Minimum distance estimation for max-stable models

R.A. Yuen† and Stilian Stoev, Department of Statistics, University of Michigan

Max-stable models

- For iid observations $Y^{(i)}(\mathbf{s}),\ i=1,2,\ldots$ at fixed points $\mathbf{s}\in S$, we consider the limit of point-wise maxima

$$\left\{\frac{1}{a_n(\mathbf{s})}\max_{i=1,\ldots,n} Y^{(i)}(\mathbf{s})-b_n(\mathbf{s})\right\}_{\mathbf{s}\in S}\xrightarrow{d}\{X(\mathbf{s})\}_{\mathbf{s}\in S},\ \text{as } n\to\infty$$

where $a_n(\mathbf{s})$ and $b_n(\mathbf{s})$ are normalization functions.

- The limiting process $X:=\{X(\mathbf{s})\}_{\mathbf{s}\in S}$ is a model for **worst case scenaria** and must be max-stable (Resnick 1987):

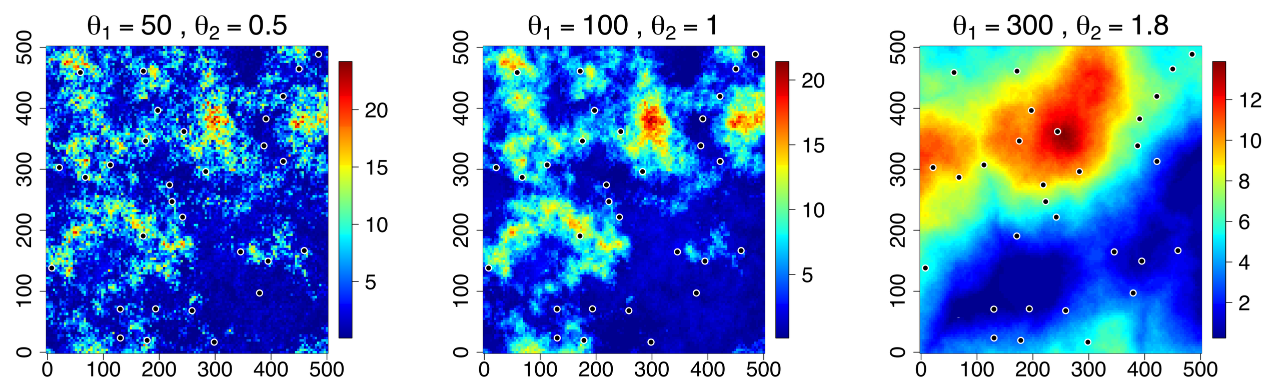
> ***Max-stable process***
>
> For independent copies $X^{(i)}(\mathbf{s}), i=1,\ldots,n$ of $X(\mathbf{s})$, there exists functions $c_n(\mathbf{s})$ and $d_n(\mathbf{s})$ such that
>
> $$\left\{\frac{1}{c_n(\mathbf{s})}\max_{i=1,\ldots,n} X^{(i)}(\mathbf{s})-d_n(\mathbf{s})\right\}_{\mathbf{s}\in S}\stackrel{d}{=}\{X(\mathbf{s})\}_{\mathbf{s}\in S}$$

- All max-stable processes have a spectral representation for their finite dimensional distribution functions (de Haan 1984):

$$\mathbb{P}(X(\mathbf{s}_1)\le x_1,\ldots,X(\mathbf{s}_d)\le x_d):=F(\mathbf{x})=\exp\left\{\int_{\mathbb{W}}\left(\max_{j=1,\ldots,d}\frac{w_j}{x_j}\right)H(d\mathbf{w})\right\}$$

where the *spectral measure* H controls the dependence structure of X.

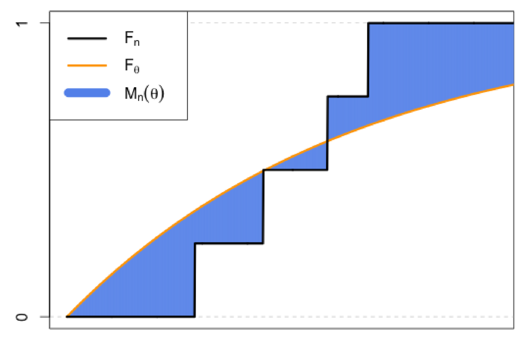
- Many flexible max-stable models have been proposed by specifiying a parametric family of spectral measures, $\{H_\theta:\theta\in\mathbb{R}^p\}$. Examples include the well known Smith storm model (Smith 1990), Brown-Resnick (Brown and Resnick 1977) and the following extremal Gaussian.

> ***Extremal Gaussian model (Schlather 2002)***
>
> $$F_\theta(\mathbf{x})=\exp\left\{-\int_{\mathbb{R}^d}\left(\max_{j=1,\ldots,d}\frac{w_j}{x_j}\right)\frac{\sqrt{|\Omega_\theta|}e^{-\frac{1}{2}\mathbf{w}^\top\Omega_\theta\mathbf{w}}}{(\sqrt{2\pi})^{d-1}}d\mathbf{w}\right\}. \qquad (1)$$
>
> where Ω_θ is the precision matrix induced by a correlation function $\rho(\mathbf{s}_i,\mathbf{s}_j|\theta)$.
>
> 
>
>
> ***Figure 1:*** *Three realizations from the extremal Gaussian model with stable correlation function:* $\rho(\mathbf{s}_i,\mathbf{s}_j|\theta)=\exp\left[-(\|\mathbf{s}_i-\mathbf{s}_j\|/\theta_1)^{\theta_2}\right], \theta_1>0,\theta_2\in(0,2]$

- Most useful max-stable models including (1) have **no tractable likelihood** when $d>2$, making inference very challenging.
- Current state of art: Bivariate maximum composite likelihood estimator (MCLE) (Padoan et. al 2010)

Minimum distance estimation

- Let $F_n(\mathbf{z})=n^{-1}\sum_{i=1}^n\mathbb{I}\{X^{(i)}(\mathbf{s}_j)\le z_j, j=1,\ldots,d\}$ be the d-dimensional empirical distribution function generated by a random sample $X^{(i)}(\mathbf{s}_1),\ldots,X^{(i)}(\mathbf{s}_d),\ i=1,\ldots,n$.
- Let μ be a tuning measure that emphasizes regions of the sample space.

> ***Minimum distance criterion***
>
> $$M_n(\theta)=\int_{\mathbb{R}_+^d}(F_n(\mathbf{z})-F_\theta(\mathbf{z}))^2\mu(d\mathbf{z})$$
>
> ***Minimum distance estimator***
>
> $$\hat{\theta}_n=\arg\min_{\theta\in\Theta}M_n(\theta)$$
>
> 
>

- Subject to mild regularity conditions, $\hat{\theta}_n\xrightarrow{p}\theta_0$ and $\sqrt{n}(\hat{\theta}_n-\theta_0)\xrightarrow{d}\mathcal{N}\left(0,J_{\theta_0}^{-1}D_{\theta_0}J_{\theta_0}^{-1}\right)$ where elements of the $p\times p$ matrices J_{θ_0} and D_{θ_0} are

$$\left(J_{\theta_0}\right)_{ij}=\int_{\mathbb{R}_+^d}\frac{\partial}{\partial\theta_i}F_\theta(\mathbf{z})\frac{\partial}{\partial\theta_j}F_\theta(\mathbf{z})\mu(d\mathbf{z})$$

and

$$\left(D_{\theta_0}\right)_{ij}=\int_{\mathbb{R}_+^d}\int_{\mathbb{R}_+^d}B_{\theta_0}(\mathbf{z}_1,\mathbf{z}_2)\frac{\partial}{\partial\theta_i}F_{\theta_0}(\mathbf{z}_1)\frac{\partial}{\partial\theta_j}F_{\theta_0}(\mathbf{z}_2)\mu(d\mathbf{z}_1)\mu(d\mathbf{z}_2)$$

with $B_{\theta_0}(\mathbf{z}_1,\mathbf{z}_2)=F_{\theta_0}(\mathbf{z}_1\wedge\mathbf{z}_2)-F_{\theta_0}(\mathbf{z}_1)F_{\theta_0}(\mathbf{z}_2)$.

Simulation Study

- Using the model (1) with stable correlation function we set $\theta_0=(100,1)$ and simulated 100 replications at $d=30$ uniformly sampled locations over a 500×500 grid. Realizations were generated using the **R** package **SpatialExtremes** (Ribatet 2012).
- For each replication, we generate samples of $n=100$ and $n=1000$. .\
- We specify an emperical measure for μ and numerically optimize the criterion $M_n(\theta)$.

Table: *Empirical mean and standard deviation from 100 replications of the minimum distance estimator. Coverage rates are calculated numerically using plug-in estimates of the matrices* $J_{\hat{\theta}_n}$ *and* $D_{\hat{\theta}_n}$.

	θ_1 (100)		θ_2 (1)	
n	100	1000	100	1000
mean	110.56	97.00	1.25	1.10
sd	113.73	32.89	0.63	0.34
.95 coverage	0.98	0.96	0.90	0.92

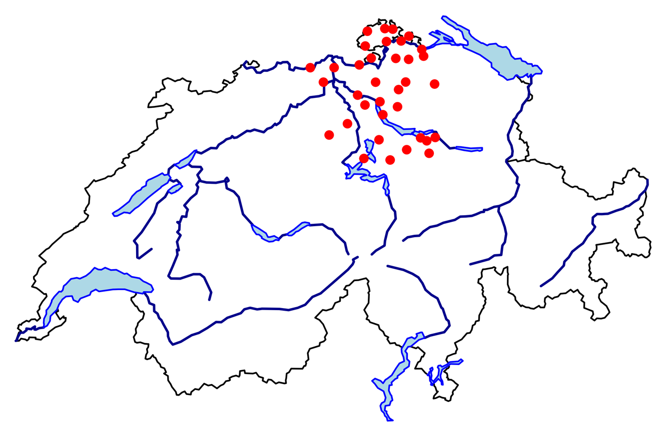
Application to Swiss rainfall

- We model maximum single day rainfall during summer months using the extremal Gaussian model (1) with Matérn correlation function

$$\rho(\mathbf{s}_i,\mathbf{s}_j|\theta)=\frac{(\|\mathbf{s}_i-\mathbf{s}_j\|/\theta_1)^{\theta_2}}{2^{\theta_2-1}\Gamma(\theta_2)}K_{\theta_2}(\|\mathbf{s}_i-\mathbf{s}_j\|/\theta_1),\quad \theta_1>0,\theta_2>0.$$

K_{θ_2} denotes the modified Bessel function of order θ_2.

- Data are recorded from 1962-2008 at 35 observation locations near Zurich, Switzerland. *(Data courtesy of* **R** *package* **SpatialExtremes***)*



- Summer maxima are first transformed to unit Fréchet margins. Each year is treated as an independent observation.
- Specifying an empirical measure for μ, we optimize the criterion $M_n(\theta)$ resulting in estimates

$$\hat{\theta}_1=18.46(25.58),\quad \hat{\theta}_2=0.5038(1.133).$$

Standard errors are in parentheses.

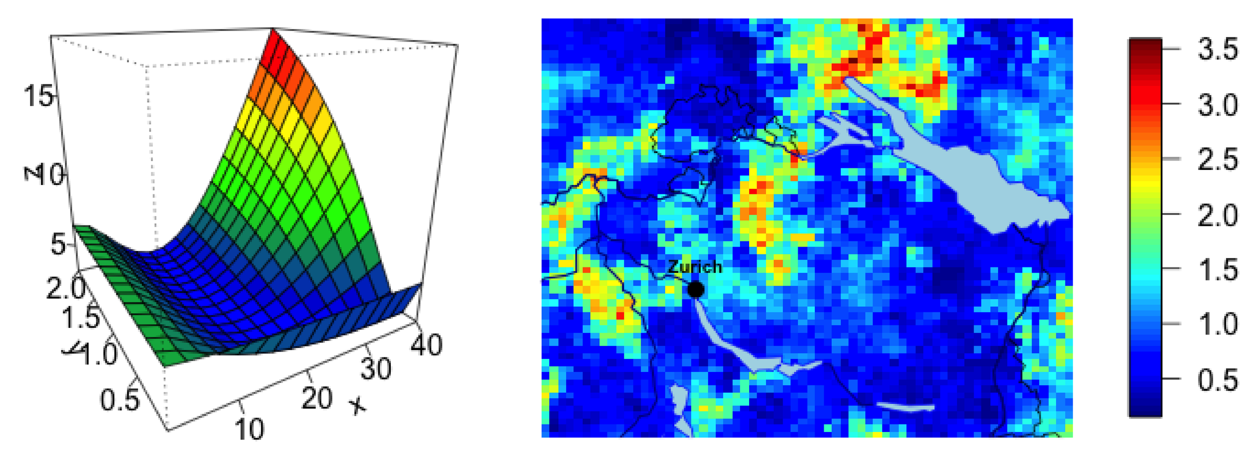


Figure 2: $M_n(\theta)$ *criterion surface for Swiss rainfall data (left). A single realization from the fitted model (right)*

For details on this research and other projects please visit:

http://www.stat.lsa.umich.edu/~bobyuen

†Partially supported by Univ. of Mich. Rackham Merit Fellowship and NSF-AGEP grant DMS 1106695.Thank you to the STATMOS research network on Statistical Methods in Oceanic and Atmospheric Sciences.