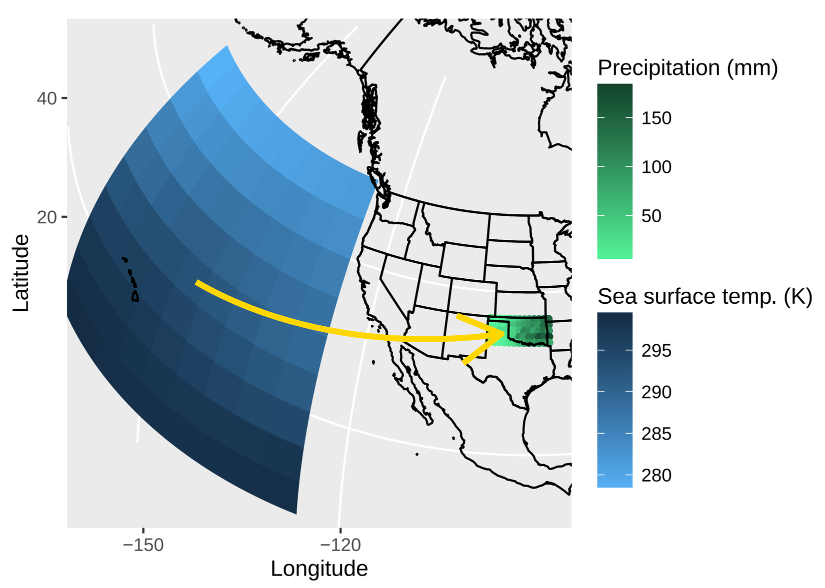
Modeling climate teleconnections

Joshua Hewitt[1], Jennifer A. Hoeting[1], and James Done[2]
[1]Department of Statistics, Colorado State University; [2]National Center for Atmospheric Research

Overview

We begin to develop a regional geostatistical model to study teleconnections—a climate phenomenon in which geographically distant areas influence regional climate patterns—at decadal time scales. Such a model could ultimately help regional planners use climate forecasts to study and prepare for local impacts of climate change.

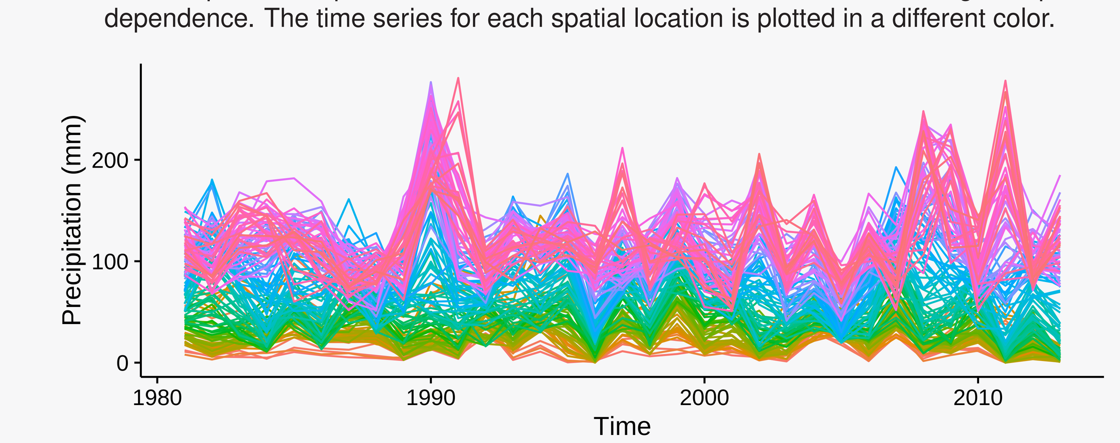


Teleconnections occur when remote covariates, like Pacific Ocean sea surface temperatures, influence regional climate variables, like average Spring precipitation in Oklahoma.

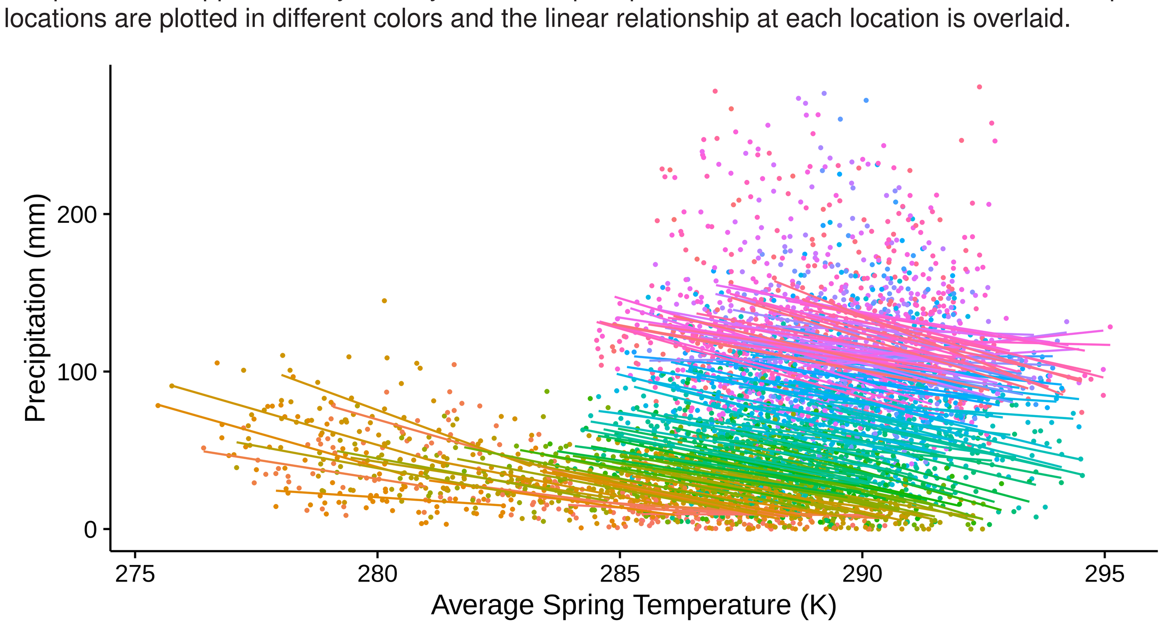
Exploratory analysis

Data: (1981-2013) Average Oklahoma and sea surface temperatures in spring from ERA-Interim Reanalysis data; PRISM precipitation (total rain and melted snow).

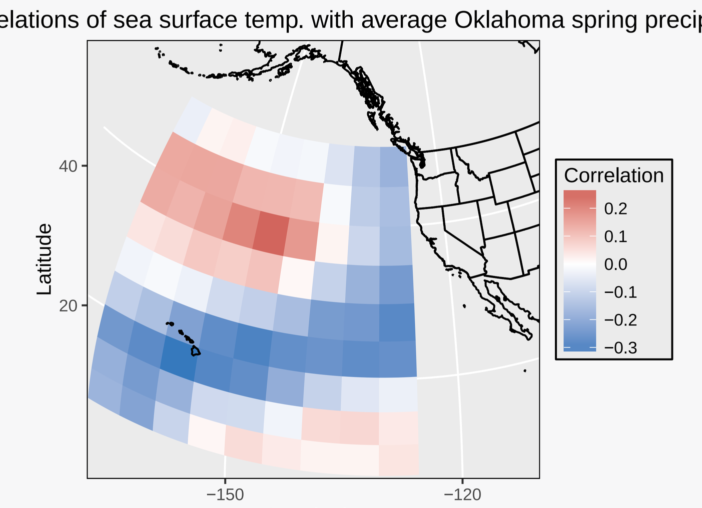
Time series plots of response variable over Oklahoma show no consistent, large temporal dependence. The time series for each spatial location is plotted in a different color.



Temperature is approximately linearly related to precipitation at each location in Oklahoma. Spatial locations are plotted in different colors and the linear relationship at each location is overlaid.



Pointwise correlations of sea surface temp. with average Oklahoma spring precipitation.



This correlation map plots the correlation between the overall average spring precipitation in Oklahoma with sea surface temperature (SST) at remote locations in the Pacific Ocean. The correlation map suggests clear, but weak teleconnection effects.

Model

We modify the spatially-varying intercept model to model teleconnection with

$$\underbrace{Y_t}_{\text{Local response}} = \underbrace{X_t\boldsymbol{\beta}}_{\text{Local effects}} + \underbrace{\left(\boldsymbol{z}_t^T\boldsymbol{\alpha}\right)\mathbf{1}_{n_s}}_{\text{Teleconnection effects}} + \underbrace{\boldsymbol{\omega}_t}_{\text{Local spatial noise}}$$

where $Y_t = \left(y_{s_1,t}, \dots, y_{s_{n_s},t}\right)^T$ and for timepoints $t \in \mathcal{T} = \{t_1, \dots, t_{n_t}\}$ and

Data	Spatial effects	Parameters
$Y_t \in \mathbb{R}^{n_s}$	$\boldsymbol{\alpha} \sim \mathcal{N}\left(\mathbf{0}_{n_r\times 1},\ \sigma_y^2\sigma_\alpha^2 R(\lambda_r)\right)$	$\boldsymbol{\beta} \sim \mathcal{N}\left(\mathbf{0}_{p\times 1},\ \Lambda\right)$
$X_t \in \mathbb{R}^{n_s \times p}$	$\boldsymbol{\omega}_t \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0}_{n_s},\ \sigma_y^2\left(H(\lambda_l) + \sigma_\varepsilon^2 I_{n_s}\right)\right)$	$\sigma^2 \sim \text{Inv-Gamma}(k,\ \theta)$
$\boldsymbol{z}_t \in \mathbb{R}^{n_r}$		$\lambda \sim \text{Uniform}(a,\ b)$

Key observations/notes:

1. We model teleconnection effects as a random intercept controlled by coefficients, $\boldsymbol{\alpha}$, that vary spatially over remote covariates, $\boldsymbol{z}_t$.
   Assumptions: time-constant assumption justified by climate system stability within decadal scales; local space-constant assumption justified by only applying model to regional data, within which effects are assumed constant.
2. Time independence of $\{\boldsymbol{\omega}_t\}$ justified by lack of large temporal trends, but could be modeled with standard spatio-temporal techniques.
3. Bounded spatial range parameters, λ, and nugget effect, σ_ε^2, ensure the model's numerical stability.
4. Scale parameters, σ^2, are parameterized for identifiability and decreased correlation during estimation.

Parameter estimation

- Estimate parameters with hybrid Gibbs algorithm applied to the model marginalized over spatial effects.
  - Adaptive random walk updates computed for:
    - $\boldsymbol{\beta}$, $\log(\sigma^2)$, $\text{logit}(\lambda)$
- Estimate teleconnection coefficients, $\boldsymbol{\alpha}$, by using posterior parameter samples to draw (in parallel) composition samples of $\boldsymbol{\alpha}$.
- Marginal likelihood complexity:
  - Likelihood evaluation in $O(n_s^3 \vee n_r^3)$ by applying Kronecker product properties and Sherman-Morrison-Woodbury formula.

Marginal model: $\left(\text{let } \Sigma_y = \sigma_y^2\left(H(\lambda_l) + \sigma_\varepsilon^2 I_{n_s}\right) \text{ and } R_\alpha = \sigma_y^2\sigma_\alpha^2 R(\lambda_r)\right)$

$$\boldsymbol{Y} = \begin{bmatrix} Y_{t_1} \\ \vdots \\ Y_{t_{n_t}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} X_{t_1} & & \\ & \ddots & \\ & & X_{t_{n_t}} \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix},\ \begin{bmatrix} \Sigma_y & & \\ & \ddots & \\ & & \Sigma_y \end{bmatrix} + \begin{bmatrix} (\boldsymbol{z}_{t_1}^T R_\alpha \boldsymbol{z}_{t_1})J_{n_s} & \cdots & (\boldsymbol{z}_{t_1}^T R_\alpha \boldsymbol{z}_{t_{n_t}})J_{n_s} \\ & \vdots & \\ (\boldsymbol{z}_{t_{n_t}}^T R_\alpha \boldsymbol{z}_{t_1})J_{n_s} & \cdots & (\boldsymbol{z}_{t_{n_t}}^T R_\alpha \boldsymbol{z}_{t_{n_t}})J_{n_s} \end{bmatrix}\right)$$

Composition-sample distribution: $\left(\text{let } R_T = R_\alpha^{-1} + \left(\mathbf{1}_{n_s}^T\Sigma_y^{-1}\mathbf{1}_{n_s}\right)\left(\sum_{t\in\mathcal{T}} \boldsymbol{z}_t\boldsymbol{z}_t^T\right)\right)$

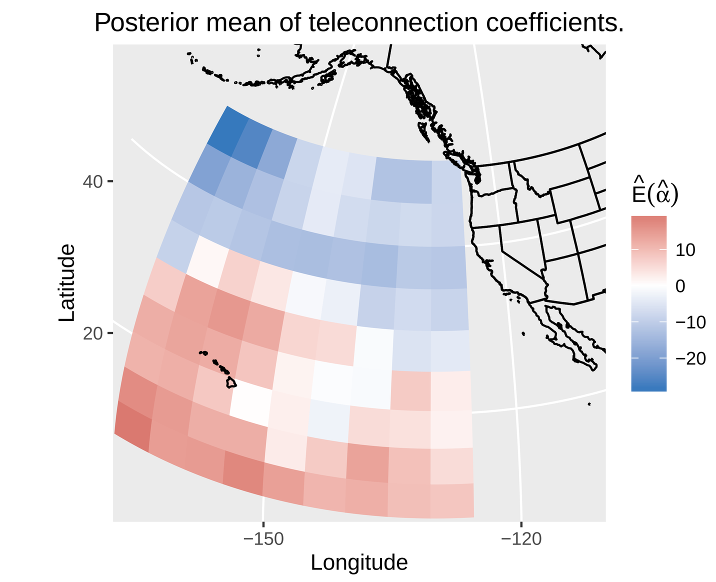
$$\boldsymbol{\alpha}|\boldsymbol{Y} \sim \mathcal{N}\left(\sum_{t\in\mathcal{T}}\left(\mathbf{1}_{n_s}^T\Sigma_y^{-1}\left(Y_t - X_t\boldsymbol{\beta}\right)\right)R_T^{-1}\boldsymbol{z}_t,\ R_T^{-1}\right)$$

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. AGS - 1419558.

Results

Posterior mean of $\alpha|\boldsymbol{Y}$ when fixing $\sigma_y^2 = 10{,}000$ and placing an informative prior on $\boldsymbol{\beta}$.



Compared to pointwise correlation maps, estimated teleconnection coefficients may yield results more consistent with known El-Niño effects, in which tropical sea surface temperatures tend to be positively correlated with spring Oklahoma precipitation.

	Posterior mean	95% HPD Interval
β_0	119	(98.3, 146)
β_1	-2.86	(-3.19, -2.52)
λ_l	7.6	(7.42, 7.74)
λ_r	1.9	(0.134, 4.17)
σ_α^2	0.311	(0.121, 0.537)
σ_ε^2	0.00424	(0.00415, 0.00432)

Posterior parameter estimates generally appear consistent with exploratory analysis.

Future work

- If our preliminary results are validated, our flexible model may produce more informative and reliable inferences than classical tools for studying teleconnections, such as correlation maps.
- Improve estimation for σ_y^2.
- Allow $\boldsymbol{\alpha}$ to vary locally as well as remotely. Challenge is to efficiently composition sample $n_s \times n_r$ teleconnection coefficients.

References

Banerjee, S., Carlin, B.P., and Gelfand, A.E., 2004, *Hierarchical Modeling and Analysis for Spatial Data*: Boca Raton, FL, Chapman and Hall/CRC.

Givens, G.H., and Hoeting, J., 2013, *Computational Statistics*: Hoboken, NJ, John Wiley & Sons, Inc.

Towler, E., Pai Mazumder, D., and Holland, G. A framework for investigating large-scale patterns as an alternative to precipitation for downscaling to local drought: Climate Dynamics *(in review)*.