Modeling maxima in climate studies

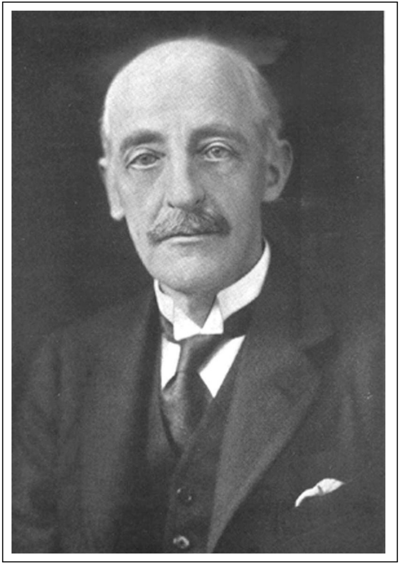
naveau@lsce.ipsl.fr
Laboratoire des Sciences du Climat et l'Environnement (LSCE)
Gif-sur-Yvette, France

AssimilEx-ANR members (Gwladys Toulemonde, Armelle Guillou, etc)
Dan Cooley (Colorado State University)
Richard Davis (Columbia University)

17 septembre 2008

Statistics and Earth sciences

"There is, today, always a risk that specialists in two subjects, using languages full of words that are unintelligible without study, will grow up not only, without knowledge of each other's work, but also will ignore the problems which require mutual assistance".



One advertisement

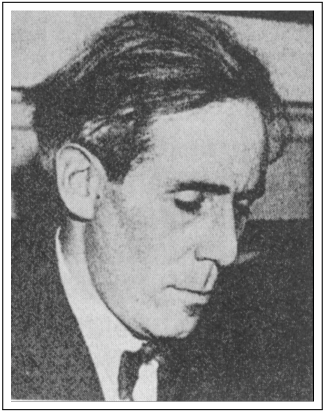
Graybill VIII
6th International Conference on Extreme Value Analysis

Home

Workshop - June 22, 2009
Conference - June 23-26, 2009

	Program Committee 2009	**2007**
Abstract Submission Registration Hotel Transportation Maps 2003-2008	Dan Cooley Richard Davis Paul Embrechts Anne-Laure Fougères Ivette Gomes Jürg Hüsler Rick Katz Claudia Klüppelberg Thomas Mikosch Philippe Naveau Liang Peng Holger Rootzén	

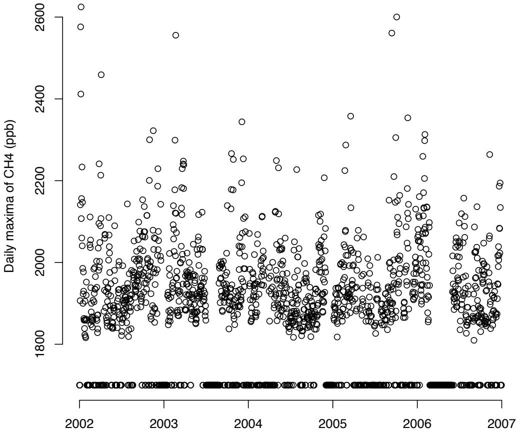
A quote from Emil Gumbel



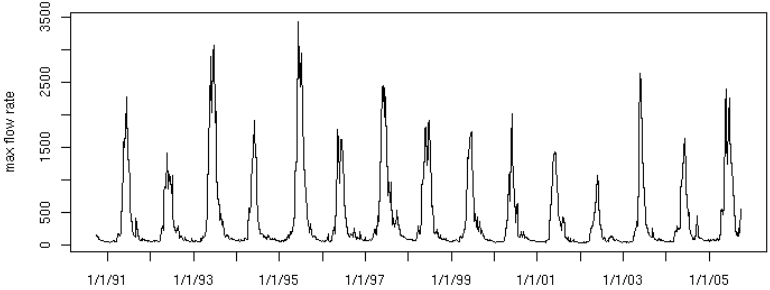
"It seems that the rivers know the theory. It only remains to convince the engineers of the validity of this analysis."

Emil Gumbel (1891-1966) 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.

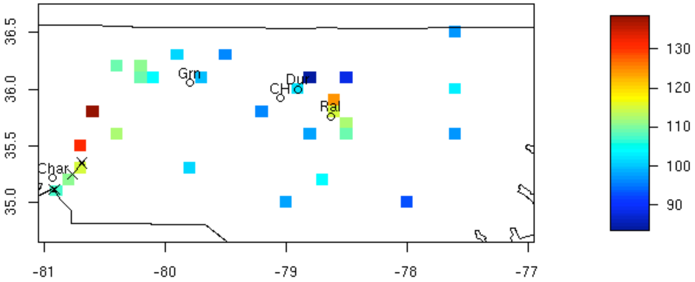
Daily maxima of CH_4 at Gif-sur-Yvette



Crystal River weekly max flow



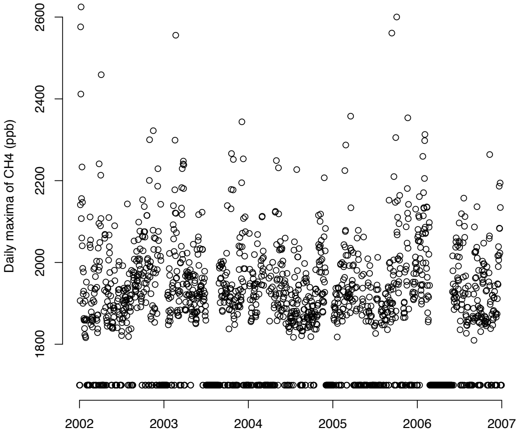
Ozone maxima in 1999



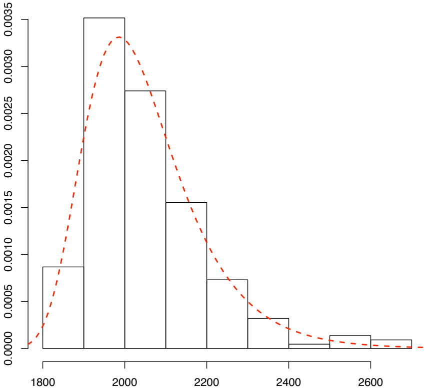
Modeling dependencies among maxima

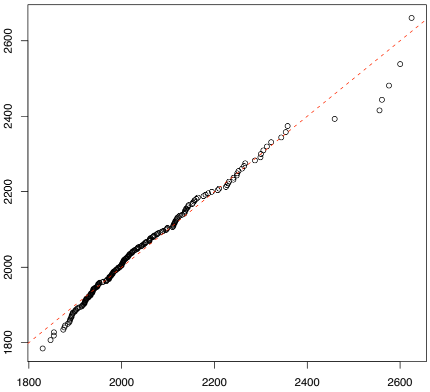
- **Max-stable processes** : Adapting asymptotic results for multivariate extremes
  e.g. Davis and Resnick (1989), Smith and Weissman (1996), Ferreira and de Haan (2006), Bacro et al. (2007), Naveau et al. (2008)
- **Bayesian or latent models** : temporal structure **indirectly** modeled via the GEV parameters distribution
  e.g., Coles & Tawn (1996), Cooley et al. (2007), Mendes, Turkman, Corte Real (2006)

Daily maxima of CH_4 at Gif-sur-Yvette

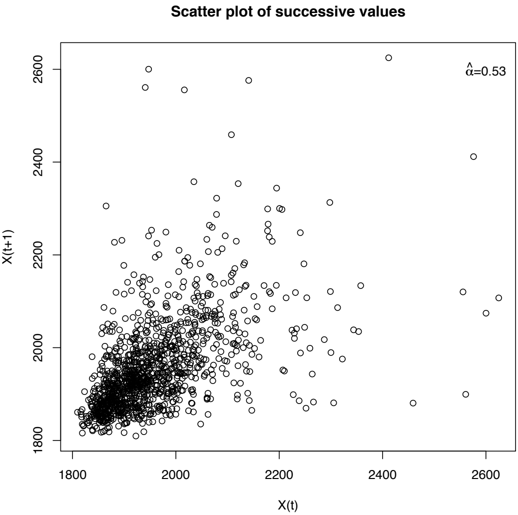


Maxima of CH_4 at Gif-sur-Yvette





Daily maxima of CH_4 at Gif-sur-Yvette



A key linear relationship (Tawn, 1990)

$$\text{Gumbel}(\mu_1 + \mu_2, \sigma/\alpha) = \mu_2 + \sigma \log S + \text{Gumbel}(\mu_1, \sigma)$$

where Gumbel(μ_1, σ) is a Gumbel r.v. and independent of S that is a positive $\alpha-$**stable** ($\alpha \in (0, 1]$) with Laplace transform

$$\mathbb{E}(\exp(-uS)) = \exp(-u^{\alpha}), \text{ for all } u > 0$$

Fougères et al. (2008)'s result

If

$$Y_t = F_t \log\left(\sum_{a \in A} c_{t,a} S_a\right) + \epsilon_t, \text{ with } t = 1, \dots, T,$$

where $\{c_{t,a} \geq 0\}$, $\{S_a, a \in A\}$ are independent positive α-stable variables, ϵ_t follows a iid Gumbel(μ_t, F_t), and all variables are mutually independent, then

$$\mathbb{P}(Y_1 \leq x_1, \dots, Y_T \leq x_T) = \prod_{a \in A} \exp\left(-\left(\sum_{t \in T} c_{t,a} e^{-\frac{x_t - \mu_t}{F_t}}\right)^{\alpha}\right)$$

Our Gumbel autoregressive model

$$\text{Gumbel}(\mu_1 + \mu_2, \sigma/\alpha) = \mu_2 + \sigma \log S + \text{Gumbel}(\mu_1, \sigma)$$

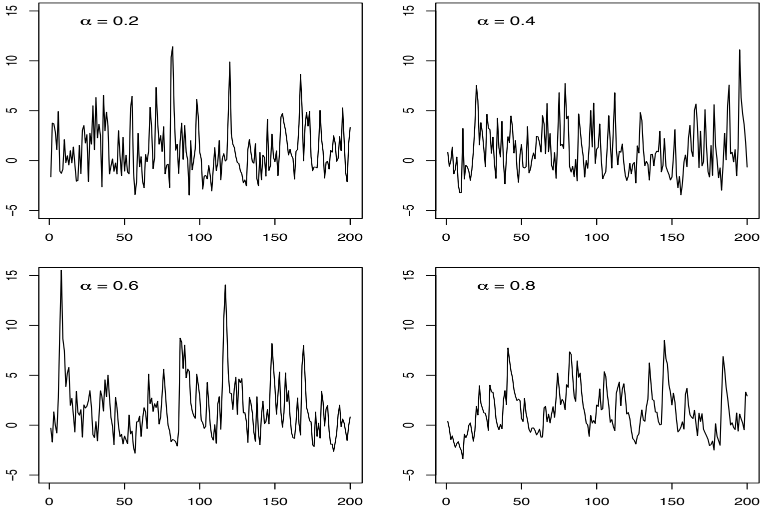
Toulemonde et al., 2008

$$X_t = \alpha X_{t-1} + \alpha\beta \log S_t$$

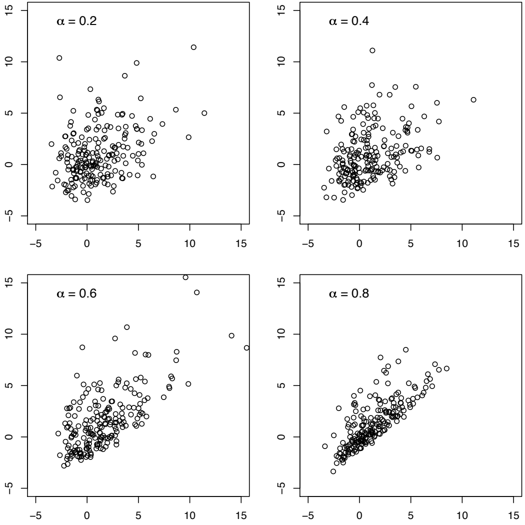
where S_t idd positive $\alpha-$**stable** noise with $\alpha \in (0, 1)$

Then X_t is Gumbel(0, β)

$X_t = \alpha X_{t-1} + \alpha\beta \log S_t$ with S_t idd positive $\alpha-$stable



$X_t = \alpha X_{t-1} + \alpha\beta \log S_t$ with S_t idd positive $\alpha-$stable

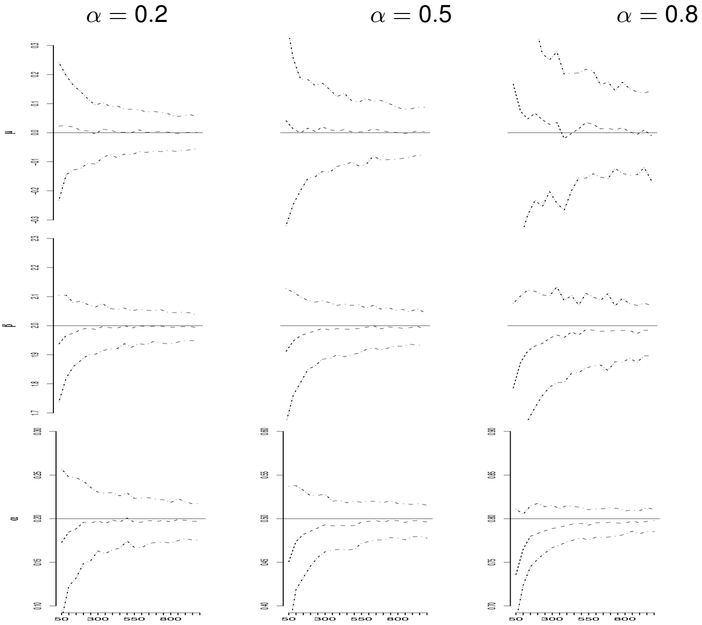


$X_t = \alpha X_{t-1} + \alpha\beta \log S_t + \mu(1-\alpha)$ with S_t idd positive $\alpha-$stable

The characteristic function of $\mathbf{X_h} = (X_t, \ldots, X_{t-h})^t$ is

$$\mathbb{E}(e^{i<u,\mathbf{X_h}>}) = e^{i\mu(\sum_{j=0}^{h} u_j)}\Gamma\left(1 - i\beta\sum_{j=0}^{h} u_j\alpha^{h-j}\right)\prod_{j=0}^{h-1}\frac{\Gamma(1 - i\beta\sum_{k=0}^{j} u_k\alpha^{j-k})}{\Gamma(1 - i\beta\sum_{k=0}^{j} u_k\alpha^{j-k+1})}.$$

$X_t = \alpha X_{t-1} + \alpha\beta \log S_t$ with S_t idd positive $\alpha-$stable



Dependencies in $X_t = \alpha X_{t-1} + \alpha\beta \log S_t$

1. $$\chi = \lim_{x\to\infty} \frac{\mathbb{P}(X_{t-1} > x, X_t > x)}{\mathbb{P}(X_{t-1} > x)}$$

   is equal to zero which corresponds to the asymptotic independence.

2. $$\overline{\chi} = \lim_{x\to\infty} \frac{2\log \mathbb{P}(X_{t-1} > x)}{\log \mathbb{P}(X_{t-1} > x, X_t > x)} - 1$$

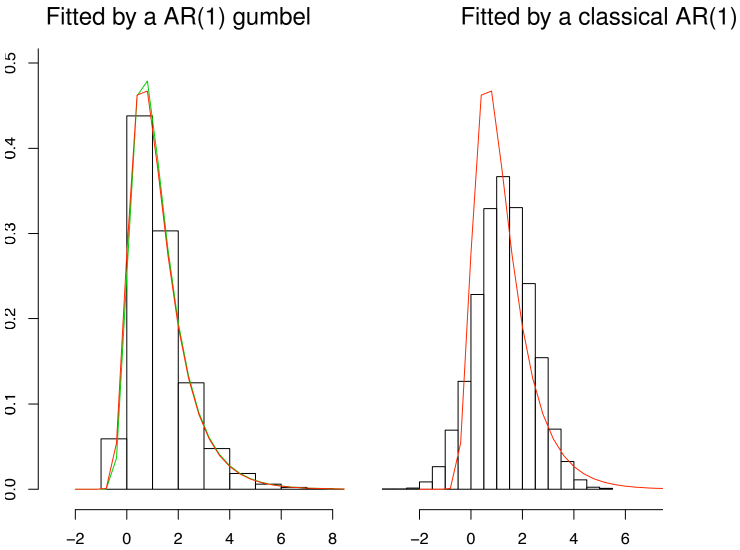
   increases with dependence strength and is equal to $\alpha/(2-\alpha) \in (0,1)$.

Asymptotic behavior of the method-of-moments estimators

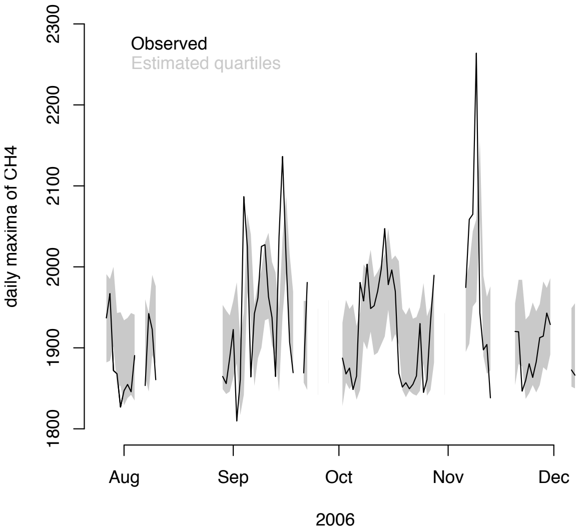
The rescaled method-of-moments estimators of μ, β and α converges to a zero-mean Gaussian vector with covariance

$$\begin{pmatrix} \frac{\pi^2\beta^2}{6}\frac{1+\alpha}{1-\alpha} - \frac{12\delta\beta^2\zeta(3)(1+\alpha+\alpha^2)}{\pi^2(1-\alpha^2)} + \frac{11\delta^2\beta^2(1+\alpha^2)}{10(1-\alpha^2)} & \frac{6\beta^2\zeta(3)(1+\alpha+\alpha^2)}{\pi^2(1-\alpha^2)} - \frac{11\delta\beta^2(1+\alpha^2)}{10(1-\alpha^2)} & -\alpha\beta\delta \\ \frac{6\beta^2\zeta(3)(1+\alpha+\alpha^2)}{\pi^2(1-\alpha^2)} - \frac{11\delta\beta^2(1+\alpha^2)}{10(1-\alpha^2)} & \frac{11\beta^2(1+\alpha^2)}{10(1-\alpha^2)} & \alpha\beta \\ -\alpha\beta\delta & \alpha\beta & 1-\alpha^2 \end{pmatrix}$$

One-step prediction with simulated data $\alpha = 0.5$



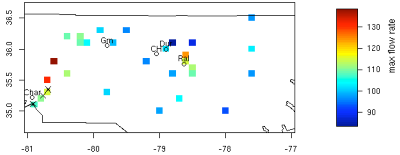
One-step prediction with CH_4 maxima

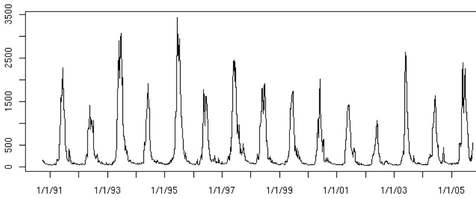


A summary of our Gumbel AR

- **+** A very simple additive model with a basic temporal dependence
- **+** Classical estimations and mathematical properties
- **+** May be useful for modeling CH_4 maxima
- **-/+** Has to be extended from Gumbel to GEV
- **-** A very specific type of temporal dependence

Ozone and flow maxima





Our strategy

1. Assume observations arise from a max-stable process
2. Find and fit parametric model for the spectral denisity
3. Approximate the **conditional density** of the unmonitored location given the “nearby” observations.

Multivariate Max-Stable Distributions

If $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_p))^T$ has a multivariate max-stable distribution with unit Fréchet margins ($\mathbb{P}(Z(\boldsymbol{x}_i) \leq z) = \exp(-z^{-1})$) then :

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

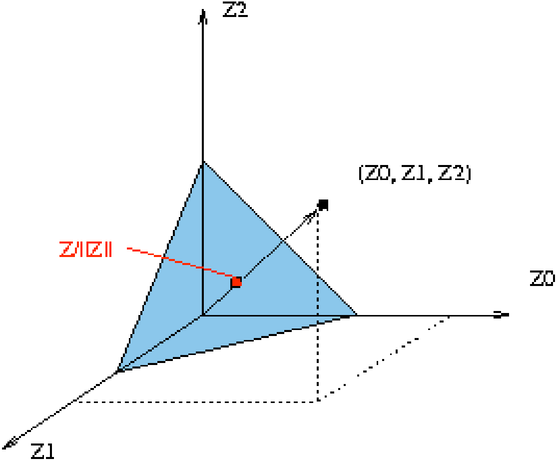
$$V(\boldsymbol{z}) = \int_{S_p} \max_i \left(\frac{w_i}{z_i} \right) dH(\boldsymbol{w}),$$

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

$$\int_{S_p} w_i dH(\boldsymbol{w}) = 1,$$

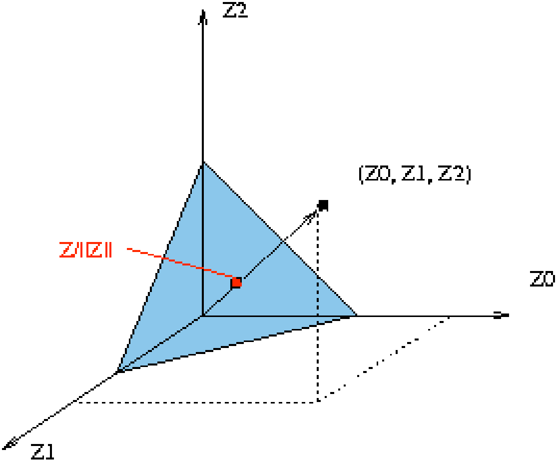
and $S_p = \{\boldsymbol{w} \in \mathbb{R}_+^p | w_1 + \ldots + w_p = 1\}$.

Multivariate Max-Stable Distributions



$V(\boldsymbol{z})$: "exponent measure function" with $\boldsymbol{z} \in \mathbb{R}^p$; $\boldsymbol{w} \in S_p$ – relates to point process intensity, $G(\boldsymbol{z}) = \exp[-V(\boldsymbol{z})]$
$H(\boldsymbol{w})$: "spectral measure" – lives on unit simplex, meets center of mass condition, $V(\boldsymbol{z}) = \int_{S_p} \max_i \left(\frac{w_i}{z_i}\right) dH(\boldsymbol{w})$
$h(\boldsymbol{w})$: "spectral density" – exists if $H(\boldsymbol{w})$ is differentiable, not Fourier !

Multivariate Max-Stable Distributions



- no parametric form for entire family
- a few useful parametric sub-families suggested

Models for Multivariate MSD's

Exponent measure function $V(\boldsymbol{z})$	Spectral density $h(\boldsymbol{w})$
- Logistic - Asymmetric Logistic (Tawn, 88) - Negative Logistic (Joe, 90)	- Dirichlet (Coles & Tawn, 91) - Dirichlet mixture (Boldi & Davison, 2006) - **Pairwise Beta**
+ Can obtain $G(\boldsymbol{z})$ – Overparametrized ? – Less flexible ?	+ More flexibility ? – Cannot directly get $G(\boldsymbol{z})$

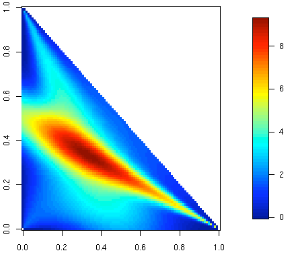
Pairwise Beta Model

$$
\begin{aligned}
h_p(\boldsymbol{w};\alpha,\boldsymbol{\beta}) &= K_p(\alpha)\sum_{i\neq j} h_{i,j}(w_i,w_j;\alpha,\beta_{i,j}), \text{ where} \\
h_{i,j}(w_i,w_j;\alpha,\beta_{i,j}) &= (w_i+w_j)^{(p-1)(\alpha-1)}(1-(w_i+w_j))^{\alpha-1}\times \\
&\quad \frac{\Gamma(2\beta_{i,j})}{(\Gamma(\beta_{i,j}))^2}\left(\frac{w_i}{w_i+w_j}\right)^{\beta_{i,j}-1}\left(\frac{w_j}{w_i+w_j}\right)^{\beta_{i,j}-1}
\end{aligned}
$$

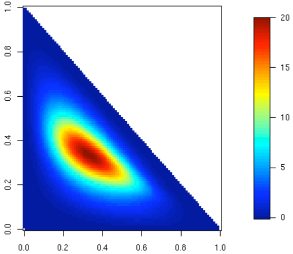
Advantages :

- no adjustment necessary to get center of mass condition
- parameters have some interpretation : α controls overall dependence, $\beta_{i,j}$'s control pairwise dependence
- largely specified by pairwise parameters

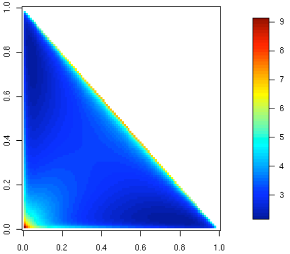
Pairwise Beta Models



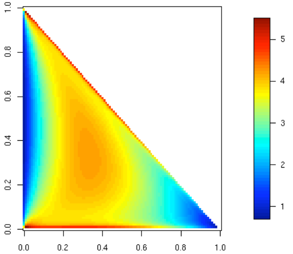
$\alpha = 1, \boldsymbol{\beta} = (2, 4, 15)$



$\alpha = 4, \boldsymbol{\beta} = (2, 4, 15)$

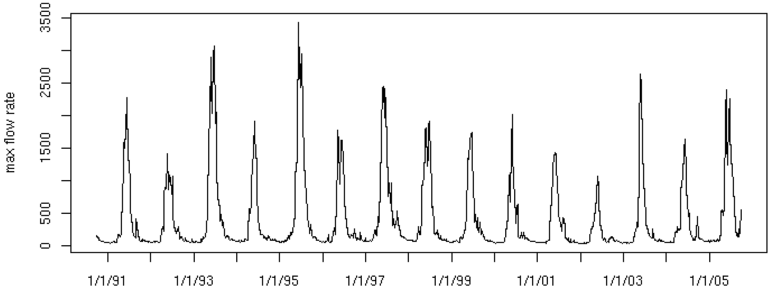


$\alpha = 1, \boldsymbol{\beta} = (2, .5, .5)$



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

River flow Example



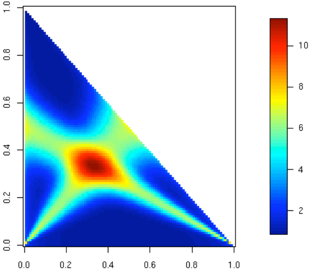
1. Data are deseasonalized.
2. ACF/PACF plots → base prediction on previous two observations.
3. GEV is fit to marginal distribution, then transformed.

Fitting the spectral density model

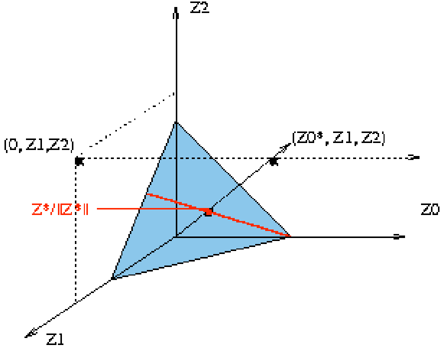
Time series broken into non-overlapping triples. Dependence measured by the extremal coefficient (Schlather & Tawn 03).

$$\phi_{1,2} = V(1, 1, \infty, \ldots, \infty); \ \phi_{i,j} \in [1, 2]$$

Ext. coefficients estimated at lags 1 and 2, $(\hat{\phi}_{-1}, \hat{\phi}_{-2}) = (1.36, 1.49)$, and pairwise beta model parameters found to match the extremal coefficient estimates. $(\hat{\alpha}; \hat{\beta}) = (1, 16, 0.7, 16)$.



Prediction : Approximating the conditional density ?



If $V(\boldsymbol{z})$ is known and differentiable, then joint density can be obtained exactly. However, we are modeling $h(\boldsymbol{w})$.

Assume Z_1, Z_2 are observed and Z_0 is unobserved. Any predictor Z_0^* will yield a point $\boldsymbol{Z}^* = (Z_0^*, Z_1, Z_2)$ which can be mapped back to S_p as $\frac{\boldsymbol{Z}^*}{\|\boldsymbol{Z}^*\|_1}$.

Approximating the conditional density ?

If $V(\boldsymbol{z}) = \mu\{(0, z]^c\}$ is small (i.e. the radius is large), then

$$G(\boldsymbol{z}) = \exp(-V(\boldsymbol{z})) \approx 1 - V(\boldsymbol{z}).$$

Using Coles and Tawn (91) result to estimate the density at $\boldsymbol{z}$:

$$g(\boldsymbol{z}) \approx \frac{\partial}{\partial z_1, \ldots, \partial z_p}[1 - V(\boldsymbol{z})] = \frac{1}{||\boldsymbol{z}||^{-(p+1)}} h\left(\frac{\boldsymbol{z}}{||\boldsymbol{z}||}\right)$$

So conditional density can be approximated by

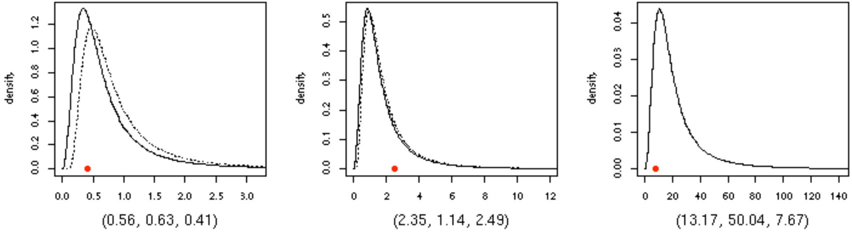
$$g_{Z_p|Z_1,\ldots Z_{p-1}}(z_p|z_1, \ldots, z_{p-1}) \approx \frac{\frac{1}{||\boldsymbol{z}||^{-(p+1)}} h\left(\frac{\boldsymbol{z}}{||\boldsymbol{z}||}\right)}{\int_0^\infty \frac{1}{||\boldsymbol{z}^*||^{-(p+1)}} h\left(\frac{\boldsymbol{z}^*}{||\boldsymbol{z}^*||}\right) d\zeta}$$

where $\boldsymbol{z}^* = (z_1, \ldots, z_{p-1}, \zeta)$.

Approximating the conditional density ?

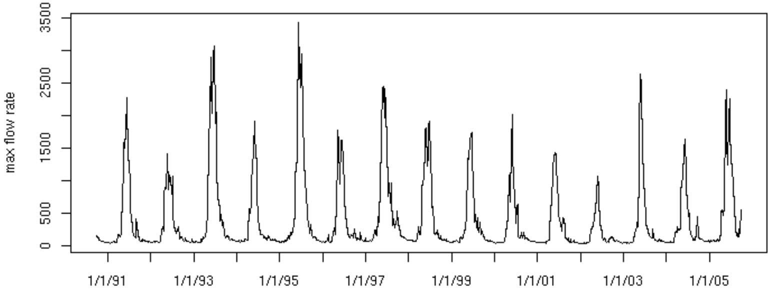
Three realizations from a trivariate symmetric logistic distribution.

True conditional density (solid line) and approximated conditional density (dotted line)



$||\boldsymbol{z}|| = 1.6$ $||\boldsymbol{z}|| = 6$ $||\boldsymbol{z}|| = 71$

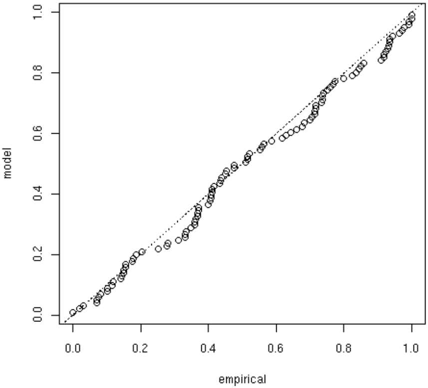
Time Series Example



1. Data are deseasonalized.
2. Prediction based on previous two observations.
3. GEV is fit to marginal distribution, then transformed.

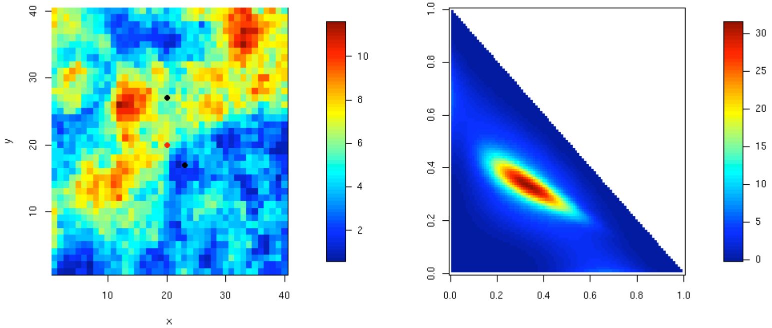
Time series prediction

75 largest triples selected for prediction. Conditional density of 3rd component given 1st and 2nd components is approximated.

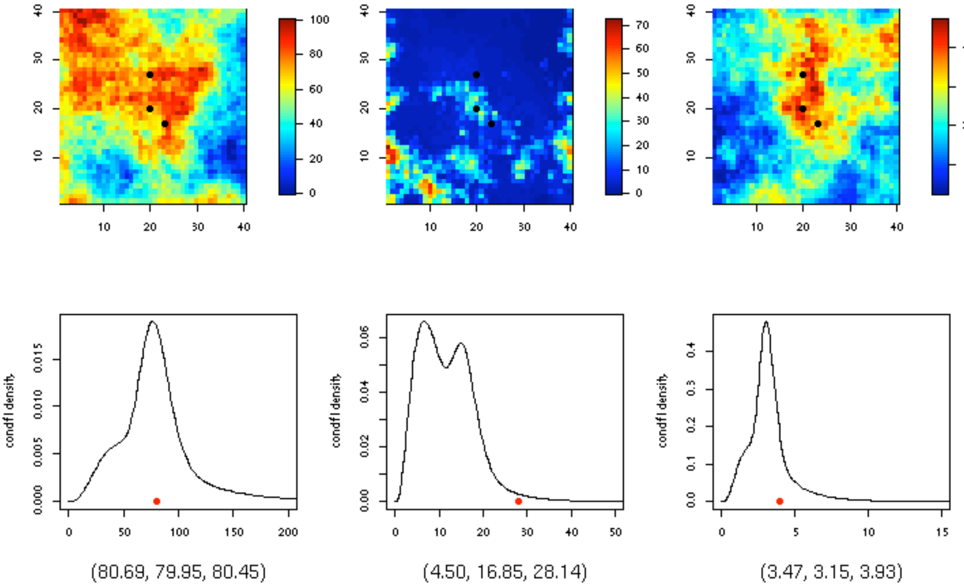


Spatial interpolation

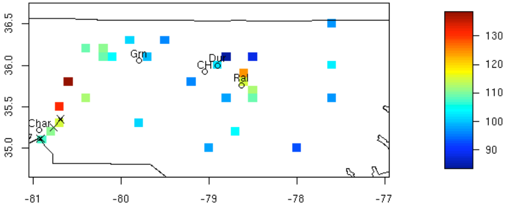
- MS fields w/ Fréchet marginals simulated (Schlather 02).
- Known bivariate dependence structure : $\phi = (1.34, 1.28, 1.22)$.
- Pairwise beta model fit as before : $(\alpha; \beta) = (4.3; 0.87, 4.4, 74)$.
- Conditional density approximated for largest 300 of 1000 simulated fields.



Spatial interpolation examples

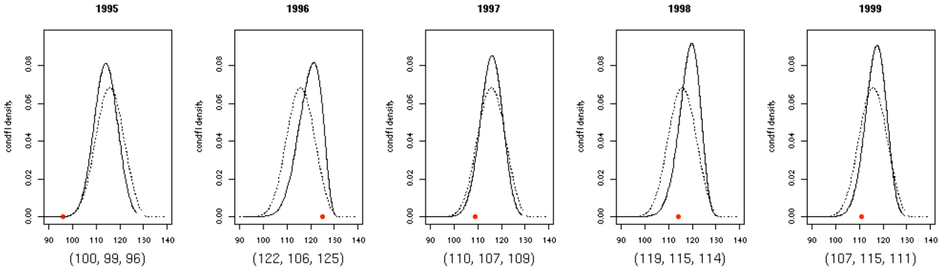


Ground level ozone



- Only five years of data.
- Marginal distributions from Gilleland, et al (2006).
- Dependence estimated as function of distance using madogram (Naveau et al 06).
- Weak dependence estimated $\phi = (1.95, 1.84, 1.67)$.

Ground level ozone



Marginal density (dotted lines) and conditional density (solid lines)

Summary of our spectral approach

- Method for approximating the conditional density of an unobserved component of a max-stable vector given the other components.
- Method designed specifically for dealing with extremes ; alternative to standard time series and spatial prediction methods which are better suited for central tendancies.
- A spectral density approach brings flexibility in modeling.
- Applied to both time series and spatial contexts.

The final thought on predictions in climate studies

"There is, today, always a risk that specialists in two subjects, using languages full of words that are unintelligible without study, will grow up not only, without knowledge of each other's work, but also will ignore the problems which require mutual assistance".—Sir Gilbert T. Walker (Walker, 1927b, page 321)

FIG. 3. *Photograph of Sir Gilbert T. Walker* (*source*: *Royal Society*; *Taylor*, 1962).

GEV autoregressive model

Toulemonde et al., 2007

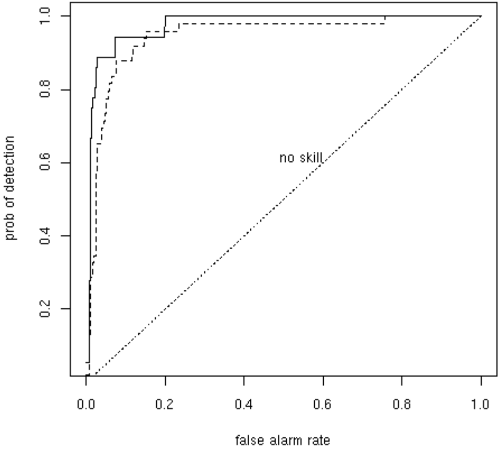
$$X_t = \left(X_{t-1} - \mu + \frac{\beta}{\xi}\right)^{\alpha} \times S_t^{\alpha\xi} \times \left(\frac{\beta}{\xi}\right)^{1-\alpha} + \mu - \frac{\beta}{\xi}$$

where S_t idd positive $\alpha-$**stable** noise with $\alpha \in (0, 1)$

Then X_t is GEV(μ, β, ξ)

Spatial interpolation results

How well does the method assess exceeding some standard ?



Repeated Simulation Results