Supplementary materials for this article are available at 10.1007/s13253-015-0210-9.

Stochastic Modeling for Velocity of Climate Change

Erin M. SCHLIEP, Alan E. GELFAND, and James S. CLARK

The velocity of climate change is defined as an instantaneous rate of change needed to maintain a constant climate. It is developed as the ratio of the temporal gradient of climate change over the spatial gradient of climate change. Ecologically, understanding these rates of climate change is critical since the range limits of plants and animals are changing in response to climate change. Additionally, species respond differently to changes in climate due to varying tolerances and adaptability. A fully stochastic hierarchical model is proposed that incorporates the inherent relationship between climate, time, and space. Space-time processes are employed to capture the spatial correlation in both the climate variable and the rate of change in climate over time. Directional derivative processes yield spatial and temporal gradients and, thus, the resulting velocities for a climate variable. The gradients and velocities can be obtained at any location in any direction and any time. In fact, maximum gradients and their directions can be obtained, hence minimum velocities. Explicit parametric forms for the directional derivative processes provide full inference on the gradients and velocities including estimates of uncertainty. The model is applied to annual average temperature across the eastern United States for the years 1963– 2012. Maps of the spatial and temporal gradients are produced as well as velocities of temperature change.

Supplementary materials accompanying this paper appear on-line.

Key Words: Climate velocity; Directional derivatives; Gaussian spatial processes.

1. INTRODUCTION

The concept of climate velocity has emerged as a useful index for evaluating the migration rates that might be required for populations to track changing climate (Loarie et al. 2009). For a single environmental variable, say, temperature, this enables the notion of *velocity of climate change*. The basic idea is to formulate an index of velocity of temperature change in km/year over a large spatial region. It is developed from spatial gradients of temperature change in °C/km and temporal rates of temperature change in °C/year. Dividing the latter

Erin M. Schliep (✉), Alan E. Gelfand and James S. Clark Department of Statistical Science, Duke University, Durham, NC, USA (E-mail: *erin.schliep@duke.edu*). James. S. Clark Nicholas School of the Environment, Duke University, Durham, NC, USA.

© 2015 International Biometric Society
Journal of Agricultural, Biological, and Environmental Statistics, Volume 20, Number 3, Pages 323–342
DOI: 10.1007/s13253-015-0210-9

by the former produces a *velocity*. We interpret the ratio as the instantaneous local velocity along the earth's surface needed to maintain constant temperature (Loarie et al. 2009). For species with small tolerances, the velocity estimates closely approximate migration speeds required to potentially avoid extinction. In this paper we introduce model-based inference, providing a formal basis for climate velocity inference and prediction. Complex topography is able to provide a spatial buffer for climate change (Peterson et al. 1997). Therefore, locations with large elevation gradients have small velocities. We show how the uncertainty in climate velocity varies depending on the spatial and temporal gradients. That is, uncertainty in the estimate of velocity scales with velocity; locations with large velocities have large uncertainties. These locations tend to have little change in topography and thus, a minimal spatial buffer for climate change.

Impacts of climate change depend heavily on the relationship between the spatial heterogeneity and temporal change in climate. Species survival depends not only on the ability to keep pace with a changing climate but also on the persistence of a suitable climate. Spatial heterogeneity under a changing climate creates climate sources and sinks, which create inaccessible regions to migrants tracking changing climates. Climate sources are areas where locally novel climates are disconnected from areas where similar climates previously occurred (Burrows et al. 2014). Sinks, on the other hand, are locations where climate conditions disappear. The continuity and diversity of habitats within a mountainous landscape provide a defense against the effects of climate change and facilitate species persistence through time (Peterson et al. 1997). Landcover changes also play a significant role in species movement. Protected areas where landscapes are less fragmented provide more suitable conditions for species to keep pace with a changing climate (Loarie et al. 2009). In fact, large protected areas can mitigate large velocities in some biomes, stressing the importance of accurate estimates of climate velocities enabling informed land management strategies.

To capture the relationship between spatial heterogeneity and temporal change in climate and to obtain formal inference on velocity, one needs a stochastic model that encompasses both spatial and temporal processes. The processes must be differentiable in order to obtain instantaneous spatial and temporal gradients and velocities. Additionally, a stochastic model allows for assessment of uncertainty in the processes, which propagates through to uncertainty in the gradients and resulting velocities.

Thus far, applications in the literature are limited to simple ratios of empirical gradients in space and time (Loarie et al. 2009; Burrows et al. 2011, 2014; Dobrowski et al. 2013). From a process perspective, the work is purely descriptive and offers no explicit modeling of the process. It therefore omits the joint linkage between temperature, time, and space. There is no error structure, no likelihood, and thus, no formal inference. To illustrate, consider temperature $\equiv T = f(t)$ where t is time, i.e., a general relationship capturing the change in temperature across time. Suppose temperature also varies with latitude, y, on a continental scale such that $T = g(y)$. In these approaches, calculations of $\frac{\partial T}{\partial t}$ and $\frac{\partial T}{\partial y}$ are not instantaneous. Rather, they are done with finite differences as the cartoon in Fig. 1 suggests.

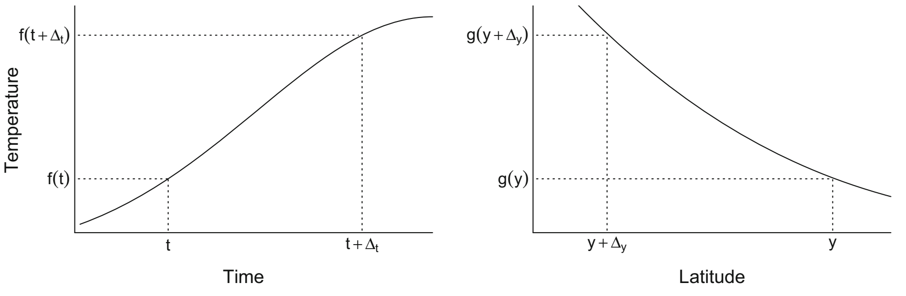

Figure 1. A cartoon of the naive velocity calculation.

That is, velocity, vel, is expressed as

$$\text{vel} = \frac{(f(t+\Delta t) - f(t))/\Delta t}{(g(y+\Delta y) - g(y))/\Delta y} = \frac{\Delta y}{\Delta t}\frac{(f(t+\Delta t) - f(t))}{(g(y+\Delta y) - g(y))}. \tag{1}$$

The fraction f/g is 1, i.e., the common ΔT, as the figure shows. Therefore Δy is determined for a common ΔT by aligning a given y with a given t and Δt. Here, Δy is the change in latitude (northings) needed to provide the change in temperature that arises as time changes from t to $t + \Delta t$. Due to the finite difference calculations done with geographic information systems (GIS) software, the foregoing papers find many ΔT's to be essentially 0 and make an arbitrary correction to obtain finite velocities. Moreover, the discretization schemes introduce additional uncertainty that will depend on scale. Lastly, ad hoc measures of uncertainty arise only from variability when an ensemble of climate scenarios is considered rather than from model mis-specification and measurement error.

The contribution of this paper is to cast the development of velocity in a fully stochastic framework. A model-based approach allows us to calculate $\frac{\partial T}{\partial t}$ and $\frac{\partial T}{\partial y}$. In modeling temperature as a function of space and time, possibilities are considerable (e.g., Craigmile and Guttorp 2011; Benth et al. 2007). However, the ability to explicitly calculate an enormous number of instantaneous derivatives limits the scope of computationally feasible models. We recognize that, at the least, we should express temperature, T, as a function of both location and time, say $T(x, y, t)$ where x denotes longitude, projected to eastings in kilometers, y denotes latitude, projected to northings in kilometers, and t denotes time. We propose a rich model for $T(x, y, t)$, accounting for elevation and incorporating spatial structure, anticipating that gradients, hence velocities, at close locations should be similar. Further, the model proves full inference; we obtain estimates of uncertainty in the processes, as well as the gradients and velocities.

A sufficiently smooth functional form for T legitimizes the notion of instantaneous velocity. In particular, the expected temperature surface, $\mathrm{E}(T(x, y, t))$, is viewed as random, or rather, as a realization of a stochastic process where process realizations are mean-square differentiable (Banerjee and Gelfand 2003). This enables us to take partial derivatives of *expected* temperatures. We model $\mathrm{E}(T(x, y, t))$ coherently and within a Bayesian framework to obtain full inference. Uncertainty in the surface yields uncertainty in the gradients, hence in the velocities. In particular, we calculate infinitesimal derivatives through a parametric specification for $\mathrm{E}(T(x, y, t))$ rather than the descriptive finite differences (1) obtained using geographic information systems (GIS) software (eight neighbor slope and aspect calculation) (Burrough and McDonnell 1998). All inference on gradients and velocities is obtained post-model fitting. Temperature gradients can be obtained at any time and location and spatial gradients at any time and in any direction. Therefore, velocities can be obtained at any location in any direction and any time. Ecologically, the direction and magnitude of the minimum velocity is meaningful as it captures what will be required for survival at that location and time. The direction of maximum velocity is not meaningful mathematically nor ecologically as there will be a direction where the spatial gradient of temperature is essentially 0, yielding essentially infinite velocity in that direction. Lastly, note that in Fig. 1, $\Delta y < 0$ so vel < 0. According to the figure, this is sensible; we need a negative change in latitude to provide a match with a positive change in time. In general, for a given location and time, both the temporal and spatial gradient can be increasing or decreasing, which means velocity can be either positive or negative.

We model climate velocity using annual temperature data for the eastern United States spanning 50 years on 7.5 min grid cells. Elevation is included to capture localized directional behavior of the spatial surface. We formulate a hierarchical model that incorporates both sources of data and measurement error in temperature. The model is fitted introducing three Gaussian processes to capture spatial dependence, two for the functional relationship for temperature given time and one for elevation. Directional derivative processes are employed to produce needed derivatives. There are two potential paths to implement the differentiation. The first is through formal directional derivative processes, i.e., limits of directional differences of a realization of a stochastic process (Banerjee et al. 2003). The other is through dimension reduction, here using the predictive process (Banerjee et al. 2008; Finley et al. 2009) which enables explicit parametric representations of the process, hence explicit differentiation. For computational reasons, we adopt the latter.

The format of the paper is as follows. In Sect. 2 we describe the data used to develop temperature velocities. We present the hierarchical model in Sect. 3. Section 4 details the calculations for both spatial and temporal gradients as well as velocities. In Sect. 5 we present the results of the data analysis, showing the broad range of inference available. We conclude with a summary and description of future work in Sect. 6.

2. THE DATA

We apply the multivariate predictive process model developed in the next section to temperature data for the eastern United States. The temperature data are from the parameter-elevation regression on independent slopes model (PRISM). The data are the annual average temperature (°C) for the period 1963–2012. The PRISM data are on 2.5 arc-min resolution, which we aggregate to 7.5 arc-min, or 1/8 degree resolution. We use the centers of these grid boxes as our observed locations and the annual average temperatures as the observations. While PRISM data have the benefit of large spatial coverage and high resolution, these

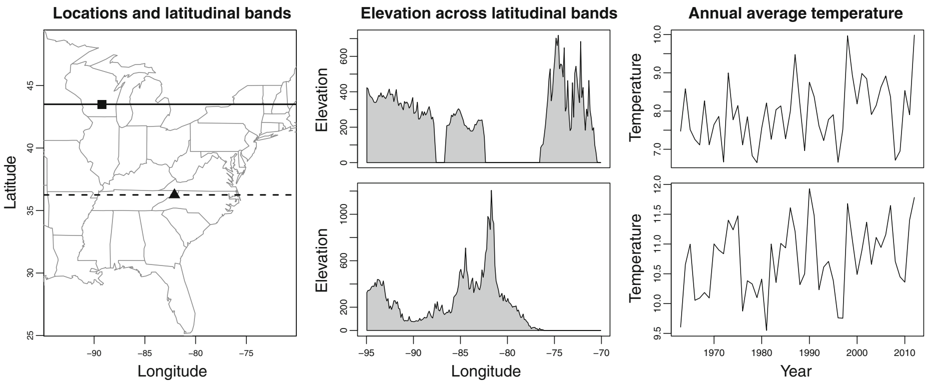

Figure 2. (*left*) Two locations chosen that vary geographically and topographically with latitudinal bands imposed, (*middle*) elevation across the latitudinal bands from north to south, and (*right*) annual average temperature trends at the locations from north to south.

data also have limitations. For example, PRISM is an interpolated data product generated from a model that incorporates multiple sources of information, including elevation. We acknowledge the possibility of bias in the data, especially at high resolutions, which may affect estimation of gradients and velocities. Nonetheless, our dataset for the eastern United States consists of 21,202 spatial locations and, to our knowledge, there is no better temperature data source available at such a large spatial scale. At a finer spatial scale, we could explore an analysis using monitoring station data and we include this as a possible direction of future work.

The elevation data are from the ETOPO1 dataset, which is a 1 arc-min global model of the earth's surface (Amante and Eakins 2009). We obtain the elevation at each of the observed temperature locations. Figure 2 shows two latitudinal bands with one location chosen on each band. Elevations across each band over the eastern United States are shown along with time series of annual average temperatures over the 50 years for the two locations. We see considerable variation in elevation as a result of the Appalachian mountain range that spans from southwest to northeast. We also see variation in annual temperature.

We use the Albers equal-area conic projection to transform the longitudinal and latitudinal coordinates of our observed locations to Albers coordinates employing the standard parallels of 29.5° and 45°. These parallels are commonly used to depict the United States. The projection yields eastings and northings in standard units (1 Albers unit $\approx$ 6376 km). Under this projection, we compute all distances as Euclidean distances. We avoid working with distances on the earth's surface due to the trigonometry associated with such distances. The representations in terms of trigonometric functions make the derivative calculations below considerably more messy and lead to ill-behaved model fitting. For instantaneous derivatives, we anticipate that the effects of the projection on the resultant posterior distributions for velocities at a given location and time will be minimal. In fact, Loarie et al. (2009) and Ordonez and Williams (2013) suggest velocities are most sensitive to spatial resolution, which the Albers equal-area conic projection preserves.

3. MODEL

We model annual average temperature using a linear mixed model with spatially correlated random effects. The model is inherently hierarchical as it combines two sources of data, annual average temperature and elevation. Let $T(x, y, t)$ be the annual average temperature for location (x, y) at time t where x is the eastings coordinate and y is the northings coordinate. Further, let $E(x, y)$ be elevation at location (x, y). We model $T(x, y, t)$ and $E(x, y)$ as

$$T(x, y, t) = \beta_0 + \beta_1 t + \beta_2 y + \beta_3 Z(x, y) + \alpha_0(x, y) + \alpha_1(x, y)t + \epsilon(x, y, t) \quad (2a)$$

$$E(x, y) = \mu + Z(x, y) + \eta(x, y), \quad (2b)$$

where $\epsilon(x, y, t) \sim N(0, \sigma_T^2)$ and $\eta(x, y) \sim N(0, \sigma_E^2)$. The coefficients β_0, β_1, β_2, and β_3 are global coefficients for the intercept, and the rate of change in annual temperature over time, latitude, and elevation, respectively. Both $\alpha_0(x, y)$ and $\alpha_1(x, y)$ in (2a) are spatial random effects that account for the local remaining spatial variation in annual average temperature and rate of change in annual temperature over time that is not accounted for by latitude and elevation. In other words, we have a spatially varying intercept and slope in our regression on temperature (Gelfand et al. 2003). The spatial random effect, $Z(x, y)$, in (2b) accounts for the spatial variation in elevation and enters as a covariate not connected to t in (2a). Complex topography is able to provide a spatial buffer for climate change, motivating the need for both gradients and velocities to be functions of elevation. The three spatial random effects, $\alpha_0(x, y)$, $\alpha_1(x, y)$, and $Z(x, y)$ must be sufficiently smooth differentiable surfaces. As such, these processes allow for gradients and velocities to be calculable at arbitrary locations. Note that the specification of $E(x, y)$ does not assume the observed elevation surface to be smooth, but rather, specifies the "centering" surface to be smooth. Apart from the fact that we need smoothness for differentiability, our focus is on continental-scale velocity. We expect that a flexible, smooth process will capture the elevation surface at a large spatial scale reasonably well. At a finer spatial resolution, a trend surface may be more appropriate for the elevation surface and we include this investigation as a direction for future work. Lastly, we do not introduce a longitudinal term in (2a) since elevation captures the majority of the longitudinal variability of temperature in our study region.

As noted in the introduction, there are two approaches for modeling the three spatial processes, $\alpha_0(x, y)$, $\alpha_1(x, y)$, and $Z(x, y)$. One is to model them as customary Gaussian processes, possibly dependent. Here, we would adopt a covariance function or cross-covariance function such that process realizations are mean-square differentiable, e.g., Matérn with $\nu \geq 1$ (Stein 1999). The spatial gradients would then be calculated as developed in (Banerjee et al. 2003). A second approach is to model them using dimension reduction, i.e., to express the surfaces as parametric linear transformations of a finite set of random variables at fixed locations. With suitably differentiable functions, such a representation enables explicit gradient calculation. We adopt the latter approach due to computational necessity as we consider temperature at more than 21,000 gridded locations. We use the predictive process for dimension reduction for each of the three spatial processes (Banerjee et al. 2008). Anticipating customary association between slope and intercept, we employ coregionalization to connect slope and intercept processes in the mean temperature model. The latent elevation process is treated as independent of these two processes.

3.1. THE PREDICTIVE PROCESS AS USED HERE

In the interest of dimension reduction, each of the foregoing spatial processes is specified through a predictive process. For the spatial process for elevation, let $\mathbf{Z} = [Z(x_1, y_1), \ldots, Z(x_n, y_n)]'$ where (x_i, y_i), $i = 1, \ldots, n$ are the observed locations. The predictive process is defined by specifying $\mathbf{Z}$ through projection onto a lower dimensional subspace. The subspace is generated by realizations of $\mathbf{Z}$ at a set of locations, referred to as "knots." Let $\mathbf{Z}^* = [Z(x_1^*, y_1^*), \ldots, Z(x_m^*, y_m^*)]'$ be the spatial process in lower dimension where (x_j^*, y_j^*), $j = 1, \ldots, m$ are the knot locations, and $m << n$. Banerjee et al. (2008) and Finley et al. (2009) offer some discussion regarding knot selection and spatial design for knot selection. Over our large spatial region, we adopted a simple geometric space filling design with equally spaced knots. Here, $\mathbf{Z}^* \sim \mathrm{GP}(\mathbf{0}, \mathbf{C}_{\mathbf{Z}^*})$. The predictive process, denoted $\widetilde{\mathbf{Z}}$, is the spatial interpolant of the process given the reduced-dimension process, $\mathbf{Z}^*$, and is induced by the original Gaussian process of interest. Realizations of the predictive process are obtained using the "kriging" equation, or, rather as the conditional mean given the values of the process at the knot locations. That is, $\widetilde{\mathbf{Z}} = \mathrm{E}(\mathbf{Z}|\mathbf{Z}^*) = \mathbf{C}'_{\mathbf{Z},\mathbf{Z}^*}(\mathbf{C}_{\mathbf{Z}^*})^{-1}\mathbf{Z}^*$ where $\mathbf{C}'_{\mathbf{Z},\mathbf{Z}^*}$ is an $n \times m$ matrix with (i, j)th element equal to the covariance between $Z(x_i, y_i)$ and $Z(x_j^*, y_j^*)$ and $\mathbf{C}_{\mathbf{Z}^*}$ is the $m \times m$ covariance matrix of $\mathbf{Z}^*$ with (i, j)th element equal to the covariance between $Z(x_i^*, y_i^*)$ and $Z(x_j^*, y_j^*)$. This implies that the predictive process, $\widetilde{\mathbf{Z}}$, is also a mean zero Gaussian process with covariance $\mathbf{C}'_{\mathbf{Z},\mathbf{Z}^*}(\mathbf{C}_{\mathbf{Z}^*})^{-1}\mathbf{C}_{\mathbf{Z},\mathbf{Z}^*}$.

We specify the covariance function for the GP as Matérn with smoothness parameter $\nu = 3/2$. Explicitly, the covariance between $Z(x_i, y_i)$ and $Z(x_j, y_j)$, for any two locations (x_i, y_i) and (x_j, y_j) is

$$\mathrm{Cov}(Z(x_i, y_i), Z(x_j, y_j)) = \tau_{\mathrm{Z}}^2 \left(1 + \phi_{\mathrm{Z}} d_{ij}\right) \exp(-\phi_{\mathrm{Z}} d_{ij}), \tag{3}$$

where d_{ij} is the distance between locations (x_i, y_i) and (x_j, y_j), ϕ_{Z} is the spatial decay parameter, and τ_{Z}^2 is the spatial variance parameter.

The spatial processes $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_1$ provide an extremely flexible model for capturing locally-linear relationship between annual average temperature, T, and time, t. Dependence between the slope and intercept processes is anticipated (Berrocal et al. 2012) so $\boldsymbol{\alpha}_0 = [\alpha_0(x_1, y_1), \ldots, \alpha_0(x_n, y_n)]'$ and $\boldsymbol{\alpha}_1 = [\alpha_1(x_1, y_1), \ldots, \alpha_1(x_n, y_n)]'$ are modeled with a bivariate predictive process using a linear model of coregionalization. The coregionalization is defined as

$$\begin{bmatrix} \boldsymbol{\alpha}_0 \\ \boldsymbol{\alpha}_1 \end{bmatrix} = [\mathbf{A} \otimes \boldsymbol{I}] \begin{bmatrix} \mathbf{W}_0 \\ \mathbf{W}_1 \end{bmatrix}. \tag{4}$$

Here, $\mathbf{W}_k = [W_k(x_1, y_1), \ldots, W_k(x_n, y_n)]'$ for $k = 0, 1$ where $\mathbf{W}_0$ and $\mathbf{W}_1$ are independent mean zero Gaussian processes with unit variance and decay parameters, ϕ_0 and ϕ_1, respectively, and

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$$

is a 2×2 lower triangular matrix with non-negative diagonal elements.

We employ predictive processes for both $\mathbf{W}_0$ and $\mathbf{W}_1$. Letting $\mathbf{W}_k^* = [W_k(x_1^*, y_1^*), \ldots, W_k(x_m^*, y_m^*)]'$ for $k = 0, 1$, then

$$\begin{aligned} \widetilde{\mathbf{W}}_k &= \mathbf{C}_k'(\mathbf{C}_k^*)^{-1}\mathbf{W}_k^* \\ \mathbf{W}_k^* &\sim \mathrm{GP}(\mathbf{0}, \mathbf{C}_k^*), \end{aligned} \tag{5}$$

where $\mathbf{C}_k'$ is the $n \times m$ cross-covariance matrix between $\mathbf{W}_k$ and $\mathbf{W}_k^*$ and $\mathbf{C}_k^*$ is the covariance matrix of $\mathbf{W}_k^*$. Again, we model covariance using the Matérn covariance function with smoothness parameter equal $\nu = 3/2$, decay parameter, ϕ_k, and scale parameter τ_k^2 fixed to 1 to enable identifiability of $\mathbf{A}$. Using $\widetilde{\mathbf{W}}_0$ and $\widetilde{\mathbf{W}}_1$ in (5), we obtain $[\widetilde{\boldsymbol{\alpha}}_0', \widetilde{\boldsymbol{\alpha}}_1']'$ by setting

$$\begin{bmatrix} \widetilde{\boldsymbol{\alpha}}_0 \\ \widetilde{\boldsymbol{\alpha}}_1 \end{bmatrix} = [\mathbf{A} \otimes \boldsymbol{I}] \begin{bmatrix} \widetilde{\mathbf{W}}_0 \\ \widetilde{\mathbf{W}}_1 \end{bmatrix}.$$

Both $\widetilde{\mathbf{W}}_0$ and $\widetilde{\mathbf{W}}_1$ are mean zero Gaussian processes which implies that both $\widetilde{\boldsymbol{\alpha}}_0$ and $\widetilde{\boldsymbol{\alpha}}_1$ are also mean zero.

Finally, using the predictive processes we model $T(x, y, t)$ and $E(x, y)$ as

$$\begin{aligned} T(x, y, t) &= \beta_0 + \beta_1 t + \beta_2 y + \beta_3 \widetilde{Z}(x, y) + \widetilde{\alpha}_0(x, y) + \widetilde{\alpha}_1(x, y)t + \epsilon(x, y, t) \\ E(x, y) &= \mu + \widetilde{Z}(x, y) + \eta(x, y). \end{aligned} \tag{6}$$

In Sect. 4 we show that the forms in (6) enable explicit calculation of gradients and velocities.

3.2. Bayesian Details

Inference is obtained for the parameters of the hierarchical model within the Bayesian framework. Let $\mathbf{T} = [\mathbf{T}_1, \ldots, \mathbf{T}_n]$ where $\mathbf{T}_t$ is the vector of observed annual average temperature for year t for all n locations such that $\mathbf{T}_t = [T(x_1, y_1, t), \ldots, T(x_n, y_n, t)]'$. Similarly, let $\mathbf{E} = [E(x_1, y_1), \ldots, E(x_n, y_n)]'$ be the vector of observed elevations and $\boldsymbol{y} = [y_1, \ldots, y_n]'$ be the vector of latitudes. Lastly, let $\mathbf{W}^* = [\mathbf{W}_0^{*'}, \mathbf{W}_1^{*'}]'$. Then, the full posterior distribution of the parameters given the data is

$$\begin{aligned} \pi(\boldsymbol{\beta}, \mathbf{Z}^*, \mathbf{W}^*, \mathbf{A}, \phi_0, \phi_1, \sigma_\mathrm{T}^2, \mu, \tau_\mathrm{Z}^2, \phi_\mathrm{Z}, \sigma_\mathrm{E}^2 | \mathbf{T}, \mathbf{E}) \propto & \prod_{t=1}^{T} \pi(\mathbf{T}_t | \boldsymbol{\beta}, \mathbf{Z}^*, \mathbf{W}^*, \mathbf{A}, \phi_0, \phi_1, \sigma_\mathrm{T}^2, \tau_\mathrm{Z}^2, \phi_\mathrm{Z}) \\ & \times \pi(\mathbf{E} | \mu, \mathbf{Z}^*, \tau_\mathrm{Z}^2, \phi_\mathrm{Z}, \sigma_\mathrm{E}^2) \\ & \times \pi(\mathbf{W}^* | \phi_0, \phi_1) \\ & \times \pi(\mathbf{Z}^* | \phi_\mathrm{Z}, \tau_\mathrm{Z}^2) \\ & \times \pi(\boldsymbol{\beta}, \mathbf{A}, \phi_0, \phi_1, \sigma_\mathrm{T}^2, \mu, \tau_\mathrm{Z}^2, \phi_\mathrm{Z}, \sigma_\mathrm{E}^2). \end{aligned} \tag{7}$$

We assign prior distributions to all hyperparameters. Non-informative conjugate priors distributions are used when possible. The coefficients β_k for $k = 0, \ldots, 3$ are assigned independent mean zero normal prior distributions with variance 10^4. The diagonal elements of the $\mathbf{A}$ matrix, a_{11} and a_{22} are assigned independent, diffuse half-normal prior distributions with parameter $\theta = 0.01$, where the mean is $1/\theta$ and variance is $(\pi - 2)/2\theta^2$. The off-diagonal element, a_{21} is assigned a diffuse mean zero normal prior with variance 10^4. The variance, σ_T^2, of the error term for the temperature model, $\epsilon(x, y, t)$, has a non-informative conjugate inverse gamma (2,2) prior distribution.

The parameters of the elevation model, μ and τ_Z^2, are also assigned conjugate non-informative priors. The mean parameter, μ, has a mean zero normal prior distributions with variance 10^4 and τ_Z^2 is again, conjugate inverse gamma (2,2). The variance of the elevation model, σ_E^2, can also be assigned a conjugate inverse gamma distribution. However, we fix $\sigma_E^2 = 10$ since the vertical accuracy of the dataset is assumed to be $\pm$ 10 m (Amante and Eakins 2009).

A uniform prior distribution is assigned to each of the spatial decay parameters, ϕ_Z, ϕ_0, and ϕ_1. Grid cell coordinates are given by eastings and northings under the Albers projection. The maximum distance between locations is 0.4655 in Albers standard units ($\approx$2970 km) and the minimum latitudinal distance is 0.0022 ($\approx$13.9 km). The values of the uniform prior distributions for the decay parameters are such that the range is less than 1/2 the max distance and greater than the minimum latitudinal distance between locations. We can do this easily by noting that, for the Matérn covariance function with smoothness $\nu = 3/2$, the relationship between the decay parameter and the range is $\phi \approx 4.7/d$. The predictive process was fitted to 532 knots, gridded across the eastern US at roughly 0.0118 (75 km) resolution.

4. CALCULATING GRADIENTS AND VELOCITIES

Here, we turn to the explicit calculations of the needed gradients. For the annual temperature model, the temporal gradient of temperature change is the *expected* change in temperature per year, whereas the spatial gradient is the *expected* change in temperature per kilometer. A spatial gradient can be calculated in an arbitrary direction and time. Attractively, the gradient in any direction can be calculated from the gradient in the easting direction, $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial x}$, and in the northing direction, $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial y}$. Let

$$\nabla \mathrm{E}(T(x, y, t)) = \begin{bmatrix} \frac{\partial \mathrm{E}(T(x,y,t))}{\partial x} \\ \frac{\partial \mathrm{E}(T(x,y,t))}{\partial y} \end{bmatrix}. \tag{8}$$

Then, the gradient in the direction $\mathbf{u}$, where $\mathbf{u} = [u_x, u_y]'$ is a unit vector, is $\nabla \mathrm{E}(T(x, y, t))'\mathbf{u}$. Evidently, the gradient in direction $-\mathbf{u}$ is the negative of the gradient in direction $\mathbf{u}$ but the magnitudes will be the same. Furthermore, the direction of maximum gradient is

$$\nabla \mathrm{E}(T(x, y, t))/||\nabla \mathrm{E}(T(x, y, t))||,$$

where $||\nabla \mathrm{E}(T(x, y, t))||$ is the magnitude of the maximum gradient.

Let us develop the details of the gradients in terms of model (6). Returning to the predictive processes, let $\mathbf{P}^*$, $\mathbf{Q}^*$, and $\mathbf{R}^*$ each be $m \times m$ correlation matrices of $\mathbf{Z}^*$, $\mathbf{W}_0^*$, and $\mathbf{W}_1^*$, respectively, i.e., $P_{jk}^* = \rho(d_{jk}; \phi_Z)$, $Q_{jk}^* = \rho(d_{jk}; \phi_0)$, $R_{jk}^* = \rho(d_{jk}; \phi_1)$. Again, ϕ_Z, ϕ_0, and ϕ_1 are decay parameters of the Matérn correlation function with $\nu = 3/2$ and $j, k = 1, \dots, m$. Further, define the $m \times 1$ correlation vectors $\mathbf{p}(x, y)$, $\mathbf{q}(x, y)$, and $\mathbf{r}(x, y)$ where the jth element of $\mathbf{p}(x, y)$ is the correlation between $Z(x, y)$ and $Z(x_j^*, y_j^*)$, similarly for $\mathbf{q}(x, y)$ and $\mathbf{r}(x, y)$.

From (6), the expected annual average temperature at location (x, y) and time t is

$$\begin{aligned}\mathrm{E}(T(x, y, t)) &= \beta_0 + \beta_1 t + \beta_2 y + \beta_3 \widetilde{Z}(x, y) + \widetilde{\alpha}_0(x, y) + \widetilde{\alpha}_1(x, y)t \\ &= \beta_0 + \beta_1 t + \beta_2 y + \beta_3 \widetilde{Z}(x, y) + [1 \quad t]\,\mathbf{A} \begin{bmatrix} \widetilde{W}_0(x, y) \\ \widetilde{W}_1(x, y). \end{bmatrix}\end{aligned} \tag{9}$$

The spatial and temporal gradients at location (x, y) and time t are computed as the derivative of the $\mathrm{E}(T(x, y, t))$ with respect to x for the eastern direction, y for the northern direction, and t for time. That is, $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial x}$ is the spatial gradient in the x direction, $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial y}$ is the spatial gradient in the y direction, and $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial t}$ is the gradient in time.

The temporal gradient is

$$\frac{\partial \mathrm{E}(T(x, y, t))}{\partial t} = \beta_1 + a_{21}\, \mathbf{q}(x, y)' \mathbf{Q}^{*-1} \mathbf{W}_0^* + a_{22}\, \mathbf{r}(x, y)' \mathbf{R}^{*-1} \mathbf{W}_1^*, \tag{10}$$

which is a spatial Gaussian process as it arises as the sum of two independent Gaussian predictive processes. In fact, its covariance structure can be obtained explicitly; we omit details. Moreover, it is free of t due to the linearity in (9), so we can display a single posterior mean surface for this process.

To obtain the spatial gradients, first write the derivatives of $\mathbf{p}_j(x, y)$, $\mathbf{q}_j(x, y)$, and $\mathbf{r}_j(x, y)$ with respect to x and y. That is, the derivative of $\mathbf{p}_j(x, y)$ with respect to x is written as

$$\begin{aligned}\frac{\partial \mathbf{p}_j(x, y)}{\partial x} &= \frac{\partial}{\partial x} \rho((x, y), (x_j^*, y_j^*); \phi_Z) \\ &= -\,\phi_Z^2 (x - x_j^*) e^{-\phi_Z \sqrt{(x - x_j^*)^2 + (y - y_j^*)^2}}.\end{aligned}$$

Similarly, the derivative with respect to y is

$$\frac{\partial \mathbf{p}_j(x, y)}{\partial y} = -\,\phi_Z^2 (y - y_j^*) e^{-\phi_Z \sqrt{(x - x_j^*)^2 + (y - y_j^*)^2}}.$$

The derivatives $\frac{\partial \mathbf{q}_j(x,y)}{\partial x}$, $\frac{\partial \mathbf{q}_j(x,y)}{\partial y}$, $\frac{\partial \mathbf{r}_j(x,y)}{\partial x}$, and $\frac{\partial \mathbf{r}_j(x,y)}{\partial y}$ can be obtained in the same fashion.

Then, the spatial gradients $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial x}$ and $\frac{\partial \mathrm{E}(T(x,y,t))}{\partial y}$ are computed as

$$\begin{aligned}\frac{\partial \mathrm{E}(T(x, y, t))}{\partial x} &= \beta_3 \frac{\partial}{\partial x}\mathbf{p}(x, y)'\mathbf{P}^{*-1}\mathbf{Z}^* + (a_{11} + a_{21}t)\frac{\partial}{\partial x}\mathbf{q}(x, y)'\mathbf{Q}^{*-1}\mathbf{W}_0^* \\ &\quad + a_{22}t\frac{\partial}{\partial x}\mathbf{r}(x, y)'\mathbf{R}^{*-1}\mathbf{W}_1^* \end{aligned} \tag{11}$$

and

$$\begin{aligned}\frac{\partial \mathrm{E}(T(x, y, t))}{\partial y} &= \beta_3 \frac{\partial}{\partial y}\mathbf{p}(x, y)'\mathbf{P}^{*-1}\mathbf{Z}^* + (a_{11} + a_{21}t)\frac{\partial}{\partial y}\mathbf{q}(x, y)'\mathbf{Q}^{*-1}\mathbf{W}_0^* \\ &\quad + a_{22}t\frac{\partial}{\partial y}\mathbf{r}(x, y)'\mathbf{R}^{*-1}\mathbf{W}_1^*. \end{aligned} \tag{12}$$

These quantities give the expected change in temperature per unit of distance in the x and y directions, respectively, and are linear in t. As with the temporal gradient, the spatial gradients are also spatial Gaussian processes, now each a sum of three predictive Gaussian processes. The cross-covariance structure can be calculated explicitly as well as the dependence between the spatial and temporal gradient processes. Again, we omit details. However, we can compute gradients in arbitrary directions from (11) and (12) as described above.

Finally, we can obtain velocities from the spatial and temporal gradients. A climate velocity for annual temperature is the ratio of the temporal gradient to the spatial gradient and is measured in km/year. The velocity in direction $\mathbf{u}$ is

$$\frac{\frac{\partial \mathrm{E}(T(x,y,t))}{\partial t}}{\nabla \mathrm{E}(T(x, y, t))'\mathbf{u}} = \frac{\frac{\partial \mathrm{E}(T(x,y,t))}{\partial t}}{\frac{\partial \mathrm{E}(T(x,y,t))}{\partial x}u_1 + \frac{\partial \mathrm{E}(T(x,y,t))}{\partial y}u_2}. \tag{13}$$

Clearly, in an absolute sense, the minimum velocity is the velocity in the direction of the maximum gradient and is

$$\frac{\frac{\partial \mathrm{E}(T(x,y,t))}{\partial t}}{||\nabla \mathrm{E}(T(x, y, t))||}.$$

As noted in the Introduction, we summarize only with minimum velocity, with justification both mathematically and ecologically. Of course, since the numerator can be negative, we will report both positive and negative velocities. As a last point, velocity in a given direction is also a spatial process. In fact, it is the ratio of predictive Gaussian processes. Such processes have been referred to as spatial Cauchy processes (See Terres (2014) for details).

In summary, large velocities indicate large changes in temperature over time relative to changes in temperature across space. Small velocities, on the other hand, indicate large changes in temperature across space relative to changes in time. Locations with a steep elevation gradient may see very small climate velocities since short distances lead to large changes in elevation, and thus, may result in large changes in temperature.

Table 1. The posterior median and 95 % credible interval for parameters.

Parameter	Median	95 % Credible interval
β_0	12.72	(12.68, 12.75)
β_1	0.022	(0.019, 0.023)
β_2	−0.862	(−0.866, −0.860)
β_3	−0.680[a]	(−0.690[a], −0.669[a])
$\mathbf{W}_0$ range	790.48[b]	(769.18[b], 814.32[b])
$\mathbf{W}_1$ range	199.21[b]	(192.04[b], 205.14[b])
a_{11}	1.636	(1.556, 1.713)
a_{21}	0.418	(0.381, 0.450)
a_{22}	0.092	(0.081, 0.100)
σ^2_{T}	0.457	(0.455, 0.458)
μ	105.49	(99.89, 112.73)
$\mathbf{Z}$ range	165.50[b]	(162.40[b], 168.69[b])
τ^2_{Z}	31,893	(30,593, 33,143)
σ^2_{E}	10	

[a] Denotes $\times 10^{-2}$.

[b] Denotes value given in kilometers.

5. RESULTS

We obtain posterior samples using Markov chain Monte Carlo (MCMC) and a hybrid Metropolis-within-Gibbs algorithm. The chains are run for 100,000 iterations and the first half are discarded as burn-in. Details on the MCMC sampling algorithm are available in the online supplementary material. See Banerjee et al. (2008) for additional information on Bayesian implementation and computational issues for multivariate predictive processes. Posterior estimates of the model parameters are given in Table 1. Overall, annual average temperature is increasing between 1963 and 2012 and decreasing with latitude and elevation. The estimates of the range of the three spatial processes, given in kilometers and functions of ϕ_0, ϕ_1, and ϕ_{Z}, indicate that the spatial correlation for the spatial process for elevation, $\mathbf{Z}$, operates at a much shorter distance than that for the spatial random intercept, $\boldsymbol{\alpha}_0$. The spatial correlation for the temporal gradient $\boldsymbol{\alpha}_1$ is a function of that of $\mathbf{W}_0$ and $\mathbf{W}_1$ under the linear model of coregionalization (4).

We present Figs. 10, 11 and 12 in the Appendix to offer some assessment of the models for temperature and elevation. In general, we should not expect to capture high resolution detail in annual average temperature and elevation due to the large spatial scale. Figure 10 gives the observed and fitted elevation surfaces indicating that our model is capturing the main features across the eastern United States. Quantiles of residuals of annual average temperature by year are shown in Fig. 11. Figure 12 shows the percent of under prediction of annual average temperature for each location across the region. Not surprisingly, the locations with the most consistent under and over prediction are primarily in the Appalachian Mountain region.

The posterior mean surface for the temporal gradient is given in Fig. 3. Between 1963 and 2012, the rate of temperature change appears greater in the northern part of the study region than in the south or west. The average increase in annual average temperature is

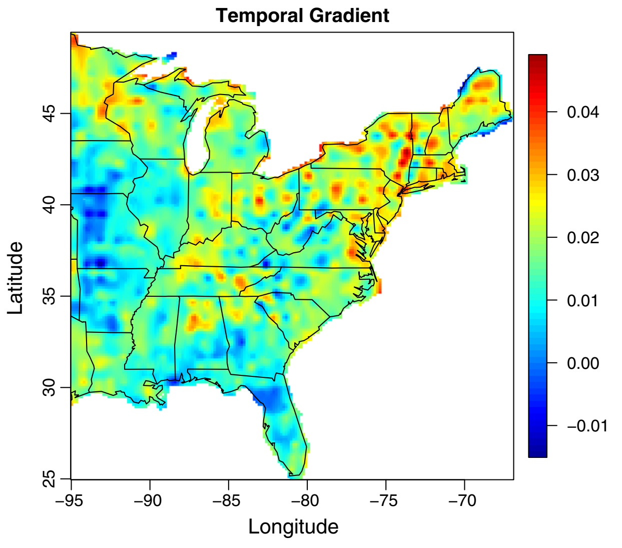

Figure 3. Posterior mean surface for the temporal gradient (°C/year).

0.016 °C/year, or 1.6 °C/century. The temporal gradient is significant at 97.67 % of the observed locations within the spatial domain where 96.68 % are significantly increasing and 0.89 % are significantly decreasing. Here, we consider an estimate "significant" if the 90 % credible interval does not include 0. We note that the credible intervals are determined marginally not simultaneously.

Across the eastern United States, the spatial gradient in the eastern direction is significant at 82.47 % of locations with 43.28 % positive and 39.18 % negative and at 89.38 % of locations in the northern direction with 15.05 % positive and 74.33 % negative (figures not shown). To illuminate spatial gradients, velocities, and uncertainties, we focus on the southeastern part of the United States. Figure 4 shows the posterior mean surface for the maximum spatial gradient. Locations along the Appalachian Mountains are seeing temperature changes as much as 0.1 °C/km. Figure 5 shows the posterior mean surface for elevation with arrows in the direction of the posterior mean maximum spatial gradient at a subset of locations within the region relative to the temporal gradient at the location. That is, under an increasing temperature trend, the arrow points in the direction one would need to travel in order to maintain a constant temperature, which is generally upwards and polewards. The maximum gradient for locations at higher elevations is influenced by the direction of maximum increase in elevation. It appears that elevation is the main driver for deviations from the northern trajectory of spatial gradients.

The posterior mean surface for velocity in the direction of the maximum spatial gradient for the year 2012 is given in Fig. 6. Velocities range from −2.07 to 21.12 km/year, with a median of 1.13 and 25th and 75th percentiles of 0.67 and 1.84 km/year, respectively. Note that while velocities can be negative or positive, for graphical purposes, velocities are shown as positive. As calculated here, a negative velocity is the result of a decrease

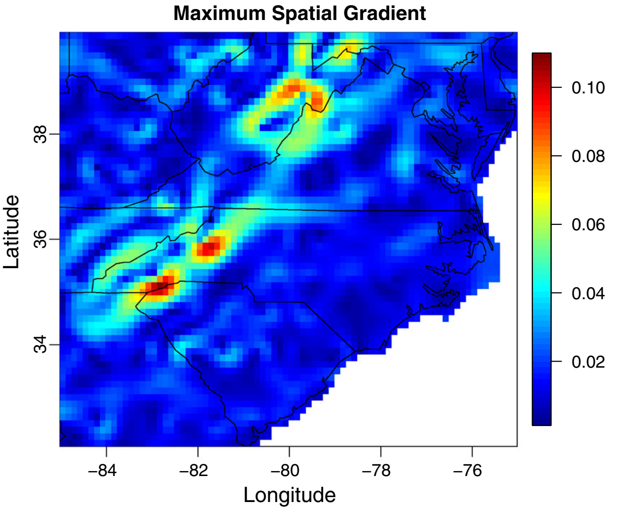

Figure 4. Posterior mean surface for the magnitude of the maximum spatial gradient (°C/km).

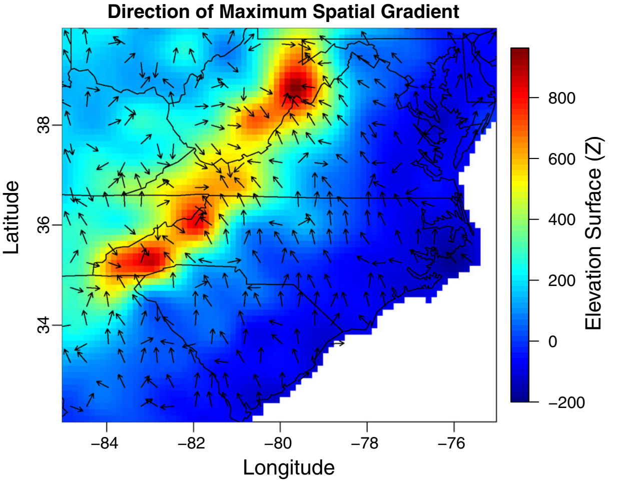

Figure 5. Posterior mean surface for the elevation process, **Z** with *arrows* pointing in the direction of maximum spatial gradient.

in temperature across time since the maximum spatial gradient is positive. Therefore, the resulting positive velocity is of the same magnitude as the negative velocity but in the opposite direction. Small velocities are more prevalent in mountainous regions, although they are also seen in locations where the temporal gradient is negative (See Fig. 3). Larger velocities are predominantly at locations with northern maximum gradients, and consequently, lower elevations or regions with little change in topography. Loarie et al. (2009) also report a strong correlation between topographic slope and the velocity of temperature change.

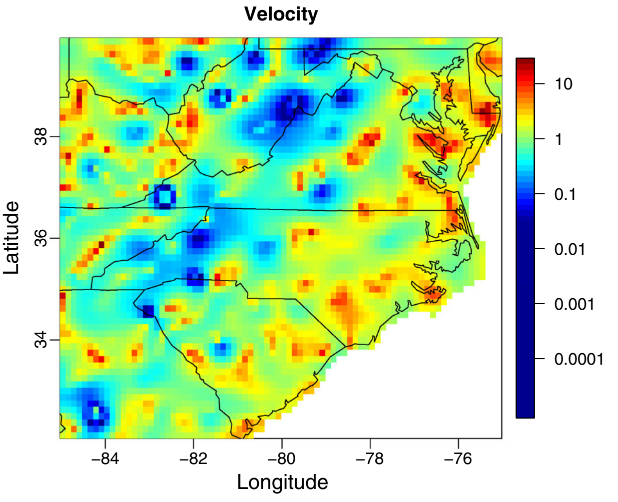

Figure 6. Posterior mean surface for velocity (km/year) in the direction of the maximum gradient evaluated for the year 2012.

As previously mentioned, the impacts of climate change depend heavily on the relationship between spatial heterogeneity and the temporal change in climate. In general, regions with smaller estimates of velocities that are in mountainous landscape are much more favorable in terms of species persistence through time. At lower elevations, such as locations within the Piedmont in North Carolina, our model yields velocity estimates greater than 1km/year, which may be faster than many estimates of plant migration. These maps, therefore, assist in highlighting extensive regions of heightened threat in terms of species survival. They can also be used to direct management and preservation efforts in establishing suitable, protected, and less fragmented landscapes for vulnerable species which have been shown to mitigate the effects of climate change by reducing the resistance in species distribution shifts.

Figure 7 shows the width of the 95 % credible interval of the posterior distribution for velocity in the direction of the maximum spatial gradient for 2012. The locations with large values of uncertainty correspond to locations with large velocities shown in Fig. 6. Locations with small velocities, specifically in the mountainous region, tend to have much less uncertainty. A time series of the posterior mean and 95 % credible interval for velocity at two illustrative locations is given in Fig. 8. The left figure is at a mountainous location and shows a slightly increasing trend in velocity over time with little change in variability. The right figure is at a non-mountainous location. Not only does this location report an increase in velocity over time but also an increase in the variability of velocity over time. This further illustrates that uncertainty in the estimates for velocity scale with velocity.

Lastly, Fig. 9 shows the posterior distribution for the directional velocities, both in magnitude and direction, for the final year in the time series for the same two locations. The length of the line depicts the magnitude of the change in temperature per kilome-

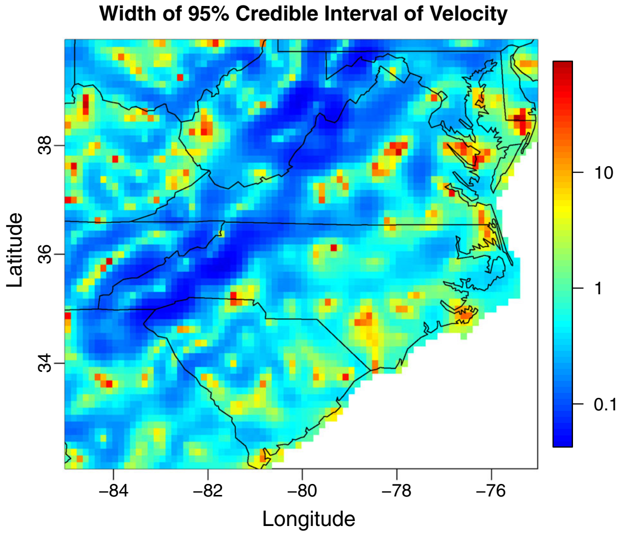

Figure 7. Width of the 95 % credible interval of the posterior distribution for velocity in the direction of the maximum gradient for the year 2012.

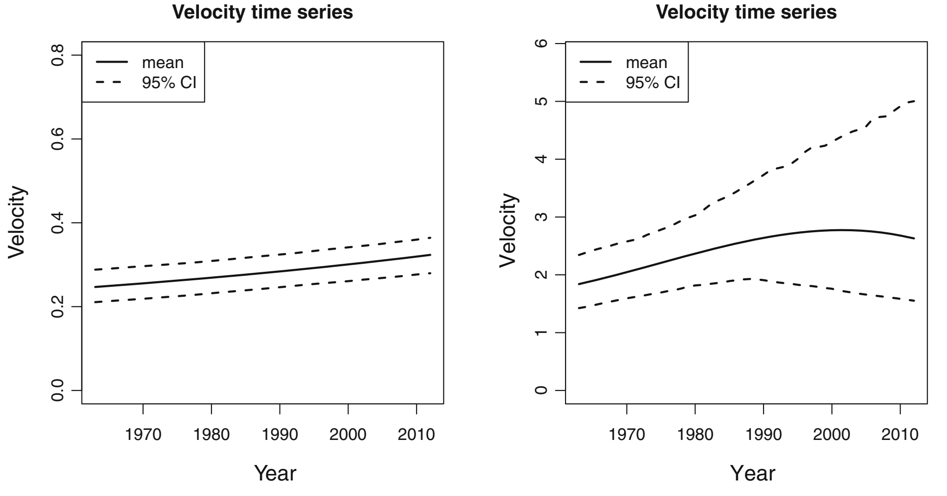

Figure 8. (*left*) Time series for the years 1963–2012 of the posterior mean velocity (km/year) and 95 % credible intervals for the two locations, mountainous (*left*) and non-mountainous (*right*).

ter and the direction of the line indicates the direction at which to travel to maintain the current temperature. That is, these posteriors distributions are of the minimum velocity. The mountainous location is on the eastern slope and has very low velocity in the uphill direction. Additionally, there is minimal angular variability in the estimate. The non-mountainous location, on the other hand, has a northward velocity that is large in terms of magnitude. It not only has much more uncertainty in the estimate for veloc-

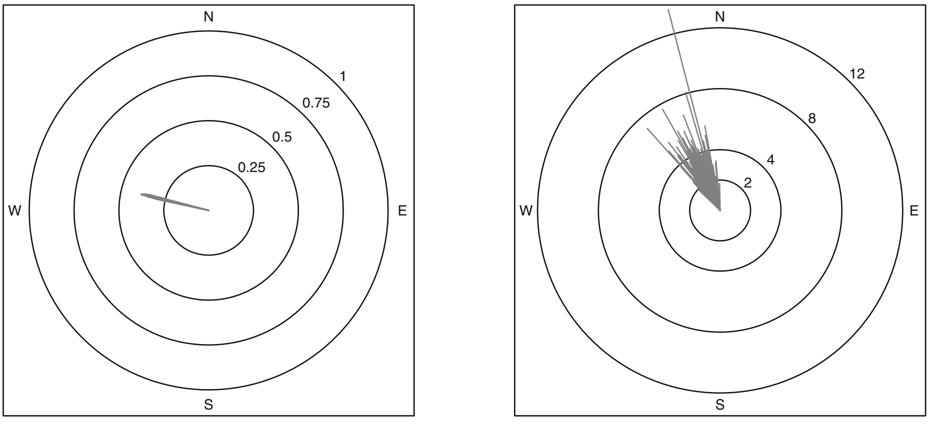

Figure 9. Posterior distribution for the directional velocities in the direction of maximum spatial gradient at the mountainous location (*left*) and non-mountainous location (*right*) for 2012. The length of the line gives the velocity (km/year) at which to travel to maintain the current temperature.

ity but also in the direction of maximum gradient signified by the high angular variability.

6. DISCUSSION AND FUTURE WORK

We have rigorously formulated a stochastic specification for annual average temperature which legitimizes calculation of infinitesimal gradients in time and space in order to create infinitesimal velocities at arbitrary locations and times. In particular, we have developed a hierarchical model that combines two sources of data, annual average temperature and elevation. The resulting expected temperature surface is viewed as a realization from a stochastic process. Directional derivative processes produce spatial and temporal gradients, and the resulting velocities, in a fully stochastic framework with inference implemented post-model fitting. Uncertainty in the spatial surfaces propagates uncertainty to all gradients and velocities. We applied the model to 50 years of annual average temperature data for the eastern United States.

There are several directions for future work and extensions of the model. In terms of temperature and velocities of climate change, we plan to develop more flexible models that include spatially varying uncertainty and non-stationary covariance functions. With climate scenarios from regional climate models (Mearns et al. 2009), we could project velocities into the future. We could also incorporate multiple sources of climate data, such as monitoring station data. In this regard, we could develop analyses at finer spatial scales using monitoring station temperature data. This would allow us to work with topography at a finer resolution and develop an elevation model using a trend surface. A finer-scale spatial model would also enable more localized gradient and velocity assessment.

Another interesting direction of future work is joint modeling of climate variables, such as temperature and precipitation. Dobrowski et al. (2013), for example, propose an ad hoc averaging method for computing velocities based upon three different climate variables. They find variability in not only the magnitude of velocities for the different climate variables but also the direction of minimum velocity. Joint process modeling in the spirit of our work here would provide a sound stochastic framework for such inference.

Lastly, the velocities presented in this work deal with climate and have no immediate connection to species. The velocities are speeds that are required to keep pace with climate change, not migration rates across space. This is, in part, due to individual species having varying tolerances and adaptability to changing climate. It would be useful to extend the modeling to formulate species-level velocities. We are currently exploring this in terms of change in intensity surfaces for point patterns of species locations, modeling the dynamics in the intensities using advection and diffusion processes.

ACKNOWLEDGEMENTS

The research of the authors was supported in part by NSF-EF-1137364. The authors thank Kai Zhu for help in obtaining PRISM annual average temperature data and Bailey Fosdick for insightful conversations in developing the MCMC sampling algorithms.

[Received October 2014. Accepted June 2015. Published Online June 2015.]

7. APPENDIX

See Figs 10, 11 and 12

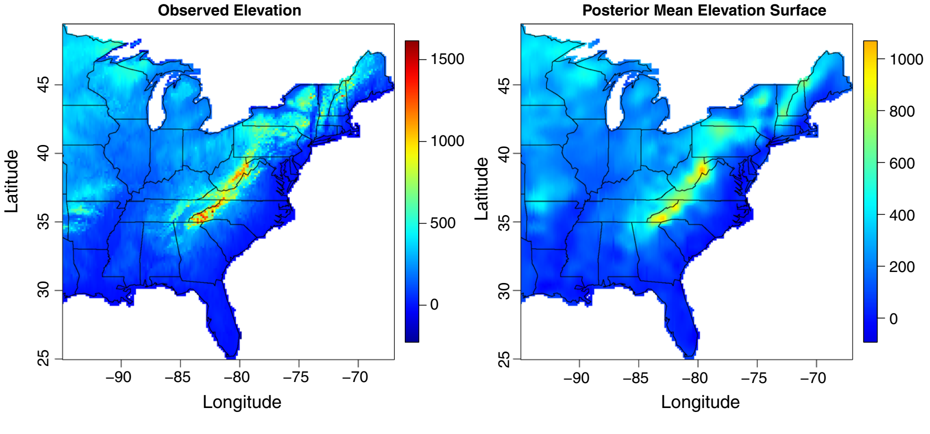

Figure 10. (*left*) Observed elevation surface and (*right*) posterior mean elevation surface indicating that the model is capturing the main features of elevation across the large spatial region. Again, our goal is not to estimate “true” elevation but rather to create a differentiable surface for elevation to improve the estimation of the spatial gradients of temperature.

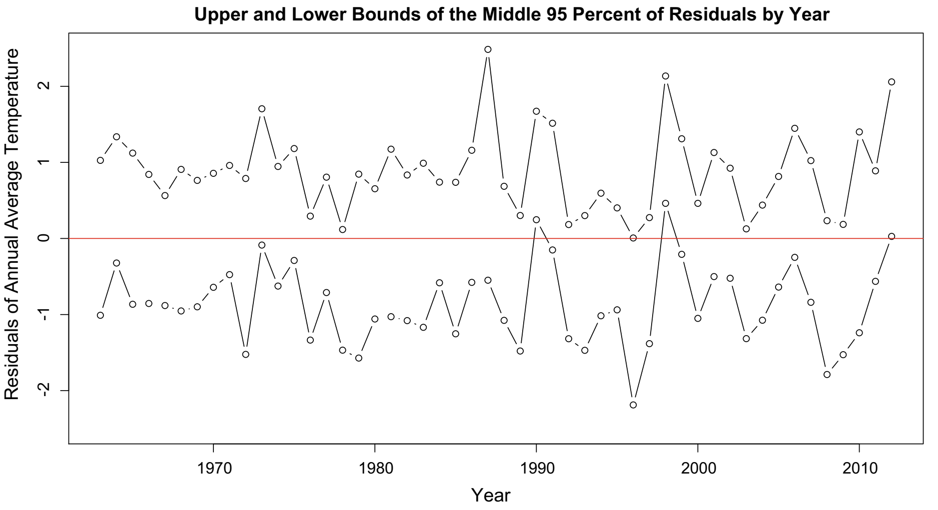

Figure 11. The *upper* and *lower* bounds of the middle 95 % of residuals by year where the residuals are computed as the observed minus the posterior mean of temperature conditional on the random effects. The model appears to underestimate annual average temperature in the years 1990, 1998, and 2012. In general, however, there does not appear to be a trend in over or under prediction through time indicating that the model adequately captures the variability in annual average temperature.

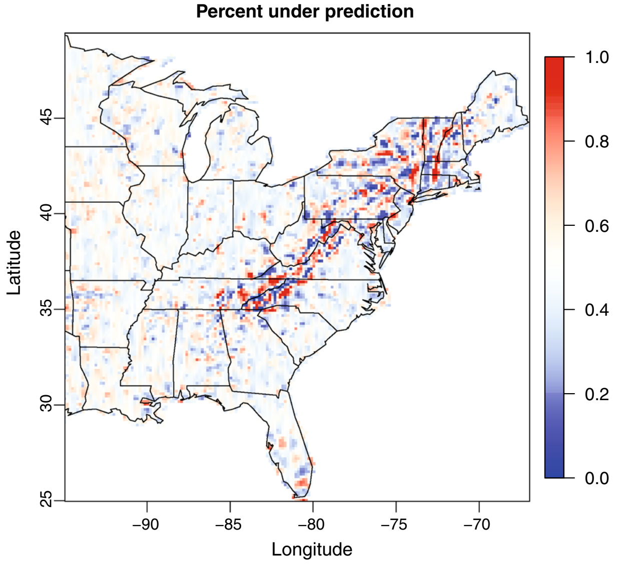

Figure 12. The percent of under predictions of annual average temperature for each location. *Dark red* (*blue*) locations indicate that the model tends to under (over) predict annual average temperature. Regions with the most consistent under and over prediction are predominantly in regions with highly variable elevation.

REFERENCES

Amante, C. and Eakins, B. W. (2009). *ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis*. US Department of Commerce, National Oceanic and Atmospheric Administration, National

Environmental Satellite, Data, and Information Service, National Geophysical Data Center, Marine Geology and Geophysics Division.

Banerjee, S. and Gelfand, A. (2003). On smoothness properties of spatial processes. *Journal of Multivariate Analysis*, 84(1):85–100.

Banerjee, S., Gelfand, A., Finley, A., and Sang, H. (2008). Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(4):825–848.

Banerjee, S., Gelfand, A. E., and Sirmans, C. (2003). Directional rates of change under spatial process models. *Journal of the American Statistical Association*, 98(464):946–954.

Benth, J. š., Benth, F. E., and Jalinskas, P. (2007). A spatial-temporal model for temperature with seasonal variance. *Journal of Applied Statistics*, 34(7):823–841.

Berrocal, V. J., Gelfand, A. E., and Holland, D. M. (2012). Space-time data fusion under error in computer model output: An application to modeling air quality. *Biometrics*, 68(3):837–848.

Burrough, P. A. and McDonnell, R. (1998). *Principles of geographical information systems*, volume 333. Oxford University Press.

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., Brown, C., Bruno, J. F., Duarte, C. M., Halpern, B. S., et al. (2011). The pace of shifting climate in marine and terrestrial ecosystems. *Science*, 334(6056):652–655.

Burrows, M. T., Schoeman, D. S., Richardson, A. J., Molinos, J. G., Hoffmann, A., Buckley, L. B., Moore, P. J., Brown, C. J., Bruno, J. F., Duarte, C. M., et al. (2014). Geographical limits to species-range shifts are suggested by climate velocity. *Nature*, 507(7493):492–495.

Craigmile, P. F. and Guttorp, P. (2011). Space-time modelling of trends in temperature series. *Journal of Time Series Analysis*, 32(4):378–395.

Dobrowski, S. Z., Abatzoglou, J., Swanson, A. K., Greenberg, J. A., Mynsberge, A. R., Holden, Z. A., and Schwartz, M. K. (2013). The climate velocity of the contiguous United States during the 20th century. *Global Change Biology*, 19(1):241–251.

Finley, A. O., Sang, H., Banerjee, S., and Gelfand, A. E. (2009). Improving the performance of predictive process modeling for large datasets. *Computational Statistics & Data Analysis*, 53(8):2873–2884.

Gelfand, A. E., Kim, H.-J., Sirmans, C., and Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98(462):387–396.

Loarie, S. R., Duffy, P. B., Hamilton, H., Asner, G. P., Field, C. B., and Ackerly, D. D. (2009). The velocity of climate change. *Nature*, 462(7276):1052–1055.

Mearns, L. O., Gutowski, W., Jones, R., Leung, R., McGinnis, S., Nunes, A., and Qian, Y. (2009). A regional climate change assessment program for North America. *Eos, Transactions American Geophysical Union*, 90(36):311–311.

Ordonez, A. and Williams, J. W. (2013). Projected climate reshuffling based on multivariate climate-availability, climate-analog, and climate-velocity analyses: implications for community disaggregation. *Climatic Change*, 119(3-4):659–675.

Peterson, D. L., Schreiner, E. G., and Buckingham, N. M. (1997). Gradients, vegetation and climate: spatial and temporal dynamics in the Olympic Mountains, USA. *Global Ecology and Biogeography Letters*, pages 7–17.

Stein, M. L. (1999). *Interpolation of spatial data: some theory for kriging*. Springer.

Terres, M. A. (2014). *Multivariate spatial process gradients with environmental applications*. PhD thesis, Duke University.