Applications of multivariate Extreme Value Theory to environmental data analysis

Philippe Naveau

Laboratoire des Sciences du Climat et l’Environnement (LSCE)
Gif-sur-Yvette, France

3 mars 2010
“Nous avons anticipé dans la mesure du possible mais on ne peut pas prévoir l’imprévisible”

Xynthia’s storm, 25th of Feb, 2010
Extreme quotes

1. “Man can believe the impossible, but man can never believe the improbable”
   Oscar Wilde (Intentions, 1891)

2. “Il est impossible que l'improbable n’arrive jamais”
   Emil Julius Gumbel (1891-1966)

Extreme events? ... a probabilistic concept linked to the tail behavior: low frequency of occurrence, large uncertainty and sometimes strong amplitude.
Patrick Galois (Meteo France) “Les tempêtes sont des phénomènes que l’on observe tous les cinq à dix ans en France en raison d’aléas climatiques. Si elle présente un caractère remarquable, Xynthia n’est pas pour autant le phénomène du siècle. Elle est ainsi moins exceptionnelle que celles de 1999 et ses vents sont moins intenses qu’en 2009. Mais son issue dramatique réside dans sa conjonction à un fort coefficient de marée sur la côte atlantique, au moment même de la marée haute.”

**Xynthia’s storm, 25th of Feb, 2010**
Important issues in Extreme Value Theory

- An asymptotic probabilistic concept
- A statistical modeling approach
- A multivariate problem
- Identifying clearly assumptions
- Assessing uncertainties
Motivation
Univariate EVT
BHM
Max-stable processes

GCM: temperatures maxima and minima

GCM: heavy precipitation

1. Relevant parameter in meteorology and climatology
2. Highly stochastic nature compared to other meteorological parameters

Emil Gumbel was born and trained as a statistician in Germany, forced to move to France and then the U.S. because of his pacifist and socialist views. He was a pioneer in the application of extreme value theory, particularly to climate and hydrology.

Waloddi Weibull was a Swedish engineer famous for his pioneering work on reliability, providing a statistical treatment of fatigue, strength, and lifetime.

Maurice Frechet was a French mathematician who made major contributions to pure mathematics as well as probability and statistics. He also collected empirical examples of heavy-tailed distributions.

**Other important names**: Fisher and Tippet (1928), Gnedenko (1943), etc
Generalized Extreme Value (GEV) distribution

\[ P \left( \frac{M_n - a_n}{b_n} < x \right) \sim \text{GEV}(x) = \exp \left\{ - \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^+ \right\} \]
Thresholding: the Generalized Pareto Distribution (GPD)

\[ \mathbb{P}\{R-u > y \mid R > u\} = \left(1 + \frac{\xi y}{\sigma_u}\right)^{-1/\xi} \]

Born in France and trained as an engineer in Italy, he turned to the social sciences and ended his career in Switzerland. He formulated the power-law distribution (or "Pareto’s Law"), as a model for how income or wealth is distributed across society.
Generalized Pareto Distribution (GPD)

\[
\mathbb{P}\{R - u > y|R > u\} = \left(1 + \frac{\xi y}{\sigma_u}\right)^{-1/\xi}
\]

Parameters

- \(u\) = predetermined threshold
- \(\sigma_u\) = scale parameter to be estimated
- \(\xi\) = shape parameter to be estimated

Advantages & Practical issues

- Flexibility to describe three different types of tail behavior
- More data are kept for the statistical inference
- Problem of threshold selection

Stability property

If the exceedance \((R - u|R > u)\) follows a \(\text{GPD}(\sigma_u, \xi)\) then the higher exceedance \((R - v|R > v)\) also follows \(\text{GPD}(\sigma_u + (v - u)\xi, \xi)\)
GPD: “From Bounded to Heavy tails”
Intro summary

**Modeling maxima : GEV**
Stability for the max operator and $X_0, X_1, \ldots X_n$ idd GEV

$$a \max(X_1, \ldots, X_n) + b = X, \text{ ie } F^n((x - b)/a) = F(x)$$

**Modeling excedances : GPD**
If exceedances $(R - u | R > u)$ follows a GPD$(\sigma_u, \xi)$ then higher exceedances $(R - v | R > v)$ also follows GPD$(\sigma_u + (v - u)\xi, \xi)$
“If an event can be produced by a number of $n$ different causes, then the probabilities of the causes given the event ... are equal to the probability of the event given that cause, divided by the sum of all the probabilities of the event given each of the causes.”

$$P(cause_i|event) = \frac{P(event|cause_i) \times P(cause_i)}{\sum_{j=1}^{n} P(event|cause_j) \times P(cause_j)}$$
Bayes’ formula = calculating conditional probability

\[ \frac{y}{x} \propto \frac{x}{y} \times [x] \]

1701(?) - 1761 “An essay towards solving a Problem in the Doctrine of Chances” (1764)
Hierarchical Bayesian Model with three levels

\[ P(\text{process, parameters}|\text{data}) \propto P(\text{data}|\text{process, parameters}) \times P(\text{process}|\text{parameters}) \times P(\text{parameters}) \]
EVT and BHM applications: a very short biblio

- **Extreme fires in Portugal**
  Turkman et al., 2007

- **Heavy rainfall**
  Cooley, Nychka and Naveau (2007)
  Coles & Tawn (1996).
  Ribatet et al, 2010

- **Biodiversity and extreme temperatures**
  Sang and Gelfand, 2009

- **Extreme snow**
  Blanchet et al., 2010

- **Lichenometry**
  Jomelli et al., 2007
Estimating return levels
Daily precipitation (April-October, 1948-2001, 56 stations)
Precipitation in Colorado’s front range

Data

- 56 weather stations in Colorado (semi-arid and mountainous region)
- Daily precipitation for the months April-October
- Time span = 1948-2001
- Not all stations have the same number of data points
- Precision: 1971 from 1/10th of an inche to 1/100

Our main assumptions

- Process layer: The scale and shape GPD parameters $(\xi(x), \sigma(x))$ are random fields and result from an unobservable latent spatial process.
- Conditional independence: precipitation are independent given the GPD parameters.

Our main variable change

$$\sigma(x) = \exp(\phi(x))$$
Hierarchical Bayesian Model with three levels

\[ P(\text{process, parameters}|\text{data}) \propto P(\text{data}|\text{process, parameters}) \times P(\text{process}|\text{parameters}) \times P(\text{parameters}) \]

Process level: the scale and shape GPD parameters \((\xi(x), \sigma(x))\) are hidden random fields
Our three levels

A) **Data layer** := $\mathbb{P}(\text{data}|\text{process, parameters}) = \mathbb{P}_\theta\{R(x_i) - u > y|R(x_i) > u\} = \left(1 + \frac{\xi_i y}{\exp \phi_i}\right)^{-1/\xi_i}$

B) **Process layer** := $\mathbb{P}(\text{process}|\text{parameters})$:

\[\text{e.g. } \phi_i = \alpha_0 + \alpha_1 \times \text{elevation}_i + \text{MVN}(0, \beta_0 \exp(-\beta_1 ||x_k - x_j||))\]

and $\xi_i = \begin{cases} 
\xi_{\text{moutains}}, & \text{if } x_i \in \text{mountains} \\
\xi_{\text{plains}}, & \text{if } x_i \in \text{plains} 
\end{cases}$

C) **Parameters layer (priors)** := $\mathbb{P}(\text{parameters})$:

e.g. $(\xi_{\text{moutains}}, \xi_{\text{plains}})$ Gaussian distribution with non-informative mean and variance
Bayesian hierarchical modeling

\[ P(R(x) > u) \]

Prior\(\alpha_0 + \alpha_1 \text{ elev}\)

\[ \beta_0 \exp(-\beta_1 ||.||) \]

\[ \xi \text{moutains} \]

\[ \xi \text{plains} \]
Climate space

foothill cities (C), plains (P), Palmer Divide (D), Front Range (F), mountain valley (V), and high elevation (H)
## Model selection

<table>
<thead>
<tr>
<th>Baseline model</th>
<th>$\tilde{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 0: $\phi = \phi$ $\xi = \xi$</td>
<td>73,595.5</td>
<td>2.0</td>
<td>73,597.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Models in latitude/longitude space</th>
<th>$\tilde{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: $\phi = \alpha_0 + \epsilon_\phi$ $\xi = \xi$</td>
<td>73,442.0</td>
<td>40.9</td>
<td>73,482.9</td>
</tr>
<tr>
<td>Model 2: $\phi = \alpha_0 + \alpha_1(\text{msp}) + \epsilon_\phi$ $\xi = \xi$</td>
<td>73,441.6</td>
<td>40.8</td>
<td>73,482.4</td>
</tr>
<tr>
<td>Model 3: $\phi = \alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$ $\xi = \xi$</td>
<td>73,443.0</td>
<td>35.5</td>
<td>73,478.5</td>
</tr>
<tr>
<td>Model 4: $\phi = \alpha_0 + \alpha_1(\text{elev}) + \alpha_2(\text{msp}) + \epsilon_\phi$ $\xi = \xi$</td>
<td>73,443.7</td>
<td>35.0</td>
<td>73,478.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Models in climate space</th>
<th>$\tilde{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 5: $\phi = \alpha_0 + \epsilon_\phi$ $\xi = \xi$</td>
<td>73,437.1</td>
<td>30.4</td>
<td>73,467.5</td>
</tr>
<tr>
<td>Model 6: $\phi = \alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$ $\xi = \xi$</td>
<td>73,438.8</td>
<td>28.3</td>
<td>73,467.1</td>
</tr>
<tr>
<td><strong>Model 7</strong>: $\phi = \alpha_0 + \epsilon_\phi$ $\xi = \xi_{\text{mtn}}$, $\xi_{\text{plains}}$</td>
<td><strong>73,437.5</strong></td>
<td><strong>28.8</strong></td>
<td><strong>73,466.3</strong></td>
</tr>
<tr>
<td>Model 8: $\phi = \alpha_0 + \alpha_1(\text{elev}) + \epsilon_\phi$ $\xi = \xi_{\text{mtn}}$, $\xi_{\text{plains}}$</td>
<td>73,436.7</td>
<td>30.3</td>
<td>73,467.0</td>
</tr>
<tr>
<td>Model 9: $\phi = \alpha_0 + \epsilon_\phi$ $\xi = \xi + \epsilon_\xi$</td>
<td>73,433.9</td>
<td>54.6</td>
<td>73,488.5</td>
</tr>
</tbody>
</table>

**NOTE:** Models in the climate space had better scores than models in the longitude/latitude space. $\epsilon \sim \text{MVN}(0, \Sigma)$, where $[\sigma]_{i,j} = \beta_{i,0} \exp(-\beta_{i,1} ||x_i - x_j||)$. 


Return levels posterior mean
Posterior quantiles of return levels (.025, .975)
Downscaling

Main statistical approaches

Coarse atmospheric data

Statistical downscaling approaches

Transfer $f^o$
Weather generator
Clustering
Geostatistics

Linear
Non-linear
Stochastic methods
Weather typing
Analogues
Kriging

Local variables (e.g., precip., temp.)
(small scale water cycle, impacts – crops, resources – etc.)
Downscaling of rainfalls

Conditional independence assumption limits

Ribatet, Cooley and Davison (2010)
Spatial Statistics for Maxima

How to describe the **spatial dependence** as a function of the **distance** between two points?
Spatial Statistics for Maxima

How to perform spatial interpolation of extreme events?
Spatial Statistics for Maxima

A few Approaches for modeling spatial extremes

- **Max-stable processes**: Adapting asymptotic results for multivariate extremes
  Schlather & Tawn (2003), Naveau et al. (2007), de Haan & Pereira (2005)

- **Bayesian or latent models**: spatial structure *indirectly* modeled via the EVT parameters distribution
  Ribatet et al., 2010

- **Linear filtering**: Auto-Regressive spatio-temporal heavy tailed processes,
  Davis and Mikosch (2007)

- **Gaussian anamorphosis**: Transforming the field into a Gaussian one
  Wackernagel (2003)
Max-stable processes

Max-stability in the univariate case with an unit-Fréchet margin

\[ F^t(tx) = F(x), \text{ for } F(x) = \exp(-1/x) \]

Max-stability in the multivariate case with unit-Fréchet margins

\[ F^t(tx_1, \ldots, tx_m) = F(x_1, \ldots, x_m), \text{ for } F_i(x_i) = \exp(-1/x_i) \]
A central question

\[ \mathbb{P} [M(x) < u, M(x + h) < v] = ?? \]
Asymptotic theory

If one assumes that we have unit Fréchet margins then

$$\lim_{n \to \infty} \mathbb{P} \left[ \frac{M_n(x) - a_n}{b_n} \leq u, \frac{M_n(x + h) - a_n}{b_n} \leq v \right] = \exp \left[ -V_h(u, v) \right]$$

where

$$V_h(u, v) = 2 \int_0^1 \max \left( \frac{w}{u}, \frac{1 - w}{v} \right) dH_h(w)$$

with $H_h(.)$ a distribution function on $[0, 1]$ such that $\int_0^1 w \, dH_h(w) = 0.5$. 
Bivariate case \((M(x), M(x + h))\)

Complex non-parametric structure

\[
V_h(u, v) = 2 \int_0^1 \max \left( \frac{w}{u}, \frac{1 - w}{v} \right) \, dH_h(w)
\]

Special case \(u = v\)

Note \(V_h(u, u) = V_h(1, 1)/u\) and define

\[
\theta(h) := V_h(1, 1)
\]

\[
P[M(x) < u, M(x + h) < u] = \exp(-\theta(h)/u)
\]

\[
= F(u)^{\theta(h)}
\]

because \(F(u) = \exp(-1/u)\)
\( \theta(h) = \text{Extremal coefficient} \)

\[
\mathbb{P} [M(x) < u, M(x + h) < u] = F(u)^{\theta(h)}
\]

**Interpretation**

- Independence \( \Rightarrow \theta(h) = 2 \)
- \( M(x) = M(x + h) \Rightarrow \theta(h) = 1 \)
- Similar to correlation coefficients for Gaussian but ...
- No characterization of the **full** bivariate dependence
Geostatistics: Variograms

Complex non-parametric structure

\[ \gamma(h) = \frac{1}{2} \mathbb{E}|Z(x + h) - Z(x)|^2 \]

- Finite if light tails
- Capture all spatial structure if \( \{Z(x)\} \) Gaussian fields
- but not well adapted for extremes
A Different Variogram

\[ \nu_h = \frac{1}{2} \mathbb{E} |F(M(x + h)) - F(M(x))| \]

Why does it work?

\[ \frac{1}{2} |a - b| = \max(a, b) - \frac{1}{2} (a + b) \]

- \( a = F(M(x + h)) \) and \( b = F(M(x)) \)
- \( \mathbb{E} a = \mathbb{E} b = 1/2 \)

\[ \mathbb{E} \max(a, b) = \mathbb{E} F(\max(M(x + h), M(x))) = \frac{\theta(h)}{1 + \theta(h)} \]
Madogram $\nu_h \Rightarrow$ Extremal coeff $\theta(h)$

$$\theta(h) = \frac{1 + 2\nu_h}{1 - 2\nu_h}$$

- The madogram $\nu_h$ gives the extremal coefficient $\theta(h)$
Schlather’s models (2003)

\[ \theta(h) = 1 + \sqrt{1 - \frac{1}{2} \left( \exp\left(-\frac{h}{40}\right) + 1 \right)} \]
Madogram $\nu_h \Rightarrow$ Extremal coeff $\theta(h)$

Schlather’s fields
The $\lambda$-madogram

\[ \nu_h(\lambda) = \frac{1}{2} \mathbb{E} \left| F^\lambda(M(x + h)) - F^{1-\lambda}(M(x)) \right| \]

**Properties**

- Defined for light & heavy tails
- Called a $\lambda$-Madogram
- Nice links with extreme value theory
- $\nu_h(0) = \nu_h(1) = 0.25$
A fundamental relationship

\[ \nu_h(\lambda) = \frac{V_h(\lambda, 1 - \lambda)}{1 + V_h(\lambda, 1 - \lambda)} - c(\lambda), \text{ with } c(\lambda) = \frac{3}{2(1 + \lambda)(2 - \lambda)} \]

Conversely,

\[ V_h(\lambda, 1 - \lambda) = \frac{c(\lambda) + \nu_h(\lambda)}{1 - c(\lambda) - \nu_h(\lambda)} \]
The $\lambda$-madogram
30-year maxima of daily precipitation in Bourgogne

146 stations of maxima of daily precipitation over 1970-1999 in Bourgogne
54-year maxima of daily precipitation in Belgium

55 stations of the Climatological network

55 stations of precipitation maxima over 1951-2005 in Belgium
Examples: fitting multivariate maxima

Air pollutants (Leeds, UK, winter 94-98, daily max) NO vs. PM10 (left), SO2 vs. PM10 (center), and SO2 vs. NO (right) (Heffernan & Tawn 2004, Boldi & Davison, 2007)

Our strategy

1. Assume observations arise from a max-stable process
2. Find and fit a flexible parametric model for the spectral density
3. Two desiderata: (A) interpretable parameters & (B) going beyond the bivariate case
Multivariate Max-Stable Distributions

If $\mathbf{Z} = (Z(x_1), \ldots, Z(x_p))^T$ has a multivariate max-stable distribution with unit Fréchet margins ($\mathbb{P}(Z(x_i) \leq z) = \exp(-z^{-1})$) then:

$$G(z) = \mathbb{P}(\mathbf{Z} \leq \mathbf{z}) = \exp[-V(\mathbf{z})],$$

where

$$V(\mathbf{z}) = p \int_{S_p} \max_i \left( \frac{w_i}{z_i} \right) dH(\mathbf{w}),$$

$H$ is a positive measure on $S_p$, s.t.

$$\int_{S_p} w_i dH(\mathbf{w}) = 1/p,$$

and $S_p = \{ \mathbf{w} \in \mathbb{R}_+^p | w_1 + \ldots + w_p = 1 \}$. 

## Models for Multivariate MSD’s

<table>
<thead>
<tr>
<th>Exponent measure function $V(z)$</th>
<th>Spectral density $h(w)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>Dirichlet</td>
</tr>
<tr>
<td>Asymmetric Logistic (Tawn, 88)</td>
<td>(Coles &amp; Tawn, 91)</td>
</tr>
<tr>
<td>Negative Logistic (Joe, 90)</td>
<td>Dirichlet mixture</td>
</tr>
<tr>
<td></td>
<td>(Boldi &amp; Davison, 2006)</td>
</tr>
<tr>
<td></td>
<td><strong>Pairwise Beta</strong></td>
</tr>
</tbody>
</table>

- Can obtain $G(z)$
  - Overparametrized?
  - Less flexible?

- More flexibility?
  - Cannot directly get $G(z)$
Dirichlet model (Coles, Tawn, 1991)

\[
h(w; \theta) = \frac{1}{p} (m \cdot w)^{-(p+1)} \prod_{j=1}^{p} m_j h^* \left( \frac{m_1 w_1}{m \cdot w}, \ldots, \frac{m_p w_p}{m \cdot w}; \theta \right)
\]

A special case: Dirichlet model

\[
h^*(w; \alpha) = \frac{\Gamma(\alpha \cdot 1)}{\prod_{j=1}^{p} \Gamma(\alpha_j)} \prod_{j=1}^{p} w_j^{\alpha_j^{-1}}, \quad \alpha_j > 0, j = 1, \ldots, p.
\]

\[
h(w; \alpha) = \frac{1}{p} \prod_{j=1}^{p} \frac{\alpha_j}{\Gamma(\alpha_j)} \frac{\Gamma(\alpha \cdot 1 + 1)}{(\alpha \cdot w)^{p+1}} \prod_{j=1}^{p} \left( \frac{\alpha_j w_j}{\alpha \cdot w} \right)^{\alpha_j^{-1}}
\]
Our Pairwise Beta Model

\[ h_p(w; \alpha, \beta) = K_p(\alpha) \sum_{i \neq j} h_{i,j}(w_i, w_j; \alpha, \beta_{i,j}), \text{ where} \]

\[ h_{i,j}(w_i, w_j; \alpha, \beta_{i,j}) = (w_i + w_j)^{(p-1)(\alpha-1)}(1 - (w_i + w_j))^{\alpha-1} \times \]

\[ \frac{\Gamma(2\beta_{i,j})}{(\Gamma(\beta_{i,j}))^2} \left( \frac{w_i}{w_i + w_j} \right)^{\beta_{i,j}^{-1}} \left( \frac{w_j}{w_i + w_j} \right)^{\beta_{i,j}^{-1}} \]

Advantages:
- no adjustment necessary to get center of mass condition
  \[ \int w_j dH(w) = 1/p \]
- parameters have some interpretation: \( \alpha \) controls overall dependence,
  \( \beta_{i,j} \)'s control pairwise dependence
- largely specified by pairwise parameters
- Middle ground between Coles & Tawn (1991) and Boldi & Davison (2007)
Pairwise Beta Models

$\alpha = 1, \beta = (2, 4, 15)$

$\alpha = 4, \beta = (2, 4, 15)$

$\alpha = 1, \beta = (2, .5, .5)$

$\alpha = 1, \beta = (2, 2, .5)$
Fitting the spectral density model


(a) have common marginals with unit tail index
(b) transform into polar coordinates and select exceedances above \( t_0 \)
(c) maximize the likelihood

\[
L(\theta; (r(i), w(i)), i = 1, \ldots, N_{t_0}) \approx \exp(-\nu(A)) \prod_{i=1}^{N_{t_0}} d\nu(r(i), w(i)) = \exp(-t_0^{-1}) \prod_{i=1}^{N_{t_0}} r(i)^{-2} h(w(i), \theta),
\]
An example with $\alpha = 1, \beta = (2, 4, 15)$
An example with $\alpha = 1$, $\beta = (2, 4, 15)$ (\ldots = asymptotic mle) 200 real * 1000
Air pollutants example

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta_{1,2}$</th>
<th>$\beta_{1,3}$</th>
<th>$\beta_{1,4}$</th>
<th>$\beta_{1,5}$</th>
<th>$\beta_{2,3}$</th>
<th>$\beta_{2,4}$</th>
<th>$\beta_{2,5}$</th>
<th>$\beta_{3,4}$</th>
<th>$\beta_{3,5}$</th>
<th>$\beta_{4,5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM10, NO</td>
<td>0.31 (0.002)</td>
<td>4.04 (0.139)</td>
<td>29.69 (1.222)</td>
<td>0.33 (0.006)</td>
<td>0.81 (0.026)</td>
<td>3.51 (0.119)</td>
<td>0.34 (0.006)</td>
<td>0.53 (0.014)</td>
<td>0.61 (0.013)</td>
<td>0.45 (0.011)</td>
</tr>
<tr>
<td>PM10, NO2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM10, O3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Scatterplots of NO vs. PM10 (left), SO2 vs. PM10 (center), and SO2 vs. NO (right).
Air pollutants example

100 largest observations.
corners = PM10 (lower right), NO (upper left), SO2 (lower left)
Summary of our spectral approach

- “Simple” and flexible spectral density with interpretable parameters
- Can be used for prediction or interpolation purposes
- Can be generalized (Ballani, Schlather, 2010)
- Can be extended to the asymptotic independent case (Qin, Smith, Ren, 2008)
BHM and bivariate max-stable (e.g., Ribatet, Cooley and Davison, 2010)

Figure 6: Comparison between one realization of the observed field and one realization of the different models analyzed: (a) observed field; (b) conditional independence model; (c) max-stable hierarchical model without adjustment; and (d) max-stable hierarchical model with adjustment. The same seed was used for all the simulations.

The benefit of the max-stable hierarchical model over the conditional independence model is that the max-stable model is able to account for local dependence. Given only fifty locations in the region, the model seems to be able to detect the true pattern of local dependence. The 95% credible intervals for the elements of $\Sigma$ are $(5.39, 8.76)$, $(-1.28, 0.67)$, $(5.58, 8.37)$ for $\sigma_{11}$, $\sigma_{12}$, and $\sigma_{22}$ respectively, which include the true values 6, 0, and 6. The fitted max-stable model provides a mechanism for producing realistic draws from the spatial process. As Figure 6 shows, a draw from the posterior distribution of the conditional independence model would be inappropriate and unrealistic for spatial phenomena such as rainfall or temperature which would produce a much smoother surface.

These results are obtained from a (near) perfect model simulation; that is, the max-stable hierarchical model fitted to the data was nearly identical to that from which the data were simulated. Nevertheless, this simulation exercise shows that the adjusted max-stable hierarchical model is able both to flexibly model marginal behavior that captures regional spatial effects and to capture local dependence via the max-stable process model. In the next section we show that it also seems to perform well on real data.

5 Application

We analyze data on maximum daily rainfall amounts for the years 1962–2008 at 51 sites in the Plateau region of Switzerland; see Figure 7. The area under study is relatively flat, the altitudes of the sites varying...
Take home messages

- Multivariate EVT may help characterizing extremes dependencies in space and time
- BHM offers flexibility and geostatistical tools can also be taken advantage of
- Physical knowledge should be integrated into the statistical analysis
- Computational issues can be arisen quickly
- Modeling trade off between parametric and non-parametric approaches
- Asymptotic independence can be an issue
- Extremes here means very rare

Two advertisements

- A two-year postdoc for the ACQWA European project
- A young researcher school on MEVT (Aussois, France, 3-5 Sep 2010)
Comparisons with other estimators

Gumbel (1960)

\[ \mathbb{P}(X \leq x, Y \leq y) = \exp \left\{ - \left[ \left( \frac{1}{x} \right)^{\frac{1}{\alpha}} + \left( \frac{1}{y} \right)^{\frac{1}{\alpha}} \right] \alpha \right\} \]

Four estimators
- Pickands’ estimator (1975)
- Deheuvels’ estimator (1991)
- Hall and Tajvidi’s estimator (2000)
- Capéraà et al. (1997) estimator
Comparisons with other estimators

\[ \alpha = 0.3 \]

\[ \alpha = 0.7 \]