Modeling Both Climate and Weather Spatial Effects for Extreme Precipitation

Dan Cooley
Department of Statistics
Colorado State University

Joint work with:
Mathieu Ribatet, Université Montpellier II
Anthony Davison, EPFL
Steve Sain, NCAR

Work supported in part by NSF-DMS 0905315

Outline

1. Introduction
   - Statistics for Extremes
   - Climate and Weather
2. Modeling Climate Spatial Effects for Extremes
   - Historical Methods
   - Bayesian Hierarchical Models
3. Models for Weather Spatial Effects for Extremes
   - Max-stable Processes
   - Fitting Max-stable Processes: Composite Likelihoods
4. Modeling Both Climate and Weather Effects
   - Possible Approaches
   - Employing a Max Stable Process Model
   - Appropriate Bayesian Inference for Composite L'hoods
   - Implementation in a BHM and Results

Extreme Value Analyses

General Idea: Distributions suggested by the asymptotic theory are fit using only data considered to be extreme.

Univariate approaches:

1. *Block maximum data*
   - GEV(μ, σ, ξ)
2. *Threshold exceedance data*
   - GPD($\tilde{\sigma}_u$, ξ, ζ)
   - PPE(μ, σ, ξ)

Things we know:

- ξ determines the type of tail and is difficult to estimate.
- always data poor in extreme value analyses.
- risk often summarized in terms of return levels.

Extreme Values in more than one dimension

Multivariate Extremes

- Theory well developed.
- Some existing parametric models for moderate dimension.
- Block maximum and threshold exceedance approaches have both been used.

Extremal Processes

- Theory well developed.
- Some existing models for max-stable processes.
- Application to threshold exceedances (Huser and Davison, 2012)
- Although process models ($d = \infty$), only the bivariate (Smith: trivariate) joint distributions known in closed form.

In practice, marginal effects and dependence structure are handled separately, often transformed to RV(1).

Spatial Extremes, Climate, and Weather

There are two spatial effects at work in the data: climate and weather effects.

Q: What's the difference?

Climate vs. Weather

"Climate is what you expect, weather is what you get"

...but this doesn't really apply when doing an EV analysis.

Instead: Climate is the *distribution* from which weather is drawn (not just the expected value).

Climate and Weather Spatial Effects

Spatial dependence in data from two sources:

- local dependence due to events that effect more than one location (weather).
- regional dependence due to similar characteristics between locations (climate).

In terms of a statistical model:

climate effects: how the marginal distribution changes with location.

weather effects: the joint behavior of multiple locations.

Outline

1. Introduction
   - Statistics for Extremes
   - Climate and Weather
2. Modeling Climate Spatial Effects for Extremes
   - Historical Methods
   - Bayesian Hierarchical Models
3. Models for Weather Spatial Effects for Extremes
   - Max-stable Processes
   - Fitting Max-stable Processes: Composite Likelihoods
4. Modeling Both Climate and Weather Effects
   - Possible Approaches
   - Employing a Max Stable Process Model
   - Appropriate Bayesian Inference for Composite L'hoods
   - Implementation in a BHM and Results

Goal of Modeling Climate Extremes

Precipitation Atlases:

NOAA's provides "official" estimates for extreme precipitation for US locations. `http://www.nws.noaa.gov/ohd/hdsc/noaaatlas2.htm`

Colorado 40°N 105°W (Boulder)

Map	Prcp (inches)	Prcp Intensity (in/hr)
2-year 6-hour	1.49	0.25
2-year 24-hour	2.13	0.09
100-year 6-hour	3.68	0.61
100-year 24-hour	5.06	0.21

Estimates based on NOAA Atlas 2, 1973.

- Precip. atlases used as inputs to hydro models.
- Other uses for climatological extremes:
  - Wind: structural design of buildings, power.
  - Sea level: sea walls, oil-drilling platforms.
  - Climate models: how will extreme precip change with climate change?

Historical Approach for Modeling Climate Extremes

Regional Frequency Analysis

- Traces roots to Dalrymple (1960).
- Thorough treatment of modern practice: Hosking and Wallis (1997)
- Basic idea is "trading space for time"
- Basic steps:
  1. Define and test homogeneous regions
  2. Normalize (block maximum) data using "index flood" – often a mean of annual maxima.
  3. Combine data and estimate parameters (L-moments)
  4. Un-normalize data by index flood to get estimates at each location
- NOAA using a variant of RFA to update its maps.

Recent Approach: Spatial Hierarchical Models

Basic idea: Assume there is a latent spatial process that characterizes the behavior of the data over the study region.

- much of early work done in epidemiology
- Diggle et al. (1998)
- Banerjee et al. (2004)
- Climate/weather applications: Cressie, Wikle, Berliner, others

Why bother?

- Latent process too complex to capture with fixed effects; covariates not rich enough.
- Intelligently borrows strength

Spatial HM General Framework

Bayesian formulation, three levels.

$$\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{y}) \propto \underbrace{\pi(\boldsymbol{y}|\boldsymbol{\theta}_1)}_{data} \underbrace{\pi(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2)}_{process} \underbrace{\pi(\boldsymbol{\theta}_2)}_{prior},$$

Data level: Likelihood characterizing the observed data *given the parameters at the process level*.
Process level: Latent process captured by spatial model for the data level parameters.
Prior level: Prior distributions put on the parameters in the process level.

Typically an assumption of *conditional independence* made at the data level and joint likelihood is the product of the individual likelihoods at each location.

Sensible for epidemiology, perhaps not for weather data.

Overview of a Case Study: Regional Climate Model Precipitation Extremes

(Cooley and Sain (2010))

Data: Output from a RCM for the western US, $\sim$ 2500 locations. Both control and future runs modeled simultaneously.

$$\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{y}) \propto \underbrace{\pi(\boldsymbol{y}|\boldsymbol{\theta}_1)}_{data} \underbrace{\pi(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2)}_{process} \underbrace{\pi(\boldsymbol{\theta}_2)}_{prior},$$

Data Level: Point process model for threshold exceedances. Conditional independence assumed.

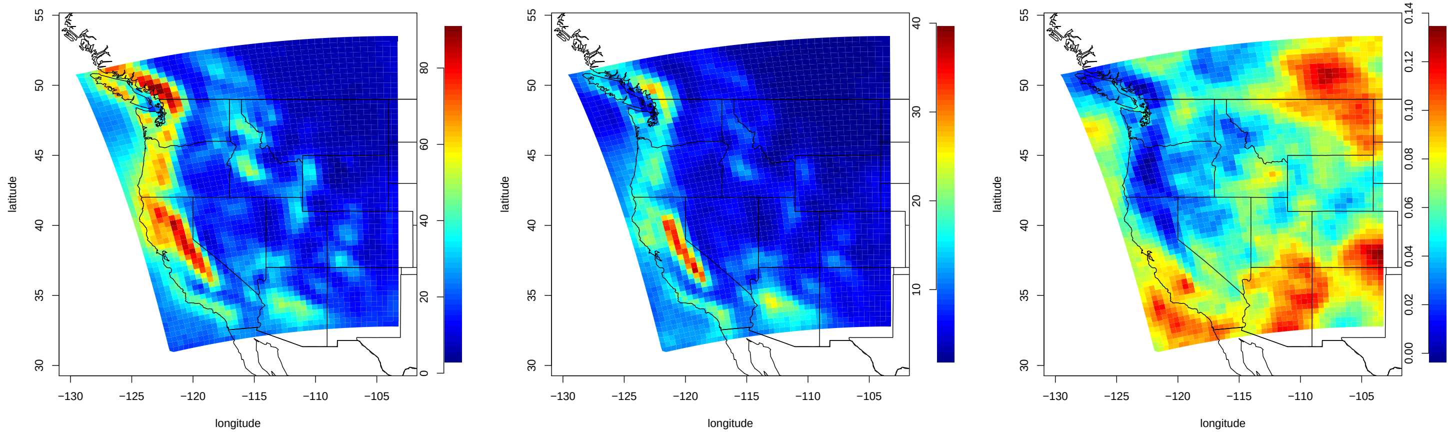
Process Level: Multivariate IAR model for $(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\xi})$.

- Q has dimension 14784 x 14784.
- 29598 (non-indep) parameters, effective number $\sim$ 4250
- Inference via Gibbs Sampler

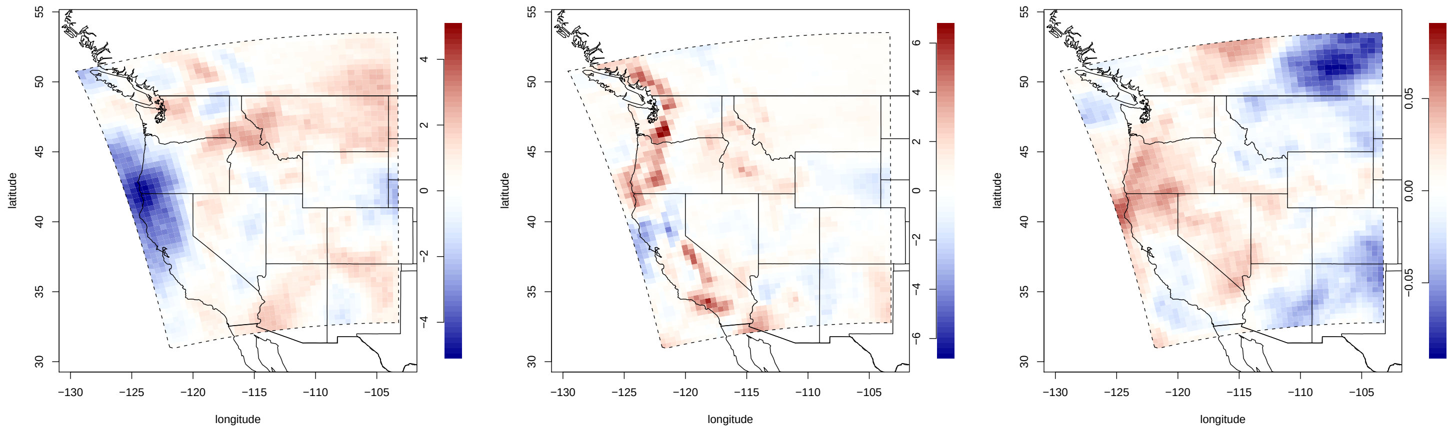
Parameter Estimates (Posterior Mean, Winter)

Control

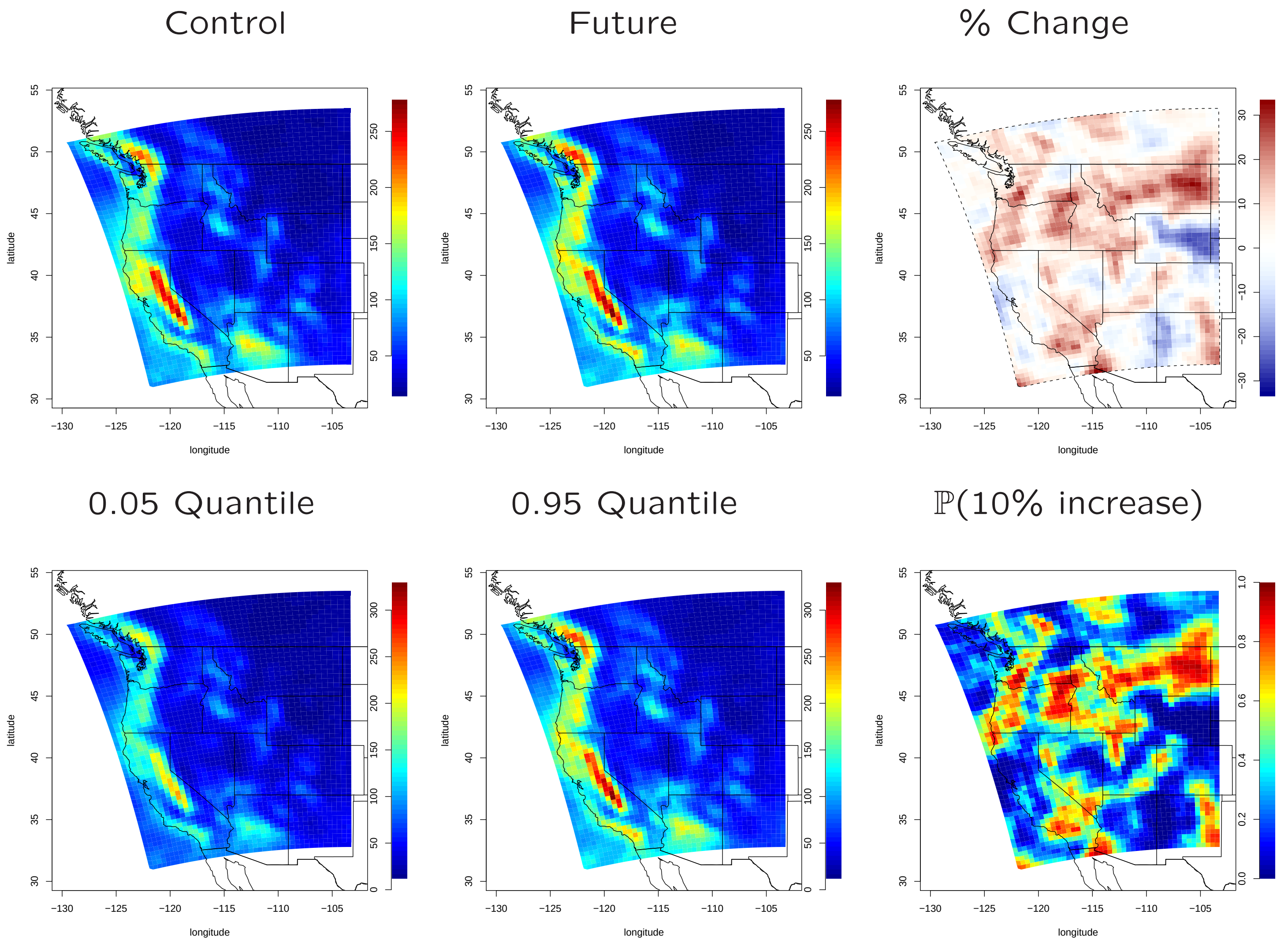
μ σ ξ



Future - Control

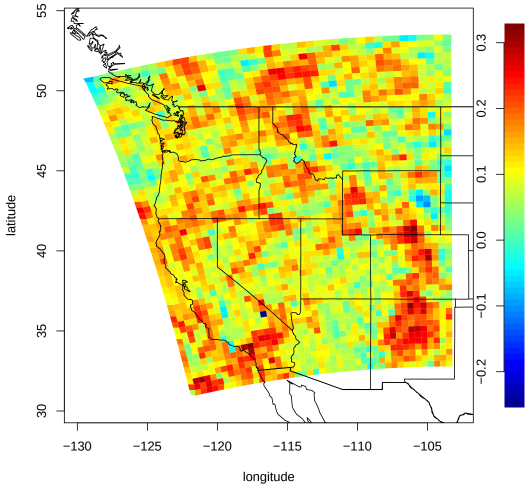


Risk Assessments: 100-year Return Levels

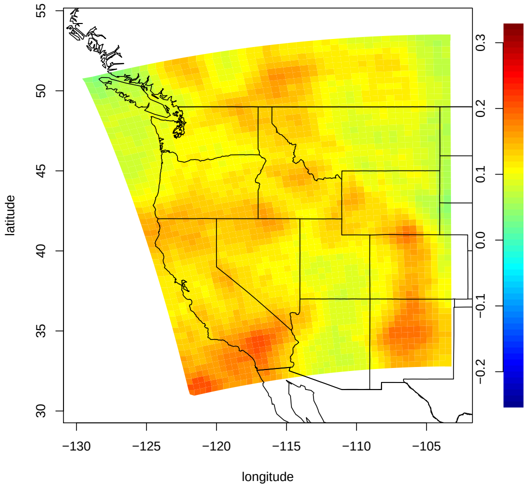


What is gained from BHM?

ξ MLE



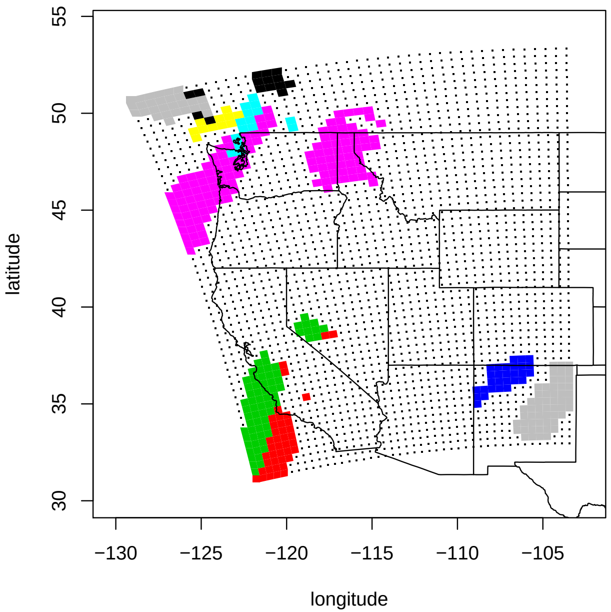
ξ BHM



- seems to capture *climate effects*
- latent spatial model needed
  - available covariates inadequate to capture effects
  - borrows strength across locations for estimates
- uncertainty easily obtained from MCMC runs

What about the weather effects?

Spatial extent of intense storms in year 1



Clearly, conditional independence assumption does not hold.

Q1: How to model extreme weather events' spatial effects?
Q2: If one is *only* interested in climate extremes questions, does one need to worry about ignoring the weather effects?

Outline

Max-stable Processes

Let $Y_m(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{D}, m = 1, \ldots, n$ be independent copies of $Y(\boldsymbol{x})$, and let $M_n(\boldsymbol{x}) = \max_m Y_m(\boldsymbol{x})$. $Y(\boldsymbol{x})$ is termed max-stable if there exist $a_n(\boldsymbol{x})$ and $b_n(\boldsymbol{x})$ such that

$$\mathbb{P}\left(\frac{M_n(\boldsymbol{x}) - b_n(\boldsymbol{x})}{a_n(\boldsymbol{x})} \leq y(\boldsymbol{x})\right) = \mathbb{P}(Y(\boldsymbol{x}) \leq y(\boldsymbol{x})).$$

- Max-stability is foundation of EVT.
- Limit process for site-wise maxima.
- Use is justified by extreme value theory.
- Suitable for modeling fields of block (annual) maximum data.
- From an EV perspective, max-stable processes are the correct answer to Q1.
- A few models have been proposed.

Point process formulation of max-stable processes

(de Haan and Ferreira, 2006, Corollary 9.4.5)

$Z(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{D}$ is max-stable with unit Fréchet marginals

iff

There exist iid positive stochastic processes $V_1(\boldsymbol{x}), V_2(\boldsymbol{x}), \ldots$ with $E[V_i(\boldsymbol{x})] = 1$ for all $\boldsymbol{x} \in \mathcal{D}$ and $E[\sup_{\boldsymbol{x} \in \mathcal{D}} V(\boldsymbol{x})] \leq \infty$ and an independent point process $\{\eta_i\}_0^\infty$ on $(0, \infty]$ with intensity measure $r^{-2}dr$ such that $Z(\boldsymbol{x}) \stackrel{d}{=} \max_{i=1,2,\ldots} \eta_i V_i(\boldsymbol{x})$.

Models for Max-stable Processes

- Smith (1990)
- Schlather (2002)
- de Haan and Ferreira (2006)
- Kabluchko et al. (2009)

Smith Model for a Max-stable Process

$$\Pr[Z(x_i) \le z_1 \quad , Z(x_j) \le z_2] = \exp\left\{ -\frac{1}{z_1}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_2}{z_1}\right) - \frac{1}{z_2}\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{z_1}{z_2}\right) \right\}$$

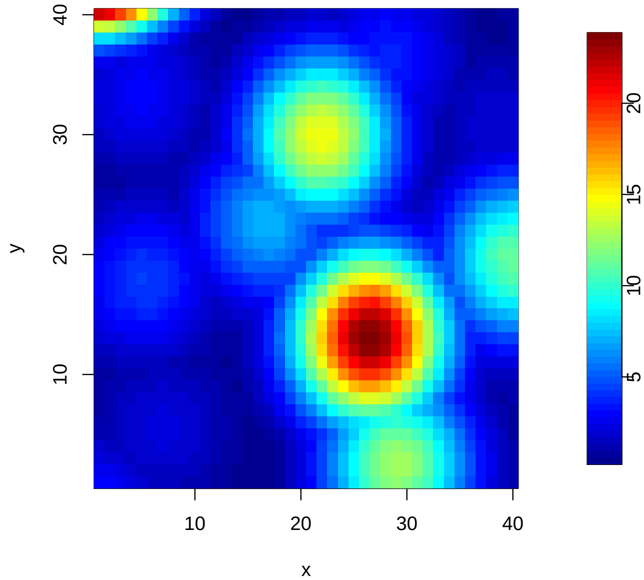
where $a^2 = (x_i - x_j)^T \Sigma^{-1} (x_i - x_j)$.

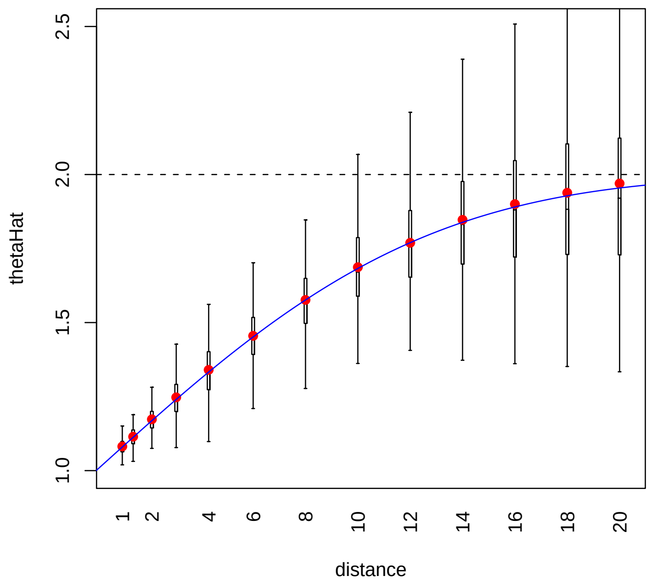
- Characterized by the parameter $\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_{2,2} \end{bmatrix}$
- Assumes the marginals are unit Fréchet: $\mathbb{P}(Z(x_i) \le z) = \exp(-z^{-1})$.
- If marginals not Fréchet, then a marginal transformation can be performed.
- Has a "storm" interpretation.
- We use this in the data level of our hierarchical model described later.

Smith Model for a Max-stable Process

$$\Pr[Z(x_i) \le z_1 \quad, Z(x_j) \le z_2] = \exp\left\{-\frac{1}{z_1}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_2}{z_1}\right)-\frac{1}{z_2}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_1}{z_2}\right)\right\}$$

where $a^2 = (x_i - x_j)^T \Sigma^{-1}(x_i - x_j)$.

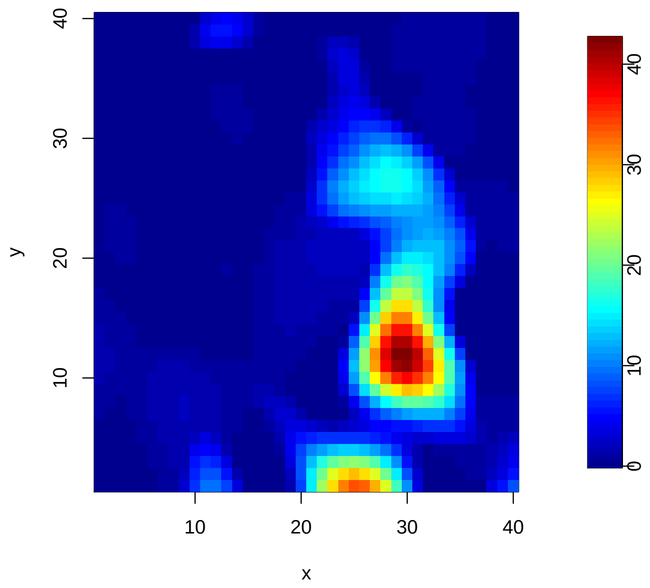


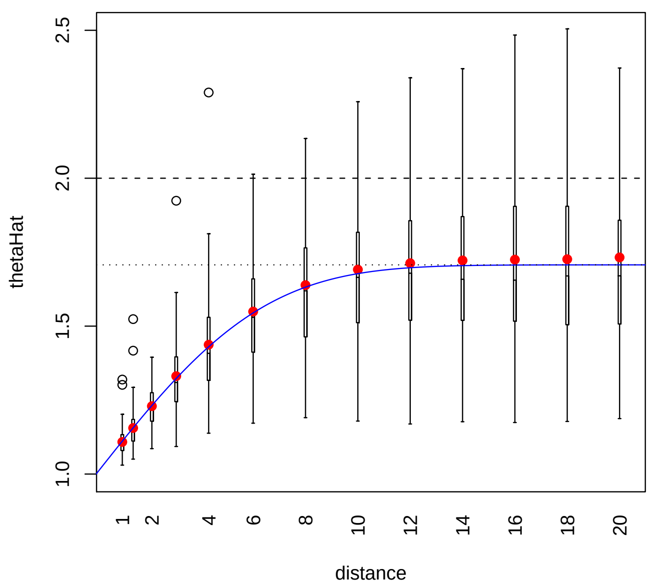


Schlather (2002) Model for a Max-stable Process

$$\Pr[Z(x_i) \le z_1, Z(x_j) \le z_2] =$$
$$\exp\left\{ -\frac{1}{2}\left(\frac{1}{z_1} + \frac{1}{z_2}\right)\left(1 + \sqrt{1 - 2(\rho(z_1 - z_2) + 1)\frac{z_1 z_2}{(z_1 + z_2)^2}}\right)\right\}$$

where ρ is the correlation function of the standard normal Gaussian process.





Brown-Resnick Model for a Max-stable Process

Kabluchko et al. (2009)

$$\Pr[Z(x_i) \le z_1, Z(x_j) \le z_2] =$$
$$\exp\left\{ - \left[e^{-z_1}\Phi\left(\frac{\sqrt{\rho(x_i - x_j)}}{2} + \frac{z_2 - z_1}{\sqrt{\rho(x_i - x_j)}}\right) + e^{-z_2}\Phi\left(\frac{\sqrt{\rho(x_i - x_j)}}{2} + \frac{z_1 - z_2}{\sqrt{\rho(x_i - x_j)}}\right)\right]\right\}$$

- Gives realistic-looking realizations.
- As distance increases, observations become independent.
- Interesting construction from Brownian motion with drift– the stationary process is constructed as a maximum of non-stationary processes.

Fitting Max-stable Processes

For the max-stable process models, only the *bivariate* distributions are known (trivariate for Smith model?). How does one fit a model to K observations?

A: Composite Likelihoods

Composite likelihoods are used to obtain estimating equations when the true likelihood is too difficult or impossible to obtain (Lindsay, 1988).

Pairwise Likelihoods

Since we have the bivariate distributions we will use the *pairwise* likelihood.

Assume independent observations $\boldsymbol{y}_m = (y_m^{(1)}, y_m^{(2)}, \ldots, y_m^{(k)})^T$, $m = 1, \ldots, m$ arise from a probability model with density $f(\boldsymbol{y}; \theta)$ which has bivariate marginals $f(y_m^{(i)}, y_m^{(j)}; \theta)$. Then the pairwise log-likelihood is given by

$$\ell_p(\theta; \mathbf{y}) = \sum_{m=1}^{n} \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \log f(y_m^{(i)}, y_m^{(j)}; \theta).$$

Things to keep in mind:

- Not a true likelihood.
- Over-uses the data – each observation appears $K-1$ times.

Frequentist Methods for Composite Likelihoods

Point estimation achieved by maximizing the composite likelihood; MCLE denoted $\widehat{\theta}_c$.

1. Estimation is unbiased.
2. Uncertainty estimates achieved via the information sandwich approach.

$$\sqrt{n}\{H(\theta_0)J(\theta_0)^{-1}H(\theta_0)\}^{1/2}(\widehat{\theta}_c - \theta_0) \longrightarrow N(0, \mathrm{Id}_p),$$

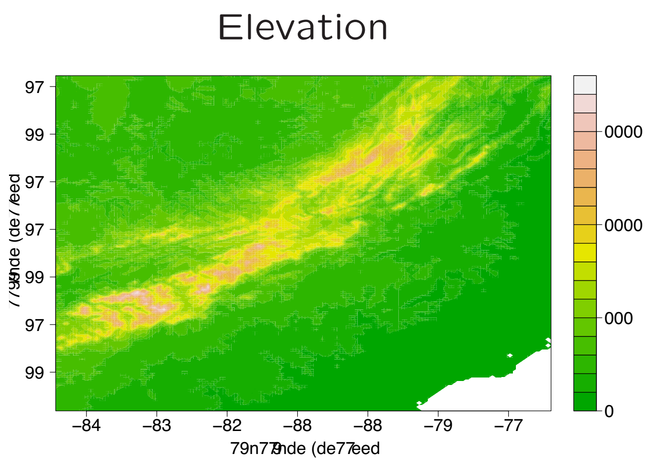
where $H(\theta_0) = \mathbb{E}[\nabla^2 \ell_c(\theta_0; Y)]$ and $J(\theta_0) = \mathrm{Var}[\nabla \ell_c(\theta_0; Y)]$.

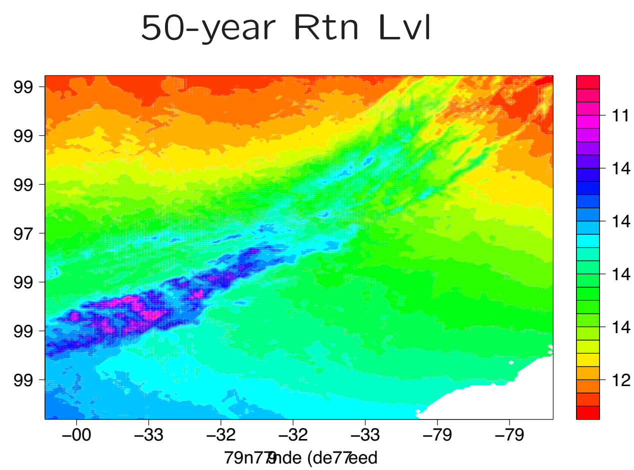
Pairwise Likelihoods for Max-stable Processes

(Padoan, Ribatet, and Sisson, 2010)

Use pairwise likelihood approach for annual maximum precipitation data.

Marginals: trend surface of longitude, latitude, and elevation.





Findings:

1. Improved modeling of joint occurrence.

2. Some reason to question the fit of marginal distributions.

Outline

1. Introduction
    - Statistics for Extremes
    - Climate and Weather
2. Modeling Climate Spatial Effects for Extremes
    - Historical Methods
    - Bayesian Hierarchical Models
3. Models for Weather Spatial Effects for Extremes
    - Max-stable Processes
    - Fitting Max-stable Processes: Composite Likelihoods
4. Modeling Both Climate and Weather Effects
    - Possible Approaches
    - Employing a Max Stable Process Model
    - Appropriate Bayesian Inference for Composite L'hoods
    - Implementation in a BHM and Results

Modeling *Both* Climate and Weather Effects

SHM Revisited

$$\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{y}) \propto \underbrace{\pi(\boldsymbol{y}|\boldsymbol{\theta}_1)}_{data} \underbrace{\pi(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2)}_{process} \underbrace{\pi(\boldsymbol{\theta}_2)}_{prior},$$

- To capture the weather effects, one needs to capture the dependence among observations due to weather events.
- Would have spatial model in *both* data and process levels.

Possible Approaches for Likelihoods

- Multivariate Extreme Value Models
- Copula Approaches
- Max-stable Process Models

Possible Approaches for Likelihoods

- Multivariate Extreme Value Models
  + Models can be applied to both block maxima or threshold exceedance data.
  – Existing models (e.g., Tawn (1990); Cooley et al. (2010)) applicable only to data of relatively low dimension ($d \sim 5$).
- Copula Approaches
- Max-stable Process Models

Possible Approaches for Likelihoods

- Multivariate Extreme Value Models
- Copula Approaches
  - Gaussian copula (Sang and Gelfand, 2010).
    - + Very natural, intuitive approach.
    - + Computationally feasible.
    - ∘ Leads to an asymptotically independent model.

$$\lim_{u \to \infty} \mathbb{P}(X_i > u | X_j > u) = 0$$

  - Dirichlet process (Fuentes et al., 2009)
    - + Both asymptotic dependence and asymptotic independence possible.
    - ∘ Unclear ties to EVT.
- Max-stable Process Models

Possible Approaches for Likelihoods

- Multivariate Extreme Value Models
- Copula Approaches
- Max-stable Process Models

Capturing the Weather Effects with a Max-stable Process

$$\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{y}) \propto \underbrace{\pi(\boldsymbol{y}|\boldsymbol{\theta}_1)}_{data} \underbrace{\pi(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2)}_{process} \underbrace{\pi(\boldsymbol{\theta}_2)}_{prior},$$

We aim to employ a likelihood from a max-stable process at the data level.

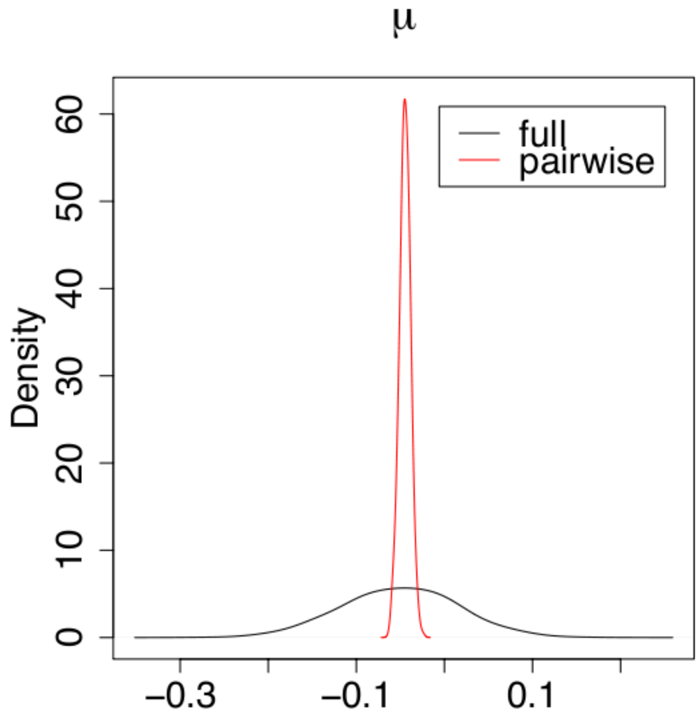
Challenge:

- Since only the bivariate distributions are known, we cannot use the correct likelihood $\rightarrow$ composite likelihood.
- Frequentist methodology exisits for composite likelihoods, can ideas be extended to obtain appropriate inference in a Bayesian setting?

Bayesian Methods for Composite Likelihoods?

posterior $\propto$ likelihood * prior

If true likelihood is replaced by pairwise likelihood, resulting posterior is too narrow.



50 replicates of a mean-zero Gaussian process with exponential covariance structure, observed at 20 locations

Appropriate Bayesian Inference

Ribatet, Cooley, and Davison (2009) investigate appropriate Bayesian inference for composite likelihoods by adjusting likelihood to obtain an appropriate posterior.

Two approaches suggested:

1. Magnitude adjustment (Also: Pauli et al. (2011))
2. Curvature adjustment

As the curvature adjustment can be shown to perform better, we focus on it here.

Curvature Adjustment

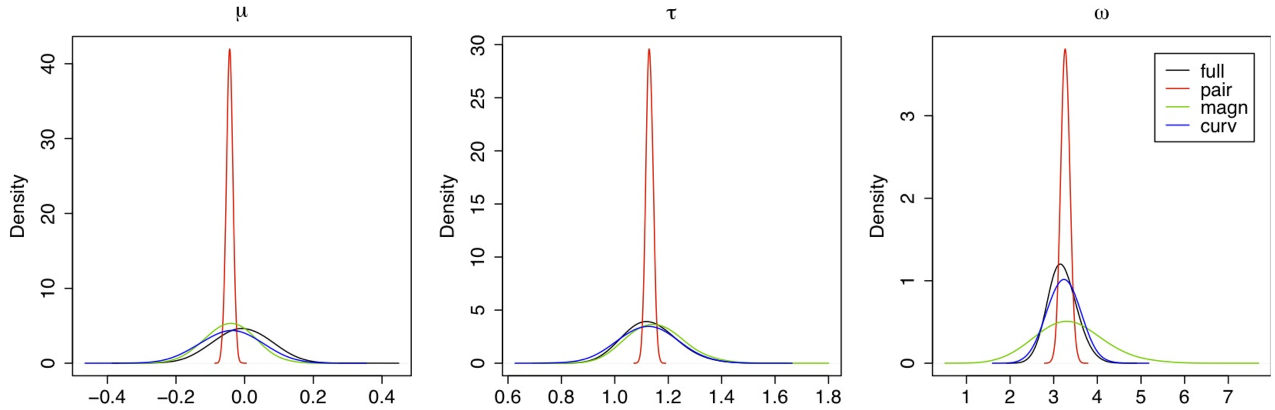
In the context of performing likelihood ratio tests (frequentist), Chandler and Bate (2007) suggest replacing the composite likelihood with an adjusted inference function:

$$\ell_A(\boldsymbol{\theta}; y) = \ell_c(\boldsymbol{\theta}^*; y),$$
$$\boldsymbol{\theta}^* = \widehat{\boldsymbol{\theta}}_c + M^{-1}M_A(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_c),$$
$$\text{where } M^T M = -H(\boldsymbol{\theta}_0),\ M_A^T M_A = H(\boldsymbol{\theta}_0)J(\boldsymbol{\theta}_0)^{-1}H(\boldsymbol{\theta}_0).$$

- $\ell_A(\widehat{\boldsymbol{\theta}}_c) = \ell_c(\widehat{\boldsymbol{\theta}}_c)$.
- the Hessian of $\ell_A(\hat{\boldsymbol{\theta}}_{\boldsymbol{c}}; y)$ is $H_A(\boldsymbol{\theta}_0) = H(\boldsymbol{\theta}_0)J(\boldsymbol{\theta}_0)^{-1}H(\boldsymbol{\theta}_0)$.
- Note: to calculate $\ell_A(\boldsymbol{\theta})$, $\widehat{\boldsymbol{\theta}}_c$ must be known.

Gaussian Process Simulation Results

Truth: $\mu = 0$, $\tau = 1$, $\omega = 3$



Coverage

	Full			Magnitude			Curvature			Unadjusted		
	μ	τ	ω	μ	τ	ω	μ	τ	ω	μ	τ	ω
$\omega = 3$	96	94	94	89	92	100	94	93	94	16	21	37
$\omega = 1.5$	94	95	96	85	93	100	94	94	93	19	22	53

Implementation in BHM models

$$\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{y}) \propto \underbrace{\pi(\mathbf{y}|\boldsymbol{\theta}_1)}_{data} \underbrace{\pi(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2)}_{process} \underbrace{\pi(\boldsymbol{\theta}_2)}_{prior},$$

Idea is straightforward: replace the true (but unattainable) likelihood with the appropriately adjusted pairwise likelihood.

$\ell_A(\boldsymbol{\theta})$ used in MH ratio, detailed balance condition met.

Implementation not straightforward: adjusting the likelihood requires knowledge of the MCLE's.

Inference for BHM models obtained via Gibbs samplers, we examine two approaches:

1. Overall Gibbs sampler: MCLE's and matrices $H(\hat{\boldsymbol{\theta}}_c)$ and $J(\boldsymbol{\theta}_c)$ obtained initially and used throughout.
2. Adapted Gibbs sampler: Data-level parameters must be drawn *conditional on the current values* of process- and prior-level parameters.

Adaptive Gibbs Sampler

Algorithm 1: Adaptive Gibbs sampler

Input : $\theta_0 \in \Theta$
Output: A Markov chain of length $N + 2$

for $t \leftarrow 0$ **to** N **do**
    for $i \leftarrow 1$ **to** K **do**
        Update $\{\widehat{\mu}(x_i), \widehat{\sigma}(x_i), \widehat{\xi}(x_i)\}$ via numerical optimization;
        Update $\{\mu(x_i), \sigma(x_i), \xi(x_i)\}$ via MH;
    end
    Update $\widehat{\Sigma}$ via numerical optimization;
    Update Σ via M.–H.;
    for $p \in \{\mu, \sigma, \xi\}$ **do**
        Update $\boldsymbol{\beta}_p$ directly;
    end
    for $p \in \{\mu, \sigma, \xi\}$ **do**
        Update τ_p directly;
    end
    for $p \in \{\mu, \sigma, \xi\}$ **do**
        Update ω_p directly;
    end
    Store all the updated values in θ_{t+1};
end
return $\{\theta_t\}_{t=0,\dots,N+1}$;

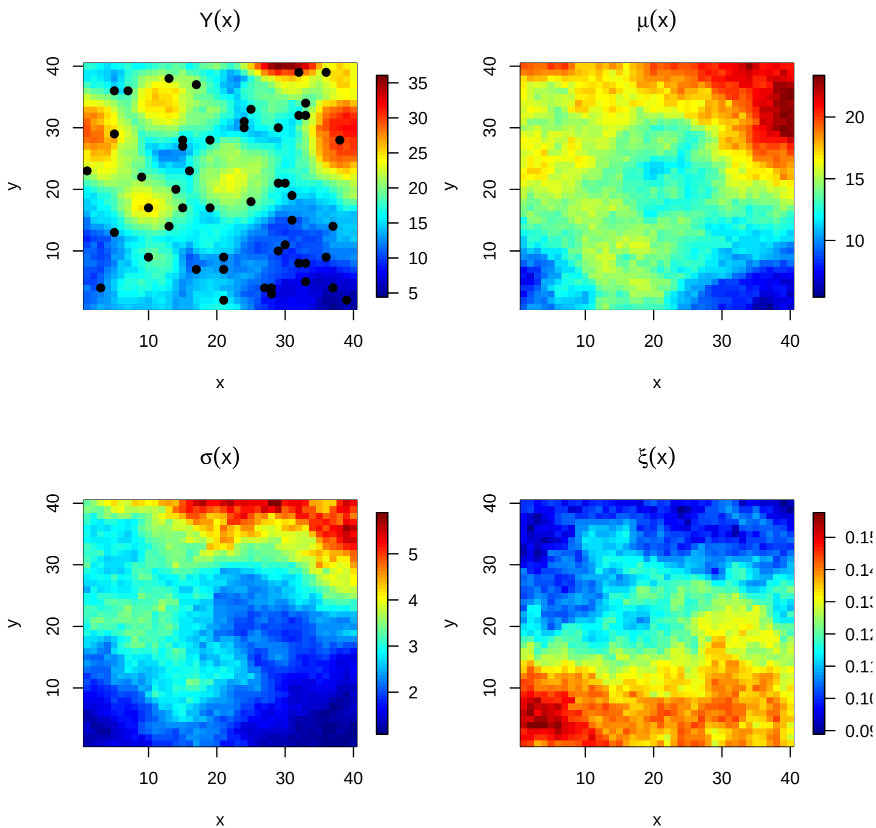
Adaptive Gibbs Sampler

- Numerical optimization at steps 3 and 6 makes things *very* slow
- Have shown via a completion argument that adaptive Gibbs sampler does converge
- However, argument does not account for *numerical* optimization
- Convergence checked using trace plots and the $\sqrt{\widehat{R}}$ statistic – all indicate convergence achieved

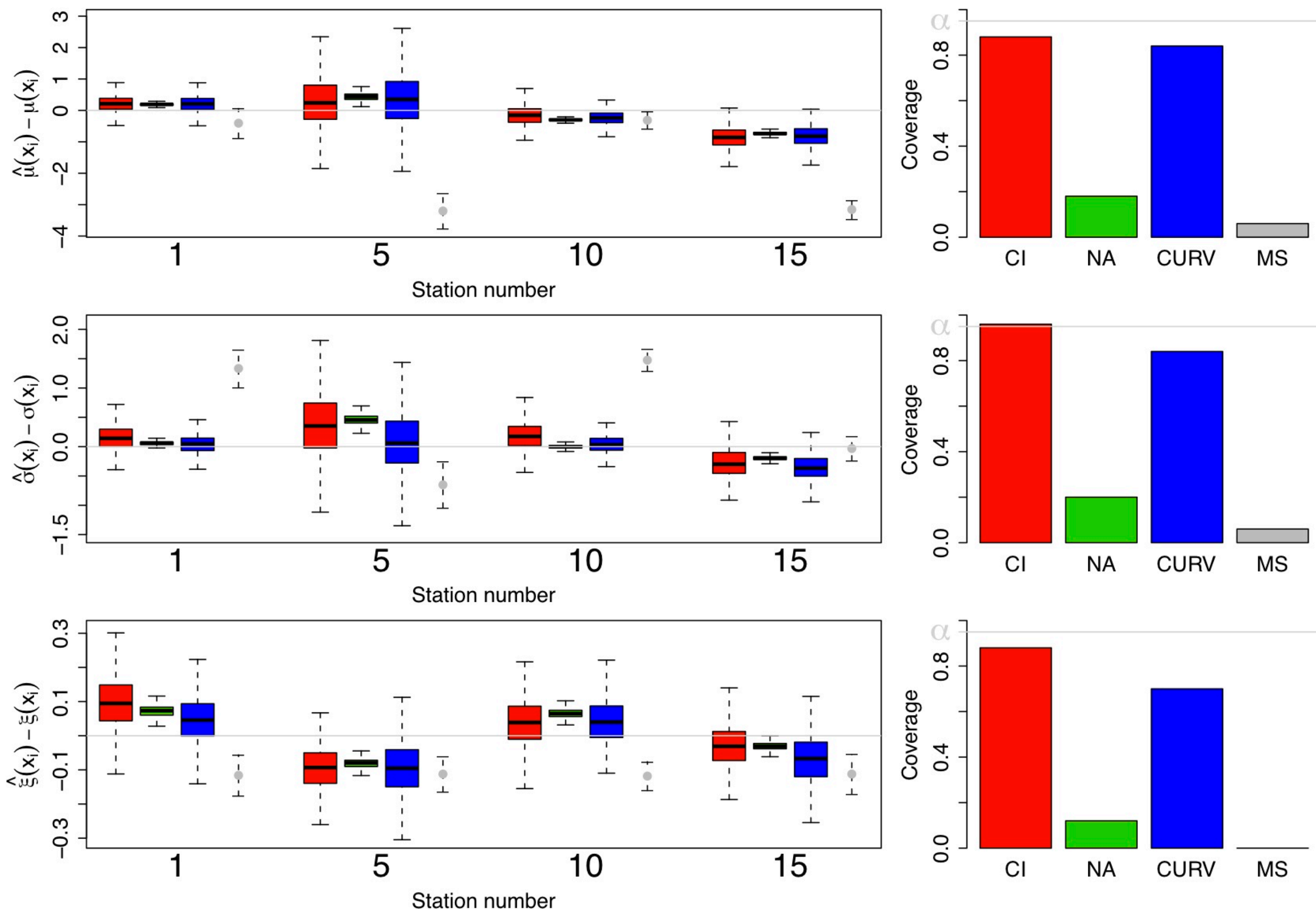
Simulation Study

$Z_j(x); j = 1, \ldots 50$ realizations of Smith's ms process. (weather)

Marginals: $Y(x) \sim GEV(\mu(x), \sigma(x), \xi(x))$;
$\mu(x), \sigma(x), \xi(x)$ are Gaussian process realizations. (climate)



Simulation Results: Estimating Marginals



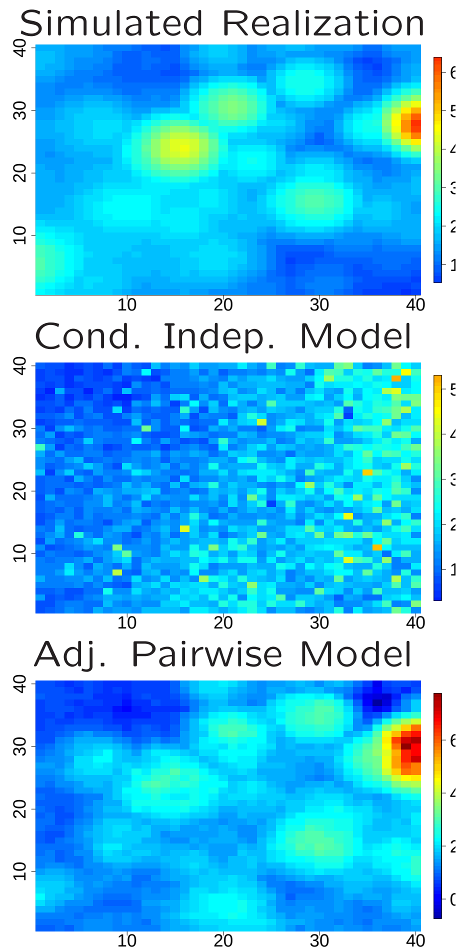
Simulation Results: Weather Dependence

Given by:

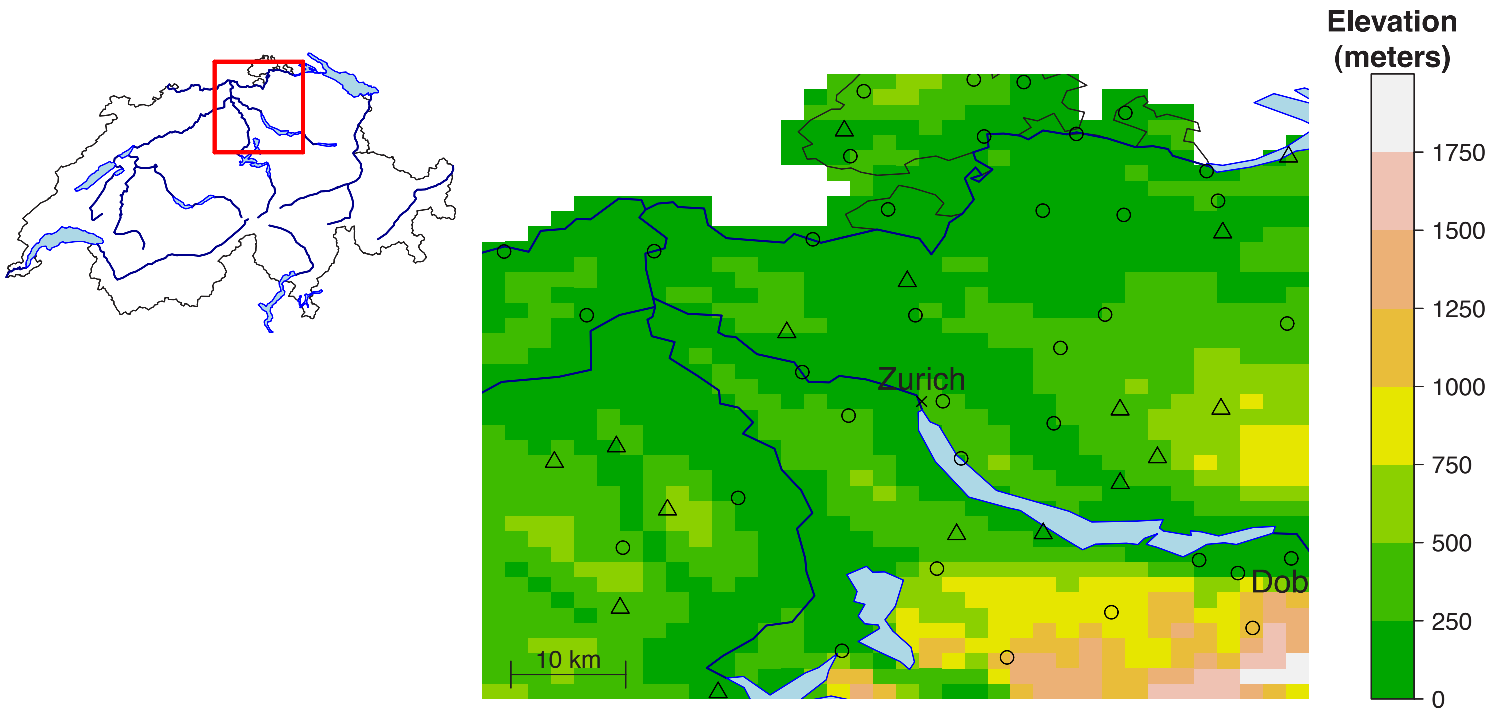
$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_{2,2} \end{bmatrix}$$

Estimates

	Simulated	95% Cred. Int.
$\sigma_{1,1}$	6	(5.39, 8.76)
$\sigma_{1,2}$	0	(-1.28, 0.67)
$\sigma_{2,2}$	6	(5.58, 8.37)



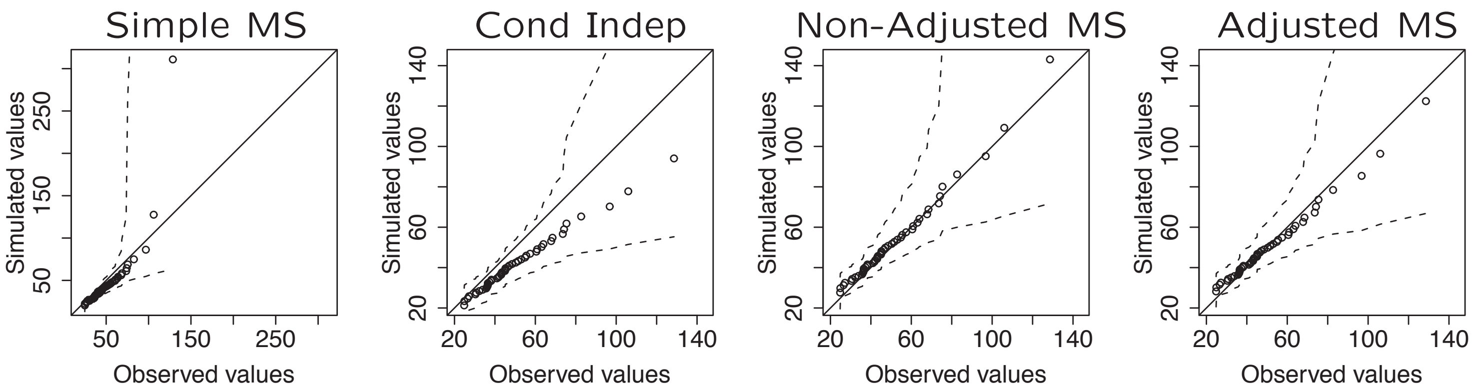
Application: Switzerland Precipitation



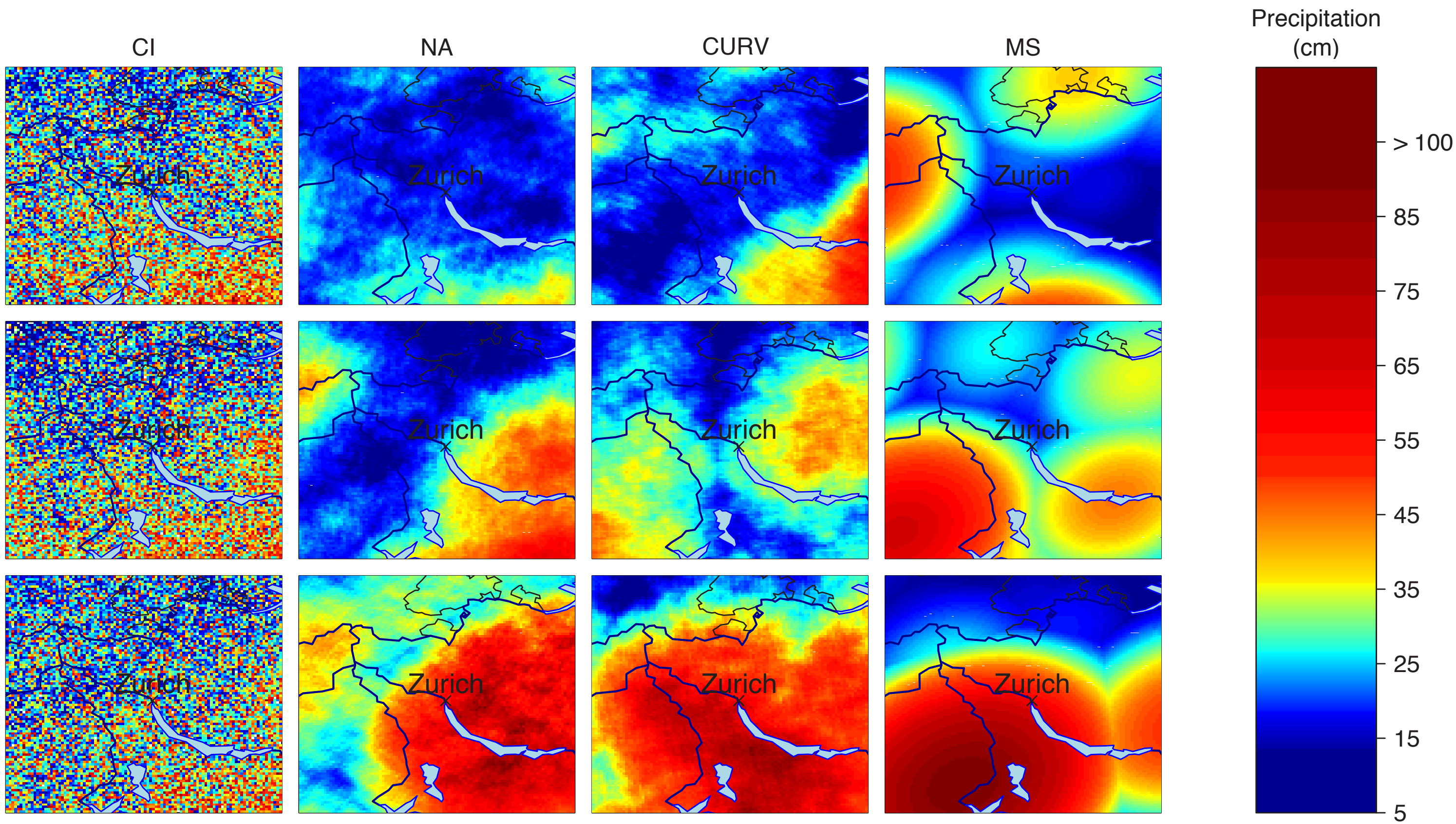
- 50 years of annual maximum data
- 35 stations used for model fitting
- 16 used for model validation

Results: Estimating Marginals

QQ plots of the annual max data at a validation station verses the model's estimated marginal.



Results: Describing Joint Behavior



Summary of this Study

By applying a max-stable model in the data level of a hierarchical model, we can account for both weather and climate spatial effects in extreme observations.

- Needed to employ a pairwise likelihood approach to fit the max-stable model.
- Needed to make adjustments to obtain appropriate inference.

+ Model at data level is grounded in extreme value theory.
+ Hierarchical approach allows for flexible modeling of climatological effects.
+ Bayesian approach simultaneously estimates marginal and dependence effects.

– As currently implemented, method is computationally feasible only for relatively small spatial problems.
  - Many avenues for speeding up computation.

References

Banerjee, S., Carlin, B., and Gelfand, A. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, Boca Raton, FL.

Chandler, R. and Bate, S. (2007). Inference for clustered data using the independence loglikelihood. *Biometrika*.

Cooley, D., Davis, R. A., and Naveau, P. (2010). The pairwise beta: A flexible parametric multivariate model for extremes. *Journal of Multivariate Analysis*, 101:2103–2117.

Cooley, D., Nychka, D., and Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102:824–840.

Cooley, D. and Sain, S. R. (2010). Spatial hierarchical modeling of precipiation extremes from a regional climate model. *Journal of Agricultural, Biological, and Envrionmental Statistics*, 15:381–402.

Dalrymple, T. (1960). Flood frequency analyses. Water supply paper 1543-a, U.S. Geological Survey, Reston, VA.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

Diggle, P., Tawn, J., and Moyeed, R. (1998). Model-based geostatistics. *JRSSC: Applied Statistics*, 47:299–350.

Fuentes, M., Henry, J., and Reich, B. (2009). Nonparametric spatial models for extremes: Application to extreme temperature data. *submitted*.

Hosking, J. R. M. and Wallis, J. R. (1997). *Regional Frequency Analysis: An approach based on L-Moments*. Cambridge, University Press, Cambridge., U.K.

Huser, R. and Davison, A. (2012). Space-time modelling of extreme events. *arXiv preprint arXiv:1201.3245*.

Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37:2042–2065.

Lindsay, B. (1988). Composite likelihood methods. *Contemporary Mathematics*, 80(1):221–39.

Padoan, S., Ribatet, M., and Sisson, S. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489):263–277.

Pauli, F., Racugno, W., and Ventura, L. (2011). Bayesian composite marginal likelihoods. *Statistica Sinica*, 21:149–164.

Ribatet, M., Cooley, D., and Davison, A. C. (2011). Bayesian inference for composite likelihood models and an application to spatial extremes. *Statistica Sinica*, 22:813–846.

Sang, H. and Gelfand, A. E. (2010). Continuous spatial process models for spatial extreme values. *Journal of Agricultural, Biological, and Environmental Statistics*, 15:49–65.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5(1):33–44.

Schliep, E., Cooley, D., Sain, S. R., and Hoeting, J. A. (2010). A comparison study of extreme precipitation from six different regional climate models via spatial hierarchical modeling. *Extremes*, 13:219–239.

Smith, R. (1990). Max-stable processes and spatial extremes. Technical report, University of North Carolina, Chapel Hill.

Tawn, J. (1990). Modeling multivariate extreme value distributions. *Biometrika*, 75:245–253.