

The Effect of Anthropogenic Land Cover Change on Pollen-Vegetation Relationships in the American Midwest

December 24, 2015

Ellen Kujawa¹

Simon Goring²

Andria Dawson^{3,4}

Randy Calcote⁵

Eric Grimm⁶

Sara Hotchkiss⁷

Elizabeth A. Lynch⁸

Jason McLachlan⁹

Jeannine-Marie St-Jacques¹⁰

Charles Umbanhowar Jr.¹¹

Jack Williams^{2, 12}

¹ Nelson Institute for Environmental Studies, University of Wisconsin - Madison

² Department of Geography, University of Wisconsin - Madison

³ Department of Statistics, University of California, Berkeley

⁴ Department of Geosciences, University of Arizona

⁵ Limnological Research Center, University of Minnesota

⁶ Illinois State Museum, Springfield, Illinois

⁷ Department of Botany, University of Wisconsin - Madison

⁸ Biology Department, Luther College

⁹ Department of Biological Sciences, University of Notre Dame

¹⁰ Prairie Adaptation Research Collaborative, University of Regina

¹¹ Departments of Biology and Environmental Studies, St. Olaf College

¹² Center for Climatic Research, University of Wisconsin - Madison

Corresponding author: Ellen Kujawa ellen.kujawa@gmail.com

Present Address: 550 N Park St, Madison WI 53706 USA

Abstract

Fossil pollen assemblages are widely used to reconstruct past vegetation community composition at time scales ranging from centuries to millennia. These reconstructions often are based on the observed relationships between the proportions of plant taxa in the source vegetation and the proportions of the corresponding pollen types in pollen assemblages collected from surface sediments. Pollen-vegetation models rely upon parameters whose values typically are assumed to be stable through time, but this assumption is largely unevaluated, due in part to the rarity of comprehensive forest data, particularly for earlier time periods.

Here we present a new dataset of early settlement-era pollen records for the upper Midwest of North America and combine it with three other pollen and forest composition datasets to assess the stability of the relationship between relative pollen composition and relative abundances of tree genera for two time periods: immediately prior to Euro-American settlement, and the late 20th Century. Over this time interval, Euro-American settlement resulted in widespread forest clearance for agriculture and logging, producing major changes to forest composition and structure and the pollen assemblages produced by these forests. These major changes provide an opportunity to test the constancy of the relationship between pollen and forest vegetation during a period of large vegetation change. Pollen-vegetation relationships are modeled, using a Generalized Linear Model, for thirteen upper Midwestern tree genera.

We find that estimates of pollen source radius for the gridded mesoscale data are 25-85 kilometers, consistent with prior studies. Pollen-vegetation relationships are significantly altered for several genera: Fagus, Betula, Tsuga, Quercus, Pinus, and Picea ($p < 0.05$). The use of contemporary pollen-vegetation relationships to model settlement era community composition significantly under-predicts the presence of Fagus, Betula, Tsuga, Quercus and Picea at all tree densities. Pinus is over-predicted at low relative proportions ($<25\%$), but under-predicted at greater abundances. The divergence of pollen-vegetation relationships appears to be greatest for late-successional taxa characterized by high shade tolerance and low fire tolerance, although the statistical power is low for this analysis.

Hence, the ongoing rapid changes in land use and ecological communities associated with the Anthropocene affect not just our ability to make confident ecological forecasts for the future, but can also modify our inferences about the past. In the Anthropocene era, characterized by its rapidly changing vegetation and climates, paleoecology must move from its traditional reliance on spatial calibration datasets assumed to represent a single "present". Instead, when possible, paleoecologists should develop calibration datasets of pollen and forest composition that are distributed across major vegetation changes in time and space. Multitemporal calibration datasets are increasingly possible given the growing length and availability of vegetation observational data and will enable paleoecologists to better understand the complex processes governing pollen-vegetation relationships and make better-informed reconstructions of past vegetation dynamics.

Introduction

Fossil pollen records and other paleoecological proxies provide ecological insights that are otherwise temporally inaccessible, expanding knowledge of community composition, environmental change, and ecological dynamics over centuries to millennia. This understanding of long-term community dynamics can be applied to understand ecological patterns and processes in a number of ways, including reconstruction of historical ecosystems (Birks et al. 2000, Morris et al. 2014), assessing climate-driven species range expansions and contractions (Davis and Shaw 2001, Morris et al. 2014), studying the causes of abrupt change in ecological systems (Booth et al. 2012), understanding the processes that give rise to the emergence of no-analog communities (Jackson and Overpeck 2000, Williams and Jackson 2007), and determining historical baselines for restoration goals and historical patterns of disturbance (Birks 1996, Foster and Motzkin 2003).

Knowledge of the relationship between pollen composition and its source vegetation is imperative for accurate reconstruction of past vegetation composition and structure. This relationship is complicated by differences among taxa in dispersal vector, pollen productivity, and dispersibility (Cain 1939, Erdtman 1943). Relative abundance data are further complicated by the Fagerlind effect, in which the abundance of one taxon depends on the abundance of other taxa at the site (Fagerlind 1952, Davis 1963, Prentice and Webb III 1986, Davis et al. 1991).

Empirical investigations into pollen-vegetation relationships typically rely upon spatial networks of paired pollen assemblages and vegetation observations (e.g. Bradshaw and Webb III 1985, Davis et al. 1991, Calcote 1995, Parshall and Calcote 2001, Williams 2002, Williams and Jackson 2003). These studies consistently show a positive relationship between pollen and plant abundances, and indicate an effect of basin size on pollen source area, with smaller basins capturing pollen from smaller source areas than larger basins (e.g. Tauber 1965, Jacobson and Bradshaw 1981, Bradshaw and Webb 1985, Prentice 1985, Sugita 1994). Estimates of pollen source area also depend on the spatial resolution of the study (Sugita 1994) with local-scale studies tending to indicate a source area on the scale of tens of meters (e.g., 40-50% of pollen in small forest hollows coming from within 50-120 meters of the pollen source (Jackson 1990, Calcote 1995, Jackson and Kearsley 1998) Regional-scale studies with spatial resolution of 1km or larger tend to produce correspondingly larger estimates of pollen source area, on the order of 10-50 km radii or larger (Bradshaw and Webb 1985, Williams and Jackson 2003, Dawson et al. in revision).

Strategies for modeling vegetation composition from pollen data are diverse and have become increasingly sophisticated. Early methods ranged from relatively simple linear regression and ratio-based models (Davis 1963, Webb III et al. 1981, Bradshaw and Webb III 1985), to extended R-value (ERC) models, which address the Fagerlind effect (Parsons and Prentice 1981, Prentice and Webb III 1986, Prentice et al. 1987). The Prentice-Sugita family of models use dispersal kernels based on the Sutton equations (Sutton 1953) for particulate dispersal for ground-level sources and information about pollen fall speeds and productivity to estimate the pollen loadings produced by plant communities around a lake (Prentice 1985, 1988, Sugita 1994). The most recent versions, LOVE and REVEALS (Sugita 2007a, 2007b), make up the Landscape Reconstruction Algorithm (Sugita et al. 2010), in

which large lakes are used to estimate regional pollen rain and small lakes are used to infer local vegetation. The LRA has become widely used to reconstruct vegetation at local to continental spatial scales (Gaillard et al. 2010). Bayesian hierarchical models have been developed recently to model pollen-vegetation relationships while also quantifying key parameters such as pollen productivity and dispersibility and the uncertainty around these estimates (Dawson et al. in revision, Paciorek and McLachlan 2009).

Regardless of form, all pollen-vegetation models are calibrated using data for one time period, then applied to make inferences for another time period. This requires, however, assuming that the relationship between pollen and source vegetation composition is constant over time. The form of this assumption varies among models. Mechanistic dispersal-based models such as REVEALS assume that the parameter estimates for dispersal, productivity, and turbulence are constant over time, while models that rely upon spatial datasets of forest and pollen relative abundance must assume that these relationships between pollen and plant abundances are stable through time. This assumption of constancy has been previously noted (e.g. Sugita 1994, Parshall and Calcote 2001), and examined. For example, pollen-vegetation relationships vary according to forest type (Seppä 1998), alterations in the structure and heterogeneity of vegetation landscapes (Bunting et al. 2004), and pollen productivity (Meltsov et al. 2011, Baker et al. 2015).

In eastern North America, because the structure and composition of contemporary forests have been heavily altered by the land use and land cover change accompanying Euro-American settlement, paleoecologists have long sought to reconstruct forest composition and pollen-vegetation relationships just prior to European settlement, known as the pre-settlement era (e.g. Grimm 1984, Jacobson and Grimm 1986, Davis et al. 1991, Hotchkiss et al. 2007). In Minnesota, the relationships between pollen and climate have shifted significantly from the pre-settlement era to the present (St-Jacques et al. 2008a) and this can affect pollen-climate reconstructions. St-Jacques et al. (2015, 2008b) recommended that pollen-based paleoclimatic inferences should be based on pre-settlement era pollen data, rather than modern pollen-climate calibration sets, to minimize the distortion and bias produced by 19th- and 20th-century land use.

Here we build upon prior work by building novel combination of four contemporary and pre-settlement pollen and vegetation datasets, and analyzing these data to 1) quantify the relationship between the relative abundances of pollen types and tree taxa for thirteen tree genera present in the upper Midwest, 2) estimate pollen source area, and 3) assess the constancy of pollen-vegetation relationships from just at or before time of Euro-American settlement (hereafter called pre-settlement) to the present. We develop a new dataset of pre-settlement fossil pollen assemblages, drawn from the Neotoma Paleoecology Database (Grimm et al. 2013), and individual data contributors. We combine this pre-settlement pollen dataset with datasets of forest composition measured at the very beginning of the Euro-American settlement era, based on Public Land Survey System (PLSS) data (Goring et al., submitted), modern pollen data from the Whitmore et al. database (Whitmore et al. 2005), and modern vegetation composition based on the U.S. Forest Service's Forest Inventory and Analysis (FIA) plots compiled in Goring et al. (submitted). We build a series of generalized linear models (GLMs; Rigby et al. 2005) to test whether the relationship

between pollen and tree relative abundances differ among taxa (as expected, based on prior work), and between the pre-settlement era and present. Note that the pairs of pollen-vegetation datasets are not perfectly temporally aligned, e.g. the ‘modern’ pollen assemblages were collected from the 1960’s to early 2000’s while the FIA data are from 2007-2011, and the pollen data are pre-settlement while the PLS data range from pre- to early settlement. However, given the large changes in vegetation composition between the present and pre-settlement eras, the effects of these temporal mismatches should be small. To estimate pollen source area, we determine which radius for averaging vegetation composition around pollen depositional sites produces the best goodness of fit in GLMs, and compare these estimates to prior work (Calcote 1995, Williams 2002). We assess the amount of divergence between the pre-settlement era and contemporary pollen-vegetation models for individual tree genera and discuss the possible role of ecological traits such as pyrophily and shade tolerance in explaining these differences.

Euro-American Land Use and Land Cover Change in the Upper Midwest

Euro-American land use in the upper Midwest was characterized by extensive land clearance for agriculture and forest harvesting for timber products. Prior to Euro-American settlement, forests in the upper Midwestern states of Michigan, Minnesota, and Wisconsin comprised several distinct community types, most notably the hardwood and boreal forests (north and east of the prairie-forest community border) and the prairies and oak savannas (south and west of the prairie-forest community border, Curtis (1959)). Compositional ecotones among forest types were steeper than at present (Goring et al. submitted). Southern oak savanna and prairie communities were largely maintained by frequent fires (Curtis 1959, Shea et al. 2014). By 1935 the oak savanna had essentially disappeared and had been replaced by agricultural fields (51% of the overall landscape) and pasture (11%) (Rhemtulla et al. 2007). The overall effect of settlement in the south was the elimination of oak savannas, the near-elimination of prairie, and the replacement of these ecosystems by agricultural land and late-successional, shade-tolerant deciduous forests.

In northern Wisconsin, upland mesic pre-settlement era forests were composed of northern hardwood species (*Acer* spp., *Betula papyrifera* and *B. allegheniensis*, *Tilia americana*, and *Ulmus* spp.) and *Tsuga canadensis*; lowland forests were dominated by *Larix laricina* and *Thuja occidentalis* (Curtis 1959). In Wisconsin in 1850, approximately 84% of this northern region was forested; by 1935 the amount of forested land decreased to 56% (Rhemtulla et al. 2007). The logging of northern forests reduced conifer biomass and community diversity and led to their subsequent replacement by deciduous forest species (Goring et al. submitted, Rhemtulla et al. 2007). Northern logging resulted in declines of *Pinus*, while widespread slash burning and continued logging led to a decrease in fire-intolerant boreal species such as *Tsuga canadensis*, and an increase in early successional genera such as *Betula* and *Populus* (Rhemtulla et al. 2007). About 1.1% of the settlement era forests remained unlogged in Michigan, Minnesota, and Wisconsin as of 1995 (Frelich 1995). Anthropogenic fire suppression after the 1930’s also has altered

forest composition, particularly in sandy areas with high fire-return intervals (Heinselman 1973, Johnson 1976b, Friedman and Reich 2005). Currently, modern fire return intervals are an order of magnitude less frequent than fires before Euro-American settlement, in most areas (Cleland et al. 2004). However, fire return intervals may have increased in regions formerly occupied by *Tsuga* forests, where pre-settlement fire return intervals likely were >1000 years. In some regions changes have led to an increase in shade-tolerant, fire-sensitive, mostly deciduous species in regions previously dominated by fire-tolerant conifers (Drobyshev et al. 2008), and a decrease in shade-intolerant, fire-tolerant, or early-successional species in areas that previously had fire-adapted vegetation (Curtis 1959, Friedman and Reich 2005, Rhemtulla et al. 2007), however regional vegetation, and species associations vary across the region, so generalizations should be used with caution. Recent estimates suggest that 29% of modern forests in the upper Midwestern US have no historical analogue when examined at a regional (64 km²) scale (Goring et al. submitted).

Abandonment of farmland during the Great Depression (1929-1939) was especially common in the north, where the land was poorly suited to agriculture (Johnson 1976b, Gough 1997). Some of this land was naturally reforested or planted with seedlings in the by public work relief programs like the Civilian Conservation Corps (Johnson 1976b), which planted mostly *Pinus resinous* seedlings on cleared land (Ahlgren and Ahlgren 1983). Reforestation has not significantly affected total area of forested land in the upper Midwest over the post-Euro-American-settlement era, because of the concurrent loss of forests to the expansion of farmland in the south, but has affected local forest composition and structure (Rhemtulla et al. 2007).

Data and Methods

Pre-Settlement Pollen Dataset

The new pre-settlement pollen dataset presented here consists of 212 sites in Michigan, Minnesota, and Wisconsin (Figure 1a, Supplemental Table 1), and is compiled from multiple sources. Of these, 127 pre-settlement era pollen samples were obtained from the Neotoma Paleoecology Database, which is a community-curated, open-source collection of paleoecological data, (Grimm et al. 2013, <http://www.neotomadb.org>). We obtained data using the *neotoma* package for R (Goring et al. 2015), searching for all sedimentary cores within the upper Midwest that had samples within the most recent 2000 years, and particularly within the last 500 years. In addition to these sites, we included 21 pre-settlement samples from a dataset gathered by Ed Cushing, that has been previously published, expanded and used elsewhere (Jacobson and Grimm 1986, St-Jacques et al. 2008a, 2008b, 2015). To these we added pre-settlement samples that have been recently published and/or contributed by researchers. We uploaded some of these to Neotoma (Supplementary Table 1; all records with Neotoma IDs >10,000), while other pre-settlement samples are part of pollen datasets that are being prepared for submission to Neotoma (Supplementary Table 1; all records with “CLH” or “CALPRE” in the ID column).

Most pre-settlement samples obtained from the Neotoma Paleoecology Database (including previously archived records and the ones newly added here) were extracted from longer fossil pollen records from cores and stratigraphic sections. For each of these records, the pre-settlement horizon had to be identified. Fortunately, the Euro-American settlement horizon is a well established biostratigraphic feature in eastern North America, largely identified by significant increases in *Ambrosia* and other associated "weedy" taxa such as *Rumex* (McAndrews 1988). We used an expert elicitation exercise to identify the settlement horizon. Expert elicitation uses expert knowledge to make estimates from (often) noisy data (Bond et al. 2012). Here four palynologists with experience in the upper Midwest were asked to identify the uppermost pollen samples prior to the settlement horizon (Dawson et al. in revision). Sites were discarded if sampling resolution was too low, defined here as a gap of >150 years between the age of the pre-settlement pollen sample and the start of settlement period. These pre-settlement data were supplemented by 48 mostly unpublished records from across northern Wisconsin (Supplemental Table 1), contributed by Calcote, Lynch, and Hotchkiss (Hotchkiss et al. 2007, Lynch et al. 2014). For these samples, pre-settlement samples were identified as the last sample before an increase in *Ambrosia* pollen percentage above background, and the decline of white pine (due to logging) which often occurs just before the *Ambrosia* rise. These samples were generally within 100 years before 1850. All pollen counts from this pre-settlement-era dataset have been made publicly available as part of the GitHub repository for this paper http://github.com/PalEON-Project/Pollen_Vegetation-Kujawa and as supplementary material.

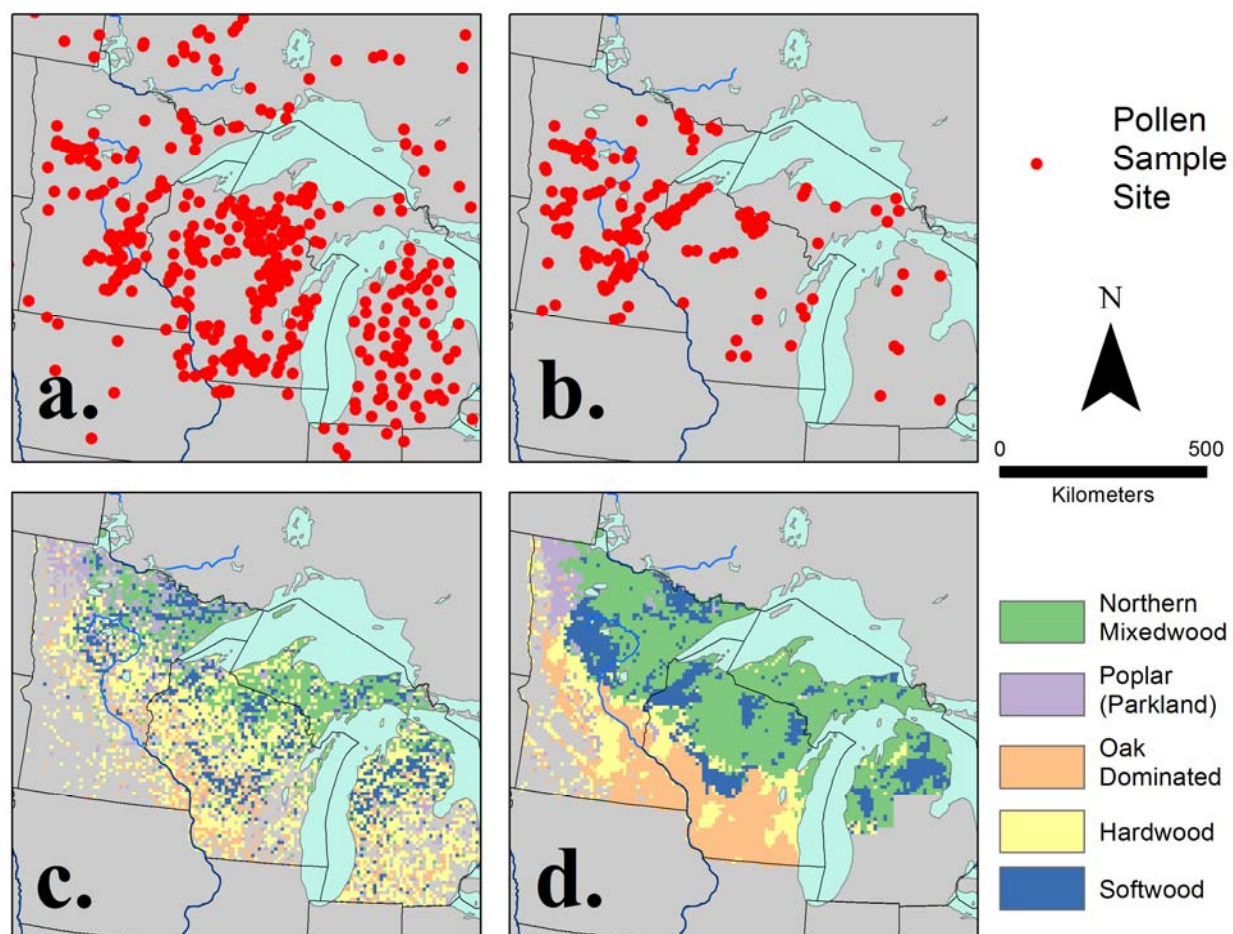


Figure 1. Map of pollen sites in Michigan, Minnesota, and Wisconsin (a, b) available for the late 20th century (a: modern era) and just prior to major Euro-American settlement and land clearance (b: pre-settlement era and modern). Forest types are classified using cluster analysis from the FIA (c) and PLSS (d) forest data (Goring et al. submitted). As discussed in Goring et al. (submitted) and elsewhere, Euro-American land use produced major shifts in forest composition and heterogeneity.

Pre-Settlement Vegetation: The Public Land Survey System

The Public Land Survey System (PLSS) began as a result of the Land Ordinance Act of 1785 and the Northwest Ordinance Act of 1787 (Carstensen 1988). Surveyors used a gridded system of six-mile-square areas, termed "townships", which were subsequently divided into 36 one-mile-square sections. At the corners and midpoints of these subplots, and at meander corners (locations where section lines intersected navigable bodies of water), surveyors recorded the species and diameter at breast height (DBH) of the nearest witness trees and their cardinal direction and distance from the point (Schulte and Mladenoff 2001). Roughly 69% of the 48 continental United States is continuously covered by this rectangle-based survey, and an additional 9% is covered intermittently (Johnson 1976a).

PLSS data represent one of the broadest and most detailed records of pre-settlement era vegetation. As such, they have been used to address a wide variety of ecological questions, and to reconstruct the community composition of pre-settlement forests (Anderson and Anderson 1975, Kline and Cottam 1979), quantify land use changes (Radeloff et al. 1999, Leahy and Pregitzer 2003), and establish restoration baselines (Fritschle 2012). Accuracy of PLSS data is affected by imprecise species identification and surveyor bias toward easily accessible, larger, or more valuable trees, particularly at fine spatial scale (Schulte and Mladenoff 2001, Fagin and Hoagland 2011, Liu et al. 2011). Nonetheless, these witness tree records provide the most complete assessment of tree composition in the western and central United States just before time of Euro-American settlement.

The data used in this analysis were sourced from a standardized dataset presented in Goring et al. (submitted). Goring et al. built on prior work, (e.g. Kronenfield and Wang 2007, Rhemtulla et al. 2009, Liu et al. 2011, Hanberry 2013) and includes novel correction factors to remedy issues of surveyor bias and sampling geometry in the original records, creating a dataset of gridded (8x8km) estimates of species composition (Figure 1d; used in our analysis), basal area, stem density and biomass. The analysis calculates stem density as a function of the distances of the two trees closest to the plot center. Corrections use a modified Morisita estimator, where stem density is calculated at the individual point, and then aggregated to the 8x8km grid cell. Estimating at the point level permits the application of point-specific corrections, to account for the fact that, for example, plot geometry varies between section and quarter-section points. The PLS data then provides regional and continuous coverage of vegetation, at, or immediately prior to major land use conversion by Euro-American settlers across Minnesota, Wisconsin, the Upper Peninsula of Michigan and the upper third of the Lower Peninsula (Goring et al. submitted).

Modern Vegetation: Forest Inventory and Analysis

The research branch of the United States Forest Service (USFS) was established in 1928 as a result of the McSweeney-McNary Forest Research (Shaw 2008, USFS 2014). The McSweeney-McNary act specifically instructed the USFS to maintain a comprehensive census of the United States' timber resources (Shaw 2008), and the USFS began Forest Surveys in 1930 (USFS 2015). The Forest Survey program has since been renamed the Forest Inventory and Analysis National Program (FIA); currently, the 48 contiguous U.S. states and parts of southern Alaska are surveyed annually (USFS 2014).

FIA data have been used in climate change research (Domke et al. 2014, MacLean et al. 2014), biodiversity conservation (Hanberry 2013, Randolph et al. 2013, Belote and Aplet 2014), and timber analysis (Domke et al. 2014). FIA data are widely applicable to ecological study because they represent a relatively complete picture of the vegetation (particularly arboreal taxa) of the United States: approximately 820 million acres are included in FIA plots and over three hundred thousand plots are subject to comprehensive vegetation surveys annually (USFS 2014).

FIA data in this study are from 2007-2011, and were prepared as in Goring et al. (submitted). Plot level data were aggregated to 8x8 km cells. The number of FIA plots per cell varies, but across the region there is a median of 3 FIA plots per cell (Goring et al.,

submitted). Within each cell, all trees with diameters > 8" were summed to generate an estimate of percent composition for that cell. This diameter cutoff was chosen for consistency with the PLSS data and estimates of minimum tree size used by Public Land Surveyors (Goring et al. submitted).

Comparisons between FIA and PLSS forest data are potentially made more difficult by differences in methods and scales between datasets. These differences are, to some degree, addressed by the common scale of aggregation (8km) used for both datasets. To further address this issue, Goring et al. (submitted) performed sensitivity analysis to understand the potential effects of differences in sampling methods and plot resolution between PLSS and FIA era data on the apparent changes to the landscape following Euro-American settlement, known as the "mixed pixel" problem (Kronenfield et al. 2010). Results showed that differences in composition are not related to the intensity of sampling within plots for the FIA or PLSS, and so, at this scale, are likely to represent true differences, rather than artifacts of differential sampling systems between the PLSS and FIA datasets (Goring et al. submitted).

Data Preparation

Data from Michigan, Minnesota, and Wisconsin were used for this analysis. Data from a set of thirteen arboreal taxa were present in all four datasets: *Abies*, *Acer*, *Betula*, *Fagus*, *Fraxinus*, *Larix*, *Tilia*, *Picea*, *Pinus*, *Populus*, *Quercus*, *Tsuga*, and *Ulmus*. We considered including a "other" category to represent other tree taxa and non-arboreal taxa, but decided that this category would be incomparable among datasets and time periods, because of differences in sampling method. To ensure consistency among all four datasets, percentages reported in all datasets represent proportions of the summed values of only these thirteen genera. The use of genera (Table 2) is necessary, given taxonomic uncertainty in both pollen and PLSS datasets. Because pollen is identified based on morphological features, and pollen morphology is strongly conserved, it is often difficult to differentiate pollen types to the species level (although it is possible, e.g. for some *Pinus* sp., *Picea*, and other taxa; Hansen and Cushing 1973, Lindbladh et al. 2002, 2003, May and Lacourse 2012, Mander and Punyasena 2014). Similarly, PLSS trees were identified to the species level by surveyors, but often recorded using common names that were locally idiosyncratic or are ambiguous, resulting in uncertainty below the genus level when trees are identified using terms such as "Pine" or "Oak". Although methods exist to differentiate PLSS data to the species level using information about local edaphic conditions (Mladenoff et al. 2002), here we opted to work with both the PLSS and pollen data at the genus level.

All datasets were projected to Albers equal-area conic projection using the R package *rgdal* (Bivand et al. 2013). For each pollen site, we calculated proportional vegetation for each of the thirteen taxa based on the average of all vegetation grid cells within a distance radius. This distance radius varied in 10 km intervals from 5 to 200 km, to encompass potential source areas found in other regional-scale pollen-vegetation studies (Bradshaw and Webb, 1985; Dawson et al. in revision).

Data Analysis

To model the relationships between pollen and vegetation and test hypotheses about the stability of pollen-vegetation relationships, we used the `gamlss` package (Rigby and Stasinopoulos 2005) to build three successively more complex generalized linear models (GLMs). The `gamlss` package is typically used to build generalized additive models (GAMs), flexible multivariate regression models that incorporate nonlinear predictive models (Hastie and Tibshirani 1986); however, `gamlss` may also be used to create GLMs, by omitting smooth splines, and contains more options for distributions than other GLM packages. We used a zero-inflated beta distribution, which is well-suited to proportional data with many zeros like ours (tree genera that are not present at any site receive a zero for that site) (Ospina and Ferrari 2010). We modeled pollen proportion as a function of vegetation proportion, with additional covariates including genus and data collection period (pre-settlement era or modern) as the predictor variables.

We tested models of increasing complexity to assess the significance of taxa and time as covariates. In Model 1, we estimate relative pollen composition as a function of vegetation, assuming that all genera and time periods behave identically. The second model, Model 2, includes taxon as a covariate of vegetation, so that the relationship is allowed to vary by taxonomic group, while the third model, Model 3, also includes an era covariate representing the time period of the vegetation dataset (either pre-settlement or the present era). Model 3 allows us to explicitly test whether there is a statistical difference in pollen-vegetation relationships for each genus between the time of Euro-American settlement and the present. We used GAIC (Generalized Akaike's Information Criterion, Rigby and Stasinopoulos 2005) to assess goodness of fit for each model and to determine the optimum radius (a proxy for pollen source area), by choosing the model and distance radius with the lowest AIC value. Vegetation is weighted using an inverse square distance weighting. Preliminary analysis indicated that this weighting provided the best fit for our models.

To quantify the magnitude and importance of the difference between pollen-vegetation models based on pre-settlement data versus models based on modern data, we calculated a divergence index for each genus, by taking the mean of absolute differences between modern and pre-settlement predictions, divided by the pollen proportion expected at the pre-settlement era. This standardization corrects for the expectation that more abundant taxa should have larger differences in their relative abundances. To assess whether shifts in disturbance history or successional stage might be affecting the observed shifts in pollen-vegetation relationships, we compared the divergence index to fire- and shade-tolerance levels for that genus (Table 1). We applied a Bonferroni correction (using the R function ``p.adjust``) to the *p*-values returned from the ANOVA comparisons.

The `ggplot2` (Wickham, 2005) and `gridExtra` (Auguie 2015) packages were used to construct figures in R (R Core Team 2014).

Results

Pollen-Vegetation Relationships

The most complex model (Model 3), which models pollen proportion as a function of vegetation proportion, taxon and era provides the best fit based on the AIC (Figure 2). For each of the three models considered, the lowest AIC (indicating best goodness-of-fit) corresponded with vegetation data that were averaged using distance radii on the order of tens of kilometers. Model 1 had the lowest AIC at a distance of 10 km, and Models 2 and 3 had the lowest AIC at a distance of 85 km respectively (Figure 2). This in turn implies that the source area for these lakes or bogs is on the order of 85 km. Spatial smoothing in the original 8km dataset and the search radius used here reduces the variability of vegetation proportions, so that few pollen sites have 0% proportions for particular tree taxa (e.g., Figure 2e), while also reducing the highest proportions for certain taxa (e.g., Figure 2b and c). At very large radii (>100 km) the estimates of vegetation revert to some mean estimate for each taxon, making a pollen-vegetation relationship difficult to recover, resulting in higher overall AICs. Model improvement in RMSE for Model 3 is approximately 25% from the single-grid-cell radius ($RMSE_{cell} = 8.9\%$) to the model employing an 85km radius ($RMSE_{85km} = 6.6\%$).

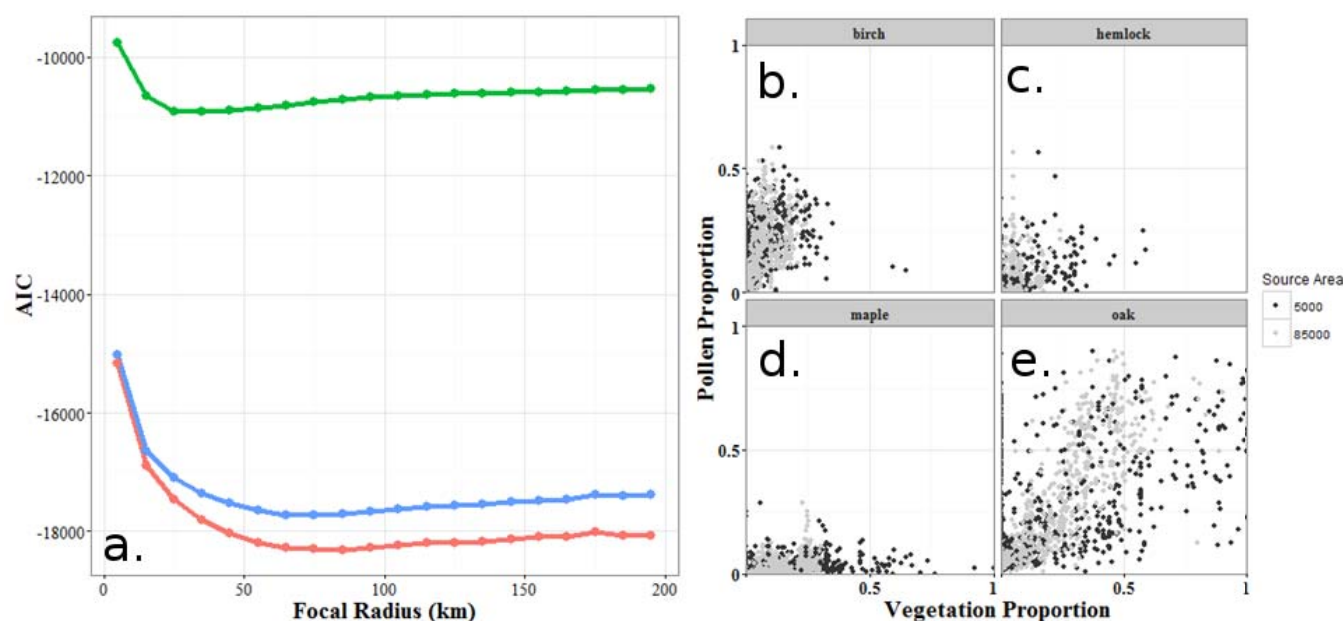


Figure 2. Aikaike's Information Criterion (AIC) values for three models at 10-km intervals from 5-200 km (panel a). Model 1 (green) has the highest goodness of fit (lowest AIC) at a distance of 10 km; models 2 (blue) and 3 (red) have the highest goodness of fit (lowest AIC) at 85 km respectively. To illustrate how the relationship between pollen and vegetation proportions is improved by increasing vegetation radius (and, by implication, a closer fit to pollen source area), panels b-e show the pollen-vegetation relationship for four selected

genera (*b*: birch, *c*: hemlock, *d*: maple, *e*: oak) in the focal grid cell (64km²) and smoothed to a radius of 85km.

Model 1, a very simple model that directly relates pollen to vegetation, regardless of genus or era, indicates that the proportion of a tree genus on a landscape is a significant predictor of its proportional composition in pollen assemblages ($p < 0.001$), confirming that pollen composition is an indicator of vegetation composition. Model 2 shows that these pollen-vegetation relationships vary widely among genera ($p \leq 0.001$), as is well-known (Davis 1963). *Betula*, *Pinus*, and *Quercus* are the most prolific pollen producers, with the steepest curves (Figure 3), while others including *Tsuga*, *Ulmus*, and *Larix* (not shown) are underrepresented genera that rarely reach high pollen proportions. The representation of *Larix* in the pollen record is complicated by its affinity for wetland environments. Hence, *Larix* is likely to be over-represented in the vegetation immediately adjacent to pollen sites in or near wetlands or bogs, but under-represented in the uplands between pollen sites. As a result, its relative proportion in the pollen record is more likely an indication of adjacent wetlands than of total landscape coverage around a sample site. For this reason, we include *Larix* in our models, but exclude it from our figures.

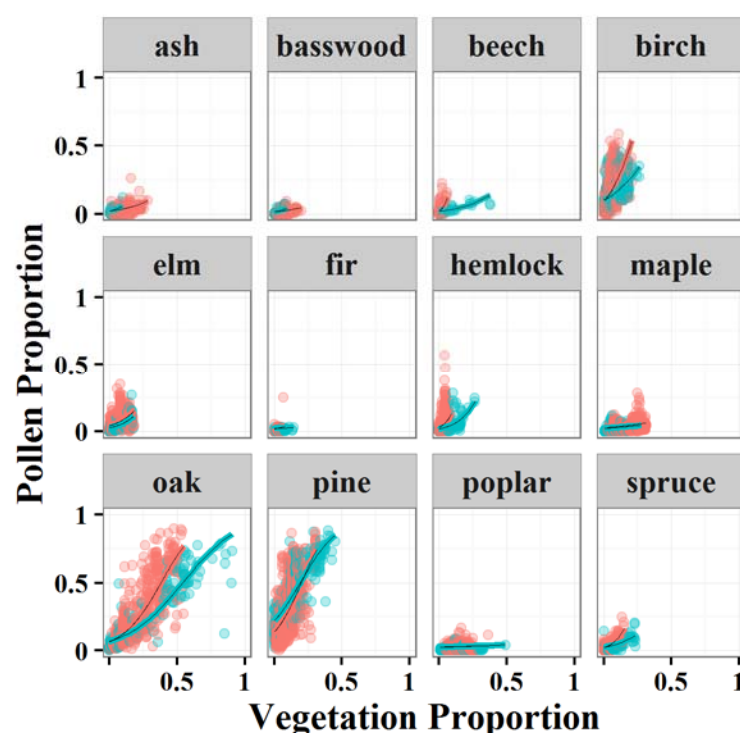


Figure 3: A scatterplot of pollen and vegetation proportions, using Model 3: pollen proportion as a function of vegetation proportion, genus, and period. Each data point corresponds to the proportion of total pollen per genus at a single pollen site at each sampling period, as a function of the proportion of total vegetation of the same genus and sampling period, averaged for all vegetation grid cells within an 85 km radius of the pollen site. Beech, birch, hemlock, oak, pine and spruce each have a statistically significant difference in the pollen-vegetation relationship for that genus between pre-settlement era (blue) and present (red).

In Model 3, the relationships between pollen and vegetation composition change significantly several genera from the pre-settlement era to the present: *Fagus*, *Betula*, *Quercus*, *Pinus*, *Tsuga*, and *Picea*. Nearly all these genera produce lower proportions of pollen relative to vegetation during the pre-settlement era than at present. The one exception is *Pinus*, which shows statistically significant different relationships for the two periods, but the two models cross, with *Pinus* over-predicted on the pre-settlement era landscape at low densities (<25%) and under-predicted at higher densities. Many of these shifts may be driven by changing vegetation proportions between pre-settlement and modern eras; for example, *Fagus* reached regional vegetation proportions of almost 50% in the pre-settlement era, but is rarely seen in proportions above 10% in modern forests. Similar declines in forested cover are seen for *Tsuga* and *Fraxinus* and, to a smaller extent, *Quercus*.

These altered pollen-vegetation relationships are also likely driven by broad-scale changes in forest composition across the region and shifts in co-occurrence between high and low pollen producers. Because pollen percentages are expressed as a relative sum, increases in the proportion of one genus can drive down the relative pollen proportions of other genera, particularly if the first genus is overrepresented in the pollen record and the other genera are underrepresented (Fagerlind 1952, Prentice and Webb III 1986). For example, the proportions of *Abies* pollen as a function of *Abies* vegetation increase strongly from the pre-settlement era to the modern, indicating that *Abies* is less under-represented in modern pollen assemblages than in the pre-settlement era. This shift in *Abies* may be caused by changes in *Pinus* vegetation proportions (a notoriously prolific pollen producer) and a weakened association between *Pinus* and *Abies* in modern forests. Correlations between *Pinus* and *Abies* vegetation decline from $r_{\text{pre-settlement}} = 0.59$ to $r_{\text{modern}} = 0.18$, indicating that rates of co-occurrence between *Abies* and the heavy pollen producer, *Pinus*, are lower in modern forests than in pre-settlement era forests.

Fagus is also under-represented in the pre-settlement era. Within the vegetation data, *Fagus* shifts from being more strongly correlated with *Quercus* ($r_{\text{pre-settlement}} = 0.27$, $r_{\text{modern}} = 0.04$) to more strongly correlated with *Betula* ($r_{\text{pre-settlement}} = 0.01$, $r_{\text{modern}} = 0.56$). Both *Fagus* and *Quercus* are heavy pollen producers, but because the proportions of *Fagus* are so limited in the modern vegetation it is difficult to know how much of this shift in relationship is driven by several outliers, or by true shifts in pollen-vegetation representivity.

Fraxinus, which is more strongly under-represented in modern samples than in pre-settlement samples, shows greater correlation to *Betula* and *Pinus* in modern forests, but also much stronger gains in correlation to weak pollen producers including *Acer* (the correlation between *Fraxinus* and *Acer* shifts from 0.17 to 0.61 between the pre-settlement era and present). Hence, its shift may be due not as much to increasing co-occurrence with heavy producers, but with weak producers.

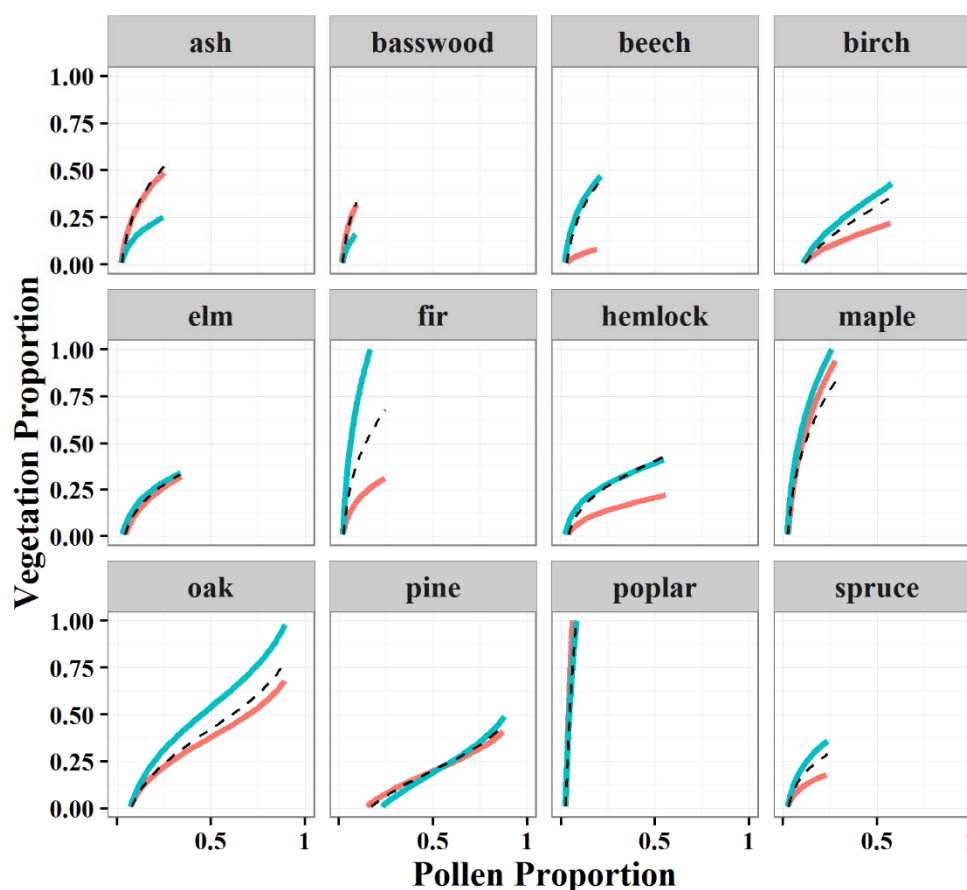


Figure 5. The inverted form of the pollen-vegetation relationship, in which vegetation proportion is expressed as a function of pollen proportion between pre-settlement (blue) and modern (red) eras using Model 3. Model 2 estimates are superimposed on the plot (black dashed line) for comparison. Where the pre-settlement (blue) line is above the red line, application of the modern pollen-vegetation relationship to fossil pollen records would lead to an underestimate of the relative abundance of that genus. The converse is true when the modern line (red) is higher than pre-settlement estimates (blue).

Table 1. An index of model divergence. The index value is the mean of absolute differences between modern and pre-settlement predictions divided by the pollen proportion expected at the pre-settlement era. Shade tolerance information is sourced from Gower (1996), with the exception of *Acer pensylvanicum* (Gabriel and Walters 1990), *Fraxinus pennsylvanica* (Kennedy 1990), *Betula nigra* (Grelen 1990), *Quercus ellipsoidalis* (Coladonato 1993), and *Quercus bicolor* (Snyder 1992). Fire tolerance information is from Thomas-Van Gundy et al. (2015). Non-native tree species are excluded from this analysis.

Genus	Divergence Index	Species names	Shade tolerance	Fire tolerance
<i>Abies</i> (fir)	1.899	<i>Abies balsamea</i> L.	Very tolerant	Pyrophobic
<i>Acer</i> (maple)	0.2275	<i>Acer saccharum</i> Marsh.	Very tolerant	Pyrophobic

		<i>A. pensylvanicum</i> L.	Very tolerant	Pyrophobic
		<i>A. nigrum</i> Michx. f.	Very tolerant	Pyrophobic
		<i>A. rubrum</i> L.	Tolerant	Pyrophobic
		<i>A. negundo</i> L.	Tolerant	Pyrophobic
		<i>A. saccharinum</i> L.	Tolerant	Pyrophobic
<i>Betula</i> (birch)	0.5533	<i>Betula papyrifera</i> Marsh.	Intolerant	Pyrophobic
		<i>B. alleghaniensis</i> Britt.	Intermediate	Pyrophobic
		<i>B. nigra</i> L.	Intolerant	Pyrophobic
<i>Fagus</i> (beech)	2.362	<i>Fagus grandifolia</i> Ehrh.	Very tolerant	Pyrophobic
<i>Fraxinus</i> (ash)	0.4275	<i>Fraxinus nigra</i> Marsh.	Intolerant	Pyrophobic
		<i>F. americana</i> L.	Intolerant	Pyrophobic
		<i>F. pennsylvanica</i> Marsh.	Intermediate	Pyrophobic
<i>Larix</i> (tamarack)	N/A	<i>Larix laricina</i> K. Koch	Very intolerant	Pyrophobic
<i>Picea</i> (spruce)	0.814	<i>Picea glauca</i> Voss	Intermediate	Pyrophobic
		<i>P. mariana</i> (Mill.) B.S.P.	Tolerant	Pyrophobic
<i>Pinus</i> (pine)	0.1166	<i>Pinus banksiana</i> Lamb.	Very intolerant	Pyrophilic
		<i>P. resinosa</i> Aiton	Intolerant	Pyrophilic
		<i>P. strobus</i> L.	Intermediate	Pyrophilic
<i>Populus</i> (poplar)	0.1166	<i>Populus deltoides</i> Bartr. ex Marsh.	Very intolerant	Pyrophilic
		<i>P. grandidentata</i> Michx.	Very intolerant	Pyrophilic
		<i>P. tremuloides</i> Michx.	Very intolerant	Pyrophilic
<i>Quercus</i> (oak)	0.4071	<i>Quercus alba</i> L.	Intermediate	Pyrophilic
		<i>Q. bicolor</i> Willd.	Intermediate	Pyrophilic
		<i>Q. macrocarpa</i> Michx.	Intermediate	Pyrophilic
		<i>Q. ellipsoidalis</i> E.J. Hill	Very intolerant	Pyrophilic
		<i>Q. rubra</i> L.	Intermediate	Pyrophilic
		<i>Q. velutina</i> Lam.	Intermediate	Pyrophilic
<i>Tilia</i> (basswood)	0.3507	<i>Tilia americana</i> L.	Tolerant	Pyrophobic
<i>Tsuga</i> (hemlock)	1.924	<i>Tsuga canadensis</i> (L.) Carr.	Tolerant	Pyrophobic
<i>Ulmus</i> (elm)	0.3416	<i>Ulmus americana</i> L.	Intermediate	Pyrophobic

<i>U. thomasi</i> Sarg.	Intermediate	Pyrophobic
<i>U. pumila</i> L.	Intermediate	Pyrophobic

The divergences between the pre-settlement and modern pollen-vegetation models varies among genera (Table 1) and are highest for *Fagus*, *Tsuga*, and *Abies* and lowest for *Populus*, *Pinus*, and *Acer*. Correlation of these divergence values to shade- and fire-tolerance characteristics is suggestive (Table 1, Figure 6) but the statistical power of this analysis is low due to the small number of taxa (12; *Larix* omitted due to known overrepresentation of pollen in wetland settings and underrepresentation elsewhere). With these caveats in mind, suggestive patterns do appear. Shade-tolerant and very tolerant taxa such as *Fagus* and *Tsuga* have stronger divergence of pollen-vegetation relationships than intolerant taxa (Table 2, $p=0.012$, corrected). We also see a possible signal of fire tolerance in the results, although without statistical significance, with pyrophilic species showing lower divergence than intolerant species ($p=0.16$). Overall, tree genera associated with mature forest show much higher rates of divergence between models than early successional taxa such as pine, poplar and maple (although successional stage varies among maple species). However, given the sample size ($n=12$), this analysis is more suggestive than conclusive.

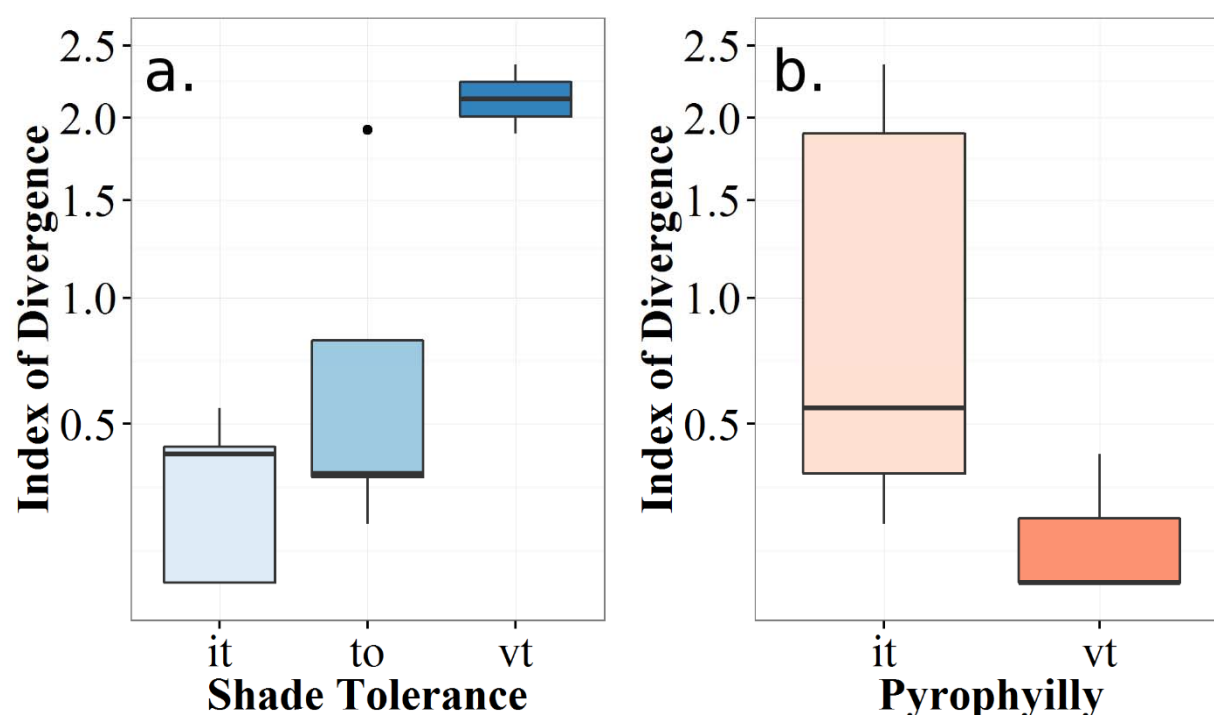


Figure 6. Index of divergence by tolerance for shade and fire. Each panel represents the mean (solid black bar), the 1st and 3rd quartiles (box limits) and the 95th percentiles for the Index values within the class. Clear trends of increasing divergence are visible for shade tolerance

(a) as tolerance increases from intolerant (it), to tolerant (to) and very tolerant (vt) species, significant at $p=0.01$ using Bonferroni correction. (b) Pyrophilly decreases divergence, indicating that species that are tolerant of fire (or that exist in fire prone landscapes; vt) have lower divergence than species that are intolerant of fire (it). These results suggest a trend, but lack significance when corrected using Bonferroni correction ($p=0.16$).

Discussion

Our work confirms our understanding that pollen and vegetation taxa compositions are highly correlated, and that these relationships differ among genera (Davis 1963, Heide and Bradshaw 1982, Prentice and Webb III 1986, Webb III 1993). Our analysis also indicates that the relationships between pollen and vegetation for some genera have shifted substantially over time. Based on these findings, reconstructions of past forest composition based on modern pollen and vegetation calibration data significantly ($p < 0.05$) under-predict the presence of some genera (*Abies*, *Betula*, *Fagus*, *Quercus*, *Tsuga*; Figure 3). These shifts in pollen-vegetation relationships are likely a response to anthropogenic land use, recent climate change, and associated changes in forest community composition.

Estimating Pollen Source Radius and Area: Comparisons to Prior Work

Our estimates of pollen source radius (Figure 2) are consistent with prior studies. Estimates of pollen source area vary by the area of the lake basin (e.g., Calcote 1995, Sugita 2007a, vegetation structure (Bunting et al. 2004) and by the spatial resolution of the study. Our estimates of pollen source area (25-85 km radius, Figure 2) are bracketed by the source area estimates developed by Sugita Sugita 2007a, 2007b for large lakes (>100 ha) and smaller lakes. Sugita (2007a) suggested that large lakes were suitable for reconstructing regional vegetation composition at scales of 104 to 105 km², (equivalent to circles with radii of 56 to 178 km). For smaller lakes, model simulations by Sugita (2007b) suggested pollen source radii on the order of 0.7 to 3.7 km.

Studies based on vegetational observation datasets with a spatial resolution of >1km and networks of moderately sized lakes (5-250 ha) located tens to hundreds of kilometers away from one other tend to find evidence for relatively large pollen source radii. Williams and Jackson 2003 found that search window half-widths of 25-75 km provided the best goodness of fit, and Bradshaw and Webb 1985 suggested that some genera accumulate significant pollen from more than 30 km away from a basin (e.g. *Pinus* and *Quercus*). Dawson et al. (in revision) estimated 50% capture radii (the maximum distance traveled by 50% of pollen sourced from a grid cell) on the order of 60 to 312 km. In contrast, modeling and empirical studies of pollen source area from forest hollows and smaller lakes tend to report smaller source areas, ranging from a few 10s of meters to a few kilometers (Calcote 1995, Jackson and Kearsley 1998, Sugita 2007b).

Implications and Causes of Changes in Pollen Vegetation Relationships

These observed changes in pollen-vegetation relationships are non-trivial. Bunting et al. (2004) show through simulation that vegetation structure on the landscape can affect the

apparent vegetation source area. Thus, the significant shift in vegetation structure and composition across the settlement boundary is likely to strongly affect pollen representivity, although here we see that it does not change the apparent source area of the optimal relationship.

The significant changes in the observed and modeled pollen-vegetation relationships between the pre-settlement era and late 20th Century (Figures 3, 4) are directly attributable to changes in vegetation composition and structure caused by human agency. Logging removed genera that were valuable timber (*Quercus*, *Pinus*), while the elimination of the natural predators of deer increased browsing and decreased the ability of some northern mesic and boreal species to establish (*Tsuga*, *Fagus*). Homogenization (Schulte et al. 2007, Rhemtulla et al. 2009) and the rise of novel forest communities (Goring et al., submitted) was also the result of human agency, resulting from regeneration following forest clearance, and much higher prevalence of pioneer species within forest communities. This homogenization also contributes to the overall "compression" of the x-axis apparent in Figure 2: i.e., each plant genus has a more reduced range of relative composition in the late 20th century than it did during the pre-settlement era. Because the compositional differences among vegetation formations is lower now (Goring et al. submitted), locations that were once dominated by one or two genera are now frequently comprised of a much more diverse genus assemblage. Note that if differences in sampling design between the FIA and PLS data were the cause, one would expect to see higher heterogeneity in the contemporary data, i.e. the opposite of the observed pattern. This is because the FIA data are sampled at individual plots while the PLS data represent 2-4 trees at points spaced ~1 mile apart, so the FIA data should be expected to be influenced more by local heterogeneity than the PLS data. Changes in vegetation heterogeneity and structure may affect pollen productivity and transport (Bunting et al. 2004), and so may be one of the causes of the changes in these pollen-vegetation relationships.

The effect of anthropogenic changes in vegetation composition and heterogeneity is magnified by the use of relative abundance data. Because the proportions of all genera at a site must sum to 1, the reduction of prolific pollen producers such as *Pinus* shifts the pollen-vegetation relationships for the remaining genera, and increases the apparent proportion of other genera on the landscape (Figure 4). Matthias and Giesecke (2014) show strong linear relationships between pollen accumulation rates and forest biomass in northeastern Germany. Much of the modern pollen data within the North American Modern Pollen Database (Whitmore et al. 2005) has no secondary information with which to estimate PAR, but the increasing ability of age-depth models to resolve uncertainty, and the more frequent use (and reporting) of exotic markers in pollen analysis mean that this may be a fruitful avenue for future research at broader regional scales.

Other factors may contribute to the apparent instability of pollen-vegetation relationships as well. Pollen productivity can vary from year to year, for example, at Arctic tree line, interannual variations in the pollen productivity of *Picea* and *Pinus* correlates to mean July temperatures of the previous year (Huusko and Hicks 2009). Other changes in forest structure and openness or meteorological conditions may also have contributed to the observed changes in pollen-vegetation relationships, by altering the dispersal of pollen grains. However, we believe that these factors are of secondary importance relative to the

widespread changes in forest composition produced by anthropogenic land use and the consequent changes in the relative proportions of overrepresented and underrepresented pollen types.

Differences in vegetation sampling design between the two eras could, in principle, cause some of these observed changes, but we view this possibility as unlikely. As noted above, the observed changes in vegetation heterogeneity are opposite to those expected given sampling design. Goring (submitted) tested for the effects of sampling design by testing whether the degree of community dissimilarity between the 8 km grid cells in the FIA and PLS datasets was influenced by the number of FIA plots per grid cell, and reported only a minor effect. Hence, we believe that the findings reported here are robust to differences in sampling design between the PLS and FIA data.

Pollen-vegetation relationships can also vary spatially. For example, *Carya*, *Quercus*, and *Ulmus* have similar pollen-vegetation relationships in both the southeast and upper Midwestern United States, but *Betula*, *Fraxinus*, *Juglans*, and *Pinus* do not (Delcourt et al. 1983). This is due in part to the different genera represented in these regions (pollen types often can only be identified to genus level; different species are present in both regions, and these species have different dispersal characteristics). Dawson et al. (in revision) report that estimates of pollen productivity and dispersal obtained by a Bayesian hierarchical model (STEPPS) fitted to the upper Midwest are in closest agreement to experimentally measured values of pollen productivity and dispersal for genera also from the upper Midwest, and do not agree well with experimental values from Europe. This geographic variability suggests a need for replication of these analyses in regions outside the upper Midwest.

Other research has begun to explore how human land use may affect the information that can be retrieved from fossil pollen record, with an emphasis on pollen-based inferences of past climates. St-Jacques et al. (2008, 2015) suggest that a pollen-climate calibration set from northern Minnesota based on pre-settlement era relationships is more accurate for reconstructing climate of the past millennium than a modern pollen-climate calibration set. Their model improvement was attributed to bias in the modern pollen record due to post-settlement increases in *Ambrosia*, *Chenopodiaceae*, and *Poaceae* and decreases in arboreal genera, particularly *Pinus* and *Quercus*. Li et al. (2014) argued that pollen-based climatic reconstructions in China for the last 1,100 years were heavily biased by human land use.

Paleoecology in the Anthropocene: New Challenges, New Opportunities

More broadly, the changing pollen-vegetation relationships documented here are part of a more general challenge of studying and understanding a rapidly changing world. Shifting baselines are a central characteristic of the Anthropocene (Jackson et al. 2001). Rates of change are accelerating, as the human imprint on ecosystems and the physical environment continues to grow (Steffen et al. 2004). Often, we use the world around us as an explicit or implicit reference point, when projecting to the future or to the past (Williams and Jackson 2007). In paleoecology and paleoclimatology, modern calibration datasets have been the backbone of quantitative research for decades (e.g. Imbrie and Kipp

1971b, Delcourt et al. 1984, Whitmore et al. 2005, Williams et al. 2006). These calibration datasets are spatially extensive but typically assumed to represent a single time period, the "present" or "modern". However, many samples used in calibration datasets were collected decades ago (Whitmore et al. 2005). Statistical transfer functions are calibrated against these modern datasets, then applied to fossil assemblages to reconstruct past climates (Viau et al. 2002), vegetation composition and structure (Williams 2002), or other variables of interest.

This work raises the possibility that at least for some taxa and regions, anthropogenic land use has significantly altered the relationship between the compositions of fossil pollen assemblages and the vegetation communities that produced them. This finding does not reduce the utility of fossil pollen data as a paleoecological and paleoenvironmental proxy, but it does indicate that pollen-based inferences about the past may be influenced by recent human activities.

A pessimistic view would note that in many areas, where the history of agriculture is longer and its spread more gradual, disentangling the role of human and climatic effects on palynological records is correspondingly more complex (Li et al. 2014, Marquer et al. 2014). Hence, in this perspective, the consequences of the Anthropocene are not just an uncertain future, caused by rapid change, but increased uncertainty about the past, as the human influence on vegetation composition and function is increasingly difficult to disentangle and vegetation and environments around us increasingly shift away from putatively stable pre-historic baselines. Even the pre-settlement era dataset developed here does not wholly remove the signal of human land use as Native Americans were of course active in the prior millennia (Muñoz et al. 2014) but reduces it by focusing on a time period prior to the onset of intensive Euro-American agriculture and timbering.

A more optimistic (and, in our opinion, better-justified) perspective is to note that we can learn much by observing and studying systems that are in flux. One key need is to develop more contemporary and historical calibration datasets from more time periods such as the pre-settlement era data presented here. Another need is to continue to collect and preserve coretop and other "modern" samples and to more precisely parse the samples in "modern" calibration datasets to reflect their actual date of sampling and thereby better link these to the environmental conditions at the time of sampling. Ultimately, we may be moving to a phase of "paleoecology in the Anthropocene" in which we explicitly recognize that even our "modern" calibration samples are being collected from a rapidly changing system. Based on this understanding, we can build calibration datasets (including both fossil samples and their associated environmental characteristics) that comprise data distributed in space and time, and across an ever-increasing range of environmental and anthropogenic conditions. By developing these calibration datasets and analytical tools we gain more power to understand the fundamental processes governing past vegetation dynamics and those governing palynological signals of these dynamics.

Conclusions

This paper presents a new dataset of pre-settlement era pollen assemblages, furthering understanding of the stability of pollen-vegetation relationships during periods of major land use and vegetation change. We document that for some genera, the relationships between relative pollen proportions and the relative proportions of the parent tree genera have changed, largely as a result of anthropogenic land use. Usage of modern pollen-vegetation relationships to reconstruct past forest composition from pollen data inaccurately predicts the relative abundance of some genera (*Fagus*, *Betula*, *Quercus*, *Pinus*, *Tsuga*, and *Picea*) on the landscape. These differences are large enough to substantively alter reconstructions of past vegetation from fossil pollen data and necessitate a need for careful consideration of pollen-based estimates of past vegetation.

These results point a new way forward, in which we can build a new generation of multi-temporal pollen-vegetation (and pollen-climate) calibration datasets with pollen samples distributed across space and time. By including samples from the pre-settlement era, the late 20th-century calibration datasets, and new surface samples, we have access to increasingly rich calibration datasets. This work supports the use of historical data in reconstructing vegetation from pollen records (Dawson et al. in revision., Jacobson and Grimm 1986, Hotchkiss et al. 2007, Paciorek and McLachlan 2009), and points to the need for more research to disentangle the effects of climate change and land use on past vegetation dynamics (Li et al. 2014). As we bring more data to bear on understanding the distributions and drivers of historical vegetation dynamics through space and time, we can achieve increasingly powerful insights into past vegetation and its relationships to climatic and anthropogenic change. This work helps us assess potential uncertainties in past reconstructions, and provides a way forward for researchers working in areas where historical data are available. In addition, it indicates the need to recognize major anthropogenic land use not only as a process that affects modern forest composition, but as something that can affect our ability to perceive the past.

Works Cited

- Ahlgren, C. E., and I. F. Ahlgren. 1983. Human impact on forest ecosystems. in S. L. Flader, editor. *The Great Lakes Forest: An Environmental and Social History*. University of Minnesota Press, Minneapolis, Minnesota.
- Anderson, R. C., and M. R. Anderson. 1975. The presettlement vegetation of Williamson County, Illinois. *Castanea* 40:345–363.
- Auguie, B. 2015. gridExtra: miscellaneous functions for “grid” graphics. R.
- Baker, A. G., M. Zimny, A. Keczyński, S. A. Bhagwat, K. J. Willis, and M. Latałowa. 2015. Pollen productivity estimates from old-growth forest strongly differ from those obtained in cultural landscapes: evidence from the Białowieża National Park, Poland. *The Holocene*

0959683615.

- Belote, R. T., and G. H. Aplet. 2014. Land protection and timber harvesting along productivity and diversity gradients in the Northern Rocky Mountains. *Ecosphere* 5:1–19.
- Birks, H. H., R. W. Battarbee, H. J. B. Birks, E. G. Bradshaw, S. J. Brooks, C. A. Duigan, V. J. Jones, G. Lemdahl, S. M. Peglar, J. O. Solem, I. W. Solhøy, T. Solhøy, and M. K. Stalsberg. 2000. The development of the aquatic ecosystem at Kråkenes Lake, western Norway, during the late glacial and early Holocene - a synthesis. *Journal of Paleolimnology* 23:91–114.
- Birks, H. J. B. 1996. Contributions of Quaternary palaeoecology to nature conservation. *Journal of Vegetation Science* 7:89–98.
- Bivand, R., T. Keitt, and B. Rowlingson. 2013. rgdal: bindings for the geospatial data abstraction library.
- Bond, C. E., R. J. Lunn, Z. K. Shipton, and A. D. Lunn. 2012. What makes an expert effective at interpreting seismic images? *Geology* 40:75–78.
- Booth, R. K., S. T. Jackson, V. A. Sousa, M. E. Sullivan, T. A. Minckley, and M. J. Clifford. 2012. Multi-decadal drought and amplified moisture variability drove rapid forest community change in a humid region. *Ecology* 93:219–226.
- Bradshaw, R. H. W., and T. Webb III. 1985. Relationships between contemporary pollen and vegetation data from Wisconsin and Michigan, USA. *Ecology* 66:721–737.
- Bunting, M. D., M.-J. Gaillard, S. Sugita, R. Middleton, and A. Broström. 2004. Vegetation structure and pollen source area. *The Holocene* 14:651–660.
- Cain, S. A. 1939. Pollen analysis as a paleo-ecological research method. *Botanical Review* 5:627–654.
- Calcote, R. 1995. Pollen source area and pollen productivity: evidence from forest hollows. *Journal of Ecology* 83:591–602.
- Carstensen, V. 1988. Patterns on the American landscape. *Publius: The Journal of Federalism* 18:31–39.
- Cleland, D. T., T. R. Crow, S. C. Saunders, D. I. Dickmann, A. L. MacLean, J. K. Jordan, R. L. Watson, A. M. Sloan, and K. D. Brosofske. 2004. Characterizing historical and modern fire regimes in Michigan (USA): a landscape ecosystem approach. *Landscape Ecology* 19:311–325.

- Coladonato, M. 1993. *Quercus ellipsoidalis*. Fire Effects Information System. United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory.
- Curtis, J. 1959. *The vegetation of Wisconsin: an ordination of plant communities*. The University of Wisconsin Press, Madison, Wisconsin.
- Davis, M. B. 1963. On the theory of pollen analysis. *American Journal of Science* 261:897–912.
- Davis, M. B., M. W. Schwartz, and K. Woods. 1991. Detecting a species limit from pollen in sediments. *Journal of Biogeography* 18:653–668.
- Davis, M. B., and R. G. Shaw. 2001. Range shifts and adaptive responses to Quaternary climate change. *Science* 292:673–679.
- Dawson, A., C. J. Paciorek, J. S. McLachlan, S. J. Goring, J. W. Williams, and S. T. Jackson. (n.d.). STEPPS2: Pollen vegetation calibration. *Quaternary Science Reviews*.
- Delcourt, P. A., H. R. Delcourt, and J. L. Davidson. 1983. Mapping and calibration of modern pollen-vegetation relationships in the southeastern United States. *Review of Palaeobotany and Palynology* 39:1–45.
- Delcourt, P. A., H. R. Delcourt, and T. I. Webb. 1984. Atlas of mapped distributions of dominance and modern pollen percentages for important tree taxa of North America. American Association of Stratigraphic Palynologists Foundation, Dallas, TX.
- Domke, G. M., C. W. Woodall, B. F. Walters, R. E. McRoberts, and M. A. Hatfield. 2014. Strategies to compensate for the effects of nonresponse on forest carbon baseline estimates from the national forest inventory of the United States. *Forest Ecology and Management* 315:112–120.
- Drobyshev, I., P. C. Goebel, D. M. Hix, R. G. I. Corace, and M. E. Semko-Duncan. 2008. Interactions among forest composition, structure, fuel loadings and fire history: a case study of red pine-dominated forests of Seney National Wildlife Refuge, Upper Michigan. *Forest Ecology and Management* 256:1723–1733.
- Erdtman, G. 1943. *An introduction to pollen analysis*. Chronica Botanica Company, Waltham, Massachusetts.
- Fagerlind, F. 1952. The real signification of pollen diagrams. *Botaniska Notiser* 1952:185–225.
- Fagin, T. D., and B. W. Hoagland. 2011. Patterns from the past: modeling Public Land Survey

- witness tree distributions with weights-of-evidence. *Plant Ecology* 212:207–217.
- Forest Inventory and Analysis national program. 2015. <http://www.fia.fs.fed.us/>.
- Foster, D. R., and G. Motzkin. 2003. Interpreting and conserving the openland habitats of coastal New England: insights from landscape history. *Forest Ecology and Management* 185:127–150.
- Frelich, L. E. 1995. Old forest in the Lake States today and before European settlement. *Natural Areas Journal* 15:157–167.
- 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.
- Fritschle, J. A. 2012. Identification of old-growth forest reference ecosystems using historic land surveys, Redwood National Park, California. *Restoration Ecology* 20:679–687.
- Gabriel, W. J., and R. S. Walters. 1990. Striped maple. in R. M. Burns and B. H. Honkala, editors. *Silvics of North America: 1. Conifers, 2. Hardwoods*. United States Department of Agriculture, Forest Service, Washington D.C., United States of America.
- Gaillard, M.-J., S. Sugita, F. Mazier, A. K. Trondman, A. Broström, T. Hickler, J. O. Kaplan, E. Kjellström, U. Kokfelt, P. Kuneš, C. Lemmen, P. Miller, J. Olofsson, A. Poska, M. Rundgren, B. Smith, G. Strandberg, R. Fyfe, A. B. Nielsen, T. Alenius, L. Balakauskas, L. Barnekow, H. J. B. Birks, A. Bjune, L. Björkman, T. Giesecke, K. Hjelle, L. Kalnina, M. Kangur, W. O. Van Der Knaap, T. Koff, P. Lageras, M. Latałowa, M. Leydet, J. Lechterbeck, M. Lindbladh, B. Odgaard, S. Peglar, U. Segerström, H. Von Stedingk, and H. Seppä. 2010. Holocene land-cover reconstructions for studies on land cover-climate feedbacks. *Climate of the Past* 6:483–499.
- Goring, S. J., A. Dawson, G. Simpson, K. Ram, R. W. Graham, E. C. Grimm, and J. W. Williams. 2015. A programmatic interface to the Neotoma Paleoecology Database. *Open Quaternary* 1.
- Goring, S. J., J. W. Williams, D. J. Mladenoff, C. V. Cogbill, S. Record, C. J. Paciorek, S. T. Jackson, M. C. Dietze, and J. S. McLachlan. (n.d.). Changes in forest composition, stem density, and biomass from settlement era to present in the upper Midwestern United States. *PLoS One*.
- Gough, R. J. 1997. *Farming the cutover: a social history of Northern Wisconsin 1900-1940*. University Press of Kansas, Lawrence, Kansas.

- Grelen, H. S. 1990. River Birch. in R. M. Burns and B. H. Honkala, editors. *Silvics of North America: 1. Conifers, 2. Hardwoods*. United States Department of Agriculture, Forest Service, Washington D.C., United States of America.
- 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.
- Grimm, E. C., R. H. W. Bradshaw, S. Brewer, S. Flantua, T. Glesecke, A.-M. Lezine, H. Takahara, and J. W. Williams. 2013. *Databases and their application*. Elsevier B.V.
- Hanberry, B. B. 2013. Changing eastern broadleaf, southern mixed, and northern mixed forest ecosystems of the eastern United States. *Forest Ecology and Management* 306:171–178.
- Hansen, B. S., and E. J. Cushing. 1973. Identification of pine pollen of late Quaternary age from the Chuska Mountains, New Mexico. *Geological Society of America Bulletin* 84:1181–1199.
- Hastie, T., and R. Tibshirani. 1986. Generalized additive models. *Statistical Science* 1:297–310.
- Heide, K. M., and R. H. W. Bradshaw. 1982. The pollen-tree relationship within forests of Wisconsin and Upper Michigan, U.S.A. *Review of Palaeobotany and Palynology* 36:1–23.
- Heinselman, M. 1973. Fire in the virgin forests of the Boundary Waters Canoe Area, Minnesota. *Quaternary Research* 3:329 – 382.
- Hotchkiss, S. C., R. Calcote, and E. A. Lynch. 2007. Response of vegetation and fire to Little Ice Age climate change: regional continuity and landscape heterogeneity. *Landscape Ecology* 22:24–41.
- Huusko, A., and S. Hicks. 2009. Conifer pollen abundance provides a proxy for summer temperature: evidence from the latitudinal forest limit in Finland. *Journal of Quaternary Science* 24:522–528.
- Imbrie, J., and N. G. Kipp. 1971. A new micropaleontological method for quantitative paleoclimatology: application to a late Pleistocene Caribbean core. Pages 71–181 in K. Turekian, editor. *The Late Cenozoic Glacial Ages*. Yale University Press, New Haven, Connecticut.
- Jackson, J. B. C., M. X. Kirby, W. H. Berger, K. A. Bjorndal, L. W. Botsford, B. J. Bourque, R. H. Bradbury, R. Cooke, J. Erlandson, J. A. Estes, T. P. Hughes, S. Kidwell, C. B. Lange, H. S. Lenihan, J. M. Pandolfi, C. H. Peterson, R. S. Steneck, M. J. Tegner, and R. R. Warner.

2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293:629–637.
- Jackson, S. T. 1990. Pollen source area and representation in small lakes of the northeastern United States. *Review of Palaeobotany and Palynology* 63:53–76.
- Jackson, S. T., and J. B. Kearsley. 1998. Quantitative representation of local forest composition in forest-floor pollen assemblages. *Journal of Ecology* 86:474–490.
- Jackson, S. T., and J. T. Overpeck. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* 26:194–220.
- Jacobson, G. L. J., and R. H. W. Bradshaw. 1981. The selection of sites for paleovegetational studies. *Quaternary Research* 16:80–96.
- Jacobson, G. L. J., and E. C. Grimm. 1986. A numerical analysis of Holocene forest and prairie vegetation in central Minnesota. *Ecology* 67:958–966.
- Johnson, H. B. 1976a. *Order upon the land: the U.S. rectangular land survey and the Upper Mississippi Country*. Oxford University Press, New York, New York.
- Johnson, J. E. 1976b. The Lake States region. Pages 81–128 in J. W. Barrett, editor. *Regional Silviculture of the United States*. John Wiley and Sons, New York, New York.
- Kennedy, H. E. 1990. Green ash. in R. M. Burns and B. H. Honkala, editors. *Silvics of North America: 1. Conifers, 2. Hardwoods*. United States Department of Agriculture, Forest Service, Washington D.C., United States of America.
- Kline, V. M., and G. Cottam. 1979. Vegetation response to climate and fire in the Driftless Area of Wisconsin. *Ecology* 60:861–868.
- Kronenfield, B. J., and Y. Wang. 2007. Accounting for surveyor inconsistency and bias in estimation of tree density from presettlement land survey records. *Canadian Journal of Forest Research* 37:2365–2379.
- Kronenfield, B. J., Y. Wang, and C. P. S. Larsen. 2010. The influence of the “mixed pixel” problem on the detection of analogous forest communities between presettlement and present in western New York. *The Professional Geographer* 62.
- Leahy, M. J., and K. S. Pregitzer. 2003. A comparison of presettlement and present-day forests in Northeastern Lower Michigan. *The American Midland Naturalist* 149:71–89.
- Li, J., Y. Zhao, Q. Xu, Z. Zheng, H. Lu, Y. Luo, Y. Li, C. Li, and H. Seppä. 2014. Human influence as a potential source of bias in pollen-based quantitative climate reconstructions.

Quaternary Science Reviews 99:112–121.

- Lindbladh, M., G. L. J. Jacobson, and M. Schauffler. 2003. The postglacial history of three *Picea* species in New England, USA. *Quaternary Research* 59:61–69.
- Lindbladh, M., R. O'Connor, and G. L. J. Jacobson. 2002. Morphometric analysis of pollen grains for paleoecological studies: classification of *Picea* from eastern North America. *American Journal of Botany* 89:1459–1467.
- Liu, F., D. J. Mladenoff, N. S. Keuler, and L. S. Moore. 2011. Broudscale variability in tree data of the historical Public Land Survey and its consequences for ecological studies. *Ecological Monographs* 81:259–275.
- Lynch, E. A., R. Calcote, S. C. Hotchkiss, and M. Tveiten. 2014. Presence of lakes and wetlands decreases resilience of jack pine ecosystems to late-Holocene climatic changes. *Canadian Journal of Forest Research* 44:1331–1343.
- MacLean, R. G., M. J. Ducey, and C. M. Hoover. 2014. A comparison of carbon stock estimates and projections for the Northeastern United States. *Forest Science* 60:206–213.
- Mander, L., and S. W. Punyasena. 2014. On the taxonomic resolution of pollen and spore records of Earth's vegetation. *International Journal of Plant Sciences* 175:931–945.
- Marquer, L., M.-J. Gaillard, S. Sugita, A.-K. Trondman, F. Mazier, A. B. Nielsen, R. Fyfe, B. V. Odgaard, T. Alenius, and H. J. B. Birks. 2014. Holocene changes in vegetation composition in northern Europe: why quantitative pollen-based vegetation reconstructions matter. *Quaternary Science Reviews* 90:199–216.
- Matthias, I., and T. Giesecke. 2014. Insights into pollen source area, transport and deposition from modern pollen accumulation rates in lake sediments. *Quaternary Science Reviews* 87:12–23.
- May, L., and T. Lacourse. 2012. Morphological differentiation of *Alnus* (alder) pollen from western North America. *Review of Palaeobotany and Palynology* 180:15–24.
- McAndrews, J. 1988. Human disturbance of North American forests and grasslands: The fossil pollen record. Pages 673–697 in B. Huntley and T. I. Webb, editors. *Vegetation History*. Springer Netherlands.
- Meltsov, V., A. Poska, B. V. Odgaard, M. Sammul, and T. Kull. 2011. Palynological richness and pollen sample evenness in relation to local floristic diversity in southern Estonia. *Review of Palaeobotany and Palynology* 166:344–351.
- Mladenoff, D. J., S. E. Dahir, E. V. Nordheim, L. A. Schulte, and G. G. Guntenspergen. 2002.

- Narrowing historical uncertainty: probabilistic classification of ambiguously identified tree species in historical forest survey data. *Ecosystems* 5:539–553.
- Morris, J. L., J. R. Mueller, A. Nurse, C. J. Long, and K. K. McLauchlan. 2014. Holocene fire regimes, vegetation and biogeochemistry of an ecotone site in the Great Lakes Region of North America. *Journal of Vegetation Science* 25:1450–1464.
- Muñoz, S. E., S. Schroeder, D. A. Fike, and J. W. Williams. 2014. A record of sustained prehistoric and historic land use from the Cahokia region, Illinois, USA. *Geology* 42:499–502.
- Ospina, R., and S. L. P. Ferrari. 2010. Inflated beta distributions. *Statistical Papers* 23:111–126.
- Paciorek, C. J., and J. S. McLachlan. 2009. Mapping ancient forests: Bayesian inference for spatio-temporal trends in forest composition using the fossil pollen proxy record. *Journal of American Statistical Association* 104:608–622.
- Parshall, T., and R. Calcote. 2001. Effect of pollen from regional vegetation on stand-scale forest reconstruction. *The Holocene* 11:81–87.
- Parsons, R. W., and I. C. Prentice. 1981. Statistical approaches to R-values and the pollen-vegetation relationships. *Review of Palaeobotany and Palynology* 32:127–152.
- Prentice, I. C. 1985. Pollen representation, source area, and basin size: toward a unified theory of pollen. *Quaternary Research* 23:76–86.
- Prentice, I. C. 1988. Records of vegetation in time and space: the principles of pollen analysis. Pages 17–42 in B. Huntley and T. I. Webb, editors. *Vegetation History*. Kluwer Academic Publishers, Dordrecht.
- Prentice, I. C., B. E. Berglund, and T. Olsson. 1987. Quantitative forest-composition sensing characteristics of pollen samples from Swedish lakes. *Boreas* 16:43–54.
- Prentice, I. C., and T. Webb III. 1986. Pollen percentages, tree abundances and the Fagerlind effect. *Journal of Quaternary Science* 1:35–43.
- R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Radeloff, V. C., D. J. Mladenoff, H. S. He, and M. S. Boyce. 1999. Forest landscape change in the northwestern Wisconsin Pine Barrens from pre-European settlement to the present. *Canadian Journal of Forest Research* 29:1649–1659.

- Randolph, K. C., A. K. Rose, C. M. Oswalt, and M. J. Brown. 2013. Status of black walnut (*Juglans nigra* L.) in the Eastern United States in light of the discovery of thousand cankers disease. *Castanea* 78:2–14.
- Rhemtulla, J. M., D. J. Mladenoff, and M. K. Clayton. 2007. Regional land-cover conversion in the U.S. upper Midwest: magnitude of change and limited recovery (1850-1935-1993). *Landscape Ecology* 22:57–75.
- Rhemtulla, J. M., D. J. Mladenoff, and M. K. Clayton. 2009. Legacies of historical land use on regional forest composition and structure in Wisconsin, USA (mid-1800s-1930s-2000s). *Ecological Applications* 19:1061–78.
- Rigby, R. A., and D. M. Stasinopoulos. 2005. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society. Series C: Applied Statistics* 54:507–554.
- Schulte, L. A., and D. J. Mladenoff. 2001. The original US Public Land Survey records: Their use and limitations in reconstructing presettlement vegetation. *Journal of Forestry*:5–10.
- Schulte, L. A., D. J. Mladenoff, T. R. Crow, L. C. Merrick, and D. T. Cleland. 2007. Homogenization of northern U.S. Great Lakes forests due to land use. *Landscape Ecology* 22:1089–1103.
- Seppä, H. 1998. Postglacial trends in palynological richness in the northern Fennoscandian tree-line area and their ecological interpretation. *The Holocene* 8:43–53.
- Shaw, J. D. 2008. Benefits of a strategic national forest inventory to science and society: the USDA Forest Service Forest Inventory and Analysis program. *iForest - Biogeosciences and Forestry* 1:81–85.
- Shea, M. E., L. A. Schulte, and B. J. Palik. 2014. Reconstructing vegetation post: pre-Euro-American vegetation for the Midwest Driftless Area, USA. *Ecological Restoration* 32:417–433.
- Snyder, S. A. 1992. *Quercus bicolor*. Fire Effects Information System. United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory.
- St-Jacques, J.-M., B. F. Cumming, D. J. Sauchyn, and J. P. Smol. 2015. The bias and signal attenuation present in conventional pollen-based climate reconstructions as assessed by early climate data from Minnesota, USA. *Plos One* 10:e0113806.

- St-Jacques, J.-M., B. F. Cumming, and J. P. Smol. 2008a. A pre-European settlement pollen-climate calibration set for Minnesota, USA: developing tools for palaeoclimatic reconstructions. *Journal of Biogeography* 35:306–324.
- St-Jacques, J.-M., B. F. Cumming, and J. P. Smol. 2008b. A 900-year pollen-inferred temperature and effective moisture record from varved Lake Mina, in west-central Minnesota, USA. *Quaternary Science Reviews* 27:781–796.
- Steffen, W., R. A. Sanderson, P. D. Tyson, J. Jäger, P. A. Matson, B. Moore III, F. Oldfield, K. Richardson, H.-J. Schellnhuber, B. L. Turner, and R. J. Wasson. 2004. *Global Change and the Earth System*. Springer, Berlin, Germany.
- Sugita, S. 1994. Pollen representation of vegetation in Quaternary sediments: theory and method in patchy vegetation. *Journal of Ecology* 82:881–897.
- Sugita, S. 2007a. Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. *The Holocene* 17:229–241.
- Sugita, S. 2007b. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. *The Holocene* 17:243–257.
- Sugita, S., T. Parshall, R. Calcote, and K. Walker. 2010. Testing the Landscape Reconstruction Algorithm for spatially explicit reconstruction of vegetation in northern Michigan and Wisconsin. *Quaternary Research* 74:289–300.
- Sutton, O. 1953. *Micrometeorology: A Study of Physical Processes in the Lowest Layers of the Earth's Atmosphere*. McGraw-Hill, New York, New York.
- Tauber, H. 1965. Differential pollen dispersion and the interpretation of pollen diagrams, with a contribution to the interpretation of the elm fall. *Danmarks Geologiske Undersøgelse* 89:69 pages.
- Thomas-Van Gundy, M. A., G. J. Nowacki, and C. V. Cogbill. 2015. Mapping the pyrophilic percentages across the northeastern United States using witness trees, with focus on four national forests.
- USDA Forest Service. 2014. Forest inventory and analysis data: fiscal year 2014 business report.
- Viau, A. E., K. Gajewski, P. Fines, D. E. Atkinson, and M. C. Sawada. 2002. Widespread evidence of 1500 yr climate variability in North America during the 14000 yr Widespread evidence of 1500 yr climate variability in North America during the past 14000 yr. *Geology* 30:455–458.

- Webb III, T. 1993. Constructing the past from late-quaternary pollen data: temporal resolution and a zoom lens space-time perspective. *Taphonomic Approaches to Time Resolution in Fossil Assemblages* 6:79–101.
- Webb III, T., S. E. Howe, R. H. W. Bradshaw, and K. M. Heide. 1981. Estimating plant abundances from pollen percentages: the use of regression analysis. *Review of Palaeobotany and Palynology* 34:269–300.
- Whitmore, J., K. Gajewski, M. Sawada, J. W. Williams, B. Shuman, P. J. Bartlein, T. Minckley, A. E. Viau, T. Webb III, S. Shafer, P. Anderson, and L. Brubaker. 2005. Modern pollen data from North America and Greenland for multi-scale paleoenvironmental applications. *Quaternary Science Reviews* 24:1828–1848.
- Williams, J. W. 2002. Variations in tree cover in North America since the last glacial maximum. *Global and Planetary Change* 35:1–23.
- Williams, J. W., and S. T. Jackson. 2003. Palynological and AVHRR observations of modern vegetational gradients in eastern North America. *The Holocene* 13:485–497.
- Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* 5:475–482.
- Williams, J. W., B. Shuman, P. J. Bartlein, J. Whitmore, K. Gajewski, M. Sawada, T. Minckley, S. Shafer, A. E. Viau, I. Webb, T. P. M. Anderson, L. Brubaker, C. Whitlock, and O. K. Davis. 2006. An Atlas of Pollen-Vegetation-Climate Relationships for the United States and Canada. American Association of Stratigraphic Palynologists Foundation, Dallas, TX.

Supplementary Table 1. *Pollen sample data for samples identified as "pre-settlement" for this study. Table includes the site name, location and PI. The ID number indicates the dataset number within the Neotoma Paleoecological Database; ID numbers that begin with CLH indicate that the sample is from the Calcote dataset, and ID numbers that begin with CALPRE come from the Calcote dataset with only a modern and settlement sample .*

SITE NAME	LOCATION	PRINCIPAL INVESTIGATOR	SAMPLE DEPTH	NEOTOMA ID
Big John Pond	47.558°N, 94.967°W	Almendinger, John C.	40	272
Hostage Lake	46.55°N, 94.133°W	Almendinger, John C.	40	1111
Mud Lake	46.867°N, 94.75°W	Almendinger, John C.	50	1778
Peterson Slough	46.967°N, 95.317°W	Almendinger, John C.	100	1891
Wentzel's Pond	46.95°N, 94.95°W	Almendinger, John C.	55	2948
Reidel Lake	46.212°N, 95.284°W	Almquist-Jacobson, Heather	496	2053
Upper Graven Lake	46.184°N, 95.307°W	Almquist-Jacobson, Heather	160	2888
Last Lake	45.452°N, 95.244°W	Alwin, Baird C.	40	12004
Wintergreen Lake	42.4°N, 85.383°W	Bailey, Robert E.	68	2961
Disterhaft Farm Bog	43.917°N, 89.167°W	Baker, Richard G.	80	3081
Lake 27	45.067°N, 84.783°W	Bernabo, J. Christopher	20	1581
Irwin Smith Bog	45.035°N, 83.64°W	Booth, Robert K.	16.5	13047
Shagawa Lake	47.917°N, 91.867°W	Bradbury, J. Platt	40	2375
Camp 11 Lake	46.667°N, 88.017°W	Brubaker, Linda B.	10	332

Lost Lake	46.72°N, 87.97°W	Brubaker, Linda B.	10	1665
Adventure	46.09°N, 91.94°W	Calcote, Randy	NA	CALPRE1
Crestview	46.274°N, 91.693°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 0
DeepPine	46.543°N, 91.35°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 1
DeerPrint	46.36°N, 91.619°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 2
FirstPine	46.651°N, 91.236°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 3
Genos	45.928°N, 92.083°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 4
Horseshoe	45.862°N, 92.279°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 5
Jakes	46.568°N, 91.3°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE1 6
LittleRound	46.013°N, 92.226°W	Calcote, Randy	NA	CALPRE1

		Lynch, Elizabeth		7
		Hotchkiss, Sara		
Menke	45.872°N, 92.341°W	Calcote, Randy	NA	CALPRE1
		Lynch, Elizabeth		8
		Hotchkiss, Sara		
Metz01	46.318°N, 91.688°W	Calcote, Randy	NA	CALPRE1
		Lynch, Elizabeth		9
		Hotchkiss, Sara		
Benach	45.927°N, 92.256°W	Calcote, Randy	NA	CALPRE2
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Perch	45.904°N, 92.061°W	Calcote, Randy	NA	CALPRE2
		Lynch, Elizabeth		0
		Hotchkiss, Sara		
Sawdust	46.571°N, 91.272°W	Calcote, Randy	NA	CALPRE2
		Lynch, Elizabeth		1
		Hotchkiss, Sara		
Tomahawk	45.924°N, 91.961°W	Calcote, Randy	NA	CALPRE2
		Lynch, Elizabeth		2
		Hotchkiss, Sara		
Wade	45.984°N, 92.168°W	Calcote, Randy	NA	CALPRE2
		Lynch, Elizabeth		3
		Hotchkiss, Sara		
Wagner	46.203°N, 91.858°W	Calcote, Randy	NA	CALPRE2
				4

		Lynch, Elizabeth Hotchkiss, Sara		
West8Mile	46.423°N, 91.48°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE2 5
WestTwin	46.684°N, 91.063°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE2 6
Arrowhead	45.906°N, 89.69°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE2 7
BigStGermain	45.934°N, 89.52°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE2 8
Eagle	45.932°N, 89.208°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE2 9
BlueJ	46.407°N, 91.528°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE3
LittleSpider	45.971°N, 89.708°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE3 0
Lost	45.966°N, 89.483°W	Calcote, Randy	NA	CALPRE3 1

		Lynch, Elizabeth		
		Hotchkiss, Sara		
Otter	45.943°N, 89.221°W	Calcote, Randy	NA	CALPRE3 2
		Lynch, Elizabeth		
		Hotchkiss, Sara		
CircleLily	46.159°N, 89.919°W	Calcote, Randy	NA	CALPRE3 3
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Round	46.172°N, 89.71°W	Calcote, Randy	NA	CALPRE3 4
		Lynch, Elizabeth		
		Hotchkiss, Sara		
BigPortage	46.126°N, 89.285°W	Calcote, Randy	NA	CALPRE3 5
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Moon	45.918°N, 89.434°W	Calcote, Randy	NA	CALPRE3 6
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Nelson	45.93°N, 89.357°W	Calcote, Randy	NA	CALPRE3 7
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Oxbow	46.241°N, 89.683°W	Calcote, Randy	NA	CALPRE3 8
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Vandercook	45.982°N, 89.685°W	Calcote, Randy	NA	CALPRE3 9

		Lynch, Elizabeth		
		Hotchkiss, Sara		
Brozie	45.96°N, 92.108°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		
		Hotchkiss, Sara		
Day	46.063°N, 89.705°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		0
		Hotchkiss, Sara		
Katinka	46.205°N, 89.793°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		1
		Hotchkiss, Sara		
Nebish	46.052°N, 89.591°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		2
		Hotchkiss, Sara		
Spruce	46.053°N, 89.567°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		3
		Hotchkiss, Sara		
Palette09F	46.067°N, 89.604°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		4
		Hotchkiss, Sara		
Allequash07X	46.037°N, 89.629°W	Calcote, Randy	NA	CALPRE4
		Lynch, Elizabeth		5
		Hotchkiss, Sara		
BigMuskie07X	45.921°N, 89.7°W	Calcote, Randy	NA	CALPRE4
				6

		Lynch, Elizabeth Hotchkiss, Sara		
Crystal07X	46.157°N, 89.178°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE4 7
Sparkling07X	46.009°N, 89.7°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE4 8
Camp2G	46.502°N, 91.411°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE5
Chase	45.927°N, 92.468°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE6
Cheney	46.382°N, 91.701°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE7
Clubhouse	45.94°N, 92.057°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE8
Corwick	45.948°N, 92.197°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	NA	CALPRE9
Eleveneses	46.382°N, 91.496°W	Calcote, Randy Lynch, Elizabeth	26.5	CLH1

		Hotchkiss, Sara		
HellHole	45.787°N, 92.219°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	16.5	CLH3
Lily	45.901°N, 92.272°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	34.5	CLH4
Lone02	45.932°N, 92.236°W	Calcote, Randy Lynch, Elizabeth Hotchkiss, Sara	17.5	CLH5
Fallison	45.994°N, 89.619°W	Calcote, Randy Hotchkiss, Sara	23.5	CLH7
Adeline	45.302°N, 91.194°W	Calcote, Randy Hotchkiss, Sara	28.5	CLH8
Eska	45.192°N, 90.608°W	Calcote, Randy Hotchkiss, Sara	17.5	CLH9
Lake of the Clouds	48°N, 91.117°W	Craig, Alan J.	62	3482
Blandin Pond	47.133°N, 93.667°W	Cushing, Edward J.	41	280
Cedar Bog Lake	45.411°N, 93.197°W	Cushing, Edward J.	190	352
Irvin Lake	47.136°N, 93.644°W	Cushing, Edward J.	25	1153
Siseebakwet South	47.133°N, 93.683°W	Cushing, Edward J.	41	2394
Pogonia Bog Pond	45.033°N, 93.633°W	Cushing, Edward J.	35	3541
Pogonia Bog Pond	45.033°N, 93.633°W	Cushing, Edward J.	330	3558

South Heron Lake	43.75°N, 95.27°W	Cushing, Edward J.	36	10970
Ocheda Lake	43.59°N, 95.64°W	Cushing, Edward J.	36	10971
Blomford Lake	45.496°N, 93.157°W	Cushing, Edward J.	40	10973
Fawn Lake	45.41°N, 93.057°W	Cushing, Edward J.	40	10974
Frovold Lake	45.38°N, 95.568°W	Cushing, Edward J.	30	10975
Goose Lake	44.995°N, 94.596°W	Cushing, Edward J.	35	10976
Olson Lake	46.514°N, 93.662°W	Cushing, Edward J.	1080	10977
Round Grove Lake	44.718°N, 94.5°W	Cushing, Edward J.	45	10983
Cottonwood Lake	44.624°N, 95.692°W	Cushing, Edward J.	25	12005
Monson Lake	45.33°N, 95.29°W	Cushing, Edward J.	80	12006
Perch Lake	46.196°N, 95.472°W	Cushing, Edward J.	80	12009
Chub Lake	44.575°N, 93.222°W	Cushing, Edward J.	195	12010
Lake Rachel	45.808°N, 95.562°W	Cushing, Edward J.	35	12011
Ziebell Lake	46.057°N, 95.266°W	Cushing, Edward J.	50	12012
Mud Lake	45.861°N, 95.466°W	Cushing, Edward J.	65	12013
Lybeck Lake	45.679°N, 95.717°W	Cushing, Edward J.	55	12014
Union Lake	45.821°N, 95.348°W	Cushing, Edward J.	45	12015
Albert Lea Lake	43.658°N, 93.289°W	Cushing, Edward J.	70	12016
Everhart Lake	43.62°N, 93.442°W	Cushing, Edward J.	35	12024
Oak Glen Lake	43.934°N, 93.088°W	Cushing, Edward J.	258	13056

Camp Lake	45.413°N, 95.398°W	Cushing, Edward J.	40	13057
Blue Mounds Creek	43.083°N, 89.867°W	Davis, Anthony M.	11	287
Tamarack Creek	44.15°N, 91.45°W	Davis, Anthony M.	13	2624
Kitchner Lake	45.667°N, 87.455°W	Davis, Margaret B.	60	1524
Spirit Lake	46.47°N, 86.958°W	Davis, Margaret B.	20	2586
Canyon Lake	46.833°N, 87.921°W	Davis, Margaret B.	18	3055
Lily Lake	45.05°N, 92.825°W	Eyster-Smith, Nancy M.	1345	1633
Weber Lake	47.472°N, 91.66°W	Fries, Magnus	60	3837
Wolverine Lake	46.429°N, 85.661°W	Futyma, Richard P.	15	3032
Little Dollar Lake	46.187°N, 85.31°W	Futyma, Richard P.	12	12001
Little Pine Lake	45.283°N, 91.483°W	Gajewski, Konrad J.	150	1643
Vestaburg Bog	43.417°N, 84.883°W	Gilliam, Jeanne A.	618	3831
French Lake	44.95°N, 94.417°W	Grimm, Eric Christopher	660	850
Wolsfeld Lake	45.008°N, 93.577°W	Grimm, Eric Christopher	714	3031
Wood Lake	45.333°N, 90.083°W	Heide, Kathleen M.	100	3035
Kellys Hollow	45.3°N, 90.35°W	Heide, Kathleen M.	21	3452
Billy's Lake	46.283°N, 94.583°W	Jacobson, George L. Jr.	220	275
Ondris Pond	46.35°N, 94.4°W	Jacobson, George L. Jr.	210	1831
Myrtle Lake	47.983°N, 93.383°W	Janssen, C. Roel	186	1786

Demont Lake	43.48°N, 85°W	Kapp, Ronald O.	60	679
Frains Lake	42.33°N, 83.63°W	Kerfoot, Wilson Charles	410	847
Warner	46.121°N, 92.039°W	Tweiten, M.	38.5	CLH6
Devils Lake	43.427°N, 89.739°W	Maher, Louis J. Jr.	22.5	684
Iola Bog	44.5°N, 89.117°W	Maher, Louis J. Jr.	90	1151
Kellners Lake	44.238°N, 87.848°W	Maher, Louis J. Jr.	60	1448
Seidel	44.45°N, 87.516°W	Maher, Louis J. Jr.	40	3620
Lake Allie	44.814°N, 94.568°W	McAndrews, John H.	50	29
Bad Medicine Lake	47.133°N, 95.383°W	McAndrews, John H.	20	223
Black Bass Lake	46.128°N, 93.704°W	McAndrews, John H.	10	241
Bob Lake	46.634°N, 92.615°W	McAndrews, John H.	20	296
Bog D	47.183°N, 95.167°W	McAndrews, John H.	70	298
Bossuot Lake	44.296°N, 93.734°W	McAndrews, John H.	60	305
Campbell Lake	44.706°N, 93.5°W	McAndrews, John H.	50	334
Cindy Pond	47.183°N, 94.917°W	McAndrews, John H.	20	382
Cottonwood Lake	45.761°N, 96.017°W	McAndrews, John H.	40	515
Crazy Pond	47.183°N, 95.733°W	McAndrews, John H.	40	521
Dalton Lake	46.174°N, 95.915°W	McAndrews, John H.	40	545
East Twin Lake	45.333°N, 93.508°W	McAndrews, John H.	30	781
Eggert Lake	44.471°N, 93.62°W	McAndrews, John H.	40	786

Faith Pond	47.283°N, 95.1°W	McAndrews, John H.	30	817
Farquar Lake	44.758°N, 93.167°W	McAndrews, John H.	50	822
Fawn Lake	46.219°N, 94.753°W	McAndrews, John H.	30	824
Fox Pond	47.165°N, 95.558°W	McAndrews, John H.	50	846
George Lake	44.233°N, 93.871°W	McAndrews, John H.	40	869
Goldsmith Lake	46.583°N, 92.467°W	McAndrews, John H.	20	967
Grice Pond	47.2°N, 95.75°W	McAndrews, John H.	30	990
Hemo Pond	47.217°N, 95.783°W	McAndrews, John H.	60	1032
Horse Pond	47.19°N, 95.872°W	McAndrews, John H.	40	1109
Kegan Lake	44.758°N, 93.117°W	McAndrews, John H.	30	1446
Kryzewinski Lake	46.353°N, 92.87°W	McAndrews, John H.	30	1552
Lake Ann	45.425°N, 93.688°W	McAndrews, John H.	60	1584
Lewis Lake	45.749°N, 93.364°W	McAndrews, John H.	40	1629
Martin Pond	47.183°N, 94.933°W	McAndrews, John H.	50	1708
McCraney Pond	47.178°N, 95.698°W	McAndrews, John H.	30	1720
Lake Minnie	47.24°N, 95.01°W	McAndrews, John H.	40	1757
Mud Lake	44.227°N, 93.831°W	McAndrews, John H.	40	1777
Murphy Lake	44.711°N, 93.341°W	McAndrews, John H.	90	1781
Lake Onamia	46.068°N, 93.68°W	McAndrews, John H.	30	1830
Pennington Lake	45.781°N, 93.273°W	McAndrews, John H.	50	1884

Pleasant Lake	44.589°N, 93.567°W	McAndrews, John H.	40	1955
Portage Lake	47.081°N, 94.113°W	McAndrews, John H.	35	1966
Reichow	47.188°N, 95.775°W	McAndrews, John H.	30	2051
Rholl Lake	47.6°N, 96.167°W	McAndrews, John H.	20	2062
Sedivy Lake	47.209°N, 95.833°W	McAndrews, John H.	60	2354
Silver Lake	44.9°N, 94.2°W	McAndrews, John H.	70	2383
Solum Lake	46.662°N, 96.25°W	McAndrews, John H.	40	2581
South Stanchfield Lake	45.725°N, 93.467°W	McAndrews, John H.	30	2584
Spectacle Lake	45.572°N, 93.408°W	McAndrews, John H.	40	2585
St. Catherines Lake	44.627°N, 93.433°W	McAndrews, John H.	50	2600
Stone Lake	45.4°N, 93.522°W	McAndrews, John H.	30	2605
Sunfish Lake	44.875°N, 93.1°W	McAndrews, John H.	40	2613
Terhell Pond	47.188°N, 95.771°W	McAndrews, John H.	40	2638
Thompson	47.2°N, 96.083°W	McAndrews, John H.	30	2640
Toenjes Lake	45.41°N, 93.981°W	McAndrews, John H.	60	2648
Twin Lakes	45.849°N, 93.168°W	McAndrews, John H.	50	2881
White Lily Lake	46.086°N, 93.103°W	McAndrews, John H.	40	2952
Lake Minnie	47.24°N, 95.01°W	McAndrews, John H.	60	3492
Jones Lake	45.429°N, 93.983°W	McAndrews, John H.	20	15274

West Olaf Lake	46.604°N, 96.198°W	Nelson, David M.	165	13051
Cub Lake	44.7°N, 84.958°W	Rasmussen, Jack B.	24	532
Andrus Lake	46.7°N, 85.04°W	Shane, Linda C. K.	15.2	205
Brown Lake	45.783°N, 89.483°W	Shane, Linda C. K.	14.2	313
Brown's Bay Lake Minnetonka	44.967°N, 93.55°W	Shane, Linda C. K.	59.5	314
Camp 12 Lake	45.986°N, 89.375°W	Shane, Linda C. K.	17.2	333
Denton Lake	45.85°N, 89.317°W	Shane, Linda C. K.	22.2	681
Dunnigan Lake	47.708°N, 91.642°W	Shane, Linda C. K.	19.8	774
Hustler Lake	48.233°N, 92.158°W	Shane, Linda C. K.	26.8	1138
McNearny Lake	46.427°N, 84.958°W	Shane, Linda C. K.	5.8	1725
Nels Lake	48.033°N, 91.9°W	Shane, Linda C. K.	20.8	1793
Otto Mielke Lake	45.933°N, 89.625°W	Shane, Linda C. K.	22.2	1840
McNearny Lake	46.427°N, 84.958°W	Shane, Linda C. K.	7.8	3485
McNearny Lake	46.427°N, 84.958°W	Shane, Linda C. K.	8.8	3486
Deans Lake	44.778°N, 93.455°W	Shane, Linda C. K.	80	12007
Lake Emily	44.314°N, 93.926°W	Shane, Linda C. K.	153	12008
Hug Lake	47.958°N, 90.933°W	Swain, Albert M.	24	1134
Hell's Kitchen Lake	46.183°N, 89.7°W	Swain, Albert M.	90	3131
Lake of the Clouds	48°N, 91.117°W	Swain, Albert M.	9	3483

Beckman Lake	45.425°N, 93.183°W	Swain, Patricia C.	83	249
Little Bass Lake	47.283°N, 93.6°W	Swain, Patricia C.	1300	1639
Beckman Lake	45.425°N, 93.183°W	Swain, Patricia C.	35	3062
Third Lake	48.108°N, 92.017°W	Swain, Patricia C.	184	3784
Third Lake	48.108°N, 92.017°W	Swain, Patricia C.	40	3786
Third Lake	48.108°N, 92.017°W	Swain, Patricia C.	60	3788
Kimble Pond	44.22°N, 93.842°W	Umbanhowar, Charles E. Jr.	1910	13032
Dogfish Lake	48.183°N, 92.183°W	Waddington, Jean C.B.	25	717
Rutz Lake	44.87°N, 93.87°W	Waddington, Jean C.B.	560	2314
Jacobson Lake	46.417°N, 92.717°W	Watts, William A.	125	3138
Gass Lake	44.05°N, 87.733°W	Webb, Sara L.	60	860
Radtke Lake	43.4°N, 88.1°W	Webb, Sara L.	52	2011
Dark Lake	45.275°N, 91.475°W	Winkler, Marjorie Green	150	546
Lake Mendota	43.1°N, 89.417°W	Winkler, Marjorie Green	65.5	1735
Ruby Lake	45.283°N, 91.458°W	Winkler, Marjorie Green	150	2309
Burntside Lake	47.924°N, 91.995°W	Wright, Herbert E. Jr.	20	325
Lake Carlson	44.801°N, 93.162°W	Wright, Herbert E. Jr.	50	1588
Kirchner Marsh	44.772°N, 93.125°W	Wright, Herbert E. Jr.	22	3455

