Paleoclimate reconstruction: looking backwards to look forward

Peter F. Craigmile  
*The Ohio State University, Columbus, Ohio*

Murali Haran  
*The Pennsylvania State University*

Bo Li  
*University of Illinois at Urbana–Champaign*

Elizabeth Mannshardt  
*North Carolina State University, Raleigh*

Bala Rajaratnam  
*University of California, Davis*

Martin Tingley  
*The Pennsylvania State University*

**CONTENTS**

33.1 Introduction ................................................................. 768
33.2 Paleoclimate reconstruction: looking backwards ........................ 769
  33.2.1 A multiproxy, multiforcing IR reconstruction .................... 770
  33.2.2 Statistical issues in paleoclimate reconstruction ................ 771
    33.2.2.1 Incorporating climate forcings in the HBM .................. 771
    33.2.2.2 Handling proxies separately .................................. 772
    33.2.2.3 Modeling temporal dependence .............................. 772
    33.2.2.4 Modeling spatial dependence and spatio-temporal reconstructions .............................................. 773
    33.2.2.5 Missing values and data augmentation ...................... 774
    33.2.2.6 Temporal uncertainty ......................................... 774
    33.2.2.7 Non-Gaussian paleoclimate reconstruction ................ 774
    33.2.2.8 Multivariate reconstructions ................................. 775
  33.2.3 Ideas and good practices .......................................... 775

33.3 Climate models and paleoclimate ..................................... 776
  33.3.1 Climate model assessment using paleoclimate reconstructions ... 776
  33.3.2 Statistical issues .................................................. 778
    33.3.2.1 Considering and embracing the lack of independence .... 778
    33.3.2.2 Refining model components ................................. 778
  33.3.3 Research directions .............................................. 779

767
33.1 Introduction

Paleoclimatology is the study of past climate. While in rare cases we may be able to study past climate from “directly observed” historical record, typically we infer past climate indirectly through the use of proxies. These proxies, such as tree rings, ice cores, corals, and pollen, have climate-sensitive characteristics that can be measured (e.g., tree ring density; the ratio of dissolved oxygen isotopes in ice cores and coral; the signature and relative abundance of pollen species). By associating these climate sensitive measurements to observed instrumental records of climate such as temperature and precipitation, we are able to predict or hindcast past climate along with, hopefully, an associated measure of uncertainty for our predictions. This task is called paleoclimate reconstruction. For example, tree ring densities of certain species of trees in the upper latitudes, suitably normalized to remove growth effects, are approximately linearly associated with temperature. By modeling the relationship between tree ring density and observed temperature, we can produce predictions of past temperature.

In paleoclimate reconstruction problems, climate scientists differentiate between two types of prediction or hindcast problems: index reconstructions (IR) and climate field reconstruction (CFR) [47]. In IR we predict climate on the basis of index series (typically averages over larger spatial scales such as Northern hemispheric or El Niño Southern Oscillation (ENSO) reconstructions; see [2] for an example). Some versions of IR are also known as “composite plus scale” (CPS) [47]. IR seeks to produce hindcasts in time, and is a simpler statistical problem compared to CFR, which seeks to produce hindcasts in both space and time. Thus, as with many other studies carried out in climate, paleoclimate reconstruction is a spatio-temporal statistical problem. This problem is challenging because the spatial and temporal domains of the proxy and instrumental records may not overlap. Instrumental records are only available more recently in time, but have reasonably good spatial and temporal coverage. Proxy measurements are expensive to collect and are typically observed sparsely in space, and proxy records typically have longer but more uncertain temporal coverage. The paleoclimate reconstruction often involves multiple proxies. Since different proxies are sensitive to climate in varied ways, the relationships between different proxy measurements and instrumental records need to be modeled distinctly and carefully. If multiple climate variables are to be reconstructed simultaneously the paleoclimate reconstruction can become a multivariate problem. For example, the reconstruction is a bivariate problem if we attempt to reconstruct temperature and precipitation together.

In this chapter we review the statistical methodologies underlying paleoclimate reconstruction, providing some commentary on existing and outstanding problems for statisticians. We also discuss more recent research that blends the study of paleoclimate with climate models, investigating how researchers should compare reconstructions to climate model output. We close with a discussion of further ideas and research in statistical paleoclimatology.
Paleoclimate reconstruction: looking backwards to look forward

33.2 Paleoclimate reconstruction: looking backwards

Informally, a paleoclimatic reconstruction produces predictions or hindcasts (and measures of uncertainty) of an underlying climate process, given direct observations of the climate system and indirect observations of climate-sensitive paleo proxies.

Formally, we introduce paleoclimate reconstruction in terms of a hierarchical Bayesian model (HBM) that link direct and indirect measurements to some underlying climate process of interest. (Section 33.2.2 contains many references to Bayesian and non-Bayesian statistical methods used for paleoclimate reconstruction.) We introduce a simplified description of a framework introduced by [91]. (Also see [43] for a review of other methods of reconstruction.) For an IR we are interested in predicting a climate process series \( \{C_t : t \in T\} \), where \( T \) is the time domain of interest (typically \( T \subset \mathbb{Z} \) for a discrete-time climate series, but it is also possible that \( T \subset \mathbb{R} \) for a continuous-time series, which can be useful for modeling time-uncertain proxy data). In the context of CFR, our climate process is \( \{C(s,t) : s \in D, t \in T\} \) for some spatial domain \( D \) and temporal domain \( T \). Commonly the spatial domain of interest is some region \( D \subset \mathbb{R}^2 \). Without loss of generality we will discuss the IR case in the definition of the HBM – with the addition of a spatial index for the models that we define below the CFR case follows similarly.

Now we introduce processes that capture the proxies and instrumental records. With one proxy and one instrumental record, let \( \{P_t : t \in T\} \) denote the proxy process and \( \{Z_t : t \in T\} \) the associated instrumental process. For example with \( \{C_t\} \) being the average latent temperature series over a given spatial domain, \( \{P_t\} \) could be the average tree ring density over that domain, and \( \{Z_t\} \) the observed average instrumental temperature, again over the same domain.

An HBM for paleoclimate reconstruction then requires the following model components to be defined:

1. The model for the instrumental process, \( f(\{Z_t\}|\{C_t\}, \theta_Z) \). This model is conditioned on the underlying climate process and depends on a vector of parameters \( \theta_Z \). In the simplest setting we could imagine that \( Z_t \) is a noisy version of \( C_t \) for each \( t \in T \): Assuming independence over time, with normal measurement errors, we have

\[
Z_t \sim \mathcal{N}(C_t, \sigma_Z^2), \quad t \in T,
\]

where \( \sigma_Z^2 \) is the measurement error. The likelihood is then

\[
f(\{Z_t\}|\{C_t\}, \theta_P) = \prod_t d\mathcal{N}(Z_t; C_t, \sigma_Z^2),
\]

where \( d\mathcal{N}(z; \mu, \sigma^2) \) is the probability density function (pdf) of a \( \mathcal{N}(\mu, \sigma^2) \) random variable (RV) evaluated at \( z \). We can see that \( \theta_Z = (\sigma_Z^2)^T \) for this simpler model. Note that the product in the likelihood for the instrumental process is over time points for which we have instrumental data (typically “recent” in time); the time points for which we will typically have proxy data will be different (“in the past” and “recent”). We discuss extensions below.

2. The model for the proxy process, \( f(\{P_t\}|\{C_t\}, \theta_P) \). This likelihood is again conditioned on the underlying climate process and depends on a vector of parameters \( \theta_P \). Writing down the likelihood for the proxy process is more involved than the model for the instrumental process as we need to specify how the proxy is related to the climate process. In paleoclimatology the functional relationship between
the proxy and climate process is called the *forward model* [23, 24, 91]. In statistical paleoclimatology we often write the forward model as a distribution that disentangles the deterministic and random components of the proxy, the latter of which can simply be measurement errors. A simple example used for tree ring density is to assume that after accounting for measurement error, there is a linear relationship between tree ring density and temperature; in this case, assuming conditional independence we suppose

\[ P_t \sim \mathcal{N}(\beta_0 + \beta_1 C_t, \sigma_P^2), \quad t \in T, \]

for regression parameters \( \beta_0 \) and \( \beta_1 \) and measurement error variance \( \sigma_P^2 \). The likelihood in this case is

\[ f(\{P_t\}|\{C_t\}, \theta_P) = \prod_t \mathcal{N}(P_t; \beta_0 + \beta_1 C_t, \sigma_P^2), \]

with \( \theta_P = (\beta_0, \beta_1, \sigma_P^2)^T \). Again we discuss more general cases below.

3. The prior distribution for the climate process, \( \pi(\{C_t\}|\theta_C) \), which depends on another vector of parameters \( \theta_C \) (we discuss the prior distributions of the other model parameters below). In the IR case, typically a prior distribution is provided that captures the smoothness of the climate process in time (e.g., an autoregressive time series model [5, 54], or uses covariate information from climate forcings). Further details are discussed below.

To complete the model, in addition to prior of the climate process we also need specific prior distributions, \( \pi(\theta) \), for the remaining parameters \( \theta = (\theta_Z, \theta_P, \theta_C)^T \). Informative priors are commonly used in the specification of this part of the model [5, 54, 92].

Given the specification of the likelihood for the instrumental and proxy processes, and the prior distributions for the climate process and other model parameters, the posterior distribution \( \pi(\{C_t\}, \theta|\{Z_t\}, \{P_t\}) \) satisfies

\[ \pi(\{C_t\}, \theta|\{Z_t\}, \{P_t\}) \propto f(\{Z_t\}|\{C_t\}, \theta_Z) f(\{P_t\}|\{C_t\}, \theta_P) \pi(\{C_t\}|\theta_C) \pi(\theta). \]

In most situations the posterior distribution is not available in a closed-form expression and Markov chain Monte Carlo is used to provide a sample-based approximation to the posterior distribution [5, 54, 92].

For an IR, the posterior distribution of climate typically has reduced variance for more recent time periods, and the uncertainty increases as we go back in time. This reflects the fact that we have more proxy and instrument records in recent time periods. The imprint of climate from proxy records through the forward model, along with the prior distribution of the climate process, drives the posterior distribution of climate as we move further and further into the past. For the CFR, spatio-temporal reconstructions have more uncertainty over space and time, due to the sparsity of instrumental and proxy records. Clearly, the selection of meaningful and accurate prior models for the climate process is key to producing more informative paleoclimate reconstructions. For such purposes, information from climate model runs can be used to form a prior that can compensate for the sparsity of data.

To further illustrate this HBM framework we next work through a detailed example of a paleoclimate reconstruction.

### 33.2.1 A multiproxy, multiforcing IR reconstruction

The power of a hierarchical Bayesian model (HBM) lies in its flexibility in modeling the complex structure or relationship among multiple data sets. The paleoclimate reconstruction
usually involves data from different sources and thus the HBM naturally provides an efficient and feasible framework for integrating paleodata of various characteristics. As an example for an index reconstruction, suppose we have different types of proxies, $P_1, \ldots, P_m$. Each proxy records the climate via a different mechanism and the statistical model should respect each proxy’s characteristics. For example, if the proxy responds to climate variation through a very slow evolving process such as the heating diffusion into the earth and rocks, then the proxy only relates to the very low frequency variation of the climate change. On the contrary, if the proxy responds quickly to annual or even seasonal variation of climate, then the proxy mainly relates to the high frequency components of climate change. For these reasons, we model the proxy data in the hierarchy as below:

$$
P_{1t} = \beta_{10} + \beta_{11} g_1(\{C_t\}) + \epsilon_{1t},
$$
$$
\vdots
$$
$$
P_{mt} = \beta_{m0} + \beta_{m1} g_m(\{C_t\}) + \epsilon_{mt},
$$
(33.1)

where $g_i(\cdot)$ for each $i = 1, \ldots, m$ are transformation functions that link the climate and proxies. These transformation functions can either be linear or nonlinear, depending on how the proxy reflects the climate evolution. The error terms $\{\epsilon_{it}\}, i = 1, \ldots, m$, are due to both the measurement errors in proxies and the imperfect linear relationship between the proxy and the transformed climate. The instrumental climate can contain a substantial amount of noise relative to the true climate. Therefore it is more appropriate to explicitly model those measurement errors in instrumental climate in order to take all major errors into account in the reconstruction:

$$
Z_t = \beta_0 + \beta_1 C_t + \zeta_t,
$$

where $\{\zeta_t\}$ may be modeled as white noise or a dependent time series process.

After modeling the likelihood of the observations, we now focus on the prior for climate, $\{C_t\}$. Depending on the purpose of the reconstruction, we can either introduce the climate forcings into the prior or simply model the climate as a stochastic process. Suppose including forcings is appropriate, and using the solar irradiance $S_t$, volcanism $V_t$, and greenhouse gases $G_t$ as examples for forcings, the prior for climate can take the form of

$$
C_t = \alpha_0 + \alpha_1 S_t + \alpha_2 V_t + \alpha_3 G_t + \eta_t,
$$
(33.2)

where $\{\eta_t\}$ are the errors in temperature that cannot be described by the variability in forcings. When it is not appropriate to include forcings, then we can simply model the prior as

$$
C_t = \eta_t.
$$
(33.3)

The models for $\{\epsilon_{it}\}$ in (33.1), and $\{\eta_t\}$ in (33.2) or (33.3) are discussed in Section 33.2.2.3.

### 33.2.2 Statistical issues in paleoclimate reconstruction

#### 33.2.2.1 Incorporating climate forcings in the HBM

The variabilities in different proxies are driven by the climate evolution and thus they serve as the noise-contaminated reflection of climate change. Forcings, on the contrary, are the internal drivers of the climate evolution, and any variation in forcings will eventually lead to fluctuation in the climate system. The major forcings include solar irradiance, volcanism, greenhouse gases, and aerosols, the first two of which are natural forcings, while the latter are anthropogenic forcings. We can also include the effect of orbital forcings [22] upon climate.
Since the forcings determine the large scale climate variation, it is very natural to employ them to improve the reconstruction. The four major forcings are estimated via different techniques. For example, the solar irradiance is derived from measurements of fluctuations of Berium $^{10}\text{Be}$ production rates which is modulated by solar magnetic variability, and the volcanic series is based on synthesis of individual ice cores and in some cases on historical records of large eruptions. Both [54] and [5] demonstrated that including forcings can make the temperature reconstruction better calibrated and more sharp. Although forcings play a very important role, when the primary interest is to use paleoclimate reconstructions to verify climate models, it is more appropriate to perform the reconstruction with no forcings included. Otherwise the modeling is circular in the sense that the forcings are also major inputs for the algorithm built in climate models.

### 33.2.2.2 Handling proxies separately

There is a large variety of climate proxies, including, measurements on tree rings, pollen assemblage, ice core, corals, boreholes, and speleothems [47, 65, 71]. The various proxies differ in their temporal resolutions. Tree rings and some corals provide an observation each year, and are accurately dated by layer counting. Pollen assemblage retain a smoothed record of climate variation due both to the persistency properties of mature plants [10] and the low-resolution dating of the lake core sediments from which pollen assemblages are extracted. Speleothem and borehole depth profiles behave as low pass filters that only retain the long term trends of climate.

The noise structures of different proxies are also characterized differently. For example, the error terms in the tree ring–climate relationship largely reflect that trees are imperfect recorders of local climate. With pollen data, however, an important source of uncertainty is due to the substantial dating errors in the radiocarbon technology. Another difference among different proxies is their relationship to climate variation. The relationship between pollen and temperature is sigmoidal, and therefore a logistic model has been developed to capture this nonlinear relationship.

Clearly, it is not optimal to treat all different proxies in the same way when trying to extract climate signal from them. [54] proposed a Bayesian framework to integrate proxies of different characteristics into one single coherent climate reconstruction. They studied the added value of tree ring, pollen, and borehole proxies representing, respectively, high, medium, and low frequency components of the climate signal. If forcings are not included in the model, they found that moderate to low frequency proxies need to be added to compensate for the information loss.

There is a growing literature on the development of scientifically-motivated forward models that describe how proxies record variations in the climate system [11, 23, 24, 97]. Hierarchical statistical modeling, combined with Bayesian inference, provides a framework for specifying and fitting a rich variety of models that are in line with current scientific understanding of the proxies.

### 33.2.2.3 Modeling temporal dependence

Climate reconstructions, either with direct or indirect regression models (see a discussion for these two frameworks in [4]), or with Bayesian hierarchical models, generally treat errors as white noise or short memory time series model such as autoregressive (AR) models with order 1 or 2 [5, 54, 90, 92, 93, 94, 95, 99, 100]. [70] and [82] demonstrate the choice of AR order could differ by the type of proxy. [36] use models based on Brownian motion and [75] use a stochastic volatility normal-inverse Gaussian model. [53] use principal components analysis (PCA) (also known as empirical orthogonal functions, EOF) in an IR context. Also worth noting is the method by [31] which models the temporal dependence nonparametrically.
Paleoclimate reconstruction: looking backwards to look forward

There is evidence [19, 44, 45, 84] that the climate dynamics is governed by a long-memory stochastic process (see, e.g., [6] for a definition), which raises an interesting question on whether long-memory error models would be helpful with the reconstruction or uncertainty quantification. A further question is which error model is most appropriate for modeling the relationship between proxy and temperature and between temperature and forcings. Under the Bayesian modeling framework, [5] investigated this question by using both the land-surface temperature and the combined land and ocean temperature data sets with the [59] proxy network. [5] reveal that when forcings are included in the reconstruction model then reconstructions are robust to the choice of different time series models for the error (there is a suggestion that a long memory model for the errors improve the accuracy of the uncertainty quantification).

33.2.2.4 Modeling spatial dependence and spatio-temporal reconstructions

Due to spatially wide-spread data availability, many paleoclimate reconstruction efforts have primarily focused on reconstructing Northern hemispheric average temperature anomalies. Such reconstructions correspond to hindcasting a single time-series; for examples (not comprehensive) see [21, 46, 52, 55, 60, 61] for an overview and discussion in [47]. [73] consider IR for each continent. There has also been considerable effort to undertake paleoclimate reconstructions spatio-temporally for particular regions, hemispheres, or for the entire globe. These entail reconstructing a random field.

Climate field reconstructions (CFR) are inherently more difficult as there is an immediate need to specify (either directly or indirectly) a spatial model to relate temperature and proxies to other temperature and proxies, which are located at different spatial and temporal points. The problem is exacerbated by the fact that the number of grid points on even a coarse spatial grid by far exceed the number of years of proxy data. Hence there is a need to obtain a parsimonious spatial representation to overcome the dimensionality issue. A commonly used approach is to use a covariance function to model the relationships in space and time [92, 93]. [92] introduce the BARCAST method that, in addition to being a multiproxy reconstruction method, introduces a spatio-temporal model for the climate process, \( \{C(s,t)\} \) (BARCAST stands for “a Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time”). The authors suppose that climate process is Gaussian with a constant mean. The covariance structure for the climate process is separable – AR(1) in time, and exponential and isotropic in space. Applications of this method include development of an imputation method to improve the estimation of climate anomalies from time series with missing data [90], a 400 year reconstruction of high northern latitude temperatures with a focus on recent extremes [94], an analysis of the divergence between temperature and tree ring records [95], and a multiproxy reconstruction of European temperatures [58]. The exponential covariance function and the assumption of separability in time and space implicit in the BARCAST methodology can be relaxed further. Section 33.2.2 discusses many of the recent state-of-the-art methods in multiproxy spatial paleoclimate reconstructions. These methods have explored more sophisticated spatial modeling approaches.

Another approach is to use regularization approaches to deal with the dimensionality issue as in Reg-EM [81]. Here the ill-conditioning of the covariance of the field is overcome by using a ridge-type penalty. [83] uses canonical correlation analysis (CCA) to undertake dimensionality reduction. Earlier work on CFR is also seen in [60] that uses inverse ordinary least squares regression in combination with principal component analysis, and [62] and [77] that apply regularized expectation maximization using truncated total least squares. Recently, [31] uses Markov random fields (MRF) to model the covariance of the spatial field. The MRF approach leverages the conditional independences in space in a natural way.
MRF models may often be parameterized with fewer parameters as compared to specifying a model for the covariance matrix directly. [56] and [57] investigated methods to compare different climate fields including CFRs.

There is also a rich literature on space-for-time substitution reconstruction methods [35, 75, 78, 88]. Assuming ergodicity of the relevant processes, the modern spatial distribution of climates and proxies is used to calibrate the proxy-climate relationship that is used for predictions. These models do not model the spatial dependence directly, but instead they use the heterogeneous proxy observed in recent time periods to describe the relationship between proxies and climate in the past.

33.2.2.5 Missing values and data augmentation

The very nature of temperature and proxy measurements gives rise to records that are often missing in both space and time. Moreover, the lengths of proxy records are quite different for the various classes of proxies. Discarding proxy time series with missing values is not a viable option as doing so can lead to a loss of valuable information. Thus paleoclimate reconstructions on real data (versus reconstructions on simulated pseudoproxy data) have to contend with missing values. Perhaps the most popular approach to deal with missing values is to use the EM algorithm, and its variant, Reg-EM [81]. Reg-EM casts the entire reconstruction problem as an imputation problem where missing values in the proxies/temperature are treated similarly, regardless of whether they appear in the calibration period or in the hindcast period. More specifically, for each year [81] uses the multivariate normal distribution to jointly model both the temperature and proxy records on a spatial grid. Reg-EM uses the EM algorithm to estimate the mean and covariance of the joint distribution of the temperature and proxies. Various regularizations can be used in sample-starved settings in order to obtain stable parameter estimates (Reg-EM is an example of this). The parameter estimates are then used to impute the missing values in the spatial field. Imputation of missing values in the hindcast period are then considered to form the reconstruction. The advantage of the EM algorithm approach is that it is rooted in established statistical theory.

33.2.2.6 Temporal uncertainty

As discussed above, the dating of proxies can be an important source of uncertainty in paleoclimate reconstructions, especially when observing long, but time-uncertain, proxy records. Significant development has been made to develop statistical models that account for time-uncertain proxies – examples include [1, 3, 33, 35, 75].

In the context of HBMs, [100] describe how dating uncertainty, for layer counted proxies such as tree rings, can be incorporated into a hierarchical framework for space-time reconstruction. The key insight is that the inferred spatial covariance of the climate variable(s) imposes a constraint on the population of chronologies associated with the time-uncertain proxies. [100] assume that, for each time-uncertain proxy, there exists an ensemble of age models that are a priori equally likely. These probabilities are then updated based on the current draw of the climate variables. In this way, simultaneous inference on both the chronologies and the space-time climate process can be achieved. The same concept can be extend to proxies dated using radiocarbon or other isotopes. Indeed, common chronology models, such as the Bayesian accumulation (BACON) model of [8], could be imbedded with the hierarchical model, as described in [100].

33.2.2.7 Non-Gaussian paleoclimate reconstruction

Gaussian spatial and temporal models are often used for characterizing proxy measurements such as tree ring width and density, and chemical measurements such as isotope ratios. This
is certainly common for hindcasting past temperatures. In more recent years there has been a growing need to work with random variables that are not continuous, or, if continuous, do not follow a Gaussian distribution.

Pollen counts or their relative abundance is an example of a proxy measurement that is discretely valued. For example, [72] consider the reconstruction of temperature based on counts of pollen from some indicator taxon. They employ a binomial distribution to model the counts, associating larger probabilities of abundance with warmer temperatures. (Also see [72] for a thorough review of pollen-based reconstruction methods.) Multivariate versions of pollen-based reconstructions include using Dirichlet-multinomial [36], zero-inflated extensions [78], and compound Poisson-Gamma processes [75].

Non-Gaussian continuous climate variables occur in paleoclimate studies of precipitation and accumulation [12, 68, 97] and the analysis of climate extremes. For example, in the latter case, [64] investigate spatially varying trends and dependencies in the parameters characterizing the distribution of extremes of a proxy dataset through the development of an HBM that implements spatially varying coefficients. This directly reconstructs extremal climate behavior by linking extremes in the proxy record to extremes in the instrumental record, rather than modeling mean climate behavior. See also, e.g., [49], [69], and [18] for other applications of extreme value theory to modeling climate proxies.

33.2.2.8 Multivariate reconstructions

An especially challenging problem is to reconstruct several climate variables at once on the basis on proxy measurements. In the HBM construction defined above we need the forward model to be a function of more than one climate variable. Statistical inference is challenging as there is often identifiability or a lack of Bayesian learning if we try to infer about multiple climate variables jointly on the basis of a single proxy. To tackle this issue we need to assume informative priors on the climate variables, including the relationship between them. Even so, the computations are non-trivial. In the context of tree ring widths, see [97]; these ideas are further explored in [96]. Another solution is to use more than one proxy in the reconstruction, each with a different relationship with the climate variables.

The space-for-time substitution methodology provides an alternative strategy to construct multiple climate variables simultaneously. For example [36] reconstruct the mean temperature of the coldest month as well as the growing degree days above 5°C in prehistoric times at a location in Ireland from fossilized pollen.

There is also a large literature focused on combining numerous climate variables into a single index, such as the Palmer Drought Severity Index [74], which are then reconstructed from climate-sensitive proxies [15, 16, 17].

33.2.3 Ideas and good practices

Many of the papers discussed here, reconstructing a single climate time series within our HBM framework on the basis of instrumental, proxy, and possibly forcing series. While there are challenges in specifying the forward model and the other time series models in the HBM, the statistical methodology required to producing reconstructions (i.e., sampling from one time series given some other time series) is straightforward. Computations are typically not particularly challenging in this fairly rich information-in-time scenario. For instance, our research group has been able to carry out computations on personal laptops, with code for a fairly routine Markov chain Monte Carlo algorithm implemented in the language R [76].

As we move to CFRs, the challenges increase. We now need to build spatio-temporal models that accurately capture the dependencies we observe in the data and the climate variables of interest. Proxy records can be sparse in space, but rich in time. Instrumen-
Handbook of Environmental and Ecological Statistics

tal records are often data products, and not raw measurements. Instrumental datasets are often massive. Spatio-temporal models for climate variables are not easy to specify. The increasing complexity of specifying and fitting the models, in combination with a greater need for more prior information about model components, leads to more demanding statistical inference. For instance, the computations would typically require specialized techniques such as sparse matrix algorithms or reduced-rank approaches for high-dimensional spatial or spatio-temporal data sets. But, this should not deter us from using the HBM approach for CFR. Understanding a complicated model via a number of conditional distributions, while fully propagating uncertainty quantification is exactly what HBMs excel at, and is what a successful paleoclimate reconstruction requires.

A key challenge moving forward is to understand how a climate process behaves spatio-temporally. In the next section, we discuss the role that climate models can play in helping us to understand climate, and their relationship to paleoclimate reconstructions.

33.3 Climate models and paleoclimate

Climate models are mathematical models used to simulate various aspects of and interactions between climate system processes, and can include the oceans, atmosphere, and the land surface. Climate models play a central role in modern climate science, and are a critical tool for projecting the future climate under various emission scenarios; see, for example, the reports of the Intergovernmental Panel on Climate Change (IPCC) ([14, 27, 51, 86]). Climate models and climate reconstructions from proxies are two different, if overlapping, sources of information on past and projected future climate. For instance, paleoreconstructions may be used to discriminate between competing climate models. Paleoreconstructions may also be used to assess or calibrate (infer the parameters of) climate models that are used to make projections about future climate (see Section 33.3.3.1). Alternatively, climate models may also be used to improve paleoreconstructions (see Section 33.3.2.1). Hence there are interesting challenges and opportunities in studying the interaction between these sources of information. Here we outline some basic questions at the intersection of the analysis of paleoclimate reconstructions and climate models, along with a discussion of some existing work in this area and avenues for future research.

33.3.1 Climate model assessment using paleoclimate reconstructions

There is a growing body of literature focused on comparing climate model output to late Holocene paleoclimate reconstructions; recent reviews include [66], [65], and [79]. As discussed in [89], a wide range of tools, of varying levels of mathematical sophistication, have been brought to bear on the problem. These include qualitative comparisons [50, 63], fuzzy logic [32], selection from an ensemble of climate simulations based on distance metrics [30], data assimilation (e.g., [29] and Chapter 35), and ideas from detection and attribution [37, 38].

Statistically rigorous approaches for studying climate models via paleoclimate reconstructions have recently been considered in [89] and in [67], the latter building upon a series of papers [39, 40, 87]. Both papers address the same question: how can the paleoclimate reconstructions be used to discriminate between different climate models and different forcing scenarios?

To fix ideas, consider a study of global temperatures from $k$ different global circulation models (GCMs) for a single volcanic forcing scenario. Both the $k$ model outputs and an
ensemble of paleoclimate reconstructions of the same global temperature time series are available to learn about the forced and unforced components of the temperature series. The ensemble of paleoclimate reconstructions is necessary to account for the uncertainty in the reconstructions – using a single “best estimate” reconstruction would ignore potentially large uncertainties in our reconstruction. Assuming a Bayesian reconstruction of the time series, for instance produced using BARCAST [92], it is straightforward to simply use the time series draws from the posterior distribution; such an ensemble is a byproduct of a standard Markov chain Monte Carlo (MCMC) simulation-based approach to Bayesian inference.

A starting assumption is that the paleoclimate reconstructions provide an approximation to the true temperature series. The ensemble accounts for uncertainties about the true temperature. Of interest is the following question: How well does the climate model-simulated temperature series “predict” (i.e., capture) the behavior of the true series? A simple approach to answer this question would be to regress the paleoclimate reconstructed series as the response with the climate model simulation as predictor. Such an approach, however, essentially treats each model-forcing combination separately and does not allow for a study of the relationships between the different components of the reconstructions and the model. In particular, what is often of interest is the component of the “true” (latent) series that may be attributable to a forcing. Hence, a useful framework would explicitly model the forced and unforced components of the climate.

An HBM approach makes it easy to specify common latent climate processes and relate them to different sets of data. For instance, this permits a model specification as follows: model how the paleoclimate reconstruction relates to the latent temperature series, which in turn is modeled as the sum of the forced and unforced components of the temperature series. In other words we assume that the latent temperature series is made up of a component due to the forcings plus a component due to the natural variability of the climate system. Similarly, the climate model-simulated series may be decomposed into forced and unforced components. Then, to relate the reconstructed and simulated temperatures, a linear regression model may be specified between their respective forced components, with the forced component of paleoclimate reconstruction as the predictor and the forced component of the model output as the response. The residual from this regression contains some information on how the forced components differ.

The unforced components for the paleoclimate reconstructions and the climate model simulations are assumed to arise from the same independent generating process. In this approach, information about the distribution of the natural variability of the climate system comes from control runs (climate model simulations that are run without forcings). More generally, because there are \( k \) different GCM runs, it is possible to tease apart the forced components from the unforced components. The methodology outlined above involves repeating the statistical analysis separately for each set of forcing combinations. As a result of such an analysis, for instance, [89] are able to show that of the three volcanic forcing scenarios they consider, one results in superior agreement with the paleoclimate reconstruction. Superiority here is measured in a number of different ways: examining the slope parameter and variance of the discrepancy in the regression model, and comparing the fraction of the variability in the forced component of the simulation that can be explained by the regression relationship.

Paleoreconstructions may be useful in principle for distinguishing among competing climate models. One caveat is that current available proxy-based reconstruction may not be informative enough to discriminate among different models. For instance, [89] find that comparisons of the spatial average of a 600 year High Northern Latitude temperature reconstruction to suites of last millennium climate simulations from the GISS 2E and CSIRO models suggest that the proxy-based reconstructions are able to discriminate only between the crudest features of the simulations within each ensemble.
33.3.2 Statistical issues

33.3.2.1 Considering and embracing the lack of independence

A subtle but potentially problematic issue is that paleoclimate reconstructions are often used to calibrate climate models and therefore the paleoclimate reconstructions used to validate the climate model are not, strictly speaking, an independent source of information. Because the calibration process used is typically informal or unknown, it is difficult to envision a rigorous way of adjusting for the fact that we may be using the paleoclimate reconstructions twice. Furthermore, as mentioned earlier some climate reconstructions have also employed the forcings in GCMs in order to improve the quality of the reconstruction [5, 54]. Although the forcings are used in very different fashion in the reconstruction procedure and the GCM formulation, such reconstructions should not be used to validate the forced response in models due to the circularity issue.

Efforts to mathematically or statistically formalize links between climate models and paleoclimate reconstructions remain relatively sparse. In HBMs involving relationships between these latent variables, there is the possibility of identifiability and confounding issues affecting the interpretability of the parameters. This issue becomes even more pronounced when data are spatially or temporally dependent (cf. [41]).

HBMs can reflect that paleoclimate proxies, instrumental data, and climate models contain different (even if overlapping) types of information about the climate. For example, proxies are imperfect recorders of local-scale climate variability, whereas climate models generally represent climate averaged over spatial areas defined by the model grid. Furthermore, the climate models are tuned to the instrumental record. Any statistically rigorous joint analysis of reconstructed and modeled climate must account for the potential spatial and temporal misalignment between the two sources of information.

33.3.2.2 Refining model components

Standard statistical modeling must be considered to ensure that the assumptions underpinning the statistical model relating paleoclimate proxies, instrumental data, and climate models to climate are reasonable. For instance, linear regression relationships may not be adequate for capturing how the two time series are related. Furthermore, there may be important lags between the two time series that are currently unaccounted for in the models such as those presented in [89].

Normal error structures may not always be appropriate, for example when a volcanic eruption impacts temperatures. It is important to account for dependence in the model appropriately. Dependent error terms are also useful in adjusting for misspecification in the model. For example, if there are trends in the residuals because the linear regression was inadequate, the dependent errors can pick them up. This is a positive when it comes to model fitting, but can be a negative when it comes to interpreting the model parameters relating the different processes of interest.

As in paleoclimate reconstruction, the more information about the underlying climate process we can provide, the better our inferences about the climate models become. [89] use information from control runs from a climate model to learn about the time series properties of temperature. Further research into statistical models for the climate process, in the framework introduced by [89], is warranted.
33.3.3 Research directions

33.3.3.1 Paleoclimate-based climate model calibration

The calibration of a model, such as a climate model, involves inferring appropriate parameter settings of the model based on observations or reconstructed observations of the physical process being modeled. Rigorous statistical calibration summarizes the information about parameters in the form of probability distributions, which in turn may be used to generate future projections from the model, while also incorporating parametric uncertainty. Going far back in time allows for the study of the impact of climatic events or periods in the Earth’s climate history where conditions were very different from conditions in the past few centuries. For example, [80] estimate climate sensitivity, an important parameter in climate models, by calibrating their model using paleoclimate reconstructions from the Last Glacial Maximum.

Given the relatively short period of time when instrumental records are available, paleoclimate reconstructions may be very valuable in that they allow researchers to examine the behavior of the models over much longer time periods than otherwise possible. For instance, [13] show how using paleoclimate reconstructions that inform the very long term behavior of the West Antarctic ice sheet result in sharper projections about the ice sheet’s future state. In particular, unrealistic simulations with overshoots in past ice retreat and projected future regrowth are eliminated. When calibration is done entirely based on modern ice sheet observations, the uncertainties related to future ice sheet projections are larger allowing for a small but non-negligible probability of no contributions (or negative contributions) to sea level rise from the ice sheet. By incorporating paleoclimate data, this possibility is virtually eliminated.

33.3.3.2 Making climate projections using paleoclimate reconstructions

Suppose we are interested in making projections based on an ensemble of models, and the goal is to form a multi-model mean of future projections, with weights based on their past agreement with the proxies. A probabilistic approach to averaging across multiple models is Bayesian model averaging (BMA) [42], which essentially provides weights based on the probability of each model given the observations. For example, [7] use historic space-time resolved temperature data to derive model weights for an ensemble of GCMs, while also incorporating the space-time dependence in the observations. This approach derives model weights, which allows for the assessment of the relative skill of the models in terms of how well they match observations. Projections of future temperatures may therefore be obtained by taking a weighted average of the GCM projections. The projections therefore incorporate historical data, and also characterize uncertainties due to space-time dependence and other sources of variation. Paleoclimate reconstructions may have much to contribute to work with multi-model ensembles as well, by incorporating paleoclimate information from far greater time ranges than is possible from instrumental records; cf [9, 34, 79] and the references therein. Instrumental records are available for mostly about 170 years, whereas paleoclimate records can span several centuries. Thus paleoclimate records have the potential to assess broad qualitative aspects of climate models, if not even more nuanced local phenomenon.

33.3.3.3 Paleoclimate reconstructions using climate models

When making weather forecasts, it is common to combine information from numerical weather forecasts based on physical models with observational data. The general idea of combining these two sources of information is often referred to as data assimilation (See Chapter 35 for a general discussion of this topic). The resulting forecasts can then, in principle, include both the knowledge of the physical system as well as empirical information
based on recent observations. There is a long history of methods for combining information from dynamic models and observations, with the Kalman filter [48] and many extensions, notably the ensemble Kalman filter [25, 26], being among the most well known. Roughly speaking, these approaches involve sequentially updating forecasts, conditional on the physical model and the observations. As the error structures become more complicated and the modeling approach becomes more flexible, more sophisticated methods like the ensemble Kalman filter and variants are useful, as well as Bayesian approaches to combining information [28]. More recently, some of these ideas have entered paleoclimate reconstructions (cf. [85]), where information from the sparse and noisy paleoclimate reconstructions is augmented with more spatially and temporally resolved climate model output, which in turn can provide more local (in space and time) reconstructions of past climate.

33.4 Discussion: looking forward

In this chapter we have discussed the important idea that in paleoclimate studies and climate studies more generally we can use statistical models to relate instrumental records, data from paleoclimate proxies, and climate model output to climate to learn about climate processes of interest. Using statistical models we are able to capture both the certainty and uncertainty inherent in relating each data source to climate. Through careful model assessment, we can understand how each data source contributes to our knowledge of climate.

Given the incomplete information about climate that are embedded in proxy records, it is worth asking what we cannot learn about climate using the statistical methods that we have discussed in this chapter. Part of the statistical difficulty in learning about the latent climate process is related to the sparseness of the measurements, but also in the fact that the measurement support for each data source can be different; to our knowledge there has been little research into methods of change-of-support applied to paleoclimate – see [91], Section 33.5.2 for a further discussion of the problem and references.

Throughout, we have reinforced the notion that successful studies tend to involve “good” information about the various processes of interest and their interaction. This requires a solid understanding of not only climate science but also the science embodied by paleoclimate proxies; for example, the study of how the interaction of trees with their environment relate to climate [96, 97, 98]. When there is a lack of knowledge about a relationship, we can build statistical models that can carefully mimic features inherent in the data [11, 20, 35, 54]. The inclusion of important covariates and climate forcings in particular are key to building reliable models for climate.

Paleoclimate reconstructions are central to climate science just as understanding our history is crucial for studying our present and future. Statistical paleoclimatology has allowed researchers to look back to learn about past climate. There are exciting opportunities to develop statistical methodologies to look back and also look forward, combining information from instruments, proxies, and climate model output to learn about the future of our climate. Given the importance of incorporating both scientific and statistical insights, the future of this research involves sustained interdisciplinary collaborations between climate scientists and statisticians.
Acknowledgments

PFC is supported in part by the US National Science Foundation (NSF) under grants NSF-DMS-1407604 and NSF-SES-1424481, and the National Cancer Institute of the National Institutes of Health under Award Number R21CA212308. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. MH was partially supported by NSF-DMS-1418090 and Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507. BL was partially supported by NSF-DPP-1418339 and NSF-AGS-1602845. BR was supported in part by the US NSF under grants DMS-CMG-1025465, AGS-1003823, DMS-1106642, and DMS-CAREER-1352656, and by the US Air Force Office of Scientific Research grant award FA9550-13-1-0043.

Bibliography


Paleoclimate reconstruction: looking backwards to look forward


Paleoclimate reconstruction: looking backwards to look forward


Paleoclimate reconstruction: looking backwards to look forward


