Research Article Environmetrics

Received: 20 April 2015, Revised: 18 September 2015, Accepted: 21 September 2015, Published online in Wiley Online Library: 29 October 2015

(wileyonlinelibrary.com) DOI: 10.1002/env.2368

Reconstruction of late Holocene climate based on tree growth and mechanistic hierarchical models

John Tipton[a*], Mevin Hooten[a,b,c], Neil Pederson[d], Martin Tingley[e,f] and Daniel Bishop[g]

Reconstruction of pre-instrumental, late Holocene climate is important for understanding how climate has changed in the past and how climate might change in the future. Statistical prediction of paleoclimate from tree ring widths is challenging because tree ring widths are a one-dimensional summary of annual growth that represents a multi-dimensional set of climatic and biotic influences. We develop a Bayesian hierarchical framework using a nonlinear, biologically motivated tree ring growth model to jointly reconstruct temperature and precipitation in the Hudson Valley, New York. Using a common growth function to describe the response of a tree to climate, we allow for species-specific parameterizations of the growth response. To enable predictive backcasts, we model the climate variables with a vector autoregressive process on an annual timescale coupled with a multivariate conditional autoregressive process that accounts for temporal correlation and cross-correlation between temperature and precipitation on a monthly scale. Our multi-scale temporal model allows for flexibility in the climate response through time at different temporal scales and predicts reasonable climate scenarios given tree ring width data. Copyright © 2015 John Wiley & Sons, Ltd

Keywords: Bayesian hierarchical model; paleoclimate; predictive validation; physical statistical model

1. INTRODUCTION

Statistical estimation of past climate is important for understanding climate change in a historical context and for predicting how climate will respond in the future (Stocker *et al.*, 2013). Ideally, one would model climate with a long time series of spatially explicit, highly precise instrumental measurements. However, the instrumental record only spans the last 100–200 years, perhaps less in many areas. Paleoclimate reconstructions allow for investigation of climate dynamics at longer timescales than instrumental records and serve as a test bed to evaluate performance of complex modern climate models. In the absence of a dense network of instrumental observations, we must rely on climate proxy data to gain a better understanding of climate history. Evans *et al.* (2013) describe a conceptual model for how proxy processes integrate physical, chemical, biological, and geological climate information to yield the observed data. Their work calls for the development of mechanistic proxy system models to describe how climate influences the proxy observations. Among available proxy data sources, many late Holocene paleoclimate reconstructions focus on tree ring widths because tree ring width data are widely available on a regional or hemispheric scale; can contain hundreds or thousands of years of observations; are sensitive to temperature, precipitation, and drought; and have a very clear annual to seasonal resolution (Jones *et al.*, 1998; D'Arrigo *et al.*, 2004; Moberg *et al.*, 2005; Mann *et al.*, 2008; Christiansen and Ljungqvist, 2011; Griffin *et al.*, 2013). Our contribution is to improve current statistical models for reconstructing late Holocene climate from annually resolved tree ring proxy data and develop a framework for future model development.

The statistical reconstruction of paleoclimate histories from tree ring width data poses many challenges (Jones *et al.*, 2009). First, tree rings are formed through a broad range of climatic, ecological, and growth allocation factors that make them noisier than instrumental records. Dendrochronologists typically process the tree ring width data in an attempt to remove the non-climate factors influencing growth (Cook

* *Correspondence to: John Tipton, Department of Statistics, Colorado State University, 1484 Campus Delivery, Fort Collins, CO 80525, U.S.A. E-mail: jtipton25@gmail.com*

a *Department of Statistics, Colorado State University, Fort Collins, CO 80523, U.S.A.*

b *Colorado Cooperative Fish and Wildlife Research Unit, US Geological Survey, Fort Collins, CO 80523, U.S.A.*

c *Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523, U.S.A.*

d *Harvard Forest, Harvard University, 324 North Main Street, Petersham, MA 01366, U.S.A.*

e *Department of Meteorology, Pennsylvania State University, 510 Walker Building, University Park, PA 16802, U.S.A.*

f *Department of Statistics, Pennsylvania State University, 510 Walker Building, University Park, PA 16802, U.S.A.*

g *Tree Ring Lab, Lamont-Doherty Earth Observatory, 61 Route 9W, PO Box 1000, Palisades, NY 10964, U.S.A.*

and Kairiukstis, 1990). After removing most of these non-climatic effects, the dendrochronologist aggregates the many tree ring widths at a site into a tree ring chronology. The final filtered tree ring chronology consists of one time series derived from a number of tree cores of a particular species exhibiting a coherent signal (Cook and Kairiukstis, 1990). The tree ring standardization procedures have been thoroughly discussed in the literature, and as such, we treat the chronologies as observed data (Melvin and Briffa, 2008).

One challenge in the reconstruction of paleoclimate from tree ring width data is that the climate signal influencing tree growth occurs in continuous time, whereas we typically summarize climate and tree ring widths in discrete time. A tree growth increment represents the integrated response of the tree to climate conditions over a growing season(s), collapsing sub-annual climate information into a univariate value, annual tree ring width (Fritts, 1976; Bradley, 2011; Carbone *et al.*, 2013). The model of tree rings we develop considers climate on a monthly time step. However, there is only one tree ring observation per year, leading to a temporal change of support problem (Gotway and Young, 2002).

Another factor to consider is that climatic influences on tree ring growth are multivariate, typically involving temperature and precipitation. The joint estimation of temperature and precipitation from a univariate tree ring width observation requires inverting a multivalued functional that has an infinite number of equally likely solutions (Tolwinski-Ward *et al.*, 2014). Without additional information or constraints, it is impossible to overcome this loss of climate information. Further complicating a multivariate climate reconstruction is the development of site-selection techniques in dendrochronology that select for univariate climate signals in tree ring chronologies, resulting in a non-random sample of chronologies (Cook and Kairiukstis, 1990).

The problem of reconstructing climate using univariate tree ring chronologies has been addressed in the literature using a variety of methods. Many authors have attempted to solve the climate reconstruction problem with linear statistical methods that allow for estimation in the presence of a rank-deficient design matrix, including regularized expectation–maximization algorithms, truncated total least squares, and multivariate calibration methods including partial least squares (Rutherford *et al.*, 2003; Zhang *et al.*, 2004; Rutherford *et al.*, 2005; Mann *et al.*, 2008; Steig *et al.*, 2009). These methods all assume a linear relationship between the observed climate and tree ring chronologies. More recently, investigators have developed new methods using correlated spatial random effects (Guillot *et al.*, 2013) or nonlinear processes (Tolwinski-Ward *et al.*, 2014) to link climate and proxy data. Other statistically rigorous work has been performed to investigate the environmental mechanisms of tree growth using dendrochronological data (Hooten and Wikle, 2007). To develop the best methodology for paleoclimate prediction, Tingley *et al.* (2012) recommend collaboration between climate scientists and statisticians to develop a Bayesian hierarchical framework that combines scientifically motivated processes with flexible spatiotemporal methods. Our contribution is to bridge the gap between the linear multivariate calibration and regularized expectation–maximization methods and the more mechanistic, ecologically motivated methods. Thus, we approach the problem of multivariate climate reconstruction in a novel way.

Our work is based on the work of Tolwinski-Ward *et al.* (2014), where we make many computational and methodological improvements. First, we propose different forms of the deterministic growth function linking climate observations to tree ring widths within a framework that allows for a statistically principled evaluation of the effects of different growth model forms. Next, we constrain the multivariate climate predictions by modeling a differential growth response for each tree species, as is common in multispecies ecological modeling. The multispecies approach ameliorates the multivalued inverse problem while gaining the additional benefit of inferring niche climate response of each tree species. We propose a computationally efficient calibration model to link observed tree ring widths to the deterministic growth model output, speeding algorithmic convergence and improving mixing during parameter estimation. We use an upscaling data model that links monthly scale climate to annual-scale observed tree ring chronologies and propose a dynamic model for downscaling annual tree ring observations to monthly climate anomalies. To facilitate climate backcasting, we propose a novel model for temperature and precipitation using a dynamic, flexible, multi-scale process. Finally, we conduct a simulation experiment to validate the model's predictive ability using a proper scoring rule that selects the optimal predictive model. We do not include a direct comparison of our model with that of Tolwinski-Ward *et al.* (2014) for a number of reasons. First, Tolwinski-Ward *et al.* (2014) model soil moisture, a nonlinear function of temperature and precipitation, instead of precipitation, making a direct comparison difficult. In addition, the model of Tolwinski-Ward *et al.* (2014) was fit to a reconstruction period of 50 years and took approximately 350 CPU hours to fit in a high-performance computing environment. We aim to reconstruct 456 years; thus, our model needs to be computationally feasible for longer timescales.

We introduce the climate data in Section 2.1, detailing the transformation of the climate data to anomaly space. In Section 2.2, we propose a calibration model to align the observed tree ring data with output from a deterministic tree ring growth model, thereby assimilating the climate and tree ring data. We describe the deterministic link function that takes climate inputs and grows synthetic tree ring widths in Section 2.3 and present a dynamic, multi-scale model that facilitates backcasting at a monthly scale in Section 2.4. Finally, we formulate the posterior distribution on which inference is desired and outline the sampling algorithm used for estimation in Section 2.5.

Ultimately, our modeling effort is successful if accurate predictions of historical climate are obtained along with associated uncertainty. In Section 3, we discuss a scoring rule commonly used in dendrochronology and describe an alternative that is proper. In Section 4, we present the pseudoproxy simulation study and interpret the results of this experiment. We conclude by presenting our reconstruction of temperature and precipitation for the Hudson Valley data in Section 5 and discuss these results in Section 6.

2. THE MODEL

2.1. Climate data model

Our instrumental period climate data are monthly Parameter–Elevation Relationships on Independent Slopes Model (PRISM) gridded data products from 1895 to 2005 (PRISM Climate Group, 2004). From PRISM, we obtain temperature and precipitation at $I = 16$ sites within the Hudson Valley of New York, USA. Monthly temperature T_{ts} and log total precipitation P_{ts} represent regional averages across the I sites for year t and month s. For our data, there are no months with zero precipitation, and the log transform is well defined, although this is not true in general. Within a given month, the temperature and log precipitation measurements are approximately normally distributed; hence,

we model them using Gaussian distributions. It is common to model the climate dynamics in anomaly space, resulting in standard normal temperature and log precipitation anomalies

$$w_{T_{ts}} = \frac{T_{ts} - \bar{T}_s}{s_{T_s}} \tag{1}$$

$$w_{P_{ts}} = \frac{P_{ts} - \bar{P}_s}{s_{P_s}} \tag{2}$$

where $\bar{T}_s, \bar{P}_s, s_{T_s}$, and s_{P_s} are the sample means and standard deviations of temperature and log precipitation for month s, respectively. We define the monthly scale anomaly temperature and log precipitation vectors for year $t, \boldsymbol{w}_{T_t} = (w_{T_{t1}}, \ldots, w_{T_{t12}})'$ and $\boldsymbol{w}_{P_t} = (w_{P_{t1}}, \ldots, w_{P_{t12}})'$, the bivariate climate anomaly vector for year $t, \boldsymbol{w}_t = (\boldsymbol{w}'_{T_t}, \boldsymbol{w}'_{P_t})'$, and the vector of all climate anomalies, $\boldsymbol{w} = (\boldsymbol{w}'_1, \ldots \boldsymbol{w}'_\tau)'$, for years $t = 1, \ldots, \tau$.

2.2. Tree ring data model

The 34 tree ring chronologies contain measurements from $J = 12$ different tree species in the Hudson Valley region of New York, shown in Table 1. Each of the tree ring chronologies is at least 160 years long, and three chronologies date back to 1453. Further details about how the tree ring chronology data were collected and processed can be found in the work of Pederson *et al.* (2013).

The tree ring observation y_{itj} represents the annual observed tree ring width from the ith location for species j at time t. We model the tree ring width data as arising from a mixture of two distributions that depend on different forms of a deterministic growth model response to climate:

$$y_{itj} | \boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{VS}}, \boldsymbol{\theta}_j^{\mathrm{Pro}}, z_j, \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\sigma}^2 \sim \begin{cases} \mathrm{N}\left(\beta_{0_j} + \beta_{1_j} f\left(\boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{VS}}\right), \sigma_j^2\right) & \text{if } z_j = 0 \\ \mathrm{N}\left(\tilde{\beta}_{0_j} + \tilde{\beta}_{1_j} \tilde{f}\left(\boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{Pro}}\right), \tilde{\sigma}_j^2\right) & \text{if } z_j = 1 \end{cases} \tag{3}$$

where $f(\boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{VS}})$ and $\tilde{f}(\boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{Pro}})$ are deterministic link functions that "grow" tree rings given the climate anomaly $\boldsymbol{w}_t$ and species-specific model parameters $\boldsymbol{\theta}_j^{\mathrm{VS}}$ and $\boldsymbol{\theta}_j^{\mathrm{Pro}}$ for the two different growth models denoted as VS and Pro to be discussed in more detail later. In Sections 2.3, 2.3.1, and 2.3.2, we describe the structure of the growth link function and the growth model parameters $\boldsymbol{\theta}_j$. The stochastic indicator variable z_j selects the growth model form appropriate for species j. Examination of the posterior distribution of $\boldsymbol{z}$ provides a statistically principled method for comparing proposed growth model forms. For $j = 1, \ldots, J$, we specify a binomial prior on z_j with prior probability 0.5 to allow each growth model to be equally likely *a priori*.

In contrast with the work of Tolwinski-Ward *et al.* (2014) where they standardize the tree ring growth model output to have the same mean and standard deviation as the observed tree ring chronology, we use (3) to calibrate the tree ring growth model output to the observed chronology. The growth model-specific parameters β_{0j} and β_{1j} ($\tilde{\beta}_{0j}$ and $\tilde{\beta}_{1j}$) center and scale the synthetic tree ring growth model output to be coherent with the observed tree ring widths up to an error with variance $\sigma_j^2(\tilde{\sigma}_j^2)$, where the symbol $\sim$ distinguishes the growth model form. We specify priors for the VS growth calibration model parameters, with those for the Pro being defined similarly. For each species $j = 1, \ldots, J$, we specify a hierarchically pooled prior across species $\beta_{0j} \sim \mathrm{N}(\mu_{\beta_0}, \sigma^2_{\beta_0})$ and $\beta_{1j} \sim \mathrm{N}(\mu_{\beta_1}, \sigma^2_{\beta_1})$ with hyperpriors $\mu_{\beta_0} \sim \mathrm{N}(1, 1)$, $\sigma^2_{\beta_0} \sim \mathrm{IG}(1, 1)$, $\mu_{\beta_1} \sim \mathrm{N}(1, 1)$, and $\sigma^2_{\beta_1} \sim \mathrm{IG}(1, 1)$. For the $j = 1, \ldots, J$ calibration standard deviation parameters, we define the prior $\sigma_j \sim \log \mathrm{N}(\mu_{\sigma^2}, \sigma^2_{\sigma^2})$ with hierarchical pooling hyperparameters $\mu_{\sigma^2} \sim \mathrm{N}(0, 10)$ and $\sigma^2_{\sigma^2} \sim \mathrm{IG}(1, 1)$. where N, log N, and IG refer to the normal, lognormal, and inverse Gamma distributions.

Table 1. Species used in the reconstruction

Species	Number of chronologies
Tsuga canadensis	3
Liriodendron tulipifera	3
Juniperus virginiana	1
Carya glabra	3
Quercus stellata	1
Betula lenta	2
Pinus rigida	1
Quercus montana	5
Quercus rubra	4
Quercus alba	5
Quercus velutina	3
Carya ovata	1
Chamaecyparis thyoides	2

2.3. Process model

The statistical learning about past climate is achieved through a deterministic tree ring growth model that uses monthly temperature and precipitation as inputs. Formation of tree ring widths occur periodically throughout the growing season, with the rate of growth influenced by the prevailing weather. However, the data and the forward model are in the form of monthly climate variables and annual tree growth increments. The monthly temporal scale of the temperature and precipitation presents a change of support problem because the observed tree ring data occur at annual, not monthly, resolution. Hence, prediction of climate at a monthly scale involves downscaling the annual resolution tree ring information into monthly increments. In years without climate observations, we use (1) and (2) to accomplish the downscaling by imposing the observed monthly climate patterns. The downscaling assumes that the monthly patterns and dynamics of temperature and precipitation within a given year are estimable from the observational period and that these patterns are representative of the reconstruction period. For example, because temperature is strongly seasonal, the pattern of warm temperatures in summer and cold temperatures in winter will be the same regardless of the absolute magnitude of the annual or decadal temperature patterns. Precipitation is less consistent, but annual variability in these patterns allows for realistic downscaling. To align the different data sources occurring at different scales, we use a discrete-time approximation of the continuous growth process on a monthly scale, thus aligning tree growth with the PRISM data. We then aggregate the monthly growth increments to an annual resolution, thereby upscaling the continuous growth from temperature and precipitation for year t, species j, and month s into the growth increments. Thus, the representation of annual tree ring growth under the VS model form (the Pro model form is defined similarly, replacing $\boldsymbol{\theta}_j^{\mathrm{VS}}$ with $\boldsymbol{\theta}_j^{\mathrm{Pro}}$, f with $\tilde{f}$ and g with $\tilde{g}$) is

$$f\left(\boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{VS}}\right) = \sum_{s=1}^{12} \chi_s \min\left(g\left(w_{T_{ts}}, \boldsymbol{\theta}_j^{\mathrm{VS}}, \bar{T}_s, s_{T_S}\right), g\left(w_{P_{ts}}, \boldsymbol{\theta}_j^{\mathrm{VS}}, \bar{P}_s, s_{P_S}\right)\right) \tag{4}$$

a weighted sum of monthly growth function responses to temperature and precipitation where the weights χ_s are the monthly average length of daylight scaled to the unit interval $(0, 1)$. Thus, χ_s scales growth to the known average amount of sunlight in a month, mimicking the physiology of tree growth. The monthly scale function g ($\tilde{g}$) downscales the marginal annual climate anomaly to a monthly value and "grows" a monthly tree ring increment given the marginal climate. These marginal monthly growth functions then are combined by taking the minimum, allowing each month's growth to be either temperature or precipitation sensitive. Hence, the tree ring growth model follows the "principle of limiting factors," which states that tree growth is constrained by the climatic variable that is limiting (Fritts, 1976). In the next sections, we describe the two forms of g and $\tilde{g}$ used in our model, which we call Vaganov–Shashkin Lite ("VS-Lite"), representing the growth model form used by Tolwinski-Ward *et al.* (2014), and "probit."

Biologically, all tree species in the tree ring network have a similar response to climate. For example, all tree species in the network have more similar climatological needs than, say, a tropical tree, which requires an entirely different climate. One explanation for what allows many tree species to grow in the same region are that different species occur in niche deviations from the overall mean response to climate. Because of this, recent research in climate reconstruction methods demonstrates that inclusion of multiple tree species improves climate reconstruction skill (García-Suárez *et al.*, 2009; Cook and Pederson, 2011). Allowing for species-specific climate responses ameliorates the difficulties of inverting the multivalued functional relationship between climate and tree ring widths by placing multiple constraints on the set of possible climate scenarios. Using multiple species is beneficial as it provides more constraints on climate and may allow for more precise estimation than in Tolwinski-Ward *et al.* (2014). However, separate growth parameters for each species increases the number of model parameters to be estimated; therefore, we borrow strength by modeling the growth parameters hierarchically to improve parameter estimation and predictive skill (Gelman and Hill, 2006; Hooten and Hobbs, 2015; Hobbs and Hooten, 2015). By treating each tree species' response to climate as a random draw from a pooled distribution, the model shares information among tree species. This borrowing of strength among species is easily incorporated into the probit growth model framework but is not straightforward in the VS-Lite framework.

2.3.1. *VS-Lite tree ring growth model*

The "VS-Lite" model represents a statistical approximation to the Vaganov–Shashkin Lite model that has been shown to create reasonable tree ring width dendrochronologies given climate (Shashkin and Vaganov, 1993; Vaganov *et al.*, 2006; Tolwinski-Ward *et al.*, 2011). The VS-Lite tree ring growth model uses the linear ramp function

$$\Psi(\eta) = \begin{cases} 0 & \text{if } \eta \leqslant 0 \\ \eta & \text{if } 0 < \eta < 1 \\ 1 & \text{if } \eta \geqslant 1 \end{cases}$$

as a link between climate and tree ring width growth shown in Figure (1). The monthly tree ring growth functions for temperature and precipitation using VS-Lite are

$$g\left(w_{T_{ts}}, \boldsymbol{\theta}_j^{\mathrm{VS}}, \bar{T}_s, s_{T_s}\right) = \psi\left(\frac{w_{T_{ts}} s_{T_s} + \bar{T}_s - T_{\min_j}}{T_{\max_j} - T_{\min_j}}\right) \tag{5}$$

$$g\left(w_{P_{ts}}, \boldsymbol{\theta}_j^{\mathrm{VS}}, \bar{P}_s, s_{P_s}\right) = \psi\left(\frac{\exp\left\{w_{P_{ts}} s_{P_s} + \bar{P}_s\right\} - P_{\min_j}}{P_{\max_j} - P_{\min_j}}\right) \tag{6}$$

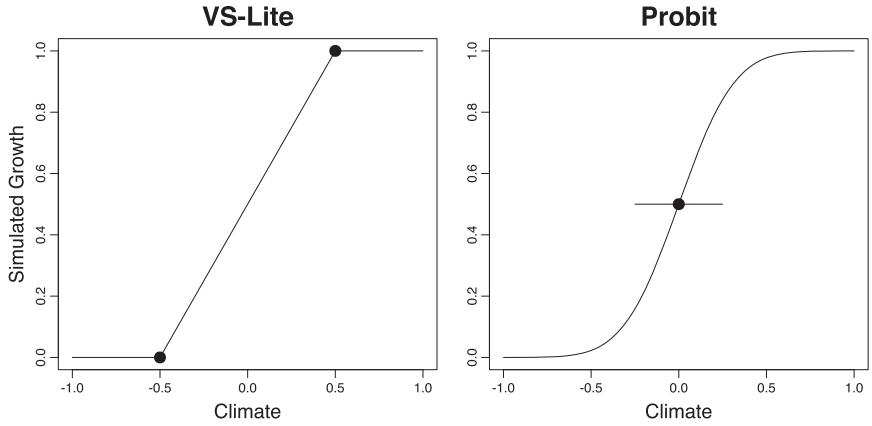

Figure 1. Example VS-Lite and probit ramp functions. The black dots on the VS-Lite plot represent the locations of $T_{\min_j}$ and $T_{\max_j}$, the temperatures below which growth is zero or above which growth is optimal (equivalently $P_{\min_j}$ and $P_{\max_j}$ for precipitation). The black dot on the probit plot represents the probit mean growth response to temperature μ_{T_j} (μ_{P_j} for precipitation) and the line shows the probit standard deviation of growth response to temperature σ_{T_j} (σ_{P_j} for precipitation)

and are combined according to the principle of limiting factors using (4). The growth parameters for the VS-Lite tree ring growth model are $\boldsymbol{\theta}_j^{\mathrm{VS}} = (T_{\min_j}, T_{\max_j}, P_{\min_j}, P_{\max_j})'$ for each species $j = 1, \ldots, J$. For temperatures below $T_{\min_j}$, there is no tree ring growth. For monthly temperatures between $T_{\min_j}$ and $T_{\max_j}$, tree ring growth increases linearly over the unit interval $(0, 1)$. When the monthly temperature exceeds $T_{\max_j}$, growth occurs at a maximum rate, taking a value of 1. The interpretation for $P_{\min_j}$ and $P_{\max_j}$ is similar.

To complete the VS-Lite growth model parameterization, we specify priors for the growth parameters $\boldsymbol{\theta}_j^{\mathrm{VS}}$ using a four-parameter Beta(α, β, L, U) distribution, a Beta(α, β) distribution that has been shifted and scaled to have a lower endpoint L and an upper endpoint U. These prior models require expert knowledge to specify; hence, we follow the recommendations of Tolwinski-Ward *et al.* (2013), recognizing that the previous work modeled soil moisture instead of precipitation under a different climate scenario than the Hudson Valley. We chose informative priors that place a majority of the probability mass in the center of the support that result in growth model priors that are reasonable given the climate of the Hudson Valley. For $j = 1, \ldots, J$, the VS-Lite parameter priors are $T_{\min_j} \sim$ Beta(9, 5, 0, 9), $T_{\max_j} \sim$ Beta(3.5, 3.5, 10, 24), $P_{\min_j} \sim$ Beta(3.5, 3.5, 65, 85), and $P_{\max_j} \sim$ Beta(3.5, 3.5, 85, 105). Despite existing knowledge of tree response to climate, the VS-Lite tree ring growth model could be sensitive to prior specification, especially if the true growth parameter values lie outside the range of prior support. In this case, the posterior probability of correctly estimating the true parameter value is exactly zero. Therefore, climate reconstruction using the VS-Lite growth model formulation has the potential to be highly influenced by misspecification of the prior support.

2.3.2. *Probit tree ring growth model*

An alternative to the VS-Lite tree ring growth model is the probit growth model, which was not examined by Tolwinski-Ward *et al.* (2014). The probit growth model replaces the linear function $\psi(\eta)$ in the VS-Lite growth model by the infinitely differentiable inverse normal cumulative distribution function $\Phi^{-1}(\eta)$. The probit growth model parameters have infinite support and are parameterized as $\boldsymbol{\theta}_j^{\mathrm{Pro}} = (\mu_{T_j}, \sigma^2_{T_j}, \mu_{P_j}, \sigma^2_{P_j})'$. There are only slight differences in the shape of the VS-Lite and probit growth model functions, as seen in Figure 1. The probit ramp function produces a smoother response to climate than the VS-Lite ramp function, but other than smoothness, the two shapes are quite similar; hence, it seems likely that the shape of the growth function alone will not significantly improve predictive performance. The motivation for the probit growth function is that our model framework takes advantage of statistical properties not available in the VS-Lite framework. First, the prior support for the probit growth model is the real line, in comparison with the VS-Lite prior support that is restricted to compact support. In practice, if the true growth model in the VS-Lite framework is not in the range of prior support, the posterior probability of correctly estimating this parameter is exactly zero, regardless of the amount and quality of data. The probit growth function does not suffer from this problem. Additionally, the probit model can be easily extended to a hierarchical pooling framework. We propose the probit growth model form to evaluate the influences of these desirable statistical probabilities on the climate reconstruction. If the predictive performance of the two models is equivalent, the probit model will be preferred because of these properties.

The monthly probit growth increments due to temperature and precipitation are

$$\tilde{g}\left(w_{T_{ts}}, \boldsymbol{\theta}_j^{\mathrm{Pro}}, \bar{T}_s, s_{T_s}\right) = \Phi^{-1}\left(\frac{w_{T_{ts}} s_{T_s} + \bar{T}_s - \mu_{T_j}}{\sigma_{T_j}}\right) \tag{7}$$

$$\tilde{g}\left(w_{P_{ts}}, \boldsymbol{\theta}_j^{\mathrm{Pro}}, \bar{P}_s, s_{P_s}\right) = \Phi^{-1}\left(\frac{\exp\{w_{P_{ts}} s_{P_s} + \bar{P}_s\} - \mu_{P_j}}{\sigma_{P_j}}\right) \tag{8}$$

where the parameters μ_{T_j} and μ_{P_j} represent the species-specific temperature and precipitation probit mean growth rate, respectively, and the parameters σ_{T_j} and σ_{P_j} control the effective range where tree growth responds to climate, the probit standard deviation. A species with a higher value of μ_{T_j} will experience better growth under warmer weather than a species with a lower value of μ_{T_j}, and a species with a larger σ_{T_j} will have a larger range of temperatures in which that tree species will grow than a species with a smaller σ_{T_j}. The interpretation of these growth model parameters is similar for precipitation.

For $j = 1, \ldots, J$, we specify the probit growth model parameter distributions for the species as $\mu_{T_j} \sim \mathrm{N}(\mu_{\mu_T}, \sigma^2_{\mu_T})$, $\sigma_{T_j} \sim \log \mathrm{N}(\mu_{\sigma_T}, \sigma^2_{\sigma_T})$, $\mu_{P_j} \sim \mathrm{N}(\mu_{\mu_P}, \sigma^2_{\mu_P})$, and $\sigma_{P_j} \sim \log \mathrm{N}(\mu_{\sigma_P}, \sigma^2_{\sigma_P})$. The pooling of these effects occurs by adding one more level in the hierarchical model by defining a hyperprior model for each of the prior parameters above with $\mu_{\mu_T} \sim \mathrm{N}(13, 4)$, $\sigma^2_{\mu_T} \sim \mathrm{IG}(2, 0.5)$, $\mu_{\sigma_T} \sim \mathrm{N}(0, 1)$, $\sigma^2_{\sigma_T} \sim \mathrm{IG}(2, 0.5)$, $\mu_{\mu_P} \sim \mathrm{N}(85, 16)$, $\sigma^2_{\mu_P} \sim \mathrm{IG}(2, 0.5)$, $\mu_{\sigma_P} \sim \mathrm{N}(0, 2)$, and $\sigma^2_{\sigma_P} \sim \mathrm{IG}(2, 0.5)$. These prior values represent likely values that cover the range of growing season temperature and precipitation values seen in the Hudson Valley while being highly flexible, thus allowing the model to estimate the growth parameters more flexibly than the VS-Lite growth model.

2.4. Dynamic multi-scale climate process

We model temperature and log precipitation anomalies jointly with a dynamic, multi-scale model, allowing prediction of unobserved temperature and precipitation when combined with the tree ring chronology data. The model is a temporal vector autoregressive process among years, given a propagator matrix $\boldsymbol{A}$, and a correlated autoregressive process among months determined by the structure of a covariance matrix $\boldsymbol{\Sigma}$. To account for trend in the temperature anomaly time series during the observational period 1895–2010, we include an intercept Δ_0 and slope Δ_1 in the model for years after $t^* = 1895$. In years before t^*, we do not have observational data, and thus, we do not model a trend. This results in the dynamic de-trended process

$$\begin{pmatrix} \boldsymbol{w}_{T_t} - \Delta_0 \boldsymbol{J} - (t - t^*)\Delta_1 \boldsymbol{J} \\ \boldsymbol{w}_{Pt} \end{pmatrix} \sim \mathrm{N}\left(\boldsymbol{A} \begin{pmatrix} \boldsymbol{w}_{T_{t-1}} - \Delta_0 \boldsymbol{J} - (t - t^*)\Delta_1 \boldsymbol{J} \\ \boldsymbol{w}_{P_{t-1}} \end{pmatrix}, \boldsymbol{\Sigma} \right) \quad \text{if } t \geqslant t^* \tag{9}$$

$$\begin{pmatrix} \boldsymbol{w}_{T_t} - \Delta_0 \boldsymbol{J} \\ \boldsymbol{w}_{P_t} \end{pmatrix} \sim \mathrm{N}\left(\boldsymbol{A} \begin{pmatrix} \boldsymbol{w}_{T_{t-1}} - \Delta_0 \boldsymbol{J} \\ \boldsymbol{w}_{P_{t-1}} \end{pmatrix}, \boldsymbol{\Sigma} \right) \quad \text{if } t < t^* \tag{10}$$

The propagator matrix $\boldsymbol{A} = \begin{pmatrix} \phi_1, 0 \\ 0, \phi_2 \end{pmatrix} \otimes \boldsymbol{I}$ defines the annual-scale autocorrelation for the temperature and log precipitation anomalies, where ϕ_1 and ϕ_2 are the annual autocorrelation parameters. In these expressions, $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{J}$ is a 12×1 vector of ones. We model the inter-annual covariance $\boldsymbol{\Sigma}$ with a temporal multivariate conditionally autoregressive (MCAR) structure, a generalization of the conditionally autoregressive (CAR) structure in time that allows for the within-year temperature and precipitation anomalies to have their own temporal autocorrelation parameters while also including a temporally explicit cross-correlation between the anomaly measurements (Mardia, 1988; Carlin and Banerjee, 2003; Gelfand and Vounatsou, 2003; Jin *et al.*, 2005). We found the MCAR model to be best fitting among a set of candidate models for temporal autocorrelation.

To construct the temporal MCAR covariance matrix $\boldsymbol{\Sigma}$, we first define a temporal CAR precision matrix that will be used as a component in building $\boldsymbol{\Sigma}$. The particular precision matrix we use, $\boldsymbol{Q}(\omega)$, specifies a process identical to an autoregressive process of order 1 in the time series literature if $|\omega| < 1$ (Cressie and Wikle, 2011, p. 170). We define the monthly temperature CAR precision matrix $\boldsymbol{Q}(\omega_T)$ with autocorrelation parameter ω_T and the monthly log precipitation CAR precision matrix $\boldsymbol{Q}(\omega_P)$ with autocorrelation parameter ω_P, allowing each climate variable to have its own monthly autocorrelation. We decompose $\boldsymbol{Q}(\omega_T) = \boldsymbol{L}_T \boldsymbol{L}'_T$ and $\boldsymbol{Q}(\omega_P) = \boldsymbol{L}_P \boldsymbol{L}'_P$ with a Cholesky decomposition (one could also use a spectral decomposition) and construct the MCAR precision matrix as

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma^2_w} \begin{pmatrix} \mathbf{L}'_T & \mathbf{0} \\ \mathbf{0} & \mathbf{L}'_P \end{pmatrix} (\boldsymbol{\Lambda} \otimes \mathbf{I}_{12}) \begin{pmatrix} \mathbf{L}_T & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_P \end{pmatrix} \tag{11}$$

where the matrix $\boldsymbol{\Lambda} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ is a correlation matrix with ρ representing the atemporal cross-correlation between temperature and precipitation, σ^2_ω is a global variance parameter, and $\otimes$ represents the Kronecker product. Thus, using (11), we model intra-annual autocorrelation, monthly autocorrelation, and cross-correlation in the climate process.

To significantly reduce computation time, we employ an empirical Bayes approach to process the climate data (Casella, 1985). We estimate the propagator matrix $\boldsymbol{A}$, the intra-annual covariance matrix $\boldsymbol{\Sigma}$, and the trend parameters Δ_0 and Δ_1 off-line using a hybrid Metropolis–Hastings and Gibbs Markov chain Monte Carlo (MCMC) sampling algorithm using $\mathrm{U}(-1, 1)$ priors for the parameters ρ, ω_T, and ω_P and an IG(1, 1) prior for the variance σ^2_w. After fitting the model, we use posterior median values of each parameter as estimates of the true processes parameters; thus, our posterior predictions do not include climate model parameter uncertainty, and our corresponding credible intervals will be overly optimistic.

2.5. Posterior

We desire inference on the posterior distribution and quantities derived from the posterior distribution. The posterior we seek to approximate with our MCMC algorithm is

$$\left[\boldsymbol{w}, \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\theta}^{\mathrm{VS}}, \boldsymbol{\theta}^{\mathrm{Pro}}, \sigma^2, z \mid \boldsymbol{y}, \boldsymbol{T}, \boldsymbol{P}\right]$$
$$\propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{t=1}^{\tau}\left[y_{itj} \mid \beta_{0_j}, \beta_{1_j}, \boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{VS}}, \sigma_j^2\right]^{(1-z_j)}$$
$$\times \left[y_{itj} \mid \tilde{\beta}_{0_j}, \tilde{\beta}_{1_j}, \boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{Pro}}\tilde{\sigma}_j^2\right]^{z_j}$$
$$\times \left[\boldsymbol{w}_t \mid \boldsymbol{w}_{t-1}, \boldsymbol{A}, \boldsymbol{\Sigma}, \Delta_0, \Delta_1\right]\left[\boldsymbol{\beta_0}\right]\left[\boldsymbol{\beta_1}\right]$$
$$\times \left[\boldsymbol{\theta}^{\mathrm{VS}}\right]\left[\boldsymbol{\theta}^{\mathrm{Pro}}\right]\left[\boldsymbol{\sigma}^2\right][z]$$

where the parameter models for the VS-Lite and probit growth model are represented as $[\boldsymbol{\theta}^{\mathrm{VS}}]$ and $[\boldsymbol{\theta}^{\mathrm{Pro}}]$, respectively. Implementation of a hybrid Metropolis–Hastings and Gibbs MCMC algorithm allows for estimation of the posterior distribution (Banerjee *et al.*, 2004; Carlin and Louis, 2011). Our model was implemented using the R software program (R Core Team, 2014), while leveraging significant portions of C++ code using RcppArmadillo (Eddelbuettel and Sanderson, 2014) to increase computation speed. The MCMC algorithm for each candidate model was run for 15,000 iterations with the first half discarded as burn-in with thinning every five observations, for three parallel chains, resulting in 4500 posterior samples for a total computation time of 1 h on a 2014 MacBook Pro. Convergence was assessed using Gelman and Rubin's (1992) $\hat{R}$ statistic and visual inspection of the trace plots.

3. MODEL EVALUATION

Traditional paleoclimate reconstructions evaluate predictive performance with out-of-sample data using the coefficient of efficiency (*CE*) (Cook *et al.*, 1994; Rutherford *et al.*, 2005; Tingley and Huybers, 2010a, 2010b). Despite the accepted use of this scoring statistic, Gneiting and Raftery (2007) suggest that skill scores like *CE* are improper in general. This implies that it is possible to have predictions that, under expectation, have better *CE* skill scores than a model that is optimal. Therefore, use of improper scoring rules can lead to incorrect inference about predictive skill among a set of predictive models. To prevent this, we use a proper scoring rule. Proper scoring rules guarantee that, under expectation, the optimal predictive model will have the best predictive score (Gneiting *et al.*, 2007; Gneiting, 2011; Hooten and Hobbs, 2015).

Our model produces a probabilistic forecast; hence, we use the continuous ranked probability score (CRPS), a proper scoring rule that accommodates both probabilistic and point forecasts. Several recent papers on late Holocene climate reconstructions have likewise made use of the CRPS (Barboza *et al.*, 2014; Werner and Tingley, 2015). Given a forecast with cumulative distribution function, F_t, at time t and out-of-sample observations $\boldsymbol{y}_{\mathrm{oos}}$, the CRPS is defined as

$$CRPS\left(\{F_t\}_{t=1}^{\tau}, \boldsymbol{y}_{\mathrm{oos}}\right) = -\sum_{t=1}^{\tau}\int_{-\infty}^{\infty}\left(F_t(y) - \boldsymbol{I}_{\{y \geq y_{\mathrm{oos},t}\}}\right)^2 \mathrm{d}y \tag{12}$$

Gneiting *et al.* (2007) show how (12) can be written alternatively as

$$CRPS\left(\{F_t\}_{t=1}^{\tau}, \boldsymbol{y}_{\mathrm{oos}}\right) = \sum_{t=1}^{\tau}\left(E_{F_t}\left|y_t - y_{\mathrm{oos},t}\right| - \frac{1}{2}E_{F_t}\left|y_t - y_t'\right|\right) \tag{13}$$

where y_t and y_t' are independent copies of a linear random variable with distribution function F_t and the expectation E is with respect to the probability density induced by F_t. The first expectation in the preceding equation measures calibration, the absolute error of the prediction relative to the out-of-sample value, and the second expectation rewards predictions that are sharp (i.e., narrow prediction intervals). Hence, the CRPS rewards probabilistic predictions that are accurate and precise.

In a Bayesian context, CRPS can be estimated after obtaining posterior samples. First, sample $\tilde{\boldsymbol{y}}^{(k)}$ from the posterior predictive distribution $[\tilde{\boldsymbol{y}}^{(k)} \mid \boldsymbol{y}, \boldsymbol{\theta}^{(k)}]$ at each post-burn-in iteration k. Then, the expression in (13) is approximated by

$$\widehat{CRPS}\left(\{\hat{F}_t\}_{t=1}^{\tau}, \boldsymbol{y}_{\mathrm{oos}}\right) = \sum_{t=1}^{\tau}\left(\frac{1}{K}\sum_{k=1}^{K}\left|\tilde{y}_t^{(k)} - y_{\mathrm{oos},t}\right| - \frac{1}{2K^2}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\left|\tilde{y}_t^{(k)} - \tilde{y}_t^{(\ell)}\right|\right) \tag{14}$$

We use the CRPS score in the simulation study to select the best model based on predictive performance. The CRPS is a negatively oriented scoring rule; therefore, the model with the lowest $\widehat{CRPS}$ is the best scoring model, and under expectation, the best predicting model.

4. SIMULATION STUDY

We consider three variants of the model presented earlier to conduct a reconstruction experiment. First, we consider the VS-Lite tree ring growth model, modifying the work of Tolwinski-Ward *et al.* (2014) to be consistent with our modeling framework using our climate model and multispecies approach. Second, we use only the probit tree ring growth model. Lastly, we use the mixture model described in (3) that allows for choice of tree ring growth model. We evaluate predictive performance of these candidate models over nine total simulation scenarios, using each of the tree ring growth models to simulate pseudoproxy data and fitting each growth model to each of the three simulated datasets.

We simulate the data as follows. First, we estimate $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$ from the instrumental climate data using (9). Using these estimates, we simulate a realization of the climate process with 446 years of pre-instrumental simulated climate variables while adding a trend with a slope of $1/110$ to the 110 years of temperature observations in the observation period, producing a current-day increase of 1°C. For the VS-Lite tree ring growth model simulation, we sample the growth parameters from uniform distributions. For each species j, $T_{\min_j} \sim \mathrm{U}(0, 9)$, $T_{\max_j} \sim \mathrm{U}(10, 24)$, $P_{\min_j} \sim \mathrm{U}(65, 85)$, and $P_{\max_j} \sim \mathrm{U}(85, 105)$. For the probit tree ring growth model, we sample the growth parameters $\mu_{T_j} \sim \mathrm{N}(16, 4)$, $\sigma_{T_j} \sim \log \mathrm{N}(\log(1), 0.5)$, $\mu_{P_j} \sim \mathrm{N}(85, 16)$, and $\sigma_{P_j} \sim \log \mathrm{N}(\log(2), 1)$ for each species $j = 1, \ldots, J$.

Next, we use each tree ring growth model to produce a noiseless tree ring chronology, standardizing each chronology to have a mean of 1 and a standard deviation of 0.2, as in the Hudson Valley dataset. Adding in noise representing measurement and processing error, we simulate noisy chronologies

$$y_{tj}^{\mathrm{sim}} \sim \mathrm{N}\left(\sqrt{(1-\sigma^2_{\mathrm{noise}})}\, f\left(\boldsymbol{w}_t, \boldsymbol{\theta}_j^{\mathrm{sim}}\right), \sigma^2_{\mathrm{noise}}\right) \tag{15}$$

The parameter $\sigma^2_{\mathrm{noise}}$ controls the signal-to-noise ratio in the simulated tree ring chronology, values of $\sigma^2_{\mathrm{noise}}$ near zero represent a high signal-to-noise ratio, while values near one represent a low signal-to-noise ratio. We let $\sigma^2_{\mathrm{noise}} = 0.75$, representing a signal-to-noise ratio of 0.58, which is at the high end of what is realistic for many tree ring chronologies (Smerdon, 2012). Figure 2 shows a realization of the simulated chronology and the observed chronology from Hudson Valley, demonstrating that our simulation methodology produces realistic tree ring chronologies. We apply the same structure of missingness to our simulated data as in the observed chronology, producing simulations as close as possible to the observed data.

The model characteristic most important to us in climate reconstruction is predictive ability. The reconstruction temperature and precipitation $\widehat{CRPS}$ values are shown in Table 2 for data simulated with the VS-Lite growth model, in Table 3 for data simulated with the probit growth model, and in Table 4 for data simulated with the mixture growth model. For each of the simulated datasets, $\widehat{CRPS}$ was estimated for both annual and growing period reconstructions, with bold scores highlighting the model that performs best for each simulated dataset and time period. As a baseline comparison, the $\widehat{CRPS}$ scores for a climatological prediction are included. The $\widehat{CRPS}$ values suggest all three models are similar in predictive ability, although there might be a slight preference for the mixture growth model. Based on this, we discuss the mixture growth model from here forward. Although the probit growth model is often not the best scoring model, the mixture growth model indicator variable, z, suggests an even split between the VS-Lite and probit tree ring growth models (the probit model is selected about 53% of the time) within the MCMC chain. An example reconstruction using the mixture tree ring growth model on the simulated data (Figure 3) demonstrates that the reconstruction performs quite well for log precipitation but is rather uninformative for temperature. Note that around 1650–1700, the uncertainty intervals in the log precipitation reconstruction increase in width, and predictive skill decreases because of increasing numbers of missing chronologies as well as the amount of replication within a given chronology decreasing. This uncertainty provides information about when the reconstruction is performing well and over what time periods the predictions are no better than climatology without relying on calibration skill measures like *CE*.

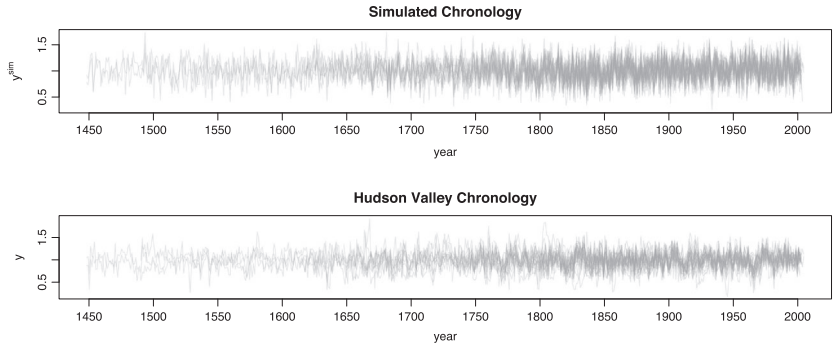

Figure 2. Simulated and observed tree ring width chronology for Hudson Valley

Table 2. Continuous ranked probability scores (CRPS) for annual temperature, growing season temperature, annual precipitation, and growing season precipitation for data simulated with the Vaganov–Shashkin Lite (VS-Lite) growth model

$\widehat{CRPS}$	Climatology	VS-Lite	Probit	Mixture
Annual temperature	531.29	**455.71**	455.75	455.98
Growing season temperature	486.52	**462.71**	463.93	464.45
Annual log precipitation	98.53	85.20	85.16	**84.86**
Growing season log precipitation	130.92	91.65	92.20	**90.63**

Bold scores highlight the model that performs best for each simulated dataset and time period.

Table 3. Continuous ranked probability scores (CRPS) for annual temperature, growing season temperature, annual precipitation, and growing season precipitation for data simulated with the probit growth model

$\widehat{CRPS}$	Climatology	VS-Lite	Probit	Mixture
Annual temperature	571.92	446.90	446.99	**446.72**
Growing season temperature	482.51	**450.02**	451.15	451.04
Annual log precipitation	101.06	90.61	**90.36**	90.69
Growing season log precipitation	129.34	**93.25**	93.59	93.61

Bold scores highlight the model that performs best for each simulated dataset and time period.
VS-Lite, Vaganov–Shashkin Lite.

Table 4. Continuous ranked probability scores (CRPS) for annual temperature, growing season temperature, annual precipitation, and growing season precipitation for data simulated with the mixture growth model

$\widehat{CRPS}$	Climatology	VS-Lite	Probit	Mixture
Annual temperature	540.23	438.65	**438.33**	438.80
Growing season temperature	489.36	440.66	441.17	**440.60**
Annual log precipitation	97.82	85.28	85.40	**84.49**
Growing season log precipitation	126.33	90.51	91.02	**90.04**

Bold scores highlight the model that performs best for each simulated dataset and time period.
VS-Lite, Vaganov–Shashkin Lite.

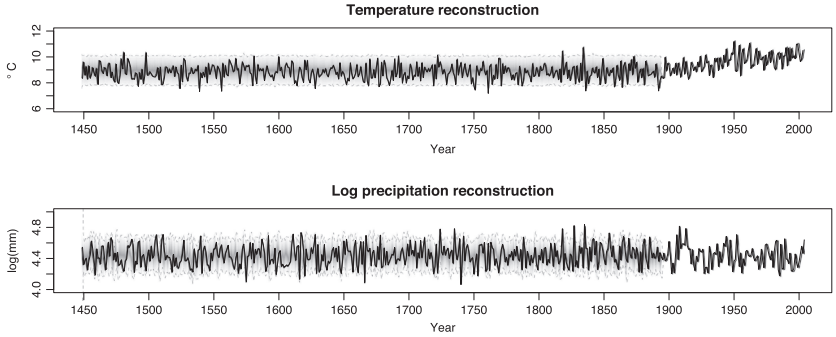

Figure 3. Reconstruction of temperature and log precipitation from data simulated with the mixture growth model and fit with the mixture growth model. The grey shading is proportional to the posterior predictive density, the dotted grey lines show the 95% posterior predictive credible interval, and the black line is the simulation truth

In addition to producing accurate climate reconstructions, all growth model types recover the simulated growth model parameters when the model is applied to data simulated with the same structure, showing that all of the models are capable of estimating the simulated growth model parameters. This result validates the growth model's usefulness for investigation of climatological niches for different tree species as well as providing accurate predictions of climate. Yet, despite accurately estimating the temperature growth parameters, the temperature reconstruction is not much more informative than a climatological prediction. Hence, the simulation suggests the failure to predict temperature patterns is not due to poor estimation of the simulated growth parameters. To further illuminate why the model is unable to reconstruct temperature, Figure 4 shows the observed monthly temperature and precipitation patterns with the probit growth model parameters shaded showing the range over which the model is sensitive to climate. We obtain very little learning about temperature at an annual scale, as the model is sensitive to temperature for, at best, two to four months of the year and only in the spring/fall with no learning about temperature during the winter (when temperature is most variable) and summer months. In fact, the effective learning about temperature occurs in much less than four months because of the multiplicative interaction between temperature and precipitation in (4). Therefore, any learning about temperature from the growth model is dominated by monthly variability in temperature. In contrast, the reconstruction of precipitation is quite accurate, especially in the near past, because the model is sensitive to precipitation values for all months of the year. This illustrates that the change of support from monthly to annual scale causes few problems for precipitation. An

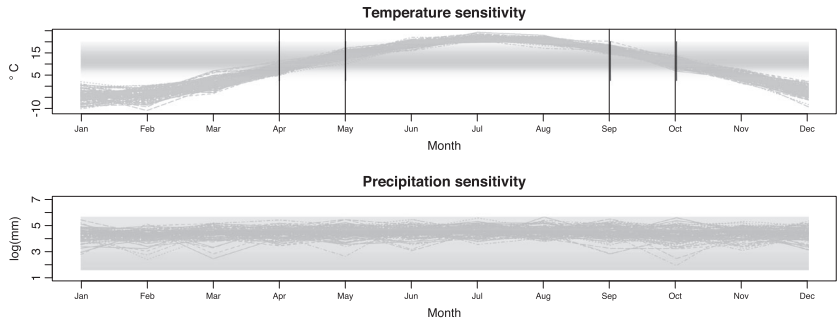

Figure 4. Reconstruction of annual temperature depends on the growth parameters. The shaded area shows where the growth parameters are sensitive to climate, and the grey lines are observed climate. The vertical lines in the temperature plot show that, at best, the annual reconstruction is dependent on four or fewer months. Thus, annual-scale information about temperature is hard to recover from only a few months, while precipitation information can be extracted across all months

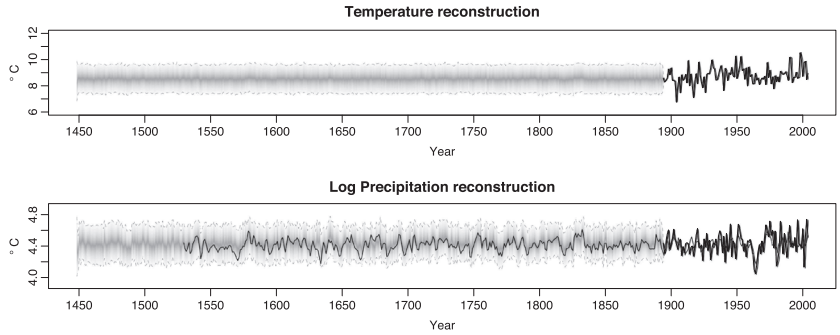

Figure 5. Climate reconstruction from the Hudson Valley chronology. The grey shading is proportional to posterior predictive density, the dotted lines show the 95% credible intervals, and the solid black line in the precipitation plot is a translated and scaled reconstruction of drought (Palmer Drought Severity Index) from Pederson *et al.* (2013)

investigation of climate scenarios with a more significant overlap between the climate and tree growth sensitivities in simulations (not shown in these results) demonstrates potential for reconstruction of both temperature and precipitation patterns, presenting opportunities for further applications to other datasets.

5. HUDSON VALLEY RESULTS

Based on the predictive results in the simulation study, we apply the mixture tree ring growth model to the Hudson Valley data. Based on our simulations and expert priors, the most obvious characteristic of the reconstruction is that we obtain only minimal learning about temperature, while the precipitation reconstruction performs well. To validate our predictive model, we translate and scale the reconstruction of Palmer Drought Severity Index (PDSI) in Pederson *et al.* (2013), a reconstruction that uses the same Hudson Valley dataset but very different methods, and compare this result with our reconstruction of log precipitation. Figure 5 shows that our log precipitation reconstruction strongly correlates ($r = 0.74$) with the previous PDSI reconstruction effort for the near past, while having the added benefit of explicitly accounting for uncertainty. It seems reasonable that PDSI would be correlated with log precipitation because PDSI is a measure of drought severity, and because there is little inter-annual variability in temperature, the primary driver of drought (and tree growth) in the Hudson Valley is the amount of precipitation (Martin-Benito and Pederson, 2015).

The uncertainty estimates for the reconstruction of log precipitation provide insight for the prediction's reliability. As we backcast in time, the 95% credible interval size approaches the size of predictive intervals derived using the observational data only, suggesting that the reconstruction lacks predictive ability before approximately 1650–1700. This loss of predictive ability, as seen in the simulation study, occurs during a time period where many tree ring chronologies are lacking enough replication to be included in the ring network. As such, deviations between our log precipitation reconstruction and the previous reconstruction of PDSI in Figure 5 before 1650–1700 can be explained in the context of prediction uncertainty.

In addition to the climate reconstructions, it is possible to obtain inference about the climatological niches that different tree species occupy from the growth parameters. Many tree species in the Hudson Valley dataset have only one or two chronologies; therefore, inference is limited because of the relatively small sample size and associated overlap in uncertainty intervals. For the VS-Lite growth model, the posterior samples show little ability to discern different ecological niches, but the posterior samples from the probit growth model posterior show many species have statistically different mean responses to climate. This demonstrates that the probit growth model can be easily adapted for ecological learning in a richer dataset where climatological prediction is not of interest. Also important is the posterior distribution of the indicator of model importance z. We find that the model has a slight preference for the probit growth model using the Hudson Valley

data, with a posterior probability $P(z = 1) = 0.54$. This is not substantially different than the prior probability of 0.5, which is expected because the two growth functions have very similar shapes. The slight preference for the probit growth form comes from improved parameter estimation through the use of a hierarchically pooled prior model.

6. DISCUSSION

We presented a methodological framework for reconstructing paleoclimate using a mechanistic Bayesian hierarchical model motivated by the work of Tolwinski-Ward *et al.* (2014). We proposed a novel probit tree ring growth model that takes advantage of the biology of tree growth and pools tree growth parameters to improve estimation and decrease sensitivity to prior specification. This new growth function was proposed in a framework that rigorously evaluates the growth model influence. Our extension to multispecies modeling of tree response to climate constrained predictive backcasts to a climate scenario consistent with the data, thereby improving predictive skill over models that do not explicitly incorporate multiple species. The introduction of a calibration model reduced computation time and improved MCMC convergence over previous standardization methods. We developed an upscaling and downscaling model to align the different data sources on the same scale and proposed a novel dynamic joint process model for temperature and precipitation accounting for temporal correlation and cross-correlation. Our pseudoproxy simulation experiment evaluated predictions with a proper scoring rule that uses the full probabilistic forecast. The simulation study provided insight about model performance and provided feedback that can be used to interpret results from the Hudson Valley data.

By using a biologically motivated Bayesian hierarchical model for reconstructing climate processes from tree rings widths, we combine statistical techniques with scientific models developed in dendrochronology. Our modeling framework explicitly accounts for uncertainties, in comparison with many previous climate reconstructions that use linear statistical methods and rely on asymptotic assumptions or bootstrap algorithms to estimate uncertainties, as discussed by Tingley and Huybers (2010b). An added benefit to using a biologically motivated model is the ability to adapt this modeling framework to make inference about ecological niches that different tree species occupy from the growth model parameter estimates.

By outlining the model shortcomings in detail, we gain a better understanding of the problem of multivariate climate reconstructions and can develop new ideas for improvement. Tree ring growth models are simplifications of the true biological process, and more scientific realism can be added. For instance, at very warm temperatures, tree growth is inhibited because of water lost by transpiration. Including more accurate biology in the growth model will increase the number of months that are sensitive to the growth process and alleviates the change of support problem for the prediction of temperature. Another option to improve temperature predictions is to modify the limiting factor approach in (4) to allow more months to contribute to the reconstruction of temperature. Generalizing the multi-scale climate process model will allow more complicated dynamics to be incorporated into the model, either through a statistical effort by allowing time-varying, non-stationarity processes or by assimilating deterministic climate forcings in the mean of the reconstruction, as in Li *et al.* (2010) and Barboza *et al.* (2014). The introduction of a spatially explicit model could lead to spatially explicit climate field reconstructions but would require a more computationally efficient method of fitting Bayesian models, like Hamiltonian Monte Carlo, variational Bayes, or integrated nested Laplace approximations. These and other potential improvements present opportunities for collaboration with dendrochronologists, climate scientists, ecologists, and statisticians to increase understanding of the relationship between climate and tree growth.

Acknowledgements

This material is based upon work carried out by the PalEON Project (paleonproject.org) with support from the National Science Foundation MacroSystems Biology program under grant no. DEB-1241856. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US government. We would also like to thank Suz Tolwinski-Ward, Kristin Broms, and Ben Bird for their input on this manuscript.

REFERENCES

Banerjee S, Gelfand AE, Carlin BP. 2004. *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press: Boca Raton, Florida.

Barboza L, Li B, Tingley MP, Viens FG. 2014. Reconstructing past temperatures from natural proxies and estimated climate forcings using short- and long-memory models. *The Annals of Applied Statistics* **8**(4):1966–2001.

Bradley RS. 2011. High-resolution paleoclimatology. In *Dendroclimatology*. Springer: Netherlands; 3–15.

Carbone MS, Czimczik CI, Keenan TF, Murakami PF, Pederson N, Schaberg PG, Xu X, Richardson AD. 2013. Age, allocation and availability of nonstructural carbon in mature red maple trees. *New Phytologist* **200**(4):1145–1155.

Carlin BP, Banerjee S. 2003. Hierarchical multivariate CAR models for spatio-temporally correlated survival data. *Bayesian Statistics* **7**:45–63.

Carlin BP, Louis TA. 2011. *Bayesian Methods for Data Analysis*. CRC Press: Boca Raton, Florida.

Casella G. 1985. An introduction to empirical Bayes data analysis. *The American Statistician* **39**(2):83–87.

Christiansen B, Ljungqvist FC. 2011. Reconstruction of the extratropical NH mean temperature over the last millennium with a method that preserves low-frequency variability. *Journal of Climate* **24**(23):6013–6034.

Cook ER, Briffa KR, Jones PD. 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *International Journal of Climatology* **14**(4):379–402.

Cook ER, Kairiukstis LA. 1990. *Methods of Dendrochronology: Applications in the Environmental Sciences*. Springer: Dordrecht, The Netherlands.

Cook ER, Pederson N. 2011. Uncertainty, emergence, and statistics in dendrochronology. In *Dendroclimatology*. Springer, 77–112.

Cressie N, Wikle CK. 2011. *Statistics for Spatio-temporal Data*. John Wiley & Sons: Singapore.

D'Arrigo RD, Kaufmann RK, Davi N, Jacoby GC, Laskowski C, Myneni RB, Cherubini P. 2004. Thresholds for warming-induced growth decline at elevational tree line in the Yukon Territory, Canada. *Global Biogeochemical Cycles* **18**(3). http://onlinelibrary.wiley.com/doi/10.1029/2004GB002249/full.

Eddelbuettel D, Sanderson C. 2014. RcppArmadillo: accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis* **71**:1054–1063.
Evans MN, Tolwinski-Ward S, Thompson D, Anchukaitis KJ. 2013. Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary Science Reviews* **76**:16–28.
Fritts H. 1976. *Tree Rings and Climate*. Elsevier: New York, New York.
García-Suárez A, Butler C, Baillie M. 2009. Climate signal in tree-ring chronologies in a temperate climate: a multi-species approach. *Dendrochronologia* **27**(3):183–198.
Gelfand AE, Vounatsou P. 2003. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* **4**(1):11–15.
Gelman A, Hill J. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press: Boca Raton, Florida.
Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science*:457–472.
Gneiting T. 2011. Making and evaluating point forecasts. *Journal of the American Statistical Association* **106**(494):746–762.
Gneiting T, Balabdaoui F, Raftery AE. 2007. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**(2):243–268.
Gneiting T, Raftery AE. 2007. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* **102**(477):359–378.
Gotway CA, Young LJ. 2002. Combining incompatible spatial data. *Journal of the American Statistical Association* **97**(458):632–648.
Griffin D, Woodhouse CA, Meko DM, Stahle DW, Faulstich HL, Carrillo C, Touchan R, Castro CL, Leavitt SW. 2013. North American monsoon precipitation reconstructed from tree-ring latewood. *Geophysical Research Letters* **40**(5):954–958.
Guillot D, Rajaratnam B, Emile-Geay J. 2013. Statistical paleoclimate reconstructions via Markov random fields. arXiv preprint arXiv:1309.6702.
Hobbs NT, Hooten MB. 2015. *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press: Princeton, New Jersey.
Hooten M, Hobbs N. 2015. A guide to Bayesian model selection for ecologists. *Ecological Monographs* **85**(1):3–28.
Hooten MB, Wikle CK. 2007. Shifts in the spatio-temporal growth dynamics of shortleaf pine. *Environmental and Ecological Statistics* **14**(3):207–227.
Jin X, Carlin BP, Banerjee S. 2005. Generalized hierarchical multivariate CAR models for areal data. *Biometrics* **61**(4):950–961.
Jones P, Briffa K, Barnett T, Tett S. 1998. High-resolution palaeoclimatic records for the last millennium: Interpretation, integration and comparison with general circulation model control-run temperatures. *The Holocene* **8**(4):455–471.
Jones P, Briffa K, Osborn T, Lough J, Van Ommen T, Vinther B, Luterbacher J, Wahl E, Zwiers F, Mann M, *et al*. 2009. High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *The Holocene* **19**(1):3–49.
Li B, Nychka DW, Ammann CM. 2010. The value of multiproxy reconstruction of past climate. *Journal of the American Statistical Association* **105**(491):883–895.
Mann ME, Zhang Z, Hughes MK, Bradley RS, Miller SK, Rutherford S, Ni F. 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences* **105**(36):13252–13257.
Mardia K. 1988. Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis* **24**(2):265–284.
Martin-Benito D, Pederson N. 2015. Convergence in drought stress, but a divergence of climatic drivers across a latitudinal gradient in a temperate broadleaf forest. *Journal of Biogeography* **42**(5):925–937.
Melvin TM, Briffa KR. 2008. A "signal-free" approach to dendroclimatic standardisation. *Dendrochronologia* **26**(2):71–86.
Moberg A, Sonechkin DM, Holmgren K, Datsenko NM, Karlén W. 2005. Highly variable northern hemisphere temperatures reconstructed from low-and high-resolution proxy data. *Nature* **433**(7026):613–617.
Pederson N, Bell AR, Cook ER, Lall U, Devineni N, Seager R, Eggleston K, Vranes KP. 2013. Is an epic pluvial masking the water insecurity of the greater New York City region? *Journal of Climate* **26**(4):1339–1354.
PRISM Climate Group. 2004. O. S. U. (4 February).
R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
Rutherford S, Mann M, Delworth T, Stouffer R. 2003. Climate field reconstruction under stationary and nonstationary forcing. *Journal of Climate* **16**(3): 462–479.
Rutherford S, Mann M, Osborn T, Briffa K, Jones PD, Bradley R, Hughes M. 2005. Proxy-based Northern Hemisphere surface temperature reconstructions: sensitivity to method, predictor network, target season, and target domain. *Journal of Climate* **18**(13):2308–2329.
Shashkin A, Vaganov E. 1993. Simulation-model of climatically determined variability of conifers annual increment (on the example of common pine in the steppe zone). *Russian Journal of Ecology* **24**(5):275–280.
Smerdon JE. 2012. Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *Wiley Interdisciplinary Reviews: Climate Change* **3**(1):63–77.
Steig EJ, Schneider DP, Rutherford SD, Mann ME, Comiso JC, Shindell DT. 2009. Warming of the Antarctic ice-sheet surface since the 1957 international geophysical year. *Nature* **457**(7228):459–462.
Stocker T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex B, Midgley B. 2013. IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.
Tingley MP, Craigmile PF, Haran M, Li B, Mannshardt E, Rajaratnam B. 2012. Piecing together the past: Statistical insights into paleoclimatic reconstructions. *Quaternary Science Reviews* **35**:1–22.
Tingley MP, Huybers P. 2010a. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: development and applications to paleoclimate reconstruction problems. *Journal of Climate* **23**(10):2759–2781.
Tingley MP, Huybers P. 2010b. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: comparison with the regularized expectation-maximization algorithm. *Journal of Climate* **23**(10):2782–2800.
Tolwinski-Ward S, Tingley M, Evans M, Hughes M, Nychka D. 2014. Probabilistic reconstructions of local temperature and soil moisture from tree-ring data with potentially time-varying climatic response. *Climate Dynamics* **44**(3-4):791–806.
Tolwinski-Ward SE, Anchukaitis KJ, Evans MN. 2013. Bayesian parameter estimation and interpretation for an intermediate model of tree-ring width. *Climate of the Past* **9**(4):1481–1493.
Tolwinski-Ward SE, Evans MN, Hughes MK, Anchukaitis KJ. 2011. An efficient forward model of the climate controls on interannual variation in tree-ring width. *Climate Dynamics* **36**(11-12):2419–2439.
Vaganov EA, Hughes MK, Shashkin AV. 2006. *Growth Dynamics of Conifer Tree Rings: Images of Past and Future Environments*, Vol. 183. Springer: Heidelberg, Germany.
Werner JP, Tingley MP. 2015. Technical note: probabilistically constraining proxy age–depth models within a Bayesian hierarchical reconstruction model. *Climate of the Past* **11**(3):533–545.
Zhang Z, Mann ME, Cook ER. 2004. Alternative methods of proxy-based climate field reconstruction: application to summer drought over the conterminous United States back to AD 1700 from tree-ring data. *The Holocene* **14**(4):502–516.

SUPPORTING INFORMATION

Supporting information including code can be found at GitHub at http://github.com//Mechanistic-Tree-Ring

