Journal of Biogeography (J. Biogeogr.) (2010) 37, 707–721

ORIGINAL ARTICLE

Bias and error in using survey records for ponderosa pine landscape restoration

Mark A. Williams* and William L. Baker

Program in Ecology and Department of Geography, Department 3371, 1000 E. University Avenue, University of Wyoming, Laramie, WI 82071, USA

ABSTRACT

Aim Public land survey records are commonly used to reconstruct historical forest structure over large landscapes. Reconstruction studies have been criticized for using absolute measures of forest attributes, such as density and basal area, because of potential selection bias by surveyors and unknown measurement error. Current methods to identify bias are based upon statistical techniques whose assumptions may be violated for survey data. Our goals were to identify and directly estimate common sources of bias and error, and to test the accuracy of statistical methods to identify them.

Location Forests in the western USA: Mogollon Plateau, Arizona; Blue Mountains, Oregon; Front Range, Colorado.

Methods We quantified both selection bias and measurement error for survey data in three ponderosa pine landscapes by directly comparing measurements of bearing trees in survey notes with remeasurements of bearing trees at survey corners (384 corners and 812 trees evaluated).

Results Selection bias was low in all areas and there was little variability among surveyors. Surveyors selected the closest tree to the corner 95% to 98% of the time, and hence bias may have limited impacts on reconstruction studies. Bourdo's methods were able to successfully detect presence or absence of bias most of the time, but do not measure the rate of bias. Recording and omission errors were common but highly variable among surveyors. Measurements for bearing trees made by surveyors were generally accurate. Most bearings were less than 5° in error and most distances were within 5% of our remeasurements. Many, but not all, surveyors in the western USA probably estimated diameter of bearing trees at stump height (0.3 m). These estimates deviated from reconstructed diameters by a mean absolute error of 7.0 to 10.6 cm.

Main conclusions Direct comparison of survey data at relocated corners is the only method that can determine if bias and error are meaningful. Data from relocated trees show that biased selection of trees is not likely to be an important source of error. Many surveyor errors would have no impact on reconstruction studies, but omission errors have the potential to have a large impact on results. We suggest how to reduce potential errors through data screening.

*Correspondence: Mark A. Williams, Program in Ecology and Department of Geography, Department 3371, 1000 E. University Avenue, University of Wyoming, Laramie, WY 82071, USA.
E-mail: mwilli58@uwyo.edu

Keywords

Measurement error, *Pinus ponderosa*, public land surveys, reconstruction, restoration, sampling bias, western USA.

INTRODUCTION

Large portions of forests in the western USA are considered to be in an unhealthy state because of past fire suppression, intensive timber harvesting, and overgrazing (e.g. Baker *et al.*, 2007). Benchmarks for assessing restoration needs come from fine-scale forest reconstruction studies, which are then applied to much larger areas (e.g. Fulé *et al.*, 1997). These studies have high precision but may miss much of the variability of forests across large areas. However, in the USA a systematic survey of

doi:10.1111/j.1365-2699.2009.02257.x

historical vegetation is available, namely the General Land Office (GLO) surveys, which can be used to reconstruct vegetation and disturbances across very large landscapes (Galatowitsch, 1990).

The GLO surveys, widely used elsewhere in the USA (Whitney, 1986; Batek *et al.*, 1999; Zhang *et al.*, 1999; Schulte & Mladenoff, 2005), have been underutilized in the West (but see White, 1976; Habeck, 1994; Arundel, 2000; Langley, 2004). Surveys were largely completed in the mid-to-late 1800s in the West, shortly after or before extensive Euro-American settlement (Galatowitsch, 1990). Surveyors recorded information on bearing trees at survey corners, including tree species, size, bearing and distance from the section corner. These data can be processed and used to estimate parameters of historical forest structure such as composition, absolute tree density, and tree diameter distribution (Anderson & Anderson, 1975; Radeloff *et al.*, 1999; Dyer, 2001; Schulte & Mladenoff, 2001; Leahy & Pregitzer, 2003). Maps of forest structure can then be analysed to provide an understanding of the variability of forests across large land areas (e.g. He *et al.*, 2000). Surveyors also recorded ecotones and disturbances encountered along survey lines. This information can be used to evaluate large changes in forest structure, such as forest type conversions and loss or gain in prairies/meadows, and to estimate disturbance rotations (Anderson & Anderson, 1975; Canham & Loucks, 1984; Schulte & Mladenoff, 2005; Andersen & Baker, 2006; Fritschle, 2008). Spatial structure and disturbance affect the productivity of trees available for future harvesting (e.g. Busse *et al.*, 2000), restoration goals relative to fire exclusion and land uses, and habitat for wildlife. The GLO surveys are one of the few sources of data available for reconstructing the historical variability of forest structure and composition, and determining disturbance history across large landscapes. Land survey data and other mapping efforts have been used in many other parts of the world to assess land cover changes over modern times, such as the crown land surveys in Canada, the public land surveys in Australia, and the Siegfried maps in Switzerland (e.g. Fensham & Fairfax, 1997; Jackson *et al.*, 2000; Leyk & Zimmerman, 2007).

The GLO was commissioned to create the public land survey system (PLSS) (Stewart, 1935). The PLSS is composed of 9.7 km by 9.7 km (6 mile by 6 mile) townships with smaller subdivisions of thirty-six 2.6-km^2 (one square-mile) sections. Surveyors placed quarter corners at 0.8 km (0.5 mile) along section lines (1.6 km or 1 mile long) and corners at the end of each section line (Fig. 1). The area around each corner was divided into quadrants of 90°. For section corners, a bearing tree was selected in each quadrant (section), and for quarter corners, one bearing tree was selected on either side of the section line (one tree in each section), irrespective of quadrant. In general, two trees were marked (inscribed with the township, range and section) and recorded at quarter corners, and four trees at corners, which is one tree in each section for which the corner stands.

There are current limitations to GLO data that need to be addressed before confidence can be placed in their widespread application. The surveys were not undertaken for scientific reasons, so it is therefore logical to test whether trees recorded at survey corners suffer from *selection bias* (Table 1), that is, whether they fail to represent an unbiased sample for any of several reasons. First, the manuals of surveying used in the West do not specify that the trees selected should each be the closest tree to the corner. Furthermore, instructions in the manuals are confusing. For example, the 1855 manual (Stewart, 1935, p. 179) states: 'from each (corner) post the course shall be taken and the distances measured to two or more adjacent trees, in opposite directions as nearly as may be...'; and later in the manual there are instructions to record a

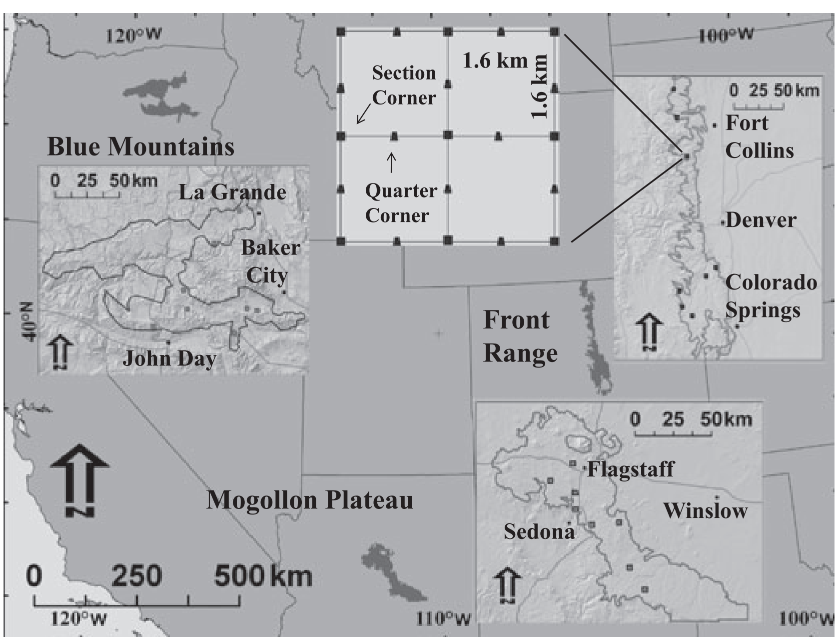

Figure 1 Study area locations in the western USA and sampling grid layout. Dark grey areas on the large map represent study area locations, which are magnified for more detail in inset maps. Squares in inset maps represent sampling grid locations. Sampling grid inset shows the arrangement of the 21 section and quarter corners sampled for each stratum.

Table 1 Eight limitations of the General Land Office surveys and the formulas we used to measure their magnitude.

Limitation	Definition	Terms, formula(s) and sources
Selection bias	Surveyor selected bearing trees preferentially, indicated by selected trees that are not the closest to the corner, so that selected trees do not represent an unbiased sample of the forest	$SBR = 100 \times (n_{nc}/n_{ct})$, where *SBR* is *selection bias rate* (%), n_{nc} is the number of trees not closest to the corner, and n_{ct} is the total number of corners or trees
Recording errors	Surveyor incorrectly recorded bearing-tree information in the field, including writing down incorrect data (e.g. NE instead of NW) in field notes or misidentifying tree species	$RER = 100 \times (n_{re}/n_{ct})$, where *RER* is *recording error rate* (%), n_{re} is the number of corners or trees with a recording error, and n_{ct} is the total number of corners or trees
Omission errors	Surveyor failed to record the required or usual number of trees at a corner, and thus did not comply with the survey instructions for the number of trees to record	$OER = 100 \times (n_o/n_{ct})$, where *OER* is the *omission error rate* (%), n_o is the number of trees omitted, and n_{ct} is the total number of available quadrants
Measurement errors	Errors in a measurement of bearing, distance, or tree diameter because of limitations in measuring devices or measuring methods (e.g. visual estimation of diameter)	Measurement Bias – measured as: $ME = (\sum(M_s - M_r))/n$, where *ME* is *mean error*, M_s is the surveyor measurement, M_r is our re-measurement, and n is the number of measurements Measurement Accuracy – measured as: $MAE = (\sum \lvert(M_s - M_r)\rvert)/n$, where *MAE* is *mean absolute error*, and other symbols are as above Relative Measurement Accuracy – measured as: $RMAE = (\sum (\lvert(M_s - M_r)\rvert/M_r))/n$, where *RMAE* is *relative mean absolute error*, and other symbols are as above
Lineage errors	(1) Surveyors or clerks incorrectly copied from original field survey notes to notes given to the surveyor general's office, or (2) modern users transcribed incorrectly from microfiche or other GLO sources to a computer database, due to illegible entries and typing mistakes	Not measured in this study; see Almendinger (1997)
Species identification ambiguities	Ambiguous or insufficiently precise species identification (e.g. only 'pine' instead of 'ponderosa pine' or 'limber pine')	No absolute measures; see Mladenoff *et al.* (2002)
Survey instruction limitations	Survey instructions imprecise or changing (e.g. did not require two trees at quarter corners and four trees at section corners, or were imprecise about the distance over which surveyors were to search for trees)	No absolute measures; see Grimm (1981) or this paper
Fraud	Surveyors did not acquire data in the field and instead created fictitious data, sometimes as part of an organized crime syndicate	No absolute measures; see Stewart (1935), Livermore (1991)

tree in each section for which the corner stands – one guideline clearly contradicted the other. Second, the bearing trees recorded were to be a lasting monument to the corner. Knowing this, the surveyors might have selected against certain species or size classes (i.e. small or very large trees) suspected to lack longevity (Grimm, 1984; Schulte & Mladenoff, 2001). Third, surveyors may have selected against certain tree species because of economic value (Lutz, 1930) or difficulty in inscribing the necessary marks (Bourdo, 1956). Finally, surveyors might have preferentially selected trees because of location, perhaps avoiding trees near quadrant or section boundaries that required additional compass measurements (Schulte & Mladenoff, 2001). All of these factors have the potential to bias the selection of trees. Conversely, there are other factors that suggest surveyors were unbiased in selecting trees. Prior to 1910, surveyors were paid according to the number of miles they surveyed, and thus might have been less likely to go out of their way to mark a tree at a greater distance (Schulte & Mladenoff, 2001). In addition, many forest types (e.g. lodgepole pine: *Pinus contorta* Douglas ex Louden) experience high-severity fires, resulting in cohorts of similar size, essentially limiting size bias. Furthermore, the number of species at a corner was probably low, especially in the West, which limited selection bias.

High rates of selection bias would be detrimental, as bias can change the species, distance and diameter components used in analysing forest structure. Knowing the tree's distance rank, or distance order from the corner, is important because the methods used to reconstruct forest parameters assume that the trees recorded by the surveyors were the closest in each section

(Cottam & Curtis, 1956; Bouldin, 2008). Statistical examination of surveyor preferences (Manies *et al.*, 2001) along with comparison of species recorded as bearing trees and in line descriptions (Almendinger, 1997) both suggest reliability, but similar testing is needed in the West, where little research has been done.

Bourdo (1956) formulated statistical methods, using chi-square analyses, to detect potential surveyor selection bias for quadrant, species, size and location. Delcourt & Delcourt (1974) improved upon these methods by using an ANOVA for distance tests for species and diameter biases. However, these methods were questioned because the basic assumptions of these tests are violated (Grimm, 1984; Bouldin, 2008), including random and equal distribution of tree species and tree diameters across the area sampled (Grimm, 1981). It is currently unknown how accurate these statistical methods are in detecting true selection bias.

In addition to surveyor selection bias, there are other limitations to survey data that are rarely considered or measured, including *recording errors*, *omission errors* and *measurement errors* (Table 1). The only way to quantify these errors and surveyor bias is by direct remeasurement of original bearing trees at survey corners. We are aware of only two small studies that have directly examined bias and error at survey corners (i.e. White, 1976; Habeck, 1994), both in the western USA, where original bearing trees are more likely to remain.

There are other limitations that have been commonly assessed (Table 1). First are *lineage* errors (Wang, 2005), which Grimm (1981) suggested might be substantial. However, Almendinger (1997) compared the original survey notes both with copies in the microfiche and with databases, and calculated an error rate of only 1–5%. Second are *species identification ambiguities*, which are particularly problematic in diverse forests, where they affect compositional analysis. Mladenoff *et al.* (2002) developed methods to minimize this limitation substantially. Third are *surveyor instruction limitations*, which include noncompliance with guidelines (Grimm, 1981). Instructions for collecting data changed over time, but the methods used by surveyors might not have changed with subsequent manuals, so that assumptions made by researchers for reconstructions could be incorrect (e.g. Grimm, 1981). Furthermore, there may be cases where surveyors altered the methods for their own advantage or ignored instructions. Finally, there are cases of outright *fraud*, which were common in many western states (Stewart, 1935; Galatowitsch, 1990; Livermore, 1991). Fraud can be detected by comparing survey plat maps with modern maps for accuracy on section and township lines (Livermore, 1991).

The goal of this study is thus to address the posited limitations of GLO data, an obstacle for widespread application in spatially comprehensive forest restoration. To do this, we directly quantified surveyor selection bias, surveyor recording errors, omission errors and measurement errors by comparing original measurements with remeasurements of surviving trees at PLSS corners in three large landscapes in the western USA. We described the average measures of trees, for which and against which there was selection, to understand the ranges of tree attributes that might be underrepresented in survey data. We also noted changes in the survey manuals over the duration of the surveys and examined the survey notes to see if surveyors promulgated new guidelines. In addition, we used our direct data to test the ability of indirect statistical methods to accurately detect the presence or absence of surveyor selection bias for quadrant, location, species and size. Quantifying error and selection bias from survey notes will help to elucidate the applicability of survey data for forest reconstructions (i.e. applicable with low bias and error), establish the need for estimator corrections (e.g. Kronenfeld & Wang, 2007), and identify any further potential sources of error.

MATERIALS AND METHODS

Study area and sampling design

To examine the geographic variability of bias and error rates for survey data, we studied large landscapes in three forested areas of the western USA. Selection of study areas was limited to ponderosa pine (*Pinus ponderosa* C. Lawson) forests with some adjacent dry mixed conifer (i.e. ponderosa pine forest mixed with other co-dominant conifers), as these forests are widespread and commonly need restoration (e.g. Baker *et al.*, 2007). Western land cover data were analysed in a GIS to identify potential study areas meeting the following criteria: (1) large contiguous forested areas >500,000 ha, (2) high percentage of public land, and (3) survey notes that were non-fraudulent (Stewart, 1935; Livermore, 1991). Using GIS, survey notes, and on-site visits, we chose three geographically distinct study areas: (1) the Mogollon Plateau, Arizona, (2) the Blue Mountains, Oregon, and (3) the Front Range, Colorado (Fig. 1).

A multipurpose design was created for sampling each study area. To span the range of forest variability, each study area was stratified based on: slope, aspect, elevation and geology (e.g. White, 1976). The first three factors are widely known to affect the distribution and growth of ponderosa pine (Cooper, 1960), and the fourth was correlated with tree volume (Aldrich, 2000). Aspect was classified into four levels, and each of slope and elevation into three levels. All combinations of levels were then created and regrouped into four classes along a temperature–moisture gradient from warm/dry (i.e. steep slope, south-west aspect, low elevation) to cool/wet (i.e. gentle slope, north aspect, high elevation). Geology was simplified into just sedimentary and igneous classes on the Mogollon Plateau and in the Blue Mountains, and into granite and gneiss in the Front Range. Temperature–moisture levels combined with geology classes resulted in eight final strata per study area.

To accommodate data collection for multiple investigations, each sampling location consisted of a 3.2 km × 3.2 km (4 square miles) grid of 21 contiguous PLSS corners (Fig. 1). The laborious nature of the work meant that only one grid could be allocated for each stratum. Its location was randomly chosen using the map of strata in GIS. Because there were 21 corners per grid and eight grids, there was the potential to assess 168 corners per study area and 504 corners in total.

Field and laboratory methods

Using coordinates for corners from the Geographic Coordinate Database (BLM http://www.gcdb.gov), a GPS was used to navigate to each approximate location, where we searched for the corner monument. If the monument itself could not be relocated, but the corner was verified by the presence of an original bearing tree, we could not assess surveyors' measurements of distance and bearing. However, we could assess bearing-tree selection bias, measurement accuracy and error of diameter estimation and species identification (potential recording errors). Corners that were re-monumented in more recent surveys but were not in the original corner location could not be used for bias and error estimation. On the Mogollon Plateau, 130 of 168 corners were relocated, whereas in the Blue Mountains and the Front Range, 145 and 109 corners were relocated, respectively. Surveys on the Mogollon Plateau, in the Blue Mountains and in the Front Range occurred from 1878 to 1904, from 1874 to 1882 and from 1867 to 1882, respectively.

We searched for each original bearing tree using its species, diameter, distance and azimuth from the GLO survey notes. For each bearing tree, we recorded: (1) current status (i.e. alive, snag type, cut stump, no evidence, etc.), (2) azimuth, using a sighting compass (resolution 0.5°), (3) distance, using a laser distance meter (Laser Technology, Englewood, CO, resolution 0.01 m), (4) diameter at stump height (d.s.h., 0.3 m) and breast height (d.b.h., 1.37 m), using a caliper, and (5) species identification. Measurements were made at both d.s.h. and d.b.h. because the height of diameter estimation was not specified in surveyor guidelines (Stewart, 1935); other studies have found that estimates were likely to have been made at stump height (White, 1976; Habeck, 1994). Diameter was measured only on one axis, generally perpendicular to a line from the corner to the tree, to approximate the diameter the surveyor may have estimated in viewing the tree, but there is no way to be sure that we replicated the surveyor measurements exactly. Prior to analysis, surveyor bearing was converted into an azimuth, surveyor diameter was converted from inches to centimetres, and surveyor distance was converted from links (a surveyor unit; 1 link = 20.12 cm or 7.92 in) into metres.

Increment cores were extracted for all live bearing trees at both d.s.h. and d.b.h. to allow estimation of diameter at the time of the surveys, and to determine which height was used by the surveyors. All other live trees, possibly of adequate size (approximately ≥10 cm) at the time of the survey, and located closer than the chosen bearing tree, were also cored and recorded to determine, in part, whether selection bias occurred. Tree cores were mounted and sanded until rings were easily separable, and then were cross-dated using existing chronologies from within each study area from the International Tree Ring Databank (http://www.ncdc.noaa.gov/paleo/treering.html). Diameters at the time of the survey were estimated using equations formulated in or near each region [Mogollon Plateau (Myers, 1963); Blue Mountains (Johnson, 1956; Spada, 1960); Front Range (Myers & Van Deusen, 1958)]. These equations predict past tree diameter using present diameters, the radial growth of each tree since the survey year, and the relationship between bark growth and radial growth. We could visually cross-date 70 to 82% of cores across three study areas, and for the others we were able only to partially cross-date the core.

Statistical approach

In our original design, we planned to analyse bias and error by surveyor, stratum, environment and geology in a nested, multi-factorial ANOVA. However, we did not find as high a percentage of trees at corners, either live or dead, as previously reported (e.g. White, 1976; Habeck, 1994), and sample size precluded this option. Moreover, because of the nature of survey contracts, it was common to have only one surveyor for each 21-point grid, and, therefore, only one surveyor per stratum. In fact, over all three study areas, 88.6% of observations within a stratum (n = 24 total strata) were, on average, from one surveyor. Therefore, an analysis of bias and error using strata was confounded by a surveyor effect, an unforeseen outcome. Further analysis was thus focused only on the effect of surveyors, with surveyors functioning as strata.

Means and standard errors were calculated using a stratified (by surveyor) systematic design. Statistics are reported for individual surveyors (where possible) and for each study area, with means and standard errors based on the proportion of surveyor composition. For example, on the Mogollon Plateau, five surveyors represented 84% of all surveyed corners. For these surveyors, sufficient data were available for estimation of bias and error rates; other surveyors were analysed in a collective stratum and an overall mean and standard error calculated for the study area.

Analysis of bias and accuracy

Mean error rates per corner and per tree were calculated to measure selection bias (Table 1) against trees because of their bearing, location, size and species. Per-corner rates are important for how error is propagated into forest metrics that use the set of trees at the corner scale. Along with a mean selection bias rate, the range of biased azimuths and diameters was calculated, as well as the average difference in distances between selected and biased trees to measure the magnitude of bias when it did occur. The frequency of each species against which there was selection was also calculated.

For purposes of comparison with our direct estimates of selection bias, Bourdo's (1956) statistical approach was implemented. The original survey data from all corners within each grid were used unless they lacked trees or were irregular corner types (e.g. standard corners or closing corners that do not represent four sections). Bourdo's methods are criticized because they assume that tree locations are randomly distributed throughout the entire sample (Grimm, 1984; Bouldin, 2008). Furthermore, to fully test for selection bias, attributes including species and diameters would also have to be randomly distributed throughout the entire sample (Grimm, 1981). Although we could not thoroughly address all of Grimm's concerns, we did quantify spatial pattern at each survey corner using Eberhardt's index (Eberhardt, 1967). To do this, at four points, one in each 90° quadrant, we recorded the distance, diameter and species of the nearest five trees. Each point was located by walking a 45° angle, down the centre of each quadrant, for approximately twice the average tree spacing, estimated by mean distance to the nearest tree in each quadrant. We then used the methods in Prayag & Deshmukh (2000) to determine whether tree spacings were random, uniform or aggregated.

Bourdo's (1956) methods analyse potential selection bias in two ways. The first test, among quadrants, uses a chi-square test of the observed vs. expected frequency of the quadrant containing the nearest tree of the two (quarter corner) or four (corner) measured trees to test the null hypothesis that there is an equal distribution of trees among quadrants. This is expected if trees are randomly located and surveyors are unbiased. Analysis was performed on a pooled set of all tree species and all corner types. The second test, within quadrants, examines the potential selection bias of trees based on their location in a quadrant. Chi-square analysis is used to test the null hypothesis that an equal number of trees occur in each degree bin, which would be expected if trees were randomly selected and distributed. For this test, we used equal divisions of quadrants (as in Manies *et al.*, 2001) with six segments of 15° bin-widths (i.e. 0–14°, 15–29°,..., 75–89°). All trees in all quadrants were analysed together, and azimuth was rounded to the nearest degree.

Distance tests were modified, following Delcourt & Delcourt (1974), to utilize ANOVA instead of chi-square. Normality and heteroscedasticity were examined visually. Because surveyors may have had a proclivity for a particular bias, the observations from any surveyor may not be independent, which we tested for specifically as bias by surveyor. The first ANOVA tested a null hypothesis of no difference in the mean distance to trees of different species. An unequal distance could indicate that surveyors travelled further to mark certain preferred species – a species selection bias. The second ANOVA tested a null hypothesis of no difference in the mean distance to trees of different diameter classes, but of the same species. A difference in mean distance could indicate selection bias for or against certain diameter classes.

Recording errors (Table 1) were also identified on a per-corner and per-tree basis. A recording error was distinguished from measurement error by visually examining the distribution of errors. Azimuth errors >10° and distance errors >3 m appeared as outliers and were designated as recording errors. The sample size for each error type varies with the number of available corners and tree evidence. Omitted trees were evaluated on a per-corner and per-available-quadrant basis. Because some surveyors did not follow the guidelines to record a tree in each section for which the corner stands, we made a distinction between the usual and required number of trees for omission-rate calculations. For example, in the Front Range, one tree at a quarter corner and two trees at a corner were the usual numbers recorded, even though the required numbers were two and four trees, respectively. We calculated omission relative to the usual number in this case.

Selection bias and recording errors were subdivided into verified and unverified categories. The unverified category means that there was insufficient evidence that a bias or error event had occurred but sufficient evidence to include it as a possibility. For example, a large tree might be located closer to the corner than the original bearing tree, but if the tree was hollow and could not be cored, we could not prove that a bias had occurred. Verified and unverified bias and error were summed to estimate total possible bias or error rate.

Measurement bias or mean error (ME) was calculated for azimuth, distance and diameter, and a *t*-test was used to test for significant measurement bias or the null hypothesis that mean error equals zero. Because most researchers using survey data cannot distinguish recording and measurement error, all data, including recording errors, were incorporated in this analysis. If estimated mean error does not differ significantly from zero, then the surveyors' ability to estimate the parameter was considered unbiased. If mean error was either significantly less than or greater than zero, measurement bias was detected. The detection of measurement bias is important because a tendency to under- or over-estimate parameters would be propagated in calculations of forest attributes (e.g. density, basal area). Mean absolute error (MAE), a measurement of accuracy or error (Table 1), was calculated for azimuth, distance and diameter, and was used over other metrics, such as mean squared error, because it is a more robust estimator of accuracy, and is less sensitive to outliers (Walther & Moore, 2005). Relative measurement accuracy, calculated as relative mean absolute error (RMAE; Table 1) was calculated for distance and diameter. RMAE scales error to the size of each estimate, which is useful for comparing surveyors or areas where tree sizes or densities differ. Tree diameter was analysed separately because diameters were visually estimated by surveyors, rather than measured. To assess the height on the tree at which surveyor estimates were made, we compared ME, MAE and RMAE (Table 1) of reconstructed diameters against surveyor estimates at both d.s.h. and d.b.h.

Other errors

Lineage errors were not calculated for this study but are probably incorporated into recording errors and measurement errors. Species identification ambiguities were evaluated for common names in the survey field notes that referred to more than one species. A list of common names that referred to more than one species was compiled. We documented changes in the survey manuals from the 1855 manual to the 1900 manual and thoroughly examined the notes to see if surveyors altered procedures with successive editions. We did not measure the rate of fraud, as we preliminarily screened study areas and townships to avoid fraud, so that other limitations could be studied.

RESULTS

Selection bias: direct estimates

Surveyor selection bias was low in all three study areas, and there was little variability among surveyors based on our direct estimates (Table 2 and Appendix S1 in the Supporting Information). Total verified selection bias per tree ranged from 1.6% of trees on the Mogollon Plateaus to 3.9% of trees in the Front Range. The total possible bias rate per tree was not much larger, ranging from only 1.8% to 4.8%. This implies that surveyors selected the closest tree within a quadrant about 95% to 98% of the time (e.g. 100%−1.8% = 98.2%). In contrast, 4.4%, 12.0% and 6.8% of the corners on the Mogollon Plateau, in the Blue Mountains, and in the Front Range contained at least one biased tree, respectively.

For the 30 trees selected with bias (3.7% of the total of 812 trees) in our study areas, we were able to assess the total effect (i.e. error propagation effect) of selection bias on 28 trees for distance, on 23 trees for diameter, and on 30 trees for species. The median distance separating the true closest tree and the selected tree was +4.2 m (mean = 6.1 m, SE = 1.1, n = 28), a median 42% increase in distance over the true closest tree (mean = 99%, SE = 33%, n = 28). The mean diameter of biased trees was not significantly different from the mean diameter of the true closest trees (t = 0.53, d.f. = 23, P = 0.60). Biased selection resulted in seven tree species changes, or 23% of the 30 bias-selected trees. Overall, though, a change in species for seven trees out of the total of 812 trees evaluated is minimal (0.9% composition change).

The 30 trees representing selection bias were chosen for particular characteristics. Ten appear to have been chosen because of diameter size, nine because of location within a section, two because of species and nine for unresolved reasons. From the trees chosen based on diameter size, trees selected by the surveyors were on average 18.3 cm (7.2 in) (SE = 1.96, n = 7; only seven had reliable diameter estimates) larger than the closest tree in the quadrant. Trees selected

Table 2 Selection bias, recording errors, measurement bias, and measurement accuracy of General Land Office survey data on bearing trees as compared with plot re-measurements in the three study areas in the western USA.

		Mogollon Plateau	Blue Mountains	Front Range
Verified selection bias (%)		3.6, 1.6	7.2, 2.9	5.5, 3.9
Total possible selection bias (%)		4.4, 1.8	12.0, 4.8	6.8, 4.7
Verified recording & omission errors	Azimuth (%)	11.9, 7.4	11.4, 6.9	18.4, 13.7
	Distance (%)	4.7, 3.2	0.0, 0.0	6.1, 4.1
	Species (%)	0.3, 0.3	0.0, 0.0	2.0, 1.4
	Omit Trees (%)	0.1, 0.1	2.1, 1.7	17.1, 14.0
Total possible recording & omission errors	Azimuth (%)	16.3, 9.3	12.9, 7.4	18.4, 13.7
	Distance (%)	8.1, 4.9	0.0, 0.0	6.1, 4.1
	Species (%)	0.3, 0.3	0.0, 0.0	2.0, 1.4
	Omit Trees (%)	0.2, 0.2	8.7, 6.2	18.9, 15.0
Survey azimuth	ME (°)	0.99	−0.11	0.02
	SE (°)	1.34	1.02	1.0
	Biased?	t = 0.74	t = 0.11	t = 0.02
		d.f.=165	d.f.=134	d.f.=48
		P = 0.46	P = 0.92	P = 0.98
	MAE (°)	5.0	4.8	4.6
	SE (°)	0.99	0.84	0.74
Survey distance	ME (m)	−0.10	0.25	−0.01
	SE (m)	0.091	0.049	0.209
	Biased?	t = 1.10	t = 5.10	t = 0.05
		d.f. = 165	d.f. = 134	d.f. = 48
		P = 0.27	**P < 0.001	P = 0.96
	MAE (m)	0.48	0.37	0.83
	SE (m)	0.074	0.042	0.171
	RMAE (%)	4.0	4.9	6.9
	SE (%)	0.4	0.6	1.4

For bias and error rates, each study area (column) has two numbers. The numbers are the means for percentage corner and percentage tree bias. For the Mogollon Plateau and the Blue Mountains, the means are weighted based on surveyor composition, but for the Front Range, sample size precluded weighted averages, and, therefore, the numbers are unweighted sample means. ME, mean or median error; SE, standard error; MAE, mean or median absolute error; RMAE, relative mean or median absolute error. **Significant at 1% level.

against because of diameter size ranged from 16.2 cm (6.4 in) to 30.0 cm (11.8 in). Trees not chosen because of location were on average 8.6° (SE = 2.1, $n = 9$) from the quadrant boundary, with a range of 1° to 17°, whereas the trees selected were on average 29.4° (SE = 3.3, $n = 9$) from the quadrant boundary. There were two documented cases of species bias. In one, a ponderosa pine was selected over a Gambel oak (*Quercus gambelii* Nutt.) and in the other a ponderosa pine was selected over a grand fir [*Abies grandis* (Douglas ex D. Don) Lindl.]. There were other cases of oak and fir being selected over ponderosa pine, but because the pine was often larger than the other species, the type of selection bias was ambiguous. Therefore, no consistent species bias was detected.

Selection bias: Bourdo's analysis

Trees at about 20% of corners deviated from a random spatial distribution and were either uniform or aggregated based on our pattern analysis. Although only a small percentage of corners were located in non-random tree stands, a non-homogeneous random distribution (e.g. non-stationary random distributions) of trees across our sampling locations is likely, and results using Bourdo's methods should be interpreted with caution (see Grimm, 1981; and Bouldin, 2008). There was no significant difference in the proportion of quadrants that contained the nearest tree (overall nearest tree) in any study area (AZ: $\chi^2 = 3.8$, $P = 0.29$; OR: $\chi^2 = 1.5$, $P = 0.67$; CO: $\chi^2 = 2.5$, $P = 0.47$). Bearing trees on the Mogollon Plateau were evenly distributed within quadrants ($\chi^2 = 3.9$, $P = 0.56$). There was, however, an unequal distribution of bearing trees within quadrants in the other two study areas (see distributions – Fig. 2). In the Blue Mountains, there were fewer bearing trees recorded in the areas nearest the quadrant boundaries and more trees located in the centre of the quadrant ($\chi^2 = 14.4$, $P = 0.014$). In the Front Range, there was no detectable pattern to the unequal distribution ($\chi^2 = 19.5$, $P = 0.002$).

Significant differences in mean tree distance by species were identified both on the Mogollon Plateau ($F_{(1,400)} = 9.2$, $P = 0.003$) and in the Blue Mountains ($F_{(2,341)} = 11.0$, $P < 0.001$). In both cases, ponderosa pine had a significantly greater mean distance than did other tree species [Mogollon: ponderosa pine mean = 15.4 m, Gambel oak mean = 10.3 m; Blue Mountains: ponderosa pine mean = 13.1 m, fir mean = 7.7 m, western larch (*Larix occidentalis* Nutt.) mean = 7.5 m]. This analysis could not be performed in the Front Range because of insufficient trees (i.e. <10 trees) of any species other than ponderosa pine. Bias for tree size using mean distance was

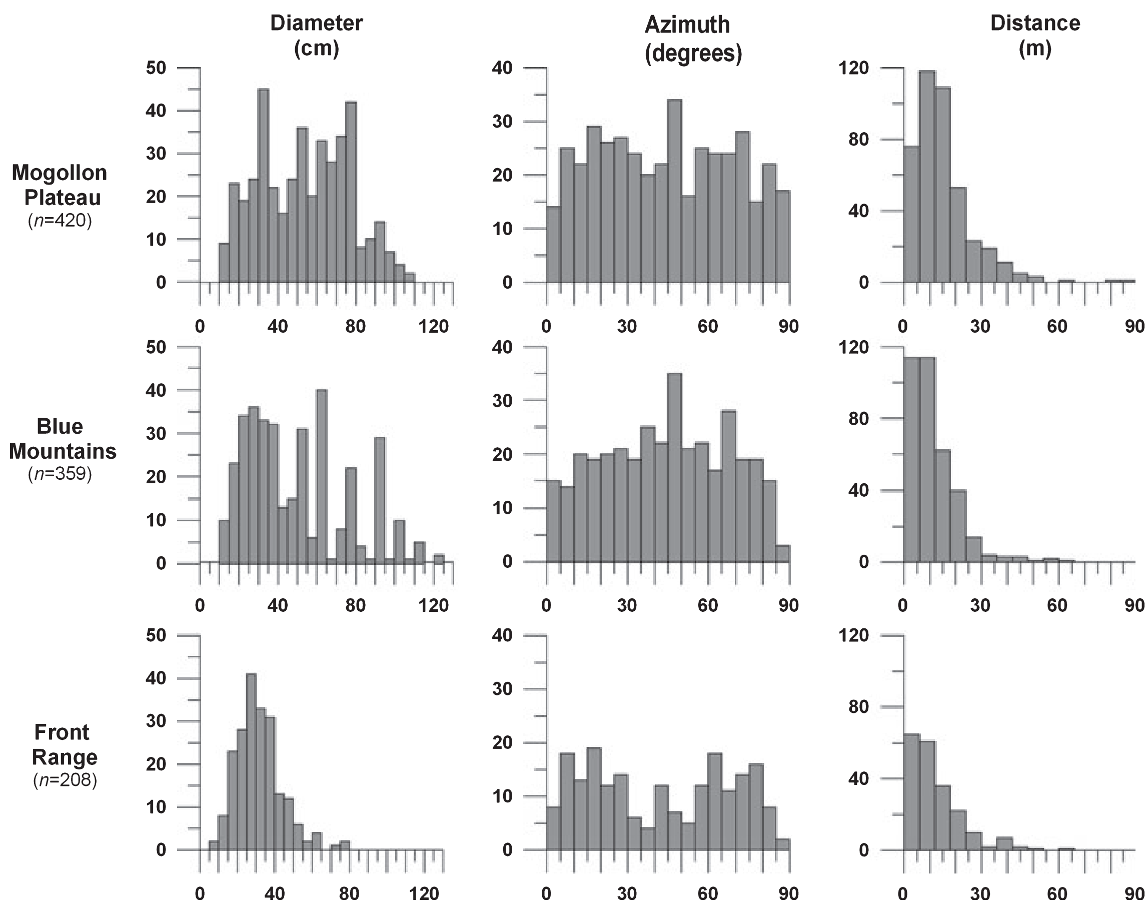

Figure 2 Distribution of historical tree data in the western USA. Data from all tree species combined. Diameters were converted from inches to centimetres, bearings were altered to range between 0 and 90° for all quadrants, and distance was converted from links to metres.

tested for ponderosa pine and Gambel oak on the Mogollon Plateau, for ponderosa pine, fir (unknown species) and western larch in the Blue Mountains, and for ponderosa pine in the Front Range. Out of six analyses, only ponderosa pine in the Front Range had a significant distance/diameter relationship ($F_{(5,152)} = 3.5$, $P = 0.005$). *Post-hoc* analysis using Tukey's honestly significant difference (HSD) test revealed that the smallest diameter class (10 cm to 20 cm) was located closer to the corner on average than the largest ($F_{(4,169)} = 5.06$, $P = 0.001$; small diameter class mean = 7.7 m, large diameter class mean = 24.7 m).

Recording and omission errors

In contrast to the paucity of selection-bias events, recording errors were common, but highly variable among surveyors (Table 2 and Appendix S1). By far the most common error was recording an incorrect bearing, which occurred for 9.3% of trees on the Mogollon Plateau, 7.4% of trees in the Blue Mountains and 13.7% of trees in the Front Range. Bearing recording error was variable among surveyors in all study areas, ranging from 0% to 26.7% of trees recorded. Many of these error values were close to the magnetic declination at the time of the surveys (see http://www.ngdc.noaa.gov/geomag/) and were likely to have been made by forgetting to add the declination to the bearing. Recording errors for distance and species were relatively rare. Only 4.9% of trees on the Mogollon Plateau, none in the Blue Mountains, and 4.1% in the Front Range had a distance error (Table 2). In each case, most of the error can be attributed to a single surveyor (Appendix S1). Only two trees (out of 812) had an incorrect species identification.

Errors from missing or omitted trees, though not uncommon (Table 2 and Appendix S1), were also largely attributable to only a few surveyors. On the Mogollon Plateau, Deputy Surveyor Secor (9.9% of quadrants surveyed had an omission) was the only surveyor who omitted trees, with a resulting 0.2% error rate per quadrant. In the Blue Mountains, Deputy Surveyor Lackland (28.1% of quadrants surveyed had an omission) was the main contributor to the overall 6.2% error rate per quadrant. Omitted tree errors were more widespread among surveyors in the Front Range, resulting in a higher 15.0% error rate per quadrant. The Front Range was different from the other study areas because the surveyors, for some unknown reason, did not consistently follow guidelines. Because we selected study areas in part to avoid widespread omission errors, we cannot estimate the true rate of omission errors in the western USA.

Measurement errors in azimuth, distance and diameter

Measurements of tree azimuth in all three study areas were statistically unbiased (i.e. ME is not significantly different from zero) overall (Table 2), and only one surveyor, Deputy Surveyor Secor, showed any significant measurement bias (Appendix S1). Azimuth accuracy was high in all three areas, and MAE only varied from 4.6° to 5.0° (Table 2). High accuracy, in the usual sense of the word, is measured here by lower errors (Table 1). Whereas the Mogollon Plateau and Front Range had unbiased tree-distance measurements (Table 2), in the Blue Mountains surveyors tended to measure distances long (ME = 0.25 m, $t = 5.10$, d.f. = 134, $P < 0.001$; Table 2). Distance accuracy varied considerably on the Mogollon Plateau among surveyors, from the best MAE of 0.3 m (1.5 links) to the worst MAE of 2.7 m (13.4 links) (Appendix S1), with an overall MAE of 0.5 m (2.5 links). There was much less variability in the Blue Mountains among surveyors, and overall MAE was 0.4 m (2 links). The Front Range showed the worst accuracy, with an MAE of 0.8 m (4 links). However, scaled by actual distance, the relative accuracy (RMAE) ranged from 4% (best) on the Mogollon Plateau to 7% (worst) in the Blue Mountains, suggesting that distance accuracy was quite high (i.e. on average 93% to 96% of the correct distance). For reference, the mean distance to a tree was 16.2 m (80.5 links) on the Mogollon Plateau, 14.2 m (70.6 links) in the Blue Mountains, and 14.9 m (74.1 links) in the Front Range.

Mean absolute error in diameter (d.s.h.) ranged from 7.0 cm (2.8 in) in the Front Range to a high of 10.6 cm (4.2 in) in the Blue Mountains, with the median absolute error in diameter (d.s.h.) generally lower, ranging from 5.9 cm (2.3 in) in the Front Range to 9.4 cm (3.7 in) on the Mogollon Plateau (Table 3). Although it appears from the MAE that surveyors in different areas estimated diameter with varying levels of accuracy, RMAE values were large, but only ranged from 24% to 27% for the reconstructed values, and, in two areas, half of the estimates were within 16%. These large RMAE values suggest that diameters might be better placed into classes. If so, confidence intervals (95%) of MAE may help guide the selection of bin width. For the Mogollon Plateau intervals ranged from 7.3 cm (2.9 in) to 10.7 cm (4.2 in); for the Blue Mountains, from 7.9 cm (3.1 in) to 13.2 cm (5.2 in); and for the Front Range, from 4.6 cm (1.8 in) to 9.4 cm (3.7 in).

Diameter estimates in all three study areas were generally made at two-inch (5.1-cm) intervals (Fig. 2). On the Mogollon Plateau, 77% of estimates were even integers; in the Blue Mountains, 86% of estimates were even; and in the Front Range, 68% of estimates were even. We did not find any trend for selection against small or large trees. There were few trees ≤10 cm (4 in) but a high percentage of large trees in most areas (Fig. 2) (e.g. on the Mogollon Plateau, 50% of the trees were >50.8 cm or 20 in).

In general, comparisons between the surveyors' estimates and the reconstructed diameters show a slightly better fit for d.s.h. than for d.b.h. MAE and RMAE values were lower for d.s.h. than for d.b.h. in all three areas, although median values were lower in only two of three areas (Table 3). Furthermore, if the surveyors were estimating diameters at d.s.h., they were unbiased in two of three areas, but if estimating at d.b.h., they were biased in two of three areas. Irrespective of height of

Table 3 Analysis of the height of diameter estimation of bearing trees by General Land Office surveyors in the three study areas in the western USA.

	Variable mean or median errors	Mean	Standard error	Median	Measurement bias?
Mogollon Plateau	ME d.s.h. (cm)	−0.7	1.5	−1.3	$t = -0.47$, $P = 0.639$
d.s.h. $n = 56$	ME d.b.h. (cm)	4.0	1.5	2.9	$t = 2.58$, $^{*}P = 0.012$
d.b.h. $n = 61$	MAE d.s.h. (cm)	9.0	0.84	9.4	
	MAE d.b.h. (cm)	9.8	1.0	6.3	
	RMAE d.s.h. (%)	27.0	2.9	23.0	
	RMAE d.b.h. (%)	43.0	7.0	19.0	
Blue Mts	ME d.s.h. (cm)	5.6	1.8	3.2	$t = 3.12$, $^{**}P < 0.01$
d.s.h. $n = 54$	ME d.b.h. (cm)	10.0	1.9	11.5	$t = 5.33$, $^{**}P < 0.01$
d.b.h. $n = 59$	MAE d.s.h. (cm)	10.6	1.3	7.5	
	MAE d.b.h. (cm)	13.9	1.4	12.1	
	RMAE d.s.h. (%)	27.0	4.1	16.0	
	RMAE d.b.h. (%)	48.0	8.1	29.0	
Front range	ME d.s.h. (cm)	−2.8	1.9	−4.7	$t = -1.52$, $P = 0.144$
$n = 21$	ME d.b.h. (cm)	2.4	1.8	3.3	$t = 1.35$, $P = 0.191$
	MAE d.s.h. (cm)	7.0	1.2	5.9	
	MAE d.b.h. (cm)	7.1	1.0	6.1	
	RMAE d.s.h. (%)	24.0	4.0	16.0	
	RMAE d.b.h. (%)	38.0	11.0	18.0	

Metrics of measurement error and accuracy for diameter estimation directed at either diameter stump height (d.s.h.) or diameter breast height (d.b.h.) from surveyors in all study areas. For all metrics the mean and median were calculated. For mean error (not tested for median), the value was compared to a null hypothesis with a mean of zero to test for significance. ME, mean or median error; MAE, mean or median absolute error; RMAE, relative mean or median absolute error. *Significant at the 5% level. **Significant at the 1% level.

measurement or study area, measurement bias was towards overestimation. To see if surveyors differed, we further examined the height of measurement for surveyors with ≥5 trees in the dataset. For those seven surveyors, four showed a better fit for estimating at d.s.h. and three for estimating at d.b.h. Although the data seem to suggest that a surveyor estimated diameter at a particular height, there is no absolute way to know if surveyors were estimating diameter at different heights on trees or if they just varied in accuracy. Knowing the height of estimation is of some importance as the mean d.s.h. exceeded the mean d.b.h. in our study areas by a range of +3.7 cm in the Front Range to +5.3 cm in the Blue Mountains. Furthermore, our reconstructed diameter has some uncertainty as it is dependent on the accuracy of the measured tree diameter, correct cross-dating, measuring of radial growth and generalized equations of tree and bark growth.

Other errors

Although species-identification ambiguities are probably low in most forests in the West, there were still ambiguous species entries. For example, 'juniper' on the Mogollon Plateau could be taken to be oneseed juniper [*Juniperus monosperma* (Engelm.) Sarg.], Rocky Mountain juniper (*Juniperus scopulorum* Sarg.), or alligator juniper (*Juniperus deppeana* Steud.).

In our study areas, surveyors appeared to be aware of changes in surveyor guidelines. For example, in the 1881 manual the number of trees at standard corners and closing corners (together referred to as double corners) was changed from two to three, and in 1894 the number of trees was reduced back to two. In our study areas surveys made prior to 1881 had two trees, surveys made in 1881 had three trees, and surveys after 1894 had only two trees at double corners. Moreover, in 1881 the suitable distance for which a surveyor was supposed to look for trees within quadrants was set at three chains (60.4 m) (White, 1983; but see circular 1864 in White). Surveys in 1881 and thereafter, report 'no trees within 3 chains' or 'no trees within limits' as opposed to 'no trees at suitable distance' or 'no trees near' when a bearing tree was not available for marking within a quadrant or section. There were also signed copies, by surveyors, of guideline changes for the Blue Mountains in the GLO notes. In contrast, there were instances where surveyors ignored the survey manuals. In the Front Range study area, few surveyors recorded the required number of trees at corners. In preliminary work leading up to the selection our study areas, we found other areas in Arizona (White Mountains) and Oregon (Ochoco Mountains) where surveyors also failed to record complete data.

DISCUSSION

Past criticism of survey data use for forest reconstructions has usually focused on the impacts of surveyor selection bias. The general argument has been that the bearing trees recorded at corners were a biased sample of the forest and were not the closest trees within each quadrant (section corner) or section (quarter corner). This argument is best stated by Grimm (1981, p. 24): 'The common assumption by ecologists that bearing trees were the closest trees to the corners is unreasonable and unwarranted.' Our data show that surveyors did record the closest tree to the corner within quadrants or sections at a rate of 95% to 98% for all corners evaluated (n = 384 corners and 812 trees). Two other studies, in Montana ponderosa pine forests, evaluated selection bias directly using a resurvey of corners, and found results similar to ours. Habeck (1994) evaluated 29 corners and found that no tree over 10.2 cm (4 in) was selected against (100% of the trees were the closest). White (1976) evaluated 37 corners and found selection bias against four trees; this is a 3.6% bias rate per tree, and thus 96.4% of the trees were the closest. Surveys prior to 1910 were undertaken through private contracts and were paid by each mile surveyed (Stewart, 1935). In more open forests, such as ponderosa pine, common wisdom suggested that surveyors would not walk further than necessary to record bearing trees (Schulte & Mladenoff, 2001), which is validated here by few cases of observed bias. The assumption by ecologists, that bearing trees were the closest trees to the corners, appears reasonable and warranted for our study areas.

Although selection bias was rare, when it did occur trees <30 cm (12 in) were most likely to be selected against, and, therefore, may be slightly under-represented in GLO data (e.g. in Fig. 2). White (1976) similarly found that the trees selected against ranged from 13 to 20 cm (5 to 8 in). We found that large trees (e.g. >80 cm) were not selected against. As hypothesized, there was no particular pattern of selection bias by species, as the number of species available at a corner was low. Surveyors avoided selecting trees near quadrant boundaries, but no particular tree characteristic appeared to be involved.

Bourdo's (1956) statistical analyses were mostly successful in detecting possible surveyor selection bias. First, a highly significant difference in mean distance to tree species was detected on the Mogollon Plateau and in the Blue Mountains, suggesting a possible species bias, but our direct data showed only two instances of species selection bias. For Bourdo's methods, differences in the mean distance to trees of different species could mean either that certain species were preferentially selected, or that certain species were generally found in denser stands (Grimm, 1981). The latter seems likely in our study areas, which is a violation of Bourdo's test assumption of equal mixing of species throughout the sampling area (Bourdo, 1954). If the user does not know whether this assumption is met, the results of this test may be misinterpreted. Second, a significant difference in mean distance to diameter class or presumed size selection bias was correctly detected by Bourdo's method in the Front Range but not correctly detected in the Blue Mountains (although results were nearly significant $F_{(8,250)} = 1.9$, $P = 0.065$). Furthermore, size selection bias was not detected on the Mogollon Plateau either through our direct method or through statistical analyses. Thus, Bourdo's methods correctly identified the presence or absence of selection bias in two of three areas. Finally, Bourdo's methods were successful in identifying a location bias within quadrants in both the Blue Mountains and Front Range and in rejecting a bias on the Mogollon Plateau. Conversely, the unequal distribution of trees within quadrants in the Front Range, which caused a rejection of the null hypothesis, could not entirely be explained by the documented location bias we found from resurveys. Selection bias was found only near quadrant boundaries in the direct methods, but Bourdo's analysis showed bias at both quadrant boundaries and in the middle of quadrants.

Bourdo's methods have been criticized as invalid because of the violation of statistical assumptions (Grimm, 1981), but the tests appear, for our data, to have identified possible selection bias correctly in most cases. This result is surprising because the assumptions of Bourdo's tests were not met in our study areas. The disadvantage of Bourdo's methods is that, although they might correctly identify the presence of selection bias, the rate and threshold of the detection of bias are unknown. We found a possible bias rate per tree of 1.8% to 4.8%, and thus surveyors chose the closest tree to the corner 95% to 98% of the time. This low bias rate, although not without associated problems, should not invalidate forest reconstructions. In short, Bourdo's analyses might not provide enough information to decide whether a particular data set has sufficient bias to invalidate results.

Recording errors, outside of species identification, are an often overlooked source of error in survey data (but see Wang, 2005). To our knowledge, only one other paper has directly documented recording errors. White (1976) documented 21 instances for which data in the survey notes did not exactly match the numbers in his remeasurements. Excluding nine cases that appear to be measurement errors, not recording errors, White's data yield a 2.7% azimuth-error rate per tree, which is much lower than our range of 7.4% to 13.7%. There are few ramifications for forest reconstructions from azimuth errors, unless a researcher is using the random pairs method (Cottam & Curtis, 1949). White (1976) also found a 2.7% distance error rate per tree, which is within the 0% to 4.9% range we found. Distance errors were relatively rare events but do have the potential to impact density calculations and have compounding impacts on other metrics, such as basal area (Bouldin, 2008).

Species-identification recording errors are rare in ponderosa pine/dry mixed conifer forests in the West. White (1976) found an error rate of <1%, similar to our range of 0% to 0.3%. The most likely reason for the low species error rates in these forests is the low number of species. In contrast, Fritschle (2007) found a 21% error rate in species identification in Redwood National Park, California, where the forest is much more diverse than other forest types in the West.

One important error with the potential to impact forest reconstructions significantly is omitted trees (Table 1). Whenever a tree is omitted, certain data are lost (e.g. species and diameter), and other data (e.g. density) must be corrected (Warde & Petranka, 1981; Dahdouh-Guebas & Koedam, 2006; Kronenfeld & Wang, 2007), albeit with unknown error. Errors of omission can substantially reduce the accuracy of reconstructed forest metrics, because calculations with corners having missing trees are adjusted (e.g. Warde & Petranka, 1981). For all available quadrants, an average of 0.2 to 15%, among study areas, contained an omitted tree. Because most of these errors were from only a few surveyors, it would be best to avoid using their surveys, if these surveyors can be identified.

Outside of recording errors, few researchers have asked how accurate the surveyors' measurements were. These have been assumed by researchers to have been either estimated by surveyors or measured using instruments. It is most likely that surveyors actually measured the bearing and distance to bearing trees using the survey instruments. Surveyors' bearing (converted to azimuth) and distance accuracies were generally high (Table 2). Scaled estimates of accuracy were quite good, with an RMAE of only 4 to 7% (93 to 96% accurate) for both azimuth and distance. Fritschle (2007) also examined surveyor accuracy, in her case using original and government resurveys. She found a lower MAE of 2.7° for azimuth and an MAE of 0.15 m (0.7 links) for distance. Our study revealed relatively accurate measurements, and only one biased measure (distance in the Blue Mountains, but with a low RMAE of 4.9%). The general absence of surveyor measurement bias means that surveyors were both under- and over-estimating parameters, and that measures should not be biased in any one direction.

None of the survey manuals that were used in our study areas specified the height at which tree diameter was to be estimated. White (1976) first suggested that surveyors were estimating diameters at stump height in the western USA. White's (1976) data (White's Table 2, p. 31) give an average error of 5.3 cm (2.1 in) when considering d.b.h. as the appropriate tree height; he states that the error for d.s.h. was lower. He also noted that the surveyors' estimates of diameter were biased (measurement bias), as 9 out of 10 were overestimates. Bourdo (1954) found that for 17 trees the actual diameter was only 1.5% less than the estimated value, although this was based on new bearing trees recently measured by government surveyors. Habeck (1994) did not assess surveyor measurement accuracy of diameter, but did core random trees at d.s.h. and d.b.h., stating that data 'confirmed' that diameter was estimated at d.s.h. These estimates of accuracy are better than the values of MAE of 7.0 cm (2.8 in) to 10.6 cm (4.2 in) and the values of RMAE of 24 to 27% we found among study areas, although estimates were unbiased in two of three areas. It appears likely that most, but not all, surveyors estimated diameter at stump height, based on data in our study and in White (1976) and Habeck (1994).

Our data also confirm that surveyors did visually estimate the diameter of trees, as they still do today (Bourdo, 1956; Grimm, 1981). We did not, however, find that surveyors selected a particular range of medium-sized trees, as suggested by other researchers (Bourdo, 1954; Grimm, 1981; but see Manies *et al.*, 2001). The relatively large RMAE for diameters (Table 3) suggests that classed data rather than raw values should be used. Because trees were most often recorded in multiples of even integers (inches), it has been suggested that diameter distributions should be analysed with 5-cm (2-in) bins. Our data suggest, however, that a 10-cm (4-in) bin would be more appropriate if higher confidence is desired, as the lowest estimated accuracy was close to 10 cm (Table 3).

Problems with changes in surveyor guidelines appear to be minimal in our study areas. Few surveying manuals were used in the West (Galatowitsch, 1990); the first surveys were initiated in the early 1850s and were mostly covered by the 1855 manual, after which there were few changes. It also appears, contrary to Grimm (1981), that the surveyors were informed of new changes in successive survey manuals. Consequently, assumptions made by researchers about survey instructions after 1855 are likely to be correct.

Although it is improbable that all sources of error can be eliminated from GLO data, there are a few ways to reduce potential problems. One way to reduce error is to screen the data for potential fraud using both the original plat maps and line description information from the survey notes. Irregularities in topography such as stream or mountain position are common in fraudulent surveys and can be compared with modern topographic maps. In some cases, a particular surveyor is recognized in subsequent resurveys to have produced fraudulent information and those surveys can be avoided. In addition to fraudulent surveys, some surveyors simply recorded fewer than the required number of trees (omission errors). These surveyors commonly record one or two fewer trees than the normal number and also often write 'no other trees within limits' in the corner description. If these occur often, particularly along forested section lines where they should be rare, the surveyor may have done poor work and might be skipped. In more meticulous surveys, actual missing trees normally occur along lines within scattered timber or near open parks. The last obvious screening technique is to check survey data for repeating diameters, azimuths or distances among trees and to see if most numbers are coarsely rounded to the nearest 5 or 10 inches, degrees or links. If the numbers are repeated throughout a township, this might imply that the surveyor fabricated data. If the surveyor appears to have severely rounded measurements, then the azimuths or distances may not have been truly measured but visually estimated or paced by foot, and the diameter not visually estimated with much rigour.

Even with a judicious screening process, some errors will always remain. Any type of selection bias by surveyors will result in decreased estimates of tree density and possibly other forest parameters, but the impacts of other errors will vary depending on specific circumstances.

Without a direct comparison of GLO survey data at PLSS corners using the original bearing trees, the rate of bias and error is unknowable (but see quadrant configuration errors in Kronenfeld & Wang, 2007). Nevertheless, all studies of direct comparison to date show that selection bias is rare and most other error rates are low. Recently, Kronenfeld & Wang (2007) developed methods to correct survey data for quadrant configuration inconsistency, bearing angle bias and species bias. Although we found that less than 1% of all trees conformed to these categories, survey data for some areas may contain anomalous amounts of bias and error. Where this is the case, we recommend that these corrections be made, if the assumptions of the correction methods are met. Furthermore, for the more common error of omitted trees, corrections are available in Warde & Petranka (1981) and Dahdouh-Guebas & Koedam (2006).

CONCLUSIONS

This study provides evidence that GLO survey data, the only data available for detailed landscape-level reconstructions of historical forest structure and disturbances, are an acceptable source for guiding forest-restoration goals. Measurements made by surveyors were generally accurate. Bearings were measured on average to within 5° of the truth, distance was measured to within 4–7% of the truth, and diameter was visually estimated to within 7–14 cm of the truth. Although rates of bias might vary with forest type, results from our three dispersed study locations suggest that the actual rate of surveyor selection bias was quite low. However, other sources of error, such as recording (0 to 14% of all trees) and omission (0 to 15% of all trees) errors, were identified by our study as limitations that could affect reconstructed forest parameters if data are not screened judiciously or corrected.

Bourdo's statistical methods correctly detected many selection biases, but did not quantify their rate. Direct analysis of resurvey data is the only effective method for quantifying bias and error rates, needed to evaluate whether detected bias and error rates are ecologically important. Indeed, surveyor selection bias, identified as a major impediment and liability for the use of survey data in the past, was not shown to be a major source of error in surveys in the western USA. In fact, data limitations posited for surveys in the eastern USA, such as species identification problems and changing survey instructions, are minor concerns in the forests we studied. Furthermore, findings from this study are likely to be applicable not only to forests in the western USA, with simpler composition and more open structure, but also to forests in the eastern USA where similar surveying guidelines were used.

ACKNOWLEDGEMENTS

This study is based on work supported by the Cooperative State Research, Education, and Extension Service, US Department of Agriculture, under Agreement No. 2007-3501-18307, and by the National Science Foundation under Grants BCS-0715070 and EPS-0447681. We appreciate reviews by Lisa Schulte, Charles Cogbill, Ken Gerow and Dan Tinker.

REFERENCES

Aldrich, M.C. (2000) *Spatial analysis of a 1900 timber inventory in northern Arizona*. MA Thesis, Northern Arizona University, Flagstaff.

Almendinger, J.C. (1997) *Minnesota's bearing tree database*. Biological Report No. 56, Minnesota Department of Natural Resources, St. Paul, MN.

Andersen, M.D. & Baker, W.L. (2006) Reconstructing landscape-scale tree invasion using survey notes in the Medicine Bow Mountains, Wyoming, USA. *Landscape Ecology*, **21,** 243–258.

Anderson, R.C. & Anderson, M.R. (1975) The presettlement vegetation of Williamson County, Illinois. *Castanea*, **40,** 345–363.

Arundel, T.R. (2000) *Using general land office survey records to determine pre-settlement forest conditions in north-central Arizona, 1878–1879*. MA Thesis, Northern Arizona University, Flagstaff.

Baker, W.L., Veblen, T.T. & Sherriff, R.L. (2007) Fire, fuels and restoration of ponderosa pine–Douglas fir forests in the Rocky Mountains, USA. *Journal of Biogeography*, **34,** 251–269.

Batek, M.J., Rebertus, A.J., Schroeder, W.A., Haithcoat, T.L., Compas, E. & Guyette, R.P. (1999) Reconstruction of early nineteenth-century vegetation and fire regimes in the Missouri Ozarks. *Journal of Biogeography*, **26,** 397–412.

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

Bourdo, E.A., Jr (1954) *A validation of methods used in analyzing original forest cover*. PhD Thesis, University of Michigan, Ann Arbor, MI.

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.

Busse, M.D., Simon, S.A. & Riegel, G.M. (2000) Tree-growth and understory responses to low-severity prescribed burning in thinned *Pinus ponderosa* forests of central Oregon. *Forest Science*, **46,** 258–268.

Canham, C.D. & Loucks, O.L. (1984) Catastrophic windthrow in the presettlement forests of Wisconsin. *Ecology*, **65,** 803–809.

Cooper, C.F. (1960) Changes in vegetation, structure, and growth of southwestern pine forests since white settlement. *Ecological Monographs*, **30,** 129–164.

Cottam, G. & Curtis, J.T. (1949) A method for making rapid surveys of woodlands by means of pairs of randomly selected trees. *Ecology*, **30,** 101–104.

Cottam, G. & Curtis, J.T. (1956) The use of distance measures in phytosociological sampling. *Ecology*, **37,** 451–460.

Dahdouh-Guebas, F. & Koedam, K. (2006) Empirical estimate of the reliability of the use of the point-centred quarter method (PCQM): solutions to ambiguous field situations and description of the PCQM+ protocol. *Forest Ecology and Management*, **228,** 1–18.

Delcourt, H.R. & Delcourt, P.A. (1974) Primeval magnolia-holly-beech climax in Louisiana. *Ecology*, **55,** 638–644.

Dyer, J.M. (2001) Using witness trees to assess forest change in southeastern Ohio. *Canadian Journal of Forest Research*, **31,** 1708–1718.

Eberhardt, L.L. (1967) Some developments in 'distance sampling'. *Biometrics*, **23,** 207–216.

Fensham, R.J. & Fairfax, R.J. (1997) The use of land survey records to reconstruct pre-European vegetation patterns in the Darling Downs, Queensland, Australia. *Journal of Biogeography*, **24,** 827–836.

Fritschle, J.A. (2007) *An intermediate-scale reconstruction of historic vegetation in Redwood National Park using the original public land survey*. PhD Thesis, University of Wisconsin, Madison, WI.

Fritschle, J.A. (2008) Reconstructing historic ecotones using the public land survey: the lost prairies of Redwood National Park. *Annals of the American Geographers*, **98,** 24–39.

Fulé, P.Z., Covington, W.W. & Moore, M.M. (1997) Determining reference conditions for ecosystem management of southwestern ponderosa pine forests. *Ecological Applications*, **7,** 895–908.

Galatowitsch, S.M. (1990) Using the original land survey notes to reconstruct presettlement landscapes in the American West. *Great Basin Naturalist*, **50,** 181–191.

Grimm, E.C. (1981) *An ecological and paleoecological study of the vegetation in the Big Woods region of Minnesota*. PhD Thesis, University of Minnesota, Minneapolis, MN.

Grimm, E.C. (1984) Fire and other factors controlling the Big Woods vegetation of Minnesota in 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.

He, H.S., Mladenoff, D.J., Sickley, T.A. & Guntenspergen, G.G. (2000) GIS interpolation of witness tree records (1839–1866) for northern Wisconsin at multiple scales. *Journal of Biogeography*, **27,** 1031–1042.

Jackson, S.M., Pinto, F. & Malcom, J.R. (2000) A comparison of pre-European settlement (1857) and current (1981–1995) forest composition in central Ontario. *Canadian Journal of Forest Research*, **30,** 605–612.

Johnson, F.A. (1956) *Use of a bark thickness–tree diameter relationship for estimating past diameters of ponderosa pine trees*. USDA Forest Service Research Note PNW-126, Pacific Northwest Forest and Range Experiment Station, Portland, OR.

Kronenfeld, B.J. & Wang, Y.-C. (2007) Accounting for surveyor inconsistency and bias in estimation of tree density from presettlement land survey records. *Canadian Journal of Forest Research*, **37,** 2365–2379.

Langley, A.N. (2004) *An assessment of the historic land survey notes and their utility in geographical studies*. MA Thesis, University of Montana, Missoula, MT.

Leahy, M.J. & Pregitzer, K.S. (2003) A comparison of presettlement and present-day forests in northeastern lower Michigan. *American Midland Naturalist*, **149,** 71–89.

Leyk, S. & Zimmerman, N.E. (2007) Improving land change detection based on uncertain survey maps using fuzzy sets. *Landscape Ecology*, **22,** 257–272.

Livermore, M. (1991) *Fictitious surveys*. P.O.B. Publishing Co., Canton, MI.

Lutz, H.J. (1930) Original forest composition in northwestern Pennsylvania as indicated by early land survey notes. *Journal of Forestry*, **28,** 1098–1103.

Manies, K.L., Mladenoff, D.J. & Nordheim, E.V. (2001) Assessing large-scale surveyor variability in the historic forest data of the original U.S. Public Land Survey. *Canadian Journal of Forest Research*, **31,** 1719–1730.

Mladenoff, D.J., Dahir, S.E., Nordheim, E.V., Schulte, L.A. & Gutenspergen, G.G. (2002) Narrowing historical uncertainty: probabilistic classification of ambiguously identified tree species in historical survey data. *Ecosystems*, **5,** 539–553.

Myers, C.A. (1963) *Estimating past diameters of ponderosa pine in Arizona and New Mexico*. USDA Forest Service Research Note RM-7, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.

Myers, C.A. & Van Deusen, J.L. (1958) *Estimating past diameters of ponderosa pine in the Black Hills*. USDA Forest Service Research Note RM-32, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.

Prayag, V.R. & Deshmukh, S.R. (2000) Testing randomness of spatial pattern using Eberhardt's index. *Environmetrics*, **11,** 571–582.

Radeloff, V.C., Mladenoff, D.J., Hong, H.S. & Boyce, M.S. (1999) Forest landscape change in northwestern Wisconsin pine barrens from pre-European settlement to present. *Canadian Journal of Forest Research*, **29,** 1649–1659.

Schulte, L.A. & Mladenoff, D.J. (2001) The original US public land survey records. *Journal of Forestry*, **99,** 5–10.

Schulte, L.A. & Mladenoff, D.J. (2005) Severe wind and fire regimes in northern forests: historical variability at the regional scale. *Ecology*, **86,** 431–445.

Spada, B. (1960) *Estimating past diameters of several species in the ponderosa pine subregion of Oregon and Washington*. USDA Forest Service Research Note PNW-181, Pacific Northwest Forest and Range Experiment Station, Portland, OR.

Stewart, L.O. (1935) *Public land surveys: history, instructions, methods*. Collegiate Press, Inc., Ames, IA.

Walther, B.A. & Moore, J.L. (2005) The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, **28,** 815–829.

Wang, Y.-C. (2005) Presettlement land survey records of vegetation: geographic characteristics, quality and modes of analysis. *Progress in Physical Geography*, **29,** 568–598.

Warde, W. & Petranka, J.W. (1981) A correction factor table for missing point-center quarter data. *Ecology*, **62,** 491–494.

White, A.S. (1976) *Evaluating land survey notes to determine the pre-settlement structure of the Lubrecht Forest, Montana*. MSc Thesis, University of Montana, MT.

White, C.A. (1983) *A history of the rectangular survey system*. USDI Bureau of Land Management, Superintendent of Documents, Washington, D.C.

Whitney, G.G. (1986) Relation of Michigan's presettlement pine forests to substrate and disturbance history. *Ecology*, **67,** 1548–1559.

Zhang, Q., Pregitzer, K.S. & Reed, D.D. (1999) Catastrophic disturbance in the presettlement forests of the Upper Peninsula of Michigan. *Canadian Journal of Forest Research*, **29**, 106–114.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Selection bias, recording errors, measurement bias, and accuracy from comparison of surveyor data and plot re-measurements for individual surveyors.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Mark Williams is a PhD candidate at the University of Wyoming. His research interests include fire and landscape ecology. His current research involves macro-scale reconstructions of historical forest data through methodological development and analysis.

William Baker is a professor at the University of Wyoming. His main interest is in landscape ecology, focusing on human and natural disturbances in southern Rocky Mountain landscapes, historical changes in landscapes, and implications for natural resource management.

Editor: Jack Williams