# Stochastic Modeling of Environmental Velocities

Alan E. Gelfand
Duke University
Joint work with Erin Schliep

# Introduction

- It is well-accepted that the range of plants and animals is changing in response to change in climate
- Large literature on species distribution models, explaining species presence and abundance, introducing notions like habitat models, climate envelopes, range limits, and niches, presence-absence or abundance surfaces.
- Also a large literature attempting to explain how species distribution will change in response to a changing climate scenario, e.g., species distribution models for trees to study climate change impacts on forest biodiversity at regional scales
- Useful in understanding this process is to relate change in climate over time to change in climate over space

# A basic idea

- Working with a single climate variable - here temperature - we can formalize the notion of **velocity of climate change**
- Informally, this index represents the instantaneous local velocity along the earth's surface needed to maintain constant temperature
- Expressed in km/yr over a large spatial region arising from spatial change in $^oC/km$ and $^oC/yr$
- Taking ratio of latter to former produces a *velocity*
- Initial work (Loarie et al., 2009) is crude. No explicit modeling of the climate process, deterministic or stochastic; it is purely descriptive.
- Fails to incorporate the joint linkage between temperature, time, and space
- Ad hoc uncertainty arising from variability in the ensemble of climate scenarios rather than from model mis-specification and measurement error.

# cont.

- Their basic idea:
- let Temp $\equiv T = f(t)$ where $t$ is time, i.e., a general relationship capturing change in temperature across time, say the past 100 years
- let Temp $\equiv T = g(y)$ where $y$ is latitude, i.e., a general relationship capturing change in temperature across change in latitude, say continental
- Suppose we calculate $\frac{dT}{dt}$ and $\frac{dT}{dy}$.
- Then the ratio, $\frac{\frac{dT}{dt}}{\frac{dT}{dy}} = \frac{dy}{dt}$ (for a common $dT$) is defined as the velocity of climate change in this case, in the latitudinal direction at a given $t$ and $y$

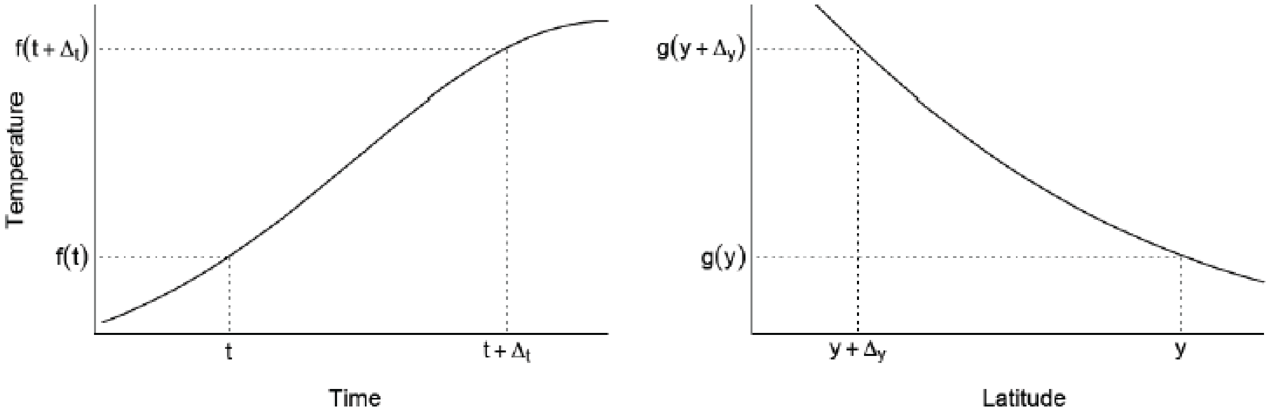
# Clarification

- We can illuminate this a bit more with finite differences and a simple figure

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

- The second fraction in the rightmost term is 1 (common $\Delta$Temp), as the suggestive figure below shows.
- Moreover, with a common "delta" temp and given a starting $t$ along with $\Delta t$, aligning with a given $y$, $\Delta y$ is determined.
- So, $\Delta y$ is the change in lat needed to provide the change in temp that arises from $t$ to $t + \Delta t$.
- With finite differences, many $\Delta$ temperatures are 0 so an arbitrary correction to obtain finite velocities

# The basic idea



# Our contribution

- We cast the development of velocity in a fully stochastic framework
- We recognize that, at the least, we should write $T(x, y, t)$, i.e., temperature is a function of both location and time.
- This legitimizes the idea of instantaneous velocity and vel= $\frac{\partial T/\partial t}{\partial T/\partial y}$
- We view the temperature surface as random, model it coherently, attach uncertainty, obtain full inference
- We specify a rich model for $T(x, y, t)$, incorporating spatial structure, anticipating that gradients, hence velocities, at close locations should be similar

## cont.

- We calculate infinitesimal derivatives through a “parametric” specification for $E(T(x, y, t))$ rather than descriptive finite difference as in GIS calculations (eight neighbor slope and aspect)
- We obtain inference about gradients and velocities as a post-model fitting exercise
- We can obtain a temperature gradient at any time and location; we can obtain a spatial gradient at any time and in any direction
- We can obtain velocity in any direction at any location and also the direction of minimum velocity
- Note that direction of maximum velocity is not meaningful mathematically or ecologically

# A temperature model

- Evidently, can build extremely complex temperature models. The one we propose is developed to capture temperature response at high spatial resolution
- We model annual average temperature using a linear mixed model with spatially correlated random effects.
- The model is inherently a hierarchical model as it combines two sources of data, annual average temperature and elevation.

# cont.

- Again, $T(x, y, t)$ is the annual average temperature for location $(x, y)$ at time $t$, $x$ is the easting coordinate, $y$ is the northing coordinate. Further, $E(x, y)$ is elevation at location $(x, y)$.
- We model $T(x, y, t) =$

$$\beta_0 + \beta_1 t + \beta_2 y + \beta_3 Z(x, y) + \beta_0(x, y) + \beta_1(x, y)t + \epsilon(x, y, t)$$

and

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

## cont.

- Here, $\epsilon(x, y, t) \sim N(0, \sigma_T^2)$ and $\eta(x, y) \sim N(0, \sigma_E^2)$.
- Both $\beta_0(x, y)$ and $\beta_1(x, y)$ are spatial random effects (intercept and slope) that account for the remaining spatial variation in annual average temperature and rate of change in annual temperature over time.
- The latent process, $Z(x, y)$ provides a spatially differentiable surface in elevation, needed since we want the gradients and velocities to be a function of elevation. So, we model $E(x, y)$ with a Gaussian process that allows explicit differentiability
- So, $Z(x, y)$ is a smooth centering surface while the $E(x, y)$ surface is not. Hopefully OK over a large spatial scale
- Remarks: Do not need a "longitude" term with coefficient; captured by elevation component
- Eastings and northings rather than longitude and latitude

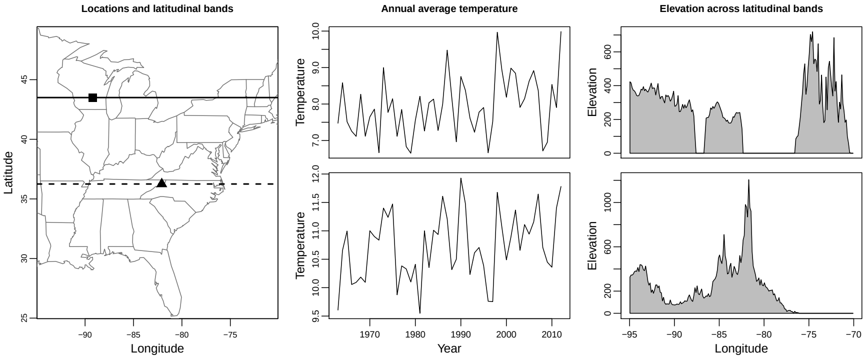
# A key point

- Two paths for the three spatial processes, $\beta_0(x,y)$, $\beta_1(x,y)$ and $Z(x,y)$.
- (i) model them as customary Gaussian processes, possibly dependent. Adopt a covariance function such that process realizations are mean square differentiable, e.g. Matérn with $\nu \geq 1$. Calculate spatial gradients (as developed in Banerjee et al. (2003))
- (ii) model them using dimension reduction, i.e., as parametric linear transformations of a finite set of random variables at fixed locations, enabling explicit gradient calculation
- Adopt the latter approach due to computational necessity (temperature at $> 21,000$ gridded locations) and use the predictive process for dimension reduction.
- Coregionalization to connect slope and intercept processes independent of latent elevation process

# The datasets

- We apply the multivariate predictive process model to temperature data for the eastern United States.
- Temperature data is from the Parameter-elevation Regression on Independent Slopes Model (PRISM) - average annual temperature (°C) for the period 1963 to 2012. Data is on 2.5 minute resolution, which we aggregate to 7.5 minute, or 1/8 degree resolution (approx 11km boxes).
- Centers of grid boxes as observed locations, average of the annual temperatures as the observations.
- Our dataset consists of 21,202 spatial locations.
- ETOPO1 elevation dataset, a 1 arc-minute global model of the earth's surface. Elevation at each of the observed temperature locations.
- Albers Equal-Area Conic projection to Albers coordinates using parallels of $29.5^\circ$ and $45^\circ$. All distances are Euclidean distances under this projection.

# Variability in annual temperature and elevation



## The predictive process

- For the spatial process for elevation, let $\boldsymbol{Z} = (Z(x_1, y_1), \ldots, Z(x_n, y_n))'$ where $(x_i, y_i)$, $i = 1, \ldots, n$ are the observed locations.
- Let $\boldsymbol{Z}^* = (Z(x_1^*, y_1^*), \ldots, Z(x_m^*, y_m^*))'$ where $(x_j, y_j^*)$, $j = 1, \ldots, m$ are the knot locations of the predictive processes. Then,

$$\widetilde{\mathbf{Z}} = C'_{\mathbf{Z},\mathbf{Z}^*}(C_{\mathbf{Z}^*})^{-1}\mathbf{Z}^*$$

$$\mathbf{Z}^* \sim GP(\mathbf{0}, C_{\mathbf{Z}^*})$$

where $C'_{\mathbf{Z},\mathbf{Z}^*}$ is an $n \times m$ covariance matrix with $(i, j)$th element equal to the correlation between $Z(x_i, y_i)$ and $Z(x_j^*, y_j^*)$ and $C_{\mathbf{Z}^*}$ is the $m \times m$ covariance matrix of $\mathbf{Z}^*$.

# cont.

- Correlation between two locations, $(x_i, y_i)$ and $(x_j, y_j)$, using the Matérn correlation function with smoothness parameter, $\nu = 3/2$, and decay parameter, $\phi_z$.
- That is, the covariance between $Z(x_i, y_i)$ and $Z(x_j, y_j)$ at any two points $(x_i, y_i)$ and $(x_j, y_j)$ is

$$\begin{aligned} Cov(Z(x_i, y_i), Z(x_j, y_j)) &= \tau_z^2 \rho(d_{ij}; \phi_z) \\ &= \tau_z^2 \left(1 + \phi_z d_{ij}\right) \exp^{-\phi_z d_{ij}} \end{aligned}$$

where $d_{ij}$ is the distance between locations $(x_i, y_i)$ and $(x_j, y_j)$ and $\tau_z^2$ is the spatial variance parameter.

# Coregionalization

- The spatial random intercept and slope processes, $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$, are modeled with a bivariate predictive process using a linear model of coregionalization (dependence anticipated).
- Let $\boldsymbol{\beta}_0$, $n \times 1$, and $\boldsymbol{\beta}_1$, $n \times 1$, be defined as

$$\begin{bmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\beta}_1 \end{bmatrix} = [A \otimes \boldsymbol{I}] \begin{bmatrix} \boldsymbol{W}_0 \\ \boldsymbol{W}_1 \end{bmatrix}$$

- Here, $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$ are independent spatial processes and $\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix}$ is a $2 \times 2$ lower triangular matrix where $a_{11}$ and $a_{22}$ are non-negative and $a_{21}$ is any real number.

## cont.

- We employ predictive processes on both $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$. Let $\boldsymbol{W}_k = (W_k(x_1, y_1), \ldots, W_k(x_n, y_n))'$ and $\boldsymbol{W}_k^* = (W_k(x_1^*, y_1^*), \ldots, W_k(x_m^*, y_m^*))'$ for $k = 0, 1$. Then,

$$\widetilde{\boldsymbol{W}}_k = C_k'(C_k^*)^{-1}\boldsymbol{W}_k^*$$
$$\boldsymbol{W}_k^* \sim GP(\mathbf{0}, C_k^*)$$

  where $C_k'$ is the covariance matrix of $\boldsymbol{W}_k$ and $\boldsymbol{W}_k^*$ and $C_k^*$ is the covariance matrix of $\boldsymbol{W}_k^*$.
- Again model correlation using the Matérn correlation function with range parameter $\phi_k$ and scale parameter $\tau_k^2$ fixed to 1 to identify of $A$.
- Then, using the predictive processes $\widetilde{\boldsymbol{W}}_0$ and $\widetilde{\boldsymbol{W}}_1$, we obtain $(\widetilde{\beta}_0, \widetilde{\beta}_1)'$ by setting $\begin{bmatrix} \widetilde{\beta}_0 \\ \widetilde{\beta}_1 \end{bmatrix} = [A \otimes \boldsymbol{I}] \begin{bmatrix} \widetilde{\boldsymbol{W}}_0 \\ \widetilde{\boldsymbol{W}}_1 \end{bmatrix}$.

# Adjusting for predictive process bias

- Predictive process systematically underestimates the variance of the spatial process at any location $(x, y)$.
- We add the adjustment term, $\zeta(x, y)$, such that

$$E(x, y) = \mu + \widetilde{Z}(x, y) + \zeta(x, y) + \widetilde{\eta}(x, y).$$

- The $\widetilde{\eta}(x_i, y_i)$'s are indep with mean 0 and variance $(C_{\mathbf{Z}})_{ii} - (C'_{\mathbf{Z},\mathbf{Z}^*}(C_{\mathbf{Z}^*})^{-1}C_{\mathbf{Z},\mathbf{Z}^*})_{ii}$
- Finally, using the predictive processes and the variance adjustment to elevation, we obtain
  $T(x, y, t) =$
  $\beta_0 + \beta_1 t + \beta_2 y + \beta_3 \widetilde{Z}(x, y) + \widetilde{\beta}_0(x, y) + \widetilde{\beta}_1(x, y)t + \epsilon(x, y, t)$
  and
  $E(x, y) = \mu + \widetilde{Z}(x, y) + \zeta(x, y) + \widetilde{\eta}(x, y).$

# Calculating gradients

- Temporal gradient: for the annual temperature model, the temporal gradient for temperature change is the *expected* change in temperature per year.
- Spatial gradient in an arbitrary direction: for the annual temperature model, the spatial gradient gives the *expected* change in temperature per kilometer.
- Can do this using the gradient in the easting direction $(\partial E(T(x,y,t))/\partial x)$, in the northing direction $(\partial E(T(x,y,t))/\partial y)$
- Let $\nabla E(T(x,y,t)) = \begin{pmatrix} \partial E(T(x,y,t))/\partial x \\ \partial E(T(x,y,t))/\partial y \end{pmatrix}$. Gradient in the direction $\mathbf{u}$, a unit vector is $\nabla E(T(x,y,t))^T \mathbf{u}$
- Max gradient direction: $\nabla E(T(x,y,t))/||\nabla E(T(x,y,t))||$
- Magnitude of the max gradient is $||\nabla E(T(x,y,t))||$

# Details

- Returning to the predictive processes, let $P^*$, $Q^*$, and $R^*$ each be $m \times m$ correlation matrices of $\boldsymbol{Z}^*$, $\boldsymbol{W}_0^*$, and $\boldsymbol{W}_1^*$, respectively, i.e.,
  $P^*_{jk} = \rho(d_{jk}; \phi_z), Q^*_{jk} = \rho(d_{jk}; \phi_0), R^*_{jk} = \rho(d_{jk}; \phi_1)$
- $\phi_z$, $\phi_0$, and $\phi_1$ are decay parameters of the Matérn correlation function with $\nu = 3/2$ and $j, k = 1, \ldots, m$.
- Further, define the $m \times 1$ correlation vectors $\mathbf{p}(x, y)$, $\mathbf{q}(x, y)$, and $\mathbf{r}(x, y)$ where the $j$th 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 $\mathbf{r}(x, y)$

## cont.

- Then, the expected annual average temperature at location $(x, y)$ and time $t$ is $E(T(x,y,t)) = \beta_0 + \beta_1 t + \beta_2 y + \beta_3 \widetilde{Z}(x,y) + \widetilde{\beta}_0(x,y) + \widetilde{\beta}_1(x,y)t = \beta_0 + \beta_1 t + \beta_2 y + \beta_3 \widetilde{Z}(x,y) + [1 \quad t] A \begin{bmatrix} \widetilde{W}_0(x,y) \\ \widetilde{W}_1(x,y) \end{bmatrix}$
- The spatial and temporal gradients at location $(x, y)$ and time $t$ are computed as the derivative of the $E(T(x,y,t))$ with respect to $x$ for the eastern direction, $y$ for the northern direction, or $t$ for time.
- That is, $\frac{\partial E(T(x,y,t))}{\partial x}$ is the spatial gradient in the $x$ direction, $\frac{\partial E(T(x,y,t))}{\partial y}$ is the spatial gradient in the $y$ direction, and $\frac{\partial E(T(x,y,t))}{\partial t}$ is the gradient through time.

# Temporal gradient

- The temporal gradient is

$$\frac{\partial E(T(x,y,t))}{\partial t} = \beta_1 + a_{21}\,\mathbf{q}(x,y)^T Q^{*-1}\boldsymbol{W}_0^* + a_{22}\,\mathbf{r}(x,y)^T R^{*-1}\boldsymbol{W}_1^*$$

- A spatial Gaussian process arising as a sum of two independent predictive processes

# Spatial gradients

- Write the derivative of $\mathbf{p}_j(x, y)$ with respect to $x$ 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.

## cont.

- Then, the spatial gradients $\frac{\partial E(T(x,y,t))}{\partial x}$ and $\frac{\partial E(T(x,y,t))}{\partial y}$ are computed as
  $\frac{\partial E(T(x,y,t))}{\partial x} = \beta_3 \frac{\partial}{\partial x}\mathbf{p}(x,y)^T P^{*-1} \boldsymbol{Z}^* + (a_{11} + a_{21}t)\frac{\partial}{\partial x}\mathbf{q}(x,y)^T Q^{*-1} \boldsymbol{W}_0^* + a_{22}t\frac{\partial}{\partial x}\mathbf{r}(x,y)^T R^{*-1} \boldsymbol{W}_1^*$
  and
  $\frac{\partial E(T(x,y,t))}{\partial y} = \beta_3 \frac{\partial}{\partial y}\mathbf{p}(x,y)^T P^{*-1} \boldsymbol{Z}^* + (a_{11} + a_{21}t)\frac{\partial}{\partial y}\mathbf{q}(x,y)^T Q^{*-1} \boldsymbol{W}_0^* + a_{22}t\frac{\partial}{\partial y}\mathbf{r}(x,y)^T R^{*-1} \boldsymbol{W}_1^*$
- These quantities give the expected change in temperature per unit of distance in the $x$ and $y$ direction
- We can compute gradients in arbitrary directions from these gradients, as described above
- Again, spatial GP's

# Finally, velocities

- A climate velocity for annual temperature is the ratio of the temporal gradient to the spatial gradient and is measured in dist/time, in our case km/yr.
- Velocity in direction $\mathbf{u}$ is $\frac{\partial E(T(x,y,t))/\partial t}{\nabla T(x,y,t)^T \mathbf{u}} = \frac{\partial E(T(x,y,t))/\partial t}{u_1 \partial E(T(x,y,t))/\partial x + u_2 \partial E(T(x,y,t))/\partial y}$
- A ratio of GP's, a Cauchy process
- Minimum velocity is velocity in direction of max gradient and is $\frac{\partial E(T(x,y,t))/\partial t}{||\nabla E(T(x,y,t))||}$
- We summarize only with minimum velocity (interpret as optimal adaptation), reduces concern regarding "0" denominators

# Again, the data

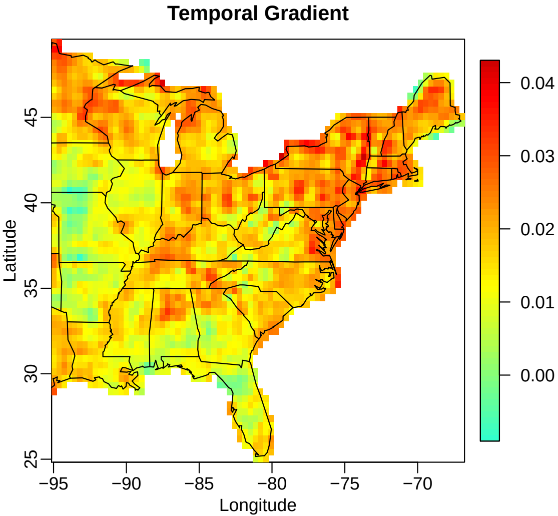
- Temperature data for the eastern United States.
- Temperature data is from the Parameter-elevation Regression on Independent Slopes Model (PRISM) - average annual temperature (°C) for the period 1963 to 2012.
- 21,202 spatial locations.
- ETOPO1 elevation dataset at each of the observed temperature locations.
- Model: $T(x, y, t) =$
  $\beta_0 + \beta_1 t + \beta_2 y + \beta_3 Z(x, y) + \beta_0(x, y) + \beta_1(x, y)t + \epsilon(x, y, t)$
  and
  $E(x, y) = \mu + Z(x, y) + \eta(x, y)$

# Parameter estimates

Table: Posterior median and 95% credible intervals

| Parameter | Median | 95% Credible Interval |
|---|---|---|
| $\beta_0$ | 12.72 | (12.68, 12.75) |
| $\beta_1(time)$ | 0.022 | (0.019, 0.023) |
| $\beta_2(lat)$ | -0.862 | (-0.866, -0.860) |
| $\beta_3(elev)$ | -0.007 | (-0.007, -0.007) |
| $\mu$ | 105.49 | (99.89, 112.73) |
| $\sigma_T^2$ | 0.457 | (0.455, 0.458) |
| $\sigma_E^2$ | 10 | |
| $\tau_Z^2$ | 31,893 | (30,593, 33,143) |

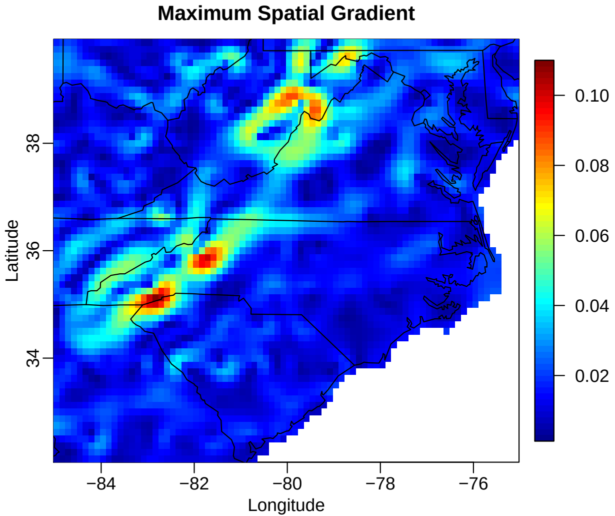
# Temperature change per year across the eastern US



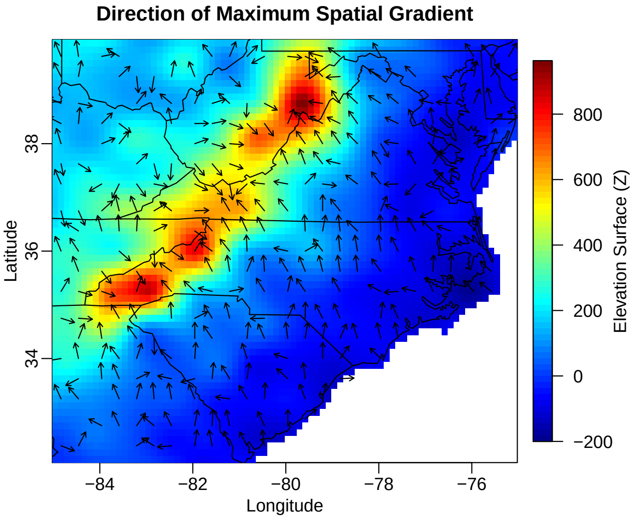

- Average increase in annual average temperature is 1.6 °C/century
- Temporal gradient is significant at 97.67% of the observed locations

# Spatial gradient of temperature across the southeast

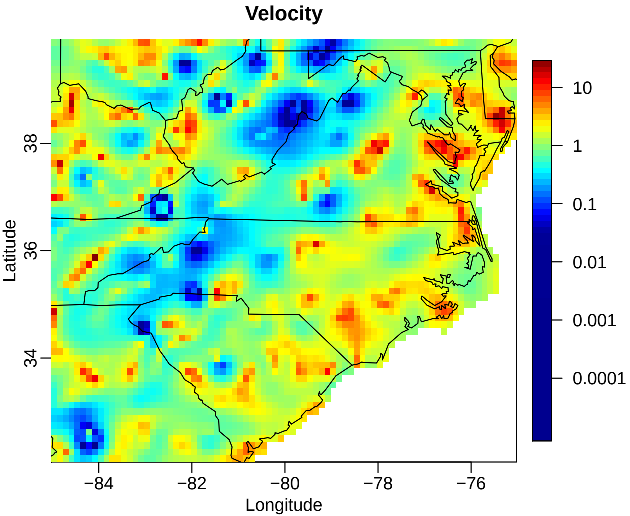
- Locations along the Appalachian Mountains are seeing temperature changes as much as 0.11 °C per km
- spatial gradient in the eastern direction is significant at 82.47% of locations (39.18% negative)
- spatial gradient in the northern direction is significant at 89.38% of locations (74.33% negative)



# Directions of maximum spatial gradient

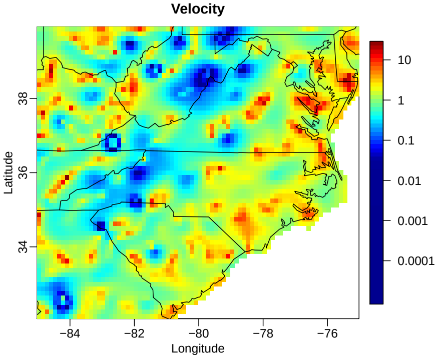


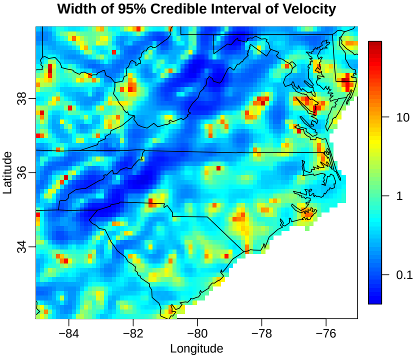
# Velocity of climate across the southeast for 2012



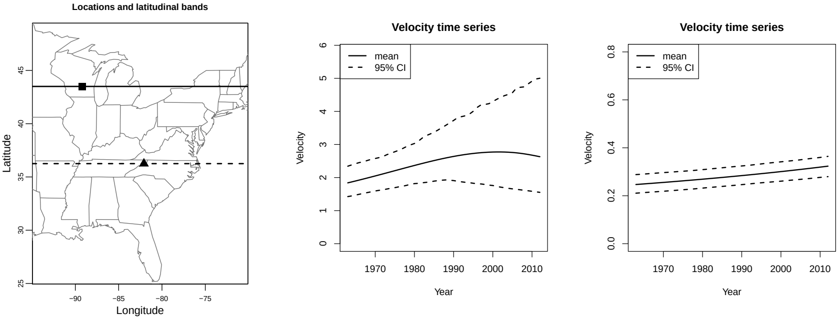

- Velocity in the direction of the maximum spatial gradient
- Positive or negative
- Range from -2.07 to 21.12 km/yr
- Quantiles: 0.67, 1.13, and 1.84 km/yr

# Uncertainty estimates of climate velocity



# Time series of velocity with credible intervals



# Posterior distribution of directional velocities for 2012