Calibrating the pollen-vegetation relationship

Andria Dawson, Christopher Paciorek, Jason McLachlan,
Simon Goring, John Williams, Stephen Jackson

June 4, 2015

1 Introduction

Understanding forest ecosystems of the past provides valuable information about how these ecosystems respond to biotic and abiotic factors at timescales outside the realm of most direct ecological observations. In particular, in a world of changing climates, reconstructing past shifts in tree distributions and forest composition provides new information about climatic processes governing forest distributions and dynamics [Goring et al., in prep.], and vegetation-atmosphere feedbacks [Matthes et al., in review.]. The fossil pollen records extracted from lakes and mires are our primary source of data about prehistoric forest composition. Rigorous estimates of past forest composition and forest dynamics from these records relies on our ability to quantify the pollen-vegetation relationship.

Key challenges to quantitative reconstructions of past vegetation from pollen data are 1) the complex nature of the pollen-vegetation relationship and 2) the heavy anthropogenic imprint on contemporary plant communities. The processes governing pollen production, transport, and deposition have been the subject of palynological study for decades [Tauber, 1965, Jackson, 1994, Jackson and Lyford, 1999, Sugita, 2007a,b, Prentice, 1988]. The pollen found in lakes and mires usually arrives by wind, from plants that vary in their pollen productivity, the dispersibility of pollen grains, and transport vector. As a result, some plant species are highly overrepresented in sedimentary archives and most are underrepresented or absent. A wide variety of pollen-vegetation models (PVMs) have been developed including modern-analog based approaches [Williams, 2003], linear-regression [Webb et al., 1981], mechanistic models of pollen transport [Sugita, 2007a,b, Prentice, 1988], and, most recently, Bayesian hierarchical models [Paciorek and McLachlan, 2009, Garreta et al., 2010]. In these Bayesian models, pollen counts from a network of ponds are modelled as a function of gridded forest composition data, taking into account pollen dispersal from both local and non-local vegetation, differential pollen production, and process uncertainty. Using a Bayesian hierarchical model allows borrowing of information across space to estimate governing parameters at a regional scale, and to estimate the uncertainty associated with these parameters.

Some PVMs, notably the LOVE/REVEALS family of models [Sugita, 2007a,b], are calibrated using information about pollen productivity and dispersal traits for individual species and diffusion-based models of particulate transport. Other PVMs are calibrated against spatial networks of pollen assemblages extracted from surface sediment samples, which are then

cross-referenced to contemporary spatial datasets of vegetation composition and abundance of the surrounding forests. These cross-referenced pollen and vegetation datasets are then use to make inferences about the processes that affect pollen production, dispersal, and deposition at the time of sampling. However, extending any inference about these processes to times prior to the time of sampling requires the assumption that pollen-vegetation relationships are stationary. This assumption of stationarity is questionable given that the contemporary landscape has been heavily modified by human land use. This land-use has transformed the composition and structure of contemporary forests and likely altered vegetation-pollen relationships for some taxa [Kujawa et al., in prep.], perhaps by altering microclimatic factors including wind dynamics. Pollen-based paleoclimate reconstructions in Minnesota have a demonstrated bias if modern calibration datasets are used instead of pre-settlement datasets [St Jacques et al., 2014]. Hence, any PVM that is calibrated using contemporary datasets may carry implicit and difficult-to-quantify biases when applied to make inferences about vegetation composition for time periods with less intensive land use. To circumvent these issues, we calibrate the pollen-vegetation relationship using newly developed datasets of forest composition and sediment pollen data that pre-dates European settlement, thereby avoiding the effects of settlement, agriculture, and industrialization.

These datasets are based on US Public Land Survey (PLS) records collected by surveyors prior to European settlement [Bourdo, 1956, Schulte and Mladenoff, 2001].

To calibrate the pollen-vegetation relationship against the PLS forest composition data requires that we identify pollen samples that correspond to the pre-settlement era. These fossil pollen samples are drawn from longer sequences that typically come from lakes or peatlands, and consist of a series of depths and associated pollen counts broken down by taxon. Most pollen records also have radiocarbon dates of organic materials scattered along the core at locations not necessarily aligned with pollen counts. Other age controls include other radiometric dates (e.g. 210Pb), biostratigraphic markers (e.g. the rise in Ambrosia and other weedy taxa during EuroAmerican settlement), and geostratigraphic markers such as the core top whose age typically corresponds with the year of sampling. These age controls are then used to constrain age-depth models that allow inference of age at any depth throughout the core. However, inferred ages are uncertain, due to analytical errors during the laboratory radiocarbon dating process [Ward and Wilson, 1978], the conversion of radiocarbon to calendar years [Reimer et al., 2013], and potential differences between age of macrofossil material and age of sediment [Blois et al., 2011]. As in ecology, Bayesian hierarchical modeling approaches to age-depth models are rapidly gaining popularity for their ability to accommodate irregular probability distribution functions, multiple sources of uncertainty, and prior information [Blaauw et al., 2011, Buck et al., 1999, Blaauw and Christen, 2005, Ramsey, 1995].

During European settlement, land-clearances led to noticeable increases in certain agricultural indicators that is reflected in the pollen records, especially *Ambrosia* [McAndrews, 1988]. This change in relative composition of *Ambrosia*, referred to as the *Ambrosia rise*, can be used to identify the pollen sample that best corresponds with pre-settlement - the pre-settlement pollen sample is the sample immediately prior to the first pollen sample that reflects the *Ambrosia rise*. The presence of an established biostratigraphic marker at the time of settlement eliminates the need to rely on uncertain age-depth models. However, time series of pollen counts are noisy, and in practice it can be difficult to precisely iden-

tify the settlement horizon. We deal with the uncertainty associated with *Ambrosia rise* identification using expert elicitation.

Here, we refine and calibrate a Bayesian hierarchical PVM, based on the STEPPS model [Paciorek and McLachlan, 2009], using newly developed settlement-era pollen and vegetation calibration datasets for the Upper Midwest. Specific advances include the building of a pre-settlement era identification of pre-settlement pollen samples from pollen time series through expert elicitation and the testing and comparison of two standard isometric dispersal kernels. We apply modelling results to develop new estimates of relative pollen productivity and source area, assess the relative importance of local and non-local pollen sources, and compare these results to prior estimates of source area and pollen productivity. This work lays the methodological foundation for rigorous reconstruction of past forest composition from fossil pollen data and all scripts are made publicly available at GitHub.

2 Data

2.1 Spatial domain

Our study area is the upper Midwestern US, including Minnesota, Wisconsin, and Michigan. This region includes two major ecotones: the "Tension Zone" between northern mixed forests and temperature broadleaved deciduous forests and the prairie/forest/savanna ecotone. Both of these ecotones have been heavily modified by contemporary land use [Goring et al., in review.].

2.2 Public Land Survey (PLS) data

Prior to major European settlement, the US General Land Office conducted a Public Land Survey (PLS) throughout much of the United States to simplify the sale of federal lands. Surveyors documented section locations using trees as landmarks, and recorded genus or species, diameter, and location (azimuth and distance from corner). This data set provides a systematic survey of the forest before settlement, and has been used by foresters, ecologists, and historians to understand ecosystem and land-use change through time. In the Upper Midwest, the survey was conducted during XXXX-XXXX. Due to the slow-growing nature of temperate forests, including those in the Upper Midwest, we can think of the PLS data as a snapshot of forest composition in time.

Survey data for the Upper Midwest has been recently digitized, and aggregated to an 8km square grid [**?**]. As a result of the sampling methodology, this aggregation results in small numbers of tree counts per grid cell. To overcome the variability arising from this we work with a smoothed version of the PLS data, based on a Bayesian spatial multinomial model [**?**].

2.3 Tree taxa

We focus on a subset of 10 tree genera and 2 functional groups that includes the most abundant taxa as well as less-abundant taxa of particular ecological importance. Our modelled

taxa are: ash, beech, birch, elm, hemlock, maple, oak, pine, spruce, larch, as well as other conifer and other hardwood groups. These latter two groups include other minor tree species identified by surveyors but not explicitly included in the aforementioned list of taxa. The separation of other hardwood and conifers follows the standard classification of tree species into plant functional types in earth system and ecosystem models [Cramer et al., 2001, Goring et al., 2015]. Additionally, the separation of other hardwood and conifers allows the calibration model to treat each group separately and tease out inherent differences between conifer and deciduous pollen production, although the variability in production and dispersal within each of these groups is still large.

2.4 Pollen data

With the push for robust and reproducible research from the scientific community, paleoecoinformatics has responded with the development of tools that make it possible to access and query large datasets. One such tool that has made this work possible is the Neotoma database (`neotomadb.org`; [**?**]), which stores a variety of types of paleoecological data, including pollen data. Accessing this data can be done using the Neotoma API (`http://apps.neotomadb.org/explorer/`), or using the R neotoma package [Goring et al., 2015]. Using R neotoma we identifed 184 sediment pollen cores in the Neotoma database in our spatial domain with the restriction that these sediment cores include at least some pollen samples believed to be from the last 2000 years. All of these pollen cores were sampled at multiple depths, often regularly spaced, and are hereafter referred to as long cores.

An additional 57 cores not available on Neotoma were contributed by independent researchers [Kujawa et al., in prep.]. Of these 57, 9 were long cores (analogous to those from neotoma), while the remaining 48 had only core top and pre-settlement samples (at least in the data file we had access to), and are therefore referred to as short cores.

Associated with each of the cores is a table containing counts by taxon for each sampled depth. These counts are not certain - classifying pollen grains based on morphological features and relating them to plant taxa is often challenging as a result of (differential) preservation. Counts between palynologists for a single sample will also differ, but for experts this difference should be negligible [**?**].

One of the challenges that arises when working with pollen data collected by several (or more) individuals is dealing with naming inconsistencies in the pollen taxonomies between sites. Both differing nomenclature and identified taxonomic resolution require that care be taken when standardizing pollen taxonomies. There arise cases where it may not be possible to associate certain pollen grains with plant taxa, and in these cases grains may be identified to type (e.g. Alnus-type), or placed into a group whose name is defined by their morphology (e.g. tricolporate). In this study, decisions regarding these poorly identified pollen grains were made on a case-by-case basis, although the total number of grains that fell into these ambiguous classes were few.

For each pollen core we are interested in the pre-settlement sample that is closest in time to the PLS data collection date (which is spatially varying). Typically, age-depth models are used to assign ages to sample depths to sediment cores. However, age-depth modelling is an uncertain process, with no consensus in the paleoecological community on a single best model or method. Instead, we rely on a panel of experts to interpret patterns in the pollen

count data in order to identify pre-settlement sample estimates. All 185 long cores were suitable for this exercise.

2.5 Identifying the EuroAmerican settlement horizon in pollen records via expert elicitation

Widespread land clearance during European settlement provided habitat for many non-arboreal colonizers, resulting in increases in non-arboreal taxa in pollen time series [McAndrews, 1988]. In the Upper Midwest, significant increases in ragweed (*Ambrosia*), docks and sorrels (*Rumex*), and grasses (*Poaceae*) are typically coincident with this settlement horizon. In principle, it is possible to identify the pre-settlement sample - the sample that falls immediately before these increases in agricultural indicator species. For many records the rises are sharp and determining the pre-settlement sample is easy, while other records are more ambiguous and placement relies on subjective judgment.

For this study, in the interest of reducing uncertainty, we want to: 1) identify pre-settlement samples using consistent methodology, and 2) assess the variability in assignment of pre-settlement among analysts. To address these questions, we asked a team of expert palynologists to identify the pre-settlement sample for 185 pollen records. Experts were provided with pollen diagrams depicting proportional changes through time as a function of depth for key indicator species and the ten most abundant arboreal taxa. Experts were prohibited from relying on stratigraphic dates (radiocarbon or other) or age-depth model estimates of sample age. In the case that there was no distinguishable pre-settlement sample, experts were instructed to report NA. In the case that experts were uncertain about their pre-settlement sample assignment, they were instructed to note this, with or without justification. Results from this exercise are used to define depths associated with pre-settlement samples. Pollen samples associated with these depths are placed into our calibration data set.

3 Calibration model

Here we describe the Bayesian hierarchical calibration model used to quantify the pollen-vegetation relationship. For the sake of clarity, we first provide an overview of the model, and then introduce the mathematical notation.

The pollen counts that we observe are a function of the vegetation on the landscape - the composition and abundance of trees on the landscape determines (in part) the composition and abundance of pollen that is deposited in the sediment. Determining how much pollen the surrounding vegetation contributes is a complex function of forest characteristics, climate (including wind), grain size and morphology, topography, and proximity of the vegetation to a deposition basin.

We work on a grid defined by the resolution of the gridded PLS data, and define each deposition site as having both local and non-local pollen contributions. Local contributions are made by the vegetation within the same grid cell as a site, and non-local contribution are made by vegetation from all other grid cells in the domain. The relative contributions of the local and non-local contributions are determined by a model parameter, with the restriction that the sum of all contributions from the pollen source area must sum to one.

Contributions from non-local grid cells are not all equal - their relative contributions are weighted using a dispersal kernel, which determines how likely it is for pollen to travel from one grid cell to another. Dispersal kernels are typically isometric, meaning that they depend only on the distance of the tree to the pond, and not the absolute position or direction. In reality, dispersal is not isometric, predominantly due to prevailing wind direction, but also as a result of other asymmetries in surface roughness and landscape structure. However, isometric dispersal kernels are are a reasonable and widely used approximation, because turbulent atmospheric mixing below the boundary layer tends to smooth out anisotropies.

Here we test two dispersal kernels: the Gaussian and the Inverse power-law. The Gaussian kernel is often used as a reference against more leptokurtic kernels such as the Inverse power law kernel, and can represent dispersal through diffusion. Although the Gaussian kernel has simpler mathematical representation, fatter-tailed kernels (such as the Inverse power-law) in general perform better than those with thinner tails (like the Gaussian or Exponential) when modelling pollen dispersal [Devaux et al., 2007, Austerlitz et al., 2004]. However, dispersal data needed to validate pollen dispersal models are generally unavailable, especially at the large spatial scales needed to adequately test long-tailed kernels [Clobert et al., 2012].

Here, we choose two kernels that are representative of the short- and long-tail kernel classes, and compare between these generalized functional forms.

To determine the total non-local contribution, we sum the contributions of all grid cells in the domain that do not contain a given pond, where a contribution is equal to the weight assigned by the kernel multiplied with the vegetation proportion at that grid cell. Then, we sum the local and non-local contributions to obtain a single vector of proportions that represent the effective contributing vegetation. Finally, each component of this vector is scaled by a taxon-specific parameter (γ) that accounts for differential production of pollen among plant species.

3.1 Model description

As previously mentioned, the spatial domain is defined as a regular grid composed of 8 km square grid cells. The grid is composed of discrete cells, but the underlying vegetation composition and dispersal spatial processes are assumed to be smooth. Spatial cells are indexed by $s = 1, \ldots, S$, where $S = 8013$.

Pollen produced by vegetation within each grid cell can be deposited locally within that same grid cell, or can be dispersed into the neighborhood around that grid cell. For a focal grid cell s_i, the pollen produced by taxon p within that cell that remains local is described by

$$\gamma \phi_p r_p(s_i) \tag{1}$$

where γ is the proportion of pollen produced in s_i that is deposited locally, ϕ_p is the scaling factor that accounts for differential production, and $r_p(s_i)$ is the proportional abundance of taxon p in s_i.

The remaining proportion $(1 - \gamma)$ of pollen produced in s_i is dispersed to other grid cells according to an isotropic dispersal kernel centered at s_i. The dispersal kernel weights

all pollen dispersing away from the focal cell as a function of the distance from s_i to any neighboring cell s_k by $w(s_i, s_k)$.

The two dispersal kernels considered here are: (i) the Gaussian and (ii) the Inverse power-law.

The Gaussian kernel, denoted $w_g(s_i, s_k)$, is written as

$$w_g(s_i, s_k) = \exp\left(-\frac{d(s_i, s_k)^2}{\psi^2}\right), \tag{2}$$

where $d(s_i, s_k)$ defines the distance between cells s_i and s_k and $\psi > 0$ is a parameter that describes the spread of the kernel.

The Inverse power-law kernel, denoted by $w_{pl}(s_i, s_k)$, is given by

$$w_{pl}(s_i, s_k) = \frac{(b-2)(b-1)}{2\pi a^2}\left(1 + \frac{d(s_i, s_k)}{a}\right)^{-b}, \tag{3}$$

where $a > 0$ and $b > 2$ are parameters that determine the kernel shape.

As expected, the weight assigned by both kernels is a decreasing function of distance - less pollen is distributed farther away.

We can then define the pollen produced by taxon p dispersing from s_i to s_k by

$$(1-\gamma)\phi_p r_p(s_i) w(s_i, s_k). \tag{4}$$

So far we have characterized the model from a source-based perspective, describing how pollen produced in grid cells is dispersed. We would like to quantify the pollen arriving at pond i located in the grid cell $s(i)$ - our data is counts of pollen grains that have been deposited in the sediment. This arriving pollen is equal to the sum of the contributions from all cells in the domain to the pollen pool in $s(i)$, which includes both the locally deposited pollen plus the pollen dispersed to $s(i)$ from all other grid cells in the domain. Therefore the pollen from taxon p arriving at pond i, referred to as $r_{i,p}^{\text{pol}}$, is given by

$$r_{i,p}^{\text{pol}} = \gamma\phi_p r_p\big(s(i)\big) + \frac{1}{C}(1-\gamma)\phi_p \sum_{s_k \neq s(i)} r_p(s_k) w\big(s(i), s_k\big), \tag{5}$$

where C is a normalizing constant equal to the sum of the weights of all the cells to which pollen can be dispersed, defined be a rectangular region that covers and extends beyond the limits of the domain.

Finally, we need to relate the arriving pollen $r_{i,p}^{\text{pol}}$ quantified in Equation 5 to the pollen count data. To do this, we must consider that the pollen data are overdispersed - the counts are more variable than we might expect if they were assumed to follow a multinomial distribution. This overdispersion is mostly attributed to heterogeneity in the pollen production and dispersal processes that is not accounted for in our mathematical characterization, but is also an artifact of comparing pollen counts at a point location (the pond) situated anywhere within a focal grid cell to a grid-based vegetation composition average. To account for this overdispersion, we model the pollen counts at pond i, denoted by $\boldsymbol{y}_i$, as dirichlet-multinomial (DM), which is a compound distribution used to account for overdispersion in multinomial

count data (and tends to the multinomial distribution in the case of no overdispersion). We have that

$$\boldsymbol{y}_i \sim DM(n_i, \boldsymbol{r}_i^{\text{pol}}) \tag{6}$$

and $\boldsymbol{r}_i^{\text{pol}} = (r_{i,1}^{\text{pol}}, \ldots, r_{i,K}^{\text{pol}})$. The precision α_i is equal to the sum of (5) over all taxa

$$\alpha_i = \sum_{p=1}^{K} \Big[\gamma \phi_p r_p\big(s(i)\big) + \frac{1}{C}(1-\gamma)\phi_p \sum_{s_k \neq s(i)} r_p(s_k) w\big(s(i), s_k\big) \Big]. \tag{7}$$

The precision parameter α_i is also affected by the proximity of a lake to the domain boundary. We do not have data outside of the defined domain, so cannot quantify the pollen dispersed from vegetation outside the domain to lakes situated close to the boundary. The repercussion of this is that the sum of the weights of the grid cells contributing pollen to a focal cell close to the domain boundary is less than the sum of the weights of the total potential contributing neighborhood, defined as C. As a result of having a reduced contributing neighborhood, the local pollen contribution is effectively up-weighted relative to the non-local contribution (see Eqn. 7).

3.2 Model priors

We assign uninformative uniform priors to model parameters over wide intervals. We define the priors $\phi_k \sim \text{uniform}(0, 300)$ for $k = 1, \ldots, K$ and $\gamma \sim \text{uniform}(0, 1)$. In the case of the Gaussian kernel we define the prior ψ uniform$(0, 2)$, while in the power law kernel case we have a uniform$(0, 500)$ and b uniform$(2, 100)$.

3.3 Model variants

As described above, neither dispersal kernel (w_{pl} or w_g) nor the parameter γ vary by taxon. In this base variant, a single dispersal kernel defines how all pollen disperses, and a single parameter determines what proportions of total pollen of any given taxon remains local. We know that in reality pollen dispersal varies with taxon, because pollen grain morphology, size, weight, and preservability are all highly variable among plant taxa. Here, we let the data inform us as to whether there is sufficient evidence to support taxon-specific dispersal kernels and localness parameters through the use of exchangeable priors ?. So, we define the same prior distribution on each of the K parameters in question, where the prior distribution is defined by shared hyperparameters. In the case where the variance hyperparameter is estimated to be relatively close to zero, we can conclude that there is not sufficient evidence to suggest that the parameter of interest varies by taxon.

For example, in the Gaussian kernel case we would like to know if there is enough support to warrant taxon specific values of ϕ. We define the exchangeable priors for taxon $k = 1, \ldots, K$

$$\log(\psi_k) \sim \text{normal}(\mu, \sigma) \tag{8}$$

where the prior on μ is uniform over the log of the prior interval used in the single-taxon case for ψ, and the prior for σ is a half-Cauchy with mean of 0 and standard deviation of 2. In the Power law kernel case, when we let a and/or b vary, their exchangeable priors are analogous to the ψ case. Note that in the Power law case, b is allowed to vary by taxon when a varies by taxon.

In the variable gamma case, we define a logit-normal prior for each γ_k, where $\mu_\gamma \sim \text{uniform}(-2, 2)$ and $\sigma_\gamma \sim \text{Cauchy}(0, 5)$ defined on $[0, \infty)$.

Allowing γ and/or the kernel to vary by taxon results in a total of 10 model variations as described in Table 5.

3.4 Model comparison

To identify the set of candidate models, we first assessed the variance estimates for the exchangeable priors. In the case where there was insufficient evidence to support the case for taxon-specific parameters, as indicated by (i) an estimated variance close to zero and (ii) overlapping credible intervals for the posterior distributions of the taxon-specific parameters, the model was not considered as a candidate model.

Then to determine which kernel and model best fit the data set we perform a formal model comparison using the Wantanabe-Akaike Information Criterion (WAIC) [Watanabe, 2010]. The WAIC provides a relative measure of loss of information (as does the more familiar AIC) in a Bayesian context, avoiding the necessity of determining the number of model parameters (as required by AIC) while accounting for the full posterior distribution as opposed to conditioning on a single point estimate (as in AIC and DIC) [Gelman et al., 2014].

3.5 Numerical implementation

We cannot directly sample from the joint posterior, so we rely on MCMC methods. With the goal of achieving more efficient sampling with respect to the effective sample size per unit time, we used the Stan statistical modeling software to estimate parameters [Stan Development Team, 2014]. Stan implements a variant of the Hamiltonian Monte Carlo method called the No-U-Turn Sampler, which is a gradient based sampling method that uses these directional derivatives to make informed decisions about how to move along (and sample from) the joint posterior surface [Hoffman and Gelman, 2011].

4 Results

4.1 Expert elicitation of biostratigraphic evidence for European settlement and site suitability

Four experts participated in the elicitation exercise. For 59 out of 185 sites, the experts were in total agreement: they all identified the same pre-settlement sample (53 cases) or agreed that no such sample could be identified (6 cases). The remaining sites varied in level of disagreement - in 79 cases, experts identified two samples; in 39 cases 3 pre-settlement

samples were identified; and in 8 cases there was no agreement. These results confirm our hypothesis that identification of biostratigraphic markers is subject to variability between analysts and is the first to quantify this uncertainty.

Of the 6 sites that experts agreed had no discernible settlement signal, two of these sites were at Rice Lake, which is an anomalous site overwhelmed by a wild rice (*Zizania*) and grass (*Poaceae*) signal, and does not show any indication of settlement. The remaining four included: Rossburg peat bog in Minnesota; Disterhaft Farm Bog in Wisconsin; Kirchner Marsh in Minnesota; and Stewarts Dark Lake in Wisconsin. Three of these four sites are bogs and marshes with local marshland non-arboreal plants growing at or near the coring location; these marshland plants may be swamping the signal from agricultural non-arboreal indicators.

Based on these results, we were subsequently faced with the challenge of establishing site suitability criteria to determine which sites to include in the calibration data set. Ideally, the now quantified pre-settlement sample uncertainty would be included into the modelling framework, but this option was rejected for now in favor of simplicity of the statistical model. Instead, we set rules for filtering suitable sites and assigning pre-settlement depths on expert determinations.

Based on these results, we were subsequently faced with the challenge of establishing site suitability criteria to determine which sites to include in the calibration data set. Ideally, the now quantified pre-settlement sample uncertainty would be included into the modelling framework. Although this is surely possible, it would increase the complexity of the statistical model. For the sake of simplicity, we filter suitable sites and assign pre-settlement depths in a systematic way based on expert determinations.

Sites were considered unsuitable if: 1) a majority chose not to assign a pre-settlement sample (12 sites), or 2) if half of the experts chose not to assign a pre-settlement sample and the remaining half identified pre-settlement samples whose modelled ages were far from the approximate time of settlement (7 sites). For this second criterion, we looked to existing age-depth models associated with the pollen cores. Modelled ages for the identified pre-settlement depths were compared to 1850 AD, which serves as an approximate year of settlement in the upper Midwest. Modeled ages were defined to be far from the expert-identified settlement horizon if discrepancies were ¿500 years All 19 unsuitable sites were excluded from further analysis.

After completion of the elicitation exercise, a follow-up examination of the stratigraphic data revealed several cores with core tops with dates much older than expected - typically the date of a true core top corresponds with the year of sampling (with some margin of error). Further investigation uncovered that three of the cores included in our analysis were missing core tops. For both Lake Mary and Green Lake, we determined that the core tops pre-date settlement, and these sites were therefore were discarded [Webb, 1971, Lawrenz, 1975]. A third site, Lake Kotiranta, had a core top corresponding to the pre-settlement sample, and this sample was retained for inclusion in the calibration data set [Wright et al., 1969]. Interestingly, for each of these cores, 1-3 experts identified a pre-settlement sample (although to be fair, they made this judgment based solely on the stratigraphic diagram and no information about site characteristics, and likely assumed that the uppermost sample was in fact a surface sample).

After the suitability screening, we were left with 165 long cores. For many cores, there was

no consensus, which left us to determine how to assign our best estimate of pre-settlement. Given the sample resolution and an even number of experts, neither the mean nor the median were options. Keeping in mind that we would rather err on the side of caution and choose a depth that was more certainly pre-settlement (deeper) than one that represented post-disturbance times (more shallow), we ordered the sample depths from deepest to shallowest, and selected the second in this list (typically the second deepest, although sometimes the two largest sample depths were equal). The additional 48 short cores from the Calcote dataset were not candidates for the elicitation exercise because they had only a core top and a pre-settlement sample recorded. The decision to include this additional data in the calibration data set was debated - the pre-settlement samples for these cores were identified using different methodology (and analysts) than the remaining sites. In the end, we opted to include these samples as part of the calibration data set, based on our confidence in the data set and analysts and the recognition that including these sites would result in a substantive increase in sample size. In total we had 213 pollen samples in the Upper Midwest settlement-era calibration data set.

4.2 Exploratory data analysis

To visually assess the relationship between sediment pollen and tree taxa, we compare pie maps which depict the relative proportions of taxa across space (Figure 1). In these maps the broadscale patterns are consistent with each other, with both pollen and PLS data capturing the major ecotones in the region [Solomon and Webb, 1985]. However, there are some striking differences. First, pine dominates the pollen records for most of the northern half of the domain. Second, in the vegetation composition data, some regions that have higher relative abundances of Hemlock, Tamarack, and Maple are not apparent or much less pronounced in the pollen data. These taxa are known to be underrepresented in the pollen record [**?**]. As expected, these pictures confirm that pollen is able to capture key vegetational gradients, but that the relationship between pollen and vegetation is complex.

To further assess spatial distributions of pollen versus PLS data, we plot the data as heat maps by taxon. Differences in extent of these distributions indicate pollen dispersal beyond species range limits. For example, for both birch and pine, both of which are known to be overrepresented in pollen records, the distributions of sediment pollen extend well beyond the boundaries of those shown in the PLS data (Figures 3 & 4).

4.3 Modelling results

A total of ten models were considered - four with the Gaussian kernel, and six with the power-law kernel. For each model, three chains were run with a warm-up of 250 iterations, followed by a sampling period of 10,000 iterations. Warm-up iterations were not used in further analysis. For a total of 30,000 iterations per model, the effective sample sizes of the joint log posterior ranged from 12,247 for the Gaussian kernel base model to XXXX for the XXX model. Effective sample sizes for the joint log posterior and parameter posteriors are given in Table **??**. Trace plots for the joint log posterior show the efficient mixing achieved by the sampler (See Figure 2 for trace plots from the Gaussian dispersed model; trace plots of other models show similar mixing, and are therefore omitted).

To compare among models, we first assess the estimates of the hyperparameters for the exchangeable priors. All parameter and hyperparameter estimates support the use of taxon-specific parameter values, except for the power-law models which allow b to vary by taxon. In these cases, the estimated variance for the exchangeable prior for $\log(b)$ is small (9e-6 for the variable a and b case; 1e-4 for the variable a, b, and gamma case). In addition, the 95 % credible intervals for the estimate of b_k for all k overlap, supporting the conclusion that the values of b_k are not statistically different. In light of this, we reject these two power-law kernel variable b models.

Out of the remaining eight candidate models (four with the Gaussian kernel; four with the power-law kernel) we compare the WAIC (Table 5). According to the WAIC the power law kernel outperforms the Gaussian kernel - even the most flexible Gaussian model (variable ψ and γ) results in a WAIC that is higher than the least flexible power law kernel model (base model). The power law kernel model with variable a and γ results in the lowest WAIC, indicating that it performs the best out of the suite of considered models. However, it is still of interest to make further comparisons between the two kernels - does the kernel formulation make a noticeable different in pollen production estimates and pollen predictions? One of the goals of this work is to use the calibrated pollen-vegetation relationship to estimate vegetation composition back through time, and therefore it is natural to think about how kernel selection affects predictions, if at all.

In light of this, we continue to compare the two base models (Gaussian and power law), as well as the models with the lowest WAIC for each kernel type - the Gaussian variable ψ and γ model and the power law variable a and γ model (hereafter referred to as the Gaussian or power law variable models).

In Figure 5, the raw pollen proportions are plotted against the vegetation proportions for each grid cell (red crosses). The relationship between the pollen and vegetation is clearly not 1:1. In particular, we see that some taxa, such as beech, maple, other conifer, and tamarack, are not prolific pollen producers - they can appear in large proportions on the landscape, but never appear in large proportion in the pollen record. Some relationships are difficult to identify, for example pine, where it varies from sparse to abundant on the local landscape, but this relative composition provides little indication about the relative abundance of pine pollen.

These relationships are complicated in part because of differential pollen production, which the calibration modelaccounts for by the taxon-specific scaling parameter ϕ_k. After scaling the local vegetation in a cell by ϕ_k, we can again compare the raw pollen proportions with these pollen predictions (Figure 5; black and blue dots). Here we show predictions only the variable gaussian and power law models, but note that the results are similar for the base models. Scaling by ϕ does improve our pollen estimates; for many taxa, the black points have shifted towards the 1:1 line. This is especially true for maple, tamarack, and other conifer, suggesting that these pollen-vegetation relationships are heavily influenced by local pollen sources. However, for most of the remaining taxa the improvement is minimal, indicating that simply scaling the local vegetation proportions is not sufficient to predict deposited pollen.

In Figure 6, we again plot the raw pollen proportions against both the vegetation (red crosses) and predicted pollen for the variable models (Gaussian: black dots; power law: blue dots), but in this case the predicted pollen is based on the local plus non-local predictions ob-

tained from the full calibration model. Now that non-local dispersal has been accounted for, the raw versus predicted pollen points fall more closely along the 1:1 line. This indicates that in our domain non-local dispersal processes account for much of the pollen that arrive at any given location. For most taxa, the improvement that we gain by using more flexible models versus the base models is not qualitatively obvious. This is not true for Hemlock - both base models underpredict Hemlock pollen at sites where there is relatively high proportion of Hemlock (although Hemlock pollen never exceeds 28% at any site in our domain). This suggests that the relative contribution of local versus non-local hemlock pollen is estimated to be less than the true value, i.e. γ is too small.

In Figure 6, we again plot the raw pollen proportions against both the vegetation (red crosses) and predicted pollen for the variable models (guassian: black dots; power law: blue dots) by taxon, but in this case the predicted pollen is based on the local plus non-local predictions obtained from the full calibration model. Now that dispersal has been accounted for, we see that the raw versus predicted pollen points fall more closely along the 1:1 line. This indicates that in our domain dispersal processes account for much of the pollen that arrive at any given location. For most taxa, the improvement that we gain by using more flexible models versus the base models is not qualitatively obvious. This is not true for Hemlock - both base models underpredict Hemlock pollen at sites where there is relatively high proportion of Hemlock (although Hemlock pollen never exceeds 28% at any site in our domain). This suggests that the relative contribution of local versus non-local pollen is estimated to be less than it should be, i.e. γ is too small.

There are two anomalous sites for which the model is unable to adequately account for the processes leading to the (relatively speaking) unusual compositional makeup of the deposited pollen. The first is Ocheda Lake, a large but shallow glacial kettle lake in Southwestern Minnesota. The pollen diagrams indicate a relative spike in deposited Maple pollen in the pre-settlement sample. The second is Tamarack Creek, a site situated in a wetland with high Tamarack pollen counts for the top 80 cm of sediment. At this site, the calibration sample is composed of 40 % Tamarack, while all other fossil pollen sites contain a maximum of 1 % Tamarack.

From each model we obtain taxon-specific estimates for the ratio of local to non-local pollen ϕ (Figure 7). If these parameter estimates truly represent differential pollen production/dispersal, then these estimated values should be consistent across model variants, which is generally the case (Figure 7). However, for several taxa including beech, Hemlock and Pine the variable power law kernel models predicts much larger ϕ values than the other models. However, in some of these cases the credible intervals are large, and overlap with competing estimates of ϕ, suggesting that these estimates may not be statistically different. Aside from the few uncertain values resulting from the variable power law model, the estimated values of ϕ form three distinct groups in ϕ parameter space: low, intermediate, and high production/dispersal (Figure 7). The group of taxa with lowest production/dispersal in decreasing order are beech, maple, other conifer, and tamarack. This low production/dispersal pattern is evident in Figure 6, where the representation of these taxa in the pollen records is consistently less than their representation on the landscape (except at a few anomalous ponds). Intermediate producers/dispersers, again in decreasing order, include other hardwood, oak, elm, hemlock, ash, and spruce. The high production/dispersal group includes pine and birch, which can also be seen in Figure 6 by their propensity to be over-represented

in the pollen record relative to the landscape. These results agree in general with [Prentice and Webb, 1986], which analyzed the pollen-vegetation relationship for sites in Wisconsin and the Upper Peninsula of Michigan. In that work, pine and birch had the largest slopes (top producers/dispersers), and maple and tamarack as limited producers/dispersers. [Jackson, 1990] also found that Pine and Birch were good dispersers with effective source areas of 1000m, while Maple was found to have a much smaller source area indicating limited dispersal.

For the base models, the local versus non-local weight parameter γ was estimated to have a mean of 0.21 for the Gaussian kernel model and 0.07 for the power law kernel model. These values indicate that 21% (or 7%) of the pollen produced by vegetation in a focal grid cell is deposited in that grid cell, while the remaining 79% (or 93%) disperses elsewhere in the domain.

When we allowed γ to vary by taxon in the variable models, the estimates for both kernels yielded similar results with overlapping 95 % credible intervals for all taxa except Pine and Oak (Table 5). Birch, Other conifer, Ash, and Other hardwood all had lower estimates of γ, with mean posterior values all less than 0.12 (mean values ranging from 0.03 - 0.11). Those taxa with larger estimated values of γ were Hemlock, Tamarack, Maple, Beech, Spruce, and Elm. For both Pine and Oak, the variable Gaussian kernel models estimated higher values of γ. It is encouraging that these estimates of localness are similar for both variable case models, but based on these estimates alone we can compare only focal-cell contributions. Non-local dispersal is determined by the product of $(1-\gamma)$ with the sum of the weights of all non-local cells, which in turn depends on the estimated kernel parameter(s).

For the base Gaussian kernel case, the spread parameter ψ was estimated to be 204 (192, 215) km. In the variable Gaussian model ψ estimates ranged from 152 for Pine to 361 for Other hardwood (Figure **??**). For many taxa, 95% credible intervals are large; this is especially true for Tamarack, with a credible interval of (128, 553). For the base power law model, the kernel parameters a and b were estimated to be $0.01\,(8.29e-3, 0.01)$ and $2.03\,(2.00, 2.12)$. For the variable power law model case, the estimated b value was almost identical to the base model $2.06\,(2.00, 2.16)$. The a estimates ranged from $2.68e-3$ for Beech to 0.203 for Elm.

The most intuitive way to assess how these parameters affect pollen dispersal is to plot the proportion of deposited pollen as a function of radius from the pollen source. So, we plot $\gamma + (1-\gamma)\sum_{s_j \neq s_i} w(s(i), s_j)$ for a hypothetical core (i) centered on our potential discretized domain, where for a defined radius, cells are included in the non-local sum if the distance between $s(i)$ and s_j is less than or equal to this radius. Then, for each taxon, we can plot cumulative density functions of the deposited pollen for each of the four considered models (Figure **??**). We note that the plotted based model functions are identical for all taxa (solid lines). For the variable models, we can see the effects of changes in magnitude in both ψ for the Gaussian kernel, and a for the variable kernel. The CDFs for the Gaussian kernel models all have a sigmoidal shape, where smaller values of ψ correspond to a steeper slope indicating that relatively more pollen deposition happens close to the source. The contrast between the two most extreme ψ values is seen the Pine (lowest) and Other hardwood (highest) variable Gaussian CDFs. For the variable power law model, for fixed b, larger values of a result in a more sigmoidal shape that resembles the Gaussian kernel, for example as seen with Elm where the variable model CDFs are nearly identical. Again for fixed b, smaller values of a

results in a concave downward curve that shows that pollen is deposited closer to the source, but at a certain distance from the source - where the Gaussian and power law CDFs intersect - the cumulative pollen deposition is equal. To help interpret these curves, we determined the radius from the source needed to capture 50% and 90% of the deposited pollen for each model and each taxon (Table 5). As seen in Figure 8, the CDFs for the two base models have two points of intersection, indicating changing relative rates of accumulation. The power law based model estimates a smaller radius needed to accumulate 50% of the pollen, however, the Gaussian base model cdf asymptotes to one more quickly (shorter-tailed kernel). For the variable models, the taxon with the smallest radius from a pollen source needed to accumulate 50% of the dispersed pollen was Hemlock (60 km for the variable Gaussian, and 28 km for the the variable power law), while Other hardwood required the largest radius to accumulate the same percent (264 km for the variable Gaussian, 252 km for the variable power law). To capture 90% of the dispersed pollen required radii of 216 km for Pine up to 448 km for Other hardwood for the variable Gaussian model, and 292 km for Hemlock to 476 km for Elm and Other hardwood for the variable power law model.

To visualize the dispersal patterns predicted by our model, we can estimate the proportion of deposited pollen by taxon for each grid cell in the domain (previously we were estimating deposited pollen only at lakes for which we had pollen data) (Figures **??** & 10). These spatial maps help us understand how the models treat dispersal. Comparing these two figures illustrates the differences between these two models - the predicted sediment pollen maps for the variable power law model are much smoother. This is due to the increased accumulation of dispersed pollen at short distances relative to the Gaussian, which predicts that the fastest accumulation rate of dispersed pollen occurs farther from the source (see Figure 8).

We can also compare the predicted pollen maps with those depicting the relative tree abundance (Figure 11), remembering that the pollen predictions have been scaled to account for differential production (so we are comparing predicted sediment pollen with vegetation composition). This scaling for differential production is particularly apparent for Tamarack. We know from the data that Tamarack pollen proportions are always lower than the vegetation proportions, and as a result, we expect that the model will predict low amounts of Tamarack pollen, which it does. To see how the model predicts dispersal by taxon, based on parameters estimated from that data, we can compare the vegetation and pollen ranges. The effects of dispersal are especially clear in the Pine and Oak cases, where we see pollen dispersing well beyond the range limits of the vegetation range boundaries.

5 Summary and Conclusions

Kernel	Model	WAIC
Gaussian	Base	13115
Gaussian	Variable γ	13053
Gaussian	Variable ψ	13038
Gaussian	Variable ψ and γ	12953
Power law	Base	12853
Power law	Variable γ	12789
Power law	Variable a	12702
Power law	Variable a and b	b EPC
Power law	Variable a and γ	12690
Power law	Variable a, b and γ	b EPC

Table 1: Wantanabe-Akaike Information Criterion (WAIC) values for eight candidate models. EPC denotes the exchangeable prior collapse for the indicated parameter. Models that use the Power law kernel have lower WAIC values.

Common name	Scientific name	Model	
		γ (Variable Gaussian)	γ (Variable Power law)
Ash	Fraxinus	6.24E-2 (7.15E-3, 1.64E-1)	3.87E-2 (3.55E-4, 1.07E-1)
Beech	Fagus	3.12E-1 (1.29E-1, 5.35E-1)	1.69E-1 (1.06E-2, 4.90E-1)
Birch	Betula	9.90E-2 (3.71E-2, 1.58E-1)	2.77E-2 (5.69E-4, 7.82E-2)
Elm	Ulmus	1.48E-1 (7.79E-2, 2.22E-1)	1.65E-1 (8.70E-2, 2.38E-1)
Hemlock	Tsuga	4.63E-1 (3.67E-1, 5.53E-1)	2.70E-1 (8.76E-2, 4.52E-1)
Maple	Acer	1.83E-1 (8.79E-2, 2.84E-1)	1.98E-1 (7.70E-2, 3.01E-1)
Oak	Quercus	2.41E-1 (1.82E-1, 3.03E-1)	5.16E-2 (2.60E-3, 1.26E-1)
Other Conifer	Abies	8.12E-2 (5.21E-3, 1.96E-1)	4.16E-2 (3.23E-4, 1.46E-1)
Other Hardwood	* See caption	4.73E-2 (3.07E-3, 8.67E-2)	5.32E-2 (1.58E-3, 1.05E-1)
Pine	Pinus	2.49E-1 (2.14E-1, 2.82E-1)	2.11E-2 (3.12E-3, 4.86E-2)
Spruce	Picea	2.37E-1 (1.46E-1, 3.29E-1)	1.44E-1 (2.41E-2, 2.86E-1)
Tamarack	Larix	4.25E-1 (2.23E-1, 6.48E-1)	3.66E-1 (2.09E-2, 6.22E-1)

Table 2: Estimates of parameter γ, indicating the proportion of locally sourced pollen for the variable Gaussian model in which ψ and γ both vary by taxon and the variable Power law kernel model in which a and γ vary by taxon. For the base models, γ was estimate to be $0.21(1.96E-1, 2.30E-1)$ for the Gaussian kernel model and $0.07(3.79E-2, 9.53E-2)$ for the Power law kernel model. * The Other Hardwood grouping is composed of: Alnus, Carya, Cornus, Cupressaceae, Gleditsia, Liquidambar, Maclura, Morus, Nyssa, Ostrya carpinus, Platanus, Populus, Robinia, Rosaceae, Salix, Tilia, Zanthoxylum, as well as any undifferentiated hardwood (e.g. pollen grains classified as Fagus backslashchar Nyssa, Betula backslashchar Corylus, etc.).

Taxon	Model			
	Gaussian		Power law	
	50%	90%	50%	90%
Hemlock	60	280	28	292
Tamarack	96	328	120	428
Pine	100	216	56	332
Birch	120	228	104	384
Beech	140	340	32	300
Oak	140	304	104	380
Other conifer	144	276	124	396
Spruce	168	356	132	408
Maple	184	368	152	424
Ash	228	404	208	452
Elm	244	440	236	476
Other hardwood	264	448	252	476

Table 3: Radii (km) from a pollen source needed to capture 50% and 90% of the dispersed pollen for the variable gaussian and variable power law models. The gaussian base model estimated radii of 136 and 292 km to capture 50% and 90% of the dispersed pollen across all taxa, while the power law model estimated radii of 92 and 376 km.

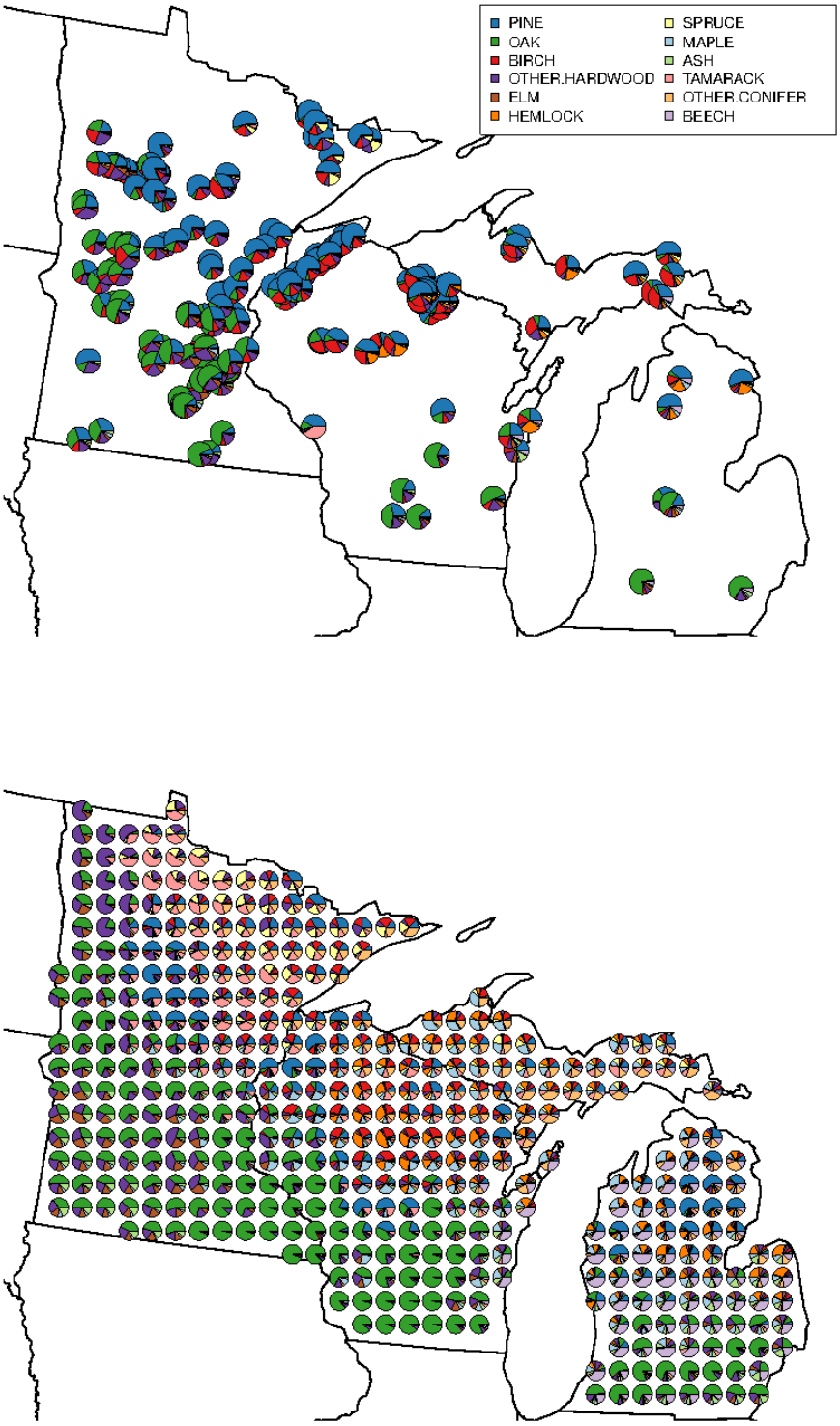

Figure 1: Pie maps depicting the relative composition of tree genera of pollen (top) and PLS vegetation (bottom) from the data. Note that the PLS data has been aggregated from 8 km to 34 km resolution for illustrative purposes.

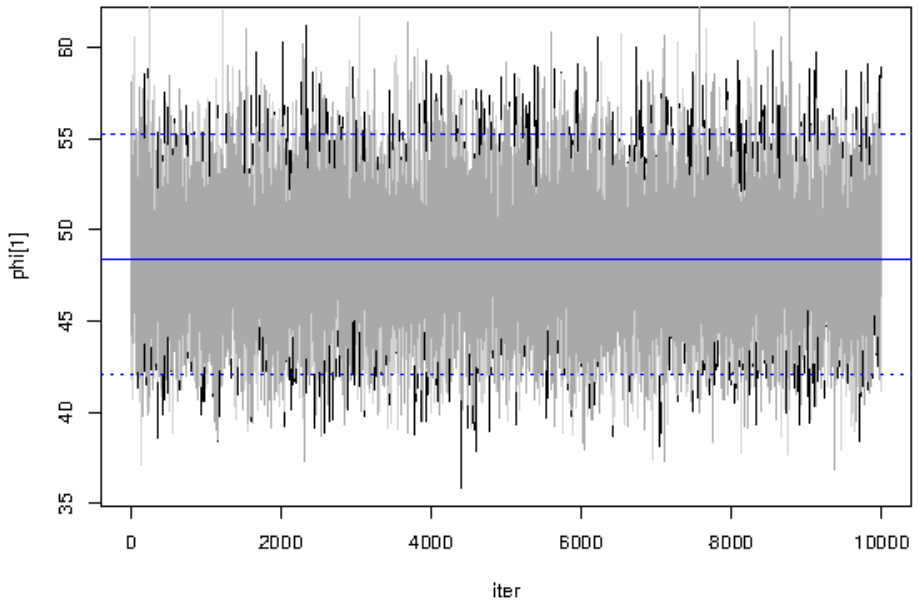

Figure 2: Trace plots for the parameters in the gaussian dispersed calibration model, including ϕ_k for $k = 1, \ldots, W$, ψ, and γ, as well as the joint log posterior, lp. Posterior estimates for each of the three runs are depicted in shades of grey. The 2.5% and 97.5% quantiles are indicated by the solid lines, while the 50% quantiles are indicated by the hashed lines.

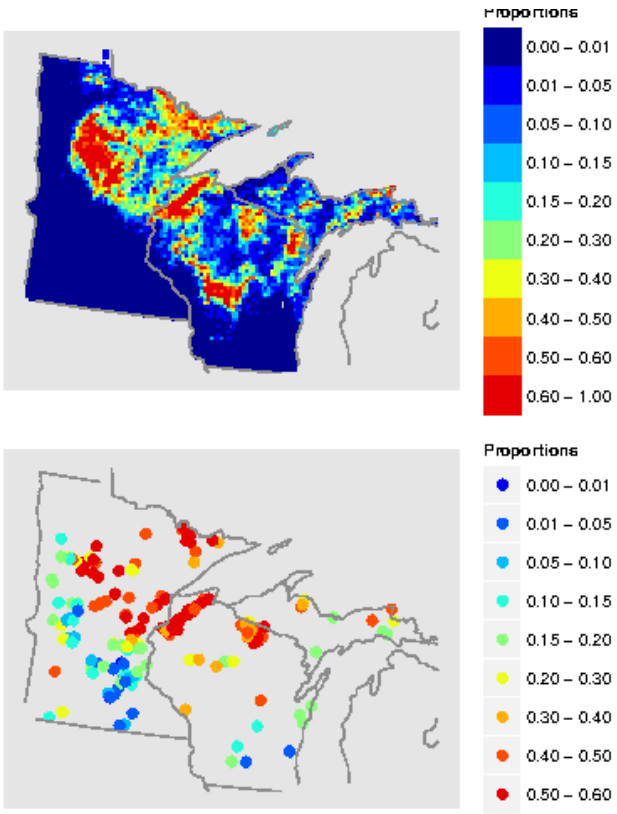

Figure 3: Heat maps showing the relative abundances of pine in the PLS composition data (top) and the pre-settlement era pollen data (bottom).

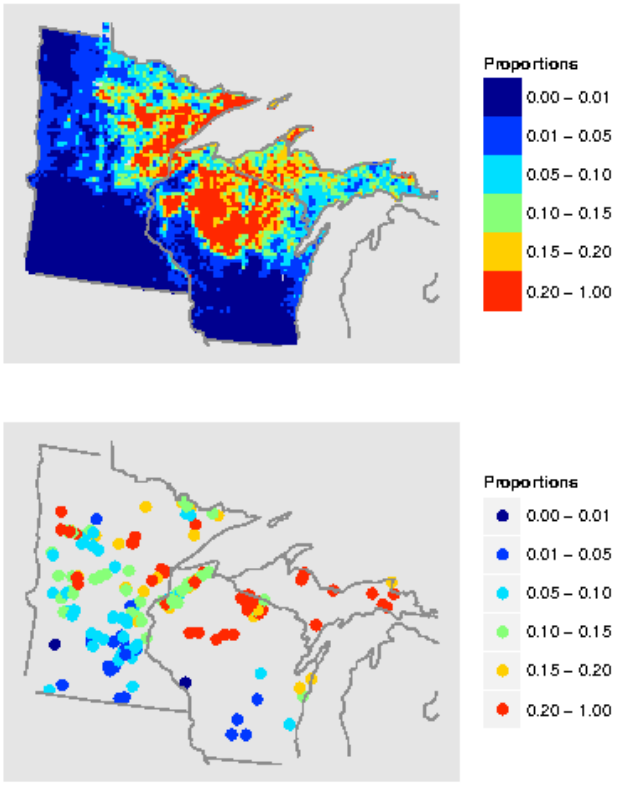

Figure 4: Heat maps showing the relative abundances of birch in the PLS composition data (top) and the pre-settlement era pollen data (bottom).

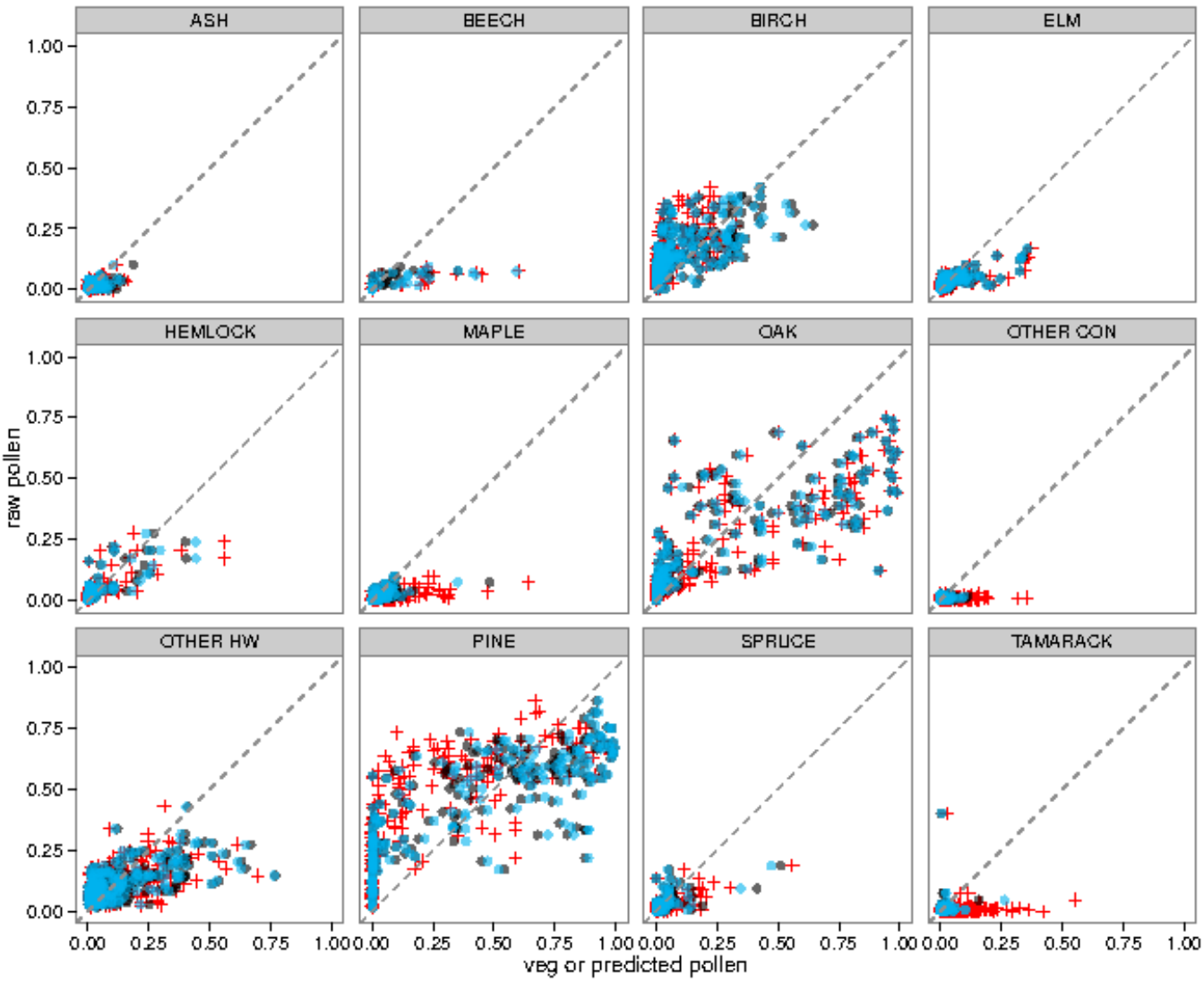

Figure 5: Pollen proportions plotted against local vegetation proportions (red crosses) or local vegetation proportion scaled by the pollen production factorϕ for the variable Gaussian kernel model (black dots) and the variable Power law kernel model (blue dots).

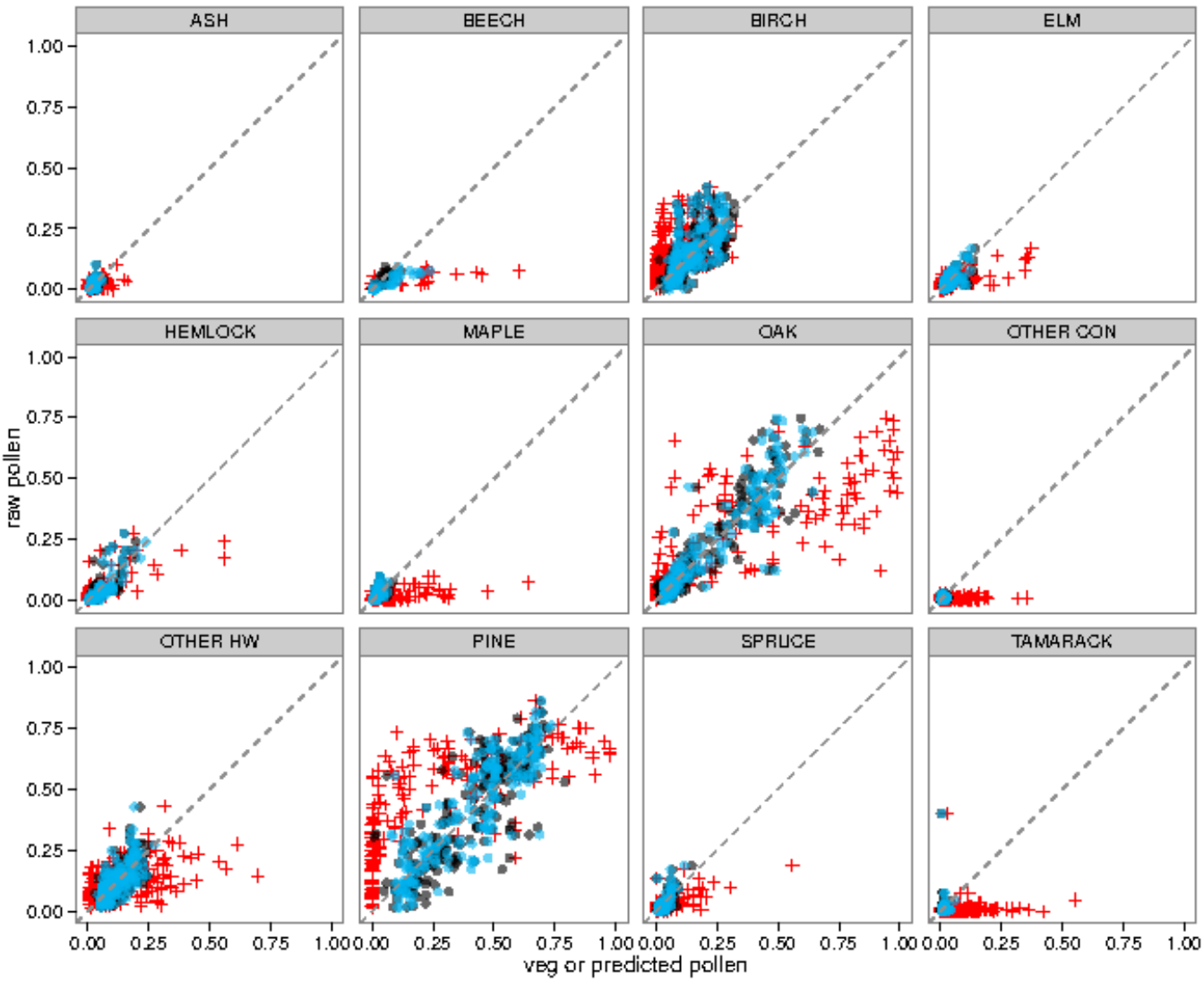

Figure 6: Pollen proportions plotted against local vegetation proportions (red crosses) or model-predicted pollen for the variable Gaussian kernel model (black dots) and the variable Power law kernel model (blue dots).

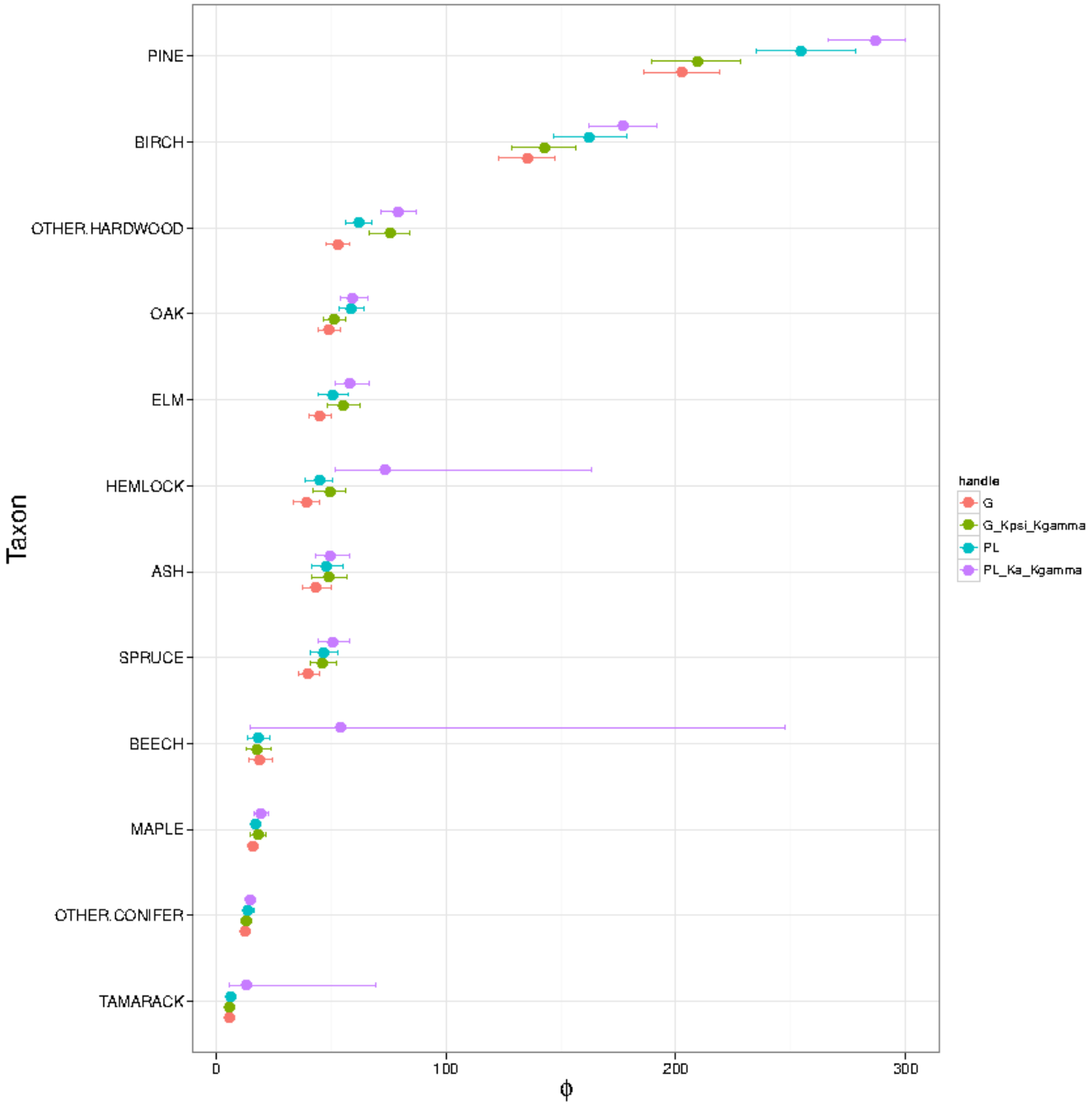

Figure 7: Mean values and 95% credible intervals for the estimated values of the differential production parameter ϕ.

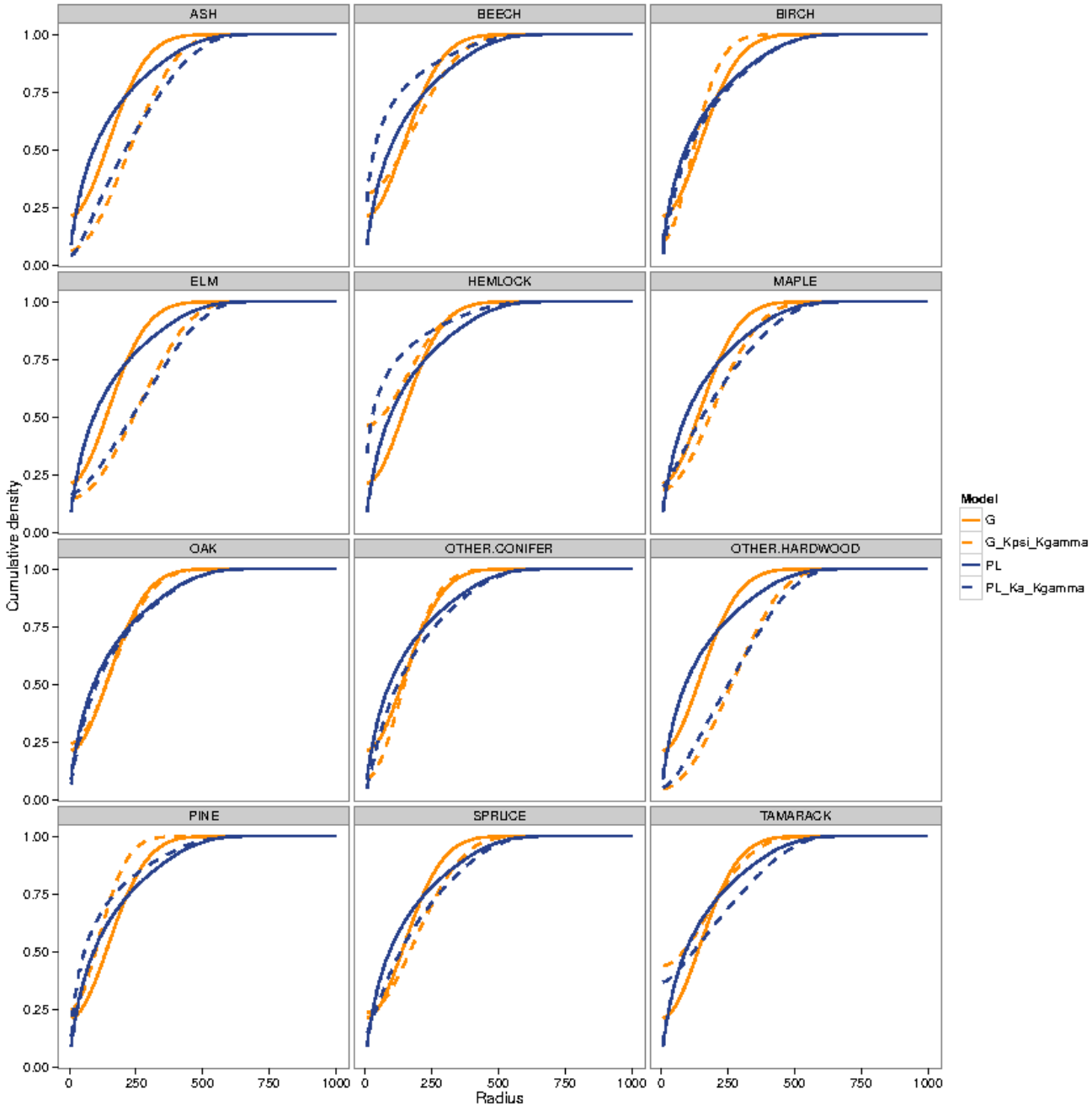

Figure 8: The proportion of deposited (or accumulated) pollen plotted as a function of radius for each of the modelled taxa and for each of the four considered models (Gaussian base and variable; Power law base and variable).

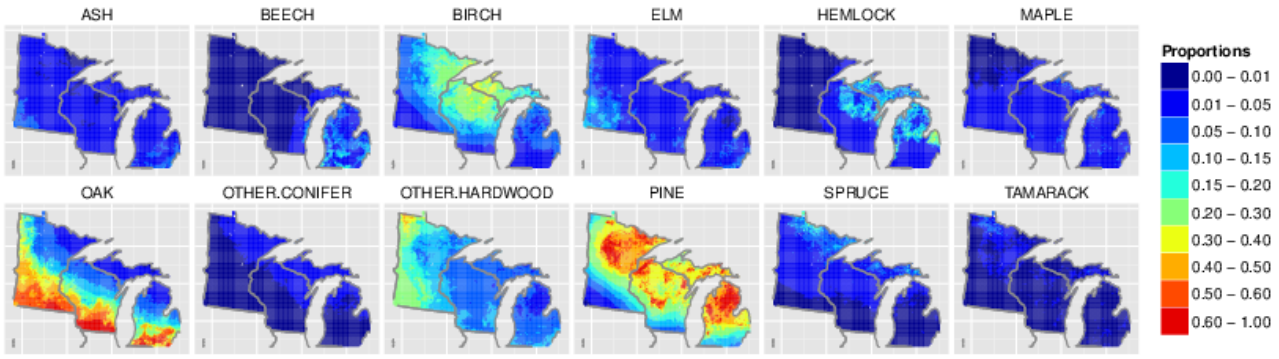

Figure 9: Heat maps of predicted pollen relative abundances from the variable Gaussian model for each grid cell in the domain, by taxon.

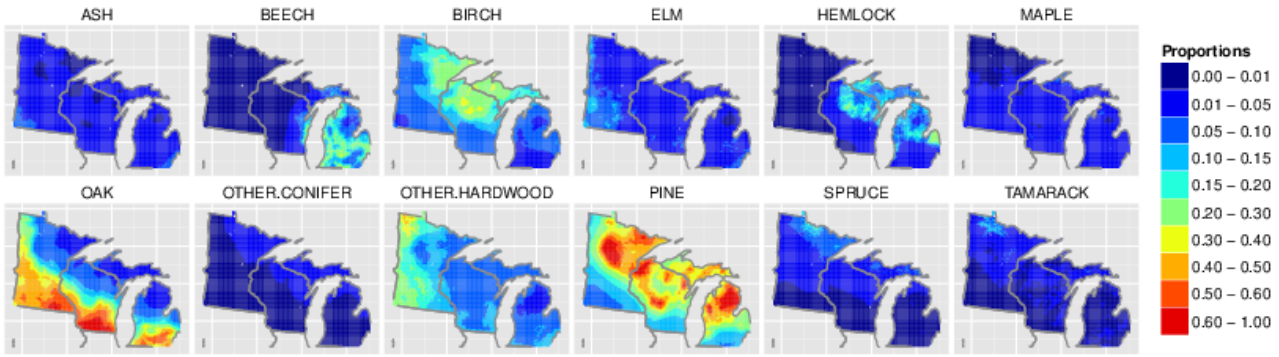

Figure 10: Heat maps of predicted pollen from the variable power law model for each grid cell in the domain, by taxon.

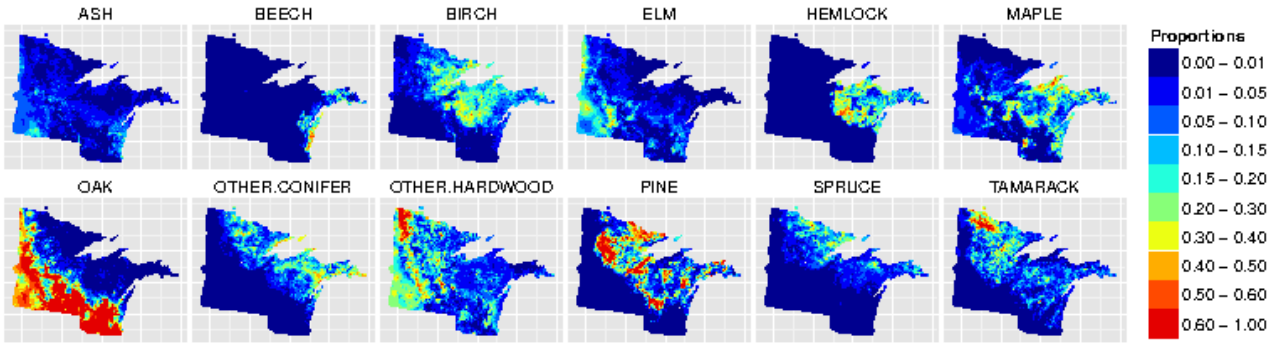

Figure 11: Heat maps of the PLS data, by taxon.

References

Frederic Austerlitz, Christopher W Dick, Cyril Dutech, Etienne K Klein, Sylvie Oddou-Muratorio, Peter E Smouse, and Victoria L Sork. Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology*, 13(4):937–954, 2004.

Maarten Blaauw and J Andrés Christen. Radiocarbon peat chronologies and environmental change. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(4): 805–816, 2005.

Maarten Blaauw, J Andrés Christen, et al. Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis*, 6(3):457–474, 2011.

Jessica L Blois, John W Williams, Eric C Grimm, Stephen T Jackson, and Russell W Graham. A methodological framework for assessing and reducing temporal uncertainty in paleovegetation mapping from late-Quaternary pollen records. *Quaternary Science Reviews*, 30(15):1926–1939, 2011.

Eric A Bourdo. A review of the general land office survey and of its use in quantitative studies of former forests. *Ecology*, pages 754–768, 1956.

Caitlin E Buck, J Andrés Christen, and Gary N James. BCal: an on-line Bayesian radiocarbon calibration tool. *Internet Archaeology*, 7, 1999.

Jean Clobert, Michel Baguette, Tim G Benton, James M Bullock, and Simon Ducatez. *Dispersal ecology and evolution*. Oxford University Press, 2012.

Wolfgang Cramer, Alberte Bondeau, F Ian Woodward, I Colin Prentice, Richard A Betts, Victor Brovkin, Peter M Cox, Veronica Fisher, Jonathan A Foley, Andrew D Friend, Chris Kucharik, Mark R Lomas, Navin Ramankutty, Stephen Sitch, Benjamin Smith, Andrew White, and Christine Young-Molling. Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models. *Global Change Biology*, 7(4):357–373, 2001.

C Devaux, C Lavigne, F Austerlitz, and EK Klein. Modelling and estimating pollen movement in oilseed rape (*Brassica napus*) at the landscape scale using genetic markers. *Molecular Ecology*, 16(3):487–499, 2007.

Vincent Garreta, Paul A Miller, Joël Guiot, Christelle Hély, Simon Brewer, Martin T Sykes, and Thomas Litt. A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model. *Climate Dynamics*, 35(2-3):371–389, 2010.

Andrew Gelman, Jessica Hwang, and Aki Vehtari. Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6):997–1016, 2014.

Simon Goring, Andria Dawson, Gavin L Simpson, Karthik Ram, Russ W Graham, Eric C Grimm, and John W Williams. neotoma: A programmatic interface to the Neotoma paleoecological database. *Open Quaternary*, 2015. doi: http://doi.org/10.5334/oq.ab.

Simon Goring et al. Xxx. *XXX*, XXX:XXX, in prep.

Simon Goring et al. Xxx. *XXX*, XXX:XXX, in review.

Matthew D Hoffman and Andrew Gelman. The No-U-Turn Sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15: 1351 – 1381, 2011.

Stephen T Jackson. Pollen source area and representation in small lakes of the northeastern United States. *Review of Palaeobotany and Palynology*, 63(1):53–76, 1990.

Stephen T Jackson. Pollen and spores in Quaternary lake sediments as sensors of vegetation composition: theoretical models and empirical evidence. In A. Traverse, editor, *Sedimentation of organic particles*, pages 253–286. Cambridge University Press, Cambridge, 1994.

Stephen T Jackson and Mark E Lyford. Pollen dispersal models in Quaternary plant ecology: assumptions, parameters, and prescriptions. *The Botanical Review*, 65(1):39–75, 1999.

Ellen Kujawa et al. Xxx. *XXX*, XXX:XXX, in prep.

R.W. Lawrenz. *The developmental paleoecology of Green Lake, Antrim County, Michigan*. PhD thesis, Central Michigan University, 1975.

J Matthes et al. Xxx. *XXX*, XXX:XXX, in review.

John H McAndrews. Human disturbance of North American forests and grasslands: the fossil pollen record. In *Vegetation History*, pages 673–697. Springer, 1988.

Christopher J Paciorek and Jason S McLachlan. Mapping ancient forests: Bayesian inference for spatio-temporal trends in forest composition using the fossil pollen proxy record. *Journal of the American Statistical Association*, 104(486):608–622, 2009.

Colin Prentice. Records of vegetation in time and space: the principles of pollen analysis. In *Vegetation history*, pages 17–42. Kluwer Academic Publishers, Dordrecht, 1988.

IC Prentice and T Webb. Pollen percentages, tree abundances and the Fagerlind effect. *Journal of Quaternary Science*, 1(1):35–43, 1986.

Christopher Bronk Ramsey. Radiocarbon calibration and analysis of stratigraphy; the OxCal program. *Radiocarbon*, 37(2):425–430, 1995.

Paula J Reimer, Edouard Bard, Alex Bayliss, J Warren Beck, Paul G Blackwell, Christopher Bronk Ramsey, Caitlin E Buck, Hai Cheng, R Lawrence Edwards, Michael Friedrich, et al. IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. *Radiocarbon*, 55(4):1869–1887, 2013.

Lisa A Schulte and David J Mladenoff. The original US public land survey records: their use and limitations in reconstructing presettlement vegetation. *Journal of Forestry*, 99(10): 5–10, 2001.

Allen M Solomon and Thompson Webb. Computer-aided reconstruction of late-Quaternary landscape dynamics. *Annual Review of Ecology and Systematics*, pages 63–84, 1985.

JM St Jacques, BF Cumming, DJ Sauchyn, and JP Smol. The bias and signal attenuation present in conventional pollen-based climate reconstructions as assessed by early climate data from Minnesota, USA. *PloS one*, 10(1):e0113806–e0113806, 2014.

Stan Development Team. Stan: A C++ library for probability and sampling, Version 2.4, 2014. URL `http://mc-stan.org/`.

Shinya Sugita. Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. *The Holocene*, 17(2):229–241, 2007a.

Shinya Sugita. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. *The Holocene*, 17(2):243–257, 2007b.

H Tauber. Differential pollen dispersion and the interpretation of pollen diagrams. Technical Report 89, 1965.

Greame K Ward and SR Wilson. Procedures for comparing and combining radiocarbon age determinations: a critique. *Archaeometry*, 20(1):19–31, 1978.

Sumio Watanabe. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *The Journal of Machine Learning Research*, 11:3571–3594, 2010.

Thompson Webb. The late-and post-glacial sequence of climatic events in wisconsin and east-central minnesota: Quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis, 1971.

TIII Webb, SE Howe, RHW Bradshaw, and KM Heide. Estimating plant abundances from pollen percentages: the use of regression analysis. *Review of Palaeobotany and Palynology*, 34(3):269–300, 1981.

John W Williams. Variations in tree cover in north america since the last glacial maximum. *Global and Planetary Change*, 35(1):1–23, 2003.

H.E. Wright, W.A. Watts, S. Jelgersma, J.C.B. Waddington, J. Ogawa, and T.C. Winter. SP-11 glacial and vegetational history of northeastern Minnesota. Technical report, Minnesota Geological Survey, 1969.