A hierarchical Gauss-Pareto model for extreme precipitation

Application to storms in southern Sweden

Robert A. Yuen[1] and Peter Guttorp[2]

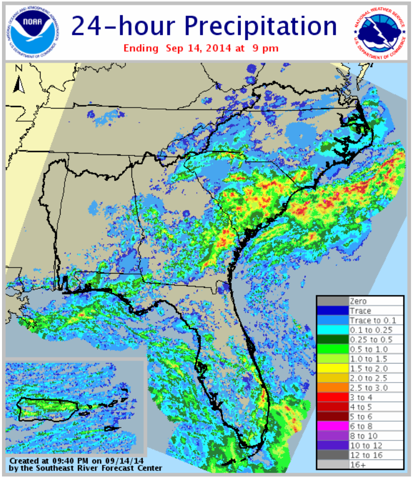
[1]University of Michigan [2]University of Washington

September 15th, 2014
UM Statistics Student Seminar
Ann Arbor, MI

Statistical modeling of extreme precipitation

Reasons for a stochastic model

- Make comparisons with weather models and remotely sensed data.
- Evaluate climate model output.
- Make predictions at unobserved sites.
- Inform decisions regarding soil saturation, landslides and potential flooding.



source: radar.weather.gov

Statistical modeling of extreme precipitation

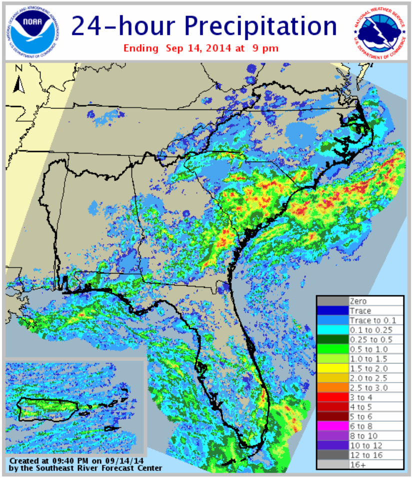
Let $\{W(s)\}_{s \in S}$ be a stochastic model for storms over region of interest $S \subset \mathbb{R}^2$.

Important characteristics of extreme precipitation data

- Non-smooth (non-differentiable).
- Left censored observations.
- Pareto tails.
- Non-trivial tail dependence:

$$\lim_{t \to \infty} \frac{P(\min\{W(s_1), W(s_2)\} > t)}{P(W(s_1) > t)} > 0.$$

Fact: Gaussian processes are tail independent.



source: radar.weather.gov

Statistical modeling of extreme precipitation

Let $\{W(s)\}_{s \in S}$ be a stochastic model for storms over region of interest $S \subset \mathbb{R}^2$.

Important characteristics of extreme precipitation data

- Non-smooth (non-differentiable).
- Left censored observations.
- Pareto tails.
- Non-trivial tail dependence:

$$\lim_{t \to \infty} \frac{P(\min\{W(s_1), W(s_2)\} > t)}{P(W(s_1) > t)} > 0.$$

Fact: Gaussian processes are tail independent.

A non-trivial model

$$V_i(s) := Z_i \exp\{B_i(s) - \gamma_i(s) + \varepsilon(s)\}$$

- Z_i - generalized Pareto distributed (GPD) with distribution function

$$G(z) = 1 - (1 + \xi z)_+^{-1/\xi},$$

  characterizes the overall intensity of the storm i.

- $\{B_i(s)\}_{s\in\mathbb{R}^2}$ - Gaussian process, independent of Z_i with semi-variogram

$$\gamma_i(s) = (\|s - \omega_i\|/\lambda_i)^{\alpha},\ \lambda > 0, \alpha \in (0,2),$$

  where $B_i(\omega_i) = 0$ a.s.

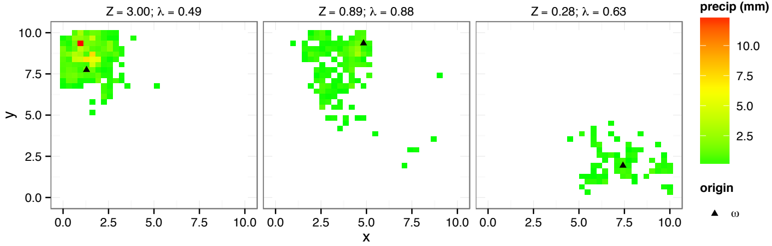
  - ω_i and λ_i are the *random* center and range of a storm i.
  - α controls smoothness of the storm profiles.

- $\{\varepsilon(s)\}_{s\in\mathbb{R}^2}$ - large scale trend surface, captures local effects such as elevation.

A non-trivial model

$$V_i(s) := Z_i \exp\{B_i(s) - (\|s - \omega_i\|/\lambda_i)^{\alpha} + \varepsilon(s)\}$$

Three realizations from the Gauss-Pareto model

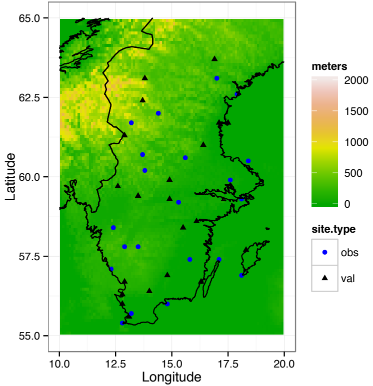


Values below 0.1mm have been censored.

Storms in south central Sweden

Extreme 24 hour precipitation data

- Swedish Meteorological and Hydrological Institute (www.smhi.se)
- Observations from 1961-2011 at 42 synoptic stations with 21 locations held out for validation.
- Select $n = 59$ independent extreme 24 hour precipitation events during summer months June, July and August.
  - Compute observed maximum for each date in record.
  - Select dates corresponding to top 5% of observed maxima.
  - Remove temporal clustering by selecting date of largest maxima within each cluster



Inference

Observe that

$$\begin{aligned} V_i(s) &:= Z_i \exp\{B_i(s) - (\|s-\omega_i\|/\lambda_i)^\alpha + \varepsilon(s)\} \\ Y_i(s) &:= \log V_i(s) \\ &\quad = B_i(s) - (\|s-\omega_i\|/\lambda_i)^\alpha + \log Z_i + \varepsilon(s) \end{aligned}$$

Hence, d-dimensional projections $\mathbf{Y}_i = (Y_i(s_1), \ldots, Y_i(s_d))$ *conditional on* Z_i are multivariate Gaussian with covariance

$$\Sigma(s_1, s_2) = \lambda^{-\alpha}\{\|s_1-\omega_i\|^\alpha + \|s_2-\omega_i\|^\alpha - \|s_1-s_2\|^\alpha\},$$

and mean

$$\mu_i(s) := \log Z_i + \varepsilon(s) - (\|s-\omega_i\|/\lambda_i)^\alpha.$$

Inference

An MCMC algorithm was developed to fit the model using three hierarchies

$$[\text{Data}] \quad \times \quad [\text{Process}] \quad \times \quad [\text{Prior}]$$

$$\prod_{i=1}^{n} p(\boldsymbol{Y}_i|Z_i, \varepsilon, \lambda_i, \omega_i, \alpha) \times p(Z_i|\xi)p(\varepsilon|\theta) \times p(\lambda_i, \omega_i, \alpha, \xi, \theta|\vartheta)$$

where

$$p(\boldsymbol{Y}_i|Z_i, \varepsilon, \lambda_i, \omega_i, \alpha) = N_d(\mu_i, \boldsymbol{\Sigma}_i).$$

$$p(Z|\xi) = GPD(1, 1, \xi).$$

$$p(\varepsilon|\theta) = N_d(0, C(\theta)).$$

Partial censoring

- Information from storm i is

$$\mathcal{D}_i = \{Y_i(s_j), j \in O_i\} \cup \{Y_i(s_j) \leq l, j \in C_i\},$$

where $C_i \subset \{1,\ldots,d\}$ denote censored locations and $O_i = \{1,\ldots,d\}\backslash C_i$.

- Likelihood contribution from storm i is

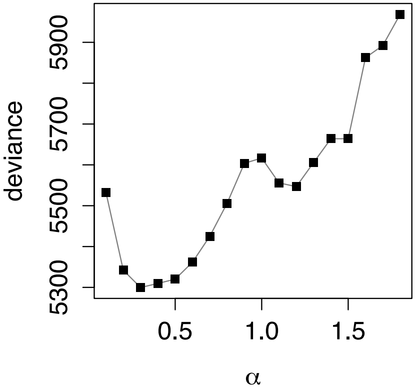
$$p(\boldsymbol{Y}_{O_i}|Z_i,\varepsilon,\lambda_i,\omega_i,\alpha)\int_{\boldsymbol{y}\leq l\boldsymbol{1}} p(\boldsymbol{y}|\boldsymbol{Y}_{O_i},Z_i,\varepsilon,\lambda_i,\omega_i,\alpha)d\boldsymbol{y}$$

where $p(\boldsymbol{Y}_{O_i}|Z_i,\varepsilon,\lambda_i,\omega_i,\alpha)$ and $p(\boldsymbol{y}|\boldsymbol{Y}_{O_i},Z_i,\varepsilon,\lambda_i,\omega_i,\alpha)$ are multivariate Gaussian densities derived from the model.

- Monte-Carlo integration is embedded in the Markov chain by sampling from the univariate truncated Gaussian at each iteration.

MCMC Results

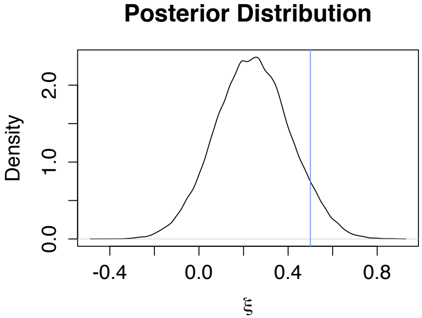
- Smoothness parameter α is difficult to estimate. Run multiple chains for $\alpha \in \{0.1, 0.2, \ldots, 1.9\}$, select best α based on *deviance score*.



MCMC Results

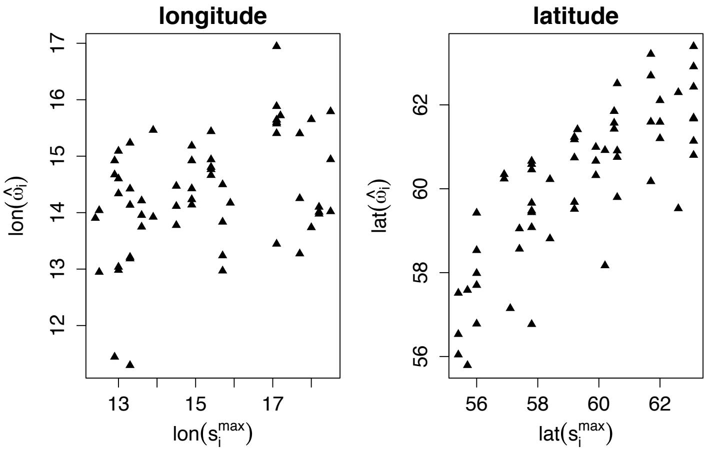
- Vague prior for ξ lead to poor mixing. Used informative prior

$$p(\xi) \propto \exp\left\{ -\frac{(\xi - 0.5)^2}{2(0.03)} \right\}.$$



Location of storms

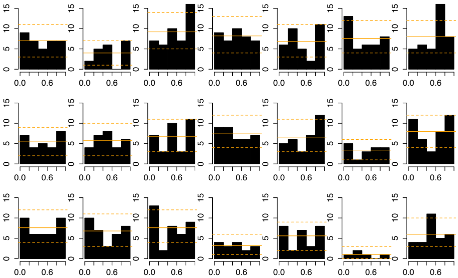
- $\hat{\omega}_i = \sum_{k=1}^{N} \omega_i^{(k)}$ estimated *center* of storm.
- $s_i^{max} = \underset{s \in \{s_1,\dots,s_d\}}{\arg\max}\ Y_i(s)$ location of observed maximum.



Prediction

- Posterior predictive distribution at unobserved location $\tilde{s}$ - sample from conditional distribution $p(Y_i(\tilde{s})|\mathbf{Y}_i, Z_i, \lambda_i, \omega_i, \alpha, \varepsilon)$ at each iteration.
- Evaluate predictions based on probability integral transform (PIT).

PIT histograms at 21 validation sites



Space-time extension

- Purely spatial model:
  - Limits in application.
  - Discarding of time dependent data.
- A Space-time extension:

$$V(s,t) = Z(t)\exp\{W(s,t) - \gamma(s,t)\}$$

  - Incorporate more information of observations clustered in time.
  - More challenging to model the time dependence in $Z(t)$.

Thanks to Stilian Stoev and Veronica Berrocal for many ideas and helpful discussions.