Paleoclimate reconstruction using statistical nonlinear forward models

Peter F. Craigmile[1,3], Martin P. Tingley[2], Jiangyong Yin[3]

[1] School of Mathematics and Statistics, The University of Glasgow, United Kingdom
[2] Department of Meteorology and Department of Statistics, Pennsylvania State University, PA, USA.
[3] Department of Statistics, The Ohio State University, Columbus, OH, USA.
Corresponding author: Peter F. Craigmile, e-mail: peter.craigmile@glasgow.ac.uk

Abstract

There is a growing interest in reconstructing climate using hierarchical models that link paleoclimate proxies and instrumental records to an underlying latent climate process, as they allow for a full propagation of uncertainties. A critical aspect of constructing such hierarchical models is to parameterize the forward model that links the proxy with the climate process. Typically these forward models are either based on simplified assumptions, such as a linear relationship between the proxy and climate, or by building physical models (e.g., models for tree growth or heat propagation down boreholes) that may involve components that are hard to parameterise and complicate the model fitting. In this work, statistical forward models that capture necessary nonlinearities under a monotonicity constraint are embedded within a Bayesian spatio-temporal hierarchical model for paleoclimate reconstruction. This methodology is then applied to the reconstruction of temperatures based on tree ring density series, and results compared to reconstructions based on a linear forward model.

Keywords: Bayesian methods, Monotonicity constraints, Spatio-temporal models, Tree ring densities.

1 Introduction

The aim of paleoclimate reconstructions is to predict a past space-time climate field (and its uncertainty) on the basis of recent instrumental data, and measurements on paleoclimatic proxies such as tree rings, ice cores, and lake floor sediments. While there are a number of algorithms available for producing spatial reconstructions, there is a growing interest in the use of hierarchical, often Bayesian, models to relate the instrumental and proxy-based data to the underlying latent climate field. For recent reviews of paleoclimate reconstructions see, e.g., Jones et al. [2009], Hughes and Ammann [2009] and Tingley et al. [2012].

A key specification of a hierarchical paleoclimate reconstruction is the so-called *forward model*, that relates the climate variable of interest to the measured proxies. In many reconstructions, a linear relationship between the climate variable of interest and the proxy is assumed [Tingley et al., 2012]. A more sophisticated approach is to embed physical understanding of each proxy's response to variations in the climate into the functional form of the forward model. For example, variants of the Vagonov-Shashkin model describe tree ring widths as a non-linear function of soil moisture and temperature [Evans et al., 2006, Tolwinski-Ward et al., 2011], while the pre-observation mean-surface air temperature model (POM-SAT) relates borehole temperatures to surface temperatures via the heat equation [Harris and Chapman, 2001, Harris, 2007, Li et al., 2010, Brynjarsdóttir and Berliner, 2011]. Defining physically-motivated forward models requires an accurate mechanistic understanding of the proxy–climate relationship. In practice, model fitting can be computationally demanding for certain choices of forward model. Also, limited understanding of certain parts of the physical relationship can complicate the reconstruction.

A compromise it to allow for nonlinear statistical relationships between the proxy and climate variable, while respecting the fact that in most practical applications it is reasonable to assume that the proxy–climate relationship is monotonic. In what follows, we relate surface temperatures to maximum latewood tree ring densities, using a scientifically motivated, non-linear forward model. We assume that tree ring growth increases monotonically with increasing temperature between a certain range of temperatures, and asymptotes

at higher or lower values – reflecting the limited capacity for a tree to grow in all temperatures [c.f. the model of Tolwinski-Ward et al., 2011]. We demonstrate the sensitivity of reconstructions to the choice of linear or non-linear forward model within a fully hierarchical Bayesian spatio-temporal model, accounting for both the uncertainty in the relationship between the proxy and climate and other uncertainties such as measurement errors. We use a version of an established Bayesian hierarchical model for paleoclimate reconstruction called BARCAST [Tingley and Huybers, 2010a,b], and compare reconstructions resulting from the linear forward model used by BARCAST to those resulting from a non-linear but monotone forward model for the proxy–climate relationship. In contrast to Hanhijärvi et al. [2013], the hierarchical approach adopted here allows for a climate reconstruction in both space and time while making a similarly general assumptions about the proxy–climate relationship. We introduce the data in Section 2, specify the model in Section 3, provide our results in Section 4, and close with a discussion in Section 5.

2 Tree ring density and temperature data

We use the $5^o \times 5^o$ gridded version of the maximum late wood tree ring density data set of Briffa et al. [2002a,b], and restrict our analysis to the North American continent. A map of the locations of the instrumental records (Figure. 1a) shows that the records in the south and west of the spatial domain are generally longer. The instrumental data set used here, the Climate Research Unit's (CRU) gridded temperature product [Brohan et al., 2006], is available on the same grid as the tree ring observations and likewise features varying spatial and temporal coverage (Figure 1b). Both data set are reported as anomalies with respect to climatological means, and both have been used extensively in reconstructions of past climate [e.g., Briffa et al., 2002a,b, Rutherford et al., 2005, Mann et al., 2007, 2008, 2009]. Figure 2 displays a scatter plot of the tree ring anomalies versus the instrumental anomalies, while the solid line shows a nonlinear regression fit using the monotone penalized spline approach of Meyer [2012], with 4 knots. Ignoring any spatial and temporal dependence in the data, we see a distinctly nonlinear relationship between the temperature and tree ring density anomalies. This nonparametric fit does not make the scientifically motivated assumption, formalized in the hierarchical Bayesian modeling framework developed in the next section, that the tree ring response for sufficiently low temperature anomalies should be flat.

3 Nonlinear paleoclimatic reconstruction

Our goal is to predict an underlying latent spatio-temporal climate field given instrumental records and paleoclimate proxy observations. More formally, let $T \subset \mathbb{Z}$ denote the time domain and $D \subset \mathbb{R}^p$ the spatial domain. Let $Y = \{Y_t(\boldsymbol{s}) : \boldsymbol{s} \in D, t \in T\}$ denote the latent climate process at time $t \in T$ and spatial location $s \in D$; $Z = \{Z_t(\boldsymbol{s}) : \boldsymbol{s} \in D, t \in T\}$ the observed instrumental record; and $P = \{P_t(\boldsymbol{s}) : \boldsymbol{s} \in D, t \in T\}$ the observed proxy records. In our application, Y is the latent space-time temperature process over North America from the year 1400, Z is the gridded space-time temperature data product over North America from 1850-2011, and P is the gridded tree ring density data product.

We follow a Bayesian hierarchical modeling framework as formalized in Tingley et al. [2012]. We assume a standard measurement error model for the temperature data product: $Z_t(\boldsymbol{s}) = Y_t(\boldsymbol{s}) + \epsilon_t(\boldsymbol{s}), \;\; t \in T, \;\; \boldsymbol{s} \in D$, where $\{\epsilon_t(\boldsymbol{s}) : t \in T, \boldsymbol{s} \in D\}$ is an independent Gaussian process with mean 0 and variance σ^2. In relating climate to the proxies, we consider the conditional distribution $f(P_t(\boldsymbol{s})|Y_t(\boldsymbol{s})$ and make the simplifying assumption of conditional independence over space and time. In some applications a non-normal distribution for f is needed – for example in relating pollen counts to climate, f may be a binomial or more generally a multinomial distribution [e.g. Haslett et al., 2006, Wahl et al., 2010]. Following earlier work [e.g., Briffa et al., 2002a,b, Rutherford et al., 2005, Mann et al., 2007, 2008, 2009, Tingley and Huybers, 2010a, Tingley et al., 2012], we assume for the tree ring densities used here that f is a normal density: $P_t(\boldsymbol{s}) = h(Y_t(\boldsymbol{s})) +$

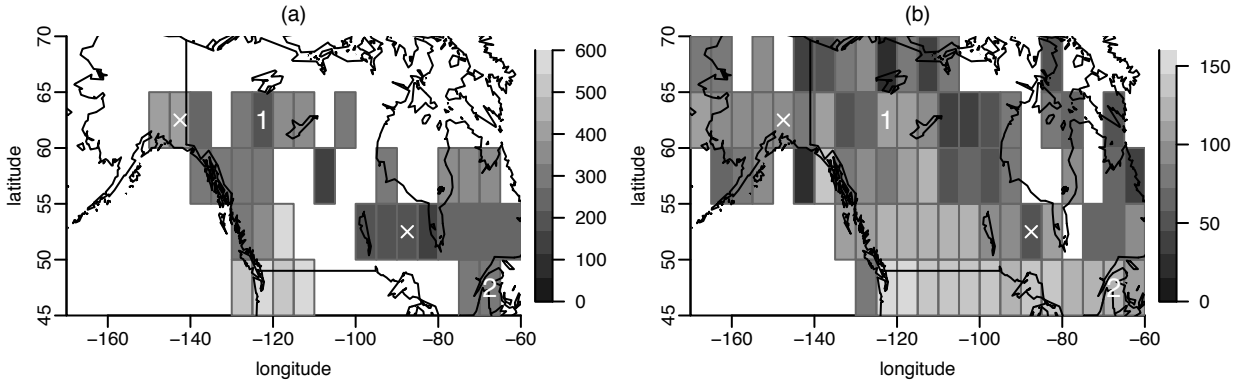

Figure 1: (a) Summary of the locations of the tree-ring density records. Shading indicates the number of years of observations at each location, the crosses indicate crossvalidation locations, and the numbered locations will be summarized later. (b) As with (a) but for the instrumental data.

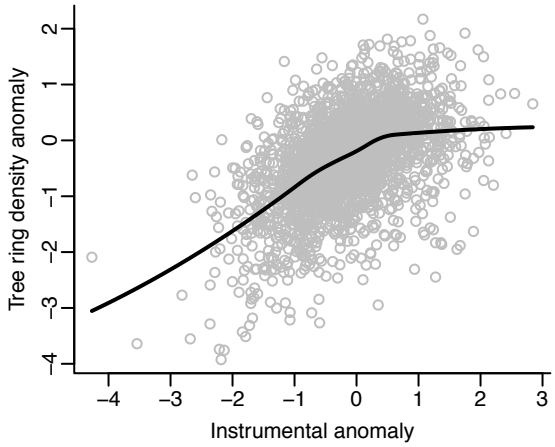

Figure 2: A scatterplot of the tree ring density anomalies versus the instrumental anomalies. The black solid line shows a monotone regression fit to these data.

$\eta_t(\boldsymbol{s}), t \in T,\ \boldsymbol{s} \in D$. The first term, $h(Y_t(\boldsymbol{s}))$, is the component of the proxy attributable to climate, and is the main focus of the discussion below, while the second term, $\{\eta_t(\boldsymbol{s})\}$, expresses the residual uncertainty in the proxy. In what follows, we make the simplifying assumption that $\{\eta_t(\boldsymbol{s}) : t \in T, \boldsymbol{s} \in D\}$ is an independent Gaussian process with mean 0 and variance τ^2, while noting that more a spatially or temporally correlated error process could be used instead [cf. Li et al., 2010].

We consider a parametric monotone forward model relating the temperatures and the tree ring density observations. More specifically, for a latent temperature y, we model $h(\cdot)$ as a sigmoid function: $h(y) = \beta_1 + \beta_2/(1 + e^{-\beta_3(y-\beta_4)})$. The sigmoid function reflects our prior assumptions that there are asymptotes at the minimum and maximum tree ring density, with the tree ring density increasing monotonically and nearly linearly with increasing temperature between these asymptotes.

To complete the statistical model, we follow Tingley and Huybers [2010a] in assuming that Y is a Gaussian spatio-temporal process with a constat mean, and a separable covariance function that is AR(1) in time and exponential in space. Letting $\boldsymbol{Y}_t$ denote the vector of temperatures at the sites of interest at time t, we assume that for each t, $\boldsymbol{Y}_t = \mu\mathbf{1} + \phi(\boldsymbol{Y}_{t-1} - \mu\mathbf{1}) + \boldsymbol{\epsilon}_t$, where $\mathbf{1}$ is a vector of ones, μ is the Y process mean, and $\boldsymbol{\epsilon}_t$ is an independent Gaussian error with mean zero and covariance $\kappa^2\boldsymbol{R}(\lambda)$, with $\boldsymbol{R}(\lambda)$ a correlation matrix with range parameter λ. For locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$, the (i, j) element of the matrix is $R_{ij}(\lambda) = \exp(-||\boldsymbol{s}_i - \boldsymbol{s}_j||)$,

where $||\boldsymbol{s}_i - \boldsymbol{s}_j||$ denotes the chordal distance between the two locations, thereby ensuring a valid correlation function on the sphere.

With $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_4)^T$ denoting the parameters of the forward model, the parameter vector of interest is $\boldsymbol{\theta} = \{Y, \sigma^2, \boldsymbol{\beta}, \tau^2, \kappa^2, \phi, \lambda\}$. The posterior distribution of $\boldsymbol{\theta}$ given $\boldsymbol{z}$ and $\boldsymbol{p}$ is then

$$\pi(\boldsymbol{\theta}|\boldsymbol{z}, \boldsymbol{p}) \propto \left[f(\boldsymbol{z}|Y, \sigma^2)\pi(\sigma^2)\right]\left[f(\boldsymbol{p}|Y, \boldsymbol{\beta}, \tau^2)\pi(\tau^2)\pi(\boldsymbol{\beta})\right]\left[f(Y|\mu, \kappa^2, \phi, \lambda)\pi(\mu)\pi(\kappa^2)\pi(\phi)\pi(\lambda)\right],$$

where $\boldsymbol{z}$ and $\boldsymbol{p}$ are the vector of instrumental and proxy observations. The density $f(\boldsymbol{z}|Y, \sigma^2)$ is the likelihood for the instrumental temperatures, $f(\boldsymbol{p}|Y, \boldsymbol{\beta}, \tau^2)$ is the likelihood for the proxy measurements, and $\pi(\cdot)$ denote prior distributions, assumed to be independent, on the parameters. With the exception of the spatial range parameter, we assume diffuse priors. The three variance parameters (σ^2, τ^2 and κ^2) are given mutually independent inverse gamma prior distribution with shape 0.01 and rate 0.01, while the β_j are given independent normal priors, with mean 0 and variance 100. The mean of the latent climate process, μ, is given a normal prior with mean zero and variance 100, while the temporal dependence parameter ϕ is given a uniform prior on $(-1, 1)$. Following the findings of Tingley and Huybers [2010a] and Mannshardt et al. [2013] we assume *a priori* that the spatial range parameter is concentrated around 2000, and specify a gamma prior prior with shape 500 and rate 1/4 for λ. There is no closed form expression for the posterior distribution, so we generalize the Markov chain Monte Carlo (MCMC) algorithm proposed in Tingley and Huybers [2010a] to allow for sampling with a nonlinear forward model.

4 Results

The results are based on two parallel chains of an MCMC algorithm [implemented in R, R Core Team, 2013], each of length 5,000 after thinning by a factor of 10 and discarding a a burn-in of length 2,000. Convergence was assessed using trace plots. We also fit a second variant of the model, assuming a linear forward model, $h(y) = \beta_1^L + \beta_2^L y$, and placing independent normal priors on β_1^L and β_2^L, each with mean zero and variance 100. All other aspects of the model remain unchanged, and posterior summaries for each model are based on the same number of posterior draws. Table 1 compares the posterior mean and 95% credible intervals (CIs) of the parameters under the two differing forward models. There are a number of notable differences between the two model fits, that extend beyond the parameters of the forward models themselves. The error variance in both the instrumental records, σ^2, and the proxy records, τ^2, is slightly higher for the nonlinear model. As a consequence, the variance for the latent climate process, κ^2, is lower for the nonlinear model. The mean of the latent climate process, μ, has a lower mean for the nonlinear forward model, while the temporal and spatial correlation parameters (ϕ, λ), and hence the correlations, themselves differ little between the two models.

Figure 3 shows posterior summaries of the forward models. Compared with the linear forward model, the nonlinear forward model predicts lower tree ring densities at lower temperatures, higher densities around zero, and lower densities for higher temperatures. To assess model performance, we calculated, using the posterior mean fits for each forward model, the root mean square error (RMSE) in predicting the tree ring density at two locations (the crosses in Figure 1) that were removed prior to model fitting. At both locations the fit was better for the nonlinear versus the linear forward model (RMSE of 0.81 versus 0.84 in the west; 0.76 versus 0.82 in the east).[1]

Figure 4 compares the reconstructed temperature anomalies at the locations denoted `1` and `2` on Figure 1, where both instrumental and paleoclimate measurements are available. Predictions from the nonlinear

[1]For a graphical summary see http://www.stat.osu.edu/~pfc/publications/documents/nonlinear_CV.pdf

Table 1: Posterior summaries of the parameters in the two Bayesian models with different forward models.

Nonlinear forward model			Linear forward model		
Parameter	Post. mean	95% post CI	Parameter	Post. mean	95% post CI
σ^2	0.07	(0.06, 0.07)	σ^2	0.03	(0.02, 0.03)
β_1	0.60	(0.49, 0.73)	β_1^L	-0.24	(-0.26, -0.21)
β_2	-4.62	(-5.17, -4.09)	β_2^L	0.68	(0.66, 0.71)
β_3	-1.12	(-1.26, -0.99)			
β_4	-1.66	(-1.83, -1.50)			
τ^2	0.42	(0.41, 0.44)	τ^2	0.39	(0.38, 0.41)
μ	-0.38	(-0.45, -0.31)	μ	-0.25	(-0.35, -0.15)
κ^2	0.49	(0.46, 0.51)	κ^2	0.77	(0.72, 0.83)
ϕ	0.49	(0.47, 0.51)	ϕ	0.51	(0.46, 0.54)
λ	2099.15	(1981.13, 2215.73)	λ	2103.37	(1975.62, 2237.98)

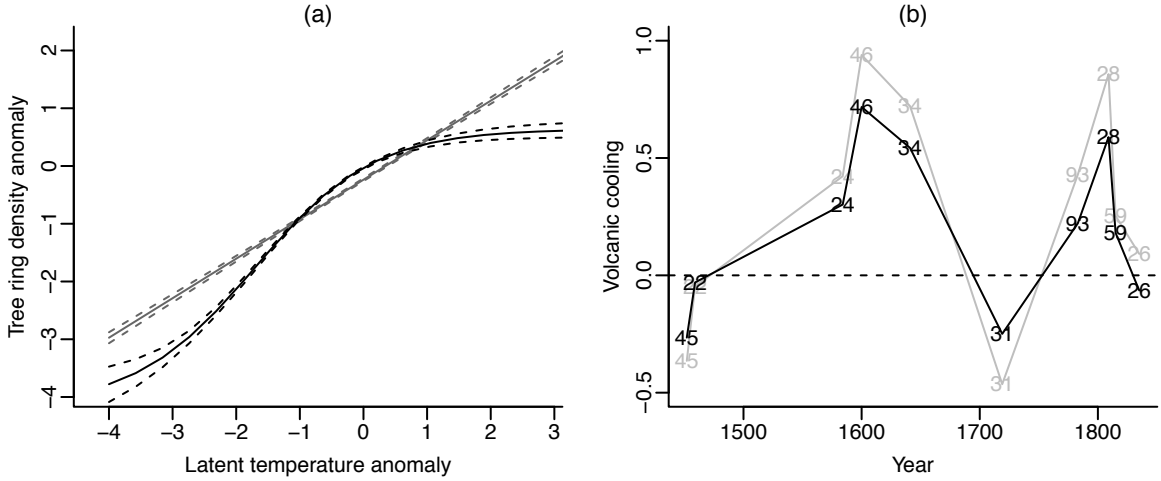

Figure 3: (a) A posterior summary of the two forward models (black: nonlinear; gray: linear). The solid line is the posterior mean and the dashed lines are pointwise 95% credible intervals; (b) A comparison of the rate of volcanic cooling for the two forward models (black: nonlinear; gray: linear). The numbers denote ice-core derived estimates of Northern Hemisphere sulphate aerosol injections, in Tg [Gao et al., 2009].

forward model are smoother in time, while the pointwise credible intervals are less variable. Both series show the effect of volcanic events, as evidenced by dips in the predictions (e.g., 1584 at Location 1, and 1816-1817 at both locations). Temporally-varying spatial summaries[2] indicate that the nonlinear forward model tends to produce cooler predictions of the latent temperatures in the past. As expected, the climate processes inferred from the two models are most similar over the last 150 years, when the spatially dense instrumental records dominate the inference. We next investigate the effect of the forward model upon the estimated cooling associated with the 10 largest pre-180 volcanic eruptions. We estimate the amount of volcanic cooling as the difference between the latent temperature for the four years preceding each volcano (averaged over space) minus the latent temperature average for year of the eruption and the subsequent year. Figure 3(b) indicates that the amount of volcanic cooling is less extreme under the nonlinear model (the black lines) versus the linear model (the gray lines). For Huaynaputina in 1600 and Tambora in 1815 the nonlinear model clearly damps the volcanic response. For Kuwae in 1453, the lack of data is likely a

[2] http://www.stat.osu.edu/~pfc/publications/documents/nonlinear_forward_maps.pdf

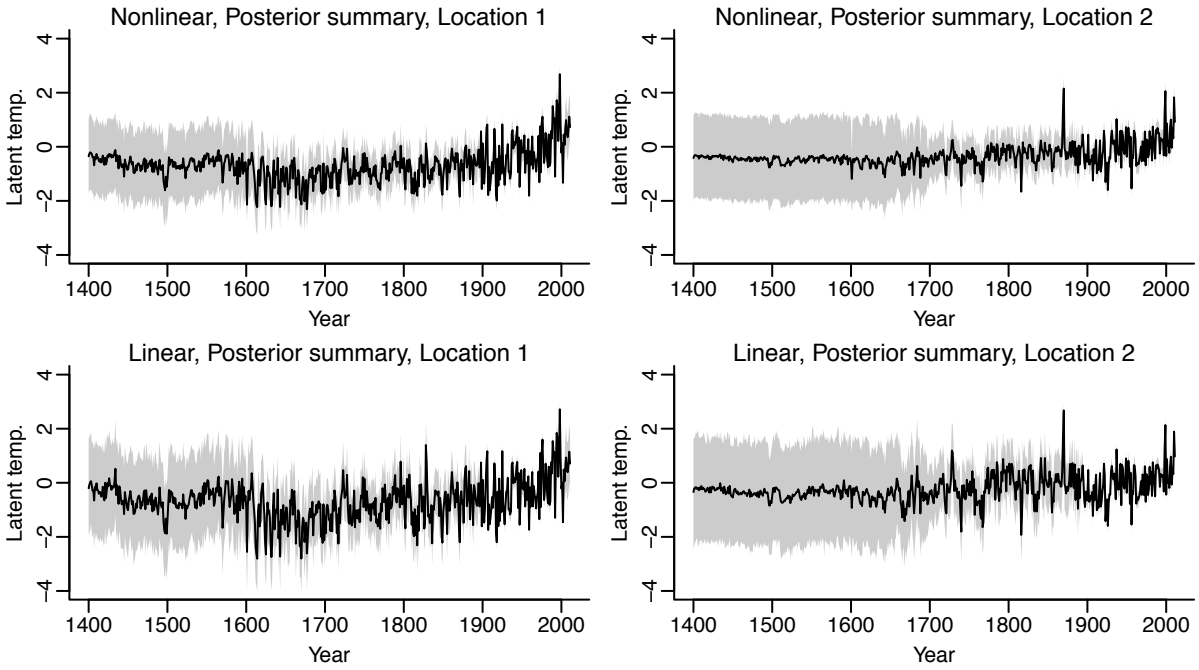

Figure 4: At two different locations containing both instrumental and tree ring density records, posterior summaries of the latent climate anomalies for the nonlinear (top panels) and linear (bottom panels) forward models. The solid line is the posterior mean and the gray region denotes pointwise 95% credible intervals.

reason for a lack of any cooling response. Over the limited spatial domain considered here, we do not find a strong correlation between the size of the eruption, according to the estimated sulphate aerosol injection [the number in Figure 3(b) Gao et al., 2009], and the amount of volcanic cooling.

5 Discussion

We have investigated the effects of nonlinear statistical forward models on paleoclimate reconstructions. Exchanging a linear forward model for a simple monotone function, within a hierarchical framework that assumes separability in the latent spatio-temporal process, results in reconstructions that differ both qualitatively and quantitatively. In future work we will further investigate the choice of statistical model and its impact upon the climate reconstruction. A natural generalization is to estimate the forward model nonparametrically using monotonic curve fitting techniques. Although traditional nonparametric function estimation is able to obtain the correct functional form asymptotically, finite sample bias is non-negligible in our application. To incorporate monotonicity constraints within the hierarchical framework, we can consider spline-based Bayesian isotonic regression, which can take a number of different forms [e.g., Neelon and Dunson, 2004, Cai and Dunson, 2007, Brezger and Steiner, 2008, Shively et al., 2009, Meyer et al., 2011]. Given a choice of basis functions, monotonicity and smoothness constraints can be easily implemented through appropriate prior distributions, and model fitting can be done using MCMC algorithms. A second generalization would let the forward model parameters vary spatially, allowing for an exploration of how the empirical tree ring – temperature response differs between regions where different climatic factors (namely, water, light, and temperature) are estimated to be most limiting to growth [Nemani et al., 2003].

References for this article can be found at `http://www.stat.osu.edu/~pfc/publications/documents/references.pdf`