A Spatial Bayesian Hierarchical Model for a Precipitation Return Levels Map

Daniel Cooley\textsuperscript{1,2}

Douglas Nychka\textsuperscript{2}, Philippe Naveau\textsuperscript{2,3}

\textsuperscript{1}Department of Applied Mathematics, University of Colorado at Boulder
\textsuperscript{2}Geophysical Statistics Project, National Center for Atmospheric Research
\textsuperscript{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  - outdated extremes techniques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
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.

Max Daily Prcp 2000

<table>
<thead>
<tr>
<th>Location</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ft. Collins</td>
<td>2.3 cm</td>
</tr>
<tr>
<td>Greeley</td>
<td>4.8 cm</td>
</tr>
<tr>
<td>Boulder</td>
<td>3.6 cm</td>
</tr>
<tr>
<td>Denver</td>
<td>4.6 cm</td>
</tr>
</tbody>
</table>

25 Year Return Level

<table>
<thead>
<tr>
<th>Location</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ft. Collins</td>
<td>9.4 cm</td>
</tr>
<tr>
<td>Greeley</td>
<td>6.1 cm</td>
</tr>
<tr>
<td>Boulder</td>
<td>8.6 cm</td>
</tr>
<tr>
<td>Denver</td>
<td>8.4 cm</td>
</tr>
</tbody>
</table>
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 – $\sigma_u$ and $\xi$ are functions of location.

- Return levels:

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

- $n_y$ is the number of observations in a year.
- $\zeta_u(x)$ 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.
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.

$$P\{Z_j(x_i) - u > z|Z_j(x_i) > u\} = \left(1 + \frac{\xi(x_i)z}{\exp \phi(x_i)}\right)^{-1/\xi(x_i)}$$
Exceedances Model - Process Level

\( \phi(x) \): Modeled with standard geophysical methods \( \rightarrow \) Gaussian process

\[
\mu_\phi(x) = f(\alpha_\phi, \text{covariates}(x)) = \alpha_{\phi,0} + \alpha_{\phi,1}(\text{elevation}) \text{ (for example)}
\]

\[
k_\phi(x, x') = \beta_{\phi,0} \ast \exp(-\beta_{\phi,1} \ast ||x - x'||_2)
\]
Exceedances Model - Process Level

\( \phi(x) \): Modeled with standard geophysical methods \( \rightarrow \) Gaussian process

\[
\mu_\phi(x) = f(\alpha_\phi, \text{covariates}(x)) \\
= \alpha_{\phi,0} + \alpha_{\phi,1}(\text{elevation}) \quad \text{(for example)}
\]

\[
k_\phi(x,x') = \beta_{\phi,0} \ast \exp(-\beta_{\phi,1} \ast \|x - x'|_2)
\]

\( \xi(x) \): Modeled in three ways

1. as a single value \( \xi \) for the whole region
2. as separate values \( \xi_{\text{mtn}}, \xi_{\text{plains}} \)
3. as a Gaussian process as above
Exceedances Model - Priors

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

$$\alpha_{\phi,\cdot} \sim \text{Unif}(-\infty, \infty)$$
Exceedances Model - Priors

**Priors for \( \alpha_{\phi} \):** Non-informative

\[
\alpha_{\phi} \sim Unif(-\infty, \infty)
\]

**Priors for \( \beta_{\phi} \):** Based on empirical information – difficult to elicit prior information

\[
\beta_{\phi,0} \sim Unif(0.005, 0.03) \\
\beta_{\phi,1} \sim Unif(1, 6)
\]
Model Schematic

\[ \phi(x) \xrightarrow{} Z_j(x) \xrightarrow{} \xi(x) \]
Model Schematic

\[ \phi(x) \rightarrow Z_j(x) \rightarrow \xi(x) \]

Prior \( \alpha_\phi \)

Prior \( \beta_\phi \)

Prior \( \alpha_\xi \)

Prior \( \beta_\xi \)
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 $\beta_{\phi,1}$. 

Ion/lat space

longitude

latitude

-106.0 -105.0

37 38 39 40

Ft. Collins
Greeley
Boulder
Denver
Colo Spgs
Pueblo

climate space

trans msp

trans elev

1 2 3 4

2 4 6 8

1500 2000 2500 3000 3500

13
Simulation experiment: Parameter bias least if threshold is chosen in the middle of the precision interval.

Threshold chosen at .45 inches for all stations
⇒ \( \zeta_u(\mathbf{x}) \) modeled spatially
⇒ 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 $\zeta(x_i)$

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

$$\mu_\zeta(x) = f_\zeta(\alpha_\zeta, \text{covariates}(x))$$

$$\text{Cov}(\zeta(x), \zeta(x')) = k_\zeta(x, x') = \beta_{\zeta,0} \ast \exp(-\beta_{\zeta,1} \ast ||x - x'||_2)$$

**Priors:**

- $\alpha_\zeta \sim \text{Unif}(-\infty, \infty)$,
- $\beta_{\zeta,0} \sim \text{Unif}(0.005, .2)$
- $\beta_{\zeta,1} \sim \text{Unif}(1, 6)$
Model Schematic for Return Levels

Exceedances Model

\[ \phi(x) \rightarrow z_T(x) \rightarrow \xi(x) \]

\[ \alpha_\phi \quad \beta_\phi \quad \alpha_\xi \quad \beta_\xi \]

Prior

Exceedance Rate Model

\[ \alpha_\zeta \quad \beta_\zeta \]

Prior

Prior

Prior
Interpolation Method

Draw values from $[\phi(x) | \phi(x_1), \ldots, \phi(x_{56}), \alpha_\phi, \beta_\phi]$. 
# Exceedance Models Tested

<table>
<thead>
<tr>
<th>Baseline Model</th>
<th>$\bar{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
</table>
| Model 0: $\phi = \phi$  
$\xi = \xi$ | 112264.2 | 2.0 | 112266.2 |

<table>
<thead>
<tr>
<th>Models in Latitude/Longitude Space</th>
<th>$\bar{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
</table>
| Model 1: $\phi = \phi + \epsilon_\phi$  
$\xi = \xi$ | 98533.2 | 33.8 | 98567.0 |
| Model 2: $\phi = \alpha_0 + \alpha_1(msp) + \epsilon_\phi$  
$\xi = \xi$ | 98532.3 | 33.8 | 98566.1 |
| Model 3: $\phi = \alpha_0 + \alpha_1(elev) + \epsilon_\phi$  
$\xi = \xi$ | 98528.8 | 30.4 | 98559.2 |
| Model 4: $\phi = \alpha_0 + \alpha_1(elev) + \alpha_2(msp) + \epsilon_\phi$  
$\xi = \xi$ | 98529.7 | 29.6 | 98559.6 |

<table>
<thead>
<tr>
<th>Models in Climate Space</th>
<th>$\bar{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
</table>
| Model 5: $\phi = \phi + \epsilon_\phi$  
$\xi = \xi$ | 98524.3 | 27.3 | 98551.6 |
| Model 6: $\phi = \alpha_0 + \alpha_1(elev) + \epsilon_\phi$  
$\xi = \xi$ | 98526.0 | 25.8 | 98551.8 |
| **Model 7:** $\phi = \alpha_0 + \alpha_1(elev) + \epsilon_\phi$  
$\xi = \xi_{\text{mtn}}, \xi_{\text{plains}}$ | **98524.0** | **26.0** | **98550.0** |
| Model 8: $\phi = \alpha_0 + \alpha_1(elev) + \epsilon_\phi$  
$\xi = \xi + \epsilon_\xi$ | 98518.5 | 79.9 | 98598.4 |

$\epsilon \sim \text{MVN}(0, \Sigma)$ where $[\sigma]_{i,j} = \beta_{.0} \exp(-\beta_{.1} |x_i - x_j|)$
Posterior Distributions

**phi**

Density

<table>
<thead>
<tr>
<th>Value</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>0.05</td>
</tr>
<tr>
<td>3.5</td>
<td>0.05</td>
</tr>
<tr>
<td>3.6</td>
<td>0.10</td>
</tr>
<tr>
<td>3.7</td>
<td>0.15</td>
</tr>
<tr>
<td>3.8</td>
<td>0.05</td>
</tr>
<tr>
<td>3.9</td>
<td>0.05</td>
</tr>
</tbody>
</table>

**xi**

Density

<table>
<thead>
<tr>
<th>Value</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.4</td>
</tr>
<tr>
<td>0.10</td>
<td>0.8</td>
</tr>
<tr>
<td>0.15</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**beta_0 (Sill)**

Density

<table>
<thead>
<tr>
<th>Value</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
</tr>
</tbody>
</table>

**beta_1 (Range)**

Density

<table>
<thead>
<tr>
<th>Value</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.0</td>
</tr>
<tr>
<td>0.010</td>
<td>0.4</td>
</tr>
<tr>
<td>0.020</td>
<td>0.8</td>
</tr>
<tr>
<td>0.030</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Spatial Coherence of $\phi$
Traditional vs Climate Space

25-year Return Level Point Estimate

Longitude

-106.0
-105.0
37 38 39 40

Latitude

37 38 39 40

Ft. Collins
Greeley
Boulder
Denver
Colo Spgs
Pueblo
Results for Model 3: Return Level Uncertainty
• 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.