# Spatial Modeling and Extremes

Richard A. Davis
Columbia University

Collaborators:
Yongbum Cho, Columbia University
Souvik Ghosh, LinkedIn
Claudia Klüppelberg, Technical University of Munich
Christina Steinkohl, Technical University of Munich

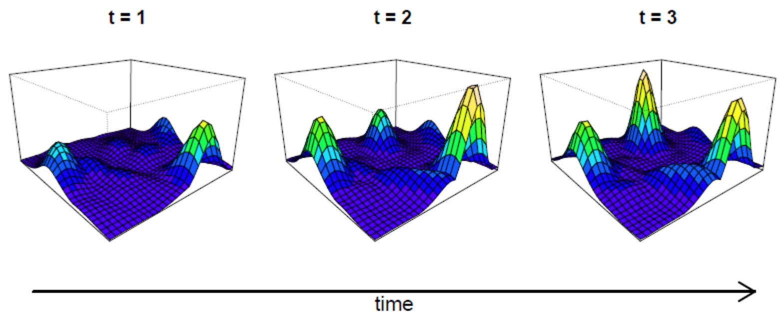
* Support of the Villum Kann Rasmussen Foundation is gratefully acknowledged.



## Game Plan

- Building a max-stable process in space-time
- Inference for Brown-Resnick process
  - pairwise likelihood
  - semi-parametric inference
- Data example
- Simulation

## Extremal Dependence in Space and Time




Space-time domain: $\{(\boldsymbol{s}, t) \in \mathbb{R}^d \times [0, \infty)\}$



## Building a Max-Stable Model in Space-Time

Building blocks: Let $Z(s,t)$ be a stationary Gaussian process on $\mathbb{R}^2 \times \mathbb{R}^+$ with mean 0 and variance 1.

- Transform the $Z(s,t)$ processes via

$$Y(s,t) = 1/-\log(\Phi(Z(s,t)),$$

where $\Phi$ is the standard normal cdf. Then $Y(s,t)$ has unit Frechet marginals, i.e., $P(Y(s,t) \le x) = \exp\{-1/x\}$.

Note: For any (nondegenerate) Gaussian process $Z(s,t)$, we have

$$\lim_{x \to \infty} P(Z(s,t) > x \mid Z(0,0) > x) = 0.$$

and hence $\lim_{x \to \infty} P(Y(s,t) > x \mid Y(0,0) > x) = 0.$

In other words,

- observations at distinct locations are asymptotically independent.
- not good news for modeling spatial extremes!

## Building a Max-Stable Model in Space-Time

Now assume $Z(s,t)$ is isotropic with covariance function

$$\mathrm{Cov}\big(Z(h,u),Z(0,0)\big) = r(|h|,u) = \exp\{-\theta_1|h|^{\alpha_1} - \theta_2|u|^{\alpha_2}\},$$

where $\theta_1,\theta_2 > 0$ are the range parameters and $\alpha_1,\alpha_2 \in (0,2]$ are the shape parameters. Note that

$$1 - r(h,u) \sim \theta_1|h|^{\alpha_1} + \theta_2|u|^{\alpha_2} =: \delta(h,u) \;\; as\; h,u \to 0,$$

(the semi-variogram $\frac{1}{2}(var(Z(h,u) - Z(0,0))$) and hence

$$\log n(1 - r(s_n h, t_n u)) \to \delta(h,u),$$

where $s_n = (\log n)^{-\frac{1}{\alpha_1}}$ and $t_n = (\log n)^{-\frac{1}{\alpha_2}}$.

It follows that

$$\mathrm{Cov}\big(Z(s_n h, t_n u), Z(0,0)\big) = r(s_n h, t_n u) \sim 1 - \delta(h,u)/\log n.$$



Then (see Kabluchko et al. (2011)),

$$Y_n(s,t) := \frac{1}{n}\bigvee_{j=1}^{n} \frac{1}{-\log\Big(\Phi\big(Z_j(s_n s, t_n t)\big)\Big)}$$
$$\to \eta(s,t)$$

on $C(\mathbb{R}^2 \times [0,\infty))$. Here the $Z_j$ are IID replicates of the GP $Z$, and $\eta$ is a *Brown-Resnick max-stable process.*

Specifically,

$$\eta(s,t) = \bigvee_{j=1}^{\infty} \xi_j \exp\{W_j(s,t) - \delta(s,t)\}$$

where $\{\xi_j\}$ pts of PPP($\xi^{-2}d\xi$), and $\{W_j\} \sim IID$ Gaussian processes with mean zero, $W(0,0) = 0$, and

i. stationary increments

ii. $Cov\big(W(s_1,t_1), W(s_2,t_2)\big) = \delta(s_1,t_1) + \delta(s_2,t_2) - \delta(s_1 - s_2, t_1 - t_2)$

## Building a Max-Stable Model in Space-Time

$$\eta_n(s,t) := \frac{1}{n}\bigvee_{j=1}^{n} -\frac{1}{\log\left(\Phi\left(Z_j(s_n s, t_n t)\right)\right)}$$
$$\to \eta(s,t)$$

Bivariate distribution function:

$$P(\eta(h,u) \le x, \eta(0,0) \le y) = \exp\{-V(x,y;\delta)\}$$

where

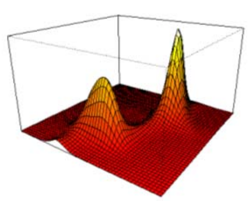
$$V(x,y;\delta) = x^{-1}\Phi\left(\frac{\log(y/x)}{2\sqrt{\delta}} + \sqrt{\delta}\right) + y^{-1}\Phi\left(\frac{\log(x/y)}{2\sqrt{\delta}} + \sqrt{\delta}\right),$$

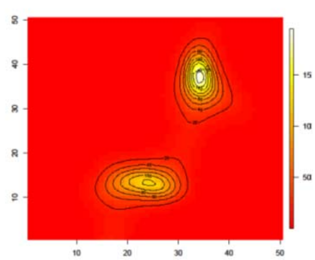
and $\delta = \delta(h,u)$.



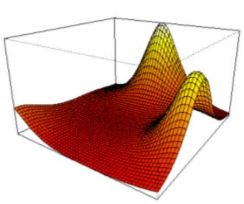
## Scaled and transformed Gaussian random fields (fixed time point)

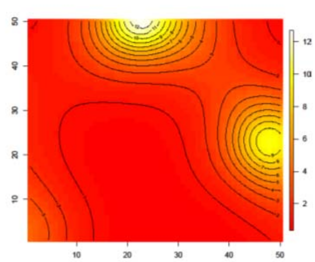
$$\frac{-1}{\log(\Phi(Z(s,t))}$$





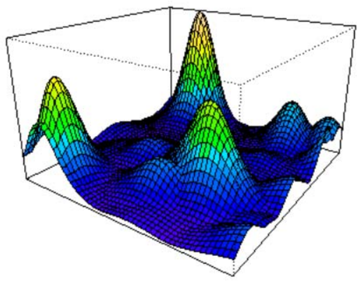
$$\frac{-1}{\log(\Phi(Z(s_n s, t_n t))}$$

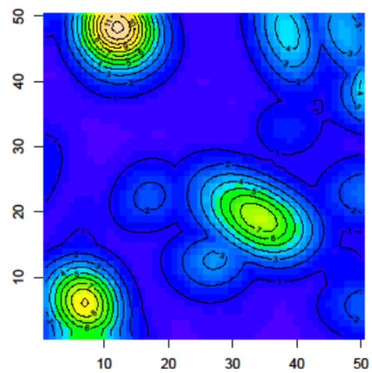




## Scaled and transformed Gaussian random fields (fixed time point)

$$\eta_n(s,t) := \frac{1}{n}\bigvee_{j=1}^{n} -\frac{1}{\log\left(\Phi\left(Z_j(s_n s, t_n t)\right)\right)}$$








## Estimation—composite likelihood approach

For dependent data, it is often infeasible to compute the exact likelihood based on some model. An alternative is to combine likelihoods based on subsets of the data.

To fix ideas, consider the following data/model setup:

(Here we have already assumed that the data has been transformed to a stationary process with unit Frechet marginals.)

Data: $Y(s_1), \ldots, Y(s_N)$ (field sampled at locations $s_1, \ldots, s_N$ )

Model: max-stable model defined via the limit process

$$\max_{j=1,\ldots,n} Y_n^{(j)}(s) \to_d X(s),$$

- $Y_n(s) = Y(s/(\log n)^{1/\beta}) = -1/\log(\Phi(Z(s/(\log n)^{1/\beta}))$
- $Z(s)$ is a GP with correlation function $\rho(|s-t|) = \exp\{-|s-t|^{\beta}/\phi\}$

## Estimation—composite likelihood approach

Bivariate likelihood: For two locations $s_i$ and $s_j$, denote the pairwise likelihood by

$$f(y(s_i), y(s_j); \delta_{i,j}) = \partial^2/(\partial x \partial y)\, F(Y(s_i) \le x, Y(s_j) \le x)$$

where F is the CDF

$F(Y(s_i) \le x, Y(s_j) \le x)$

$$= \exp\{-(x^{-1}\Phi(\log(y/x)/(2\sqrt{\delta}) + \sqrt{\delta}) + y^{-1}\Phi(\log(x/y)/(2\sqrt{\delta}) + \sqrt{\delta}))\},$$

and $\delta_{i,j} = |s_i - s_j|^{\beta}/\phi$ is a function of the parameters $\beta$ and $\phi$.

Pairwise log-likelihood:

$$PL(\phi, \beta) = \sum_{\substack{i=1 \\ i \ne j}}^{N} \sum_{j=1}^{N} \log f(y(s_i), y(s_j); \delta_{i,j})$$



## Estimation—composite likelihood approach

Potential drawbacks in using all pairs:

- Still may be computationally intense with $N^2$ terms in sum.
- Lack of consistency (especially if the process has long memory)
- Can experience huge loss in efficiency.

Suppose we have observations: $\eta(s_i, t_k), i = 1 \ldots, M; j = 1, \ldots, T.$ Then the weighted composite likelihood is given by

$$PL^{M,T}(\psi) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \sum_{k=1}^{T-1} \sum_{l=k+1}^{T} w_{i,j}^{M} w_{k,l}^{T} \log f_{\psi}\,(\eta(s_i, t_k), \eta(s_j, t_l))$$

where $\psi = (\theta_1, \alpha_1, \theta_2, \alpha_2)$ and the weights are band limited,

$$w_{i,j}^{M} = 1_{|s_i - s_j| \le r}, \qquad w_{i,j}^{M} = 1_{|t_k - t_l|| \le p}.$$

Estimate $\psi$ by maximizing $PL^{(M,T)}(\psi)$.

## Estimation—composite likelihood approach

Asymptotic properties: Under ergodic, mixing, and identifiability conditions on the max-stable process (see Davis, Klüppelberg, and Steinkohl (2013), then

$$\sqrt{MT}\ (\hat{\psi}-\psi) \to_d\ N(0 F^{-1}\Sigma F^{-T}).$$



## Simulation Examples

Simulation setup:

- Simulate 1600 points of a spatial (max-stable) process Y(s) on a grid of 40x40 in the plane.
- Choose a distance $r$ = 9, 15, 25 (number of neighbors used)
- Maximize

$$PL(\phi,\beta)=\sum_{i=1}^{N}\sum_{j:|s_j-s_i|\le D_i}\log f(y(s_i),y(s_j);\delta_{i,j})$$

with respect to $\phi$ and $\alpha$

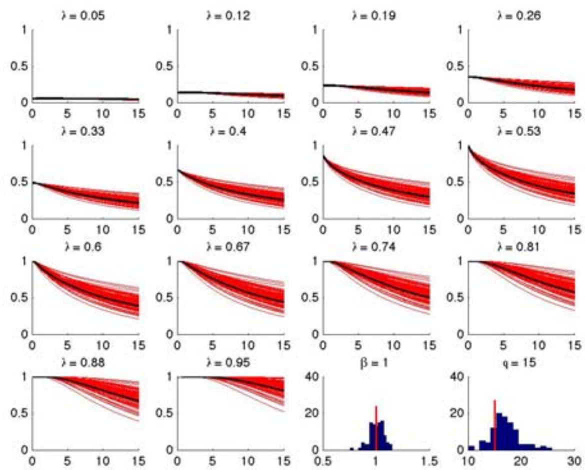
- Calculate summary dependence statistics:

$\rho(\lambda; |s-t|) = \lim_{n\to\infty} P(Y_n(s) > n(1-\lambda) \mid Y_n(t) > n\lambda)$.

## Simulation Examples

Process: Limit process $\delta = |s-t|^{\beta}/\phi$, $\beta = 1$ , $\phi = 5$; 9 nearest neighbors

$\rho(\lambda; |s-t|) = \lim_{n\to\infty} P(Y_n(s) > n(1-\lambda) \mid Y_n(t) > n\lambda)$

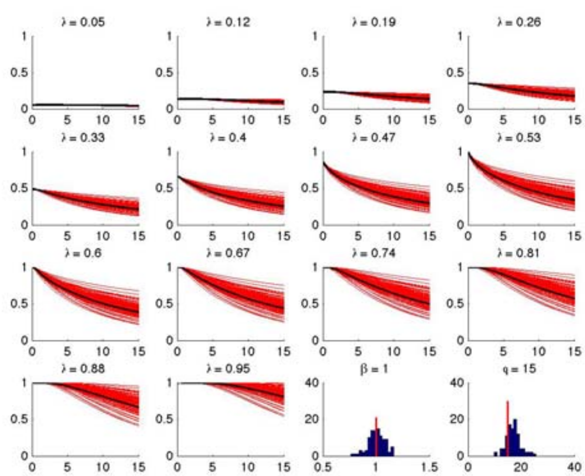





## Simulation Examples

Process: Limit process $\delta = |s-t|^{\beta}/\phi$, $\beta = 1$ , $\phi = 15$; 25 neighbors

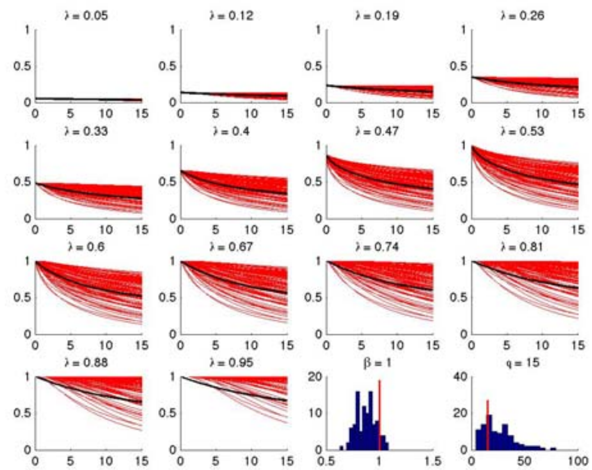
$\rho(\lambda; |s-t|) = \lim_{n\to\infty} P(Y_n(s) > n(1-\lambda) \mid Y_n(t) > n\lambda)$



# Simulation Examples

Schlather Process r(h) =exp(-h/20); 25 neighbors

$\rho(\lambda; |s-t|) = \lim_{n\to\infty} P(Y_n(s) > n(1-\lambda) \mid Y_n(t) > n\lambda)$

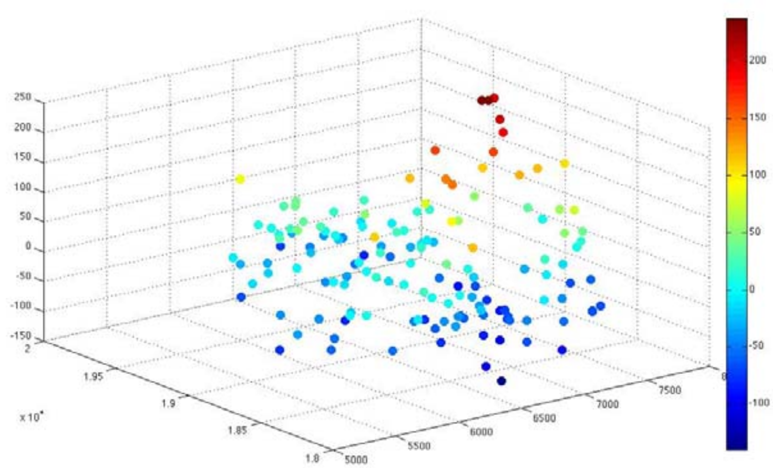






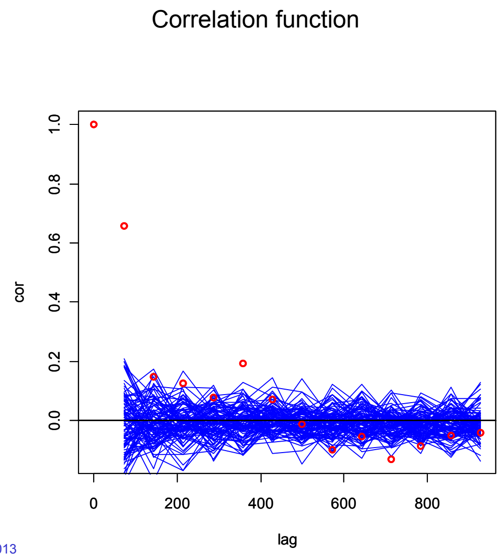
# Illustration with French Precipitation Data

Data from Naveau et al. (2009). Precipitation in Bourgogne of France; 51 year maxima of daily precipitation. Data has been adjusted for seasonality and orographic effects.



## Illustration with French Precipitation Data

Estimated spatial correlation function before transformation to unit Frechet.

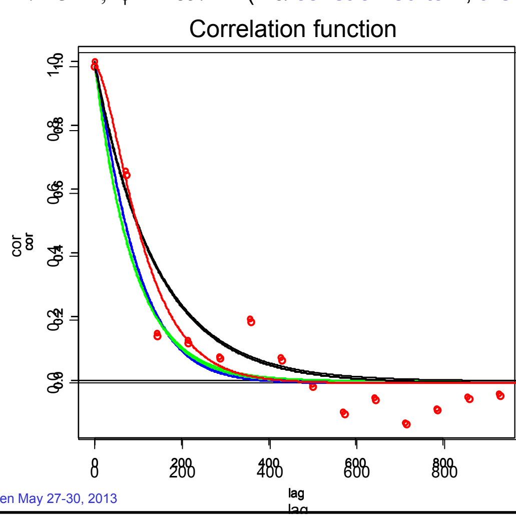





## Illustration with French Precipitation Data

After transforming the data to unit Frechet marginals, we estimated $\alpha$ and $\phi$ using pairwise likelihood (nearest neighbors with r = 12).

$\alpha = 1.143$ ; $\phi = 185.1$ (if $\alpha$ constrained to 1, then $\phi = 89.2$)




Green: $\alpha = 1.143$ ; $\phi = 185.1$

Blue: $\alpha = 1$ ; $\phi = 89.2$

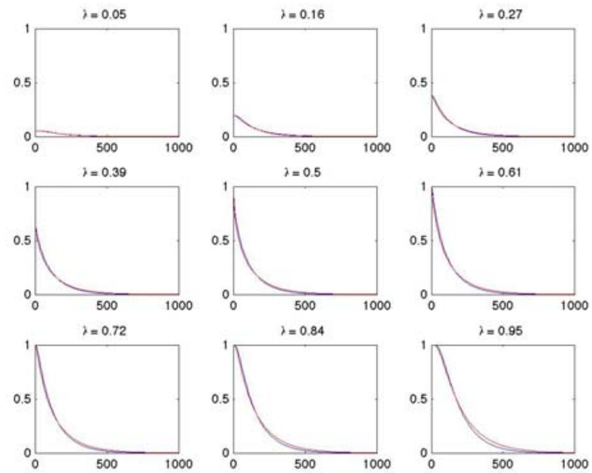
Black: $\alpha = 1$ ; $\phi = 139.6$ (MLE)

Red: $\alpha = 1.17$ ; $\phi = 147.8$ (Matern)

(MLE based on GP likelihood)

## Illustration with French Precipitation Data

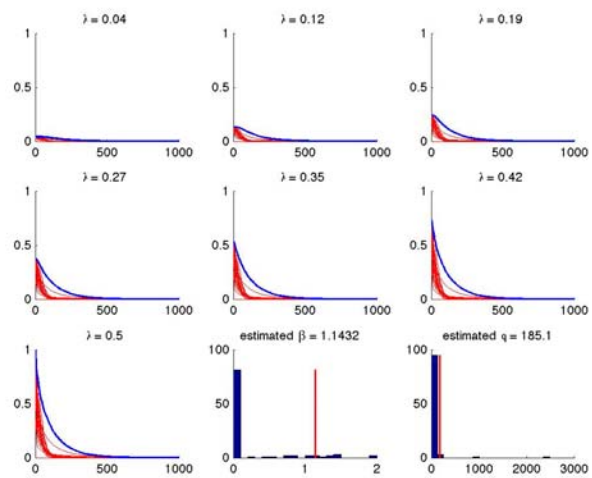
Plot of $\rho(\lambda; |s-t|) = \lim_{n\to\infty} P(Y_n(s) > n(1-\lambda) \mid Y_n(t) > n\lambda)$






## Illustration with French Precipitation Data

Results from random permutations of data—just for fun. Since random permutations have no spatial dependence, $\rho$ should be 0 for $|s-t| > 0$.



## Inference for Brown-Resnick Process

Recall the spatial extremogram given by.

$$\rho_{A,B}(h,u) = \lim_{x\to\infty} P(\eta(s+h,t+u)\,\epsilon\, xB \mid \eta(s,t)\,\epsilon\, xA)$$

For the special case, $A = B = (1,\infty)$,

$$\chi(h,u) = \lim_{x\to\infty} P(\eta(s+h,t+u) > x \mid \eta(s,t) > x)$$

For the Brown-Resnick process described earlier

$$\chi(h,u) = 2(1-\Phi(\sqrt{\theta_1 h^{\alpha_1} + \theta_2 u^{\alpha_2}})),$$

we find that

$$2\log(\Phi^{-1}(1-\tfrac{1}{2}\chi(h,0)) = \log\theta_1 + \alpha_1 \log h \text{ and}$$

$$2\log(\Phi^{-1}(1-\tfrac{1}{2}\chi(0,u)) = \log\theta_2 + \alpha_2 \log u$$



## Inference for Brown-Resnick Process

Semi-parametric: Use nonparametric estimates of the extremogram and then regress function of extremogram on the lag.

Regress: $2\log(\Phi^{-1}(1-\tfrac{1}{2}\hat{\chi}(h,0)))$ on $1$ and $\log h$

$2\log(\Phi^{-1}(1-\tfrac{1}{2}\hat{\chi}(0,u)))$ on $1$ and $\log u$,

The intercepts and slopes become the respective estimates of $\log\theta_i$ and $\alpha_i$. Asymptotic properties of spatial extremogram derived by Cho et al. (2013).

## Bias-correction

Recall: Generally, we need to center the empirical extremogram by the pre-asymptotic extremogram. How do we get consistent estimates for the semi-parametric estimates?

Temporal: For the BR process, it turns out that one can center with the actual extremogram—have asymptotic equivalence of the centering b/ pre-asymptotic extremogram and extremogram).

Spatial: The PA-extremogram $\chi_m(r,0)$, in the spatial direction can be written as

$$\chi_m(r,0) \sim \chi(r,0) + \frac{1}{4n_m}(\chi(r,0)^2 - \chi(r,0))$$

Bias corrected estimate becomes

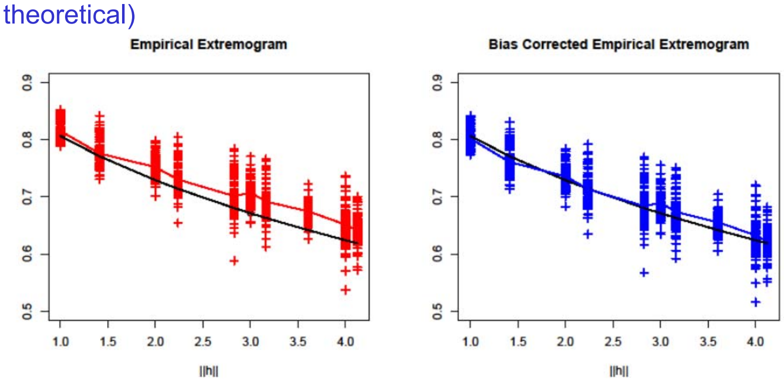
$$\tilde{\chi}(r,0) = \hat{\chi}(r,0) - \frac{1}{4n_m}(\hat{\chi}(r,0)^2 - \hat{\chi}(r,0))$$



## Bias-correction

Remark: In Davis, Klüppelberg, Steinkohl (2013), work out asymptotics for $\tilde{\chi}(r,0)$.

Simulation: Empirical extremogram (left); bias corrected (right) for 100 simulated max-stable processes w/ $\delta(h,0) = .06|h|$ (black is theoretical)



## Semi-parametric estimates

Asymptotics for spatial parameter: Let $\psi = (\log\theta_1, \alpha\_1)$ be the parameter vector and $\hat{\psi}^c$ its *constrained* weighted least squares estimate. Then

$$\left(\frac{m^2}{n_m}\right)^{\frac{1}{2}}\left(\hat{\psi}^c - \psi\right) \to_d \begin{cases} Z_1, & \alpha_1 < 2, \\ Z_2, & \alpha_1 = 2, \end{cases}$$

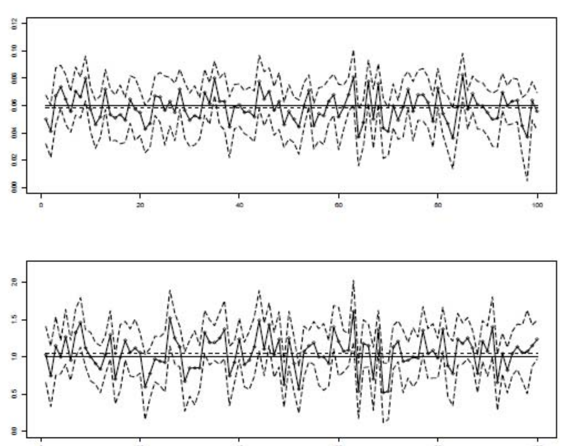
where $Z_1 \sim N(0, \Sigma_1)$ and $Z_2$ has a constrained distribution (Andrews (1999)..

Bootstrapping: Bootstrapping also works here, but one needs to take care of the constraint properly (Andrews (2000)).

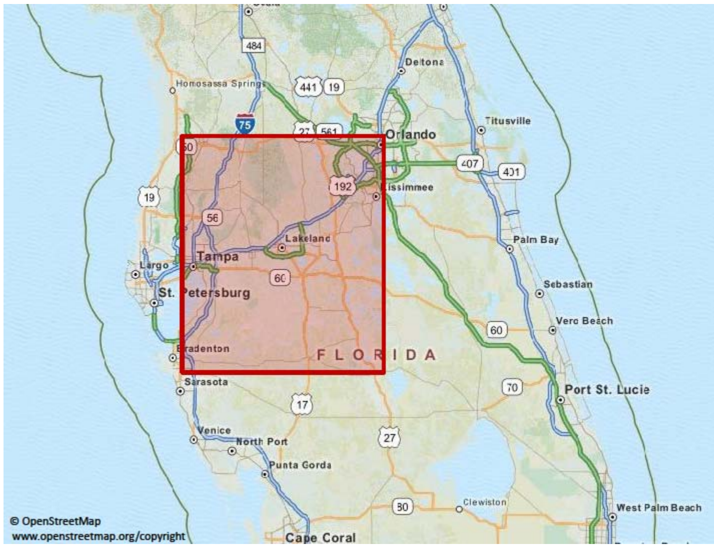


## Semi-parametric estimates

Estimates of $\theta_1$(top) and $\alpha_1$ (bottom) for 100 simulated max-stable processes with 95% CIs via BS (middle line true, dotted is average)



## Data Example:  extreme rainfall in Florida






## Data Example:  extreme rainfall in Florida

Radar data:

Rainfall in inches measured in 15-minutes intervals at points of a spatial 2x2km grid.

Region:

120x120km, results in 60x60=3600 measurement points in space.

Take only wet season (June-September).

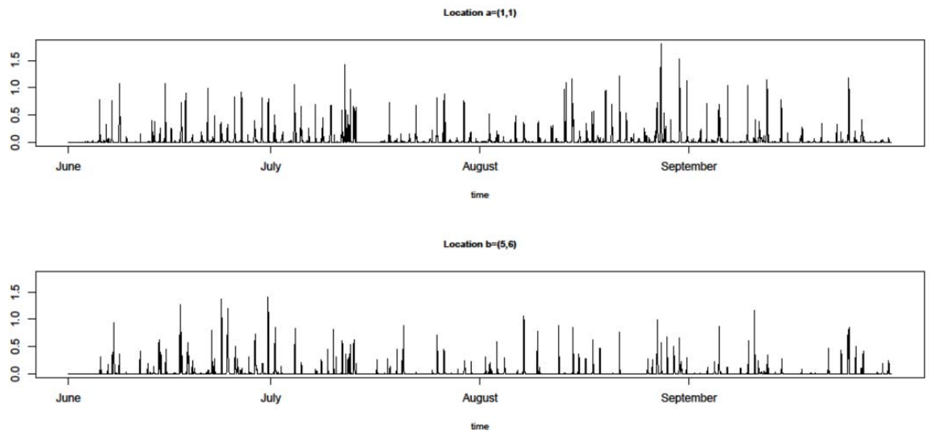
Block maxima in space: Subdivide in 10x10km squares, take maxima of rainfall over 25 locations in each square. This results in 12x12=144 spatial maxima.

Temporal domain: Analyze daily maxima and hourly accumulated rainfall observations.

Fit extremal space-time model to daily/hourly maxima.

## Data Example: extreme rainfall in Florida

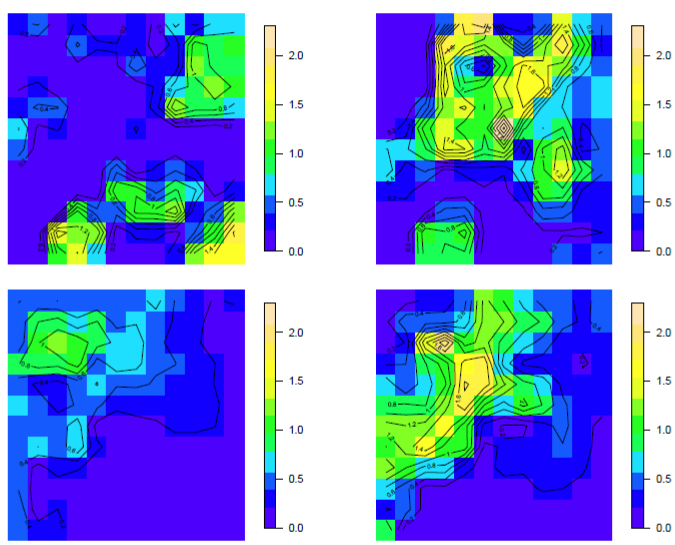
Hourly accumulated rainfall time series in the wet season 2002 at 2 locations.





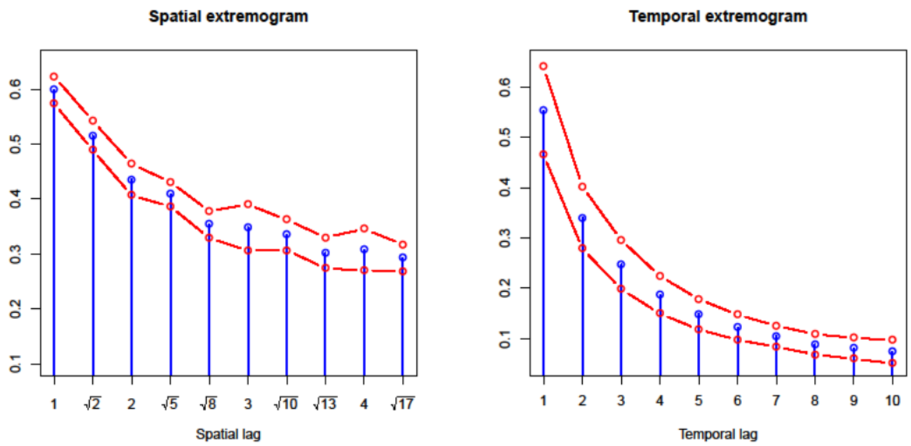

## Data Example: extreme rainfall in Florida

Hourly accumulated rainfall fields for four time points.



## Data Example: extreme rainfall in Florida

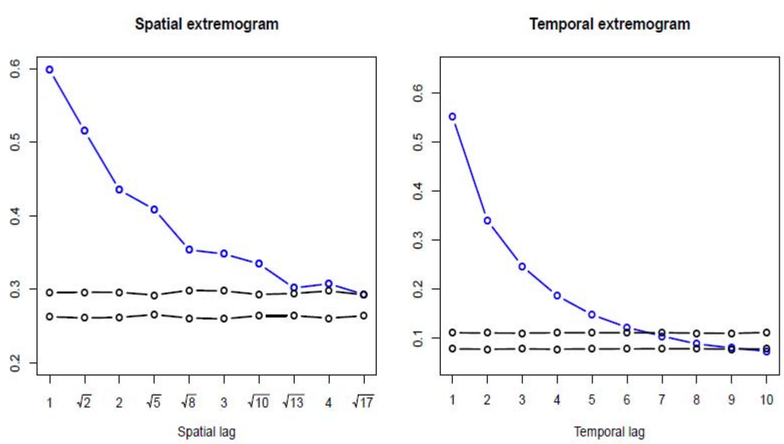
Empirical extremogram in space (left) and time (right)






## Data Example: extreme rainfall in Florida

Empirical extremogram in space (left) and time (right):
spatial indep for lags > 4; temporal indep for lags > 6.



## Data Example: extreme rainfall in Florida

Empirical extremogram in space (left) and time (right)

### Semiparametric estimates

| | | | | |
|---|---|---|---|---|
| Estimate | $\hat{\theta}_1$ | 0.2987 | $\hat{\alpha}_1$ | 0.9664 |
| Bootstrap-CI | | [0.2469,0.3505] | | [0.8407,1.0921] |
| Estimate | $\hat{\theta}_2$ | 0.4763 | $\hat{\alpha}_2$ | 1.0686 |
| Bootstrap-CI | | [0.2889,0.6637] | | [0.8514,1.2859] |

### Pairwise likelihood estimates

| PL estimates | $\hat{\theta}_1$ | $\hat{\alpha}_1$ | $\hat{\theta}_2$ | $\hat{\alpha}_2$ |
|---|---|---|---|---|
| | 0.3353 | 0.9302 | 0.4845 | 1.0648 |



## Data Example: extreme rainfall in Florida

Computing conditional return maps.

Estimate $z_c(s,t)$ such that

$$P(Z(s,t) > z_c(s,t) \mid Z(s^*,t^*) > z^*) = p_c,$$

where $z^*$ satisfies $P(Z(s^*,t^*) > z^*) = p^*$ is pre-assigned.

A straightforward calculation shows that $z_c(s,t)$ must solve,

$$p_c = 1 - \frac{1}{p^*}\exp\left\{-\frac{1}{z_c(s,t)}\right\} + \frac{1}{p^*}F_{(BR)}(z_c(s,t), 1-p^*),$$

100-hour return maps ($p_c = .01$): $s^* = (5{,}6)$, time lags = 0,2,4,6 hours (left to right on top and then right to left on bottom), quantiles in inches.