Extreme Value Theory and state space modeling for environmental data

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

Gwladys Toulemonde and Anne Sabourin

21 juin 2011

Outline

1. How to be extremes ?
2. Basics about the EVT distributions
3. State-space modeling for elliptical distributed extremes (Anne Sabourin)
4. State-space modeling for Gumbel distributed maxima (Glwadys Toulemonde)

Gauss-Laplace distribution function

A very nice distribution

- The assumption of normality (or mixture of normals) is very prevalent in the theoretical and applied statistical research
- Asymptotic justification : Central Limit Theorem
- Nice properties of Gaussian vectors

- Completely characterized by its first two moments
- Stability under linearity
- Stability under summation
- Stability under conditioning

Two central questions

How to analyze extremes without losing the advantages of the Gaussian distributions ?

- Stability under linearity ?
- Stability under addition ?
- Stability under conditioning ?

How to take advantage of Extreme Value Theory in a state-space modeling context ?

- *Nous avons anticipé dans la mesure du possible mais on ne peut pas prévoir l'imprévisible"* **Xynthia's storm, 25th of Feb, 2010**
- *"Il est impossible que l'improbable n'arrive jamais"* Emil Julius Gumbel (1891-1966)

Max-stability

Let $M_n = \max(X_1, \ldots, X_n)$ with X_i iid with distribution F.

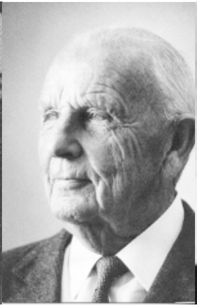
Definition : ***F*** **max-stable if**

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} < x\right) = F^n(a_n x + b_n) = F(x)$$

An example

Gumbel $F(x) = \exp(-\exp(-x))$ for all real x. Then $a_n = 1$ & $b_n = \log n$

An historical perspective of Extreme Value Theory

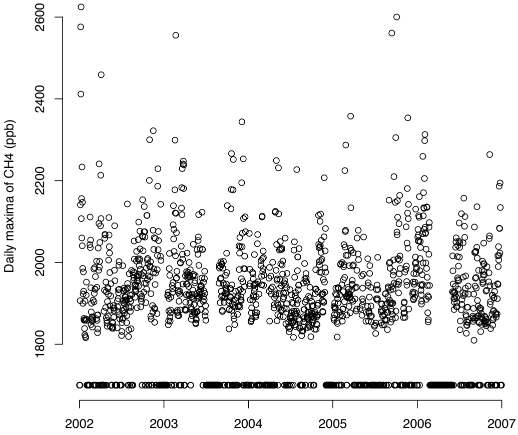


Gumbel (1891-1966) **Weibull (1887-1979)** **Fréchet (1878-1973)**

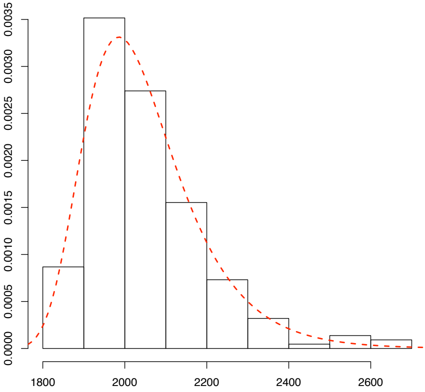
- Emil Gumbel was born and trained as a statistician in Germany, forced to move to France and then the U.S. because of his pacifist and socialist views. He was a pioneer in the application of extreme value theory, particularly to climate and hydrology.
- Waloddi Weibull was a Swedish engineer famous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

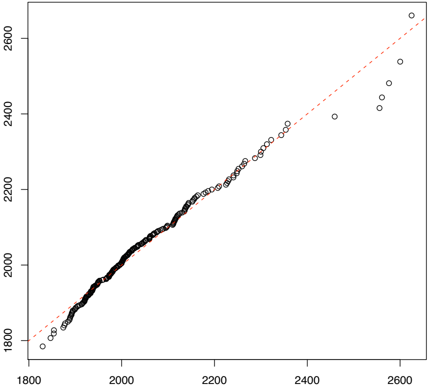
Other important names : Fisher and Tippet (1928), Gnedenko (1943), etc

Daily maxima of CH_4 at Gif-sur-Yvette (Toulemonde et al., 2009, Environmetrics)

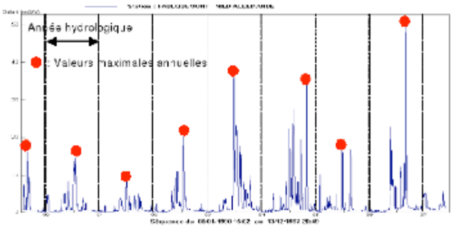


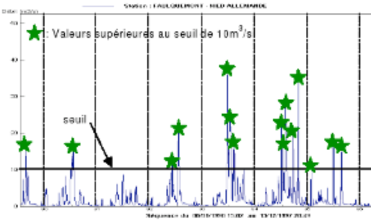
Maxima of CH_4 at Gif-sur-Yvette (Toulemonde et al., 2009, Environmetrics)





Peak over Threshold (POT)

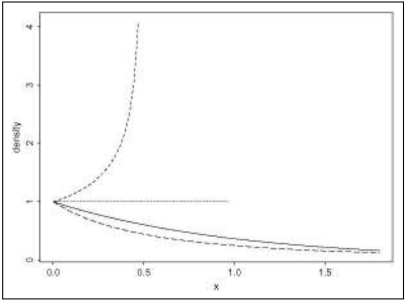




Thresholding : the Generalized Pareto Distribution (GPD)

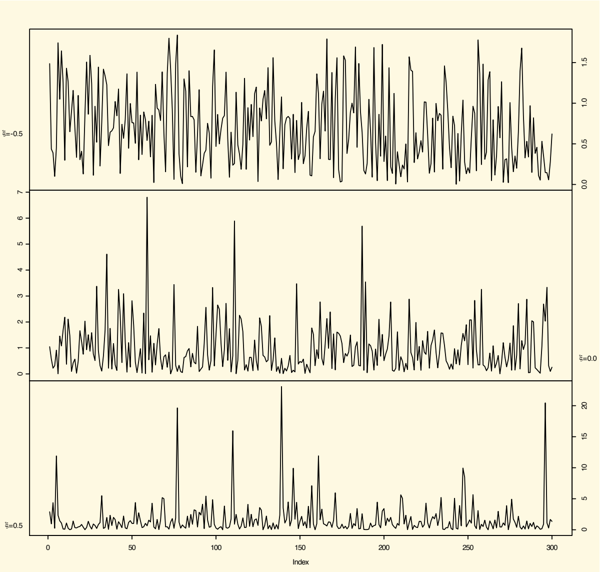
$$\mathbb{P}\{\mathbf{R}-u > y | \mathbf{R} > u\} = \left(1 + \frac{\xi\, y}{\sigma_u}\right)_+^{-1/\xi}$$

Vilfredo Pareto : 1848-1923



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.

GPD : "From Bounded to Heavy tails"



Elliptical distributions

A wide class, allowing for bounded or heavy tailed laws.

Definition

A random vector : $X \in \mathbb{R}^n$ with density f is elliptical with

- *Parameters* : $\mu \in \mathbb{R}^n$, $\Sigma \in \mathcal{M}_{n\times n}(\mathbb{R})$ a positive definite symmetric matrix
- *Density generator* g such that $\int_0^{+\infty} t^{n/2-1} g(t)dt < \infty$,

iff

$$f(x) = c_n |\Sigma|^{-1/2} g((x-\mu)'\Sigma^{-1}(x-\mu)),$$

$$c_n = \frac{\Gamma(n/2)}{\pi^{n/2}\int_0^{+\infty} t^{n/2-1} g(t)dt}$$

Gaussian vectors : a specific case of elliptical vectors with generator $g(s) = \exp(-\frac{s}{2})$ (see *e.g* [5] or [7])

Elliptical distributions

Any elliptical vector can be written as :

$$X = \mu + RA'U$$

where

- $U \in \mathbb{R}^n$ is uniformly distributed on the unit sphere
- $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is such that $A'A = \Sigma$
- R (called the radial variable) is a positive real random variable, independent from U and with density

$$h(r) = \frac{2}{\int t^{n/2-1} g(t)dt} r^{n-1} g(r^2) I_{[0,\infty[}(r)$$

An easy way to simulate elliptical distributions. see e.g [5] or [7]

Notations for conditioning

Crucial for filtering data !

Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $X_1 \in \mathbb{R}^p$, $X_2 \in \mathbb{R}^{n-p}$
Corresponding blocks for μ and Σ

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

see [5]

Margins, still elliptical

$$X_1 \sim \mathcal{E}_p(\mu_1, \Sigma_{11}, g_{(1)})$$

with

$$g_{(1)}(s) = \int_0^{+\infty} w^{\frac{n-p}{2}-1} g(s+w)dw$$

Conditioning, still elliptical

$$X_2 | (X_1 = x_1) \sim \mathcal{E}_{n-p}(\mu_{2|1}, \Sigma_{2|1}, g_{2|1})$$

with :

$$\begin{aligned}
\mu_{2|1} &= \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \\
\Sigma_{2|1} &= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \\
g_{2|1}(s) &= g(q_1 + s), q_1 = (X_1 - \mu_1)'\Sigma_{11}^{-1}(X_1 - \mu_1)
\end{aligned}$$

Same equations as for conditionals from Gaussian laws !

Elliptical distributions and Pareto generator

GP Distribution : Commonly used in hydrology to model threshold exceeding.

$$\begin{aligned} g_{\sigma,\xi}(s) &= \mathbb{P}(Y_{\sigma,\xi} > s) \text{ where } Y_{\sigma,\xi} \sim GPD(\sigma,\xi) \\ &= \left(\left(1+\frac{\xi s}{\sigma}\right)^{+}\right)^{\frac{-1}{\xi}} \text{ with } s \geq 0 \end{aligned}$$

- $\xi > 0$: heavy tailed distributions
- $\xi < 0$: bounded ones.
- Hereafter : $\xi > 0$ (case $\xi < 0$ is similar)

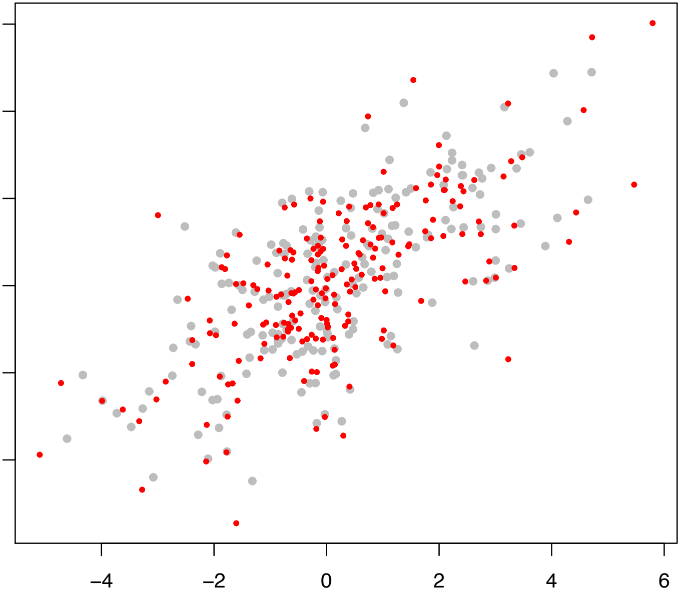
Elliptical distributions and Pareto generator

Fundamental property

$$g_{\sigma,\xi}(s+u) = g_{\sigma+\xi u}(s) g_{\sigma}(u)$$

A key to obtain explicit expressions for conditional and margins

Pareto versus exponential generators



AR(1) $X_t = FX_{t-1} + \epsilon_t$

Lemma

Elliptical innovations $\epsilon_t = (X_t - FX_{t-1}|x_{t-1})$ have GPD generator with parameters :

$$\tilde{\sigma} = \frac{\sigma + \xi q_{t-1}(x_{t-1})}{1 - \alpha\xi}, \qquad \tilde{\xi} = \frac{\xi}{1 - \alpha\xi}$$

Note : $q_{t-1}(x_{t-1})$ is as in (5)
Upper bound for $\tilde{\xi}$: $\tilde{\xi}_{sup} = \frac{1}{n}$

AR(1) $X_t = FX_{t-1} + \epsilon_t$

for $\xi > 0$

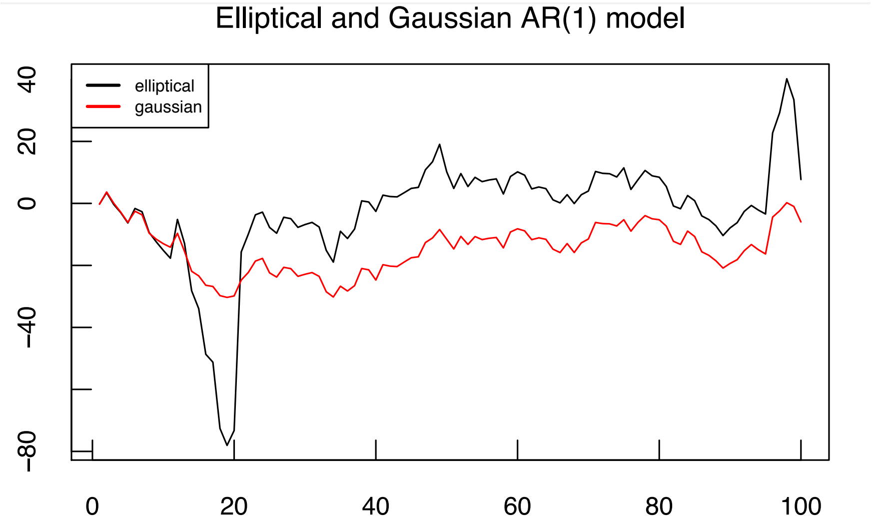
$$H_t(R) = \text{pbeta}_{\left(\frac{n}{2}, \frac{1}{\xi} - \frac{n(T-1)}{2} - \frac{n}{2}\right)}\left(\frac{\xi R^2}{\sigma + \xi(q_{t-1}(x_{t-1}) + R^2)}\right)$$

$$H_t^{-1}(u) = \sqrt{\left(\frac{\sigma}{\xi} + q_{t-1}(x_{t-1})\right)\frac{w_t^X(u)}{1 - w_t^X(u)}}$$

where

$$w_t^X(u) = pbeta^{-1}_{\left(\frac{n}{2}, \frac{1}{\xi} - \frac{n(T-1)}{2} - \frac{n}{2}\right)}(u)$$

AR(1) $X_t = FX_{t-1} + \epsilon_t$

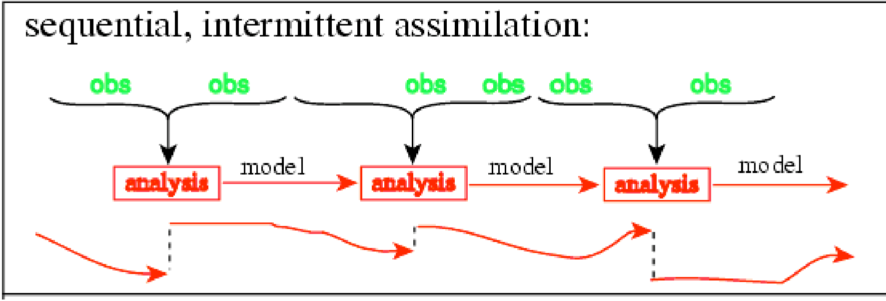


xi = 0.0187 ; sigma= 0.1869

max eigen value for noise = 4

Data assimilation

Basic principles

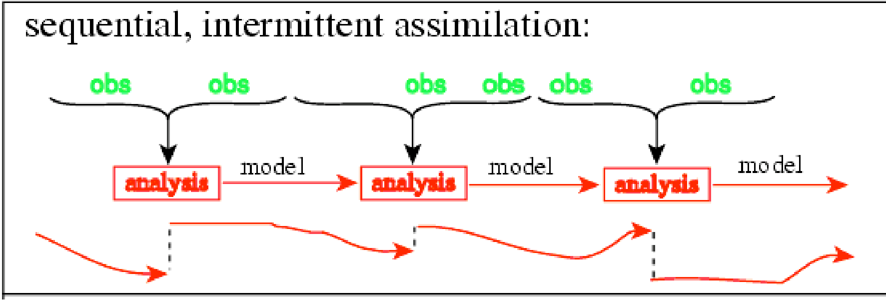


Dynamical structures :

- Kalman filter, Ensemble Kalman filters, Bayesian approaches

State-space modeling

Data assimilation techniques / State-space models



$$\text{Observational-equation} : Y_t = F_t(X_t, \epsilon_t),$$

$$\text{State-equation} : X_t = G_t(X_{t-1}, \eta_t),$$

where η_t and ϵ_t correspond to noises

State-space modeling

Classical state-space models used in geophysics :

$$Y_t = F_t X_t + \epsilon_t,$$
$$X_t = G_t X_{t-1} + \eta_t,$$

where η_t and ϵ_t correspond to **Gaussian** noises

Kalman filters

- Ensemble Kalman filters (Anderson, 2001)
- Mixture ensemble Kalman filters (Bengtsson, 2002)
- Particle filtering (Doucet at al., 2001)

A possible Kalman filter definition

A recursive procedure for inference about the state vector X_t (Meinhold and Singpurwalla 1993).

Kalman filters and elliptical distributions

- Elliptical innovations $\epsilon_t = (X_t - FX_{t-1}|x_{t-1})$ and $\nu_t = (Y_t - GX_t|x_t)$ with

$$\begin{aligned} \epsilon_t &\sim \mathcal{E}\left(0, \Sigma^{\epsilon}, g^{\epsilon}_{t,x_{t-1}}\right) \\ \nu_t &\sim \mathcal{E}\left(0, \Sigma^{\nu}, g^{\nu}_{t,x_t}\right) \end{aligned}$$

- Finite time process : $t \in \{0 : T\}$
- Elliptical global vector

$$\begin{aligned} W &= (X_0', X_1' \ldots X_T', Y_1' \ldots Y_T')' \\ &\sim \mathcal{E}_{nT+p}\left(0, \Sigma^W, g^W\right) \end{aligned}$$

- Result : Equations for estimates $\hat{x}_t$ and $\hat{\Sigma}_{X_t}$ are similar to those of the gaussian filter. Additional equations for conditional generators.

Kalman filters and elliptical distributions

block matrices in Σ^W

$$
\begin{aligned}
\Sigma_{X_t} &= \Sigma_\epsilon + F\Sigma_{X_{t-1}}F' \\
\Sigma_{X_t,X_{t-k}} &= F^k\Sigma_{X_{t-k}} \\
\Sigma_{Y_t,X_{t-k}} &= GF^k\Sigma_{X_{t-k}} \\
\Sigma_{Y_t} &= G\Sigma_{X_t}G' + \Sigma_\nu \\
\Sigma_{Y_t,Y_{t-k}} &= GF^k\Sigma'_{X_{t-k}}
\end{aligned}
$$

Kalman filters and elliptical distributions

Generators

$$
\begin{aligned}
g_t^{\nu}(s) &= \int_0^{+\infty} w^{\frac{nT+p-n}{2}-1} g^W(s+w+q_t(x_t)) \\
g_t^{\epsilon}(s) &= \int_0^{+\infty} w^{\frac{nT-p}{2}-1} g^W(s+w+q_{t-1}(x_{t-1}))
\end{aligned}
$$

with $q_t(x_t) = x_t'(\Sigma_{X_t})^{-1} x_t$

Kalman filters : bringing the GPD

Choose $g^W(s) = g_{\sigma,\xi}(s)$ as a global generator for W.
Upper bound for ξ : $\xi_{\text{sup}} = \frac{2}{nT+p}$

Lemma

Elliptical innovations $\epsilon_t = (X_t - FX_{t-1}|x_{t-1})$ *and* $\nu_t = (Y_t - GX_t|x_t)$ *have GPD generator with parameters :*

$$\sigma^\epsilon = \frac{\sigma+\xi q_{t-1}(x_{t-1})}{1-\alpha^\epsilon\xi} \qquad \xi^\epsilon = \frac{\xi}{1-\alpha^\epsilon\xi}$$
$$\sigma^\nu = \frac{\sigma+\xi q_t(x_t)}{1-\alpha^\nu\xi} \qquad \xi^\nu = \frac{\xi}{1-\alpha^\nu\xi}$$

with $\alpha^\epsilon = \frac{nT-p}{2}, \quad \alpha^\nu = \frac{nT+p-n}{2}$

Kalman filters : bringing the GPD

Proposition 4.13. *Inverse conditional uni variate cdf's for centered conditional variables* $\left((X_{t,i}|y_{1:t}) - \mu_{t+1}^{\theta_i}\right)$ *are:*

For $\xi > 0$

$$F^{-1}_{X_{t,i}|y_{1:t}}(u) = \begin{cases} \sqrt{\frac{\tilde{\sigma}_t \Sigma_{X_t,[ii]}}{\tilde{\xi}} \frac{qbeta_{\frac{1}{2},\frac{1}{\tilde{\xi}_t}-\frac{1}{2}}(2u-1)}{1-qbeta_{\frac{1}{2},\frac{1}{\tilde{\xi}}-\frac{1}{2}}(2u-1)}} & \text{if } u \geq \frac{1}{2} \\ -F^{-1}_{X_{t,i}|y_{1:t}}(1-u) & \text{if } u < \frac{1}{2} \end{cases}$$

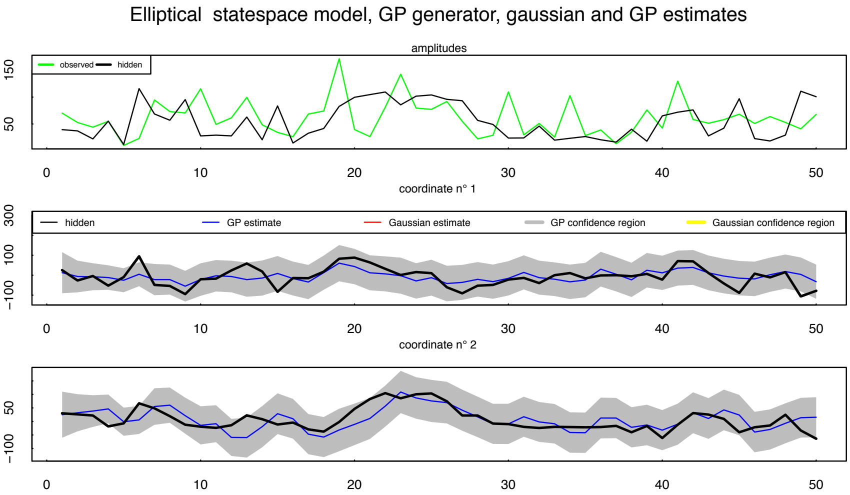
For $\xi < 0$*:*

$$F^{-1}_{X_{ti}|y_{1:t}}(u) = \begin{cases} \sqrt{\frac{\tilde{\sigma}_t \Sigma_{X_t,[ii]}}{-\tilde{\xi}_t} qbeta_{\frac{1}{2},-\frac{1}{\tilde{\xi}_t}+1}(2u-1)} & \text{if } u \geq \frac{1}{2} \\ -F^{-1}_{X_{t;i}|y_{1:t}}(1-u) & \text{if } u < \frac{1}{2} \end{cases}$$

with $\alpha = \frac{nT+p-t(n-p)-1}{2}$, $\tilde{\sigma}_t = \frac{\sigma+\xi Q_t(y_{1:t})}{1-\alpha\xi}$ $\tilde{\xi}_t = \frac{\xi}{1-\alpha\xi}$

Simulations

FIGURE 5.1. $\xi > 0$, infinite second moments

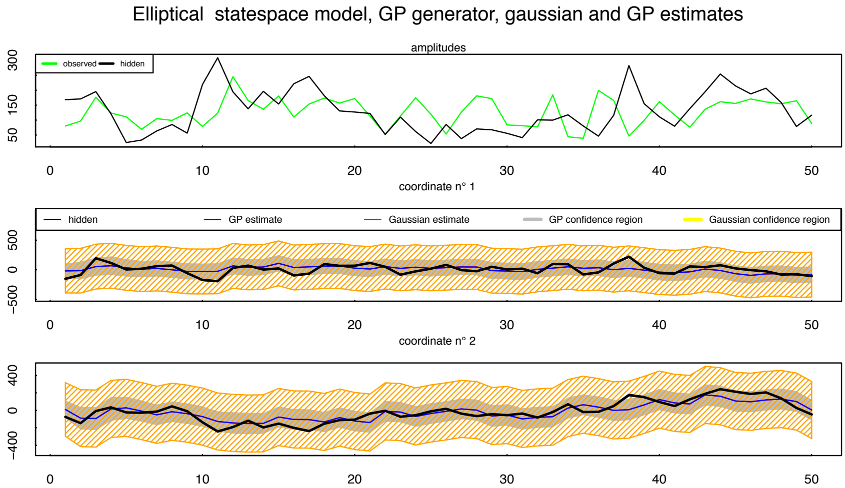


joint xi= 0.0079 ; sigma= 1 ; univariate xi= 0.7
0.95 % confidence regions; radial quantile = 0.705

max eigen value for hidden vector's noise = 13.325
for observable vector's noise = 15.136

Simulations

FIGURE 4.1. $\xi > 0$, extreme radial quantile

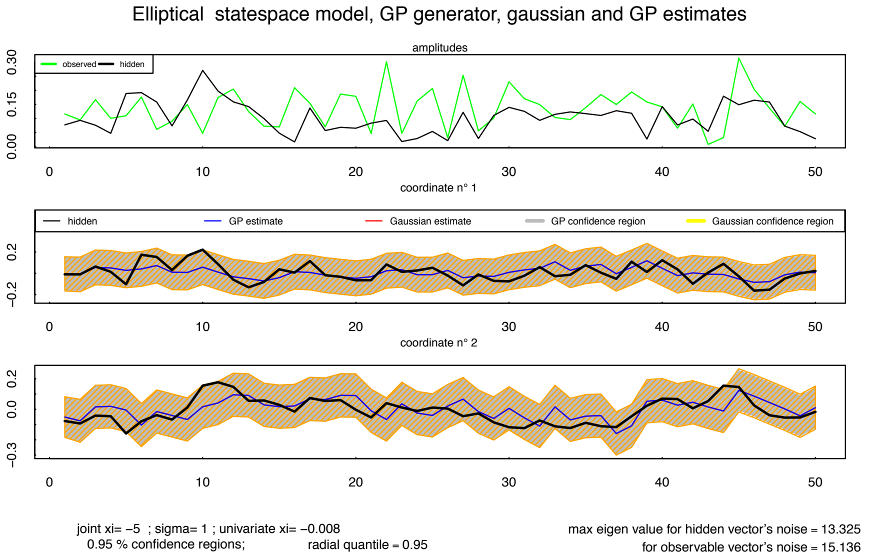


joint xi= 0.0079 ; sigma= 1 ; univariate xi= 0.66
0.95 % confidence regions; radial quantile = 0.9

max eigen value for hidden vector's noise = 13.325
for observable vector's noise = 15.136

Simultations

FIGURE 4.3. $\xi < 0$



Conclusions Part I

- Elliptical distributions with GPD generators provide explicit KF equations
- It can handle bounded, Gaussian and heavy tails
- Restricted to finite times series
- Looking for applications with symmetrical distributions

[7] E. Gómez, M.A. Gómez-Villegas, and J.M. Marín. Continuous elliptical and exponential power linear dynamic models. *Journal of Multivariate Analysis*, 83(1):22 – 36, 2002.
[8] E. Gómez, M.A. Gómez-Villegas, and J.M. Marín. A survey on continuous elliptical vector distributions. *Rev. Mat. Complut*, 16:345–361, 2003.

Non-symetrical Gumbel maxima

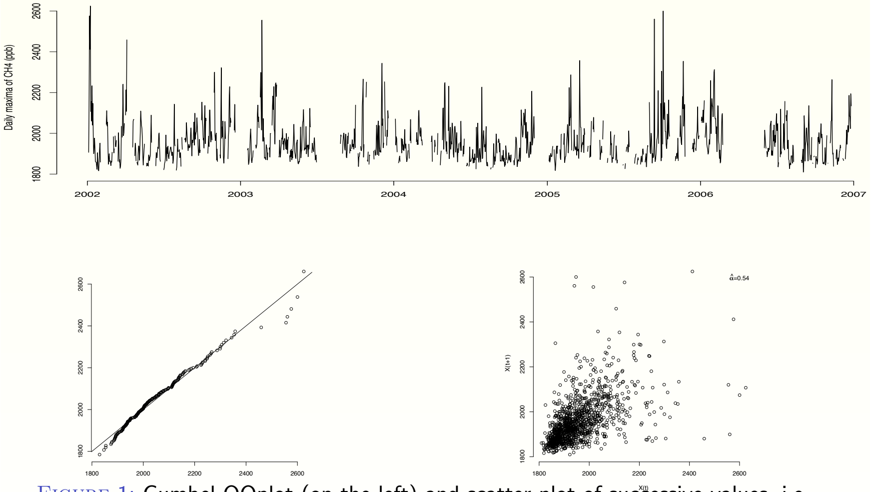


FIGURE 1: Gumbel QQplot (on the left) and scatter plot of successive values, i.e. (X_t, X_{t+1}) (on the right) corresponding to the daily maxima of CH4.

Gumbel maxima and alpha-stable

A key linear relationship

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

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

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

- A random variable S is said to be (α)-stable if and only if for all $k > 1$ there exist $c_k > 0$ and d_k such that $S_1 + \ldots + S_k \stackrel{d}{=} c_k S + d_k$ where S_1, S_2... are iid copies of S.
- Examples and special cases where one can write down explicit expressions for the density : Gaussian, Cauchy, Levy distributions.

AR(1) Gumbel maxima model

Result (1)

Let $\{X_t, t \in \mathbb{Z}\}$ *be a stochastic process defined by the recurrence relation*

$$X_t = \alpha X_{t-1} + \alpha\sigma \log S_t \qquad (1)$$

where $\sigma \in \mathbb{R}_*^+$.

Equation (1) has a unique strictly stationary solution,

$$X_t = \sigma \sum_{j=0}^{\infty} \alpha^{j+1} \log S_{t-j} \qquad (2)$$

and X_t *follows a Gumbel*$(0, \sigma)$ *distribution,* $\forall\, t \in \mathbb{Z}$.

State-space Gumbel maxima model

•Naveau and Poncet (2007) proposed the following Gumbel state-space model :

$$Y_t = F_t \log U_t + \varepsilon_t$$
$$U_t = G_t U_{t-1} + S_t$$

where ε_t corresponds to an iid Gumbel noise and where S_t represents an iid positive α-stable noise.

•This implies

$$Y_t = F_t \log \left(\sum_{i=0}^{\infty} c_{t,i} S_{t-i} \right) + \varepsilon_t$$

where $c_{t,i} = 0$ for $i \geq t$, $c_{t,i} = 1$ for $i = 0$ and $c_{t,i} = \prod_{j=0}^{i-1} G_{t-j}$ otherwise.

•Y_t are **Gumbel** distributed, the state-space model is **linear** and $\mathbb{P}(Y_1 \leq x_1, \ldots, Y_T \leq x_T)$ is **explicit** (Fougères et al., 2009).

•Problem : the estimation of U_t given the observations Y_t has not been solved yet.

Another state-space Gumbel maxima model

• The random variable $V = \mu + \sigma \log S_\alpha$ will be called Exponential-Stable with parameters α, μ, σ : $V \sim ExpS(\alpha, \mu, \sigma)$.

The model can be rewritten as

$$Y_t = \nu_t + H_t Z_t + \eta_{t,\alpha_2} \qquad \text{(observational equation)}$$
$$Z_t = \alpha_1 Z_{t-1} + \varepsilon_{t,\alpha_1} \qquad \text{(state equation)}$$

with

$$Z_0 \sim Gumbel(0, \sigma)$$
$$Z_t | Z_{t-1} \sim ExpS\left(\alpha_1, \alpha_1 x_{t-1} - \delta\sigma(1-\alpha_1), \alpha_1\sigma\right)$$
$$Y_t | Z_t \sim ExpS\left(\alpha_2, \nu_t + H_t z_t - H_t \frac{\delta\sigma}{\alpha_2}(1-\alpha_2), H_t\sigma\right)$$

The random variables $S_{t,\alpha}$ have no explicit density but it is possible numerically to compute them, to compute quantiles and to simulate values for V (see Nolan, 1997).

- Filtering density : $p(Z_t | Y_{1:t})$? ?

Filtering

$$\begin{array}{c} \hspace{7.5cm} Y_k \\ \hspace{7.5cm} \downarrow \\ p(Z_{k-1}|Y_{1:k-1}) \xrightarrow[p(Z_k|Z_{k-1})]{prediction} p(Z_k|Y_{1:k-1}) \xrightarrow[p(Y_k|Z_k)]{correction} p(Z_k|Y_{1:k}) \end{array}$$

Prediction and filtering densities

$$p(Z_k|Y_{1:k-1}) = \int p(Z_k|Z_{k-1})p(Z_{k-1}|Y_{1:k-1})dZ_{k-1} \qquad \text{(Prediction step)}$$

$$p(Z_k|Y_{1:k}) = \frac{p(Y_k|Z_k)p(Z_k|Y_{1:k-1})}{\int p(Y_k|Z_k)p(Z_k|Y_{1:k-1})dZ_k} \qquad \text{(Correction step)}$$

Comparing filtering techniques

- Classical Kalman filter (KF)
- Sampling Importance Resampling (SIR)
- Auxiliary Particle Filter (APF-ps) (Pitt Sheppard, 1999)
- Auxiliary Particle Filter with optimal weights (APF-ow)

A simulation

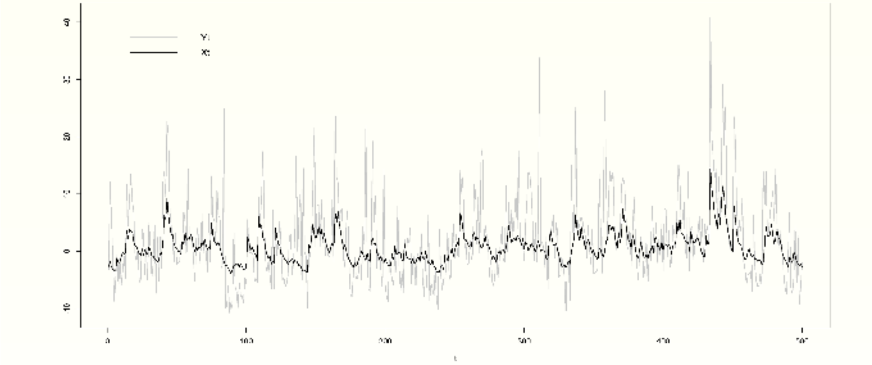


FIGURE 3: Simulated series for X_t and Y_t with $\alpha_1 = 0.8$, $\alpha_2 = 0.7$, $H_t = 2$, $\mu_t = 0$, $\sigma = 2$.

SIR vesus APF-ps

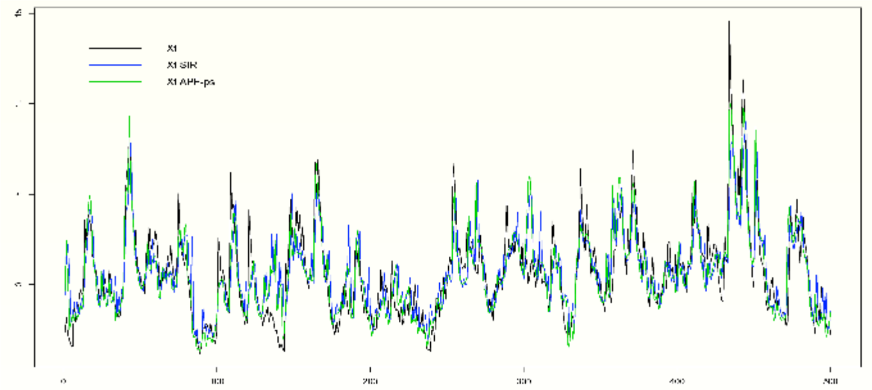
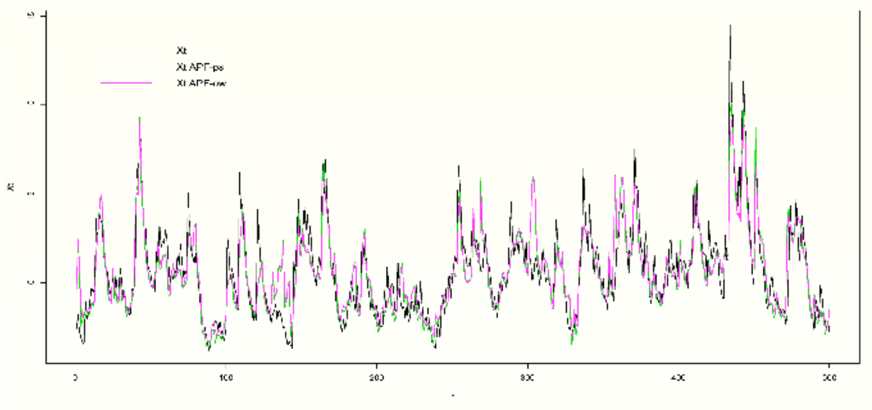
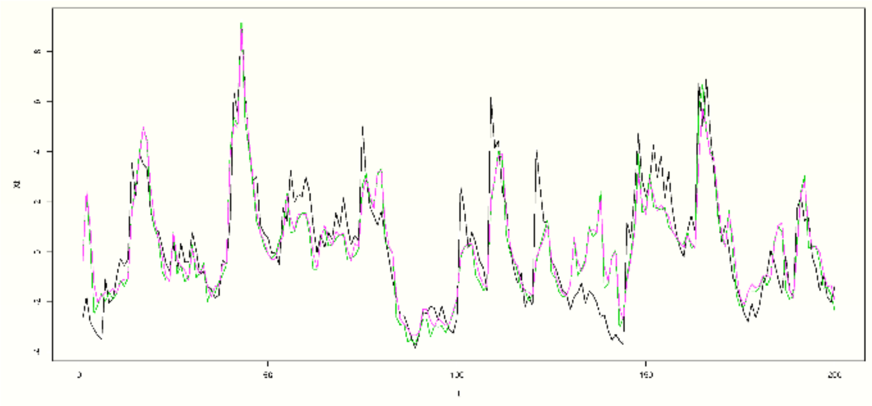


FIGURE 5: $\alpha_1 = 0.8$, $\alpha_2 = 0.7$, $H_t = 2$, $\mu_t = 0$, $\sigma = 2$ and $N = 60000$.

APF-ps vesus APF-ow



APF-ps vesus APF-ow (Zoom)



MSE

$$MSE = \frac{1}{500}\sum_{k=1}^{500}\left(\widehat{Z}_k - Z_k\right)^2$$

MSE KF	MSE SIR	MSE APF-ps	MSE APF-ow
2.52	2.34	1.71	1.69

But only one realization...

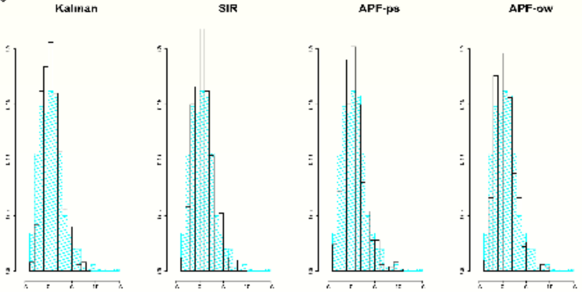


Figure 8: Histograms of filtered state (black) and simulated series (blue)

Conclusions Part II

- We propose a state-space model for Gumbel distributed maxima
- Inference about the hidden state is possible by using particle filtering techniques
- Optimizing the weights does not seem to bring a strong improvement
- Looking for nice applications

Particle filtering

We suppose that the set of weighted particles $\{\xi_{k-1}^i, w_{k-1}^i\}_{i=1}^N$ approximates the filtering density $p(Z_{k-1}|Y_{1:k-1})$.
The empirical distribution corresponding to this approximation is

$$p^N(Z_{k-1}|Y_{1:k-1}) = \sum_{i=1}^{N} w_{k-1}^i \delta_{\xi_{k-1}^i}(Z_{k-1}).$$

An approximation of the prediction density and the filtering density follows :

$$p^N(Z_k|Y_{1:k-1}) = \sum_{i=1}^{N} w_{k-1}^i p(Z_k|\xi_{k-1}^i)$$

$$p^N(Z_k|Y_{1:k}) \, \alpha \, \sum_{i=1}^{N} w_{k-1}^i p(Y_k|Z_k) p(Z_k|\xi_{k-1}^i)$$

The objective is to obtain a set of weighted particles $\{\xi_k^i, w_k^i\}_{i=1}^N$ that approximate $p(Z_k|Y_{1:k})$.
ξ_k^i will be sampled according to the importance density $q(.|\xi_{k-1}^i, Y_k)$ and the associated importance weight will be computed with the relation

$$w_k^i \, \alpha \, w_{k-1}^i \frac{p(Y_k|\xi_k^i) p(\xi_k^i|\xi_{k-1}^i)}{q(\xi_k^i|\xi_{k-1}^i, Y_k)}$$

Particle filtering

We suppose that the set of weighted particles $\{\xi_{k-1}^i, w_{k-1}^i\}_{i=1}^N$ approximates the filtering density $p(Z_{k-1}|Y_{1:k-1})$.
The empirical distribution corresponding to this approximation is

$$p^N(Z_{k-1}|Y_{1:k-1}) = \sum_{i=1}^{N} w_{k-1}^i \delta_{\xi_{k-1}^i}(Z_{k-1}).$$

An approximation of the prediction density and the filtering density follows :

$$p^N(Z_k|Y_{1:k-1}) = \sum_{i=1}^{N} w_{k-1}^i p(Z_k|\xi_{k-1}^i)$$

$$p^N(Z_k|Y_{1:k}) \, \alpha \sum_{i=1}^{N} w_{k-1}^i p(Y_k|Z_k) p(Z_k|\xi_{k-1}^i)$$

The objective is to obtain a set of weighted particles $\{\xi_k^i, w_k^i\}_{i=1}^N$ that approximate $p(Z_k|Y_{1:k})$.
ξ_k^i will be sampled according to the importance density $q(.|\xi_{k-1}^i, Y_k)$ and the associated importance weight will be computed with the relation

$$w_k^i \, \alpha \, w_{k-1}^i p(Y_k|\xi_k^i)$$

Particle filtering

Bootstrap filter

At time $t = t_0$

$$\xi_{t_0}^{1:N} \overset{iid}{\sim} p(X_{t_0})$$

At time $t_0 < k \leq T$,

1) Propagation
$\xi_k^i \sim p(X_k | \xi_{k-1}^i)$ for $i = 1, ..., N$

2) Computation of the weights for $i = 1, ..., N$
$w_k^i \leftarrow p(Y_k | \xi_k^i)$
$w_k^i \leftarrow \frac{w_k^i}{\sum_{i=1}^N w_k^i}$

3) Selection step
$\xi_k^{1:N} \leftarrow \text{resample}(w_k^{1:N}, \xi_k^{1:N})$

Particle filtering

Auxiliary particle filter (APF)

At time $t = t_0$

$$\xi_{t_0}^{1:N} \overset{iid}{\sim} p(X_{t_0})$$
$$w_{t_0}^{1:N} \leftarrow \frac{1}{N}$$

At time $t_0 < k \leq T$,

1) Selection step

$$\beta_k^i \leftarrow w_{k-1}^i \widehat{p}(Y_k | \xi_{k-1}^i)$$
$$j^{1:N} \leftarrow \text{resample}(\beta_k^{1:N}, 1:N)$$

2) Propagation

$$\xi_k^i \sim p(X_k | \xi_{k-1}^{j^i}) \quad \text{for } i = 1, ..., N$$

3) Computation of the weights for $i = 1, ..., N$

$$w_k^i \leftarrow \frac{p(Y_k | \xi_k^i)}{\widehat{p}(Y_k | \xi_{k-1}^{j^i})}$$
$$w_k^i \leftarrow \frac{w_k^i}{\sum_{i=1}^N w_k^i}$$

Particle filtering

Pitt and Shepard (1999) propose to consider $\widehat{p}(Y_k|\xi_{k-1}^i) = p(Y_k|Z_k = \mu_k^i)$ with $\mu_k^i = \mathbb{E}(Z_k|\xi_{k-1}^i)$. This filter will be noted APF-ps.

Let recall the Gumbel state-space model we consider

$$Y_t = \nu_t + H_t Z_t + \eta_{t,\alpha_2} \qquad \text{(observational equation)}$$
$$Z_t = \alpha_1 Z_{t-1} + \varepsilon_{t,\alpha_1} \qquad \text{(state equation)}$$

with

$$Z_0 \sim Gumbel(0, \sigma)$$
$$Z_t|Z_{t-1} \sim ExpS\left(\alpha_1, \alpha_1 x_{t-1} - \delta\sigma(1-\alpha_1), \alpha_1\sigma\right)$$
$$Y_t|Z_t \sim ExpS\left(\alpha_2, \nu_t + H_t z_t - H_t \frac{\delta\sigma}{\alpha_2}(1-\alpha_2), H_t\sigma\right)$$

In this case, it is possible to compute the optimal weights and to propose an adapted APF to our model (APF-ow).

A very short biblio

- J. ANDERSON, An ensemble adjustment Kalman filter for data assimilation, *Mont. Weath. R.*, **129**, 2001.
- AZZALINI, A., DALLA VALLE A., The multivariate skew-normal distribution, *Biometrika*, **83**, 715-726 (1996).
- T. BENGTSSON, D. NYCHKA, C. SNYDER A frame work for data assimilation and forecasting in high-dimensional non-linear dynamical systems. (2002).
- DOUCET A., N. FREITAS, N. GORDON, (2001). Sequential Monte Carlo Methods in Practice.
- WIKLE, C.K. AND N. CRESSIE, (1999) A dimension reduced approach to space-time Kalman filtering. *Biometrika* , 86 , 815-829.