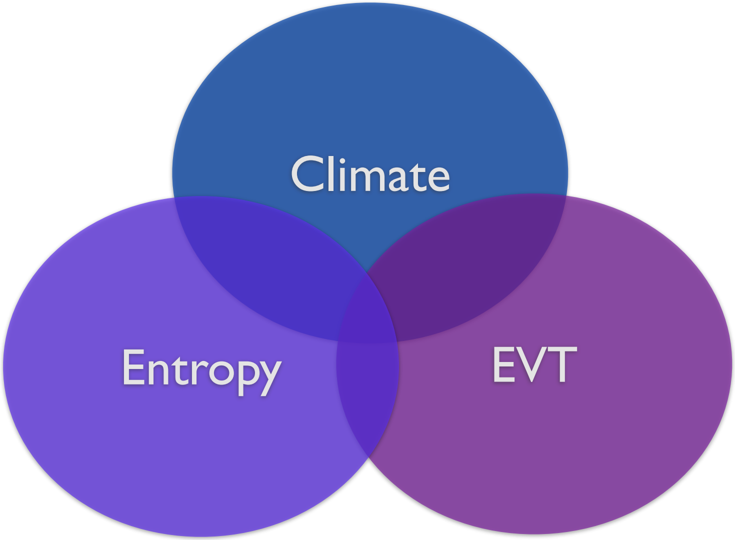
Special topics for extremes

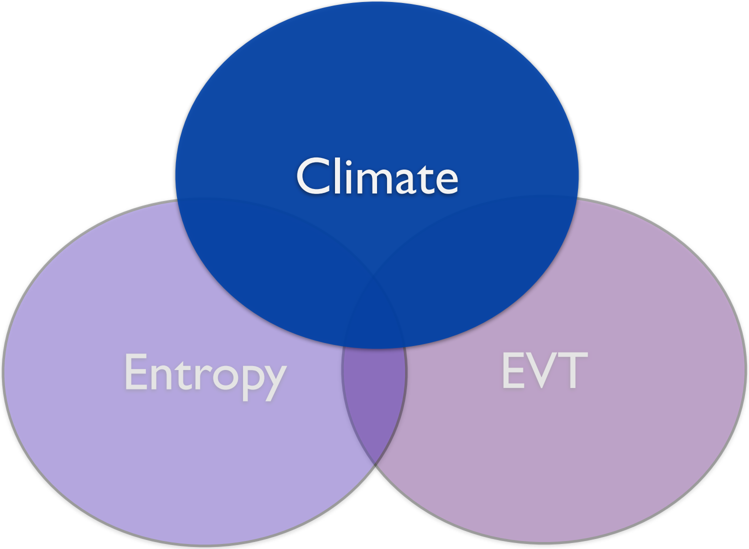
Philippe Naveau

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

ANR-McSim, ExtremeScope, LEFE-MULTI-RISK

A simple and fast tool to deal with non-stationarity in GPD

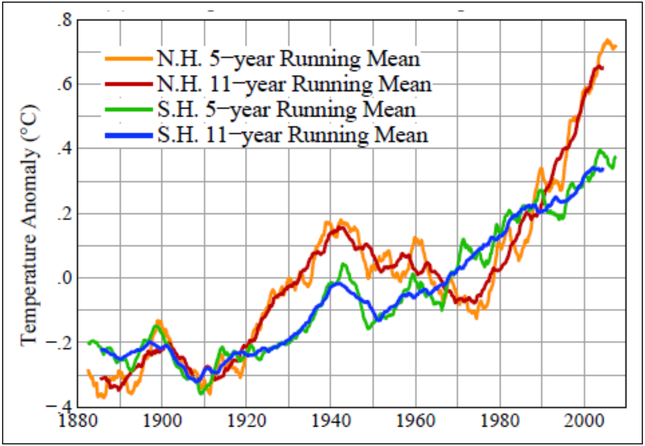



Climate
Entropy
EVT

Objectives

- Detecting changes in time, e.g. are the last 30 year extreme temperatures different from earlier periods ?
- Detecting changes in space, e.g. are extreme temperatures in Paris different from the ones recorded in Trieste ?
- We don't deal with the attribution problem here (see Francis Z.)

Hemispheric mean temperatures (source GISS-NASA 2010)



Temperatures anomalies (1961-1990)

Source : Abarca-Del-Rio and Mestre, GRL. (2006)

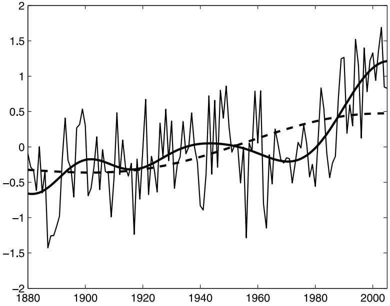
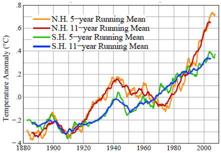


Figure 5. Time series of annual mean temperature anomalies over France (black), (A6 + D5 + D6) reconstruction (solid black), and A6 (dashed).

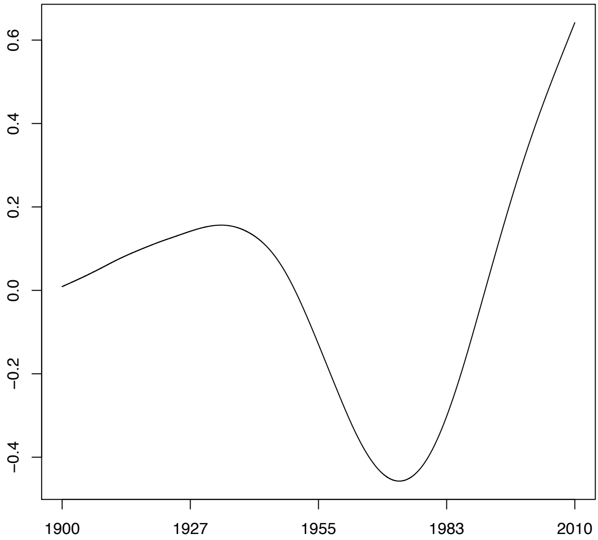


An illustration

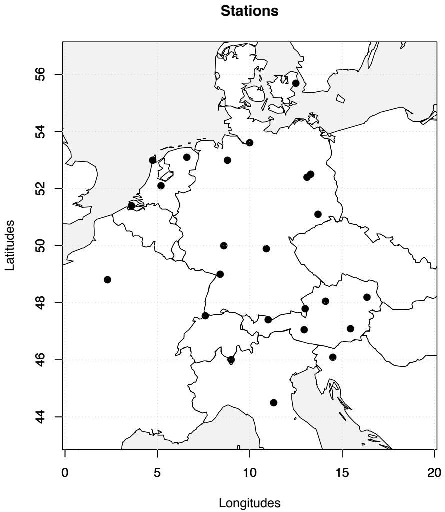
Paris, station Montsouris

- Daily maxima of temperatures
- Years from 1900 to 2010
- 40 515 maxima
- Urban island effect (?)
- Temperature maxima are often of Weibull type (precip light Fréchet)
- Clustering versus declustering
- 30 years = the climatology yard stick

Paris, smooth trend



Beyond Paris

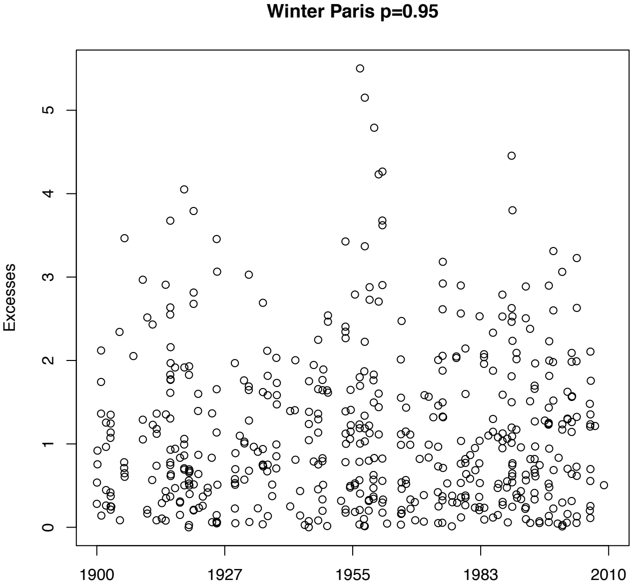


European stations with long time series

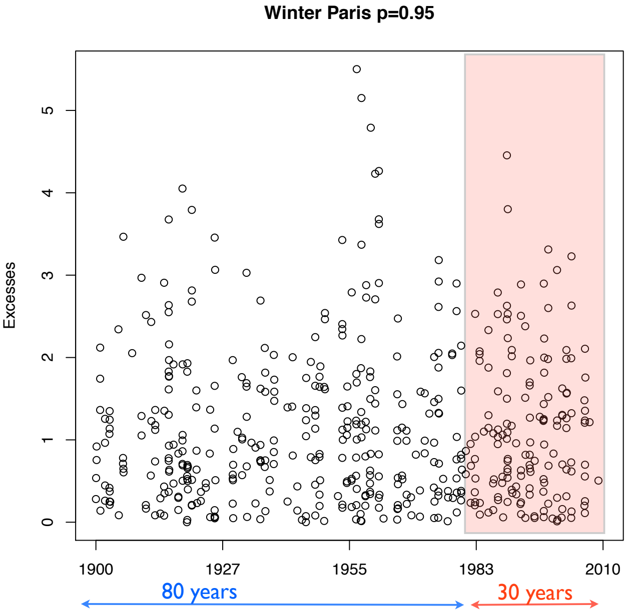
Table 1. Characteristics of 24 weather stations from the European Climate Assessment & Dataset project `http://eca.knmi.nl/dailydata/predefinedseries.php`. The heights are expressed in meters.

Austria						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Kremsmunster	+48:03:00	+014:07:59	383	1876	2011	-
Graz	+47:04:59	+015:27:00	366	1894	2011	2
Salzburg	+47:48:00	+013:00:00	437	1874	2011	5
Sonnblick	+47:03:00	+012:57:00	3106	1887	2011	-
Wien	+48:13:59	+016:21:00	199	1856	2011	2
Denmark						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Koebenhavn	+55:40:59	+012:31:59	9	1874	2011	-
France						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Montsouris	+48:49:00	+002:19:59	77	1900	2010	-
Germany						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Bamberg	+49:52:31	+010:55:18	240	1879	2011	-
Berlin	+52:27:50	+013:18:06	51	1876	2011	1
Bremen	+53:02:47	+008:47:57	4	1890	2011	1
Dresden	+51:07:00	+013:40:59	246	1917	2011	-
Frankfurt	+50:02:47	+008:35:54	112	1870	2011	1
Hamburg	+53:38:06	+009:59:24	11	1891	2011	-
Karlsruhe	+49:02:21	+008:21:54	112	1876	2011	2
Potsdam	+52:22:59	+013:04:00	100	1893	2011	-
Zugspitze	+47:25:00	+010:58:59	2960	1901	2011	1
Italy						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Bologna	+44:30:00	+011:20:45	53	1814	2010	-
Netherlands						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
De Bilt	+52:05:56	+005:10:46	2	1906	2011	-
Den Helder	+52:58:00	+004:45:00	4	1901	2011	-
Eelde	+53:07:24	+006:35:04	5	1907	2011	-
Vlissingen	+51:26:29	+003:35:44	8	1906	2011	2
Slovenia						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Ljubjana	+46:03:56	+014:31:01	299	1900	2011	5
Switzerland						
Station name	Latitude	Longitude	Height	First year	Last year	Missing years
Basel	+47:33:00	+007:34:59	316	1901	2011	-
Lugano	+46:00:00	+008:58:00	300	1901	2011	-

Paris, analyzing excesses per season



A main question about the last "30 year" extremes



A current approach in the climate community about detecting changes in extremes

- Fitting a GEV or a GPD based model to describe extremes
- Investigating how the GEV or GEV parameters change in space or time.

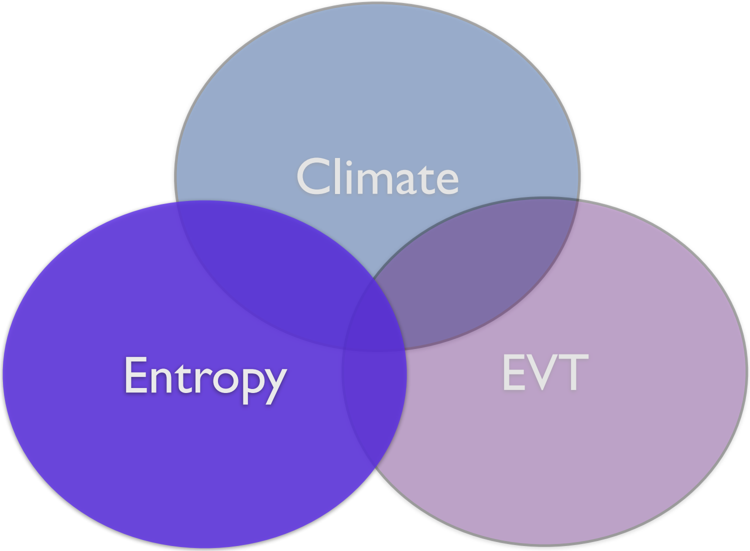
E.g., Jarusková and Rencová (2008), Fowler and Kilsby, (2003), Kharin et al., (2007)

Desiderata of our statistical approach

- Very few assumptions (neither imposing a GPD nor a GEV)
- Fast computations
- Good statistical properties of our estimators
- Interpretability
- Cross discipline tools (statistics and climatology)

Our assumptions

- Trends and annual cycles have been removed
- High detrended excesses can be considered stationary (or even iid)
- The two periods of interest belongs to the same domain of attraction
- The upper end points of both periods are equal
- We do **not** assume that the two periods have necessarily the same shape parameter


Climate
Entropy
EVT

Definitions (Kullback, 1968)

The Kullback-Leibler directed divergence

$$I(f;g) = \mathbb{E}_f \left(\log \left(\frac{f(\mathbf{X})}{g(\mathbf{X}}) \right) \right),$$

where f and g pdfs and $\mathbf{X}$ random vector with density f.

The Kullback-Leibler divergence

$$D(f;g) = I(f;g) + I(g;f)$$

a symmetrical measure relative to f and g.

Kullback-Liebler divergence

Advantages

- Interpretability
- A single and simple summary
- Cross discipline tools (statistics and climatology)
- It is not an index
- Links with model selection criteria

Kullback-Liebler divergence

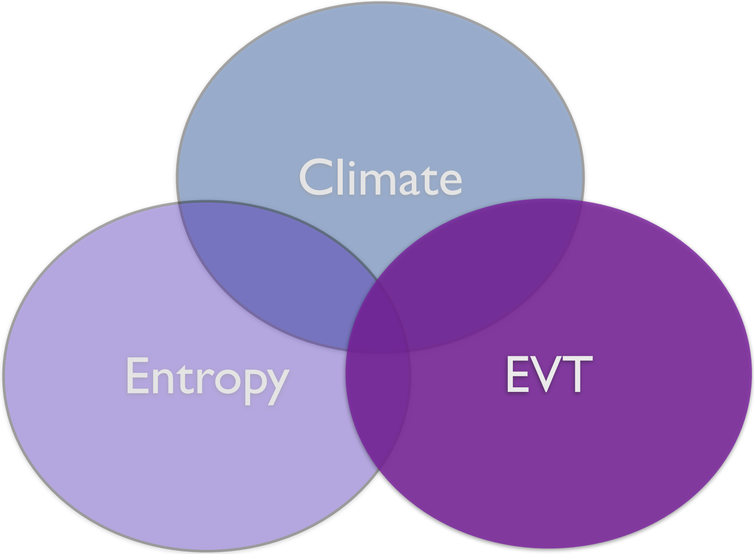
$$I(f;g) = \mathbb{E}_f\left(\log\left(\frac{f(\mathbf{X})}{g(\mathbf{X})}\right)\right), \text{ and } D(f;g) = I(f;g) + I(g;f)$$

How to compute and infer the divergence ?

- Assume f and g have explicit expressions, e.g. GP
- Plug in the mle estimates

Drawbacks

- Strong assumptions : f and g have explicit expressions, e.g. GP
- Difficult to work with the likelihood when the dimension is large (composite likelihoods, etc)


Climate
Entropy
EVT

The “Old” life

The “Modern” style

Distributions (cdf)

Densities (pdf)

$I(f;g) = \mathbb{E}_f \left(\log \left(\frac{f(\mathbf{X})}{g(\mathbf{X})} \right) \right)$

A general comment

The divergence is based on the **probability density functions**, **f** and **g**, but extreme behaviors are better captured by **cumulative distribution functions**, **F** and **G**, or their tails, $1 - \mathbf{F}$ and $1 - \mathbf{G}$.

..., but what is the question for this application ? It is not

- to infer GPD parameters

Entropy for excesses above u

Notations

Let X and Y be two abs. cont. r.v.'s with identical upper end-points.
For any u, we define the random vector $[X|\ X > u]$ by its density

$$f_u(x) = \frac{f(x)}{\overline{F}(u)}, \text{ for all } x_F > x > u.$$

Divergence definition

$$\begin{aligned} I(f_u; g_u) &= \mathbb{E}_{f_u}\left(\log\left(\frac{f_u(X_u)}{g_u(X_u)}\right)\right) \\ &= \frac{1}{\overline{F}(u)} \int_u^{x_F} \log\left(\frac{f_u(x)}{g_u(x)}\right) f(x)dx \end{aligned}$$

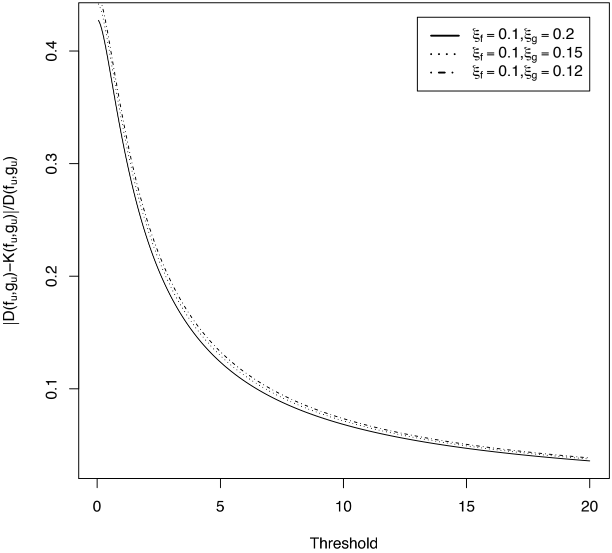
Entropy for excesses above u

An approximation of $I(f_u; g_u) = \mathbb{E}_{f_u}\left(\log\left(\frac{f_u(X_u)}{g_u(X_u)}\right)\right)$

Under mild assumptions, the divergence $D(f_u; g_u) = I(f_u; g_u) + I(g_u; f_u)$ is equivalent to the quantity $K(f_u; g_u) = -L(f_u; g_u) - L(g_u; f_u)$ where

$$L(f_u; g_u) = \mathbb{E}_f\left(\log \frac{\overline{G}(X)}{\overline{G}(u)} \middle| X > u\right) + 1$$

Approximation for the GPD



Necessary condition for applying our divergence approximation

Under mild conditions, the divergence $D(f_u; g_u) = I(f_u; g_u) + I(g_u; f_u)$ is equivalent to the quantity $K(f_u; g_u) = -L(f_u; g_u) - L(g_u; f_u)$ where

$$L(f_u; g_u) = \mathbb{E}_f \left(\log \frac{\overline{G}(X)}{\overline{G}(u)} \middle| X > u \right) + 1$$

Inference

One advantage of $\mathbb{E}_f\left(\log \frac{\overline{G}(X)}{\overline{G}(u)}\middle| X > u\right)$ **over** $\mathbb{E}_{f_u}\left(\log\left(\frac{f_u(X_u)}{g_u(X_u)}\right)\right)$

The Kullback Leibler divergence approximation can be estimated by

$$\widehat{L}(f_u; g_u) = \frac{1}{N_n}\sum_{i=1}^{n}\log\frac{\overline{G_m}(X_i \vee u)}{\overline{G_m}(u)}$$

with G_m is the empirical cdf of the Y_i's and $N_n := \#\{X_i, X_i \geq u\}$

Distribution under the null hypothesis : $\overline{F} = \overline{G}$

Suppose $n = m$

$$\overline{G_n}(x) = \overline{F_n}(x) \text{ in distribution}$$

Distribution under the null hypothesis : $\overline{F} = \overline{G}$

Suppose $n = m$

$$\overline{G_n}(x) = \overline{F_n}(x) \text{ in distribution}$$

The statistic

$$\widehat{L}(f_u; f_u) = \frac{1}{N_n} \sum_{i=1}^{n} \log \frac{\overline{F_n}(X_i \vee u)}{\overline{F_n}(u)}$$

does not depend on the original F and it is a distribution free statistic.

Simulations results

Coming tables

Number of false positive (wrongly rejecting that $f = g$) and negative out of **1000** replicas of two samples of sizes $n = m$ for a 95% level where f and g GP densities with shape parameter ξ_f and ξ_g, respectively.

Bounded tails

Weibull-Weibull case	$\xi_f = -.1$				
n \ ξ_g	$-.2$	$-.15$	$-.1$	$-.08$	$-.05$
50	**107**	**550**	**54**	**816**	**93**
100	**3**	**233**	**58**	**680**	**5**
200	**0**	**35**	**45**	**469**	**0**
500	**0**	**0**	**62**	**94**	**0**
1000	**0**	**0**	**50**	**5**	**0**

Bounded tails : comparing with the Kolmogorov-Smirnov test

Weibull-Weibull case	$\xi_f = -.1$									
n \ ξ_g	$-.2$		$-.15$		$-.1$		$-.08$		$-.05$	
50	262	**107**	691	**550**	26	**54**	889	**816**	241	**93**
100	42	**3**	445	**233**	50	**58**	824	**680**	24	**5**
200	0	**0**	110	**35**	30	**45**	627	**469**	0	**0**
500	0	**0**	0	**0**	52	**62**	189	**94**	0	**0**
1000	0	**0**	0	**0**	42	**50**	17	**5**	0	**0**

Heavy-tails : comparing with the Kolmogorov-Smirnov test

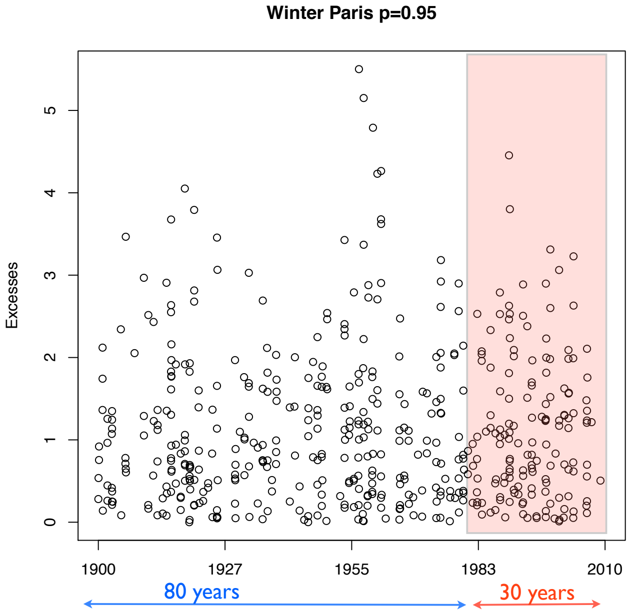
Fréchet-Fréchet case	$\xi_f = .1$									
n \ ξ_g	.05		.08		.1		.15		.2	
10	986	**946**	991	**939**	15	**49**	986	**954**	963	**950**
100	973	**935**	960	**941**	43	**49**	953	**930**	925	**899**
1000	937	**740**	946	**926**	45	**49**	943	**773**	740	**285**
10000	648	**9**	911	**597**	41	**55**	688	**16**	14	**0**

Gumbel case : comparing with the Kolmogorov-Smirnov test

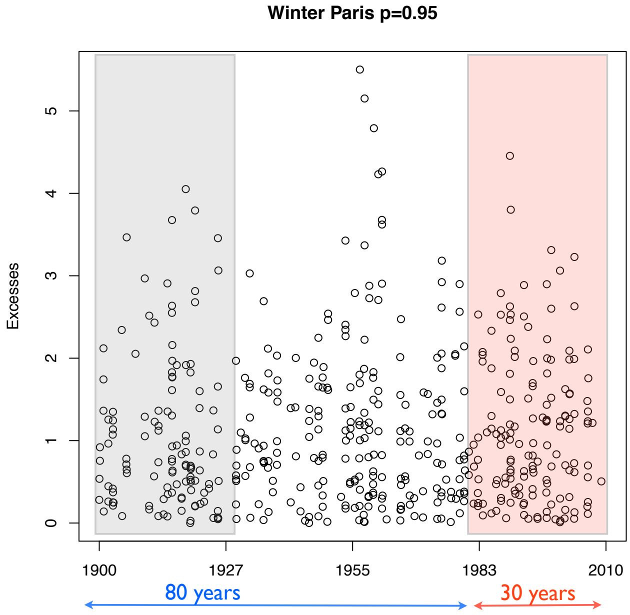
Gumbel-Weibull case	$\xi_f = 0$									
n \ ξ_g	$-.5$		$-.4$		$-.3$		$-.2$		0	
50	804	**155**	868	**377**	921	**628**	949	**799**	28	**48**
100	513	**3**	727	**37**	896	**231**	933	**599**	38	**50**
200	63	**0**	348	**0**	709	**12**	888	**252**	37	**60**
500	0	**0**	1	**0**	134	**0**	616	**4**	58	**53**
1000	0	**0**	0	**0**	0	**0**	229	**0**	36	**54**

Gumbel-Fréchet case	$\xi_f = 0$									
n \ ξ_g	0		$.2$		$.3$		$.4$		$.5$	
50	33	**63**	943	**853**	945	**746**	919	**589**	885	**473**
100	36	**74**	923	**695**	907	**470**	844	**253**	733	**100**
200	30	**57**	902	**412**	807	**120**	606	**18**	392	**1**
500	43	**55**	706	**52**	335	**0**	101	**0**	13	**0**
1000	45	**59**	423	**1**	34	**0**	0	**0**	0	**0**

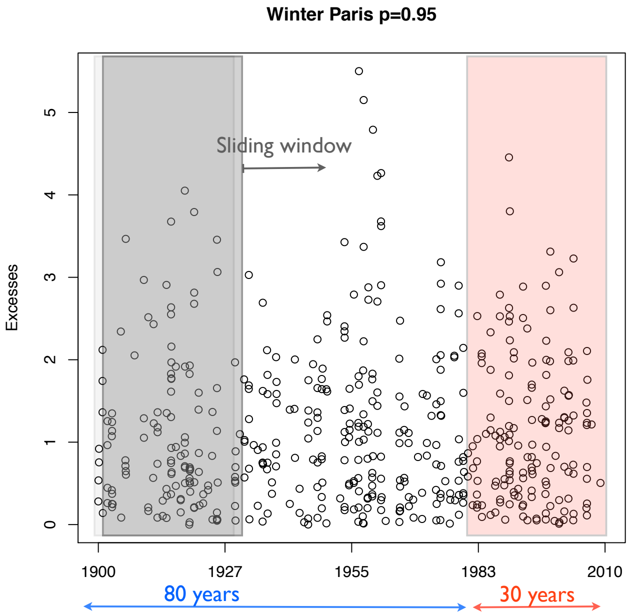
Coming back to our temperatures maxima recorded in Paris



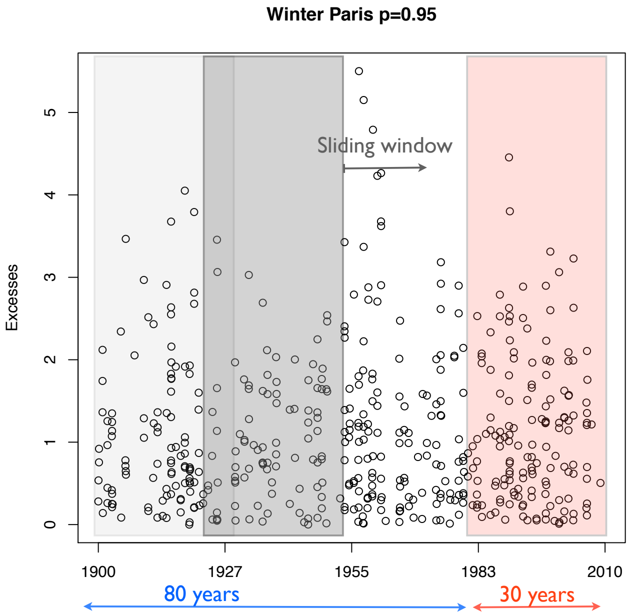
Step A (sliding window of length 30 years)



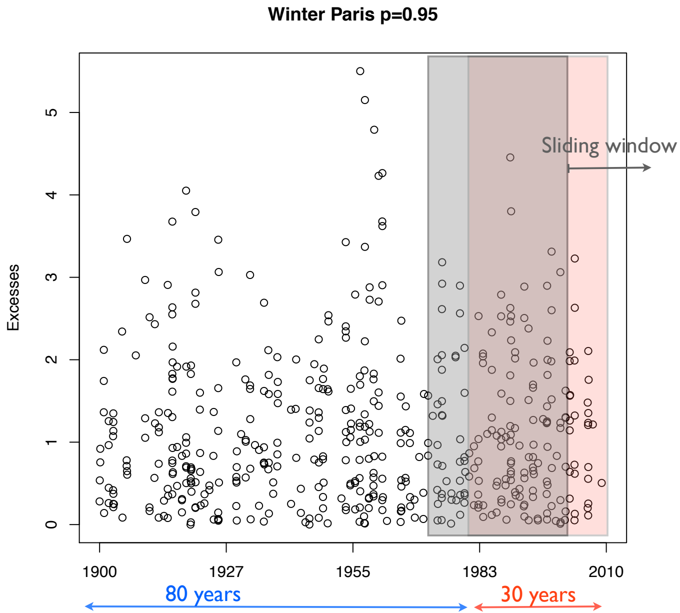
Step A (sliding window of length 30 years)



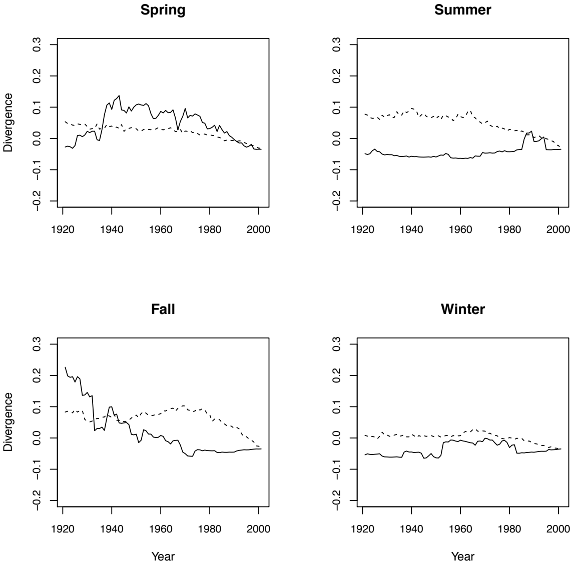
Step A (sliding window of length 30 years)



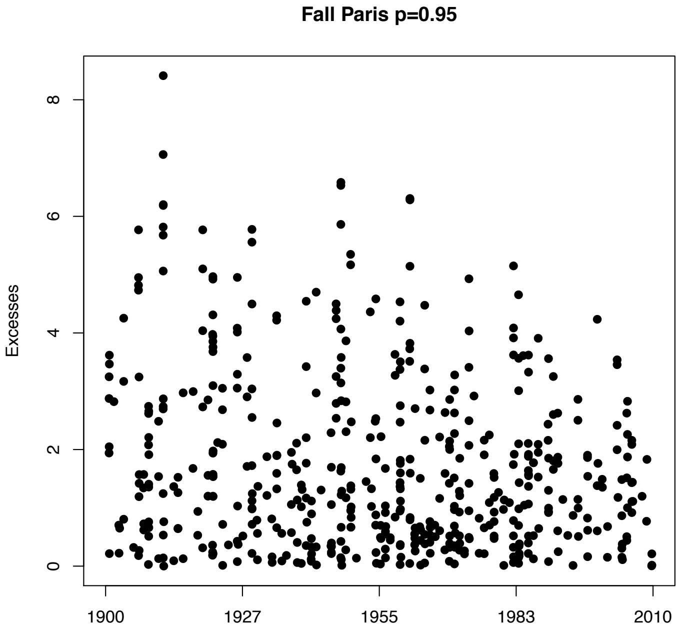
Step A (sliding window of length 30 years)

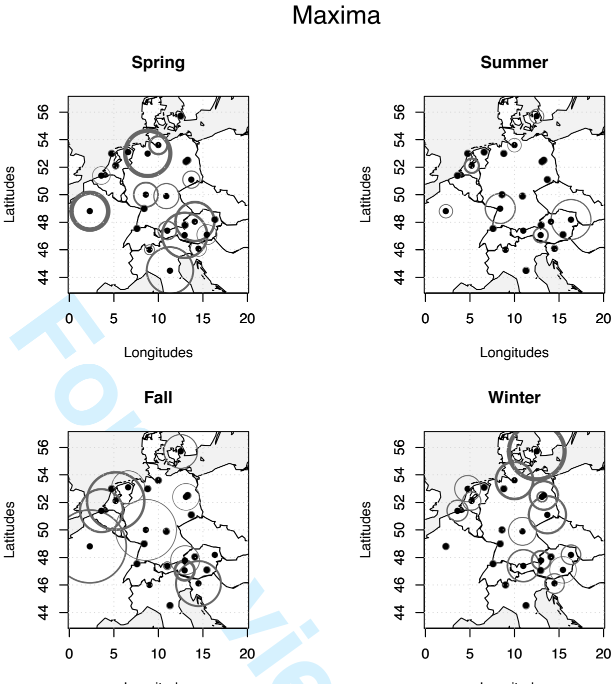


Back to Paris

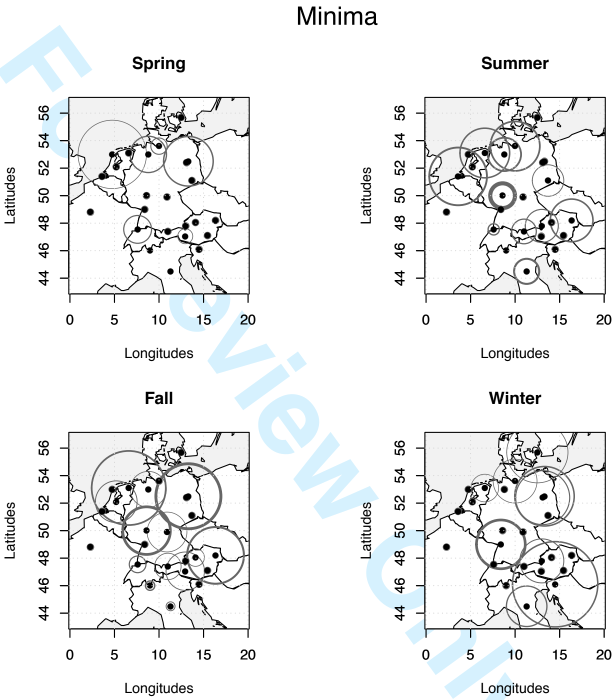


Paris Montsouris





Maxima
Spring
Summer
Fall
Winter
Latitudes
Longitudes
44
46
48
50
52
54
56
0
5
10
15
20

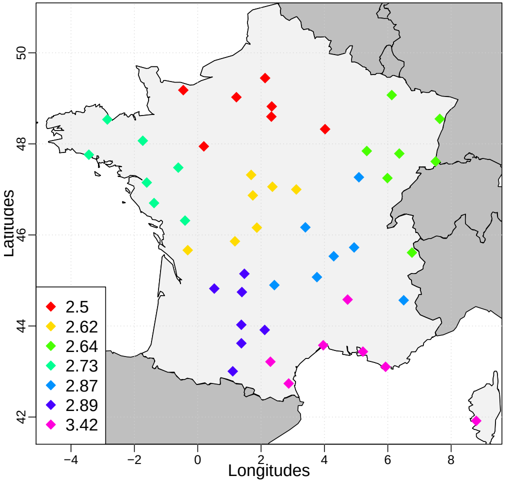

Minima
Spring
Summer
Fall
Winter
Latitudes
Longitudes
44 46 48 50 52 54 56
0 5 10 15 20

Take home messages for detecting with the entropy

- Extremes here mean very rare
- Connections between EVT and Kullback-Leibler divergence
- No need to choose a explicit density
- Fast algorithm
- Limited to one question
- Significant changes over Europe, especially for minima
- More research needed for the multivariate case
- More applications to large datasets

WEEKLY MAXIMA OF HOURLY RAINFALL IN FRANCE (FALL SEASON, 1993-2011)

How to measure the dependence within each cluster[1] of size $d = 7$



1. Bernard, Naveau, Vrac, and Mestre, JOC (2013)

Multivariate Extremes Framework [2]

Margins (unit-Fréchet)

$$F(x) = P(X_1 \leq x) = \cdots = P(X_d \leq x) = \exp(-1/x)$$

Max-stability property

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

2. de Haan and Ferreira, 2006, Resnick 1987, Embrechts, Kluppelberg and Mikosch (1997)

Max-stable multivariate vector

Max-stability in the univariate case

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

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

$$F^t(tx_1, \dots, tx_d) = F(x_1, \dots, x_d)$$

Multivariate distribution

$$\mathbb{P}\{X_1 \leq x_1, \dots, X_d \leq x_d\} = \exp\{-V(\mathbf{x})\}$$

with

$$V(t\mathbf{x}) = t^{-1} V(\mathbf{x})$$

Modeling and inference options

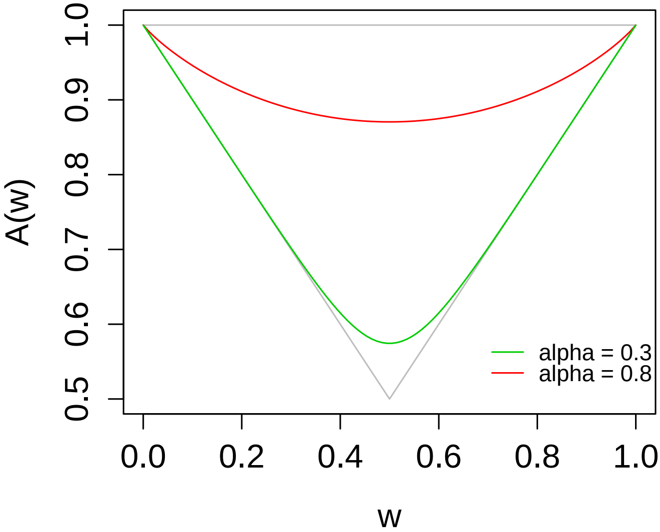
- Working with $V(\mathbf{x})$
- or the **spectral measure** $H(d\mathbf{w})$

$$V(\mathbf{x}) = \int_{S_d} \bigvee_{i=1}^{d} \left(\frac{w_i}{x_i} \right) H(d\mathbf{w})$$

- or the **Pickands dependence function** $A(\mathbf{w})$

$$V(\mathbf{x}) = \left(\frac{1}{x_1} + \ldots + \frac{1}{x_d} \right) A(\mathbf{w})$$

Pickands Dependence Function $A(\mathbf{w})$ for $d = 2$



$A(.)$ Pickands Dependence Function

Polar coordinates on the simplex

$$r = \sum_{i=1}^{d} x_i \qquad w_i = \frac{x_i}{r}$$

$$\mathcal{S}_{d-1} := \left\{ (w_1, \ldots, w_{d-1}) \in [0,1]^{d-1} : \sum_{i=1}^{d-1} w_i \leq 1 \right\}$$

Constraints on the Pickands function

P1) $A(\mathbf{w})$ convex

P2) $A(\mathbf{w}) \geq \max(w_1, \ldots, w_{d-1}, w_d)$ and $1/d \leq A(\mathbf{w}) \leq 1$

P3) $A(\boldsymbol{e_i}) = 1$ and $A(\mathbf{0}) = 1$

How to extend beyond the 2d case ? Recall about the bivariate case

L1-distance (Madogram estimator)

$$\nu = \mathbb{E}|U - V|$$

Special case : $U = \max(F_1(X_1), F_2(X_2))$ **and** $V = (F_1(X_1) + F_2(X_2))/2$

$$V(\mathbf{1}) = \frac{1 + 2\nu}{1 - 2\nu}$$

Inference

$$\hat{V}(\mathbf{1}) = \frac{1 + 2\hat{\nu}}{1 - 2\hat{\nu}}$$

L1-distance (Madogram estimator)

$$\nu = \mathbb{E}|U - V|$$

$$\nu(\mathbf{w}) = \mathbb{E}\left(\bigvee_{i=1,\dots,d} \left\{ F_i^{1/w_i}(X_i) \right\} - \frac{1}{d} \sum_{i=1,\dots,d} F_i^{1/w_i}(X_i) . \right), \quad \mathbf{w} \in \mathcal{S}_{d-1}$$

L1-distance (Madogram estimator)

$$\nu = \mathbb{E}|U - V|$$

$$\nu(\mathbf{w}) = \mathbb{E}\left(\bigvee_{i=1,\dots,d} \left\{ F_i^{1/w_i}(X_i) \right\} - \frac{1}{d} \sum_{i=1,\dots,d} F_i^{1/w_i}(X_i) \,. \right), \quad \mathbf{w} \in \mathcal{S}_{d-1}$$

Proposition

$$\nu(\mathbf{w}) = \frac{V(1/w_1, \dots, 1/w_d)}{1 + V(1/w_1, \dots, 1/w_d)} - c(\mathbf{w}),$$

where $c(\mathbf{w}) = d^{-1} \sum_{i=1}^{d} w_i/(1 + w_i)$.

$$A(\mathbf{w}) = \frac{\nu(\mathbf{w}) + c(\mathbf{w})}{1 - \nu(\mathbf{w}) - c(\mathbf{w})}.$$

Multivariate Madogram

$$\nu(\mathbf{w}) = \mathbb{E}\left(\bigvee_{i=1,\dots,d} \left\{ F_i^{1/w_i}(X_i) \right\} - \frac{1}{d} \sum_{i=1,\dots,d} F_i^{1/w_i}(X_i) \, . \right), \quad \mathbf{w} \in \mathcal{S}_{d-1}$$

MULTIVARIATE MADOGRAM

$$\nu(\mathbf{w}) = \mathbb{E}\left(\bigvee_{i=1,\dots,d} \left\{ F_i^{1/w_i}(X_i) \right\} - \frac{1}{d} \sum_{i=1,\dots,d} F_i^{1/w_i}(X_i) \, . \right), \quad \mathbf{w} \in \mathcal{S}_{d-1}$$

A natural estimator of the multivariate madogram is given by

$$\widehat{\nu}_n(\mathbf{w}) = \frac{1}{n} \sum_{m=1}^{n} \left(\bigvee_{i=1,\dots,d} \left\{ \hat{F}_i^{1/w_i}(X_{m,i}) \right\} - \frac{1}{d} \sum_{i=1,\dots,d} \hat{F}_i^{1/w_i}(X_{m,i}) \right),$$

where $\hat{F}_i$ is the empirical cdf for the ith coordinate.

$$\nu(\mathbf{w}) = \mathbb{E}\left(\bigvee_{i=1,\ldots,d} \left\{ F_i^{1/w_i}(X_i) \right\} - \frac{1}{d} \sum_{i=1,\ldots,d} F_i^{1/w_i}(X_i)\, . \right), \quad \mathbf{w} \in \mathcal{S}_{d-1}$$

A natural estimator of the multivariate madogram is given by

$$\widehat{\nu}_n(\mathbf{w}) = \frac{1}{n} \sum_{m=1}^{n} \left(\bigvee_{i=1,\ldots,d} \left\{ \hat{F}_i^{1/w_i}(X_{m,i}) \right\} - \frac{1}{d} \sum_{i=1,\ldots,d} \hat{F}_i^{1/w_i}(X_{m,i}) \right),$$

where $\hat{F}_i$ is the empirical cdf for the ith coordinate.

The Pickands function can be then estimated by

$$\widehat{A}_n^{MD}(\mathbf{w}) = \frac{\widehat{\nu}_n(\mathbf{w}) + c(\mathbf{w})}{1 - \widehat{\nu}_n(\mathbf{w}) - c(\mathbf{w})}, \qquad \mathbf{w} \in \mathcal{S}_{d-1}.$$

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}_n$ with some non-parametric estimator (e.g. MD)

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}_n$ with some non-parametric estimator (e.g. MD)
3. Derive the projection of $\hat{A}_n$ in the space $\mathcal{A}$ of functions satisfying the desired properties

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}_n$ with some non-parametric estimator (e.g. MD)
3. Derive the projection of $\hat{A}_n$ in the space $\mathcal{A}$ of functions satisfying the desired properties
4. It can be done solving the optimization problem

$$\tilde{A}_n = \arg\min_{A \in \mathcal{A}} \|\widehat{A}_n - A\|_2,$$

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}_n$ with some non-parametric estimator (e.g. MD)
3. Derive the projection of $\hat{A}_n$ in the space $\mathcal{A}$ of functions satisfying the desired properties
4. It can be done solving the optimization problem

$$\tilde{A}_n = \arg\min_{A \in \mathcal{A}} \|\widehat{A}_n - A\|_2,$$

BUT there is no closed solution

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}_n$ with some non-parametric estimator (e.g. MD)
3. Derive the projection of $\hat{A}_n$ in the space $\mathcal{A}$ of functions satisfying the desired properties
4. It can be done solving the optimization problem

$$\tilde{A}_n = \arg\min_{A \in \mathcal{A}} \|\widehat{A}_n - A\|_2,$$

BUT there is no closed solution

5. Consider a nested sequence $\mathcal{A}_k \subseteq \mathcal{A}$ of constrained multivariate Bernstein-Bézier polynomial families

$$\mathcal{A}_k = \{B_A(\mathbf{w}; k) = \mathbf{b}_k(\mathbf{w})\boldsymbol{\beta}_k : \mathbf{C}_k\boldsymbol{\beta}_k \geq \mathbf{c}_k$$

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}_n$ with some non-parametric estimator (e.g. MD)

3. Derive the projection of $\hat{A}_n$ in the space $\mathcal{A}$ of functions satisfying the desired properties

4. It can be done solving the optimization problem

$$\tilde{A}_n = \arg\min_{A \in \mathcal{A}} \|\widehat{A}_n - A\|_2,$$

BUT there is no closed solution

5. Consider a nested sequence $\mathcal{A}_k \subseteq \mathcal{A}$ of constrained multivariate Bernstein-Bézier polynomial families

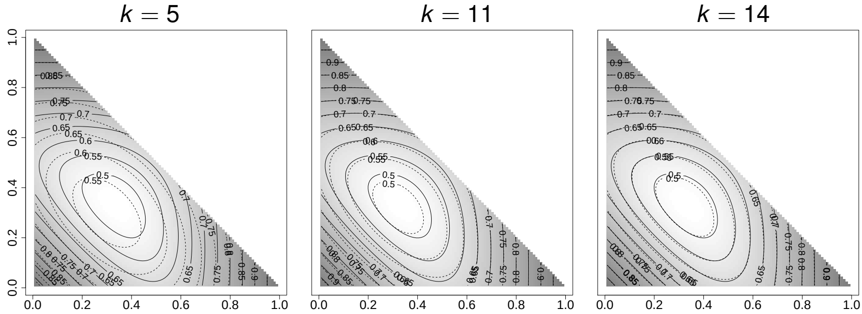
$$\mathcal{A}_k = \{B_A(\mathbf{w}; k) = \mathbf{b}_k(\mathbf{w})\boldsymbol{\beta}_k : \mathbf{C}_k\boldsymbol{\beta}_k \geq \mathbf{c}_k$$

➡ The approximate projection estimator is given by the solution of

$$\tilde{A}_{n,k} = \arg\min_{B_A \in \mathcal{A}_k} \|\widehat{A}_n - B_A\|_2$$

NUMERICAL RESULTS (d=3) Increasing k

- True dependence function of the **Symmetric Logistic Model** (solid line)
- dependence parameter $\alpha = 0.3$
- $n = 20$



Numerical Results (d=3) MISE

$$\text{MISE}(\widehat{A}_n, A) = \mathbb{E}\left\{\int_{\mathcal{S}_{d-1}} \left(\widehat{A}(\mathbf{w}) - A(\mathbf{w})\right)^2 d\mathbf{w}\right\},$$

Parameter α	Estimator	Poly's degree k	Sample size 50	100	200
0.3	HT		5.71×10^{-4}	5.27×10^{-4}	5.25×10^{-4}
	MD		4.94×10^{-4}	2.70×10^{-4}	1.94×10^{-4}
	BP-HT	14	6.77×10^{-5}	6.39×10^{-5}	6.32×10^{-5}
	BP-MD	17	5.66×10^{-5}	4.68×10^{-5}	2.52×10^{-5}
0.6	HT		3.80×10^{-3}	2.85×10^{-3}	9.69×10^{-4}
	MD		2.73×10^{-3}	1.36×10^{-3}	6.83×10^{-4}
	BP-HT	3	7.15×10^{-4}	5.08×10^{-4}	2.79×10^{-4}
	BP-MD	3	6.74×10^{-4}	3.53×10^{-4}	1.75×10^{-4}
0.9	HT		6.98×10^{-3}	4.32×10^{-3}	2.31×10^{-3}
	MD		4.81×10^{-3}	3.17×10^{-3}	1.51×10^{-3}
	BP-HT	3	1.57×10^{-3}	1.04×10^{-3}	6.05×10^{-4}
	BP-MD	3	1.12×10^{-3}	5.39×10^{-4}	4.13×10^{-4}

Confidence Bands

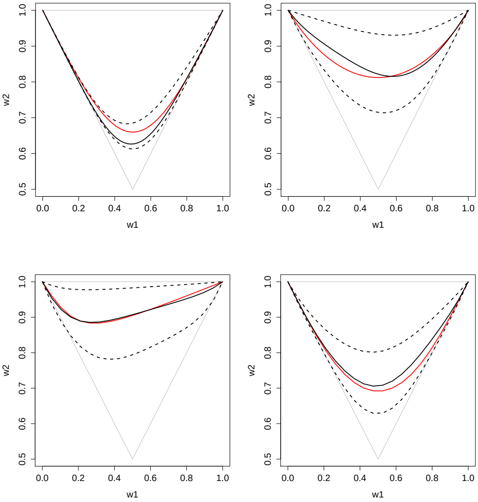
- Use bootstrap confidence bands (say $R = 500$ or $2,000$ replicates)
- Derive non-parametric $(100 - \alpha)\%$ individual confidence bands from the quantiles of the bootstrap sampling $\tilde{\beta}^*_\ell$
- Construct them, for each β_ℓ, individually

$$\mathbb{P}\left\{\tilde{\beta}_{\ell(R\alpha/2)}^{*} \leq \beta_\ell \leq \tilde{\beta}_{\ell(R(1-\alpha/2))}^{*}\right\} = 1 - \alpha \qquad \forall \ell \in L_k,$$

- Define a bootstrap simultaneous $(1 - \alpha)$ confidence band specifying the lower $\tilde{A}^L_{n,k}(\mathbf{w})$ and upper $\tilde{A}^U_{n,k}(\mathbf{w})$ limits as

$$\left[\sum_{l \in L_k} \tilde{\beta}_l^{*\lceil r(\alpha/2)\rceil} b_l(\mathbf{w}; k);\ \sum_{l \in L_k} \tilde{\beta}_l^{*\lceil r(1-\alpha/2)\rceil} b_l(\mathbf{w}; k)\right], \quad \mathbf{w} \in \mathcal{S}_{d-1},$$

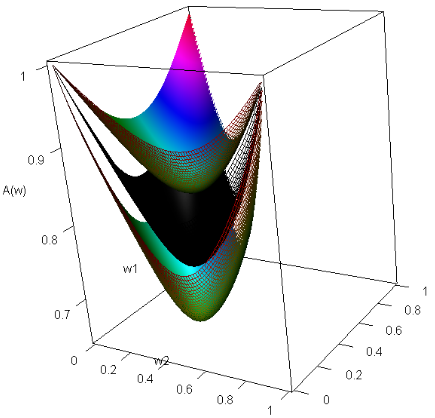
Numerical Results (d=2)

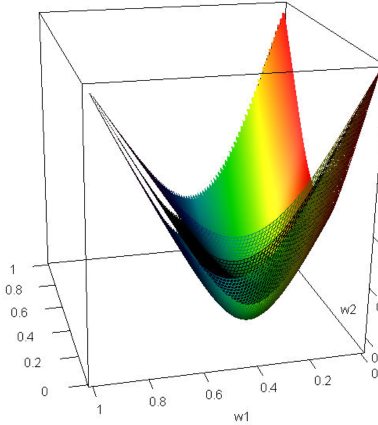


True dependence function of the models (red) :

- **Symmetric Logistic Model** with dependence parameter 0.4 and 0.7
- **Asymmetric Logistic Model** with dependence parameter 0.4 and $t_1 = 0.2$, $t_2 = 0.8$
- **Hüsler-Reiss Model** with $\lambda = 1$
- $n = 50$
- $ngrid = 200$
- $\tilde{A}(\mathbf{w})$ (solid black), Confident Bands (dotted black)

Confident Bands (d=3)



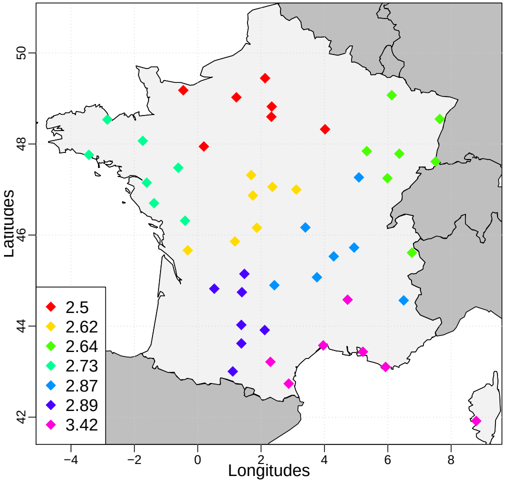


CONFIDENT BANDS (D=3)

Parameter α	Estimator	Poly's degree k	Coverage Probability $[\tilde{A}^L_{n,k},\ \tilde{A}^U_{n,k}]$
0.3	BP-HT	14	0.929
	BP-MD	17	0.930
0.6	BP-HT	3	0.904
	BP-MD	3	0.928
0.9	BP-HT	3	0.903
	BP-MD	3	0.916

WEEKLY MAXIMA OF HOURLY RAINFALL IN FRANCE (FALL SEASON, 1993-2011)

$$\widehat{\theta} = 7\,\tilde{A}^{MD}_{n,k}(1/7,\ldots,1/7)$$



Convergence results

Order statistics and empirical copulas

Bergaus B., A. Bucher and H. Dette (2013).
Mimimum distance estimators of the Pickands dependence function and related tests of MEVT dependence
SFDS 154.

Fils-Villetard, A., A. Guillou, and J. Segers (2008).
Projection estimators of pickands dependence functions.
The Canadian Journal of Statistics 36(3), 369–382.

Gudendorf, G. and J. Segers (2012).
Nonparametric estimation of multivariate extreme-value copula.
Journal of Statistical Planning and Inference 142(3073–3085).

Marcon G., Padoan S.A., Naveau P. and Muliere P. (2014).
Multivariate Nonparametric Estimation of the Pickands Dependence Function using Bernstein Polynomials.
Submitted (ArXiv).

Conclusions

Multivariate max-stable vectors

- A well-developed theory with a lot of inference approach
- Parametric versus non-parametric
- Still a lot of open questions and room for improvement (spatio-temporal, non-stationary, extreme versus non-extremes, covariates, dynamics, dimension reduction, ...)
- for recent advances in MEVT, see the journal Extremes

Conclusions

Multivariate max-stable vectors

- A well-developed theory with a lot of inference approach
- Parametric versus non-parametric
- Still a lot of open questions and room for improvement (spatio-temporal, non-stationary, extreme versus non-extremes, covariates, dynamics, dimension reduction, ...)
- for recent advances in MEVT, see the journal Extremes

Coming events

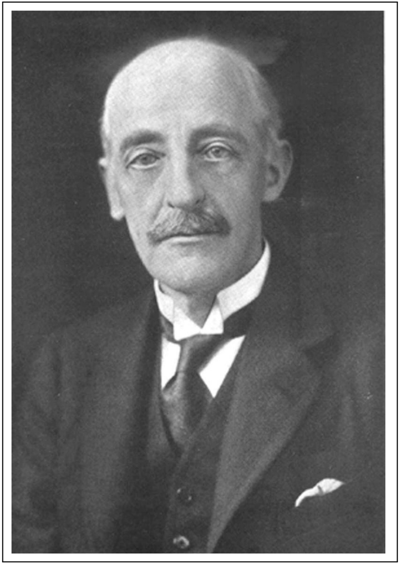
- Workshop on stochastic weather generators, Avignon. Sep, 17-19, 2014
- Workshop on Extreme Value Theory, Spatial and Temporal Aspects, in Besançon, November 3-5, 2014
- EVA (Extreme Value Analysis), Michigan, June 15-19, 2015

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".

QUIZ

- *(A) Gilbert Walker*
- *(B) Ed Lorenz*
- *(C) Rol Madden*
- *(D) Francis Zwiers*



Necessary condition for applying our divergence approximation

$$\lim_{u \to x_F} \int_u^{x_F} \left(\log \frac{f(x)}{\overline{F}(x)} - \log \frac{g(x)}{\overline{G}(x)} \right) (f_u(x) - g_u(x))\, dx = 0$$

Necessary condition for applying our divergence approximation

$$\lim_{u \to x_F} \int_u^{x_F} \left(\log \frac{f(x)}{\overline{F}(x)} - \log \frac{g(x)}{\overline{G}(x)} \right) (f_u(x) - g_u(x))\, dx = 0$$

Necessary condition for applying our divergence approximation

$$\lim_{u \to x_F} \int_u^{x_F} \left(\log \frac{f(x)}{\overline{F}(x)} - \log \frac{g(x)}{\overline{G}(x)} \right) (f_u(x) - g_u(x))\, dx = 0$$

If $\left(\log \frac{f(x)}{\overline{F}(x)} - \log \frac{g(x)}{\overline{G}(x)} \right) \sim$ constant for large x, then

$$\begin{aligned} \text{constant} \times \int_u^{x_F} (f_u(x) - g_u(x))\, dx &= \int_u^{x_F} f_u(x) dx - \int_u^{x_F} g_u(x) dx, \\ &= 1 - 1 = 0 \end{aligned}$$

Checking $\lim \int_u^{x_F} \left(\log \frac{f(x)}{\overline{F}(x)} - \log \frac{g(x)}{\overline{G}(x)}\right)(f_u(x) - g_u(x))\, dx = 0$

Basically, this condition is satisfied for all classical distributions and more.

PROPOSITION 4. *Suppose that $X_u \geq_{st} Y_u$ for large u and define*

$$\alpha(x) = \log\left(\frac{f(x)}{\overline{F}(x)}\right) - \log\left(\frac{g(x)}{\overline{G}(x)}\right).$$

If $\mathbb{E}(X_u)$, $\mathbb{E}(\alpha(X_u))$ and $\mathbb{E}(\alpha(Y_u))$ are finite and the derivative $\alpha'(\cdot)$ is monotone and goes to zero as $x \uparrow \tau$, then (2) is satisfied.