewart, 1935).

Tree diameters are necessary for both determining the distance from the corner and setting a diameter limit. Regardless of the height at which the diameter is measured, the corner-to-tree distance as measured by the surveyor does not indicate the actual distance to the tree position—the center of the tree. Half the tree's diameter at whatever height the diameter was measured must be added to the surveyor's distance to give the complete post-to-tree-center distance (Anderson et al., 2006; Ashby, 1972; Bouldin, 2008; Bourdo, 1956). Beyond establishing total distances and a reference minimum diameter, diameter is not directly involved in density calculations as density represents individual trees regardless of size. Diameters and their uncertainties, however, are crucial to the calculation of basal area, biomass, or diameter frequency distributions (Bouldin, 2008; Cogbill et al., 2018; Paciorek et al., 2021).

Diameter biases arise when surveyors choose or avoid certain trees as witness trees because of their size. All exploratory examinations and quantitative summaries have shown that small trees are omitted to some degree in PLS surveys (e.g., Bouldin, 2008; Bourdo, 1956; Dyer, 2001; Friedman & Reich, 2005; Grimm, 1981; Hanberry, Yang, et al., 2012; Leahy & Pregitzer, 2003; Liu et al., 2011; Tulowiecki, 2014; Williams & Baker, 2010). Small trees are neither easily marked (i.e., blazed, notched, scribed), particularly prominent, nor considered permanent, but questions remain about just how small was deemed unsuitable. Using different diameter limits, quantitative analyses have shown statistical differences in density and distances for small trees (Bouldin, 2010; Bourdo, 1956; Dyer, 2001; Grimm, 1981; Kronenfeld, 2014; Tulowiecki, 2014; Van Deelen et al., 1996; Williams & Baker, 2010).

## Witness tree diameter frequency

The empirical evidence for PLS diameter bias is the actual frequency of diameters recorded by the surveyors. For example, the distribution of diameters in northern Wisconsin shows a strong unimodal frequency (Figure 12). The diameters at 2-tree and 4-tree corners display very similar distributions with a modal abundance at approximately 25-cm diameter (10 inches). This typical pattern is repeated across the Midwest, with the peak position shifting from 17 cm (7 inches) at corners in northern Minnesota, to 25 cm (10 inches) in northern Wisconsin and northern Michigan, to 30 cm (12 inches) in the witness trees in southern Michigan and Indiana. The frequency of trees smaller than the mode, across all PLS corners, decreases dramatically, exactly the opposite expected in the "reverse-J" distribution of old-growth structure where small trees are most numerous (Oliver & Larson, 1990). After reaching a peak at medium-sized diameters, the diameter frequency decreases quasi-exponentially (reverse-J) to a long tail of very low values beyond 60-cm diameter (Figure 12; Bouldin, 2010; Rhemtulla et al., 2009).

Although the empirical diameter data are relative frequencies (proportions), they can be scaled using congruent patterns of absolute frequencies (densities) in modern, old forests. The PLS empirical distribution was matched to modern landscape distribution in eight well-developed, mixed hardwood forests with little human alteration, presumably structurally similar to historical forests (Appendix S3: Table S1). The modern standard is expressed as an envelope of the minimum and maximum absolute density by diameter classes (Figure 13). This empirical envelope is analogous to the fitting of curves to individual diameter frequencies (Bouldin, 2010; Hanberry & He, 2015; Hanberry et al., 2015; Kronenfeld, 2014; Rhemtulla et al., 2009), but is more general and incorporates the structural variability found across landscapes. The PLS relative diameter distributions were converted to absolute densities by scaling the frequency above 20 cm to equal the density estimate over all corner types (2-tree: 285 trees/ha; 4-tree: 327 trees/ha) for the same set of trees in Wisconsin (Figure 13). The densities above 20-cm diameter generally fit within the modern envelope, but smaller trees were increasingly below the envelope. These results indicate a partial bias against trees smaller than a particular diameter limit, herein termed the "veil-line." The medium-sized trees (20–35 cm diameter), including the modal class nominally at 25 cm (10 inches), were at

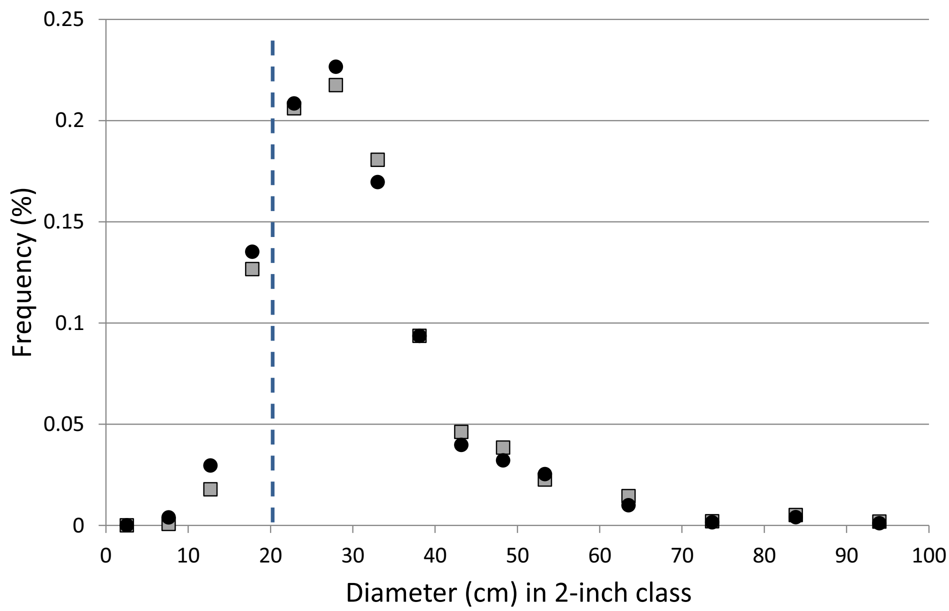


**FIGURE 12** Empirical relative frequency distribution of diameters for witness trees over all corner types in 68 townships in northern Wisconsin public land survey. Diameter in centimeters is plotted at mid-point of 2-inch classes. Data over all corner types at 5633, 2-tree corners and 811, 4-tree corners. Circles are for 11,233 trees at 2-tree corners and shaded squares are for 3196 trees at 4-tree corners. Dashed line is the "veil-line" below which trees were partially censored.

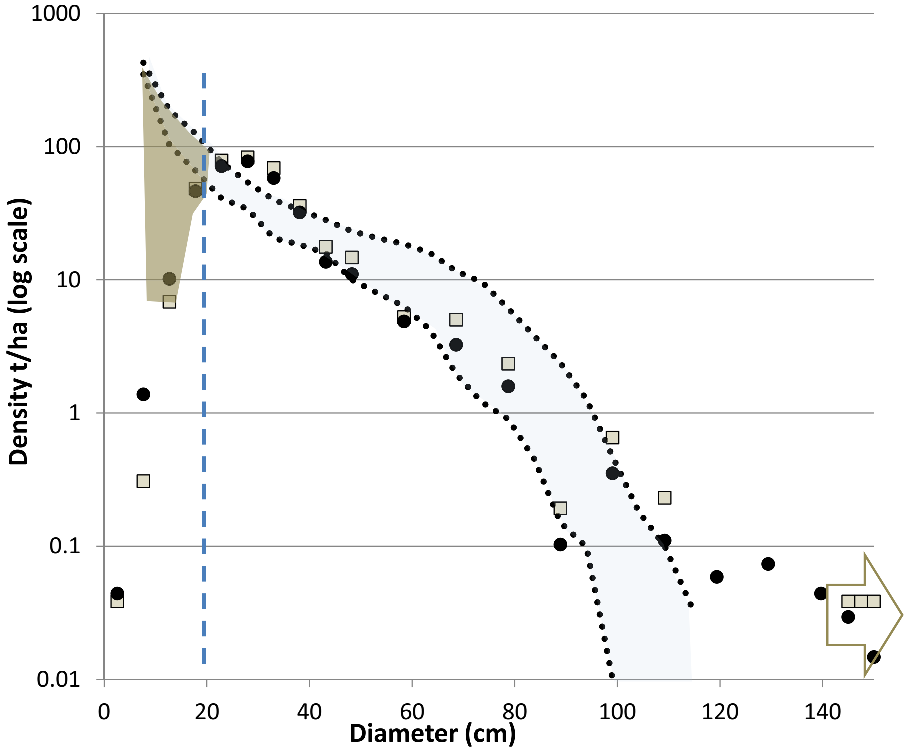


**FIGURE 13** Empirical public land survey diameter frequency of trees over eight northern Wisconsin counties fitted into diameter envelope modeled from modern stands. Same symbols as in Figure 12 for 23,332 trees at interior-quarter, 2-tree corners and 9948 trees at exterior-section, 4-tree corners, plotted on a semi-log scale. Dotted lines are the maximum and minimum values from eight old, well-developed northern hardwood forests (Appendix S3: Table S1) forming the light shaded envelope. The composite envelope encompasses extremes of density (163–345 trees/ha ≥20 cm) and basal area (18.1–51.5 $m^2$/ha). Dashed line is the 20-cm "veil-line" below which trees were partially censored. Dark shading is an area of partially censored or "missing" trees.

or slightly higher densities than the maximum at that diameter in the envelope. Although at low densities, very large trees (>100 cm [40 inches]) were historically more abundant and well beyond the envelope from modern samples. Evidently the landscape represented by the historical witness tree diameter distributions is consistent with a combination of both young regeneration areas with dense small- to medium-sized stems and a mix of old low-density areas containing scattered large trees.

Although the actual veil-line varies somewhat by forest type and surveyor, it is important to base density estimates on a single reference diameter limit through which all PLS and modern samples can be compared. Surveying practice and empirical PLS diameter distributions concur that 20 cm is a reasonable limit below which trees were undersampled (Figure 13; Dyer, 2001; Rhemtulla & Mladenoff, 2010). The 20-cm value is at the upper end of diameters commonly cited as diameter limits (Bourdo, 1956; Manies et al., 2001; Schulte et al., 2007; Zhang et al., 2000). A regional 20-cm value is practical because any lower limit is completely contained within the 20-cm truncation and the generally broad mode (level peak) minimizes the differential from higher values. Also, the veil-line is only slightly below the 25-cm (10 inches) cutoff for canopy trees in modern forestry practice.

## Small tree bias

There are three logical scenarios where small trees might have been included in PLS surveys. First, the local forest may be composed entirely of small trees with no larger trees available as witness trees. Second, even if there are large trees, small trees are usually more abundant than larger trees and tend to grow in relatively dense groups. It might be difficult to ignore (or bypass) multiple small trees or those small, close trees could be obscuring a larger, more distant tree. Third, in a low-density forest or patch, trees larger than the diameter limit might be at too far a distance, and a nearer, albeit smaller, tree might be a necessary substitute. Empirical data from throughout the Midwest indicate the presence of all three scenarios. For example, in northern Wisconsin, 9.1% of the 2-tree corners had both trees <20 cm compared with 4.1% expected based on a random distribution of the proportion of small trees. Surveyors also consistently sampled proportionally more small (<20 cm) witness trees at either short or long distances from the post (Table 4). Analyses of the frequencies of witness trees in distance classes demonstrate a highly significant statistical difference in distances (density) between trees below 20-cm diameter (veil-line) and those ≥20 cm (Table 4). The proportion of trees by diameter class illustrates the differences in sizes with larger trees being more numerous at 2–10 m (2-tree) or 4–10 m (4-tree) and smaller trees being relatively more abundant either near the post (0–2 m) or at far distance (>10 m). The patterns are similar at 2-tree and 4-tree corners, with the size differential most marked among very close trees (0–1 m) at 2-tree corners or among distant trees (>4 m) at 4-tree corners.

At low densities (distances >20 m), the few large (≥20 cm) trees are notably farther (difference: 4.9 m for 2-tree, 29.3 m for 4-tree) from the post than small trees (Table 5). The relative nearness of small (<20 cm) trees at far distances is consistent with their convenient substitution for larger, even more distant trees. This use of small trees at low densities has a significant effect on density calculations as it forms a bimodal frequency of small trees over distance. On average, small witness trees were more frequent in dense patches, or in low-density forests when large trees were unavailable. Removal of any small trees (or their corners) from the calculation will substantially lower the density estimate as the distance to the proper large tree is farther.

## Diameter bias correction (ϕ)

Surveyors generally avoided small trees due to their unsuitability as witness trees, but in certain situations apparently favored them when larger trees were problematic. Small tree bias is not easily defined as the surveyor was both avoiding (bypassing) and favoring (selectively choosing) small trees. Thus, diameter bias is first determining the diameter below which trees were subjectively chosen. The correction (ϕ) is not only a countering of surveyors' bias but technically also a normalization of the density to conform to a diameter limit imposed after the fact by the analyst.

Two methods have been proposed to account for surveyor small diameter bias: truncating the biased frequency distribution below a cutoff (e.g., Dyer & Hutchinson, 2019; Goring et al., 2016; Hanberry, Yang, et al., 2012; Schulte et al., 2007; Zhang et al., 2000); or filling in missing smaller diameter trees by extrapolating from larger diameters based on theoretical curves or modern stand tables (Bouldin, 2010; Bourdo, 1956; Rhemtulla et al., 2009; Williams & Baker, 2011). Regardless of the reason for sampling a tree below a diameter limit, a tree above that limit exists at a greater distance. If the empirical distances are used, the calculated density must necessarily be more than when using the distances to the trees actually larger than the limit. Not having the actual missing distances makes determining the degree of overestimate of density challenging. Rather than trying to calculate the adjusted, unbiased value directly, the original biased calculation can be construed as giving the correct density for that particular mixture of diameters, including some smaller than a standard limit. The proportion of this mixture containing diameters below any standard is then used to eliminate the same proportion of trees from the calculated density.

## Northern Wisconsin diameter correction

In northern Wisconsin, 20.9% of the interior quarter-section corner witness trees were <20 cm in diameter (Table 6). Thus, the calculated density can be

**TABLE 4** Frequency (freq.) of trees broken down by diameter limit of 20 cm (8″) from 3585 2-tree interior quarter-section (intqtr) and 601 4-tree exterior-section (extsec) public land survey corners in nine counties in northern Wisconsin.

| Witness trees | Corner-to-tree distance (m) 0–1 | 1–2 | 2–4 | 4–10 | 10–20 | >20 | Relative freq. | Relative density |
|---|---|---|---|---|---|---|---|---|
| 2-tree intqtr | | | | | | | | |
| No. | 138 | 1135 | 2794 | 2684 | 289 | 130 | 7170 | |
| <20 cm (%) | 30 | 27 | 20 | 18 | 27 | 25 | 20.9 | 24.4 |
| ≥20 cm (%) | 70 | 73 | 80 | 82 | 73 | 75 | 79.1 | 75.6 |
| 4-tree extsec | | | | | | | | |
| No. | 39 | 199 | 635 | 1305 | 192 | 27 | 2397 | |
| <20 cm (%) | 13 | 24 | 22 | 18 | 20 | 26 | 20.5 | 23.8 |
| ≥20 cm (%) | 87 | 76 | 78 | 82 | 80 | 74 | 79.5 | 76.2 |

*Note*: Analyses of the frequencies of sampled trees above and below 20 cm diameter (veil-line) in distance classes demonstrate a statistical difference in distances between small and large trees (chi-square test: $\chi^2 = 74.0$, df = 5, $p < 0.001$ for 2-tree corners; $\chi^2 = 20.4$, df = 5, $p < 0.001$ for 4-tree corners).

**TABLE 5** Mean distance to corner, size, and density of witness trees broken down by diameter limit of 20 cm (8″) from 3585 2-tree interior quarter-section (intqtr) and 601 4-tree exterior-section (extsec) public land survey corners in nine counties in northern Wisconsin.

| | Corner-to-tree distance (m) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Witness trees** | **0–1** | **1–2** | **2–4** | **4–10** | **10–20** | **>20** | **Mean distance (m)** | **Mean dbh (cm)** | **Morisita λ (trees/ha)** |
| 2-tree intqtr | | | | | | | | | |
| All | 0.7 | 1.5 | 2.9 | 5.8 | 13.2 | 39.1 | 5.19 | 27.0 | 326 |
| <20 cm | 0.7 | 1.4 | 2.8 | 5.6 | 13.2 | 35.5 | 4.79 | 15.0 | 381 |
| ≥20 cm | 0.7 | 1.5 | 2.9 | 5.8 | 13.2 | 40.4 | 5.36 | 30.1 | 311 |
| 4-tree extsec | | | | | | | | | |
| All | 0.7 | 1.5 | 2.9 | 6.2 | 12.6 | 47.0 | 5.83 | 28.2 | 393 |
| <20 cm | 0.6 | 1.5 | 2.9 | 6.2 | 13.2 | 25.3 | 5.79 | 15.4 | 455 |
| ≥20 cm | 0.7 | 1.5 | 2.9 | 6.2 | 12.3 | 54.6 | 5.84 | 31.5 | 377 |

"corrected" for the variable diameter limit simply by multiplying the calculated density by the complement of small tree frequency ($\phi = 0.791$). This correction eliminates sampled trees less than the veil-line from the tally and explicitly estimates the density for trees only ≥20 cm dbh. At 4-tree exterior-section corners, the ≥20 cm correction is $\phi = 0.795$.

At northern Wisconsin interior-quarter corners, small trees <20 cm were $\sigma = 20.9\%$ $(1 - \phi)$ of the sample. In contrast, the density of small trees (12.7 cm < diameter < 20 cm dbh) in modern old-growth forests in the Upper Peninsula in Michigan is $\mu = 81.1\%$ of that ≥20 cm implying an expected 64.2% $(\phi \times \mu)$ small trees <20 cm (Eyre & Zillgitt, 1953). These frequencies indicate that only 32.5% $(\sigma/[\phi \times \mu])$ of small trees 12.7 cm ≤ dbh < 20 cm were actually sampled. Additionally, surveyors bypassed 30.2% $([\phi \times \{1 + \mu\} - 1]/[\phi \times \{1 + \mu\}])$ of all trees ≥12.7 cm. Thus, the diameter bias can be viewed as the surveyor avoiding (bypassing) 30.2% of all trees because of their small size and favoring (selecting) 32.5% of the trees below the veil-line despite their small size.

Importantly, the diameter correction $(\phi)$ of the Morisita II density estimate by relative frequency of small trees assumes that the proportion of small trees is independent of density. This is not the case in PLS surveys, as there regularly is a correlation between tree size and density. The strength of the relationship, however, varies by region, corner type, surveyor, and metric (e.g., Bourdo, 1956; Kronenfeld, 2014; Manies et al., 2001; Tulowiecki, 2014; Williams & Baker, 2010). For example, in northern Wisconsin, the Pearson correlation (Sokal & Rohlf, 1981) between witness tree diameter and its distance from the post is $r = 0.048$ ($n = 7170$, $p < 0.05$) for 2-tree and $r = 0.059$ ($n = 2397$, $p < 0.05$) for 4-tree corners. The diameter–distance relationship is positive and statistically significant despite the low variance explained ($r^2 < 0.4\%$) (Table 5). Small trees are generally found at short distances, but their presence at long distances (replacing "too far" large trees) weakens the linear relationship. The more pertinent correlation is between the tree basal area and the base Morisita density at their corner with $r = -0.093$ ($n = 7170$, $p < 0.05$) and $-0.135$ ($n = 2397$, $p < 0.05$) at 2- and 4-tree corners, respectively. The basal area–density relationship is negative and stronger than the diameter–distance correlation, because it incorporates size and integrates data from multiple trees using an average density metric at the corner. Over the Midwest, the mean correlation between total basal area and density at corners is $r = -0.095$ ($p < 0.05$) for 2-tree and $r = -0.152$ ($p < 0.05$) for 4-tree corners. Overall, roughly 1% (2-tree) to 3% (4-tree) of the variance in the density is explained by the variance in tree size. Thus, diameter (or tree size) is not independent of distance and the relationship, although relatively weak, is strengthened with more appropriate metric (density), integrated over the entire corner, or at 4-tree corners with better density estimates.

## Diameter bias alternative: Small tree absolute density

While the proportional frequency method for the elimination of small trees is straightforward, incorporating the covariance of tree size and density gives an alternative estimate of the influence of small trees on density estimation. Weighting each tree by its corner-specific density $(\propto 1/\{\sum_{j=1}^{k} r_j^2/k\})$ produces a size frequency adjusted for the local difference in Morisita density (Table 4; Cogbill et al., 2018). Furthermore, the ratio of the density-weighted trees below the veil-line to all trees approximates the diameter bias within the empirical density estimate attributable only to trees <20 cm (Appendix S3: Section S1). This relative absolute density

**TABLE 6** Estimates of 17 bias types and summary of empirical bias, corrections, and density at 3807 northern Wisconsin interior quarter-section corners.

| Bias type | Symbol | Base λ correction[a] | Relative % bias[b] | Corrected λ (trees/ha) |
|---|---|---|---|---|
| Six basic biases (3579 corners) | | | | |
| Design | κ | 1.000 | 0.0 | 326 |
| Pair angle | θ | 1.083 | 8.3 | |
| Azimuthal | ζ | 1.052 | 5.2 | |
| Near-post | ν | 1.142 | 14.2 | |
| Species | $C$ | 1.000 | 0.0 | |
| Union basic positional[c] | ψ | 1.255 | 25.5 | 409 |
| Diameter ≥20 cm | ϕ | 0.791 | −20.9 | 258 |
| Net composite basic[d] | All | 0.993 | −0.7 | 323 |
| Alternative estimates (3579 corners) | | | | |
| Transect >1 m f/post (azimuth) | ξ− | 1.066 | 6.6 | 347 |
| Ratio second/first (near-post) | MoR | 1.169 | 16.9 | 381 |
| Quadrant (pair angle/design) | SAD | 1.105 | 10.5 | 360 |
| Size–density covar (diameter) | ϕ− | 0.756 | −24.4 | 246 |
| Morisita II InH$^2$ bias (estimator) | PDE | 0.980 | −2.0 | 319 |
| Alternative with transect and covar[e] | κ, θ, ξ−, ν, $C$, ϕ− | 0.976 | −2.4 | 318 |
| Supplementary bias (3579 corners) | | | | |
| Quadrant − pair angle | SAD − θ | 1.022 | 2.2 | |
| Azimuthal >15° | ζ+ | 1.006 | 0.6 | |
| Transect >2 m wide | ξ+ | 1.008 | 0.8 | |
| Transect >3.9 m not in θ | ξ − θ | 1.013 | 1.3 | |
| Near-post >1 m | ν+ | 1.005 | 0.5 | |
| Ratio second/first >1 m | MoR − ν | 1.023 | 2.3 | |
| Tree-based | $C$+ and δ | ~1.000 | ~0.0 | |
| Union additional positional[f] | ψ+ | 1.066 | 6.6 | 437 |
| Size–density covar additional | ϕ+ | 0.965 | −3.5 | |
| Morisita II InH$^2$ bias | PDE | 0.980 | −2.0 | |
| Net supplemental[g] | | 1.008 | 0.8 | 328 |

Abbreviations: covar, covariance; PDE, plotless density estimator; SAD, same-sided:adjacent-across:diagonal.

[a]Correction = 1/bias.

[b]Bias causing error relative to the true density = {correction − 1}; if negative, overestimation; if positive, underestimation.

[c]Equal to the union of all spatial biases adjusted for their random intersection.

[d]Net correction is the union of individual corrections assuming independence and equivalent to the product of the biases.

[e]Alternative composite uses additional methods for azimuth (transect >1 m) and diameter (size–density covariance) biases. Transect and near-post biases are here treated as mutually exclusive together with the four other primary biases.

[f]Union of supplemental positional additions due to transect width beyond azimuth, quadrant beyond SAD, and MoR beyond near-post biases.

[g]Supplemental composite is the union of basic, supplemental additional positional bias, diameter covariance, and estimator bias assuming mutual exclusivity.

also accounts for differential densities across the landscape (Bouldin, 2010). Northern Wisconsin interior-quarter corners have an estimated 24.4% <20 cm dbh witness trees weighted by their corner-specific densities (Table 4). The associated correction (ϕ− = 0.756) for 2-tree diameter bias explicitly accommodates the size–density correlation and includes an additional −3.5% effect on density over the conservative simple proportional frequency (ϕ = 0.791) (Table 6; Appendix S3: Table S2). Similarly, at 4-tree exterior-section corners in northern Wisconsin, the density correction can be adjusted by an average of −3.2% (ϕ− = 0.763) to accommodate the excess density of small trees due to size–density covariance (Table 5).

The primary density corrections ($\phi = 0.791$ at 2-tree corners, $\phi = 0.795$ at 4-tree corners) are based on the simple relative frequency of <20 cm dbh witness trees. Meanwhile, the relative density proportion corrections ($\phi- = 0.753$ 2-tree, $\phi- = 0.762$ 4-tree) adjust for the size–density correlation of small tree frequency. Although tree size and density are not strictly independent, the subjective choice of small trees and the use of a blended density estimate uncorrected for surveyor biases at each corner make the weighted density method supplementary to the proportional frequency method of calculating small tree bias (Tables 4 and 5; Appendix S3: Section S1).

### Fill-in diameter correction

The third method to accommodate small tree bias is to fill in the censored part of the density between the veil-line and the lower diameter limit (dark shading below 20 cm in Figure 13). In practice, expanding the unknown bias against trees below a diameter limit requires curve fitting or an external analogue, but this is not needed when truncating the diameter distribution below that limit. The density of bypassed small trees is estimated by deviation from an expectation based on density between 12.7 and 20 cm dbh equaling 81.1% of that ≥20 cm in nearby Michigan old growth (Eyre & Zillgitt, 1953). This extrapolation is approximate and the increased density is above a new diameter limit (i.e., ≥12.7 cm). Inclusion of the latter bypassed trees in the bias-adjusted base density in northern Wisconsin expands to a density of 586 trees/ha, congruent with the traditionally assumed (≥12.7 cm dbh) diameter limit (Appendix S3: Table S2). This correction shows surveyors favoring some <20 cm trees (20.9% of all witness trees) and the avoidance of other small trees (67.4% of trees 12.7 cm < dbh < 20 cm).

### Large tree bias correction ($\delta$)

A surveyor's preference for medium-sized trees presumes an avoidance of large trees. Indeed, bias against large trees has been proposed (e.g., Bouldin, 2010; Friedman & Reich, 2005; Hushen et al., 1966; Manies et al., 2001; Rhemtulla et al., 2009; Schulte & Mladenoff, 2001; White, 1976) for some of the same reasons as the discrimination against small trees (e.g., difficulty of marking, low survival). This avoidance, however, would require the surveyor to bypass the large tree to use a smaller tree at a farther distance. The low frequency of large trees is actually above that expected in undisturbed landscapes (Figure 13) and consistent with a negative exponential or even more attenuated Weibull function (Bouldin, 2010; Hanberry & He, 2015; Hanberry et al., 2015; Rhemtulla et al., 2009; Zhang et al., 2000). This empirical evidence from Midwest witness tree diameters indicates a putatively unbiased distribution ($\delta = 1.00$) for the portion of the diameter distribution greater than the veil-line (Appendix S3: Section S3; Bouldin, 2010; Bourdo, 1956; Grimm, 1981; Tulowiecki, 2014; Van Deelen et al., 1996; Williams & Baker, 2010). The tendency to bypass small trees ($\phi$) and the natural low abundance of large trees (unbiased $\delta$), not a preference for medium-sized trees, result in the observed unimodal PLS diameter distributions.

## Species bias

Echoing anecdotal testimony of old-time and modern surveyors (Bourdo, 1956; Gordon, 1969; Lutz, 1930), there is a persistent belief that surveyors favored certain species as witness trees (Grimm, 1984; Liu et al., 2011; Manies et al., 2001; Whitney, 1994). Yet, if all chosen trees were the nearest to the corner in their sectors, the sample is by definition unbiased for species composition. If the surveyors were biased toward (or against) a species, they would have to go farther to find the preferred species (or another) (Kenoyer, 1930). Because the surveyor would have to select a different species at a greater distance, the bias would affect both the species composition and the density estimates. Researchers have often looked for species bias using the increased distance, mostly following Bourdo's (1956) method of testing (ANOVA, *t* tests) for a statistical difference among the mean distances to various species (or among subgroups such as diameters or surveyors). Many of these PLS studies have shown no overall statistical differences in distances among species (e.g., Delcourt & Delcourt, 1974; Dorney & Dorney, 1989; Dyer, 2001; Hushen et al., 1966; Kline & Cottam, 1979; Van Deelen et al., 1996). Yet, other studies have found scattered preferences for sundry species (e.g., Almendinger, 1996; Delcourt, 1976; Fralish & McArdle, 2009; Kronenfeld & Wang, 2007; Manies et al., 2001; Siccama, 1971; Williams & Baker, 2010). Predictably, the few significant differences in distance were for species common in either dense (short distances) or open (long distances) forests.

Unfortunately, the Kenoyer/Bourdo differential distance procedure for determining bias is dependent on the implicit assumption that all subsets come from a uniform density, essentially following CSR (Grimm, 1984). In reality, different species are expected to occur naturally in different stands or forest types with different inherent densities. Because the distance parameter in PLS surveys is primarily an indicator of density, any concurrent use as an indicator of bias requires the additional assumption that overall density is homogeneous (stationary). Thus, species bias based on differences in raw distances (Bourdo, 1956; Manies et al., 2001) or densities estimated from raw distances (Kronenfeld & Wang, 2007; Manies et al., 2001) have questionable statistical validity (Grimm, 1984).

## Species bias from distance ratios

The empirical evidence for species bias in historical PLS samples is a recently proposed index of the differences in each species' relative distance at each corner (Dyer & Hutchinson, 2019; Kronenfeld, 2014; Tulowiecki, 2014). By replacing the post-to-tree distance for species $i$ ($d_{ik}$) with its relative value ($\tau$) at each corner $k$ $\left(\tau_{ik} = d_{ik}/\overline{d_{*k}}\right)$, the assumption of regional homogeneity of density is relaxed. The correction for each species' actual composition under species bias based on relative distance is then:

$$p_i' = [p_i/\overline{\nu_i}]/\left[\sum_{i=1}^{s} p_i/\overline{\nu_i}\right], \tag{3}$$

where $p_i$ is the observed (biased) proportion of species $i$, $p_i'$ is a corrected proportion for each species absent bias, $\mathrm{v}_i = d_{ik}^2/\overline{d_{*k}^2}$, is the relative area occupied by tree of species $i$ at corner $k$, and * indicates all distances at corner (method from Kronenfeld, 2014).

Half of the summation of the differences (every preference is paired with an avoidance) between this corrected frequency and the observed frequency becomes an index of surveyor bias over all species:

$$\Delta_\tau = \left[\sum_{i=1}^{s} |p_i - p_i'|/2\right]. \tag{4}$$

This index ($\Delta_\tau$) can be framed as the percentage of trees that have "switched" species identity (Kronenfeld, 2014), or equivalently have been replaced by a farther tree of a different species. If bias is found in an overall ANOVA of relative distances ($\tau$) among species, the degree of species bias is estimated from those species with statistically significant different (post hoc $t$ test for least significant difference {$\mathrm{LSD}_{0.05}$} from $\tau = 1$) relative distances. To maximize the sensitivity of the tests, uninformative single species corners are eliminated from the analysis and only species with more than 50 corners are included in the comparisons. The sum of bypasses for species with LSD-test long relative distances ($\Delta_{\mathrm{sig}}$) yields an approximation of surveyor species preference.

## Species bias correction ($C_\tau$)

The surveyors' overwhelming choice was the nearest tree with a limited number of next nearest trees possible. Fortuitously, the relative distance analysis (Kronenfeld, 2014) estimates the percentage of trees ($\Delta_\tau$), which have been bypassed or "switched" species. Theoretically and also confirmed with R simulations, any density estimated using the assumed nearest tree that is in reality the second nearest will correspond to half the actual density (Thompson, 1956; R code in *Data Availability Statement*). The quantitative underestimate of density due to distance to the second nearest tree results in a theoretical per bypass multiplier factor of 2.00 for the Morisita PDE (Equation 1). The surveyor compositional bias can then be converted into a correction for density without species bias: $C_\tau = 1 + \Delta_\tau$. A larger bias would result if a greater than second nearest (higher rank) tree had been chosen, but that would require multiple trees to be bypassed. The choice of a different species (or any other tree-based bias) would be more likely if the bypassed and favored species (or other conditions) were at nearly the same distance. Then the resultant species bias would be more difficult to detect and the density underestimate would be less than if the analysis were based entirely on species switches. In the literature, the new relative distance tests have found little significant overall species bias ($\Delta_{\mathrm{sig}}$) totaling 0.8%–1.5% significantly favored ($F_\tau$ test, $p < 0.001$): *Quercus* L. spp. (oak), *Pinus strobus* L. (white pine), and *Pinus resinosa* Ait. (red pine) in Minnesota (Kronenfeld, 2014); 2.2% significantly favored ($F_\tau$ test, $p < 0.05$): *Fagus grandifolia* Ehrh. (beech) in western New York (Tulowiecki, 2014); but no statistical difference in Ohio (Dyer & Hutchinson, 2019).

The raw distance differential and the relative distance differential methods for determining species bias give very different results. Analyses of raw densities among species (i.e., ANOVA on $\overline{d_i^2}$; Bourdo, 1956) in Midwest PLS surveys show significant (most $p < 0.001$) differences among species in 78 of 86 (91%) subregion/tree number units (Appendix S4: Table S1). In contrast, the relative distance method (i.e., ANOVA on $\overline{\tau_i}$; Kronenfeld, 2014) yields limited statistical significance (20 of 86 units {23%} $p < 0.05$; Appendix S4: Table S1). These results confirm that density variability is at the root of most species' distance differences. The traditional Kenoyer/Bourdo method (differences in raw distance) for determining species bias is misleading and most of the species differences found in previous studies are nullified. In the northern Wisconsin case, the recent Kronenfeld (relative distance) test indicates absolutely no significance in relative distances among species (total $\Delta_\tau = 1.4\%–2.0\%$; $F_{\tau(14,3000)} = 1.30–1.56$, $p > 0.20$) and infers a density correction for species bias of $C_\tau \approx 1.00$.

## Overlap of species and tree size biases

Species bias is intimately connected to small tree bias. In contrast to long relative distances from the post indicating surveyor preference, short distances do not indicate species avoidance. When a species (or other tree condition such as size) is avoided, its data are unrecorded; the species identity (or condition) and its distance are, by definition, censored. Neither distance nor relative distance tests can identify those trees avoided because of their species, size, or condition. The prominence of species with significantly short relative distances ($\tau \ll 1$; e.g., *Ostrya* Scop. or *Carpinus* L. (ironwood), *Populus* cf. *tremuloides* Michx. (aspen), *Betula* L. spp. (birch), *Carya* Nutt. spp. (hickory), *Abies balsamea* (L.) Mill. (fir), various shrubs; PalEON data; Kronenfeld, 2014) appears to be a function of the tree size and density rather than surveyors' avoidance of the species. Therefore, significant short relative distances are predominantly related to small tree bias. In turn, trees with long relative distances are dependent on co-occurrence with small trees at short relative distances. For example, if the 20 subregion units with significant relative distance ANOVA tests are analyzed by removing corners containing a small tree (<20-cm diameter), only four (5% of 86 subregions) retain statistically significant differences in relative distances among species (Appendix S4: Table S1). Even these units with statistically significant species bias have limited estimated gross bias ($\Delta_\tau = 2.4$) and reduced significance (long distances) for ostensibly preferred species ($\Delta_{sig} = 1.2\%$).

## Species bias in the Midwest

Given the obvious surveyor biases in survey design, tree size, and the position of witness trees, there is remarkably little evidence for additional bias for species. Over the Midwest domain, the gross species relative distance differential averages $\Delta_\tau = 2.2 \pm 1.2\%$ SD, but the weighted average of statistically significant relative long distances indicates a slight density underestimate of $\Delta_{sig} = 0.54\%$ (Appendix S4: Table S1; Table 7). Thus, the relative distance test for species preference indicates an average density correction of $C_\tau = 1.0054$, while the residual preference after removal of small tree bias is minimal ($\Delta_{sig}$ = maximum 1.7% over 4 units) with $C_\tau = 1.00058$. Moreover, the relative distance test at subregion scale lacks statistical power as it is based on small sample size for the number of species and is insensitive to low levels of bias (Appendix S4: Table S1; Kronenfeld, 2014). Contrary to persistent beliefs, species bias in the Midwest surveys is rarely detectable and overall negligible (<0.1%; Appendix S4: Table S1).

## Species bias alternative: Line trees

Species (and diameter) biased choice of witness trees have often been estimated by assuming the "line

**TABLE 7** Composite weighted average by number of corners for bias, density corrections, and density estimates from Midwest public land surveys aggregated over all corner types into state-scale regions.

| Region | No. corners | Base $\lambda_M$ (trees/ha) | Pair angle, $\theta$ | Azimuth, $\zeta$ | Near-post, $\nu$ | Union bias, $\psi$ | Bypass (%) | Species, $C$[a] | Corrected Morisita, $\lambda$ | >20 cm, $\phi$ | Diameter corrected, $\lambda$ > 20 cm |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ohio | 2719 | 200 | 1.08 | 1.01 | 1.06 | 1.15 | 10 | 1.000 | 232 | 0.88 | 203 |
| Indiana | 56,363 | 254 | 1.09 | 1.02 | 1.06 | 1.17 | 14 | 1.008 | 298 | 0.84 | 246 |
| Illinois | 28,215 | 122 | 1.08 | 1.06 | 1.06 | 1.21 | 16 | 1.000 | 144 | 0.84 | 117 |
| S Michigan | 29,532 | 219 | 1.09 | 1.06 | 1.06 | 1.19 | 13 | 1.004 | 265 | 0.85 | 225 |
| N Michigan | 34,310 | 305 | 1.19 | 1.18 | 1.12 | 1.39 | 27 | 1.007 | 424 | 0.79 | 327 |
| S Wisconsin | 15,826 | 147 | 1.07 | 1.08 | 1.08 | 1.20 | 11 | 1.005 | 181 | 0.85 | 151 |
| N Wisconsin | 28,887 | 216 | 1.10 | 1.08 | 1.10 | 1.25 | 17 | 1.003 | 273 | 0.85 | 230 |
| S Minnesota[b] | 6014 | 52 | 1.04 | 1.03 | 1.10 | 1.19 | 9 | 1.000 | 61 | 0.80 | 49 |
| N Minnesota | 19,898 | 287 | 1.11 | 1.09 | 1.11 | 1.27 | 20 | 1.012 | 364 | 0.56 | 201 |
| 2-tree | 221,764 | 225 | 1.10 | 1.08 | 1.08 | 1.23 | 17 | 1.005 | 280 | 0.81 | 221 |
| N Wisconsin | 4791 | 254 | NA | 1.20 | 1.09 | 1.28 | 19 | 1.002 | 325 | 0.89 | 289 |
| N Minnesota | 7257 | 278 | NA | 1.23 | 1.11 | 1.32 | 22 | 1.011 | 371 | 0.57 | 212 |
| 4-tree | 12,048 | 268 | NA | 1.22 | 1.10 | 1.30 | 21 | 1.008 | 350 | 0.74 | 239 |
| Midwest | 233,812 | 227 | NA | 1.08 | 1.08 | 1.24 | 17 | 1.005 | 284 | 0.80 | 222 |

*Note*: Weighted mean values do not necessarily align due to unequal representation and covariance among parameters. The design correction ($\kappa = 1$ or 2) is incorporated in the base density as appropriate for the corner type. Pair angle bias is not applicable at 4-tree corners, so their union bias is for azimuthal and near-post only.
Abbreviations: N, north; S, South.
[a]Weighted average of significant species at long relative distance ($\Delta_{sig}$) prorated over all corners (Appendix S4: Table S1), if small trees (<20 cm) are removed, generally $C \approx 1.000$.
[b]Raw bearing data are incomplete, so angle and azimuthal biases and bypass estimates are approximate.

trees" cited along the survey lines between corners are an unbiased sample (e.g., Almendinger, 1996; Bjorkman & Vellend, 2010; Hanberry, Yang, et al., 2012; Janke et al., 1978). Although most line trees were blazed, on many survey lines the surveyors tallied only one or two line trees in each half mile (Bourdo, 1956; White, 1983; PalEON data). Unfortunately, in the Midwest, only a small percentage of trees intersecting the section line were recorded in field notes (5%–32%, Appendix S4: Section S1: Line tree bias). Thus, the cited line trees are not an objective sample of trees randomly "chosen" by the survey line. These records contain substantial surveyor choice compromising any calculation of species composition, tree density, or size frequency distribution from them (cf. Almendinger, 1996; Grimm, 1984; Hanberry, 2021; Kronenfeld, 2014; Liu et al., 2011).

## Union spatial bias (ψ correction)

Biases are initially derived separately; the reciprocal (1/bias) of each is expressed as an individual correction factor ($\kappa$, $\theta$, $\zeta$, $\nu$, $\phi$, $C$). The basic design factor corrected ($\kappa$) for mismatch with PDE defines a fixed sampling geometry from which the other spatial biases deviate. The three spatial biases (pair angle, azimuth, near-post) are not mutually exclusive as the same tree can be within 1 m of the post, within the azimuth exclusion angle, and/or outside the pair sample design. The composite sectors for the near-post and azimuth biases include a central circle 2 m in diameter, wedges within 15° of the line-of-travel beyond 1 m, and possibly the perpendicular section lines (Celtic Cross bias). In addition, pair angle bias adds to the positional bias when witness trees are bypassed in the equivalent of a narrow pie-shaped sector. Despite the unequal domains that decrease the statistical intersections and the nonrandom interactions that increase the intersections, the spatial corrections are treated as independent variables. The individual spatial biases are combined in a single additive "union spatial bias" of multiple factors with random intersection. In this case, the resultant net spatial bias correction ($\psi$) is approximated as the union of the probabilities:

$$\psi = 1 + (\theta - 1) + (\zeta - 1) + (\nu - 1) - (\theta - 1) \times (\zeta - 1) - (\theta - 1) \times (\nu - 1) - (\zeta - 1) \times (\nu - 1) + (\theta - 1) \times (\zeta - 1) \times (\nu - 1). \tag{5}$$

The final two corrections ($\phi$, $C$) are from bypassing due to tree-based nonspatial biases. The species bias is calculated from the proportion of bypassed trees for favored species, or factors covariant with species, such as size or ease of blazing. Any large tree bias ($\delta$) or species bias ($C$) potentially contributes additional corrections and combine with the positional union, but in northern Wisconsin, both are negligible (=1.000). Diameter bias ($\phi$) also incorporates species bias when particular species are naturally small (Hanberry, Yang, et al., 2012; Kronenfeld, 2014; Tulowiecki, 2014). Diameter bias is a combination of a traditional surveyor bias against trees below the veil-line size and a post hoc cancellation of that bias by the analyst. The normalization eliminates partial bias for small trees and is unique because it adjusts density for an overestimate ($\phi < 1$) after the fact.

The conversion of spatially bypassed trees to a correction factor is not strictly proportional (Table 7). The mean corrections are equivalent to a weighted harmonic mean of the bias, measured as the proportion of bypassed trees. The harmonic mean is necessarily less than the arithmetic mean of bypasses as it conveniently downweights outlier high bypass values. Moreover, the trees bypassed because of near-post (or MoR) bias are at short distances from the post and have disproportionately high density. Therefore, the near-post density correction for the same proportion of bypassed trees is larger than the azimuth or pair angle biases.

The composite of the six enumerated biases ($\kappa$, $\theta$, $\zeta$, $\nu$, $C$, $\phi$) approximates the total surveyor bias in PLS surveys (e.g., Goring et al., 2016; Kronenfeld & Wang, 2007). Importantly, there is a deterministic interaction between species and diameter biases (see above), which is subsumed in the diameter bias. Moreover, the diameter and net positional biases have different bases and countervailing signs, so they are treated as mutually exclusive variables. The union of spatial, species, and diameter biases potentially deducts the intersections (overlap) among biases from their sum, but the positional biases ($\theta$, $\zeta$, $\nu$) already include their intersection in their sum ($\psi$). Therefore, the product of the design, spatial, species, and diameter biases produces a composite correction factor for the base density (unbiased ≥20 cm, $\lambda = \lambda_{m\ \Pi} \times \kappa \times \psi \times C \times \phi$).

# RESULTS

## Surveyor empirical protocols

Analyses of distance and bearing parameters elucidate the Midwest PLS survey methodology regardless of the instructions, historical compilations, or surveyor tradition (e.g., Bourdo, 1956; Paciorek et al., 2021; Stewart, 1935; White, 1983). Laying the survey chain involved significant marking and clearing of trees along the section line. The corner vicinity was cleared further for unmistakable monumenting of the corner. At each section or quarter-section corner, trees in opposing semicircles were marked (blazed, notched, and scribed) by the axemen (White, 1983). These witness trees accurately perpetuated the corner and provided on-site identification of the section in which they resided. If a witness tree was close to overlapping the corner, it was converted to a corner tree, replacing the post. The trees used to witness the corner were not necessarily the nearest to the corner because surveyors restricted the sectors containing witness trees. Interestingly, the perpendicular quarter-section lines, and to a minor degree the yet to be surveyed section line ahead of the corner, had less surveyor bias than the corner itself or the previously surveyed section line.

This procedure appears to have been implemented from the earliest PLS surveys in Ohio as two witness trees were consistently recorded, although a surveyor rarely noted "no tree in opposite direction for bearing tree" (Ohio T7 R9 Ohio River Base 1797 outlines, NARA RG 49). An 1833 clarification stipulated that bearing trees be in different sections, "on each side of the section line." In addition, instructions called for exterior corner trees to be inside the surveyed township on the same side of the line. These 2-tree protocols remained unchanged through the 1850s in Illinois, Indiana, Michigan, and Ohio. In 1846, the instructions for Wisconsin and Iowa (including what is today Minnesota) changed to include four witness trees at township and section corners and two trees in opposite sections at quarter-section corners (White, 1983). The section corners had larger posts, more information inscribed on witness trees, and after about 1846 more witness trees at each corner than at quarter-section corners (White, 1983). These patterns represent the survey design and every Deputy Surveyor, and often the chainmen, axemen, and flagmen solemnly swore and certified that they had faithfully and strictly followed the instructions of the Surveyor General (e.g., affidavits in field notes: 1806 T10S R1E, Illinois; 1825 T1S R1E, Michigan; 1858 T41N R3W, Wisconsin; Grimm, 1981; White, 1983). The surveyor's oath is strong motivation, but given personal idiosyncrasies and the minimal quality control, there is no a priori assurance as to how closely the ideal design was followed.

Significantly the actual PLS sampling designs did not follow previous researchers' assumptions of either trees in the two nearest quarters (2nQ) (e.g., Anderson & Anderson, 1975; Bouldin, 2008; Delcourt, 1976; Goring et al., 2016; Hanberry et al., 2011; Manies et al., 2001; Ward, 1956) or of sampling based on quadrants (e.g., Bouldin, 2008; Bourdo, 1956; Hanberry, Yang, et al., 2012; Kronenfeld & Wang, 2007; Manies et al., 2001). Previous references to "quadrants" at 2-tree corners are misconstrued as they are conflated with the surveyors' quarter-bearing notation (e.g., cited as S53° E) or allusions to the point-centered-quarter sampling of four trees (Cottam et al., 1953; Cottam & Curtis, 1956). Surveyors generally followed three sampling templates (2-tree opposite: 2oH; 2-tree same side: 2sH; 4-tree sections: PCQ) depending on date and corner type (exterior vs. interior and section vs. quarter-section). The earliest surveys, generally before 1837, invariably used two witness trees per corner. Exterior corners initially had both trees inside the township. At quarter-section corners, witness tree orientation was either in two opposing sections (82.6% adjacent-across plus diagonal corners) or in opposing halves regardless of sections (15.0% same-sided corners). All 2-tree empirical designs are concordant with equal halves sampling, with employing the unaltered Morisita II estimator, and with modification for bias. Essentially this is a single design fulfilling the ca. 1805 instruction of "two … trees in opposite direction as nearly as may be" (White, 1983) and assured the clearest witnessing of the corner. Four-tree corners were rare, except in Wisconsin and Minnesota after 1846 where the four trees defined four different sections.

Importantly, witness trees not only served to allow relocation of the corner, but also lay unambiguously within, and identified, different sections on the ground. Witness trees were generally ≥20 cm in diameter and thus easily scribed, but those less than 20 cm were utilized when small trees were difficult to ignore or when a larger tree was at too great a distance. At 4-tree corners in later surveys, surveyors showed greater bias near the line-of-travel and cleared (or avoided) more trees immediately near the post (Table 7). In addition, there were fewer small witness trees near the post at 4-tree corners than at 2-tree corners, indicating the consistent use of larger trees, with more scribing, at the prominent section (or township) corners. These procedures all validate that the surveyors did not always use the nearest witness tree, instead creating design, azimuthal, near-post, pair angle, and diameter biases.

## Northern Wisconsin case study

The focus on northern Wisconsin (*Data Availability Statement*: Northern Wisconsin Case Study) allows a comprehensive analysis of surveyor bias in PLS surveys. In the 68 northern Wisconsin townships detailed above, multiple biases affecting density estimates are combined for the distances, directions, and diameters from 3604 interior quarter-corner pairs of witness trees. The analyses indicate surveyors bypassed a gross 16.4% of the nearest trees (7.7% due to pair angle design, 4.3% azimuth avoidance, and 4.4% for near-post bias) for spatial bias.

The combination of 15.6% of trees (incorporating random intersection) bypassed due to position and the mutually exclusive 30.2% due to small diameters (see *Diameter bias*) totals 45.8% bypassed trees. Stated another way, 46% of the nearest trees >12.7-cm diameter did not become witness trees. Estimates of specific corrections for pair angle bias ($\theta = 1.08$), azimuthal bias ($\zeta = 1.05$), and near-post bias ($\nu = 1.14$) result in a union surveyor correction ($\psi = 1.25$). Both intended sampling design and species bypassing reflect negligible surveyor bias as the sampling design was opposite halves ($\kappa \approx 1.00$) and species bias was not statistically significant ($C \approx 1.00$). The composite correction for base density at interior quarter-section corners in northern Wisconsin is $\kappa \times \psi \times C = 1.255$ (+25%) due to the surveyors bypassing a significant proportion of the nearest trees to the post.

Using northern Wisconsin interior quarter-section corners, the base density calculated from the witness tree distances (observed plus radius) using the Morisita II PDE estimator is 325.8 trees/ha (Cogbill et al., 2018; Morisita, 1957). This estimate corrected for six surveyor biases is an unbiased 409 trees/ha as occurred on the ground (Table 6). The unbiased density estimate corrected for spatial bias is 7%–9% less than previous literature estimates from the same area (i.e., Manies et al., 2001: 441 trees/ha; Hanberry & Dey, 2019: 447 trees/ha). This density, however, has an indefinite diameter limit and can be normalized by the removal of small trees ($1 - \phi = 0.209$) together with their bias. The adjusted density estimate of 323.4 trees/ha ≥20 cm is coincidentally nearly the same as the raw base density (325.8 trees/ha). The underestimates for spatial biases and the overestimates for density normalization offset each other with a net bias of −0.73% (Table 6). The diameter-specific density estimate is 26%–28% less than literature values. Moreover, corrected for the putative diameter limit of previous studies (12.7 cm), the estimated density (585.7 trees/ha ≥12.7 cm) is 31%–33% greater than literature values. All four density estimates (326 trees/ha base; 409 trees/ha unbiased; 323 trees/ha ≥20 cm; and 586 trees/ha ≥12.7 cm) are valid in their own context. The robust PDE, expanded consideration of bias, explicit diameter limit, and lack of ancillary assumptions assure that the ultimate unbiased estimate of 323 trees/ha ≥20 cm for presettlement northern Wisconsin interior-quarter corners is comprehensive and minimally subjective.

## Supplemental biases in northern Wisconsin

The six corrections detailed above are conservative estimates that address the primary, widespread, and well-documented biases included in Midwest PLS density estimates. Although requiring additional assumptions, simulations, and estimation of the relationship between tree size and density, the alternative calculations for tangent section line ($\xi$) and density accommodating size–density covariance ($\phi$ plus $\phi+$) can be substituted for azimuthal ($\zeta$) and density proportion ($\phi$), respectively. The alternative net correction, albeit less assured, is 0.976 (1.7% less than the primary correction) of the base density, resulting in an unbiased density estimate of 318 trees/ha ≥20 cm (Table 6).

Various secondary biases also produce additional underestimates (Table 6). Three spatial ($\zeta$, $\nu$, $\theta$) biases use subjective boundaries to approximate the extent of the bias. The 15° azimuth angle or transect 1 m from the section line (Figure 6), within 1 m of the post (Figure 11), and deviating from the opposite halves (2oH) pair angle line (Figure 5) are imposed to standardize the estimate of bias. They all, however, have potential supplementary positional corrections beyond the boundaries. The supplemental corrections ($\psi+ = +6.6\%$) are partially balanced by the independent diameter-covariance supplemental bias ($\phi+ = -3.5\%$) (Table 6). There probably are statistical intersections among bypasses, which make the net union of the positional and small tree normalization (≈+2.9%) a maximum adjustment. Additionally, error in the Morisita II estimator bias for $\text{lnH}^2$ dispersion (up to −2.0%) results in a potential net supplementary correction of ≈+0.8% yielding 328 trees/ha ≥20 cm.

The corrected estimate for basic biases (323 trees/ha), calculated with alternative techniques (318 trees/ha), and including supplementary biases (328 trees/ha), confirm a robust composite density estimate. Given the unknowns, assumptions, and minimal effect, the adjunct biases are best viewed as providing an uncertainty of less than ±2% in the density corrected for the six primary biases.

In northern Wisconsin, 8.6% of all interior quarter-section corners do not have two opposite witness trees. If the nonstandard designs are reduced to a nearest tree or a proxy distance (i.e., 100 m) substituted for trees "too far," there is an additional latent correction ($\eta \approx -2.4\%$) for the missing trees in generally sparse densities (Appendix S5: Table S1). Due to the missing distances, uncertain design, and inexact methodology, corners other than 2-tree are excluded from the unbiased diameter-specific density estimate. Nevertheless, if all primary, supplementary, and noncompliant biases are assumed mutually exclusive and are combined with nontreed corners ($\lambda = 0.00$), the resultant 306 trees/ha ≥20 cm forms a potential complete landscape estimate over northern Wisconsin (Appendix S5: Table S1).

## Surveyor bias across the Midwest

This study expands the northern Wisconsin case study by evaluating and calculating bias in Midwest PLS surveys covering 80 representative subregional units (Appendix S6: Table S1). These regional analyses demonstrate a remarkably widespread and consistent occurrence of surveyor bias in PLS surveys (Table 7). Reiterating findings in northern Wisconsin, 2-tree corners over the Midwest have a gross weighted average of 17.7% (mean 10.2% pair angle, 5.1% azimuthal, 2.4% near-post) of the nearest trees bypassed. Regionally, pair angle bias was most important in early surveys in the lower Midwest, while near-post bias was notable in later surveys from northern states (Table 7). The 4-tree corners do not support any pair angle bias but have a gross total of 20.5% bypassed trees (mean 18.1% azimuthal, 2.5% near-post). The 2-tree corners (hourglass bias) have less than a third of azimuthal bias of the 4-tree corners (Celtic Cross bias). The grand total biases with the addition of pair angle bias, however, yield roughly comparable bypassing (2-tree: 17.7%, 4-tree: 20.5%) at both corner types (Table 7). Extrapolation of the partial bias for trees <20 cm indicates that the surveyors included only 30.2% of the trees between 12.7 and 20 cm as witness trees and bypassed 31.2% of all the trees ≥12.7 cm because of their small size. Combining the average 17.0% of bypassed trees (with random intersection) at all corners because of positional bias and the 31.2% of bypassed trees because of their size means that a total of 48.2% of trees >12.7 cm nearest the post in the Midwest were bypassed (Table 7).

Over all corner types in the Midwest, the incorporation of disproportional weight of bypasses and reductions for random intersection (overlap) of spatial biases yields a weighted average union correction of $\psi = 1.238$ (Table 7). In addition to the spatial biases, the scattered putative bias for favored species ($\Delta_{sig}$) over the Midwest after removal of small tree bias is a weighted average of $C = 1.00058$ (see *Species bias*). In total, the explicit empirical estimate of surveyor bias in the Midwest results in a grand average underestimate (80.7%) of the base density, including restricted survey design, avoiding cardinal directions, clearance near the line-of-travel and post, and preference for species (correction $\psi \times C \times \delta = 1.239$, Table 7). The size bias correction may either increase or decrease the unbiased density estimate depending on the diameter limit set by the analyst. If adjusted to the commonly cited ≥12.7 cm, the grand correction to the Morisita base density over the Midwest is a gross 1.80 multiplier ($\psi \times C \times \delta \times \phi \times 1.811$) of the base density. Normalized to the ≥20-cm standard, however, the unbiased composite correction is virtually equal to the base density ($\psi \times C \times \delta \times \phi = 0.995$).

In the Midwest, the grand weighted average Morisita base density estimate is 227 trees/ha (Table 7). The average union positional correction increases the estimate to an unbiased 284 trees/ha. The removal of diameter bias ($\phi = 0.803$) results in an average of 222 trees/ha ≥20 cm, an accurate density corrected for spatial, tree-based, and diameter biases over all corner types averaged over the entire Midwest. If expanded to a 12.7-cm-diameter limit, the estimate would rise to 411 trees/ha ≥12.7 cm. Even when the base density estimate is numerically similar to the corrected density estimate, the corrected value is more rigorous because it incorporates surveyor bias and has an explicit diameter limit. As in northern Wisconsin, the union of positional biases involving the sampling of non-nearest trees that give underestimates ($\theta$, $\zeta$, $\nu$, $C$; all ≥1) is offset by the diameter normalization for trees ≥20 cm that give overestimates ($\phi \leq 1$). The balance between spatial underestimates and diameter overestimates varies by region, corner type, and surveyor, but ranges from a net correction of +26 trees/ha (+8%) in northern Michigan to −81 trees/ha (−28%) in northern Minnesota.

The quantity of each type of bias varies depending on corner type, surveyor, underlying density, and ecosystem type. At 2-tree corners in Midwest PLS surveys, the documented corrections for spatial bias are spread roughly equally over pair angle (domain mean $\theta = 1.10$), azimuthal ($\zeta = 1.08$), and near-post ($\nu = 1.08$) biases (Table 7). Significantly, at 2-tree corners, azimuthal correction at section corners ($\zeta = 1.11$) is greater than at quarter corners ($\zeta = 1.06$) and pair angle correction at interior corners ($\theta = 1.10$) is greater than exterior corners ($\theta = 1.05$). The composite union of surveyor biases in subregions (county-scale or state divisions) ranges from $\psi = 1.04$ (Iowa-Lafayette Cos., Wisconsin) to $\psi = 1.56$ (Alpena County, Michigan). The variability of average subregional corrections ranges from pair angle ($\theta = 1.07$ to 1.11), azimuthal ($\zeta = 1.01$ to 1.08), near-post ($\nu = 1.06$ to 1.12), to a union composite ($\psi = 1.14$ to 1.26+). Northern Michigan is a singularly extreme region with an elevated total for all spatial biases and an outlying composite ($\psi = 1.39$). This result is due in part to the surveyors' unusual inclusion of azimuthal bias in all four cardinal directions (Maltese Cross) at 2-tree corners (e.g., Figure 8), and the exacting spatial censoring. Small tree normalization is highly variable, ranging from $\phi = 0.28$ (Lake of the Woods, County, Minnesota) to $\phi = 0.95$ (Taylor-Marathon Counties, Wisconsin). In northern Minnesota, the small size of trees appears to be an outlier of forest structure rather than any surveyor bias.

## Comparison with previous studies

The determination of density in previous PLS publications has been profoundly affected by a mixture of assumed sampling designs, estimators used, surveyor biases considered, and corner types used (Table 1). In practice, the analyst errors in PLS density determinations are largely caused by mistaken assumptions about sampling design and using the Cottam (Cottam & Curtis, 1956) or Pollard (1971) density estimators that yield a relative root mean square error (RRMSE) of 76%, $n = 41$ (Appendix S7: Table S1; Cogbill et al., 2018). In contrast, recent use of the Morisita estimator, in various forms and sometimes aggregated over a small number of corners, dramatically reduces this controllable analytical error to 7.7% RRMSE, $n = 10$ (Appendix S7: Table S1; Cogbill et al., 2018; Levine et al., 2017; Morisita, 1957).

### Analyst bias

Decisions made by modern analysts influence calculations of forest density for PLS surveys even before any historical surveyor bias can be considered. Misestimates in early studies, commonly accentuated by the Ward assumption, are ubiquitous (Appendix S7: Table S1). For example, the use of the single corner Cottam (i.e., Shanks) PDE estimator with the Ward (2nQ) assumption by Manies et al. (2001) found a weighted mean density of 467 trees/ha in the northern Wisconsin case study area. Meanwhile, the present study using the Morisita estimator for the same area found 300 trees/ha raw base density (Appendix S6: Table S1), indicating an overall 56% overestimate in the Manies study (Cogbill et al., 2018).

Even if the same PDE is used, the estimate of density can vary due to analysts' different assumptions about the application of that estimator. Both Hanberry (2020) and the present study applied the Morisita (1957) estimator over the northern two thirds of Michigan PLS but used different methods of calculation with different results. Hanberry found a base density of 252 trees/ha while the present study finds 305 trees/ha from the same database (Table 7). Before any correction for surveyor bias, Hanberry produced a 17% underestimate of base density. This is apparently due to the mismatch of the known design and the Morisita estimator: applying the Ward assumption (2nQ) to opposite half (2oH) design; reduction of 4-tree corners to three nearest quadrants; and not accommodating same-side (2sH) exterior corners (Hanberry, 2020; Hanberry, Yang, et al., 2012).

Multiple errors under the control of the analyst compromise density calculations. Overall, there are four prominent sources of analyst bias that appear in most previous studies that calculate density from PLS data. Analysts have used biased estimators, applied them to unknown or unrealistic sampling designs, omitted adding radius to the distance, and consistently failed to cite or impose a minimum diameter to account for the surveyors avoiding inappropriate small trees (Appendix S7: Table S1). When all errors and assumptions are combined, estimator and quadrants used (76%), design ($\kappa = +5\%$), added radius ($\rho = -14\%$), and diameter limit ($\phi = -19\%$ to $+52\%$) create a large uncertainty in these density estimates. Thus regardless of any additional surveyor bias, virtually all past PLS density values are unreliable as a baseline for comparisons with past or modern values unless corrected.

### Surveyor bias

The most extensive assessment and correction for surveyor bias in PLS density estimation is the methodology of Hanberry, Yang, et al. (2012) applied in 11 publications covering Missouri, Minnesota, Wisconsin, northern Michigan, and small areas of Mississippi, Oregon, Washington, and South Dakota (Hanberry, 2020; Hanberry & Dey, 2019; Hanberry & He, 2015; Hanberry et al., 2014, 2015, 2016, 2018; Hanberry, Kabrick, et al., 2012; Hanberry, Palik, et al., 2012; Hanberry, Justice, et al., 2020; Tatina & Hanberry, 2022). This methodology combines a rank-based approach (essentially assuming a modeled average rank of witness tree distance) and a bias-based approach (using quadrants occupied, Maltese Cross azimuthal bias, and empirical "line" tree composition and size). The adopted rank-based assumption of 1.8 average rank order alone would result in a theoretical correction equivalent to 80% second ranked ($\Delta_2$) trees or $C_2 = 1.80$ correction (Hanberry, Yang, et al., 2012). The total corrections for the identical surveyor bias-based assumptions are cited as 2.33 in Missouri and 1.66 in northern Michigan (Table 1; Hanberry, 2020; Hanberry, Yang, et al., 2012). The majority of the Hanberry bias-based method is based on deviation from species and diameters of line trees and from a 2:1 adjacent:diagonal ratio (2nQ design), both of which have questionable validity (see above; Appendix S7: Table S1).

Correcting for the same biases ($\kappa$, $\zeta$, $C$) over the same exact northern Michigan surveys as the Hanberry (2020) study, the present study found a composite 1.19 correction for surveyor bias (Table 7, PalEON data). The Morisita base density, adjusted for these surveyor biases, yields an unbiased density of 359 trees/ha. The Hanberry estimated surveyor bias is 3.5 times greater and the bias-corrected density estimate is 17% greater than in the present study. Similarly, the identical Hanberry methodology yielded a 15%–39% overestimate of the bias-corrected density in four ecoregions in Minnesota and Wisconsin (Hanberry & Dey, 2019; Hanberry, Palik, et al., 2012; Table 7). Combining the density and bias estimates, Hanberry (2020) projected an unbiased density of 420 trees/ha ≥12.7 cm in northern Michigan. The present study using the same exact PLS raw data (Table 7) produces a density estimate of 616 trees/ha ≥12.7 cm, indicating an overall 32% underestimate in the Hanberry study.

## CONCLUSIONS

This paper heeds Bourdo's (1956) warning to investigate the characteristics of PLS sampling to determine the possibility and degree of any bias before estimating forest density. Biases inherent in both the estimator equations and the underlying forest pattern can be minimized by using an appropriate PDE. All four parameters (distance, bearing, species, diameter) recorded by the PLS surveyors can reveal potentially biased choices defined simply as not sampling the nearest trees to the corner post. Because the surveyors' precise sampling design and the exact trees chosen are preserved in the recorded empirical data, any assumption of a sampling design is unnecessary.

The crux of this paper rests on six primary sources of surveyor bias as identified in PLS surveys conducted from 1786 to 1866 in the Midwest. In addition to four biases (design, azimuthal, species, diameter) investigated in previous papers (Table 1), this study introduces two novel surveyor biases (pair angle, near-post). Furthermore, the previously recognized quadrant (design) bias and diameter bias for medium-sized trees are expanded and subsumed under the newly framed design and diameter biases, respectively. Diameter bias, selectively choosing small trees, is accommodated by truncating all trees below a veil-line, simultaneously yielding an explicit lower diameter limit for the density. Even the conceptually unchanged azimuthal and species biases are evaluated with new adjunct techniques (transect bias, relative distance).

This paper makes no assumptions of CSR spatial patterns of trees or random sampling by surveyors that have clouded past density estimation in PLS studies (Grimm, 1984). Significant previous uncertainties in density estimation are overcome if the analyst: uses a PDE robust to non-CSR spatial patterns; uses a PDE congruent with the surveyor's sampling design; and corrects relevant biases. The Morisita II PDE with its inherent design of the nearest witness trees in opposite halves (2oH) facilitates the derivation of the base density ($\lambda_{MII}$). The Morisita II plotless density estimator (Equation 1 with $g = 1$: distance rank nearest, $k = 2$: 2-tree) relaxes any assumption of homogenous density. Because the Morisita PDE has been found to be robust to underlying nonrandom spatial patterns (Cogbill et al., 2018), violation of the random dispersion assumption is inconsequential.

Deviations of empirical observations from theoretical expectations of the nearest tree to the post define the various spatial biases and their correction. The assumption of an isotropic pattern of witness tree directions from the corner results in a prediction of random bearings for three spatial biases (design, azimuthal, pair angle). These three bias indices plus diameter bias adopt a nearest tree expectation that is independent of the spatial pattern. Four other bias indices (near-post, species, transect, MoR), in which expectations are affected by the underlying spatial pattern, are based on a simulated inhibited-inhomogeneous ($InH^2$) expectation, more realistic than CSR. This approach obviates the need for the assumption of random dispersion (CSR) of trees.

The correction factors adjust the base density estimate derived from empirical (perhaps biased) distance measurements to an unbiased density. Significantly, four biases (design geometry, pair angle, azimuthal, near-post) essentially restrict the spatial area from which the witness tree is chosen. The density corrections for the biases are quantified in three ways: the transposition of the witness trees to the nearest in an adjusted sector (design, pair angle, azimuthal, transect, quadrant, and SAD biases); the use of the next (second) nearest tree for bypassed trees in the same sector (near-post, species, and MoR biases); or the complete truncation of small-sized trees regardless of distance rank (diameter bias). The bypassed trees, albeit the nearest, are absent from the surveyors' empirical sample and their presence is reconstructed by comparison with expected geometric probability, computer simulations at various densities and spatial patterns, and/or analogue reference frequencies. The base density underestimate is "corrected" by quantifying the trees bypassed by the surveyor (extending Kronenfeld & Wang, 2007). The sampling of the nearest tree, even if transposed to a new sector, obviates the need for the assumption of surveyors' random sampling of trees.

Due to the inappropriateness of small trees as witness trees, there was a universal undersampling of trees less than 20-cm diameter. Contrariwise, the surveyors also had a partial preference for small trees in particular situations (see above). The previous widespread opinion that surveyors were biased for certain sizes and species (e.g., Bourdo, 1956; Grimm, 1984; Kronenfeld & Wang, 2007; Manies et al., 2001; Nelson, 1997; Tulowiecki, 2014) was, principally, a result of surveyors selectively sampling small trees. This study proposes a normalization of density using only trees with a diameter above a veil-line. This correction represents not as much a surveyor bias against small trees, as it is a post hoc elimination of the preferred subsample of small trees actually sampled. To my knowledge, this procedure has previously been applied only in PalEON-connected studies (e.g., Goring et al., 2016; Knight et al., 2020; Paciorek et al., 2021). All other PLS studies simply cite a minimum diameter, thus underestimating density by the partial omission of small trees (bypassed trees between the minimum diameter and the veil-line).

The warnings by Bourdo (1956) and Grimm (1984) about surveyors' nonrandom sampling of witness trees were prophetic. Surveyors regularly countered the "nearest tree conjecture" and bypassed nearly half of the trees nearest the corner. The reasons for the surveyors' biased choices of witness trees were overwhelmingly pragmatic: replacing trees removed in preparing the line or corner; choosing trees from a more efficacious position; or selecting small trees when larger ones were less available. The surveyors avoided roughly 70% of the nearest small (12.7 cm $\leq$ diameter < 20 cm) trees and bypassed 17% of the larger potential witness trees (Table 7). Undoubtedly, individual surveyors were occasionally biased for species, medium size, or tree condition, but these tree-based biases postulated by Bourdo (1956) and Grimm (1984) were negligible. In contrast, the surveyors made other nonrandom choices of witness trees, albeit correctable for density estimation. The elimination of small tree bias through a veil-line truncation ($\phi = 0.80$) offset the average quantifiable surveyor bias in PLS Midwest surveys ($\psi = 1.24$).

Pivotally, this paper provides a framework of analytical techniques for deriving unbiased estimates of forest structure before Euro-American settlement. Across the Midwest (Appendix S6: Table S1; Goring et al., 2016; Paciorek et al., 2021), this method produces a historical forest density substantially different from previous studies. Applying this comprehensive methodology allows an accurate estimate and potential reconsideration of tree density in any PLS survey regardless of nonrandom tree dispersion, nonstationary densities, and/or nonrandom choice of witness trees by the surveyor. This formulation resolves the 70-year-old conundrum of surveyor bias and nonrandomness in density estimation for PLS surveys.

## ACKNOWLEDGMENTS

This material is based upon work carried out for the PalEON Project (paleonproject.org) with support from the National Science Foundation MacroSystems Biology program under grants DEB-1065702, DEB-1065656, and DEB-1241868. I thank the PalEON participants, particularly Jody Peters, Chris Paciorek, Jack Williams, Jason McLachlan, and David Mladenoff, for their support and discussions over the years. Earlier versions of this manuscript were improved by comments made by John Battles, Chris Paciorek, Rachel Walker Cogbill, and two anonymous reviewers. I am deeply indebted to all the surveyors and crews who conducted the original surveys, the archivists and curators who preserved the public land survey records, and the researchers, especially Jody Peters, Eric Grimm, Jim Almendinger, David Mladenoff, Ted Sickley, Dennis Albert, Ed Schools, Jim Dyer, Jack Williams, and undergraduates in the McLachlan Lab at the University of Notre Dame and in the Williams Lab at the University of Wisconsin-Madison, who transcribed, collated, digitized, and shared databases of the survey records.

## CONFLICT OF INTEREST STATEMENT

The author declares no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The raw survey data from which all the analyses and summaries are derived are available via the LTER Network Data Portal: Indiana (McLachlan, 2020a; https://doi.org/10.6073/pasta/c3e2404f5b34204b5871a743ebce3c51); Illinois (McLachlan, 2020b; https://doi.org/10.6073/pasta/b8fdabd7cfba3a2b3d55fe4c1dc5383f); Northern Michigan (McLachlan & Williams, 2020a; https://doi.org/10.6073/pasta/3760eec82562e0a8b7cd493c0a3e3ef4); Southern Michigan (McLachlan & Williams, 2020b; https://doi.org/10.6073/pasta/8d033c1cfadca42bf060f9f38940c81e); Southeastern Michigan (McLachlan, 2020c; https://doi.org/10.6073/pasta/409ec6dfb218b6a3e98022916d2b4438); Wisconsin (Mladenoff et al., 2020; https://doi.org/10.6073/pasta/c3e680e51026e74a103663ffa16cb95d); Minnesota (Williams & McLachlan, 2020; https://doi.org/10.6073/pasta/f55f6b7f4060a9b4f07374e7db8443cd); Ohio (Cogbill, 2022a; https://doi.org/10.6073/pasta/ab7ff96c75a62e0d05b6efb8c02eae29); Northern Wisconsin Case Study (Cogbill, 2022b; https://doi.org/10.6073/pasta/00e0b5842fb0a64a2191a4afb93031cc). Simulations of dispersion patterns and virtual plotless sampling were done with R (R Core Team, 2019, version R-3.6.2) in RStudio (version 1.3.959). Codes for R simulations and plotless sampling under different forest dispersions and sampling designs are in the program PlotlessPatternSimulationver-4.R (Peters, 2023), available from Zenodo: https://doi.org/10.5281/zenodo.7871803.

## REFERENCES

Abrams, M. D., G. J. Nowacki, and B. B. Hanberry. 2022. "Oak Forests and Woodlands as Indigenous Landscapes in the Eastern United States." *Bulletin of the Torrey Botanical Society* 149(2): 101–121.

Almendinger, J. L. 1996. *Minnesota's Bearing Tree Database*. Biological Rept. No. 56. St. Paul, MN: Minnesota Department of Natural Resources.

Anderson, R. C., and M. R. Anderson. 1975. "The Presettlement Vegetation of Williamson County, Illinois." *Castanea* 40: 345–363.

Anderson, R. C., S. L. Jones, and R. Swigart. 2006. "Modifying Distance Methods to Improve Estimates of Historical Tree Density from General Land Office Survey Records." *Journal of the Torrey Botanical Society* 133: 449–459.

Ashby, W. C. 1972. "Distance Measurements in Vegetation Study." *Ecology* 53(5): 980–81.

Baker, W. L., C. T. Hanson, M. A. Williams, and D. A. DellaSala. 2023. "Countering Omitted Evidence of Variable Historical Forests and Fire Regime in Western USA Dry Forests: The Low-Severity-Fire Model Rejected." *Fire* 2023(6): 146. https://doi.org/10.3390/fire6040146.

Bjorkman, A. D., and M. Vellend. 2010. "Defining Historical Baselines for Conservation: Ecological Changes since European Settlement on Vancouver Island, Canada." *Conservation Biology* 24(6): 1559–68.

Blankenship, K., R. Swaty, K. R. Hall, S. Hagen, K. Pohl, A. S. Hunt, J. Patton, L. Frid, and J. Smith. 2021. "Vegetation Dynamics Models: A Comprehensive Set for Natural Resource Assessment and Planning in the United States." *Ecosphere* 12(4): e03484. https://doi.org/10.1002/ecs2.3484.

Bollinger, J., L. A. Schulte, S. N. Burrows, T. A. Sickley, and D. J. Mladenoff. 2004. "Assessing Ecological Potentials of Wisconsin (U.S.A.) Using Historical Landscape Reconstructions." *Restoration Ecology* 12: 124–142.

Bouldin, J. 2008. "Some Problems and Solutions in Density Estimation from Bearing Tree Data: A Review and Synthesis." *Journal of Biogeography* 35(11): 2000–2011.

Bouldin, J. 2010. "Issues in Estimates of Relative Metrics of Historic Forest Conditions from Bearing Tree Data." *Ecological Applications* 20(4): 1183–87.

Bourdo, E. A., Jr. 1954. "A Validation of Methods Used in Analyzing Original Forest Cover." PhD thesis, The University of Michigan.

Bourdo, E. A., Jr. 1956. "A Review of the General Land Office Survey and of Its Use in Quantitative Studies of Former Forests." *Ecology* 37: 754–768.

Bowles, M., J. McBride, and L. Bell. 1998. "Landscape Vegetation Pattern, Composition and Structure of DuPage County, Illinois, as Recorded by the U.S. Public Land Survey (1821-1840)." Report to the DuPage County Forest Preserve District, The Morton Arboretum, Lisle, Illinois. www.plantconservation.us/plspub.html.

Clark, D. C. 2000. "The Surveyors Surveyed: Investigations of Bias in General Land Office Surveyor Data for Northeastern Illinois, 1837–1840." MA thesis, Northeastern Illinois University.

Cogbill, C. 2022a. "Witness Tree Database for 4 Survey Districts in Southeastern Ohio, Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/ab7ff96c75a62e0d05b6efb8c02eae29.

Cogbill, C. 2022b. "Witness Tree Database for 68 Townships in Northern Wisconsin, Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/00e0b5842fb0a64a2191a4afb93031cc.

Cogbill, C. V., A. L. Thurman, J. W. Williams, J. Zhu, D. J. Mladenoff, and S. J. Goring. 2018. "A Retrospective on the Accuracy and Precision of Plotless Forest Density Estimators in Ecological Studies." *Ecosphere* 9(4): e02187. https://doi.org/10.1002/ecs2.2187.

Cottam, G. 1949. "The Phytosociology of an Oak Woods in Southwestern Wisconsin." *Ecology* 30: 271–287.

Cottam, G., and J. T. Curtis. 1949. "A Method for Making Rapid Surveys of Woodlands by Means of Pairs of Randomly Selected Trees." *Ecology* 30: 101–4.

Cottam, G., and J. T. Curtis. 1956. "The Use of Distance Measures in Phytosociological Sampling." *Ecology* 37: 451–460.

Cottam, G., J. T. Curtis, and B. W. Hale. 1953. "Some Sampling Characteristics of Population of Randomly Dispersed Individuals." *Ecology* 34: 741–757.

Cox, L. E., and J. L. Hart. 2015. "Two Centuries of Forest Compositional and Structural Changes in the Alabama Fall Line Hills." *American Midland Naturalist* 174(2): 218–237.

Delcourt, H. R. 1976. "Presettlement Vegetation of the North of Red River Land District, Louisiana." *Castanea* 41(2): 122–139.

Delcourt, H. R., and P. A. Delcourt. 1974. "Primeval Magnolia-Holly-Beech Climax in Louisiana." *Ecology* 55: 638–644.

Delcourt, H. R., and P. A. Delcourt. 1977. "Presettlement Magnolia-Beech Climax of the Gulf Coastal Plain: Quantitative Evidence from the Apalachicola River Bluffs, North-Central Florida." *Ecology* 58: 1085–93.

Dorney, C. H., and J. R. Dorney. 1989. "An Unusual Oak Savanna in Northeastern Wisconsin: The Effect of Indian-Caused Fire." *American Midland Naturalist* 122: 103–113.

Dyer, J. M. 2001. "Using Witness Trees to Access Forest Change in Southeastern Ohio." *Canadian Journal of Forest Resources* 31: 1708–18.

Dyer, J. M., and T. F. Hutchinson. 2019. "Topography and Soils-Based Mapping Reveals Fine-Scale Compositional Shifts over Two Centuries within a Central Appalachian Landscape." *Forest Ecology and Management* 433: 33–42.

Edgin, B., and J. E. Ebinger. 1997. "Barrens and the Forest-Prairie Interface in Presettlement Crawford County, Illinois." *Castanea* 62(4): 260–67.

Eyre, F. H., and W. M. Zillgitt. 1953. *Partial Cuttings in Northern Hardwoods of the Lake States: Twenty-Year Experimental Results*. USDA Technical Bulletin No. 1076. Broomall, PA: U.S. Department of Agriculture, 1–124.

Fralish, J. S., A. Carver, R. C. Anderson, R. G. Thurau, S. Sauer, C. Ruffner, D. D. Close, R. Swigart, and E. M. White. 2002. *Presettlement, Present, and Projected Forest Communities of the Shawnee National Forest, Illinois*. Harrisburg, IL: USDA Forest Service, Shawnee National Forest. 145 pp.

Fralish, J. S., and T. G. McArdle. 2009. "Forest Dynamics across Three Century-Length Disturbance Regimes in the Illinois Ozark Hills." *American Midland Naturalist* 162: 418–449.

Friedman, S. K., and P. B. Reich. 2005. "Regional Legacies of Logging: Departure from Presettlement Forest Conditions in Northern Minnesota." *Ecological Applications* 15: 726–744.

Gordon, R. B. 1969. *The Natural Vegetation of Ohio in Pioneer Days*. Columbus: Ohio State University. Ohio Biological Survey Bulletin New Series, Vol. 3, no. 2, 1–113.

Goring, S. J., D. J. Mladenoff, C. V. Cogbill, S. Record, C. J. Paciorek, S. T. Jackson, M. C. Dietze, et al. 2016. "Novel and

Lost Forests in the Upper Midwestern United States, from New Estimates of Settlement-Era Composition, Stem Density, and Biomass." *PLoS One* 11: e0151935. https://doi.org/10.1371/journal.pone.0151935.

Grimm, E. C. 1981. "An Ecological and Paleoecological Study of the Vegetation in the Big Woods Region of Minnesota." PhD diss., University of Minnesota.

Grimm, E. C. 1984. "Fire and Other Factors Controlling the Big Woods Vegetation of Minnesota the Mid-Nineteenth Century." *Ecological Monographs* 54: 291–311.

Habeck, J. R. 1994. "Using General Land Office Records to Assess Forest Succession in Ponderosa Pine/Douglas-Fir Forests in Western Montana." *Northwest Science* 68: 69–78.

Hanberry, B. B. 2020. "Baseline to Novel Ecosystems in Michigan, USA, with a Quantitative and Qualitative Assessment." *Écoscience* 27(4): 255–268.

Hanberry, B. B. 2021. "Forest Disturbance Types and Current Analogs for Historical Disturbance-Independent Forests." *Land* 10: 136. https://doi.org/10.3390/land10020136.

Hanberry, B. B., D. C. Bragg, and H. D. Alexander. 2020. "Open Forest Ecosystems: An Excluded State." *Forest Ecology and Management* 472: 118256.

Hanberry, B. B., K. Coursey, and J. S. Kush. 2018. "Structure and Composition of Historical Longleaf Pine Ecosystems in Mississippi, USA." *Human Ecology* 46: 241–48.

Hanberry, B. B., and D. C. Dey. 2019. "Historical Range of Variability for Restoration and Management in Wisconsin." *Biodiversity and Conservation* 28: 2931–50.

Hanberry, B. B., S. Fraver, H. S. He, Y. Jian, D. C. Dey, and B. J. Palik. 2011. "Spatial Pattern Corrections and Sample Sizes for Forest Density Estimates of Historical Tree Surveys." *Landscape Ecology* 26(1): 59–68.

Hanberry, B. B., and H. S. He. 2015. "Effects of Historical and Current Disturbance on Forest Biomass in Minnesota." *Landscape Ecology* 30: 1473–82.

Hanberry, B. B., H. S. He, and S. R. Shifley. 2016. "Loss of Aboveground Forest Biomass and Landscape Biomass Variability in Missouri, US." *Ecological Complexity* 25: 11–17.

Hanberry, B. B., D. C. Justice, and D. C. Powell. 2020. "Discovering Douglas-Fir Woodlands in the Historical Forests of Umatilla National Forest, Eastern Oregon and Washington." *Forests* 11(10): 1122. https://doi.org/10.3390/f11101122.

Hanberry, B. B., J. M. Kabrick, and H. S. He. 2014. "Densification and State Transition across the Missouri Ozarks Landscape." *Ecosystems* 17: 66–81.

Hanberry, B. B., J. M. Kabrick, and H. S. He. 2015. "Potential Tree and Soil Carbon Storage in a Major Historical Floodplain Forest with Disrupted Ecological Function." *Perspectives in Plant Ecology, Evolution and Systematics* 17: 7–23.

Hanberry, B. B., J. M. Kabrick, H. S. He, and B. J. Palik. 2012. "Historical Trajectories and Restoration Strategies for the Mississippi River Alluvial Valley." *Forest Ecology and Management* 280: 103–111.

Hanberry, B. B., B. J. Palik, and H. S. He. 2012. "Comparison of Historical and Current Forest Surveys for Detection of Homogenization and Mesophication of Minnesota Forests." *Landscape Ecology* 27: 1495–1512.

Hanberry, B. B., J. Yang, J. M. Kabrick, and H. S. He. 2012. "Adjusting Forest Density Estimates for Surveyor Bias in Historical Tree Surveys." *American Midland Naturalist* 167(2): 285–306.

Hushen, T. W., R. O. Kapp, R. D. Bogue, and J. T. Worthington. 1966. "Presettlement Forest Patterns in Montcalm County, Michigan." *Michigan Botanist* 5: 192–211.

Janke, R. A., D. McKaig, and R. Raymond. 1978. "Comparison of Presettlement and Modern Upland Boreal Forests on Isle Royale National Park." *Forest Science* 24: 115–121.

Kenoyer, L. A. 1930. "Ecological Notes on Kalamazoo County, Michigan, Based on the Original Land Survey." *Michigan Academy of Science Arts, and Letters* 11: 211–17.

Kline, V. M., and G. Cottam. 1979. "Vegetation Response to Climate and Fire in the Driftless Area of Wisconsin." *Ecology* 60: 861–68.

Knight, C. A., C. V. Cogbill, M. D. Potts, J. A. Wanket, and J. J. Battles. 2020. "Settlement-Era Forest Structure and Composition in the Klamath Mountains: Reconstructing a Historical Baseline." *Ecosphere* 11(9): e03250. https://doi.org/10.1002/ecs2.3250.

Kronenfeld, B. 2014. "Validating the Historical Record: A Relative Distance Test and Correction Formula for Selection Bias in Presettlement Land Surveys." *Ecography* 38(1): 41–53.

Kronenfeld, B. J., and Y.-C. Wang. 2007. "Accounting for Surveyor Inconsistency and Bias in Estimation of Tree Density from Presettlement Land Survey Records." *Canadian Journal of Forest Research* 37: 2365–79.

Leahy, M. J., and K. S. Pregitzer. 2003. "A Comparison of Presettlement and Present-Day Forests in Northeastern Lower Michigan." *American Midland Naturalist* 149(1): 71–89.

Leitner, L. A., C. P. Dunn, G. R. Guntenspergen, F. Stearns, and D. M. Sharpe. 1991. "Effects of Site, Landscape Features, and Fire Regime on Vegetational Patterns in Presettlement Southern Wisconsin." *Landscape Ecology* 5: 203–217.

Levine, C. R., C. V. Cogbill, B. M. Collins, A. J. Larson, J. A. Lutz, M. P. North, C. M. Restaino, H. D. Safford, S. L. Stephens, and J. J. Battles. 2017. "Evaluating a New Method for Reconstructing Forest Conditions from General Land Office Survey Records." *Ecological Applications* 27(5): 1498–1513.

Liu, F., D. J. Mladenoff, N. S. Keuler, and L. S. Moore. 2011. "Broad-Scale Variability in Tree Data of the Historical Land Survey and Its Consequences for Ecological Studies." *Ecological Monographs* 81(2): 259–275.

Lutz, H. J. 1930. "Original Forest Composition in Northwestern Pennsylvania as Indicated by Early Land Survey Notes." *Journal of Forestry* 28(8): 1098–1103.

Manies, K. L. 1997. "Evaluation of the General Land Office Survey Records for Analysis of the Northern Great Lakes Hemlock-Hardwood Forests." MS thesis, University of Wisconsin.

Manies, K. L., and D. J. Mladenoff. 2000. "Testing Methods to Produce Landscape-Scale Presettlement Vegetation Maps from the U.S. Public Land Survey Records." *Landscape Ecology* 15: 741–754.

Manies, K. L., D. J. Mladenoff, and E. V. Nordheim. 2001. "Assessing Large-Scale Surveyor Variability in the Historic Forest Data of the Original US Public Land Survey." *Canadian Journal of Forest Research* 31(10): 1719–30.

McLachlan, J. 2020a. "Settlement Trees, Indiana Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/c3e2404f5b34204b5871a743ebce3c51.

McLachlan, J. 2020b. "Settlement Trees, Illinois Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/b8fdabd7cfba3a2b3d55fe4c1dc5383f.

McLachlan, J. 2020c. "Settlement Trees, Southeastern Michigan Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/409ec6dfb218b6a3e98022916d2b4438.

McLachlan, J., and J. Williams. 2020a. "Settlement Trees, Northern Michigan Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/3760eec82562e0a8b7cd493c0a3e3ef4.

McLachlan, J., and J. Williams. 2020b. "Settlement Trees, Southern Michigan Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/8d033c1cfadca42bf060f9f38940c81e.

Meunier, J., and M. E. Shea. 2020. "Applying the Usual Rules to an Unusual Ecological Situation: Fire Rotation in Great Lakes Pine Forests." *Forest Ecology and Management* 472(2020): 118246.

Mladenoff, D., J. Williams, and J. McLachlan. 2020. "Settlement Trees, Wisconsin Level 0 Ver 0." Environmental Data Initiative. https://doi.org/10.6073/pasta/c3e680e51026e74a103663ffa16cb95d.

Morisita, M. 1954. *Estimation of Population Density by Spacing Method (in English)*. Fukuoka: Memoirs of the Faculty of Science, Kyushu University, Series E, (Biology), Vol. 1, no. 1.

Morisita, M. 1957. "A New Method for the Estimation of Density by Spacing Method Applicable to Non-Randomly Distributed Populations." *Physiology and Ecology* 7: 134–144. Seiri Seitai 7: 134–144 (in Japanese with English Summary) also English Translation at http://people.hws.edu/mitchell/Morisita1957.pdf.

Nelson, J. C. 1997. "Presettlement Vegetation Patterns along the 5th Principal Meridian, Missouri Territory, 1815." *American Midland Naturalist* 137(1): 79–94.

Oliver, C. D., and B. C. Larson. 1990. *Forest Stand Dynamics*. New York: McGraw-Hill Inc.

Paciorek, C. J., C. V. Cogbill, J. A. Peters, S. J. Goring, J. W. Williams, D. J. Mladenoff, A. Dawson, and J. S. McLachlan. 2021. "The Forests of the Midwestern United States at Euro-American Settlement: Spatial and Physical Structure Based on Contemporaneous Survey Data." *PLoS One* 16(2): e0246473. https://doi.org/10.1371/journal.pone.0246473.

Peters, J. 2023. "PalEON-Project/Bias-Estimation-Forest-Density: Bias Estimation for Forest Density (v1.0.0)." Zenodo. https://doi.org/10.5281/zenodo.7871804.

Pielou, E. C. 1959. "The Use of Point-to-Plant Distances in the Study of the Pattern of Populations." *Journal of Ecology* 47: 607–613.

Pollard, J. H. 1971. "On Distance Estimators of Density in Randomly Distributed Forests." *Biometrics* 27: 991–1002.

Price, E. T. 1995. *Dividing the Land*. Chicago, IL: University of Chicago Press.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Raiho, A. M., C. J. Paciorek, A. Dawson, S. T. Jackson, D. J. Mladenoff, J. W. Williams, and J. S. McLachlan. 2022. "8000-Year Doubling of Midwestern Forest Biomass Driven by Population- and Biome-Scale Processes." *Science* 376: 1491–95.

Rhemtulla, J. M., and D. J. Mladenoff. 2010. "Relative Consistency, Not Absolute Precision, Is the Strength of the Public Land Survey: Response to Bouldin." *Ecological Applications* 20(4): 1187–89.

Rhemtulla, J. M., D. J. Mladenoff, and M. K. Clayton. 2009. "Historical Forest Baselines Reveal Potential for Continued Carbon Sequestration." *Proceedings of the National Academy of Sciences of the United States of America* 106(15): 6082–87.

Schulte, L. A., and D. J. Mladenoff. 2001. "The Original US Public Land Survey Records." *Journal of Forestry* 99(10): 5–10.

Schulte, L. A., D. J. Mladenoff, T. R. Crow, L. C. Merrick, and D. T. Cleland. 2007. "Homogenization of Northern US Great Lakes Forests Due to Land Use." *Landscape Ecology* 22(7): 1089–1103.

Sears, P. B. 1921. "Vegetation Mapping." *Science* 53: 325–27.

Sears, P. B. 1925. "The Natural Vegetation of Ohio I: A Map of the Virgin Forest." *Ohio Journal of Science* 25: 139–149.

Shanks, R. E. 1938. "The Original Vegetation of a Part of the Lake-Plain of Northwestern Ohio: Wood and Henry Counties." PhD diss., The Ohio State University.

Shen, G., X. Wang, and F. He. 2020. "Distance-Based Methods for Estimating Density of Nonrandomly Distributed Populations." *Ecology* 101(10): e03143. https://doi.org/10.1002/ecy.3143.

Siccama, T. G. 1971. "Presettlement and Present Forest Vegetation in Northern Vermont with Special Reference to Chittenden County." *American Midland Naturalist* 85: 153–172.

Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*, 2nd ed. San Francisco, CA: W. H. Freeman and Company.

Stewart, L. O. 1935. *Public Land Surveys: History, Instructions, Methods*. Ames, IA: Collegiate Press.

Tatina, R. E., and B. B. Hanberry. 2022. "Historical Forests of the Black Hills, South Dakota, USA, Determined Using General Land Office Surveys." *Silva Fennica* 56(3): 17. https://doi.org/10.14214/sf.10754.

Thompson, H. R. 1956. "Distribution of Distance to Nth Neighbour in a Population of Randomly Distributed Individuals." *Ecology* 37(2): 391–94.

Tulowiecki, S. J. 2014. "Assessing Native American Influences upon Pre-European Tree Species Composition in Western New York: An Approach Using Original Land Survey Records and Species Distribution Models." PhD thesis, University at Buffalo, SUNY.

Tulowiecki, S. J., C. P. S. Larsen, and Y.-C. Wang. 2015. "Effects of Positional Error on Modeling Species Distributions: A Perspective Using Presettlement Land Survey Records." *Plant Ecology* 216(1): 67–85.

Van Deelen, T. R., K. S. Pregitzer, and J. B. Haufler. 1996. "A Comparison of Presettlement and Present-Day Forests in Two Northern Michigan Deer Yards." *American Midland Naturalist* 135(2): 181–194.

Wang, Y.-C. 2004. "A Geographic Information Approach to Analyzing and Visualizing Presettlement Land Survey Records for Vegetation Reconstruction." PhD thesis, State University of New York at Buffalo.

Ward, R. T. 1954. "A Phytosociological Study of the Beech Forests of Wisconsin." PhD thesis, University of Wisconsin-Madison.

Ward, R. T. 1956. "The Beech Forests of Wisconsin—Changes in Forest Composition and the Nature of the Beech Border." *Ecology* 37: 407–419.

Warde, W., and J. W. Petranka. 1981. "A Correction Factor Table for Missing Point-Center Quarter Data." *Ecology* 62: 491–94.

White, A. S. 1976. "Evaluating Land Survey Notes to Determine the Pre-Settlement Structure of the Lubrecht Forest, Montana." MS thesis, University of Montana.

White, C. A. 1983. *A History of the Rectangular Survey System*. Washington, DC: USDI Bureau of Land Management, Superintendent of Documents.

Whitney, G. G. 1994. *From Coastal Wilderness to Fruited Plain: A History of Environmental Change in Temperate North America from 1500 to the Present*. Cambridge: Cambridge University Press.

Whitney, G. G., and J. R. Steiger. 1985. “Site-Factor Determinants of the Presettlement Prairie-Forest Border Areas of North-Central Ohio.” *Botanical Gazette* 146: 421–430.

Williams, J., and J. McLachlan. 2020. “Settlement Trees, Minnesota Level 0 Ver 0.” Environmental Data Initiative. https://doi.org/10.6073/pasta/f55f6b7f4060a9b4f07374e7db8443cd.

Williams, M. A., and W. L. Baker. 2010. “Bias and Error in Using Survey Records for Ponderosa Pine Landscape Restoration.” *Journal of Biogeography* 37(4): 707–721.

Williams, M. A., and W. L. Baker. 2011. “Testing the Accuracy of New Methods for Reconstructing Historical Structure of Forest Landscapes Using GLO Survey Data.” *Ecological Monographs* 81(1): 63–88.

Zhang, Q. F., K. S. Pregitzer, and D. D. Reed. 2000. “Historical Changes in the Forests of the Luce District of the Upper Peninsula of Michigan.” *American Midland Naturalist* 143(1): 94–110.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Cogbill, Charles V. 2023. “Surveyor and Analyst Biases in Forest Density Estimation from United States Public Land Surveys.” *Ecosphere* 14(8): e4647. https://doi.org/10.1002/ecs2.4647