PROBABILISTICALLY CONSTRAINING PROXY AGE–DEPTH MODELS WITHIN A BAYESIAN HIERARCHICAL RECONSTRUCTION MODEL

Johannes P. Werner
Department for Earth Science and
Bjerknes Centre for Climate Research
University of Bergen, P.O. Box 7803
5020 Bergen, Norway
johannes.werner@uib.no

Martin P. Tingley
Department of Meteorology and Department of Statistics
Pennsylvania State University
University Park, PA 16802
USA

Abstract

Reconstructions of the late-Holocene climate rely heavily upon proxies that are assumed to be accurately dated by layer counting, such as measurements of tree rings, ice cores, and varved lake sediments. Considerable advances could be achieved if time-uncertain proxies were able to be included within these multi-proxy reconstructions, and if time uncertainties were recognised and correctly modelled for proxies commonly treated as free of age model errors.

Current approaches for accounting for time uncertainty are generally limited to repeating the reconstruction using each one of an ensemble of age models, thereby inflating the final estimated uncertainty – in effect, each possible age model is given equal weighting. Here, we demonstrate how to exploit the inferred space-time covariance structure of the climate to re-weight the possible age-depth models. Although they are a priori all deemed to be equally correct, the probabilities associated with the age models are formally updated within the Bayesian framework, thereby reducing uncertainties. Numerical experiments show that updating the age model probabilities decreases uncertainty in the resulting reconstructions, as compared with the current de facto standard of sampling over all age models, provided there is sufficient information from other data sources. This approach can readily be generalised to non-layer-counted proxies, such as those derived from marine sediments.

1 Introduction

Current large-scale climate reconstructions over the Common Era rely on accurately and precisely dated climatic proxies with annual resolution. These mostly comprise of layer counted data, such as tree rings, varved sediments, and annually layered ice cores [e.g., Jones et al., 2009]. As there are a number of records that do not fulfil these strict criteria [see e.g. PAGES2k Consortium, 2013], considerable progress could be made by including time-*uncertain* proxy records, using methodologies that allowed for accurate propagation of uncertainties. This would allow for radiometrically dated proxies, such as many marine or lacustrine sediment archives and speleothems, to be included in high-resolution reconstructions. As these records are believed to better preserve the low-frequency climate variability than annually resolved proxies [e.g., Jones et al., 2009], they could aid in characterising multi-centennial to millennial scale climate variability over the late Holocene. Including these records also potentially widens the spatial and temporal scope of feasible climate field reconstructions.

Often, age–depth relations in sediment cores are narrowed down by maximising some similarity metric, such as a correlation, between time-uncertain records [e.g., Lisiecki and Raymo, 2007] or between a time-uncertain record and a known, often orbital, signal [e.g., Shackleton et al., 1990]. Most of such studies average numerous time-uncertain proxy series to explore temporal variability [Huybers and Wunsch, 2004], or spatial variability at distinct time slices [e.g., Curry and Oppo, 2005]. Our aim is, however, to use a combination of time-certain and time-uncertain proxies to reconstruct spatio-temporal climate variability over the late Holocene.

In a recent publication we have shown how to formalise the use of time-uncertain proxies within a Bayesian hierarchical model for climate field reconstruction [Werner and Tingley, 2015]. Our approach is a data-derived compromise between selecting a single, optimal age model by some metric, or treating all age models as equally likely and iterating over them [Tierney et al., 2013; Comboul et al., 2014]. The resulting statistical model is hierarchical, and model fitting exploits conditional dependencies by sequentially updating estimates of the climate conditional on the currently selected age model, and then updating the probabilities associated with members of an ensemble of age–depth models based on the current estimate of the climate.

The presented method focuses on banded climate archives, such as tree rings, annual layers in ice cores, varves in lake sediments or corals, that feature dating errors caused by skipping or over-counting layers. Currently, van der Bilt et al. [2015] apply it on high resolution sediment archives to validate the existence of a regionally representative signal in periglacial lake sediments from South Georgia Island.

Although we will work with the established method BARCAST [Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time Tingley and Huybers, 2010] as used by Tingley and Huybers [2013] and Werner et al. [2013], our results are readily transferable to any other Bayesian hierarchical model for inferring past climate.

We first provide background information on BARCAST in Section 2, before we describes the technical modifications required to update the probabilities of the ADMs. Results of numerical experiments characterising and illustrating the advantages of our approach are presented in Sect. 3. Our core idea of updating the probabilities associated with time-uncertain proxy data by including an ADM level within a Bayesian hierarchical model is general, and discussions provided in Sect. 4 touch upon on how the core ideas can be extended, including e. g. radiometrically derived ADMs.

2 Bayesian hierarchical models for climate field reconstructions

Bayesian hierarchical modelling is a natural framework for inferring past climate from proxy observations [Tingley et al., 2012]. It involves disentangling assumptions made about the climate system from assumptions made about the distribution of observations consistent with a given climate state. Hierarchal modelling allows sophisticated models to be developed via the specification of a series of simpler, interlinked conditional probability statements [Wikle et al., 2001].

In the paleoclimate context, such models consist of a process level, a simple, parametric model describing the stochastic variability of the target climate process, and a data level, a description of the observations conditional on the climate. Thus, forward models for proxy data [Evans et al., 2013] are the natural way to formulate the latter step of the hierarchy. Finally, more or less informative prior distributions, encoding pre-analysis beliefs, must be specified for all unknown parameters.

BARCAST [Tingley and Huybers, 2010] models the target climate process $\vec{C}_t$ as a first-order autoregressive [AR(1)] process in time, with mean μ and persistence α, with multivariate normal innovations featuring exponentially decaying spatial covariance. Although based on a relatively simple process-level model, numerous studies have shown that BARCAST works well in practice for reconstructing temperature variations [e.g., Werner et al., 2013, 2014; Tingley and Huybers, 2013]. The processes level, the evolution of (the latent, i. e. never observed without error) spatial climate anomalies $\vec{C}_t$ in time, takes the form

$$\begin{aligned} &\vec{C}_{t+1} - \mu = \alpha \left(\vec{C}_t - \mu \right) + \vec{\epsilon}_t \\ &\vec{\epsilon}_t \sim \mathcal{N}(\vec{0}, \mathbf{\Sigma}) \quad \text{(independent)} \\ &\mathbf{\Sigma}_{i,j} = \sigma^2 \exp\left(-\phi |x_i - x_j|\right). \end{aligned} \tag{1a}$$

The innovations capture spatial persistence in the form of an exponential decreasing correlation as a function of separation between locations x_i and x_j, with e-folding distance $1/\phi$. The resulting shared information in space and time is critical in constraining age models for time-uncertain proxies.

At the data level, BARCAST specifies a separate linear forward model for each type of observation:

$$\begin{aligned} &\vec{O}_t = \mathbf{H}_t(\beta_0 + \beta_1 \cdot \vec{C}_t + \vec{e}_t) \\ &\vec{e}_t \sim \mathcal{N}(\vec{0}, \tau^2 \cdot \mathbf{I}) \quad \text{(independent)}. \end{aligned} \tag{1b}$$

The parameters $(\beta_0, \beta_1, \tau^2)$ are assumed to be different for each type of observation (e.g., tree ring widths, ice cores), but are sometimes taken to be common for all observations of a given type. Furthermore, the instrumental observations are assumed to be unbiased and on the correct scale ($\beta_0 = 0$ and $\beta_1 = 1$). The selection matrix $\mathbf{H}_t$ is composed of zeros and ones, and selects out at time step t the locations for which there are proxy observations of a given type. Inference on the parameters and the latent climate process proceeds via Markov chain Monte Carlo [MCMC; e.g., Gelman et al., 2003]. While we refer the reader to Tingley and Huybers [2010] for the technical details, we note that a core principle of MCMC is to estimate the joint probability distribution of all unknowns by iteratively sampling from each unknown conditional on the current values of all other unknowns. For example, we draw from the distribution of the climate process, conditional on the parameters, and then update the parameters conditional on the climate. One of the main shortcomings of BARCAST, shared by most other climate field reconstruction methods [Schneider, 2001; Smerdon et al., 2011; Luterbacher et al., 2004; Guillot et al., 2014], is the inability to incorporate data with dating uncertainty in a statistically rigorous manner. Since hierarchical models such as this are specified through a series of simple models (Eq. 1a and b), they are naturally modular and amenable to modification. In particular, (1b) can be viewed as *conditional on* the correct age model, and can be generalised to include updating of the probabilities associated with an ensemble of age models, *conditional on* the climate.

2.1 Including time-uncertain data in BARCAST

We now augment the basic framework to permit inclusion of annual-resolution, layer-counted proxies that feature ADM errors. Layers may be miscounted when the annual bands are weak, [e.g., corals Comboul et al., 2014], or when hiatuses in the record are misdated [e.g., speleothem data; Osete et al., 2012]. An ensemble of possible ADM errors is shown in Fig. 1: after a perfectly dated top section counting errors accumulate and lead to a substantial spread in possible dates for the lowest layers. For the purposes of exposition, and to simplify notation, we consider the case of a single time-uncertain, layer-counted proxy; the formalism can then be repeated for each time-uncertain proxy.

Associated with each time-uncertain proxy record is an ensemble of possible ADMs, $\{\mathcal{T}_k, k = 1, \ldots, M\}$, all of which are equally likely. This ensemble is generated based on understanding of the proxy archive and should reflect an honest assessment of possible uncertainties. We here take this ensemble as given, and seek to update the conditional posterior probabilities associated with the ADMs *conditional on* the current draw of the climate and parameters in the MCMC algorithm

We rewrite the data-level model of BARCAST for the time-uncertain proxy: Whereas Eq. (1b) relates all observations throughout space at time t to the concurrent climate field, it is more convenient in this case to relate $\vec{O}_s$, the time series of the time-uncertain proxy observations at location s, to $\vec{C}_s$, the co-located time series of the estimated climate process. The dependence of the proxy observations conditional on the climate time series and a particular ADM $\mathcal{T}$ then takes the form

$$\begin{aligned} \vec{O}_s|\mathcal{T}, \vec{C}_s &= \beta_0 + \beta_1 \cdot \mathbf{\Lambda}_s^{\mathcal{T}} \cdot \vec{C}_s + \vec{e}_s \\ \vec{e}_s &\sim \mathcal{N}(\vec{0}, \tau^2 \cdot \mathbf{I}) \quad \text{(independent)}, \end{aligned} \tag{2}$$

with the local observation error time series $\vec{e}_s$. Analogous to $\mathbf{H}_t$ in Eq. (1b), the ADM-dependent selection matrix $\mathbf{\Lambda}_s^{\mathcal{T}}$ picks out the elements of the vector $\vec{C}_s$ which correspond to elements of $\vec{O}_s$.

The only dependence in either process or data level model on the ADM enters in eq. (2) through the matrix $\mathbf{\Lambda}_s^{\mathcal{T}}$. The resulting conditional likelihood of the time series of proxy observations at location s conditional on climate and selected ADM is multivariate normal, with a diagonal covariance matrix

$$\mathcal{L}\left(\vec{O}_s|\mathcal{T}, \vec{C}_s\right) \sim \mathcal{N}\left(\beta_0 + \beta_1 \cdot \mathbf{\Lambda}_s^{\mathcal{T}}, \tau^2 \cdot \mathbf{I}\right) \tag{3}$$

and, assuming equal prior probabilities for the ADMs, $\pi(\mathcal{T} = \mathcal{T}_k) = 1/M$, the conditional posterior probabilities for the candidate ADMs $\{\mathcal{T}_k\}$ are

$$p\left(\mathcal{T} = \mathcal{T}_k|\vec{C}_s, \vec{O}_s\right) \propto \mathcal{L}\left(\vec{O}_s|\mathcal{T}_k, \vec{C}_s\right) \cdot \pi(\mathcal{T}_k). \tag{4}$$

This is again a normal distribution, and the problem is in theory solved. We can now sample the climate process and scalar parameters as described by Tingley and Huybers [2010], and at each step of the MCMC additionally select an ADM according to the conditional posterior probabilities in Eq. (4). In practice the sampler is slow to explore the full probability space of the ADMs. Intuitively, the strong interdependence between the currently selected ADM and the current draw of the climate field favours retaining the current ADM, and the algorithm wanders around the local optimum.

This problem can be overcome using parallel tempering and Metropolis-Coupled MCMC [$(\text{MC})^3$, Altekar et al., 2004]: Several chains of the MCMC sampler are run in parallel, each at a different "temperature" and subsequently coupled, their states being swapped in a Metropolis step. The additional "heating" allows MCMC to more easily escape local optima. We only need to modify the posterior probabilities of the ADMs as described by Altekar et al. [2004]:

$$p_\theta\left(\mathcal{T} = \mathcal{T}_k|\vec{C}_s, \vec{O}_s\right) \propto \mathcal{L}\left(\vec{O}_s|\mathcal{T}_k, \vec{C}_s\right)^\theta \cdot \pi(\mathcal{T}_k), \tag{5}$$

where the parameter $\theta \in [0, 1]$ also called "inverse temperature". For $\theta = 1$ (no heating) the normal posterior is recovered, while $\theta = 0$ results in a unity likelihood and the posterior equals the (here: flat) prior.

After a predetermined number of iterations, the states of two chains of different temperatures can be swapped in a Metropolis step. The probability of swapping the states of two chains j and k, with heatings θ_j and θ_k conditional on the state of the chain, that is, selected ADM and estimated climate, [Altekar et al., 2004]

$$\begin{aligned} &p(j \leftrightarrow k|\mathcal{T}_{j,k}, \vec{C}_{j,k}, \vec{O}) = \\ &\min\left(1, \prod_s \frac{p_{\theta_k}\left(\mathcal{T}_j|\vec{C}_j, \vec{O}\right) \cdot p_{\theta_j}\left(\mathcal{T}_k|\vec{C}_k, \vec{O}\right)}{p_{\theta_j}\left(\mathcal{T}_j|\vec{C}_j, \vec{O}\right) \cdot p_{\theta_k}\left(\mathcal{T}_k|\vec{C}_k, \vec{O}\right)}\right). \end{aligned} \tag{6}$$

The product is over all locations with time-uncertain proxies, and note that $\mathcal{T}_{j,k}, \vec{C}_{j,k}, \vec{O}$ all depend on the spatial location s. This procedure allows more diverse ADMs to be selected by the heated chains, while the unheated chain retains the correct stationary distribution, and is thus evaluated in the end. We select the the heating parameters θ_i of the chains as a geometric series between θ_{min} and $\theta = 1$, with a smallest heating parameter such that in an uncoupled experiment rapid exploration of the ADM space can be seen.

3 Simulation experiments

We use a set of numerical experiments to show how the algorithm uses the information shared between proxies across space and time. The data are constructed using the BARCAST process and data-level models (Eqs. 1a and 1b; parameter values Table 1). We consider a small spatial domain consisting only of 30 grid points with a distance of about 500 km. Experiments are run over 1300 time steps (years), with high-precision "instrumental" data available over the most recent 150 years. The two proxy time series consists of

Table 1. Model parameters for the experiments. Proxy 1 is time-certain and Proxy 2 features an uncertain age–depth model. The signal-to-noise ratio is set to about 0.25 [cf. Werner et al., 2013]. The measurement noise for the instrumental data is given by τ_I.

Process Level		Data Level	
α	0.6	τ_I	0.05
μ	0.0	$\tau^2_{P,1}$	0.4
σ^2	0.9	β_0	0
ϕ^{-1}	1000 km	β_1	0.5

annually resolved data, one correctly dated, the other with a 2 % probability [Comboul et al., 2014] of miscounted layers. An ensemble of 1000 such ADMs is shown by the grey lines in Fig. 1.

The experiments are run with differently located time-certain and time-uncertain proxies. The study is designed to test the effect of spatial distance between the two records on the quality of the reconstruction. As already pointed out in an earlier study [Werner and Tingley, 2015], our new BARCAST+AMS method falls back to random sampling should the information shared between the two localities be too low.

To evaluate the results over the earliest 500 time steps to stress the effect of the time uncertainty (shading in Fig. 1) we use the average cross correlation and root mean square error (RMSE) between the target and the reconstruction ensemble. In addition, we calculate the continuous ranked probability score [CRPS; Gneiting and Raftery, 2007], as it are more suitable evaluations of ensemble estimates. CRPS is a combined measure of the sharpness (uncertainty) of the ensemble reconstructions – the potential average CRPS ($\overline{\mathrm{CRPS}}_{\mathrm{pot}}$) – and an estimate how well the nominal coverage rates of the ensemble correspond to the empirical ones – termed the average reliability score $\overline{\mathrm{Reli}}$. They are in the units of the evaluated variables. In contrast to the often employed coefficient of efficiency and the reduction of error [Cook et al., 1994] they offer no simple threshold indicating a “skillful” reconstruction. However, the coefficient of efficiency and the reduction of error are no proper scoring rules [Gneiting and Raftery, 2007], and thus not suitable for evaluating ensemble predictions – even if they are convenient and widely used.

Results of the experiments are summarised in Figure 2, using both BARCAST+AMS (implemented using MC^3) and randomly selecting from the ADMs within BARCAST. All measures for both analyses indicate a better reconstruction at the location of the time-certain proxy, where results are comparable between the two analysis choices (not shown). The added value of our method can however be seen when comparing the results at the location of time-*uncertain* proxies. Of course, no difference can be seen when the two proxies are exactly co-located. However, as the distance between the proxies increases, the cross-correlation of the BARCAST+AMS experiments remains larger than that

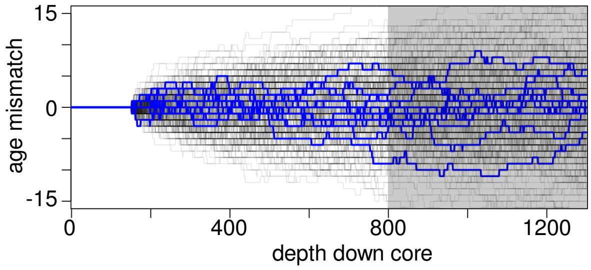

Figure 1. Trace plots of ADMs used. Shown is the mismatch of layer number (experimental date) vs. true date, for each of the ensemble of ADMs. A perfect ADM would be a straight horizontal line at zero. All ADMs (black), ADMs selected (blue) with a distance of 2/3 decorrelation length between the proxies. Shading denotes range over which skill is evaluated.

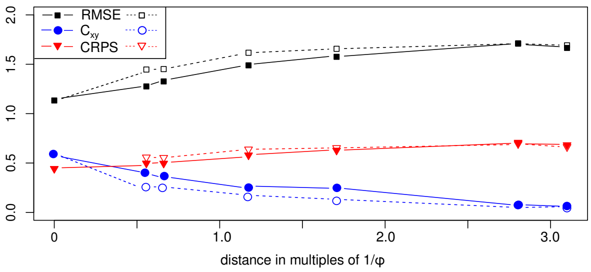

Figure 2. Experimental results for BARCAST+AMS (solid lines, closed symbols) and random ADM selection (dashed, open symbols), evaluated at the location of the time-uncertain proxy. Spatial decorrelation length scale is 1000 km. The reconstructions are evaluated over the earliest 500 time steps. The standard deviation of the local target climate signal is about 1.3.

of the random sampling experiments. The RMSE and the $\overline{\mathrm{CRPS}}_{\mathrm{pot}}$ remain lower, indicating both a lower standard error and a sharper reconstruction. As distance further increases, however, the results from both lines of experiments approach each other.

We also evaluate how the algorithm can in fact narrow down the ADMs from the prior ensemble. In the trace plot Fig. 1, the ADMs that are selected at a inter-proxy distance of about 3/4 decorrelation length (blue lines). Both the spread between the selected ADMs and the average age mismatch between the ADM and the correct ADM (horizontal line at 0 in Fig. 1) are greatly reduced. In Figure 3 we show the distribution of drawn ADMs (by their L1 distance from the true ADM) for three different experiments: a distance between the proxies of 0, 2/3 and 2 decorrelation length. As the distance between the time certain and the time uncertain proxy increases, the amount of shared information is reduced and the BARCAST+AMS algorithm can no longer narrow down the ADM selection.

These results clearly demonstrate the improvements afforded by formally updating the probabilities associated with the ADMs. We note that the ability to learn about the posterior distribution of the ADMs is a strong function of the amount of nearby (as measured by the spatial decorrelation length scale) information that is

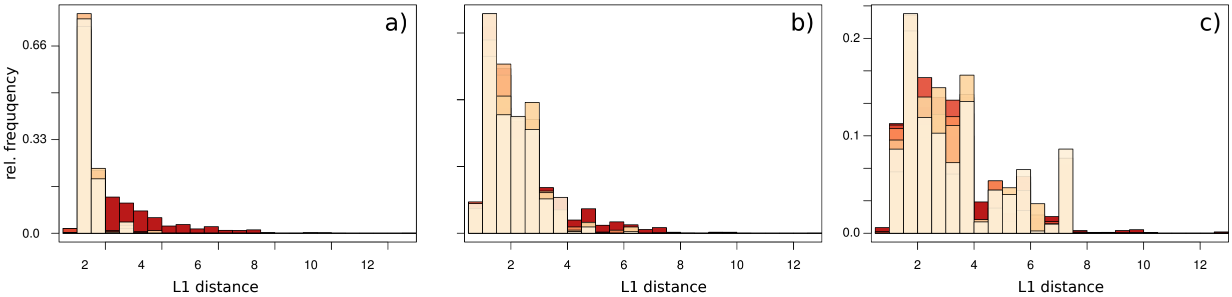

Figure 3. Draws of ADMs (by mismatch) for one perfectly dated and one time-uncertain proxy. Colours from beige (unheated) chain to red (warmest) chain. **(a–c)** The separation between the two proxies is 0, 750, 2084 km.

available to constrain the climate at the location of the time-uncertain proxy. Indeed, when the time-certain proxy is more than twice the spatial correlation length from the time-uncertain proxy there is little gain over random ADM selections. This illustrates the need for a shared signal between the proxies that can be used to correct for the misdating in either. The exact source of the shared signal is not important, be it through spatial closeness as constructed here, or through a long-range teleconnection.

4 Discussions and extensions

We have described and implemented an extension of a Bayesian hierarchical model for climate field reconstructions that accounts for uncertainties due to misspecified age–depth models in annually resolved proxy records. Although we have focused on a particular type of time-uncertainty (miscounting of annual layers) and the BARCAST reconstruction algorithm of Tingley and Huybers [2010], the methodology we outline is broadly applicable. To achieve adequate mixing in the MCMC, we make use of Metropolis-Coupled MCMC with parallel tempering [Altekar et al., 2004]. These techniques increase computational demands by a factor given roughly by the number of coupled chains, but are necessary to ensure adequate exploration of the probability space.

As demonstrated with simulation experiments, our method places higher posterior probability on ADMs with low L1 distance to the correct ADM (Figs. 1 and 3). Moreover, reconstructions that update the probabilities associated with the ADMs based on the current draw of the climate feature better score and skill metrics than reconstructions that feature random ADM selection (Fig. 2).

A number of useful extensions to the general framework we have proposed are possible. Many proxy archives, such as sediment cores, do not form annual layers. Proxies derived from these archives are generally measured as an average over a depth increment of sediment. The time boundaries of each increment are, in turn, determined by an ADM that is generally constrained using radiometric dating. If each proxy observation now represents an average over some number of time points, then each row of $\mathbf{\Lambda}_s^{\mathcal{T}}$ features a segment composed of the corresponding averaging weights, determined by the ADM, instead of ones and zeroes.

We have assumed the a priori existence of an ensemble of possible ADMs for each time-uncertain proxy. In many cases, such an ensemble of ADMs is not available with the proxy record – though recent efforts to define standards for proxy metadata suggest that the original dating information, such as radiometric ages and uncertainties, be included along with the proxy observations [see PAGES2k Consortium, 2014]. Recent work in radiocarbon dating has focused on stochastic modelling of the sedimentation process, and development of Bayesian models to constrain possible depositional histories, and therefore ADMs, conditional on a set of imperfect age control points [e.g., Ramsey, 2008; Blaauw and Christen, 2011]. Given dating information, any of the existing ADM construction algorithms can be used to produce an ADM ensemble. Such an algorithm could also be embedded within the climate reconstruction method, where instead of simply drawing from possible ADMs, a new ADM could be generated, which could then be accepted or rejected in a Metropolis [Gelman et al., 2003] step. This would increase the number of candidate ADMs, and might speed up exploration of the ADM space.

There are many potential benefits to including time-uncertain proxy records within annually resolved, late-Holocene climate field reconstructions. Inclusion of time-uncertain observations would increase the number of proxy records and the diversity of proxy types available to such analyses, and increase the spatial coverage of the proxy network. Furthermore, lower-resolution, time-uncertain records may permit improved inference on low-frequency climate variability [e.g., Jones et al., 2009], and have the potential to extend the time span of reconstructions.

Acknowledgements

Discussions at the PAGES Advances in Climate Field Reconstructions workshop held at WHOI, 15–16 April 2014 gave the impetus for this work. J. P. Werner is supported by the Centre for Climate Dynamics at the Bjerknes Centre. M. P. Tingley is supported in part by NSF P2C2 grant 1304309.

References

Altekar, G., Dwarkadas, S., Huelsenbeck, J. P., and Ronquist, F. (2004). Parallel metropolis coupled markov chain monte carlo for bayesian phylogenetic inference. *Bioinformatics*, 20(3):407–415.

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

Comboul, M., Emile-Geay, J., Evans, M. N., Mirnateghi, N., Cobb, K. M., and Thompson, D. M. (2014). A probabilistic model of chronological errors in layer-counted climate proxies: applications to annually banded coral archives. *Climate of the Past*, 10(2):825–841.

Cook, E. R., Briffa, K. R., and Jones, P. D. (1994). Spatial regression methods in dendroclimatology – a review and comparison of 2 techniques. *Int. J. Climatol.*, 14:379–402.

Curry, W. and Oppo, D. (2005). Glacial water mass geometry and the distribution of δ^{13}c of σCO_2 in the western atlantic ocean. *Paleoceanography*.

Evans, M. N., Tolwinski-Ward, S., Thompson, D., and Anchukaitis, K. J. (2013). Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary Science Reviews*, 76:16–28.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003). *Bayesian Data Analysis*. Chapman & Hall, 2 edition.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

Guillot, D., Rajaratnam, B., and Emile-Geay, J. (2014). Statistical paleoclimate reconstructions via markov random fields. *Annals of Applied Statisitcs*. Submitted.

Huybers, P. and Wunsch, C. (2004). A depth-derived pleistocene age model: Uncertainty estimates, sedimentation variability, and nonlinear climate change. *Paleoceanography*, 19(1):PA1028.

Jones, P. D., Briffa, K. R., Osborn, T. J., *et al.* (2009). High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *The Holocene*, 19(1):3.

Lisiecki, L. and Raymo, M. (2007). Plio-Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. *Quaternary Science Reviews*, 26:56–69.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H. (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, 303:1499–1503.

Osete, M.-L., Martín-Chivelet, J., Rossi, C., Edwards, R. L., Egli, R., Muñoz-García, M. B., Wang, X., Pavón-Carrasco, F. J., and Heller, F. (2012). The blake geomagnetic excursion recorded in a radiometrically dated speleothem. *Earth and Planetary Science Letters*, 353:173–181.

PAGES2k Consortium (2013). Continental-scale temperature variability over the common era. *Nature Geosciences*, 6:339346.

PAGES2k Consortium (2014). Pages 2k proxy database. Technical report, IGBP Pages 2k.

Ramsey, C. (2008). Deposition models for chronological records. *Quaternary Science Reviews*, 27:42–60.

Schneider, T. (2001). Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate*, 14:853–871.

Shackleton, N., Berger, A., and Peltier, W. (1990). An alternative astronomical calibration of the lower pleistocene timescale based on odp site 677. *Trans. R. Soc. Edinburgh Earth Sci*, 81:251–261.

Smerdon, J. E., Kaplan, A., Zorita, E., González-Rouco, J. F., and Evans, M. (2011). Spatial performance of four climate field reconstruction methods targeting the common era. *Geophysical Research Letters*, 38(11).

Tierney, J. E., Smerdon, J. E., Anchukaitis, K. J., and Seager, R. (2013). Multidecadal variability in east african hydroclimate controlled by the indian ocean. *Nature*, 493(7432):389–392.

Tingley, M. P., Craigmile, P. F., Haran, M., Li, B., Mannshardt-Shamseldin, E., and Rajaratnam, B. (2012). Piecing together the past: Statistical insights into paleoclimatic reconstructions. *Quaternary Science Reviews*, 35:1–22.

Tingley, M. P. and Huybers, P. (2010). A Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time. Part I: Development and Applications to Paleoclimate Reconstruction Problems. *J. Climate*, 23(10):2759–2781.

Tingley, M. P. and Huybers, P. (2013). Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature*, 496:201–205.

van der Bilt, W. G. M., Bakke, J., and Werner, J. P. (2015). Constraining the controls on Southern Ocean glacier-climate over the past millennium. In Review.

Werner, J. P., Smerdon, J. E., and Luterbacher, J. (2013). A Pseudoproxy Evaluation of Bayesian Hierarchical Modelling and Canonical Correlation Analysis for Climate Field Reconstructions over Europe. *J. Climate*.

Werner, J. P. and Tingley, M. P. (2015). Technical Note: Probabilistically constraining proxy agedepth models within a Bayesian hierarchical reconstruction model. *Clim. Past*, 11:533–545.

Werner, J. P., Toreti, A., and Luterbacher, J. (2014). Stochastic models for climate reconstructions – how wrong is too wrong? *NOLTA Proc.*, 24:528–531.

Wikle, C. K., Milliff, R. F., Nychka, D., and Berliner, L. M. (2001). Spatiotemporal hierarchical Bayesian modeling tropical ocean surface winds. *Journal of the American Statistical Association*, 96(454):382–397.