Quaternary Science Reviews 31 (2012) 17–29

Contents lists available at SciVerse ScienceDirect

Quaternary Science Reviews

journal homepage: www.elsevier.com/locate/quascirev

Invited review

Review of probabilistic pollen-climate transfer methods

Christian Ohlwein [a,*], Eugene R. Wahl [b]

[a] Meteorological Institute, University of Bonn, Auf dem Hügel 20, 53121 Bonn, Germany
[b] National Oceanic and Atmospheric Administration, National Climatic Data Center (NOAA-NCDC), Paleoclimatology Branch/World Data Center for Paleoclimatology, Boulder, CO, USA

ARTICLE INFO

Article history:
Received 2 May 2011
Received in revised form
2 November 2011
Accepted 3 November 2011
Available online 16 December 2011

Keywords:
Paleoclimate reconstructions
Transfer functions
Fossil pollen
Lake sediments
Wisconsin
Southern California
North America
Bayesian modeling
BHM
Indicator taxa
Mutual climatic range
Modern analogues
Response surfaces
Plant functional types
Biomisation
Ensemble post-processing
Generalized linear model

ABSTRACT

Pollen-climate transfer methods are reviewed from a Bayesian perspective and with a special focus on the formulation of uncertainties. This approach is motivated by recent developments of spatial multi-proxy Bayesian hierarchical models (BHM), which allow synthesizing local reconstructions from different proxies for a spatially complete picture of past climate. In order to enhance the pollen realism in these models we try to bridge the gap between spatial statistics and paleoclimatology and show how far classical pollen-climate transfer concepts such as regression methods, mutual climatic range, modern analogues, plant functional types, and biomes can be understood in novel ways by refining the data models used in BHMs. As a case study, we discuss modeling of uncertainty by introducing a new probabilistic pollen ratio model, which is a simplified variation of the modern analogue technique (MAT) including the concept of response surfaces and designed for later inclusion in a spatial multiproxy BHM. Applications to fossil pollen data from varved sediments in three nearby lakes in west-central Wisconsin, USA and for a Holocene fossil pollen record from southern California, USA provide local climate reconstructions of summer temperature for the past millennium and the Holocene respectively. The performance of the probabilistic model is generally similar in comparison to MAT-derived reconstructions using the same data. Furthermore, the combination of co-location and precise dating for the three fossil sites in Wisconsin allows us to study the issue of site-specific uncertainty and to test the assumption of ergodicity in a real-world example. A multivariate ensemble kernel dressing approach derived from the post-processing of climate simulations reveals that the overall interpretation based on the individual reconstructions remains essentially unchanged, but the single-site reconstructions underestimate the overall uncertainty.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past century numerous approaches to paleoclimate reconstruction have been made. Due to different types of proxy data, different reconstruction methods, and spatially and temporally inhomogeneous distribution of paleo archives, obtaining a coherent picture of past climate variability remains a complicated task. In order to address this problem, we identify two common goals when analyzing paleoclimate data, which are to:

- combine proxies for a spatially complete picture of past climate in order to inform physical understanding of the climate system in terms of dynamics and response to forcings; and
- estimate the uncertainty in climate reconstructions, which is naturally large due to the approximation of a complex proxy-climate relation.

A significant advance in the attempt to provide a spatially complete picture of past climate has been the development of systematic climate field reconstruction methods, or CFRs. CFRs were first developed from eigenvector/singular value empirical orthogonal function (EOF) methods used in modern instrumental climatology. In paleoclimatology, several variants of EOF-based CFRs have been employed, primarily with annually resolved proxy data (of multiple kinds) over the past two millennia: all use the proxy information to estimate the temporal loadings of a truncated basis set of the EOFs of the instrumental climate field of interest to be reconstructed back in time. Once established through direct (e.g. Luterbacher et al., 2004) or inverse regression relationships (e.g. Mann et al., 1998, 1999; Wahl and Ammann, 2007),

* Corresponding author.
E-mail address: christian.ohlwein@gmx.net (C. Ohlwein).

0277-3791/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.quascirev.2011.11.002

these estimated loadings are then applied to the retained EOF patterns (through the appropriate expansion formula for EOF derivation) during both the instrumental and reconstruction periods to form estimated values of the entire climate field, conditional on the proxy information. A variant of this method uses the regularized expectation maximization (RegEM) algorithm to estimate the temporal loadings from the proxy information, and then reconstructs the EOFs as described (e.g., Mann et al., 2009); RegEM can also be used to impute the reconstruction field directly (e.g., Rutherford et al., 2005; Smerdon et al., 2010). A closely-related method forms a truncated basis set from the joint proxy-instrumental cross-covariance for the same purpose in canonical correlation analysis (CCA) (cf. Cook et al., 1994; Smerdon et al., 2010). Although it is not a systematic field reconstruction method per se, the point-by-point regression (PPR) method developed by Cook et al. (2004, 2007, 2010) has been used successfully to reconstruct fields of the Palmer Drought Stress Index (PDSI), a key hydroclimate variable that integrates both temperature and precipitation in a soil moisture model, in both North America (2004, 2007) and Monsoon Asia (2010).[1]

CFR methods are a large step forward in reconstructing past climate in a spatially complete manner, however in some cases they have not been accompanied by systematic estimation of uncertainty at the field level, as opposed to index values of regional, hemispheric or global means (e.g., Luterbacher et al., 2004; Mann et al., 2008, 2009). Mann et al. (2009) used the unexplained variability in the validation period multiplied by the calibration period decadal variance at each grid cell as a measure of the squared decadal standard error of reconstruction. Gebhardt et al. (2008) present a different approach where a probabilistic data-based method for local reconstructions is combined with a dynamic constraint on the reconstructed climate parameter, which leads to climatological fields being optimized with respect to both the proxy data and to the prescribed dynamics in a statistically consistent way. Wahl and Smerdon (submitted) apply a simplified version of a technique for generating probabilistic ensemble draws of spatial mean temperature reconstructions (Li et al., 2007) to one of the truncated EOF CFR methods described above; thereby generating ensemble draws of the estimated temporal loadings associated with the retained EOFs and in turn yielding a probabilistic field reconstruction ensemble for the target domain of western North America. This method, while a valuable step forward, still underestimates the full uncertainty inherent in the reconstruction process because it does not include a statistical model for uncertainty in the proxy data. A more complete characterization of uncertainty would involve developing such models and incorporating them into the reconstruction process, to which we now turn.

At the forward edge of these methods, with special focus on the common goals of incorporating multiple proxies and quantifying uncertainties in a rigorous manner, is, e.g., Li et al. (2010). They developed a Bayesian hierarchical model (BHM) to reconstruct the northern hemisphere mean temperature from different sources, such as proxies with different temporal resolution as well as external forcings such as solar irradiance, volcanic aerosols, and greenhouse gas concentrations. Their method is applied to synthetic climate data taken from a global climate model and synthetic proxy data also derived from the simulation. Tingley and Huybers (2010) provide an important extension to this approach by reconstructing space–time temperature processes.

However, as mentioned, these methods have been applied to pseudo-proxies. On the one hand, this is a well-justified and natural approach to test the capability of a BHM for paleoclimate reconstructions, as emphasized in the comments on Li et al. (2010) by Cressie and Tingley. It allows estimating the characteristics with respect to spatio-temporal processes before incorporating more complex climate-proxy relations. Similar tests a have been performed, e.g., by von Storch et al. (2009) where they compare three different reconstruction methods, the inverse regression method by Mann et al. (1998), a direct principal components regression, and the composite plus scaling method, in the virtual reality of the ECHO-G climate simulation. The result, that the methods display different performance depending on the noise model and the size of the pseudo-proxy network, shows that test-bed studies are an important informative tool. This utility is confirmed in a growing body of reconstruction simulation experiments (e.g., Smerdon et al., 2010).

On the other hand, additional uncertainty arises from the question how well the test-bed studies relate to the performance of the same methods using real data. It should be noted that the way pseudo-proxies are created provides a degree of caution to the application of an assumption that test-bed results and real-data results using the same method and targeting the same spatial domain will necessarily strongly parallel each other, even if the test-bed climate realistically represents the real-world climate. This is an issue that needs to be addressed in applying these approaches to fossil proxy data, including the main focus of this article, which is on pollen. Although the aforementioned BHMs account for different noise models, the underlying assumption is a linear relation between proxies and temperature. We discuss the aspect of linearization for a limited case in Section 3. However, glancing at the large number of approaches to pollen-based climate reconstruction, or so-called transfer functions (Section 2), it is obvious that a linear link between pollen counts and environmental variables such as near surface temperature is not generally applicable for pollen as a paleo-proxy. From these observations come the idea and motivation for this study, which is to enhance the realism of spatial hierarchical models for use of pollen proxies in paleoclimate reconstruction.

The general challenge in pollen-based paleoclimate reconstructions is that there is not a strong direct functional or mechanistic relationship between pollen spectra and climatic variables. This complexity arises because pollen production is affected by the interaction of a large number of nonlinear processes. Let us consider the ideal case of a single tree species and its dependence on one climate parameter, e.g. average near surface temperature. In a very simplified picture, one might think of a unimodal response of pollen counts on temperature, describing a maximum climatic range with increasing pollen production toward optimum conditions somewhere within. While this assumption appears to be appropriate for the abundances of whole organisms (such as chironomids, used in methods related to pollen-based paleoenvironmental reconstruction), it should be noted that pollen production is a sub-organism-level, secondary metabolic process—each grain containing the male gamete component of plant reproductive biology. It includes irregularities like mast years, during which plants produce heavy seed crops, while in other years there is no, poor, or only moderate seed production. Furthermore, climate-pollen relations exhibit a second order effect beyond vegetation

[1] A number of criteria have been used to determine the retained EOF/CCA spatial patterns; the reader is pointed to the above references and citations therein for more information. It is important to note that stationarity of the retained EOF/CCA basis set over time is a key assumption in the truncated EOF/CCA methods described, along with the temporal stationarity of the proxy-EOF/CCA temporal loading relationship. The former assumption has been examined in a paleoclimate model simulation environment by Wahl and Smerdon (submitted), who found it reasonable based on comparison of 11th–20th century global temperature EOFs 1–5 to 17th century (coolest) and 20th century (warmest) EOFs 1–5. The entire-grid RMSEs of these comparisons were larger than parallel comparisons to EOFs 1–5 generated from 1000 random-century draws only 8% of the time, concentrated in global EOF2 (4%). Stationarity of point-wise proxy/PDSI relationships over time is also assumed in the PPR method.

differences: in some situations pollen production under stress, i.e. at the edge of a plant's climatic range, might even be higher than average. We do not address tertiary effects on local pollen assemblages such as wind-drift, but a key process that needs to be accounted for is plant competition, which makes it difficult to treat different taxa independently. In addition, the combination of taxa itself can be disturbed by cultivation of land, leading to so-called no-modern-analogue situations. A concrete example is given in Neumann et al. (2007) for the Eastern Mediterranean, which is heavily prone to cultivation during the late Holocene. We will detail the basic concepts of pollen-based climate reconstructions from a more paleobotanical perspective in Section 2, but the brief considerations above indicate how much pollen-based climate reconstructions naturally involve multivariate, nonlinear processes and considerable uncertainties.

Uncertainties in paleoclimate transfer methods can also be understood from a more fundamental point of view. In climate science and risk management the terms epistemic and aleatory uncertainty are used (e.g. Troccoli et al., 2008). Epistemic uncertainty follows from limited process knowledge or modeling capabilities and relates to most of the problems addressed in the previous paragraph. Aleatory uncertainty is given by the stochastic nature of the climate system and its subsystems, atmosphere, ocean, cryosphere, biosphere, and lithosphere, and hence is not reducible. Using properties of one subsystem to deduce properties of another subsystem, as is done by botanical-climatological transfer methods for instance, naturally involves random processes. This motivates the recent trend from a deterministic point of view toward probabilistic methods where traditional transfer functions are understood as joint or conditional probability distributions (e.g. Kumke et al., 2004). A very natural framework to deal with this point of view, i.e. various levels of random effects, is a BHM, which leads to the objectives of this article. These new and upcoming approaches, as mentioned above, are typically derived from the perspective of spatio-temporal statistics, rather than from pollen-based research, and there is therefore a need to re-consider classical pollen-climate transfer functions (as far as this term still holds) in the development of this wider, probabilistic perspective.

In Section 2 we discuss the background of pollen-climate transfer functions in the context of a Bayesian framework in order synthesize the existing approaches. A case study is given in Section 3 by developing a probabilistic version of the pollen-ratio method for potential inclusion in a spatial multiproxy BHM. Finally we go beyond single-site reconstructions and analyze uncertainty in reconstruction from three nearby lake sediment records, by the means of ensemble post-processing (Section 4).

2. Probabilistic pollen-climate transfer methods

2.1. Bayesian framework

The history of pollen-climate transfer methods dates back at least 65 years, leading to a number of different approaches. A general synthesis is given in Bartlein et al. (2010) and Birks et al. (2010). Since most of the methods have been developed from a point of view focused on the paleo archives, they often employ mechanistic relations or expectation values. In order to put these into the general aims of this paper, in particular to account for the stochastic nature of both the climate and proxy systems, we discuss how far recent approaches already involve a continuous formulation of random effects (Schölzel, 2006; Garreta, 2010). For this purpose we introduce three, and later four, multivariate random variables: the climate variables $\mathbf{C}$, such as averaged near-surface temperature, annual precipitation, mean temperature of the coldest month, etc.; the pollen spectra $\mathbf{P}$, considered here as raw counts of, typically, many pollen types (percentage and proportion representation are also common representations of pollen data); and a set of parameters $\boldsymbol{\Theta}$ that characterize the statistical model. Later we add an optional random variable V to express the intermediate vegetation step used by the different pollen-climate transfer methods.

In a general sense, the reconstruction problem is the joint probability density function $[\mathbf{P},\mathbf{C},\boldsymbol{\Theta}]$. A well-known method to split this complicated model into separate levels is a BHM (e.g. Li et al., 2010):

$$[\mathbf{P},\mathbf{C},\boldsymbol{\Theta}] = [\mathbf{P}|\mathbf{C},\boldsymbol{\Theta}][\mathbf{C}|\boldsymbol{\Theta}][\boldsymbol{\Theta}] \tag{1}$$

where $[\cdot|\cdot]$ denotes the conditional probability density. The first level in the hierarchy on the right-hand side of equation (1) is the data stage $[\mathbf{P}|\mathbf{C},\boldsymbol{\Theta}]$, which models the statistical relation between pollen proxies and climate variables. The remaining two levels are the process stage $[\mathbf{C}|\boldsymbol{\Theta}]$, modeling the (space–time) climate process, and the prior distributions $[\boldsymbol{\Theta}]$ (described in Section 2.7). The numbered equations here are designed to conform with those described in Li et al. (2010). The import of Bayes' rule becomes more obvious when plugging the left-hand side of equation (1) into the form $[\mathbf{C},\boldsymbol{\Theta}|\mathbf{P}] = [\mathbf{P},\mathbf{C},\boldsymbol{\Theta}]/[\mathbf{P}]$. Since $[\mathbf{P}]$ is the marginal distribution of the data, the Bayesian formulation of (1) can thus be stated as $[\mathbf{C},\boldsymbol{\Theta}|\mathbf{P}] \propto [\mathbf{P},\mathbf{C},\boldsymbol{\Theta}]$, indicating how the joint distribution of climate (the reconstruction target) and the model parameters, conditional on the proxies, is proportional to the product of the data stage, process stage, and prior distributions. Depending on the space-time processes to be modeled the specification of the different layers and sampling from the posterior distribution can be complicated. We omit detailed description here and refer to Li et al. (2010) and also to Tingley and Huybers (2010) for an improved model with respect to the spatio-temporal processes. Tingley et al. (2010) provide a comprehensive description of the palaeoclimate reconstruction problem using the language and tools of modern statistics.

As pointed out in the comments on Li et al. (2010), a BHM are is a powerful tool to account for various sources of uncertainties, but all these uncertainties have to be expressed through probability distributions. Following Section 1, this especially applies to the data stage. In terms of the purposes of this paper, the question can be put in terms of how far $[\mathbf{P}|\mathbf{C},\boldsymbol{\Theta}]$ can be represented by the existing pollen-climate transfer methods.

For a more complete characterization of relevant uncertainties let us introduce the additional vegetation step:

$$[\mathbf{P},\mathbf{V},\mathbf{C},\boldsymbol{\Theta}] = [\mathbf{P}|\mathbf{V},\mathbf{C},\boldsymbol{\Theta}][\mathbf{V}|\mathbf{C},\boldsymbol{\Theta}][\mathbf{C}|\boldsymbol{\Theta}][\boldsymbol{\Theta}] \tag{2}$$

where the dependence of pollen counts on climate is split into the conditional density $[\mathbf{V}|\mathbf{C},\boldsymbol{\Theta}]$, which describes the dependence of vegetation (e.g. taxa, biomes, analogues, etc.) on climate, and $[\mathbf{P}|\mathbf{V},\mathbf{C},\boldsymbol{\Theta}]$ which characterizes the link between pollen and vegetation (cf. Sugita, 2007). Under the simplification that pollen production is conditionally independent of climate given the vegetation state, the density $[\mathbf{P}|\mathbf{V},\mathbf{C},\boldsymbol{\Theta}]$ can be reduced to $[\mathbf{P}|\mathbf{V},\boldsymbol{\Theta}]$. The latter assumption requires that vegetation and climate are in equilibrium, which is not necessarily true but is implicitly or explicitly incorporated in several pollen-climate transfer methods; as a corollary, it also means that short-term changes in climate that do not affect the vegetation composition cannot affect pollen productivity. Nearly all the following methods can be characterized by the way this additional random variable V and the conditional densities are defined, allowing a useful, general way of comparing and evaluating the methods.

2.2. Regression methods (linear)

Some of the first mathematical transfer methods for using pollen to reconstruct climate were linear multiple regression

functions (e.g., Bernabo, 1981; Howe and Webb, 1983; Bartlein et al., 1984), in deterministic inverse model forms with stochastic errors $\mathbf{C} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{P}' + \boldsymbol{\varepsilon}$, ($\mathbf{P}'$ represents pollen abundances of the taxa included in the model or transformations of these abundances to linearize them in relation to a climate parameter of interest). These models could exhibit high explained variability in calibration (e.g., Bartlein et al., 1984; Gajewski, 1988; Bartlein and Whitlock, 1993), but were found to be non-general across climate and geographic space (e.g., Bartlein and Webb, 1985; Gajewski, 1988; Bartlein and Whitlock, 1993), potentially leading to degrees of arbitrariness in terms of defining appropriate regions for development of specific prediction equations (but see Bartlein and Webb, 1985; Bartlein and Whitlock, 1993 for use of quantitative analyses to define appropriate calibration regions) and requiring an additional assumption that such spatial definitions would hold over the relevant period of past being reconstructed (cf. Bartlein et al., 2010). This temporal threshold itself could vary in region-specific ways. For these reasons, and more generally because the method "proved too cumbersome", these kinds of transfer relationships were abandoned (Bartlein et al., 2010). It is interesting to note that some reconstructions developed by these methods are largely corroborated using the new Bayesian generalized linear model (GLM) framework presented below in section 3 (Wahl et al., in press, cf. Appendix), a result noted more generally by Bartlein et al. (2010). Classical prediction interval estimation in this context is described in Howe and Webb (1983) and good examples of its use in paleoenvironmental reconstruction are in the same article and Bartlein and Whitlock (1993). A good examination of conformance with/violation of the standard statistical assumptions associated with this kind of regression model is also provided in Howe and Webb (1983). Additional description and examination of this and other regression-based transfer methods, such as weighted-average methods, are presented in, e.g., ter Braak and Juggins (1993), Birks (1998), and Bartlein et al. (2010).

In terms of the aforementioned BHM, the additional vegetation step is omitted and a direct relation between pollen proxies and climate variables is assumed. The implementation of a linear regression into a BHM via $[\mathbf{P}|\mathbf{C},\boldsymbol{\Theta}]$ is straightforward—at least for the pollen-climate relation—whereas the corresponding space-time model can be complex. The assumption of a linear relation between pollen counts and climate is actually used in the test-bed studies by Li et al. (2010) and Tingley and Huybers (2010).

2.3. Indicator taxa and mutual climatic range

The indicator taxa approach dates back to Iversen (1944) and Hintikka (1963) and is based on the concept of defining subdomains in a bivariate climate space, also known as envelopes, that represent the bioclimatic range of a species. The climate reconstruction is a graphical method based on the combination of taxa existing in the paleo record, i.e. the overlap of all taxon-specific climate ranges. Grichuk (1969) extended the approach to closed subdomains and introduced the term mutual climatic range (MCR) method. In order to relate the graphical method to the Bayesian framework introduced above, the additional vegetation step is expressed by the presence of plant taxa $\mathbf{V} = (V_1,\ldots,V_m)$ where the corresponding pollen counts are required to be above a predefined threshold p_j:

$$[\mathbf{V}|\mathbf{C},\boldsymbol{\Theta}] = \Pi_j [V_j|\mathbf{C},\boldsymbol{\Theta}], \quad \text{for all } j \text{ with } P_j > p_j. \tag{3}$$

The classical MCR method does not account for pollen counts and hence $[V_j|\mathbf{C},\boldsymbol{\Theta}]$ equals one where the taxon is present in the climate space, e.g. obtained from the recent spatial distribution, and zero otherwise. Furthermore, conditional independence for the presence of different taxa given climate is assumed.

On the one hand, the graphical definition of the climatic ranges as sharp boundaries does not allow for uncertainty estimates, which occurs for two reasons. First, the original version of the MCR method implicitly assumes discrete distributions, i.e., uniform distributions in the univariate case. Second, these climatic ranges may have an arbitrary shape, mostly indicating dependence on the recent climate distribution, not only on the physical bioclimatic range of the respective taxa. Enhancements of the MCR concept are given by Pross et al. (2000) and Klotz et al. (2004), with so-called probability mutual climatic spheres (PCS). These enhancements address the common problem of non-overlapping mutual climatic ranges for a high number of taxa, which is caused by the previously mentioned overfitting of uniform distributions.

On the other hand, the approach can be extended by introducing parametric probability densities of the taxon-specific likelihoods $[V_j|\mathbf{C},\boldsymbol{\Theta}]$. Recent applications of the so-called probabilistic indicator taxa approach or pdf method include bi- and tri-variate random vectors describing winter temperature, summer temperature and annual precipitation. The conditional probabilities $[V_j|\mathbf{C},\boldsymbol{\Theta}]$ can be estimated directly e.g., via generalized linear models (GLMs) with polynomials of continuous covariates (Schölzel, 2006) or indirectly via Bayes' theorem $[V_j|\mathbf{C},\boldsymbol{\Theta}] = [\mathbf{C}|V_j,\boldsymbol{\Theta}][V_j,\boldsymbol{\Theta}]/[\mathbf{C},\boldsymbol{\Theta}]$. In the latter case, $[\mathbf{C}|V_j,\boldsymbol{\Theta}]$ is described by a Gaussian mixture model (Kühl et al., 2002) or by a copula model (Kühl et al., 2007; Litt et al., 2009; Schölzel and Friederichs, 2008) in order to allow for non-Gaussian marginal distributions, e.g., precipitation values. Parameter estimates are typically obtained from taxonomic atlases (Schölzel et al., 2002) and gridded climate data sets, following the assumption of ergodicity, which means that the average long-term state of the relation between a particular kind of vegetation and climate at a single location is the same as its average state across different locations at a single time (Dieckmann et al., 2000).

Thus, probabilistic versions of the classical indicator taxa or mutual climatic range concept can be directly incorporated into Eq. (2). Since this approach entirely neglects information about pollen counts, it is comparably robust against the no-modern-analogue problem (Neumann et al., 2007). This robustness arises from the idea that the presence of taxa in the fossil sample is mainly conditioned on climate and independent of other taxa (Eq. (3)). Plant competition and other non-climatic effects may have a stronger influence on pollen counts than on presence/absence information. As a consequence, pollen percentages have first to be transferred into presence/absence data by defining thresholds, which might be difficult to determine on an objective basis and with uncertainty estimates (Latalowa and van der Knaap, 2006). Another effect of neglecting pollen counts is that the estimated posterior distributions are typically wider than for other approaches.

2.4. Modern analogues

The modern analogue technique (MAT) uses dissimilarity coefficients, i.e. distance metrics in pollen space, to measure the difference between fossil pollen samples and potential modern analogue samples (cf. Overpeck et al. 1985; Gavin et al., 2003; Jackson and Williams, 2004). Paleoclimate reconstructions follow from the weighted averages of the climate at the recent pollen sites, e.g. with the weights derived from the inverses of the distance metrics, or from selection of the "best" analogue, defined to be the modern sample that is the least dissimilarity distance from a given fossil sample (the "nearest-neighbour" sample). Various dissimilarity coefficients have been studied, ranging from unweighted coefficients, which are strongly influenced by large-valued pollen

taxa, to equally weighted coefficients, which are more sensitive to variations among rare taxa. The major strength of the MAT is, first, that it uses information from many taxa in the multivariate pollen spectrum (or in some cases the clearly dominant taxa, cf. Lytle and Wahl, 2005), and second, that it is not limited to parametric taxon-climate relations. However, an important limitation of the MAT is that it cannot handle no-modern-analogue situations (Williams, 2003), and thus generally is likely to be of less and less value the further back in time reconstruction is desired.

The problem of no-modern-analogues can be reduced to a degree by introducing response surfaces (e.g., Bartlein et al., 1986; Bartlein and Whitlock, 1993). A response surface is essentially a smoothed representation of the pollen data associated with a gridded version of the modern (uni- or multi-variate) climate space used for analogue selection in the MAT. The surface can be projected outside the envelope of modern climates, allowing reasonable extrapolation—at least in the region "near" the known modern climate-pollen space. The response surface method thus can be viewed as a variant of the MAT that allows some potential for extrapolation into no-analogue situations and that presumably removes a significant amount of the spatial variability in the modern pollen information that is non-climatic (Bartlein and Whitlock, 1993).

In both the "straight" and response surface versions of the MAT, a simple method to allow uncertainty estimation in the reconstruction process is to define a *k*-nearest-neighbor (*k*NN) limit for the number of analogs selected, as opposed to choosing the single-nearest-neighbor analogue data point or location on the response surface. Using the *k*NN limit, the weighted average climate value(s) described above is (are) treated as the "expected value" climate reconstruction and the weighted standard deviation(s) of this (these) value(s) provide(s) a measure of reconstruction uncertainty. Clear problems with this approach arise from the definition of the nearest-neighbour limit, which can potentially remove useful climatic information that might be contained in the non-selected pollen assemblages, and also from the fact that the spread of the climate value(s) of the selected analogues may be quite restricted or quite wide simply due to the characteristics of the modern data set used (Bartlein and Whitlock, 1993). The response surface method can further restrict variability because it selects from smoothed representations of the pollen-climate relationship, which can be problematic if climate extremes are a desired target of reconstruction (P. Bartlein, per. comm.).

A refinement of the MAT to deal with the *k*NN issue establishes a cutoff value of the dissimilarity metric, below which good quality analogues are defined to occur. A recent extension of this method uses receiver operating characteristic (ROC) analysis to determine optimizing cutoff values of the dissimilarity metric(s), explicitly addressing the tradeoff between introducing false positives (Type I error) and false negatives (Type II error) that necessarily occurs in such situations (cf. Gavin et al., 2003; Wahl, 2004; Sawada, 2006). In this implementation of the MAT, *k* is not a determined value, with the exception that a minimum number of analogues needed to reconstruct may be specified (e.g., Wahl, 2003), but rather represents the number of analogues below the determined dissimilarity threshold for each fossil sample for which reconstruction is attempted. No-modern-analogue situations occur when there is no modern data point or location on the response surface within the determined cutoff distance of the fossil samples of interest. It should be noted that while this process objectifies threshold selection and allows explicit consideration of the balance between Type I and Type II errors (unless joint minimization of both errors is chosen; cf. Wahl, 2004), it does not eliminate the problem that estimated uncertainties can be strongly affected by idiosyncratic characteristics of the modern data set. Another possibility regarding analogue selection could be to simply eliminate ROC optimization or ways of selecting a set value for *k*, and instead use all potential analogues with a strongly declining weighting as a function of increasing dissimilarity. Such a method could potentially eliminate possible artefacts that might be introduced into reconstruction by use of a specific threshold for analogue selection. Examination of such a method in relation to the ROC optimization process can potentially determine if it could extract more useful climate information and eliminate more non-useful climate information than ROC-determined thresholds. Additional examination of the analytical choices involved in implementing the MAT in its traditional form is provided by Sawada et al. (2004), Sawada (2006), and Williams and Shuman (2008).

Gavin et al. (2003) use ROC methods to evaluate several distance metrics for analogue assignments, and then apply a simple Bayesian statistical approach to examine posterior probabilities that dissimilarity distances of a specific metric (squared chord distance) are associated with particular vegetation-types (biomes, see below) relative to all other vegetation types examined. For a full probabilistic version of the MAT, the vegetation step V in Eq. (2) is described as set of multivariate random vectors V_j in terms of the pollen spectra of the modern analogues. By integrating out the vegetation step, the posterior distribution follows as a mixture distribution

$$[\mathbf{P},\mathbf{C},\mathbf{\Theta}] = \sum_j [\mathbf{P}|V_j,\mathbf{\Theta}]\,[V_j|\mathbf{C},\mathbf{\Theta}]\,[\mathbf{C}|\mathbf{\Theta}][\mathbf{\Theta}],\ \text{for all } j = 1,\ldots,m, \quad (4)$$

where $[V_j|\mathbf{C},\mathbf{\Theta}]$ represents the recent climate of the modern analogue in a forward sense. It can be easily seen that $[\mathbf{P}|V_j,\mathbf{\Theta}]$ is the probabilistic equivalent to the distance metric: the likelihood of the fossil pollen spectrum $\mathbf{P}$ given the modern spectrum V_j.

A hierarchical Bayesian formulation of the MAT that includes the concept of response surfaces is already used in Haslett et al. (2006). Their approach relates to the Bayesian response models by Vasko et al. (2000), Toivonen et al. (2001), and Korhola et al. (2002), which have been developed for direct whole-organism-based reconstructions, not pollen proxies. The likelihood $[\mathbf{P}|V_j,\mathbf{\Theta}]$ follows a multinomial distribution $[\mathbf{P}|V_j^*]$ where the pollen spectra of the analogues V_j are replaced by the event probabilities V_j^*. Consequently, $[V_j|\mathbf{C},\mathbf{\Theta}]$ is replaced by a Dirichlet distribution $[V_j^*|\Phi_j]$ with expected values Φ_j, which is a practical choice as it is conjugate to the multinomial. By using a smooth function of climate for the expected values $\Phi_j(\mathbf{C})$, which is obtained from the modern analogues V_j, the approach can be understood as the Bayesian version of the classical response surface technique.

It is important to note the clear multimodality (one global maximum and up to several local ones) exhibited in the posterior climate reconstructions by Haslett et al. (2006; cf. their Fig. 8) for the test site at Glendalough, Ireland. This outcome highlights a key issue in pollen-based paleoenvironmental reconstruction: it arises as an artefact of the interaction between the lack of species-specific pollen morphology for many pollen types and the wide spatial/climatological range of the modern pollen assemblages used in formulating Haslett et al.'s response surface model. The authors comment explicitly on this outcome, noting as an example the wide climatological range associated with *Juniperus*-type pollen in the modern data set; this pollen type cannot be further distinguished at the sub-generic level, yet it incorporates eight species that range from Arctic to Mediterranean environments in Europe. Relatively high proportions of *Juniperus* pollen in the fossil archive at Glendalough at certain times during the reconstruction period then necessarily lead to multimodal reconstruction outcomes – thereby reflecting the inherent uncertainty in the climate-pollen relationship created by the pollen morphology constraints. A classical approach to eliminate this problem in pollen-based

paleoenvironmental reconstruction is to limit the spatial range of the modern data used for reconstruction to that nearer the sits(s) to be reconstructed and/or across well-established bio-regional boundaries (e.g., Williams and Shuman, 2008). However, this approach does not necessarily eliminate the problem; e.g., in mountainous regions distinct plant species with identical pollen morphologies but different environmental ranges can have modern ranges quite near each other (Wahl, 2002). More generally, such limitation necessarily involves a (lesser or greater) amount of ad hoc decision-making regarding appropriate spatial boundaries, and by construction operates against the goal of making very generally applicable reconstruction models.

The pollen ratio model introduced in Section 3 is also a special case of the MAT that includes the concept of response surfaces. With the assumption of a binomial response, it is the equivalent of a multinomial in a reduced bivariate pollen space. In relation to response surfaces, the logistic link function provides the smoothed relation to the climate of the modern analogues—albeit for a more limited climatic range.

2.5. Plant functional types and biomisation

Plant functional types (PFTs) are defined as groups of taxa with similar biological characteristics such as leaf form, habit, phenology, or bioclimatic range (e.g. Prentice et al., 1992, 1996). Each pollen taxon can be empirically assigned to at least one PFT. The affinity of a recent or fossil pollen sample with each PFT is then measured by an "affinity score", which is the sum of the square roots of all pollen percentages above predefined thresholds defined into each PFT. Similar to the distance metrics in the MAT, transformation by the square root and use of a threshold have been introduced to increase the signal-to-noise ratio of the affinity scoring process. A biome is a more general classification and defined as a combination of PFTs. Hence, the biome affinity score is calculated from the sum of the corresponding PFT scores. Climate reconstructions can in turn be obtained by mapping PFTs (or biomes) to bioclimatic ranges. For instance, Peyron et al. (1998) and Tarasov et al. (1998) employ artificial neural networks (ANN) to estimate the relationships between the PFT scores of modern pollen spectra and the climatic parameters estimated at the modern pollen sites.

The concept of PFTs and biomes is another way to enhance the realism of spatial hierarchical models for pollen proxies. With respect to the probabilistic framework above, the vegetation step $\mathbf{V}$ is expressed in terms of PFT or biomes, e.g., the presence/absence of a PFT derived from its affinity score. Following a Bayesian formulation of a biome-based climate reconstruction (Schölzel, 2006), equation (2) then would have the same general form as equation (4) for the MAT, reading

$$[\mathbf{P},\mathbf{C},\mathbf{\Theta}] = \sum_j [\mathbf{P}|V_j,\mathbf{\Theta}]\,[V_j|\mathbf{C},\mathbf{\Theta}]\,[\mathbf{C}|\mathbf{\Theta}][\mathbf{\Theta}],\ \text{for all}\, j = 1,\ldots,m \qquad (5)$$

with the vegetation step again integrated out. As in the case of MCRs, the PFT-climate relation $[V_j|\mathbf{C},\mathbf{\Theta}]$ can be estimated, e.g., by a GLM with binary response, logistic link, and polynomials of continuous covariates, which is equivalent to a feedforward single-layer ANN. The pollen-PFT relation $[\mathbf{P}|V_j,\mathbf{C},\mathbf{\Theta}]$ can be derived from the affinity scores and hence simplified to $[\mathbf{P}|V_j,\mathbf{\Theta}]$, so that pollen is assumed to be directly determined only by vegetation.

The freedom of choice in defining the affinity scores—which can be compared to distance metrics and analogue assignments in MAT (e.g., Gavin et al., 2003)—and their interpretation as probabilities is problematic: differing definitions can affect uncertainty evaluation and are not derived from general statistical principles, such as use of the multinomial and Dirichlet within a more general MAT framework. Gachet et al. (2003) and Gritti et al. (2004) introduce a probabilistic approach to the use of pollen indicators for plant attributes and biomes, where affinity scores are replaced by probabilities, but a strict BHM formulation would require that all uncertainties have to be expressed through parametric probability distributions. A possible solution would be to define V_j as a multivariate random vector V_j and $[\mathbf{P}|V_j,\mathbf{\Theta}]$ as multinomial and Dirichlet, formally identical to Bayesian versions of the MAT above (Section 2.4). To our knowledge such an approach has not been applied to PFT or biomes so far.

2.6. Process-based vegetation models

Process-based vegetation models simulate the quantitative properties of PFTs in dependence on various environmental parameters like climate, soil properties, and atmospheric CO_2 content. Typical representatives are BIOME3 (Haxeltine and Prentice, 1996) or BIOME4 (Kaplan et al., 2003), which are equilibrium models. Dynamic vegetation models like the LPJ-GUESS model (Smith et al., 2001) also take into account competition and mortality. An advantage of dynamic vegetation models is that they allow simulation of a process-based vegetation history, whereas most of the methods described above implicitly assume temporal independence in the fossil pollen samples. This advantage further incorporates the ecological reality that climate and vegetation do not necessarily have to be in equilibrium, as state-of-the-art vegetation models account for competition, growth, and mortality. Since process-based models are forward models, paleoclimate reconstructions are obtained by inversion (Guiot et al., 2000). A description of vegetation modeling and Bayesian inversion is given in Guiot et al. (2009).

The model inversion is closely related to the framework above. A major difference is that a fixed mechanistic model is applied to random draws from the climate parameter space to generate the posterior distribution. In contrast, the methods above treat the model parameters as random variables as well, i.e. including their stochastic uncertainty. Garreta et al. (2010) introduce an important change toward this direction by treating the vegetation step as random variable. However, since the inversion of a dynamical vegetation model is already numerically complex, additional parameter inference related to incorporating the vegetation model into a large spatio-temporal BHM is probably not feasible at the current stage of development. With respect to the overall focus of enhancing the pollen realism in a multiproxy BHM framework, posterior probabilities from processed-based vegetation models could be incorporated by an empirical Bayesian approach.

2.7. Selection of prior distributions

Finally, the selection of prior distributions should be briefly addressed here, since priors are used in various ways in paleoclimate reconstruction and with a profound effect on the interpretation of the results obtained. Every calculation of a posterior probability or probability density directly depends on the selection of the corresponding prior probability or probability density (Eqs. (1) and (2)), so that posteriors are only interpretable in the context of priors. As the priors have to be given in advance, the concept of Bayes' theorem refers to "learning from data", transforming prior knowledge into posterior knowledge. The posterior distribution can be seen as a compromise between data and prior information (Gelman et al., 2003). For instance, in case the posterior is identical to the prior, there is no information contribution from the data, i.e., the likelihood function which incorporates information from the data (e.g., in section 2.4 on the MAT, $[\mathbf{P}|V_j,\mathbf{\Theta}]$ is the likelihood of the fossil pollen spectrum $\mathbf{P}$ given the modern spectrum V_j) could be replaced by a uniform distribution.

In a sensitivity analysis on the selection of prior distributions in Bayesian paleoclimate temperature reconstructions Erästö and Holmström (2006) demonstrate how crucial the selection of a prior probability distribution for the past values can be, particularly when the data are noisy and the statistical model used is complex. Highly informative prior distributions, for example modern temperatures in the prior parameters of past temperatures, as in Toivonen et al. (2001) and Korhola et al. (2002), can to a large extent prescribe the reconstructed temperature (Erästö and Holmström, 2006). Although imposing strong prior beliefs is consistent with Bayes' theorem, the practical use is open to question as it will be difficult to interpret the results in a broader sense.

Hence, we propose to use non-informative priors on the reconstructed variable whenever possible or restrict to physically plausible constraints only (e.g. Neumann et al., 2007), in order to maximize the information in the posterior that comes from the data or to indicate that the data do not have strong information content. Moreover, MCMC methods require that a prior distribution is specified for each parameter. Finding proper prior distributions for all model parameters can be a complicated task. Since there is a vast field of literature on this topic, a description is omitted here and we refer to e.g. Gilks et al. (1995) for details.

An alternative use of priors, which underlines the idea of Bayesian learning, is shown in Garreta et al. (2010) by using posterior distributions from other reconstruction methods in order to test for improvements of a new model with respect to this reference model. In that light, it should be noted that by definition the prior information has to be independent of the data used in the likelihood/data stage; i.e., the reference model has to be built on an entirely different data set.

3. Case study of the pollen-ratio model

3.1. Method

As mentioned in section 2.4, we introduce here a simplified variation on the MAT that includes the concept of response surfaces, using a generalized linear model (GLM) with binomial response and logistic link function. This statistical formulation is well-expressed theoretically in that pollen count data $\mathbf{P} = (P_1, P_2)$ for two groups of taxa of the form $P_i = \sum_j P_{ij}$ for the individual taxa j in group i are binomially distributed, conditional on additional covariates and the overall pollen sum across taxonomic groupings. The name pollen-ratio model originated from an alternative but equivalent description by using the ratio $P_1/(P_1 + P_2)$ conditional on the included count values' sum. The logistic formulation of the modeled climate-to-pollen relationship is both physically and mathematically appropriate. It is a forward model, expressing the physically correct direction of causation and the binomial parameters and/or ratio values are constrained by construction to be limited to the range [0,1]. Once a relation of this kind is estimated on modern data using Bayesian (or classical) GLM methods, it can be inverted and used with fossil pollen data for paleoclimate reconstruction. Moreover, and importantly, the Bayesian posterior distributions of the GLM parameters coupled with the binomial distribution of the fossil pollen count values allow explicit statistical characterization of uncertainty via a two-way Monte Carlo random draw process. Taxonomic specification and classical estimation of the model using data for northeastern North American forest ecosystems is provided in Wahl et al. (2010), cf. included SOM; modern pollen data are from Whitmore et al. (2005). A summary and Bayesian estimation is provided here.

The modern data forward model is:

$$P_1^m | N_m, \pi \sim \mathrm{Bin}(N_m, \pi)$$
$$P_2^m | N_m, \pi \sim \mathrm{Bin}(N_m, 1-\pi)$$
$$\pi = \mathrm{logit}^{-1}(\alpha + \beta C^m)$$

Here, $\mathbf{P}^m = (P_1^m, P_2^m)$ are the modern-site pollen counts for the two groups of taxa, $P_i^m = \Sigma_j P_{ij}^m$ for the taxa j in each group i; $N_m = P_1^m + P_2^m$ is the overall pollen count; and C^m is the climate value at each modern data site corresponding to a specific set of pollen counts.

In a first step, random samples of the GLM parameters α and β are drawn. This can be done by either Bayesian or by classical GLM methods, for example using parametric distributions and maximum likelihood estimates $\widehat{\alpha}$ and $\widehat{\beta}$ including their standard deviations. In the following, we especially focus on posterior draws of $\tilde{\alpha}$ and $\tilde{\beta}$ from a Bayesian GLM (Hadfield, 2010). Reasonably flat standard prior distributions with zero mean and a variance of 10^8 were chosen for the GLM parameters.

In a second step, inversion is done by solving the GLM equations for the climate variable C as a function of α, β, and the binomial parameter π. Note that this solution is in terms of the modern values of the climate variable used to estimate the GLM, and not in terms of paleoclimate values, which in strict terms comprise a separate variable. To interpret this solution in terms of paleoclimate, the inversion model needs to be re-expressed as follows, where C^f denotes the reconstructed climate, $\tilde{\alpha}$ and $\tilde{\beta}$ denote random draws from the first step, and $\mathbf{P}^f = (P_1^f, P_2^f)$ denotes fossil pollen information:

$$C^f = \left(\mathrm{logit}\left(\tilde{\pi}_f\right) - \tilde{\alpha}\right) \Big/ \tilde{\beta}$$
$$\tilde{\pi}_f \Big| \mathbf{P}^f \sim \mathrm{Beta}\left(1 + P_1^f, 1 + P_2^f\right)$$

In order to use this form, the binomial parameters $\tilde{\pi}_f$ are drawn from the beta distribution, which naturally provide the posterior probability of this parameter after observing P_1^f "successes" (realization of the first group of taxa) and P_2^f "failures" (realization of the second group of taxa). Note that the mode $\tilde{\pi}_f$ is given by the fossil data ratio $P_1^f/(P_1^f + P_2^f)$. Under this assumption, the above formulas yield the random draws of reconstructed paleoclimate, given the observed fossil pollen counts and the posterior draws of the GLM parameters.

More generally, C^f is a random variable with three components of stochastic variation: $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\pi}_f$. The first two components of variation, defined by the joint posterior distribution of α and β, arise from the randomness in the modern pollen ratio-temperature relationship modeled by the binomial-logistic GLM, which is also assumed to hold during the reconstruction time period. The third component of variation arises from the fact that measured values of $\mathbf{P}^f$ from a sediment core are in fact a draws from a stochastic distribution of pollen counts that can reasonably be assumed to occur across the depositional environment during an examined time step. Thus, the randomness of the binomial parameter can be modeled as a beta process in which the observed fossil ratio is assumed to asymptotically represent the binomial parameter. Note that the beta distribution and the binomial distribution are each special cases of the Dirichlet distribution and the multinomial distribution, respectively, which links the pollen ratio model to probabilistic versions of the MAT with continuous responses. (c.f. Section 2.4).

3.2. Results and comparison to MAT

Application of the inverted Bayesian binomial GLM model to fossil pollen data from varved (annually-laminated) sediments in

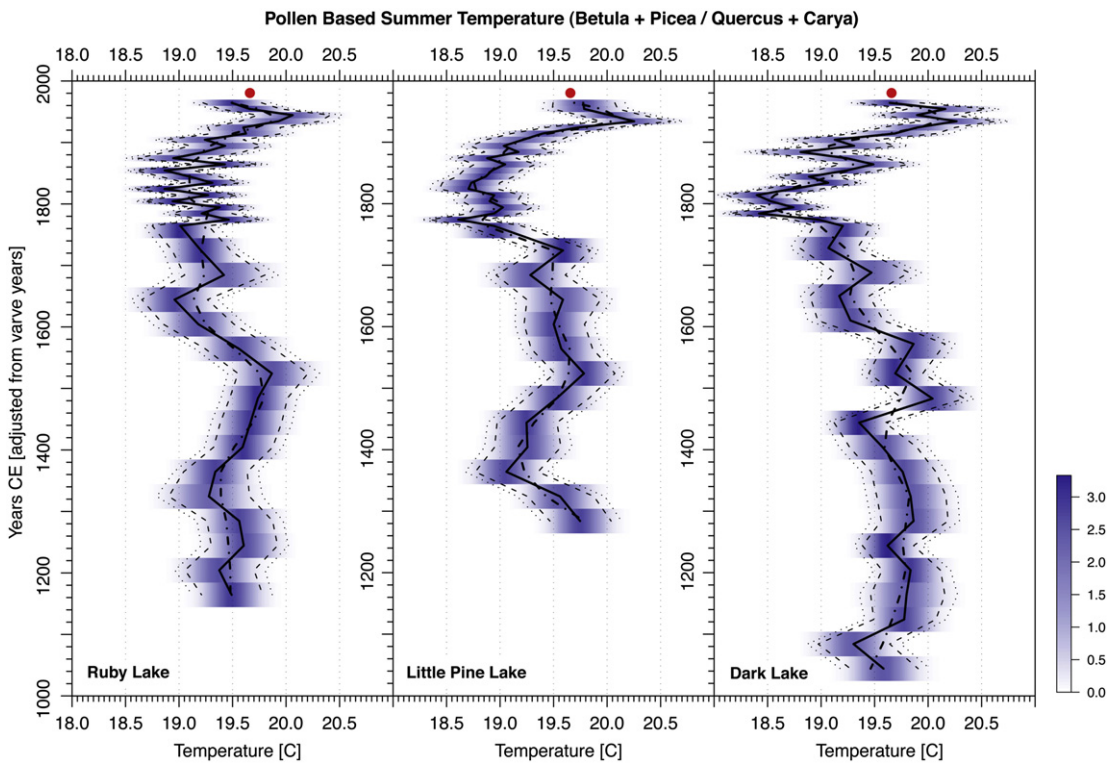

Fig. 1. Reconstruction of summer (JJA) surface temperature in central North America over the past millennium, based on pollen preserved in varved sediments at Ruby Lake, Little Pine Lake, and Dark Lake, Wisconsin, USA. Blue shading shows the estimated probability density for temperature; dashed/dotted lines indicate the 95/99% probability ranges of the reconstructions; the red dot represents the modern long-term mean (1961–1990) in the vicinity of the lakes; the smoothed (dashed–dotted) line is a lowess fit of the ensemble median reconstruction (solid line). (Cf. Wahl et al., in press.)

three nearby lakes in west-central Wisconsin, USA (Gajewski et al., 1985), is shown in Fig. 1, for reconstruction of mean summer (June–July–August) temperature. For these three basins, the modern summer temperature climatology values (1961–1990; from Whitmore et al., 2005) are well inside the 95% probability ranges of the inversion reconstructions (cf. Fig. 1), which provides a minimum validation test for the methodology, and is coupled with high explained deviance (analogous to explained variation in linear models) of the modern pollen fit (0.79, from a classical fit of the GLM to the same data for reconstruction of mean July temperature; Wahl et al., 2010, SOM).[2] The high explained deviance of the model yields reconstructions whose 99% probability ranges allow clear distinction of a cooler Little Ice Age (LIA) in relation to both earlier (Medieval Climate Anomaly, MCA) and later times (by $\sim 1.5^{\circ}$, compared to a mean 99% reconstruction probability range $\leq \sim \pm 0.45^{\circ}$). This overall interpretation using the information from the three lakes individually remains essentially unchanged when the estimated probability distributions from the individual lakes are combined through application of ensemble post-processing methods (cf. section 4 below) (Fig. 6). The most likely excursion of LIA cooling in the combined case reduces to $\sim 1.0^{\circ}$ and the width of the overall 99% probability range expands to a maximum of $\sim \pm 1.15\ ^{\circ}$C (it is generally $\leq \pm 0.85\ ^{\circ}$C). Wahl et al. (in press) provide a detailed examination of this reconstruction from a paleoclimatic perspective, which is beyond the scope of the current article: including timing of the onset of LIA cooling (in terms of the rapidity of pollen response to temperature changes), and the meaning of the early (MCA) portion of the reconstruction in terms of a possible modern analogue analysis related to larger-scale atmospheric circulation regimes (pressure and drought).

A second application of the inverted Bayesian binomial model to a much longer (full Holocene) fossil pollen record from southern California, USA (Wahl, 2002, 2003) has also been done, and the results compared to an ROC-optimized MAT reconstruction from the same data (Wahl, 2002, 2003, 2004). GLM estimation for July mean temperature is shown in Fig. 2, and both the GLM-inversion and MAT reconstruction results are shown in Fig. 3. In comparison, there is more variability in the GLM-based reconstruction, with narrower comparable probability ranges, but the "overall picture" is the same across the two reconstructions: both indicate strong apparent coolness just after the end of the Younger Dryas/beginning of the Holocene (or perhaps indistinguishable from this transition depending on the length of a possible lag of vegetation response to climate change; cf. Wahl et al., in press) and temperatures generally near today's since then (with some variability over time). It is worth noting an advantage of the GLM ratio-based reconstructions shown by this comparison; the model can generate reconstructions in cases when there were judged to be too few modern analogues to reconstruct (lowest brown data point and gap-filling blue data point in right panel of Fig. 3). It should also be

[2] Where both the numerator-specific taxa, *Betula* (birch) and *Picea* (spruce), and the denominator-specific taxa, *Quercus* (oak) and *Carya* (hickory), are represented above trace levels in the pollen spectra so the model could be reasonably evaluated, the inverted GLM ratio model was employed with high-resolution fossil data over the last 1,000 years from other lake basins in temperate northeastern North America. The modern climatology values for these three basins lie far outside the derived 99% probability ranges, and thus this information was not considered validated. This situation could be due to ecological differences between forests in the region of the three west-central Wisconsin lakes and those in the regions of the other lakes (cf. Gajewski et al., 1985), which reduce or perhaps negate the validity of the specified taxonomic model for use at those sites. Similar region-specific reconstruction success and other issues using pollen-based regression models (cf. section 2.2 above) are discussed in Bartlein et al. (2010), cf. section 2.3 in that article.

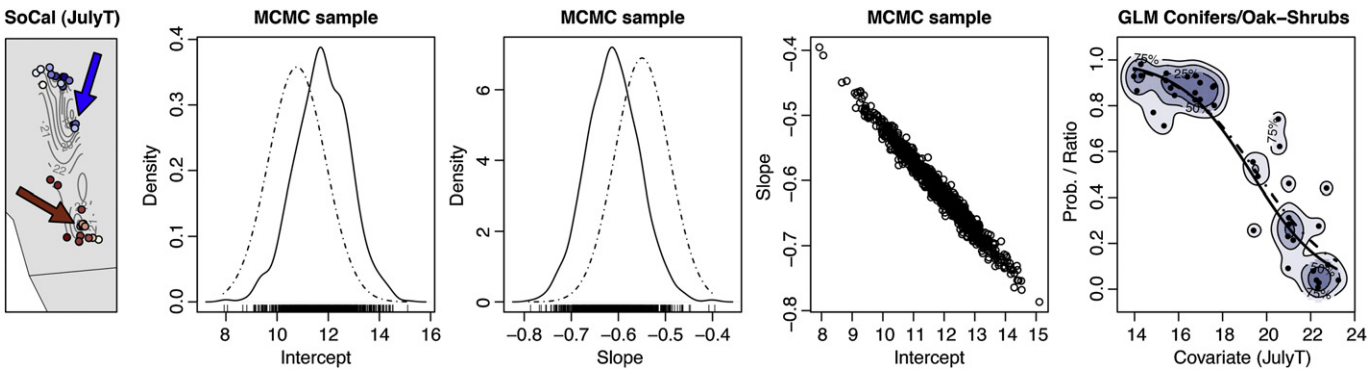

Fig. 2. Binomial-logistic GLM fit for July mean temperature for the mesic (western) side of the coastal southern California mountains. The region and modern pollen-climate data sites used in the estimation are shown in the far-left panel (note the USA-Mexico border, 32° 32′ N at the adjacent Pacific Ocean, in the lower part of the panel). Darker blue/darker red dots for the sites indicate cooler/warmer modern temperature; arrows note the sites of the fossil data used in reconstruction, shown in Fig. 3 (~1440 m elevation, brown; ~2400 m elevation, blue). The marginal posterior distributions of α and β (solid lines; dashed lines show comparable distributions from classical estimation) and their strongly linked relationship are shown in the three centre panels. The scatter diagram of pollen ratios [(all conifers)/(all conifers + oak + important shrubs)] vs. July T and fitted modal Bayesian logistic curve (solid line) are shown in the far-right panel (classical fitted EV curve is shown by the dashed line). The explained deviance (from classical fit) is 0.79.

noted that the probability ranges estimated from the inverted GLM model would almost certainly be significantly wider if multiple nearby fossil deposition sites were available for reconstruction and ensemble post-processing of the individual reconstruction distributions could be done as in the case of the three Wisconsin lakes mentioned above (cf. section 4 below).

A disadvantage of the GLM-inversion reconstruction method as described and employed here is that it cannot be used simultaneously to reconstruct more than one climatic variable; whereas the MAT method to which it is compared can reconstruct multiple variables simultaneously, based on the general MAT assumption that the environmental conditions associated with the selected modern analogue pollen sites should be similar to those of the fossil sites reconstructed. An example of the alternative use of the GLM-inversion model to fit and reconstruct Holocene annual precipitation in southern California (rather than temperature) is presented in Figs. 4 and 5, and compared to the Wahl (2002) MAT-based reconstruction in Fig. 5.

It should be noted in the case of precipitation that the requirement of reasonable reconstruction of modern conditions is not met for the GLM-based reconstructions, and thus the comparison with the MAT-based reconstructions is potentially not as meaningful as that shown for July temperature. However, unlike the situation for the lakes in northeastern North America mentioned above, in which possible ecological differences among forests in this larger region could potentially reduce or negate the validity of the specific taxonomic model used, the offset in the much more spatially-restricted southern California case can be directly attributed to an unavoidable effect of the GLM estimation procedure. This effect occurs because both fossil sites and their very nearby modern data sites represent relatively extreme cases of the climate-pollen relationship for their kinds of vegetation in the modern data set; the

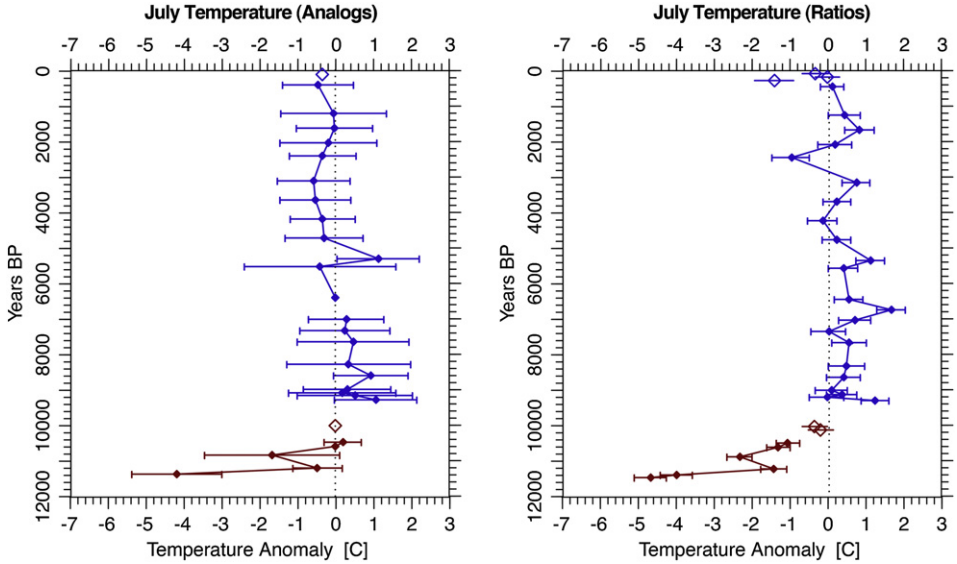

Fig. 3. Reconstruction of July surface temperature over the Holocene (based on pollen preserved in wet meadow sediments) in southern California, USA. ROC-optimized MAT reconstruction from Wahl (2002) is shown at left; GLM-inversion reconstruction reported in this article [based on pollen ratios of (all conifers)/(all conifers + oak + important shrubs)] is shown at right (cf. Fig. 2). The blue and brown reconstructions are from the two sedimentary fossil sites noted in left panel of Fig. 2, with temporal data coverages as shown. Vertical line indicates modern climatology (0° anomaly) at the two sites. Estimated 67% ("one sigma") probability ranges of reconstruction are indicated. Reconstructions for modern data sites in very near proximity to the fossil data sites are shown by open diamonds. Timing of fossil pollen samples is based on calibrated ^{14}C dating reported in Wahl (2002); "BP" indicates years before CE 1950.

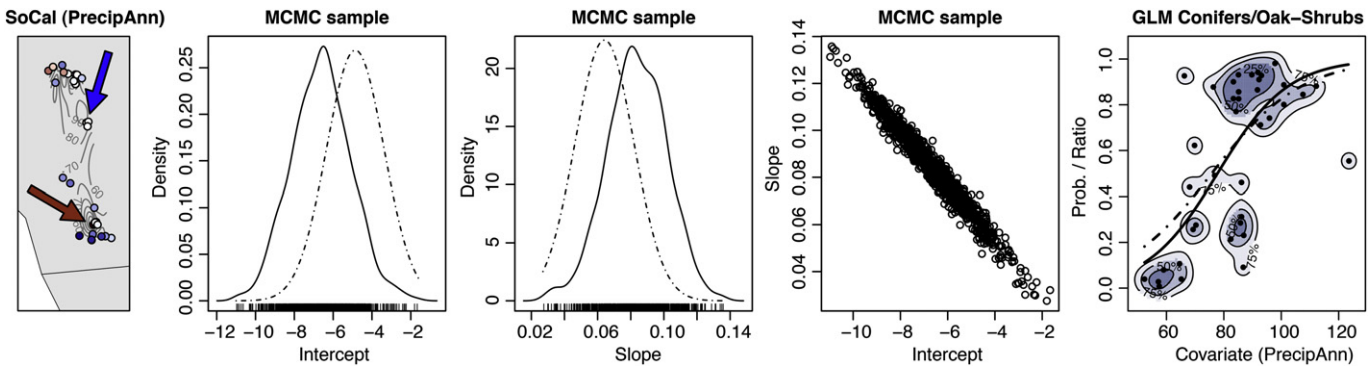

Fig. 4. Binomial-logistic GLM fit for annual precipitation for the mesic (western) side of the coastal southern California mountains. The region and modern pollen-climate data sites used in the estimation are shown in the far-left panel (note the USA-Mexico border, 32° 32′ N at the adjacent Pacific Ocean, in the lower part of the panel). Darker blue/darker red dots for the sites indicate lower/greater modern precipitation; arrows note the sites of the fossil data used in reconstruction, shown in Fig. 5 (~1440 m elevation, brown; ~2400 m elevation, blue). The marginal posterior distributions of α and β (solid lines; dashed lines show comparable distributions from classical estimation) and their strongly linked relationship are shown in the three centre panels. The scatter diagram of pollen ratios [(all conifers)/(all conifers + oak + important shrubs)] vs. annual precipitation and fitted modal Bayesian logistic curve (solid line) are shown in the far-right panel (classical fitted EV curve is shown by the dashed line). The explained deviance (from classical fit) is 0.33.

higher-elevation "blue" site has low precipitation for highly conifer-dominated forest sites and the lower-elevation "brown" site has quite high precipitation for mixed conifer-oak forest sites (Wahl, 2003). Thus, it is possible to predict in this case that the GLM-based reconstructions will tend to over-reconstruct/strongly under-reconstruct modern climatology at the higher/lower sites, respectively, which is what the results show in Fig. 5.

With this set of caveats in mind, the "overall picture" is similar across the two kinds of precipitation reconstructions for the higher-elevation site, especially when the systematic positive offset in the reconstructions of the related modern sample sites is taken into account. For the lower-elevation site, the precipitation reconstructions diverge, especially when the systematic negative offset in the GLM-based reconstruction of the related modern sample sites is accounted for. In both methods there is no indication of a quite dry early Holocene, even if the modern sample offset is ignored. Climatologically, this represents a key result: it suggests—in conjunction with data from Davis (1995) for the Sierra Nevada Mountains of northern California and from Anderson (2000) for lowland coastal southern California—that there was a sharp moisture gradient (dry to the west and north/wet to the east and south) just west and north of montane southern California in the early Holocene (cf. Wahl, 2002).

4. Discussion of site-specific uncertainty

An obvious next step in probabilistic reconstruction modeling is to go beyond single-site reconstructions. The reconstructions from the three nearby lakes in west-central Wisconsin, USA (Section 3) allow discussion of different sources of uncertainty and also to relate to the use of ensemble simulations in modern climatology and numerical weather prediction. Ensembles of climate simulations are presently the only way to sample uncertainties arising from initial conditions, parameter selection, and structural uncertainties in model design. The important question is how ensembles can be translated into probabilistic information, which is typically

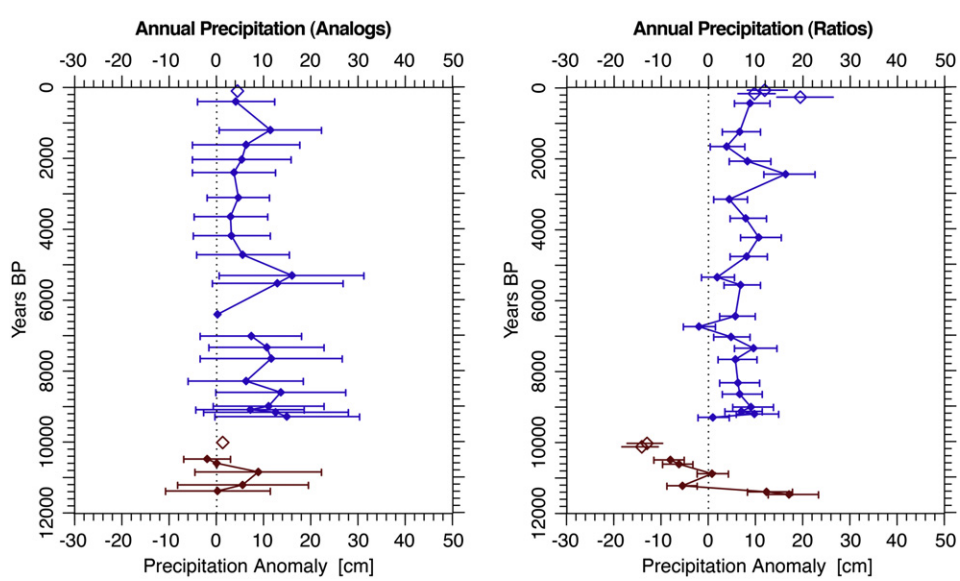

Fig. 5. Reconstruction of annual precipitation over the Holocene (based on pollen preserved in wet meadow sediments) in southern California, USA. ROC-optimized MAT reconstruction from Wahl (2002) is shown at left; GLM-inversion reconstruction reported in this article [based on pollen ratios of (all conifers)/(all conifers + oak + important shrubs)] is shown at right (cf. Fig. 4). The blue and brown reconstructions are from the two sedimentary fossil sites noted in left panel of Fig. 4, with temporal data coverages as shown. Vertical line indicates modern climatology (0 cm anomaly) at the two sites. Estimated 67% ("one sigma") probability ranges of reconstruction are indicated. Reconstructions for modern data sites in very near proximity to the fossil data sites are shown by open diamonds. Timing of fossil pollen samples is based on calibrated ^{14}C dating reported in Wahl (2002); "BP" indicates years before CE 1950.

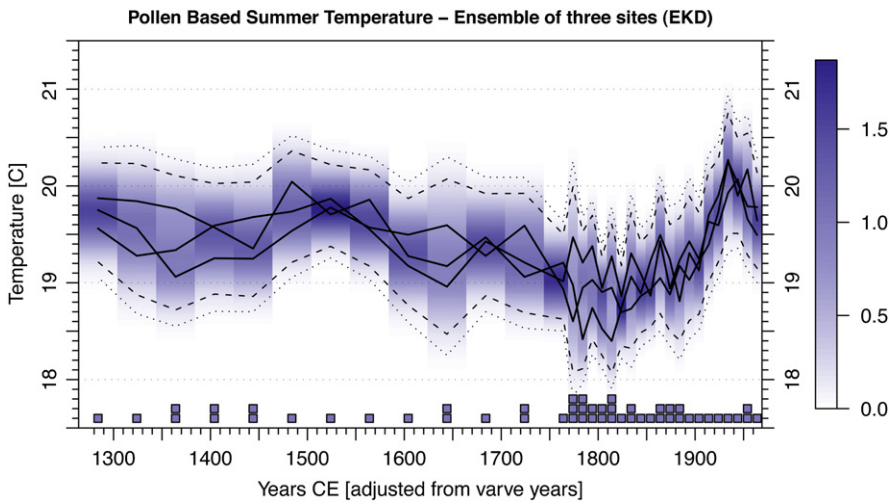

Fig. 6. Estimated mixture model probability distribution of the reconstructions for the three lakes shown in Fig. 1. The distribution follows from an ensemble kernel dressing approach (EKD) as described in Schölzel and Hense (2011). The blue shading shows the posterior probability density for temperatures; dashed/dotted lines indicate the 95/99% probability ranges. The number of boxes along the time axis estimates the optimum number of mixture model components at each time period reconstructed, using the Bayesian Information Criterion (BIC). Two or three boxes support a conclusion that there is significant site-specific representation of pollen in relation to summer temperature across the individual lakes; implying that the uncertainty in reconstruction is underestimated at the individual sites. (Cf. Wahl et al., in press.)

referred to as ensemble post-processing or ensemble interpretation (e.g. Bröcker and Smith, 2008; Wilks and Hamill, 2007). Here, we try to emphasize the link to these problems, as they are actually the same in paleoclimate, albeit at lower model complexity.

In order to test the consistency of site-specific uncertainty modeling in a real-world example, the reconstructions for the three sites in Wisconsin are interpreted as a three-member ensemble. With respect to the preciseness of the climate data involved, the three sites can be considered co-located, so that no additional spatial modeling of climate is required. Fig. 6 shows estimates for the optimum number of clusters according to the Bayesian information criterion (BIC) for a parameterized Gaussian mixture model (Fraley and Raftery, 2002). If the differences between the three sites would be correctly described by their individual uncertainty estimates, there should mainly be one mixture component. Most of the time, there are two or three components, indicating that the uncertainty in the reconstruction from the ratio model at each site—although we account for effects like uncertainty in pollen counts—is typically underestimated. The same applies for a multivariate cluster analysis where the full time series is interpreted as an n_t-dimensional random vector.

More realistic estimates of uncertainty are obtained from ensemble post-processing methods. The approach used here is related to ensemble kernel dressing (EKD) and fully described in Schölzel and Hense (2011). It basically comprises a multivariate Gaussian kernel density estimation with temporal error correlation, i.e., the whole time series is treated as a multivariate random vector, which enables drawing time series from the combined posterior. The dressing covariance, which is the multivariate extension of the bandwidth, is estimated from the differences of all ensemble members. As a variation from the original method, the limitation to a Toeplitz structure for the covariance matrix has been dropped, since the assumption of a non-stationary error covariance does not hold for irregular time steps.

The resulting posterior distribution is wider when accounting for inter-site differences (Fig. 6), even though the inverse pollen-ratio model already accounts for different sources of uncertainty: (1) the uncertainty in pollen counts, which should be represented by the draws from the Beta distribution; and (2) the uncertainties due to the limited statistical model and representation of local climate, simulated by the MCMC draws for the GLM parameters. For this experiment we conclude that there must be a systematic underestimation of the model uncertainties due to a spatial correlation of unaccounted explanatory variables; from a paleoenvironmental standpoint, one of these could be basin-specific sedimentary characteristics that lead to differential pollen representation. Formally speaking, the assumption of ergodicity does not hold, at least in a strong form, which is a natural drawback of the reduced model complexity. Less sensitive but more robust methods like the probabilistic indicator taxa approach described show more consistent reconstructions of uncertainties, as can be seen for two nearby sediment cores from lake Holzmar and lake Meerfelder Maar, Germany (Litt et al., 2009), which allow for a similar comparison of posterior distributions like the three lakes from Wisconsin.

5. Conclusions

Compared to recent advances in modern climatology, reconstructing paleoclimate from proxy data is still a complicated task. The picture of past climate involves a patchwork of different proxy variables, environmental variables, and spatial and temporal scales using a large number of statistical methods. An important challenge to the paleoclimatological community going forward is to synthesize local reconstructions from different proxies for a spatially complete picture of past climate. This includes full estimation of the uncertainties involved in the reconstruction process, which are naturally large due to the complexity of the proxy-climate relationships. Spatial multi-proxy BHM approaches are a promising development toward these aims. However, at the current stage, they do not access the large accumulated scientific knowledge of pollen-climate transfer methods.

We therefore have sought to provide perspective on how to enhance the pollen-based part in multi-proxy paleoclimate reconstructions and to encourage thinking about classical approaches in a more mathematically rigorous way such as a BHM implementation, which allows specification of the different modeling assumptions in a unified framework. With respect to a strict probabilistic formulation, it is shown how far classical pollen-climate transfer concepts like regression methods, MCR, MAT, and PFT can be understood in such a framework, or have already been implemented. Almost all of the existing approaches are characterized by the way an additional random variable,

representing the vegetation step, is defined. Whether a mixture or ensemble of different pollen-based transfer methods is advisable, analogous to well-established multi-model ensembles in numerical climate simulations, remains unclear. Problems may arise with conflicting modeling assumptions and the availability of independent recent calibration data.

In order to discuss the aspect of uncertainty in a real-world example, a probabilistic version of a pollen-ratio model, which is a simplified variation on the MAT using the binomial-logistic formulation of a GLM, has been introduced and examined in two reconstruction situations: 1) to estimate last-millennium summer temperature from pollen in varved sediments in nearby lakes in west-central Wisconsin, USA; and 2) to estimate full-Holocene July temperature and annual precipitation from pollen in wet meadow (fen) sediments in southern California, USA. Compared to traditional MAT-based reconstructions using the same full-Holocene pollen data, there is more variability in the GLM-based reconstruction, but the overall picture for July temperature is the same. For an alternative use of the GLM model to reconstruct precipitation, the results can be similarly characterized, except that in the very early Holocene they diverge. An advantage of the ratio-based reconstructions is that the model can generate reconstructions in cases when there were too few modern analogues to reconstruct from the MAT. A disadvantage is that the GLM model as described cannot be used simultaneously to reconstruct more than one climatic variable, and thus an analyst would have to choose to reconstruct either temperature or precipitation from the examples shown. Furthermore, an ensemble reconstruction experiment reveals that the overall interpretation based on the individual reconstructions remains essentially unchanged, but the single-site reconstructions underestimate the uncertainty. Whether this also applies to other methods as well has not been tested, but is an important aspect to be considered for pollen-based transfer methods, which are typically built under the (implicit) assumption of ergodicity.

Both the theoretical considerations and the case study of uncertainty modeling in a practical application provide insights into inclusion of pollen proxies into a larger multi-proxy BHM, which is currently under development.

Acknowledgements

Support for the research reported in sections 2.4 and 3 was provided by NOAA (Wisconsin summer temperature reconstructions). Additionally, this material is based upon work supported by the National Science Foundation under Grant No. (0724619) (Wisconsin summer temperature reconstructions) and Grant No. (9801449) (California July temperature and annual precipitation reconstructions). Additional acknowledgment is extended to Dr. Margaret Bryan Davis, Regent's Professor Emerita, the University of Minnesota, and to Dr. Douglas Nychka, Director, Institute of Mathematics Applied to Geosciences, the National Center for Atmospheric Research.

References

Anderson, R.S., 2000. Late Quaternary paleoenvironments of southwestern California. In: Geological Society of America, GSA Abstracts with Programs, vol. 32, pp. A-403.
Bartlein, P.J., Harrison, S.P., Brewer, S., Connor, S., Davis, B.A.S., Gajewski, K., Guiot, J., Harrison-Prentice, T.I., Henderson, A., Peyron, O., Prentice, I.C., Scholze, M., Seppa, H., Shuman, B., Sugita, S., Thompson, R.S., Viau, A.E., Williams, J., Wu., H., 2010. Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. Climate Dynamics, 1–28. doi:10.1007/s00382-010-0904-1.
Bartlein, P.J., Prentice, I.C., Webb III, T., 1986. Climatic response surfaces from pollen data for some eastern North American taxa. Journal of Biogeography 13, 35–57.
Bartlein, P.J., Webb III, T., 1985. Mean July temperature at 6000 yr B.P. in eastern North America: Regression equations for estimates from fossil-pollen data. Syllogeus 55, 301–342.
Bartlein, P.J., Webb III, T., Fleri, E., 1984. Holocene climatic change in the northern Midwest: Pollen-derived estimates. Quaternary Research 22, 361–374.
Bartlein, P.J., Whitlock, C., 1993. Paleoclimatic interpretation of the Elk Lake pollen record. In: Bradbury, J.P., Dean, W.E. (Eds.), Evidence for rapid climate change in the North-Central, United States: The Elk Lake record. Geological Society of America Special Paper, vol. 276, pp. 275–293.
Bernabo, J.C., 1981. Quantitative estimates of temperature changes over the last 2700 years in Michigan based on pollen data. Quaternary Research 15, 143–159.
Birks, H.J.B., 1998. D.G. Frey & E.S. Deevey review #1—numerical tools in palaeolimnology—progress, potentialities, and problems. Journal of Paleolimnology 20 (4), 307–332.
Birks, H.J.B., Heiri, O., Seppa, H., Bjune, A.E., 2010. Strengths and weaknesses of quantitative climate reconstructions based on Late-Quaternary biological proxies. Open Ecology Journal 3, 68–110.
Bröcker, J., Smith, L., 2008. From ensemble forecasts to predictive distribution functions. Tellus A 60 (4), 663–678.
Cook, E.R., Briffa, K.R., Jones, P.D., 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. International Journal of Climatology 14, 379–402.
Cook, E.R., Anchukaitis, K.J., Buckley, B.M., D'Arrigo, R.D., Jacoby, G.C., Wright, W.E., 2010. Asian monsoon failure and megadrought during the last millennium. Science 328, 486. doi:10.1126/science.1185188.
Cook, E.R., Seager, R., Cane, M.A., Stahle, D.W., 2007. North American drought: reconstructions, causes, and consequences. Earth Science Reviews 81 (1–2), 93–134.
Cook, E.R., Woodhouse, C., Eakin, C.M., Meko, D.M., Stahle, D.W., 2004. Long-term aridity changes in the western United States. Science 306, 1015–1018.
Davis, O.K., 1995. Climate and vegetation patterns in surface samples from arid western U.S.A.: application to Holocene climatic reconstructions. Palynology 19, 95–117.
Dieckmann, U., Law, R., Metz, J.A.J. (Eds.), 2000. The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, Cambridge, UK, 564 pp.
Erästö, P., Holmström, L., 2006. Selection of prior distributions and multiscale analysis in Bayesian temperature reconstructions based on fossil assemblages. Journal of Paleolimnology 36 (1), 69–80.
Fraley, C., Raftery, A.E., 2002. Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association 97, 611–631.
Gachet, S., Brewer, S., Cheddadi, R., Davis, B., Gritti, E., Guiot, J., 2003. A probabilistic approach to the use of pollen indicators for plant attributes and biomes: an application to European vegetation at 0 and 6 ka. Global Ecology and Biogeography 12, 103–118.
Gajewski, K., 1988. Late-Holocene climates of eastern North America estimated from pollen data. Quaternary Research 29, 255–262.
Gajewski, K., Winkler, M., Swain, A.M., 1985. Vegetation and fire history from three lakes with varved sediments in northwestern Wisconsin (U.S.A.). Review of Palaeobotany and Palynology 44, 277–292.
Garreta, V., 2010. Bayesian approach of pollen-based palaeoclimate reconstructions: toward the modelling of ecological processes, Ph.D. Thesis, University of Aix-Marseille III, France.
Garreta, V., Miller, P., Guiot, J., Hély, C., Brewer, S., Sykes, M., Litt, T., 2010. A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model. Climate Dynamics 35, 371–389.
Gavin, D.G., Oswald, W.W., Wahl, E.R., Williams, J.W., 2003. A statistical approach to evaluating distance metrics and analog assignments for pollen records. Quaternary Research 60 (3), 356–367.
Gebhardt, C., Kühl, N., Hense, A., Litt, T., 2008. Variational reconstruction of Quaternary temperature fields. Climate Dynamics 30, 421–437.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, second ed. Chapman and Hall/CRC, 696 pp.
Gilks, W.R., Richardson, S., Spiegelhalter, D., 1995. Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC, 512 pp.
Grichuk, V.P., 1969. An attempt of reconstruction of certain climatic indexes of the northern hemisphere during the Atlantic stage of the holocene. In: Neishtadt, M. (Ed.), Proceedings of the 3rd International Palynological Congress, pp. 41–57. Golotsen, Nauka, Moscow.
Gritti, E.S., Gachet, S., Sykes, M.T., Guiot, J., 2004. An extended probabilistic approach of plant vital attributes: an application to European pollen records at 0 and 6 ka. Global Ecology and Biogeography 13, 519–533.
Guiot, J., Torre, F., Jolly, D., Peyron, O., Borreux, J.J., Cheddadi, R., 2000. Inverse vegetation modeling by Monte Carlo sampling to reconstruct paleoclimate under changed precipitation seasonality and CO2 conditions: application to glacial climate in Mediterranean region. Ecology Modelling 1, 119–140.
Guiot, J., Wu, H.B., Garreta, V., Hatté, C., Magny, M., 2009. A few prospective ideas on climate reconstruction: from a statistical single proxy approach towards a multi-proxy and dynamical approach. Climate of the Past 5, 571–583.
Hadfield, J., 2010. MCMC methods for multi-response Generalised linear mixed models: the MCMCglmm R Package. Journal of Statistical Software 33 (2), 1–22.
Haslett, J., Whiley, M., Bhattacharya, S., Salter-Townshend, M., Wilson, S., Allen, J.R.M., Huntley, B., Mitchell, F.J.G., 2006. Bayesian paleoclimate reconstruction. Journal of the Royal Statistical Society, Series A (Statistics in Society) 169, 395–438.
Haxeltine, A., Prentice, I., 1996. BIOME3: an equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability and competition among plant functional types. Global Biogeochemical Cycle 10, 693–709.
Hintikka, V., 1963. Über das Großklima einiger Pflanzenareale in zwei Klimakoordinatensystemen dargestellt. Annales Botanici Societatis Zoologicae Botanicae Fennicae, Vanamo, 1–64.

Howe, S., Webb III, T., 1983. Calibrating pollen data in climatic terms: improving the methods. Quaternary Science Reviews 2, 17–51.
Iversen, J., 1944. Viscum, Hedera and Ilex as climate indicators. Geologiska Foereningens Stockholm Foerhandlingar 66, 463–483.
Jackson, S.T., Williams, J.W., 2004. Modern analogs in Quaternary paleoecology: here today, gone yesterday, gone tomorrow? Annual Review of Earth and Planetary Sciences 32 (1), 495–537.
Kaplan, J.O., Bigelow, N.H., Prentice, I.C., Harrison, S.P., Bartlein, P.J., Christensen, T.R., Cramer, W., Matveyeva, N.V., McGuire, A.D., Murray, D.F., Razzhivin, V.Y., Smith, B., Walker, D.A., Anderson, P.M., Andreev, A.A., Brubaker, L.B., Edwards, M.E., Lozhkin, A.V., 2003. Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, and future projections. Journal of Geophysical Research 108 (D19), 8171.
Klotz, S., Müller, U., Mosbrugger, V., de Beaulieu, J.L., Reille, M., 2004. Eemian to early Würmian climate dynamics: history and pattern of changes in Central Europe. Palaeogeography, Palaeoclimatology, Palaeoecology 211, 107–126.
Korhola, A., Vasko, K., Toivonen, H., Olandor, H., 2002. Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modeling. Quaternary Science Reviews 21, 1841–1860.
Kühl, N., Gebhardt, C., Litt, T., Hense, A., 2002. Probability density functions as botanical–climatological transfer functions for climate reconstruction. Quaternary Research 53 (3), 381–389.
Kühl, N., Schölzel, C., Litt, T., Hense, A., 2007. Eemian and Early Weichselian temperature and precipitation variability in northern Germany. Quaternary Science Reviews 26, 3311–3317.
Kumke, T., Schölzel, C., Hense, A., 2004. Transfer Functions for Paleoclimate Reconstructions – Theory and Methods, The KIHZ Project: Towards a Synthesis of Holocene Proxy Data and Climate Models. Springer Verlag, pp. 229–244.
Latalowa, M., van der Knaap, W.O., 2006. Late Quaternary expansion of Norway spruce Picea abies (L.) Karst. in Europe according to pollen data. Quaternary Science Reviews 25, 2780–2805.
Li, B., Nychka, D.W., Ammann, C.M., 2007. The "Hockey Stick" and the 1990s: a statistical perspective on reconstructing hemispheric temperatures. Tellus 59, 591–598.
Li, B., Nychka, D.W., Ammann, C.M., 2010. The value of multi-proxy reconstruction of past climate (with discussions and rejoinder). The American Statistical Association 105, 883–895.
Litt, T., Schölzel, C., Kühl, N., Brauer, A., 2009. Holocene vegetation and climate history in the Western Eifel Volcanic Field (Germany) based on annually laminated lacustrine maar sediments. BOREAS 38 (4), 679–690.
Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., Wanner, H., 2004. European seasonal and annual temperature variability, trends and extremes since 1500. Science 303, 1499–1503.
Lytle, D., Wahl, E., 2005. Palaeoenvironmental reconstructions using the modern analog technique: the effects of sample size and decision rules. Holocene 15 (4), 554–566.
Mann, M.E., Bradley, R.S., Hughes, M.K., 1999. Northern Hemisphere temperatures during the past millennium – inferences, uncertainties, and limitations. Geophysical Research Letters 26, 759–762.
Mann, M.E., Bradley, R.S., Hughes, M.K., 1998. Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392, 779–787.
Mann, M.E., Zhang, Z., Hughes, M.K., Bradley, R.S., Miller, S.K., Rutherford, S., Ni, F., 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. Proceedings of the National Academy of Sciences of the United States of America 105, 13252–13257. doi:10.1073/pnas.0805721105.
Mann, M.E., Zhang, Z., Rutherford, S., Bradley, R.S., Hughes, M.K., Shindell, D., Ammann, C., Faluvegi, G., Ni, F., 2009. Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. Science 326, 1256–1260. doi:10.1126/science.1177303.
Neumann, F., Schölzel, C., Litt, T., Hense, A., Stein, M., 2007. Holocene vegetation and climate history of the northern Golan heights (Near East). Vegetation History and Archaeobotany 16, 329–346.
Overpeck, J.T., Webb, T., Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. Quaternary Research 23, 87–108.
Peyron, O., Guiot, J., Cheddadi, R., Tarasov, P., Reille, M., De Beaulieu, J., Bottema, S., Andrieu, V., 1998. Climatic reconstruction in Europe for 18,000 yr BP. from pollen data. Quaternary Research 49, 183–196.
Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A., Solomon, A.M., 1992. A global biome model based on plant physiology and dominance, soil properties and climate. Journal of Biogeography 19, 117–134.
Prentice, I.C., Guiot, J., Huntley, B., Jolly, D., Cheddadi, R., 1996. Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. Climate Dynamics 12, 185–194.
Pross, J., Klotz, S., Mosbrugger, V., 2000. Reconstructing palaeotemperatures for the Early and Middle Pleistocene using the mutual climatic range method based on plant fossils. Quaternary Science Reviews 19, 1785–1799.
Rutherford, S., Mann, M.E., Osborn, T.J., Bradley, R.S., Briffa, K.R., Hughes, M.K., Jones, P.D., 2005. Proxy-based Northern Hemisphere surface temperature reconstructions: sensitivity to method, predictor network, target season, and target domain. Journal of Climate 18, 2308–2329.
Sawada, M., 2006. An open source implementation of the Modern Analog Technique (MAT) within the R computing environment. Computers & Geosciences 32, 818–833.
Sawada, M., Viau, A.E., Vettoretti, G., Peltier, W.R., Gajewski, K., 2004. Comparison of North-American pollen-based temperature and global lake-status with CCCma AGCM output at 6 ka. Quaternary Science Reviews 23, 225–244.
Schölzel, C., 2006. Palaeoenvironmental transfer functions in a Bayesian framework with application to Holocene climate variability in the Near East. In: Bonner Meteorologische Abhandlungen, vol. 62, Bonn, Germany, 104 p.
Schölzel, C., Hense, A., 2011. Probabilistic assessment of regional climate change in Southwest Germany by ensemble dressing. Climate Dynamics 36, 2003–2014.
Schölzel, C., Hense, A., Hübl, P., Kühl, N., Litt, T., 2002. Digitization and georeferencing of botanical distribution maps. Journal of Biogeography 29, 851–856.
Schölzel, C., Friederichs, P., 2008. Multivariate non-normally distributed random variables in climate research – introduction to the copula approach. Nonlinear Processes in Geophysics 15, 761–772.
Smerdon, J.E., Kaplan, A., Chang, D., Evans, M.N., 2010. A Pseudoproxy evaluation of the CCA and RegEM Methods for reconstructing climate fields of the last millennium. Journal of Climate 23, 4856–4880.
Smith, B., Prentice, I.C., Sykes, M.T., 2001. Representation of vegetation dynamics in modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space. Global Ecology and Biogeography 10, 621–637.
Sugita, S., 2007. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. The Holocene 17, 243–257.
Tarasov, P., Cheddadi, R., Guiot, J., Bottema, S., Peyron, O., Belmonte, J., Ruiz-Sanchez, V., Saadi, F., Brewer, S., 1998. A method to determine warm and cool steppe biomes from pollen data; application to the Mediterranean and Kazakhstan regions. Journal of Quaternary Science 13, 335–344.
ter Braak, C.J.F., Juggins, S., 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental parameters from species assemblages. Hydrobiologia 269-270, 485–502.
Tingley, M.P., Craigmile, P.F., Haran, M., Li, B., Mannshardt-Shamseldin, E., Rajaratnam, B., 2010. Piecing Together the Past: Statistical Insights into Paleoclimatic Reconstructions. Technical Report No. 2010-09. Department of Statistics, Stanford University.
Tingley, M.P., Huybers, P., 2010. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: Development and applications to paleoclimate reconstruction problems. Journal of Climate 23, 2759–2781.
Toivonen, H.T.T., Mannila, H., Korhola, A., Olander, H., 2001. Applying Bayesian statistics to organism-based environmental reconstruction. Ecological Applications 11 (2), 618–630.
Troccoli, A., Harrison, M., Anderson, D., Mason, S., Stephenson, D., 2008. An introduction to probability forecasting. In: Seasonal Climate: Forecasting and Managing Risk. NATO Science Series. Springer, Netherlands, pp. 235–257.
Vasko, K., Toivonen, H.T.T., Korhola, A., 2000. A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. Journal of Paleolimnology 24, 243–250.
von Storch, H., Zorita, E., Gonzalez-Rouco, F., 2009. Assessment of three temperature reconstruction methods in the virtual reality of a climate simulation. International Journal of Earth Science 98, 67–82.
Wahl, E.R., 2002. Paleoecology and testing of paleoclimate hypotheses in southern California during the Holocene. Ph.D. Dissertation, University of Minnesota, 248 pp. (Relevant portion for citations herein available in "pdf" format from author.)
Wahl, E.R., 2003. Assigning climate values to pollen surface sample sites and validating modern analog climate reconstructions in the southern California region. Madroño 50 (4), 271–285.
Wahl, E.R., 2004. A general framework for determining cutoff values to select pollen analogs with dissimilarity metrics in the modern analog technique. Review of Palaeobotany and Palynology 128, 263–280.
Wahl, E., Ammann, C., 2007. Robustness of the Mann, Bradley, Hughes reconstruction of Northern Hemisphere surface temperatures: examination of criticisms based on the nature and processing of proxy climate evidence. Climatic Change 85, 33–69.
Wahl, E., Schölzel, C., Williams, J., Tigrek, S., 2010. Comment: the use and value of Bayesian Hierarchical modelling for paleoclimate reconstruction: including a new pollen-climate data model. The American Statistical Association 105 (491), 900–905.
Wahl, E.R., Diaz, H.F., Ohlwein, C. A pollen-based reconstruction of summer temperature in central North America and implications for circulation patterns during medieval times. Global and Planetary Change, doi:10.1016/j.gloplacha.2011.10.005.
Wahl, E.R., Smerdon, J. Comparative performance of spatially-explicit paleoclimate reconstructions derived from real proxies and simulated proxies. Geophysical Research Letters, submitted.
Wilks, D., Hamill, T., 2007. Comparison of ensemble-MOS methods using GFS reforecasts. Monthly Weather Reviews 135, 2379–2390.
Whitmore, J., Gajewski, K., Sawada, M., Williams, J.W., Shuman, B., Bartlein, P.J., Minckley, T., Viau, A.E., Webb III, T., Shafer, S., Anderson, P., Brubaker, L., 2005. Modern pollen data from North American and Greenland for multi-scale paleoenvironmental applications. Quaternary Science Reviews 24, 1828–1848.
Williams, J.W., 2003. Variations in tree cover in North America since the last glacial maximum. Global and Planetary Change 35 (1–2), 1–23.
Williams, J.W., Shuman, B., 2008. Obtaining accurate and precise environmental reconstructions from the modern analog technique and North American surface pollen dataset. Quaternary Science Reviews 27, 669–687.