A Spatial Bayesian Hierarchical Model for a Precipitation Return Levels Map

Daniel Cooley[1,2]

Douglas Nychka[2], Philippe Naveau[2,3]

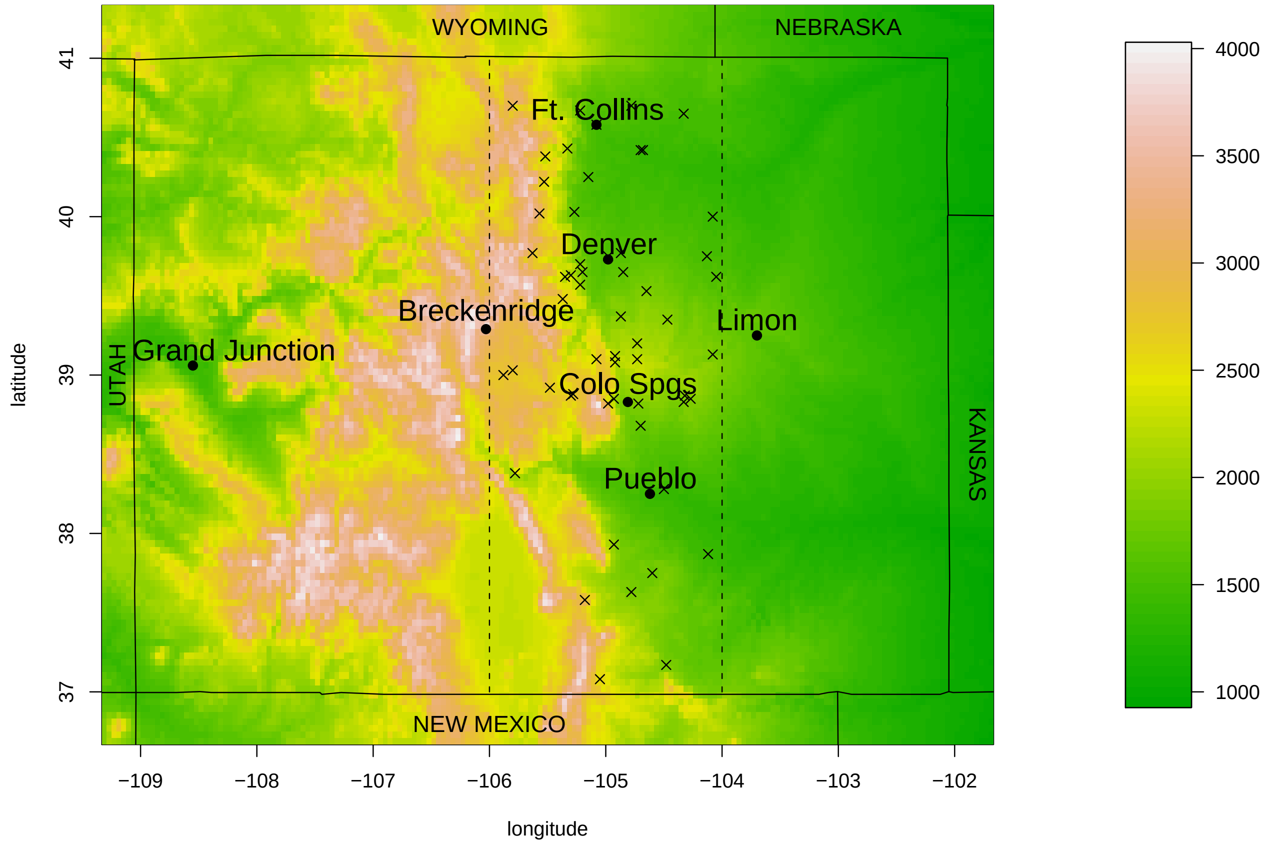
[1]Department of Applied Mathematics, University of Colorado at Boulder
[2]Geophysical Statistics Project, National Center for Atmospheric Research
[3]Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS, Gif-sur-Yvette, Fr

Project Background

Goal: To produce a map which describes potential extreme precipitation for Colorado's Front Range.

- Part of a larger NCAR project on flooding
- 1973 NOAA/NWS Precipitation Atlas is currently used
  - no uncertainty estimates
  - outdated extremes techniques
  - 30 more years of data
- Current NWS effort to produce updated maps
  - maps produced for two US regions (not Colorado)
  - using RFA methodology of Hoskings and Wallis
- Precipitation atlases provide return levels measures

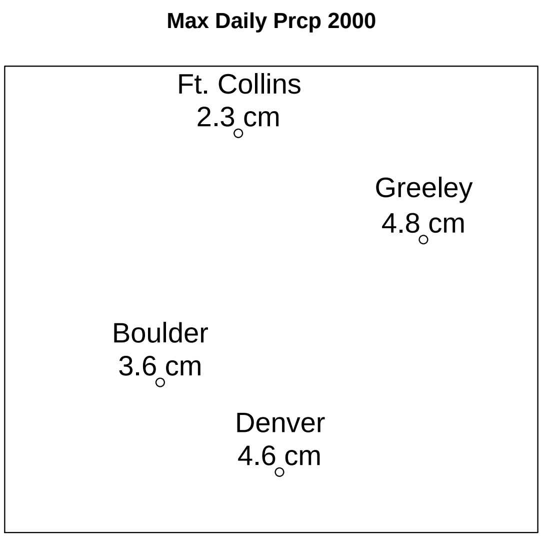
Study Region



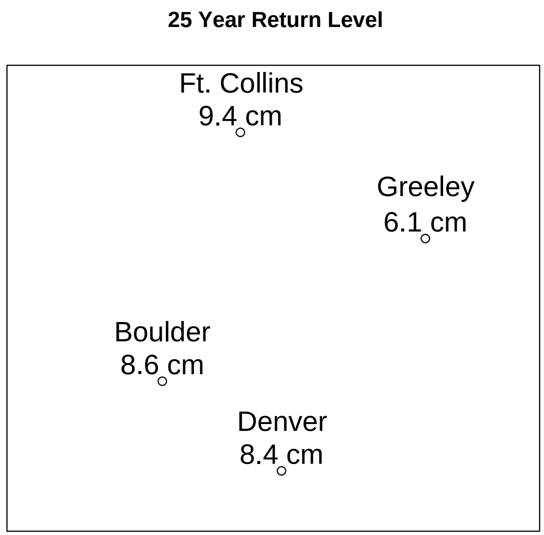
Data: 56 weather stations, 12-53 years of data/station, Apr 1 - Oct 31, 24 hour precipitation measurements

Weather, Climate and Spatial Extremes

- Modeling observations.
- Characterizing spatial dependence in the data.
- Short-range dependence.



- Modeling return levels.
- Characterizing dependence in the distributions.
- Longer-range dependence.



Modeling Climatological Extremes

- We use a POT approach, and assume exceedances over a threshold u are described by a $GPD(\sigma_u, \xi)$.
- We assume the climatological extreme precipitation is characterized by a latent process – σ_u and ξ are functions of location
- Return levels:

$$z_r(\boldsymbol{x}) = u(\boldsymbol{x}) + \frac{\sigma_u(\boldsymbol{x})}{\xi(\boldsymbol{x})} \left[(r n_y \zeta_u(\boldsymbol{x}))^{\xi(\boldsymbol{x})} - 1 \right].$$

  – n_y is the number of observations in a year.
  – $\zeta_u(\boldsymbol{x})$ is the probability an observation exceeds u.

Model Goals

- Utilize extreme value theory (GPD)
- Pool the data from the stations into one model – different from RFA
- Model should have spatial coherence
- Should utilize available covariates – elevation and mean Apr-Oct precipitation
- Should be flexible enough to be able to compare models
- Produce measures of uncertainty

Spatial Bayesian Hierarchical Models:
Independent 3-layer (data, process, prior) models for threshold exceedances and exceedance rates.

Exceedances Model - Data Level

Let $Z_j(x_i)$ be the precipitation amount recorded at the station located at x_i on day j. We assume that precipitation events $Z_j(x_i)$ which exceed a threshold $u = .45$ inches are GPD, whose parameters depend on the station's location.

$$P\{Z_j(x_i) - u > z | Z_j(x_i) > u\} = \left(1 + \frac{\xi(\boldsymbol{x}_i) z}{\exp \phi(\boldsymbol{x}_i)}\right)^{-1/\xi(\boldsymbol{x}_i)}$$

Exceedances Model - Process Level

$\phi(\boldsymbol{x})$: Modeled with standard geophysical methods $\longrightarrow$ Gaussian process

$$\begin{aligned} \mu_\phi(\boldsymbol{x}) &= f(\boldsymbol{\alpha}_\phi, covariates(\boldsymbol{x})) \\ &= \alpha_{\phi,0} + \alpha_{\phi,1}(elevation) \text{ (for example)} \end{aligned}$$

$$k_\phi(\boldsymbol{x}, \boldsymbol{x}') = \beta_{\phi,0} * exp(-\beta_{\phi,1} * ||\boldsymbol{x} - \boldsymbol{x}'||_2)$$

Exceedances Model - Process Level

$\phi(\boldsymbol{x})$: Modeled with standard geophysical methods $\rightarrow$ Gaussian process

$$\begin{aligned}\mu_\phi(\boldsymbol{x}) &= f(\boldsymbol{\alpha}_\phi, covariates(\boldsymbol{x})) \\ &= \alpha_{\phi,0} + \alpha_{\phi,1}(elevation) \text{ (for example)}\end{aligned}$$

$$k_\phi(\boldsymbol{x}, \boldsymbol{x}') = \beta_{\phi,0} * exp(-\beta_{\phi,1} * ||\boldsymbol{x} - \boldsymbol{x}'||_2)$$

$\xi(\boldsymbol{x})$: Modeled in three ways

1. as a single value ξ for the whole region
2. as separate values ξ_{mtn}, ξ_{plains}
3. as a Gaussian process as above

Exceedances Model - Priors

Priors for $\alpha_{\phi,\cdot}$: Non-informative

$$\alpha_{\phi,\cdot} \sim Unif(-\infty, \infty)$$

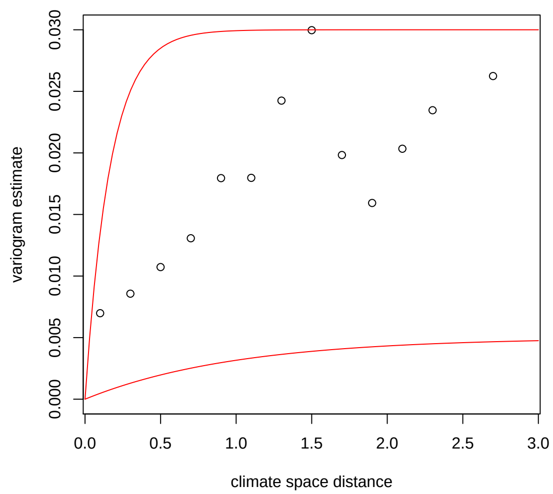
Exceedances Model - Priors

Priors for $\alpha_{\phi,\cdot}$: Non-informative

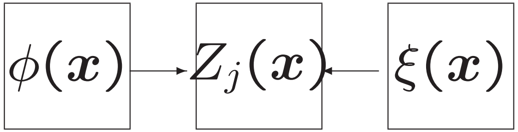
$$\alpha_{\phi,\cdot} \sim Unif(-\infty, \infty)$$

Priors for $\beta_{\phi,\cdot}$: Based on empirical information – difficult to ellicit prior information

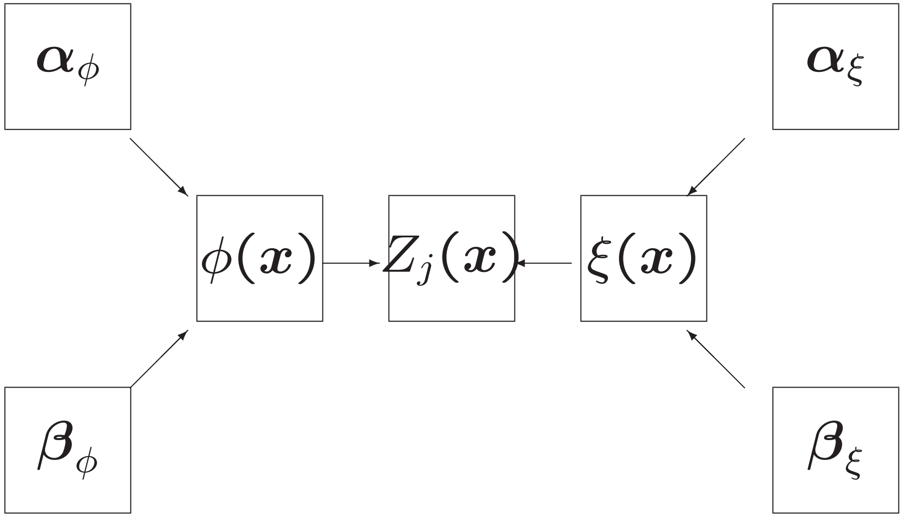
$$\begin{aligned}\beta_{\phi,0} &\sim Unif(0.005, 0.03)\\ \beta_{\phi,1} &\sim Unif(1, 6)\end{aligned}$$



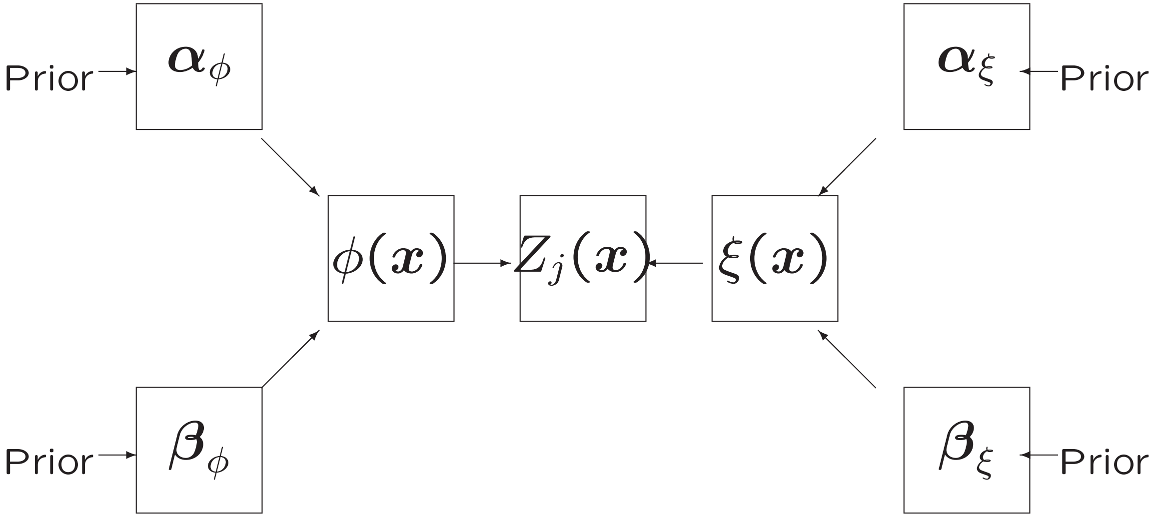
Model Schematic



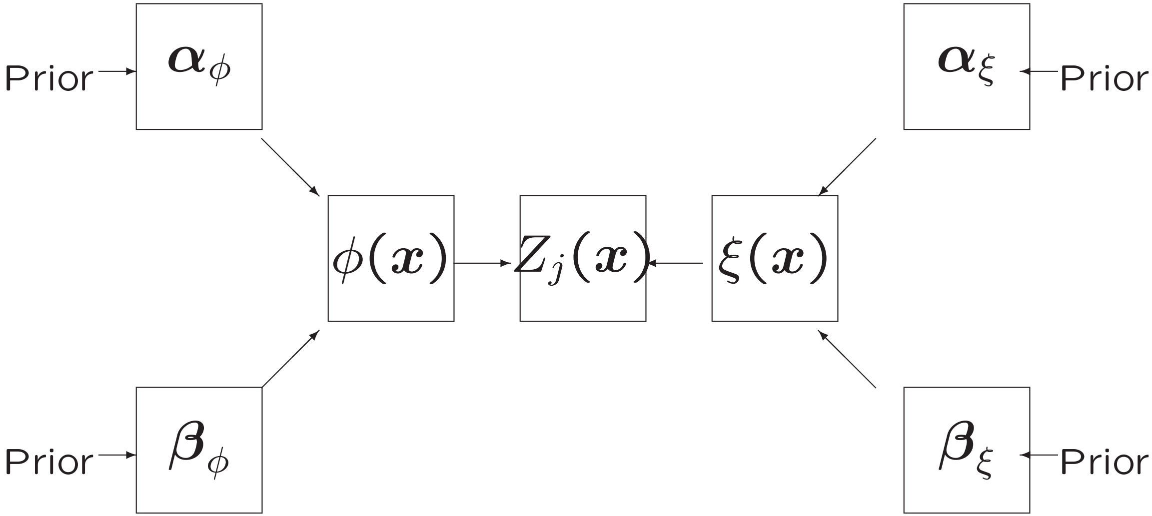
Model Schematic



Model Schematic



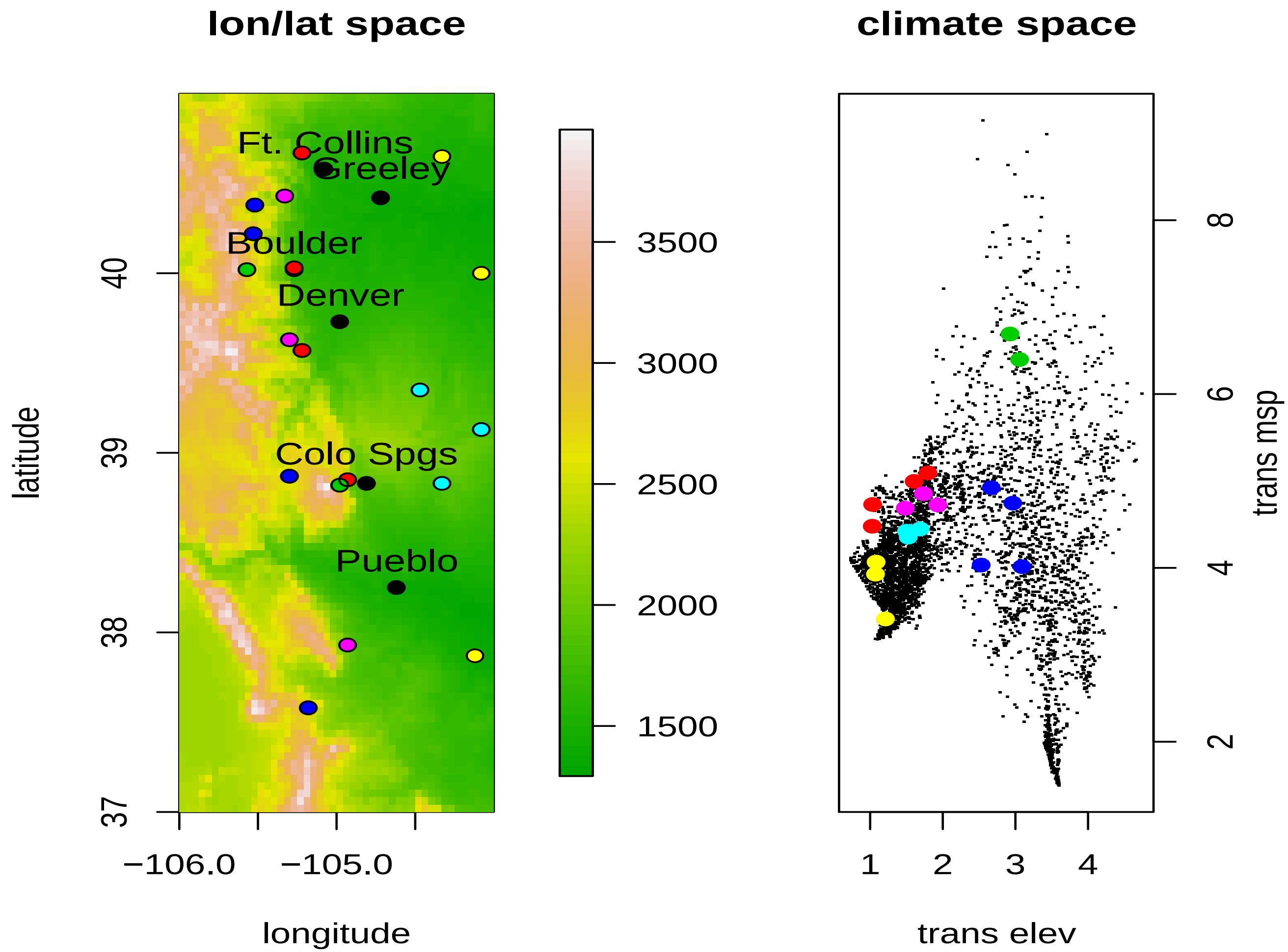
Model Schematic



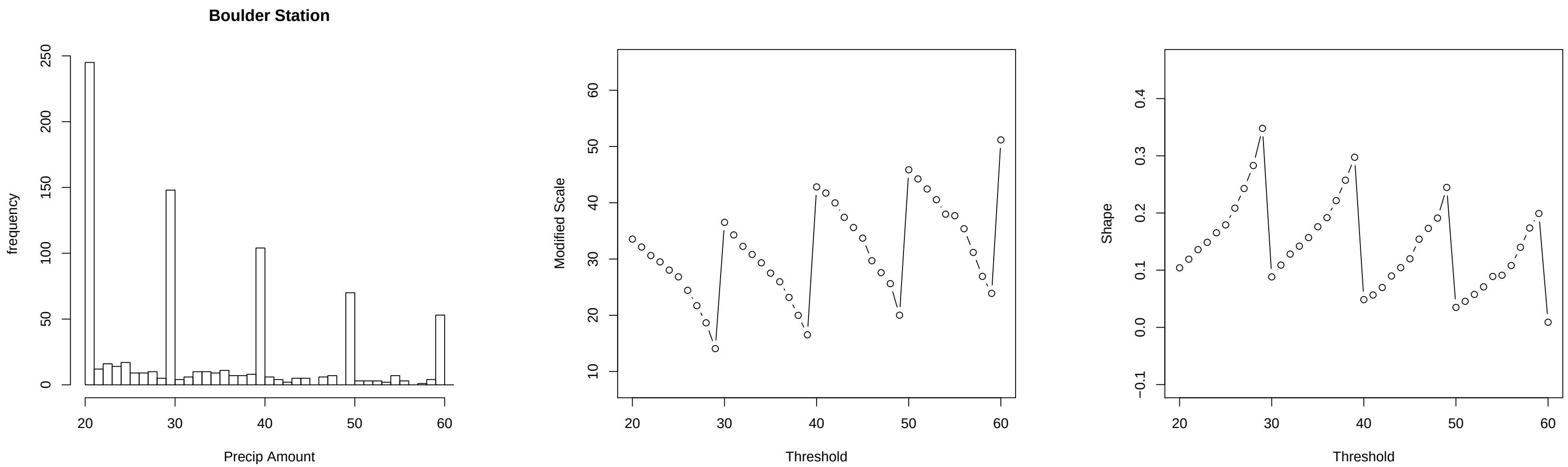
Model assumes that the observations are temporally and spatially independent (conditional on the stations' parameters).

Climate Space

Problem: Difficult to obtain convergence for $\beta_{\phi,1}$.



Threshold Selection: Quantile or Level?



Simulation experiment: Parameter bias least if threshold is chosen in the middle of the precision interval.

Threshold chosen at .45 inches for all stations

$\Rightarrow \zeta_u(\boldsymbol{x})$ modeled spatially

$\Rightarrow$ Exceedance Rate Model

Exceedance Rate Model

Data Layer: Assume each station's number of exceedances N_i is a binomial random variable with m_i trials each with a probability of $\zeta(\boldsymbol{x}_i)$

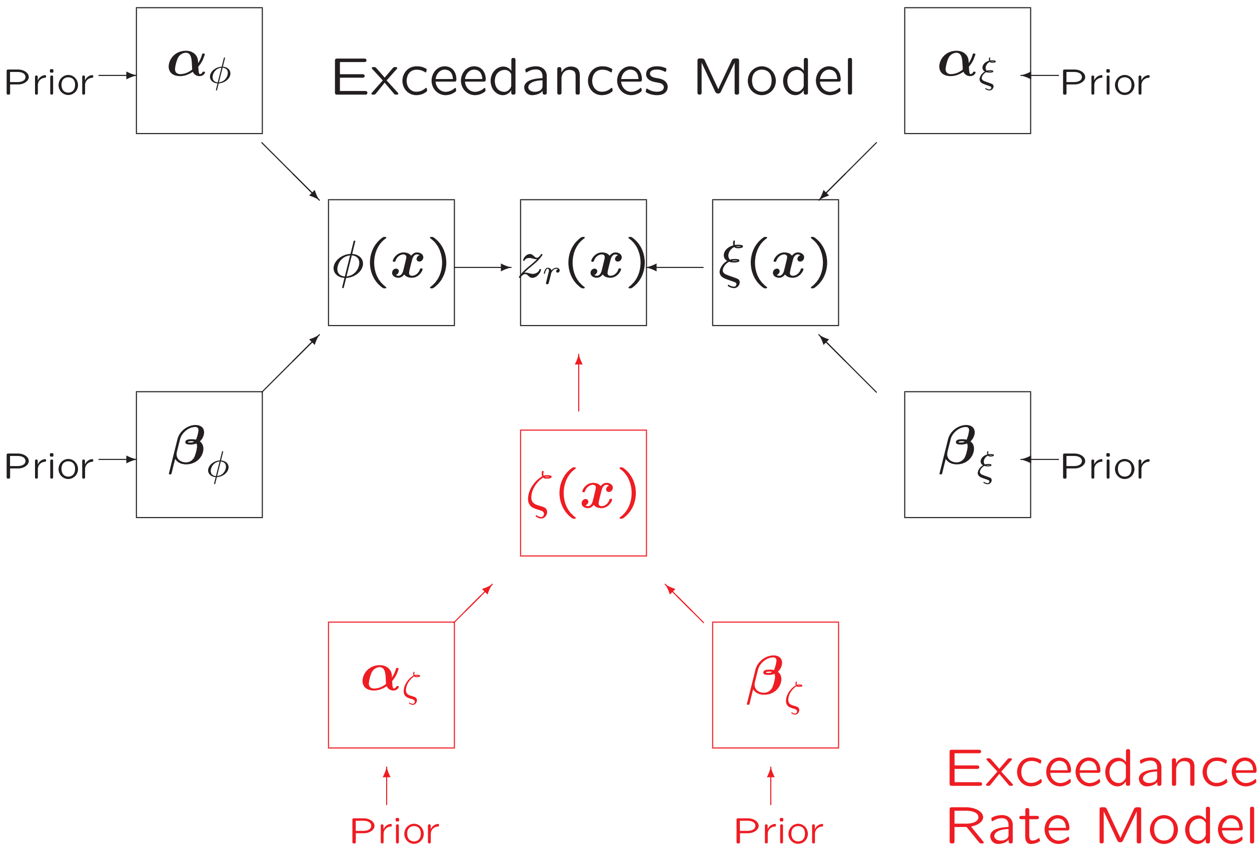
Process Layer: Assume $logit(\zeta(\boldsymbol{x}))$ is a Gaussian process, with mean and covariance

$$\begin{aligned} \mu_\zeta(\boldsymbol{x}) &= f_\zeta(\boldsymbol{\alpha}_\zeta, covariates(\boldsymbol{x})) \\ Cov(\zeta(\boldsymbol{x}), \zeta(\boldsymbol{x}') &= k_\zeta(\boldsymbol{x}, \boldsymbol{x}') = \beta_{\zeta,0} * exp(-\beta_{\zeta,1} * ||\boldsymbol{x} - \boldsymbol{x}'||_2) \end{aligned}$$

Priors:

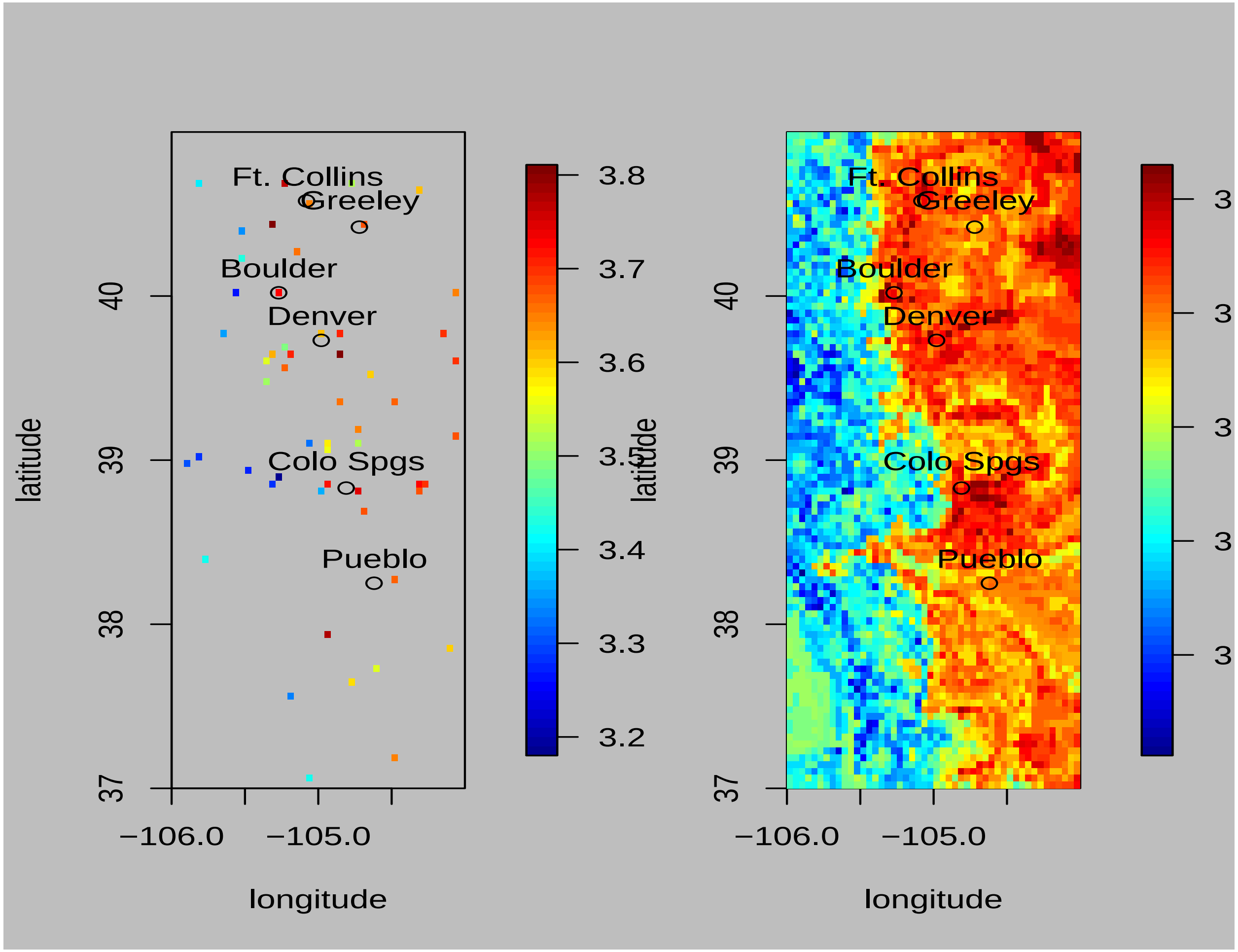
- $\alpha_{\zeta,\cdot} \sim Unif(-\infty, \infty)$,
- $\beta_{\zeta,0} \sim Unif(0.005, .2)$
- $\beta_{\zeta,1} \sim Unif(1, 6)$

Model Schematic for Return Levels



Interpolation Method

Draw values from $[\phi(\boldsymbol{x})|\phi(x_1), \ldots, \phi(x_{56}), \boldsymbol{\alpha}_\phi, \boldsymbol{\beta}_\phi]$.

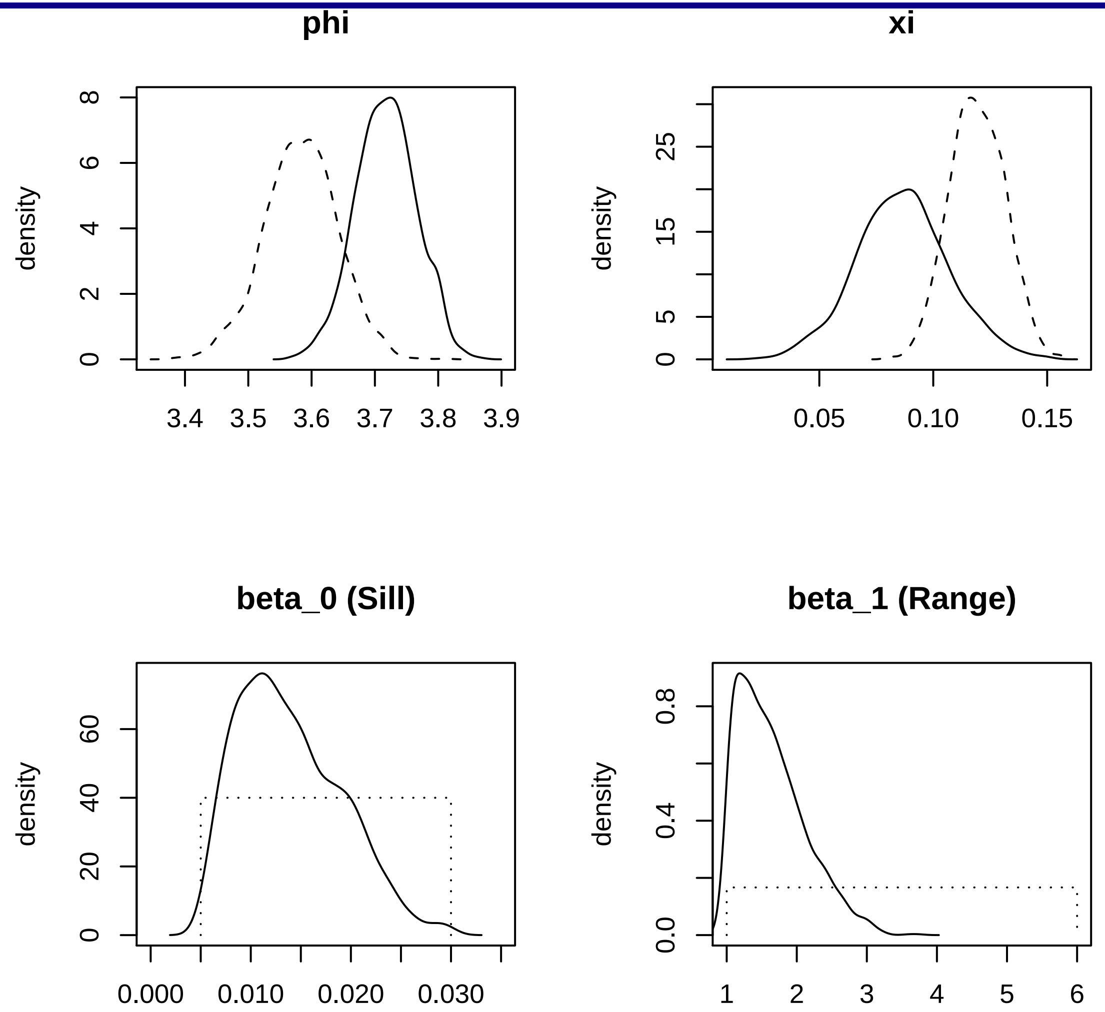


Exceedance Models Tested

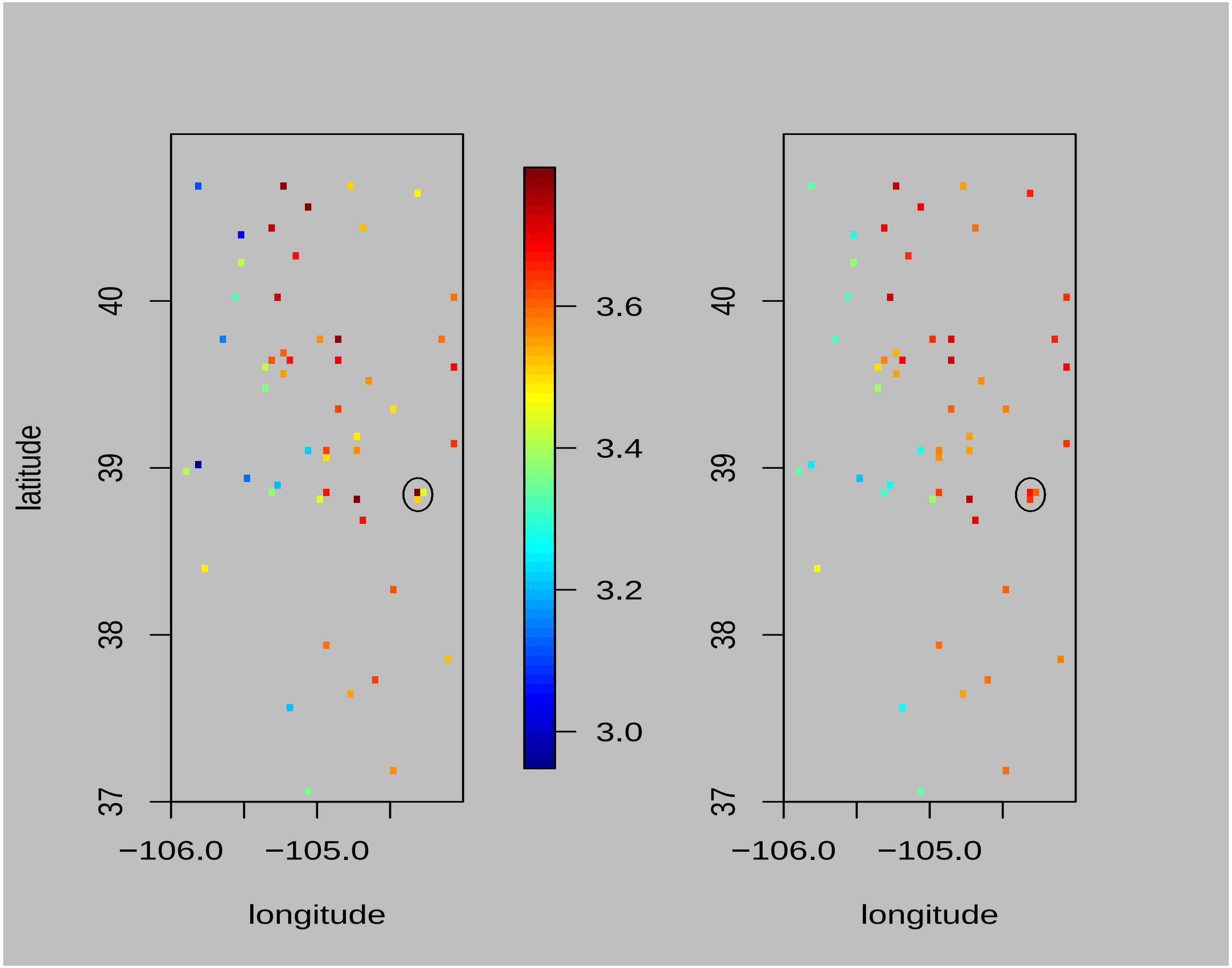
Baseline Model			$\bar{D}$	p_D	DIC
Model 0:	$\phi =$	ϕ	112264.2	2.0	112266.2
	$\xi =$	ξ			
Models in Latitude/Longitude Space			$\bar{D}$	p_D	DIC
Model 1:	$\phi =$	$\phi + \epsilon_\phi$	98533.2	33.8	98567.0
	$\xi =$	ξ			
Model 2:	$\phi =$	$\alpha_0 + \alpha_1(\text{msp}) + \epsilon_\phi$	98532.3	33.8	98566.1
	$\xi =$	ξ			
Model 3:	$\phi =$	$\alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$	98528.8	30.4	98559.2
	$\xi =$	ξ			
Model 4:	$\phi =$	$\alpha_0 + \alpha_1(\text{elev}) + \alpha_2(\text{msp}) + \epsilon_\phi$	98529.7	29.6	98559.6
	$\xi =$	ξ			
Models in Climate Space			$\bar{D}$	p_D	DIC
Model 5:	$\phi =$	$\phi + \epsilon_\phi$	98524.3	27.3	98551.6
	$\xi =$	ξ			
Model 6:	$\phi =$	$\alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$	98526.0	25.8	98551.8
	$\xi =$	ξ			
Model 7:	$\phi =$	$\alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$	**98524.0**	**26.0**	**98550.0**
	$\xi =$	$\xi_{mtn},\ \xi_{plains}$			
Model 8:	$\phi =$	$\alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$	98518.5	79.9	98598.4
	$\xi =$	$\xi + \epsilon_\xi$			

$\epsilon_{\cdot} \sim MVN(0, \Sigma)$ where $[\sigma]_{i,j} = \beta_{\cdot,0} \exp(-\beta_{\cdot,1}||\boldsymbol{x}_i - \boldsymbol{x}_j||)$

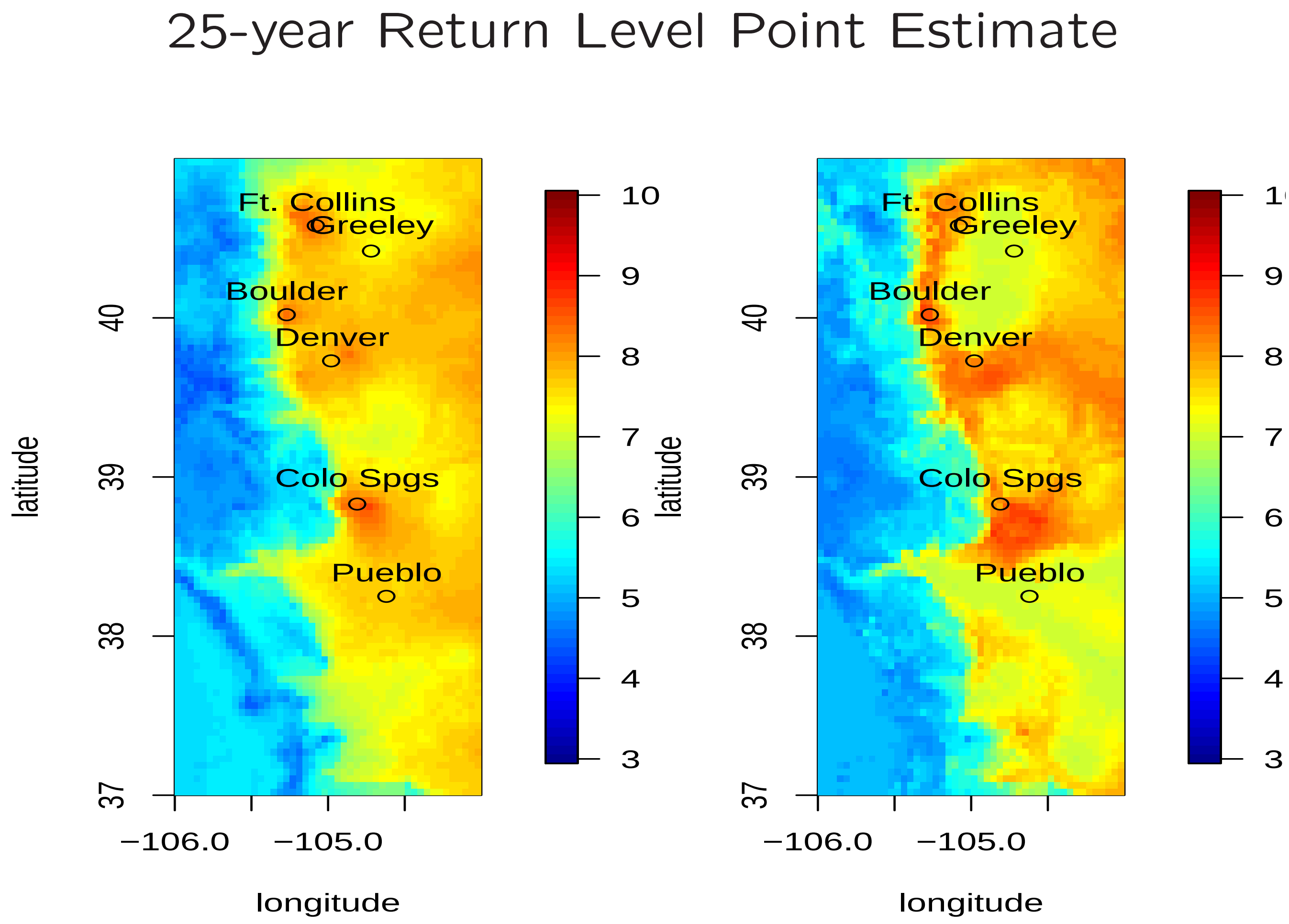
Posterior Distributions



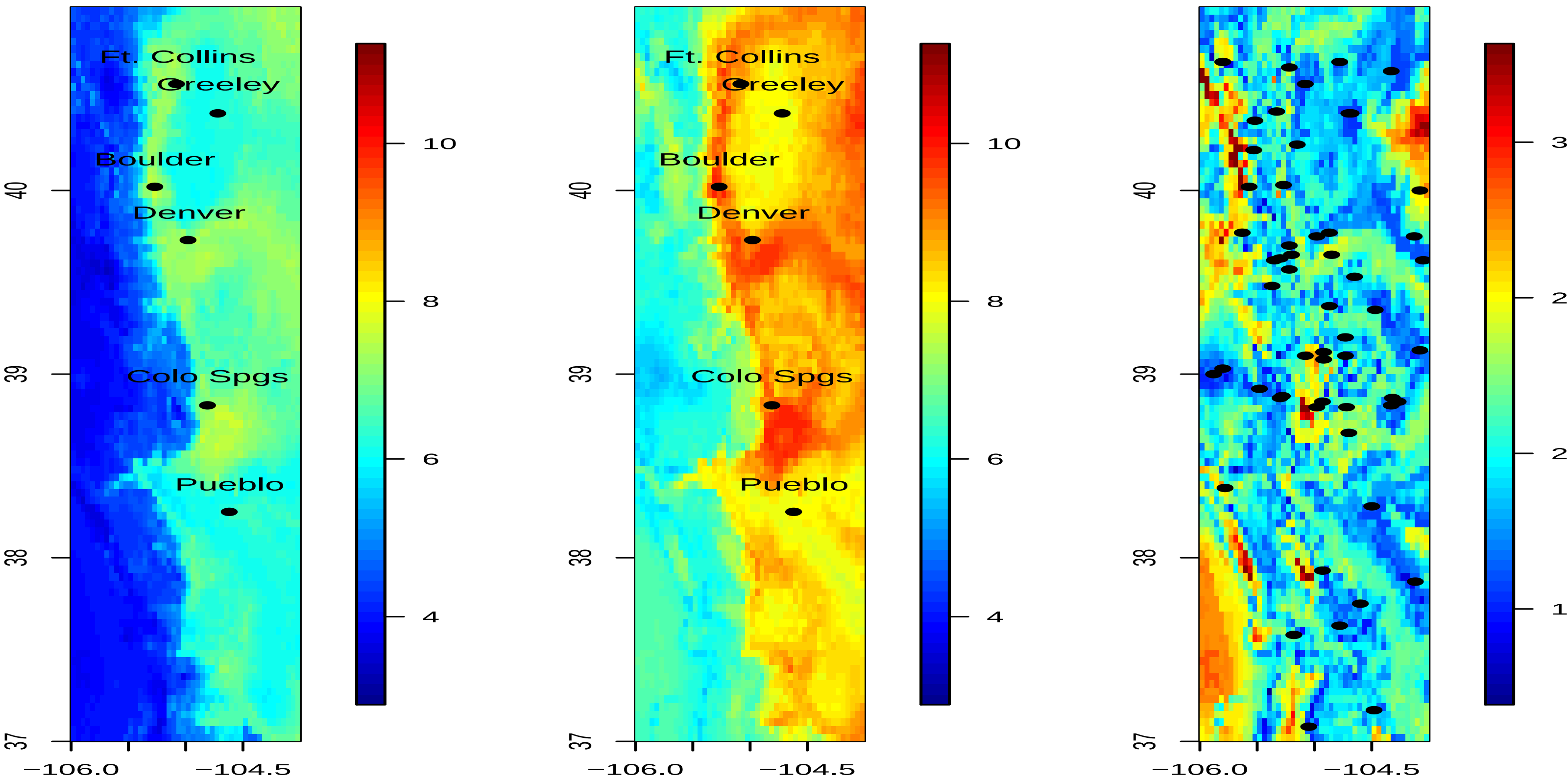
Spatial Coherence of ϕ



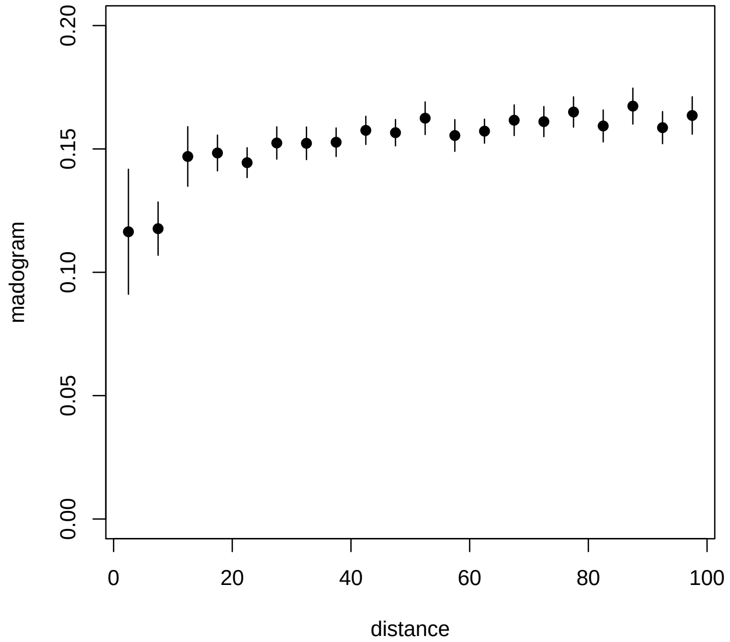
Traditional vs Climate Space



Results for Model 3: Return Level Uncertainty



Madogram on Colorado Data



- Renormalized *annual max* data.
- Shows very short range dependence in annual max observations.
- Like to apply the madogram to GPD data.

Conclusions and Future Work

- Used a Bayesian hierarchical model to produce maps which characterize *climatological* extreme precipitation.
- Method accounts for uncertainty due to both parameter estimation and interpolation.
- Dealt with issue of low precision by modeling exceedance rate spatially.
- Performed spatial analysis in a non-traditional climate space.

- Extend idea to other data (ozone levels).
- Model for all duration periods.