Extreme value theory and applications in climate

Théo Rietsch

Ph.D funded by the GIS-PEPER project

GEV approach

Theorem (Fisher, 1928)

Let $(X_1, \ldots, X_n)$ be an i.i.d. sample of random values, if there exists $(a_n)_n > 0$ and $(b_n)_n \in \mathbb{R}$, and a non degenerate distribution H such that

$$\lim_{n\to\infty} \mathbb{P}\left(\bigvee_{i=1}^{n} \frac{X_i - b_n}{a_n} \leq x \right) = H(x)$$

then H can be written

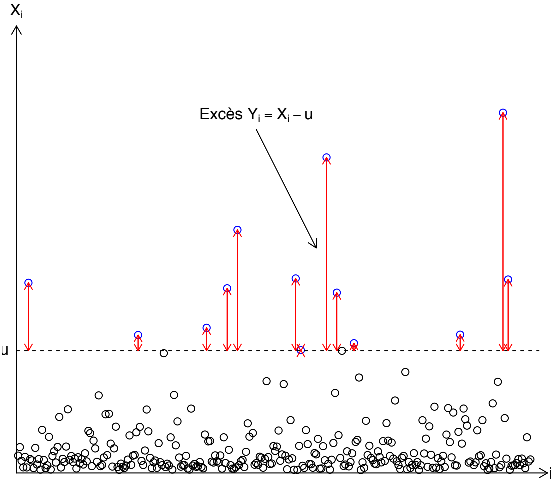
$$H_{\mu,\sigma,\xi}(x) := \exp\left\{ - \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]_{+}^{-\frac{1}{\xi}} \right\}$$

Domains of attraction

- $\xi > 0 \Rightarrow \tau_F = +\infty$: Fréchet
- $\xi < 0 \Rightarrow \tau_F < +\infty$: Weibull
- $\xi = 0 \Rightarrow \tau_F < \infty$ or $\tau_F = \infty$: Gumbel

with $\tau_F = \sup\{x : F(x) < 1\}$. We denote $F \in \mathcal{D}(H_\xi)$.

POT approach (Peaks over threshold)



$$F_u(x) = \mathbb{P}(Y \leq x \mid X > u) = \frac{F(u+x)-F(u)}{1-F(u)}$$

POT approach (Peaks over threshold)

Theorem (Pickands, 1975)

F belongs to the domain of attraction $\mathcal{D}(H_\xi)$ if and only if there exists a real ξ and a function σ such that

$$\lim_{u\to\tau_F} \sup_{0<x<\tau_F-u} \mid F_u(x) - G_{\sigma(u),\xi}(x) \mid = 0,$$

where

$$\overline{G}_{\sigma(u),\xi}(x) := \begin{cases} \left(1+\xi\frac{x}{\sigma(u)}\right)^{-1/\xi} & \text{for } \xi \neq 0 \text{ and } 1+\xi\frac{x}{\sigma(u)} > 0 \\ e^{-x/\sigma(u)} & \text{for } \xi = 0. \end{cases}$$

Motivation

Question: Did the distribution of the extremes changed during the last century ?

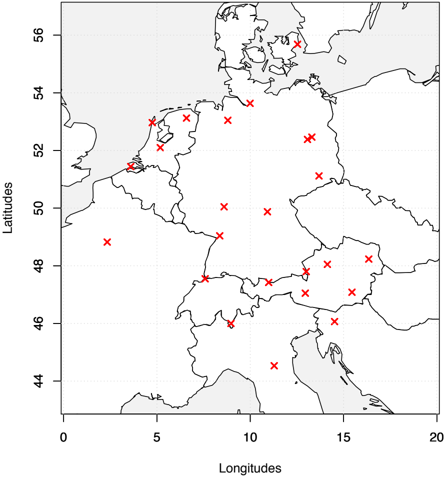


Figure 1: Stations.

Motivation

Question: Did the distribution of the extremes changed during the last century ?

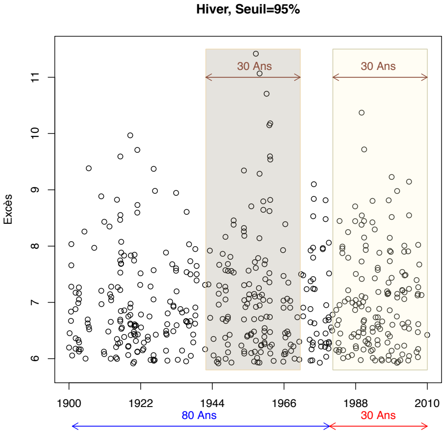


Figure 1: Excesses from 1900 to 2010, Winter.

Existing approaches

• Analyse the evolution of the frequency and intensity of some extreme weather indicators (consecutive dry days, warmest days...)

$\hookrightarrow$ Alexander *et al.* (2006)

• Study tne evolution of the parameters of the distribution of the extremes :

$\hookrightarrow$ Jaruskova et Rencova (2008)

$\Longrightarrow$ Drawbacks :

1. GEV approach : we have to suppose that the distribution is a GEV and can only consider maxima
2. In In both cases : even if the shape parameter is a good indicator of the behaviour of the tail, there are two parameters remaining

Kullback-Leibler divergence

Kullback-Leibler information or entropy (Kullback, 1968)

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

Divergence:

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

Symmetrical measure (not a distance) with respect to *f* and *g*.

Advantages of the divergence

- Simple and fast to compute
- Gives only one value that sums up the distance between the densities
- Pluridisciplinar tool (stat and climatology)
- Metric, not an index
- Link with model section criteria (AIC, BIC)

Negative point of the Kullback Leibler divergence

The divergence is based on the density functions f and g but it is easier to work with the survival functions $\overline{F}$ and $\overline{G}$ since they capture the behaviour of the tail.

Entropy and excesses

Notations: We consider $X \sim f$ and $Y \sim g$ two random variables with the same endpoint τ. For all u, define $X_u = [X|X > u]$ with density

$$f_u(x) = \frac{f(x)}{\overline{F}(u)} \quad \forall x \in (u, \tau).$$

Definition of the divergence for extremes

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

Approximation of the divergence

Proposition 1

Under the hypothesis

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

and

$$\mathbb{E}_f\left\{\log\frac{\overline{G}(X)}{\overline{G}(u)}\middle| X > u\right\} < \infty \quad \textit{and} \quad \mathbb{E}_g\left\{\log\frac{\overline{F}(X)}{\overline{F}(u)}\middle| Y > u\right\} < \infty,$$

we have that when $u\uparrow\tau$, the divergence $D(f_u;g_u) = I(f_u;g_u) + I(g_u;f_u)$ is equivalent to

$$K(f_u;g_u) = -L(f_u;g_u) - L(g_u;f_u)$$

where

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

Approximation of the divergence

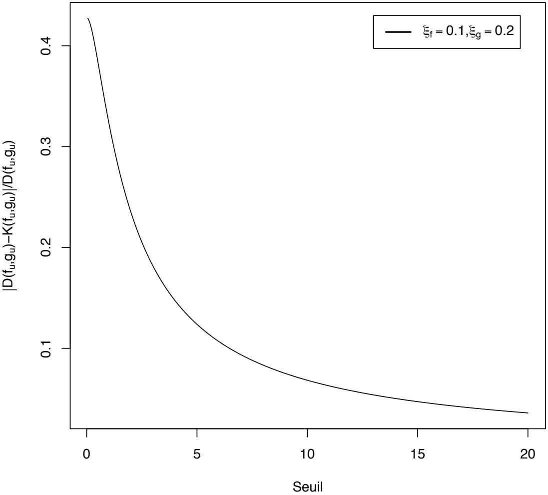


Figure 2: Relative difference between the true divergence $D(f_u; g_u)$ and its approximation $K(f_u; g_u)$ as a function of the threshold.

Classical examples : Fréchet

Proposition 2

Suppose that X and Y are in the Fréchet domain of attraction. It means that $\tau = \infty$ and $\overline{F} \in \mathcal{RV}_{-\alpha}$ and $\overline{G} \in \mathcal{RV}_{-\beta}$.

We also impose the following second order condition :

$$\lim_{t\to\infty} \frac{\frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-\alpha}}{q_F(t)} = x^{-\alpha}\frac{x^{\rho}-1}{\rho} \ \textit{et} \ \lim_{t\to\infty} \frac{\frac{\overline{G}(tx)}{\overline{G}(t)} - x^{-\beta}}{q_G(t)} = x^{-\beta}\frac{x^{\eta}-1}{\eta}$$

with $\rho < 0$ and $\eta < 0$. If the functions

$$B(x) = \frac{xf(x)}{\overline{F}(x)} - \alpha \ \textit{and} \ C(x) = \frac{xg(x)}{\overline{G}(x)} - \beta$$

are asymptotically monotone, then Proposition 1 holds.

Classical : Weibull

Proposition 3

Suppose that X and Y are in the Weibull domain of attraction. It means that $\tau < \infty$. Proposition 1 holds under the same hypotheses than in the Fréchet case, except that the hypotheses are made on the survival functions $\overline{F_*}(x) := \overline{F}\left(\tau - x^{-1}\right)$ and $\overline{G_*}(x) := \overline{G}\left(\tau - x^{-1}\right)$ rather than on $\overline{F}(x)$ and $\overline{G}(x)$.

General result

Proposition 4

Consider two random variables X and Y, with the same endpoint τ. Let $u > 0$ and suppose $X_u \geq_{st} Y_u$ for u sufficiently large, which means

$$\mathbb{P}(X_u > t) \geq \mathbb{P}(Y_u > t), \quad \forall t \geq u.$$

Under suitable hypotheses on the function

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

Proposition 1 holds.

Example : $X \sim \mathcal{E}(1)$ and $Y \sim \mathcal{N}(0,1)$

Estimator of the divergence

$$\mathbf{X} = (X_1, \ldots, X_n)^T \quad \text{and} \quad \mathbf{Y} = (Y_1, \ldots, Y_m)^T$$

Define

$$\widetilde{\overline{G}}_m(t) \quad := \quad 1 - \frac{1}{m+1} \sum_{j=1}^{m} \mathbb{1}_{\{Y_j \leq t\}}$$

and

$$\widehat{L}(f_u; g_u) := 1 + \frac{1}{N_n} \sum_{i=1}^{n} \log \left(\frac{\widetilde{\overline{G}}_m(X_i \vee u)}{\widetilde{\overline{G}}_m(u)} \right)$$

where $N_n := \# \{X_i, X_i \geq u\}$.

Main result

Theorem

Suppose F and G are continuous with same endpoint τ. *Set* $u < \tau$, $\frac{n}{m} \to c \in (0,\infty)$ *and*

$$\mathbb{E}_f\left(\log\left(\frac{\overline{G}(X \vee u)}{\overline{G}(u)}\right)^2\right) < \infty$$

We suppose that there exists two positive and decreasing sequences k_n/n *and* ℓ_m/m, *that verify*

$$k_n \geq \max\left(\log n, 8n\overline{F}\left(\overline{G}^{\leftarrow}\left(\frac{\ell_m}{m}\right)\right)\right), \quad \frac{k_n}{n}\log n \to 0 \quad et \quad \frac{\ell_m}{\log\log m} \to \infty.$$

Then we have

$$\widehat{L}(f_u; g_u) - L(f_u; g_u) = o(1)\ p.s.$$

Main result

Corollary

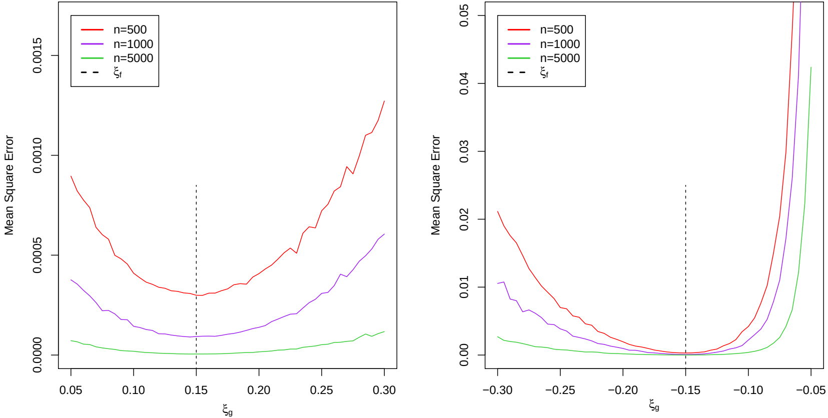
Under the same hypotheses as in theorem 1, adding

$$\mathbb{E}_g\left(\log\left(\frac{\overline{F}(Y\vee u)}{\overline{F}(u)}\right)^2\right)<\infty,$$

we have

$$\widehat{K}(f_u;g_u)-K(f_u;g_u)=o(1)\ \ p.s.$$

Simulations



$f = GPD(\mu = 0, \sigma = 1, \xi = \xi_f)$ et $g = GPD(\mu = 0, \sigma = 1, \xi = \xi_g)$

$f = GPD(\mu = 0, \sigma = \frac{1}{\xi_f}, \xi = \xi_f)$ et $g = GPD(\mu = 0, \sigma = \frac{1}{\xi_g}, \xi = \xi_g)$

Figure 3: MSE of $\widehat{K}$ as a function of ξ_g. Fréchet domain of attraction on the left, Weibull on the right.

Simulations

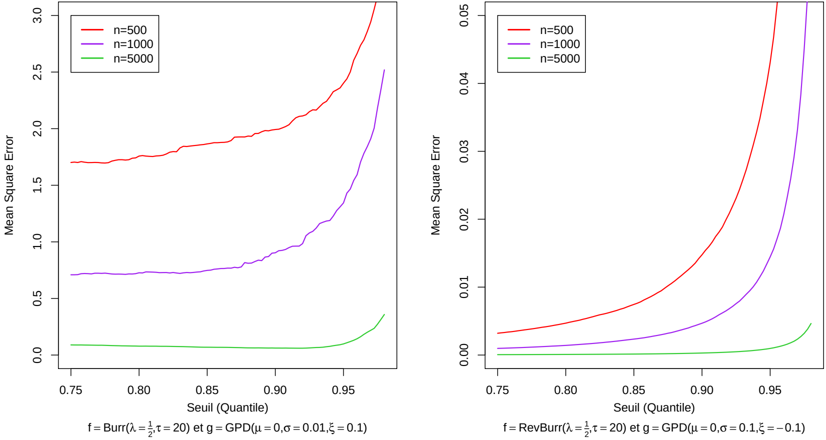


Figure 4: MSE of $\widehat{K}$ as a function of the threshold.

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities f and g

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities *f* and *g*
- Resample within the sample $(\mathcal{F}, \mathcal{G})$

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities f and g
- Resample within the sample $(\mathcal{F}, \mathcal{G})$
- Split the sample into two parts of size 1000

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities f and g
- Resample within the sample $(\mathcal{F}, \mathcal{G})$
- Split the sample into two parts of size 1000
- Compute the divergence between the two samples

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities f and g
- Resample within the sample $(\mathcal{F}, \mathcal{G})$
- Split the sample into two parts of size 1000
- Compute the divergence between the two samples
- Do that 100 times $\Rightarrow$ 100 values of the divergence under the null hypothesis

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities f and g
- Resample within the sample $(\mathcal{F}, \mathcal{G})$
- Split the sample into two parts of size 1000
- Compute the divergence between the two samples
- Do that 100 times $\Rightarrow$ 100 values of the divergence under the null hypothesis
- Take the 95% quantile q_{95} of the sample of these 100 divergences

Hypothesis testing

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

- Simulate two samples $\mathcal{F}$ and $\mathcal{G}$ of size 1000 from two densities f and g
- Resample within the sample $(\mathcal{F}, \mathcal{G})$
- Split the sample into two parts of size 1000
- Compute the divergence between the two samples
- Do that 100 times $\Rightarrow$ 100 values of the divergence under the null hypothesis
- Take the 95% quantile q_{95} of the sample of these 100 divergences
- Compute the divergence between $\mathcal{F}$ and $\mathcal{G}$ and compare it to q_{95}

Hypothesis testing

Weibull-Weibull case	$\xi_f = -.1$															
ξ_g n	$-.15$				$-.1$				$-.08$				$-.05$			
50	691	605	563	**550**	26	41	48	**54**	889	845	841	**816**	241	153	123	**93**
100	445	318	266	**233**	50	70	69	**58**	824	739	719	**680**	24	15	11	**5**
200	110	56	32	**35**	30	49	47	**45**	627	530	476	**469**	0	0	0	**0**
500	0	0	0	**0**	52	57	51	**62**	189	131	95	**94**	0	0	0	**0**
1000	0	0	0	**0**	42	48	57	**50**	17	10	4	**5**	0	0	0	**0**

Hypothesis testing

Fréchet-Fréchet case	$\xi_f = .1$															
ξ_g n	.05				.08				.1				.2			
10	986	955	953	**946**	991	952	960	**939**	15	41	44	**49**	990	957	963	**950**
100	973	944	945	**935**	960	944	950	**941**	43	59	59	**49**	947	925	925	**899**
1000	937	923	901	**740**	946	957	955	**926**	45	34	34	**49**	854	835	740	**285**
10000	648	701	470	**9**	911	903	859	**597**	41	54	54	**55**	12	182	14	**0**

Hypothesis testing

Gumbel-Weibull case	$\xi_f = 0$															
ξ_g n	$-.4$				$-.3$				$-.2$				0			
50	868	847	757	**377**	921	893	847	**628**	949	934	912	**799**	28	45	50	**48**
100	727	730	556	**37**	896	853	743	**231**	933	906	862	**599**	38	53	45	**50**
200	348	534	280	**0**	709	722	502	**12**	888	870	777	**252**	37	48	52	**60**
500	1	159	19	**0**	134	424	138	**0**	616	712	490	**4**	58	55	52	**53**
1000	0	12	0	**0**	0	139	7	**0**	229	452	178	**0**	36	55	44	**54**

Hypothesis testing

Gumbel-Fréchet case	$\xi_f = 0$															
ξ_g n	0				.2				.3				.4			
50	33	57	58	**63**	943	931	921	**853**	945	924	898	**746**	919	861	794	**589**
100	36	54	62	**74**	923	900	860	**695**	907	866	794	**470**	844	813	690	**253**
200	30	53	52	**57**	902	886	794	**412**	807	768	626	**120**	606	634	389	**18**
500	43	40	48	**55**	706	737	541	**52**	335	505	222	**0**	101	306	69	**0**
1000	45	40	51	**59**	423	541	262	**1**	34	213	31	**0**	0	71	0	**0**

Protocol

Protocol to compare two periods of 30 years: For each season

- Reference sample : $[1981, 2010]$
- Moving window of 30 years : $[1900+t, 1929+t]$ for $t \in \{1, \ldots, 80\}$
- Threshold u : mean of the 95% quantiles of the two periods

$\Rightarrow$ estimator of the divergence between the reference sample and the moving window for all t

Allows us to see the evolution of the behaviour of the extremes.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

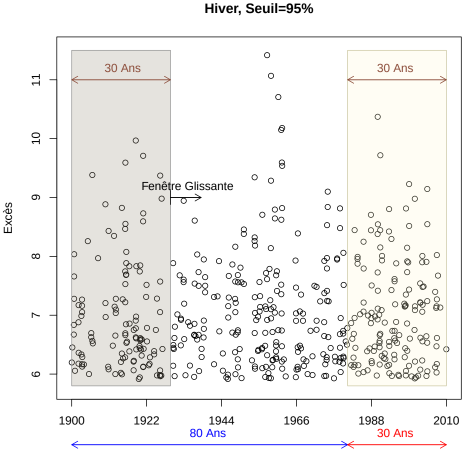


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

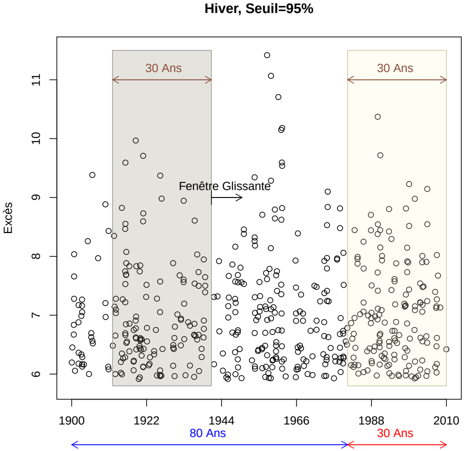


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

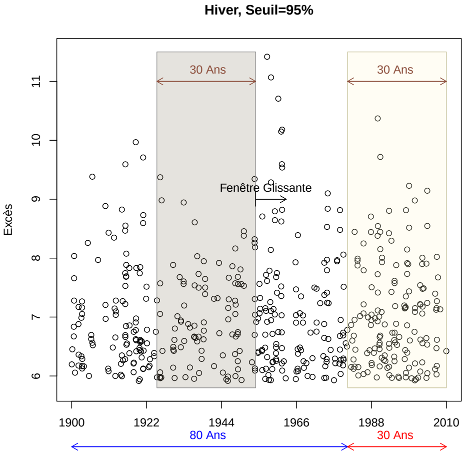


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

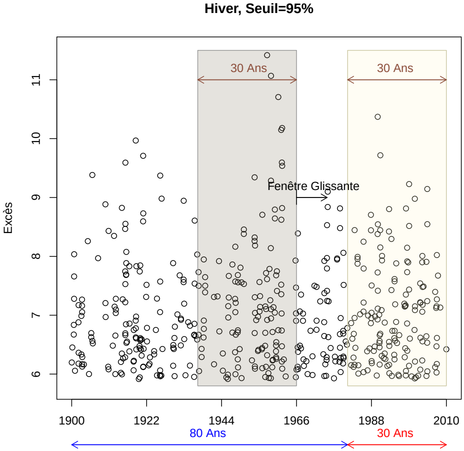


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

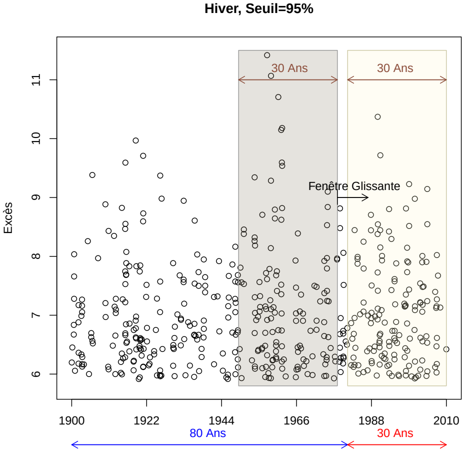


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

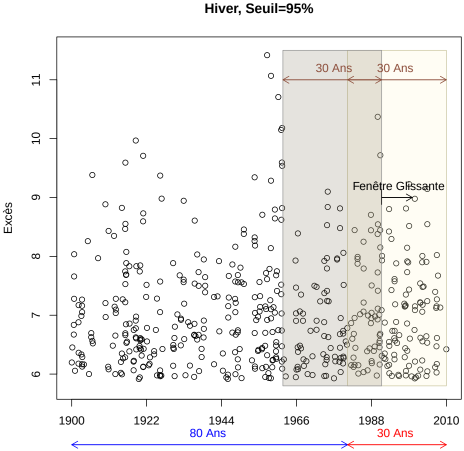


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

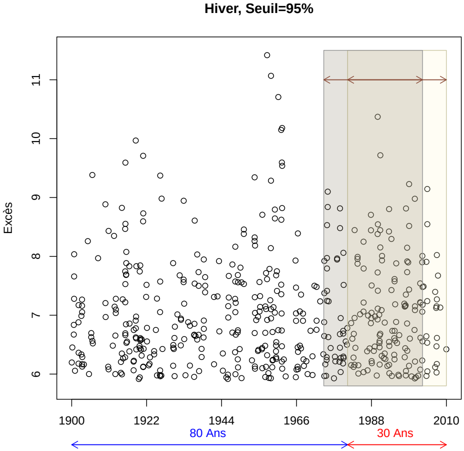


Figure 5: Excesses from 1900 to 2010, Winter.

Protocol

Question: Did the distribution of the extremes changed during the last century ?

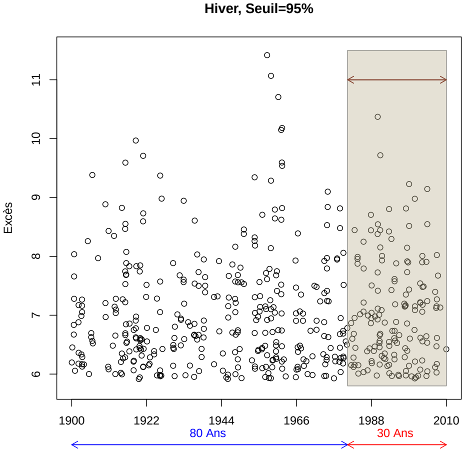


Figure 5: Excesses from 1900 to 2010, Winter.

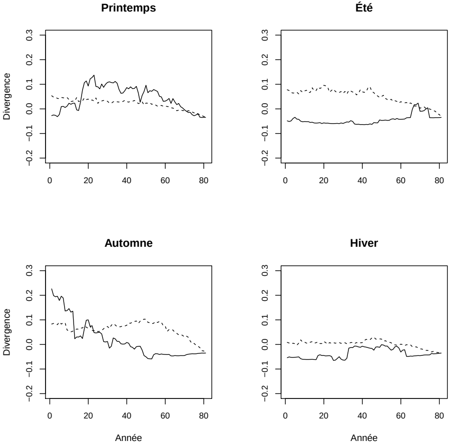
Protocol

If there is no modification in the distribution of the extremes, resampling randomly in the dataset should not have any influence on the result

In order to compare the divergence with a reference value, we adopt the following method for each season (Davis *et al.*, 2012) :

- The size of the sample is 111 years. We resample these years randomly.
- We reapply the protocol to this new sample.
- We repeat these two operations 100 times and take the 95% quantile of the sample of the 100 estimations obtained.

Montsouris data



We detect a difference in the behaviour of the extremes in Spring and Fall.

Maxima

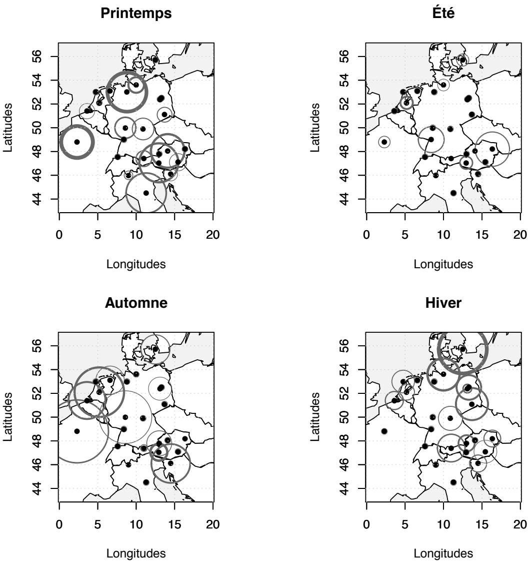


Figure 6: Results for the maxima of temperature.

Minima

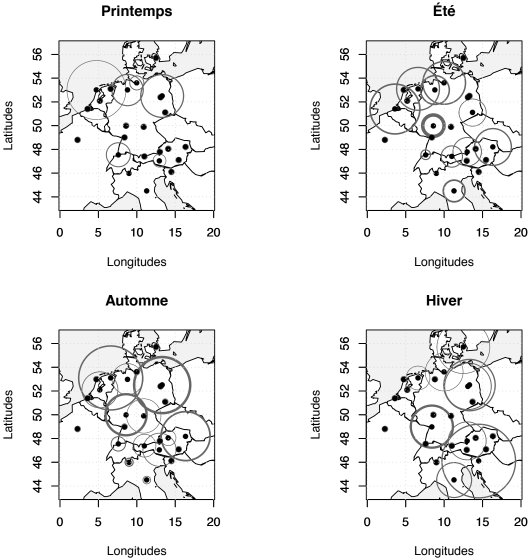
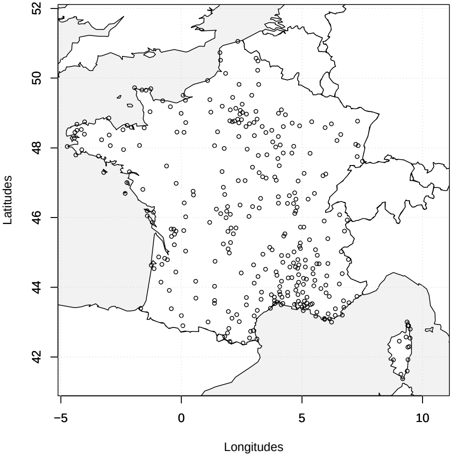


Figure 6: Results for the minima of temperature.

Network of stations



Problem

Questions

1. How to model the behaviour of the extreme precipitation on the territory ?

Problem

Questions

1. How to model the behaviour of the extreme precipitation on the territory ?
2. How to spot the stations that bring the least information about the behaviour of the extremes if we want to narrow the network ?

Problem

Questions

1. How to model the behaviour of the extreme precipitation on the territory ? $\Rightarrow$ Extreme value theory and neural networks
2. How to spot the stations that bring the least information about the behaviour of the extremes if we want to narrow the network ?

Problem

Questions

1. How to model the behaviour of the extreme precipitation on the territory ? ⇒ Extreme value theory and neural networks
2. How to spot the stations that bring the least information about the behaviour of the extremes if we want to narrow the network ?
   ⇒ Query by Committee

Hypothesis

We suppose that the extremes follow a GPD with parameters $\sigma(\mathbf{x})$ and $\xi(\mathbf{x})$ on each point $\mathbf{x}$.

$\Rightarrow$ The problem is reduced to the interpolation of the two functions $\mathbf{x} \mapsto \sigma(\mathbf{x})$ and $\mathbf{x} \mapsto \xi(\mathbf{x})$.

Estimation of the parameters

We get the following map of estimations

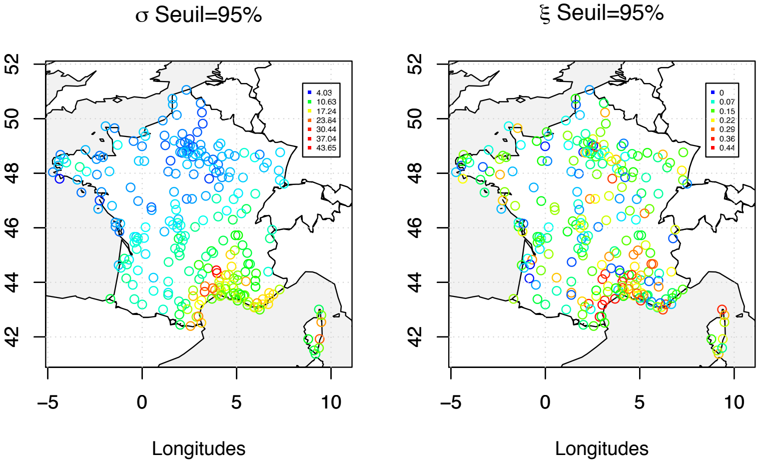


Figure 7: Estimations of the parameters on each station of the network.

Interpolation

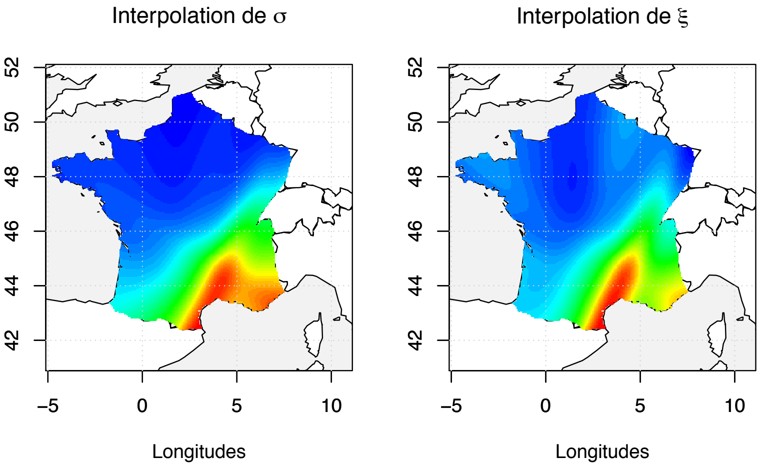


Figure 8: Interpolation of σ and ξ on the whole territory.

1. Extreme value theory
2. Kullback-Leibler divergence and extreme values
   - Problem
   - Divergence
   - Simulations
   - Real data
3. Query by Committee and optimal design
   - Motivation
   - Behaviour of the extremes
   - Query by Committee
   - Simulations
   - Real data
4. Perspectives

Presentation

Query by committee (Seung *et al.*, 1992)

- Algorithm that comes from machine learning theory
- Used to choose where to make new observations
- Particularly useful when it is costly to obtain new observations

1D Illustration

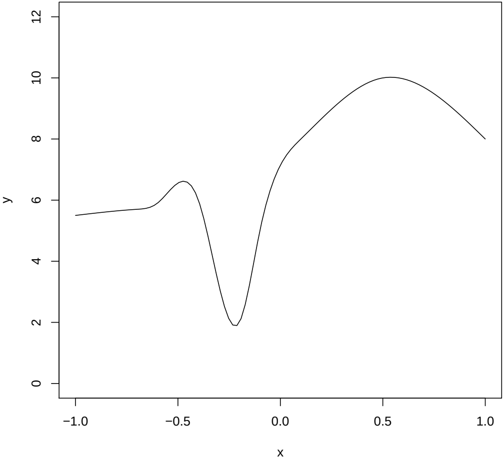
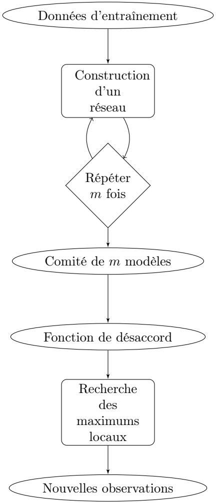


Figure 9: Function that we want to estimate.

1D Illustration



Données d'entraînement

Construction d'un réseau

Répéter m fois

Comité de m modèles

Fonction de désaccord

Recherche des maximums locaux

Nouvelles observations

Figure 10: Protocol.

1D Illustration

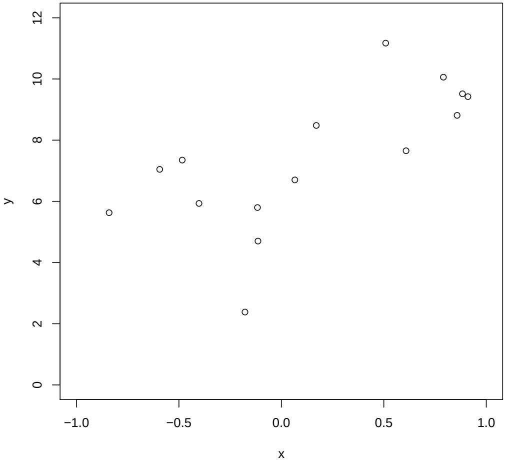
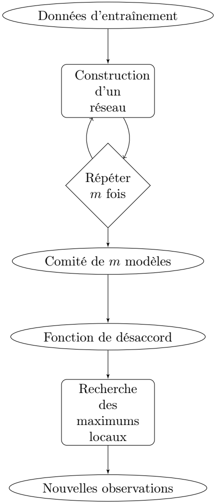


Figure 10: Data.

1D Illustration



Données d'entraînement

Construction d'un réseau

Répéter m fois

Comité de m modèles

Fonction de désaccord

Recherche des maximums locaux

Nouvelles observations

Figure 10: Protocol.

1D Illustration

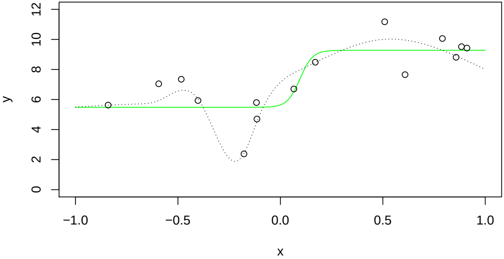


Figure 10: Model.

1D Illustration

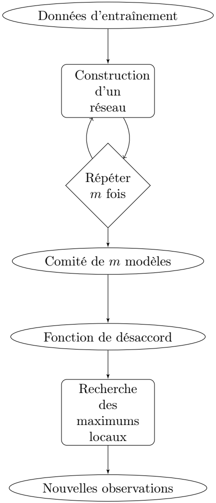


Figure 10: Protocol.

1D Illustration

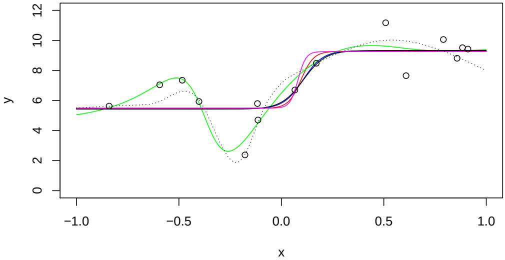
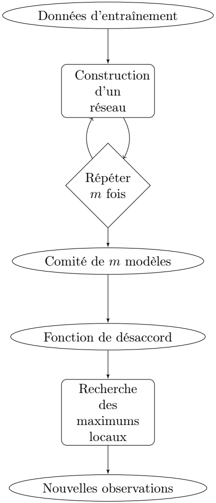


Figure 10: Committee.

1D Illustration



Données d'entraînement

Construction d'un réseau

Répéter m fois

Comité de m modèles

Fonction de désaccord

Recherche des maximums locaux

Nouvelles observations

Figure 10: Protocol.

1D Illustration

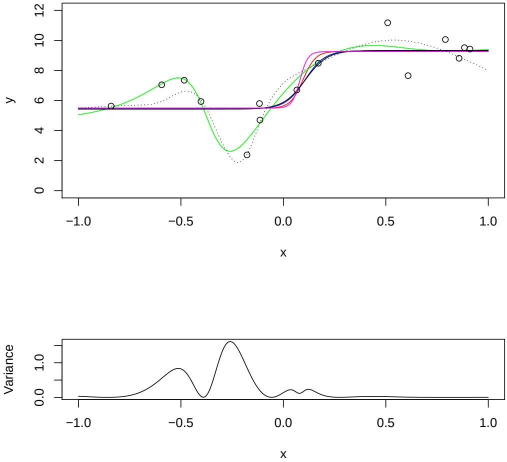


Figure 10: Disagreement function.

1D Illustration

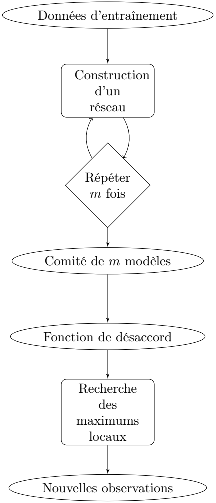


Figure 10: Protocol.

1D Illustration

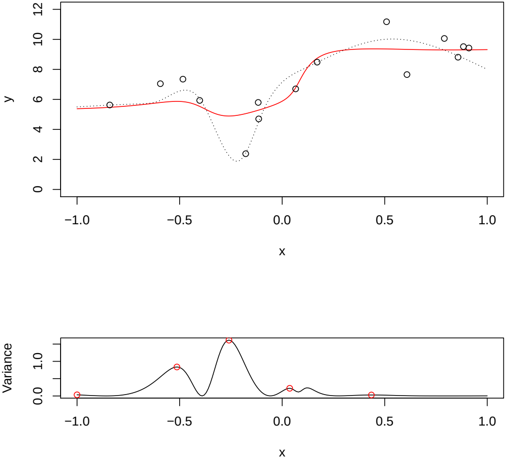
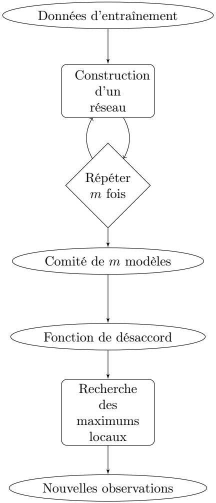


Figure 10: Maxima.

1D Illustration



Données d'entraînement

Construction d'un réseau

Répéter m fois

Comité de m modèles

Fonction de désaccord

Recherche des maximums locaux

Nouvelles observations

Figure 10: Protocol.

1D Illustration

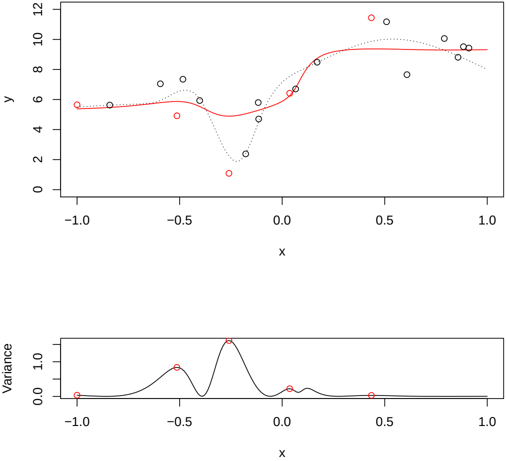


Figure 10: Design of experiment.

1D Illustration

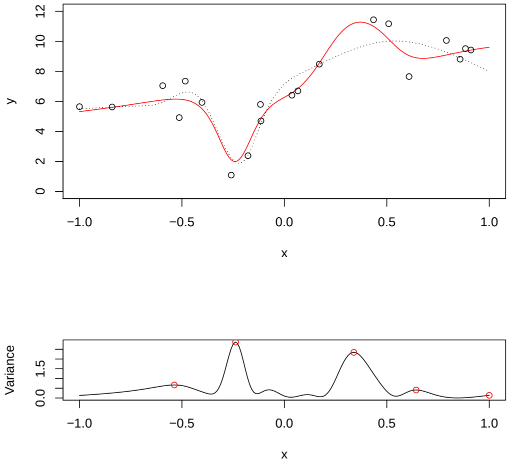


Figure 10: Step 2.

1D Illustration

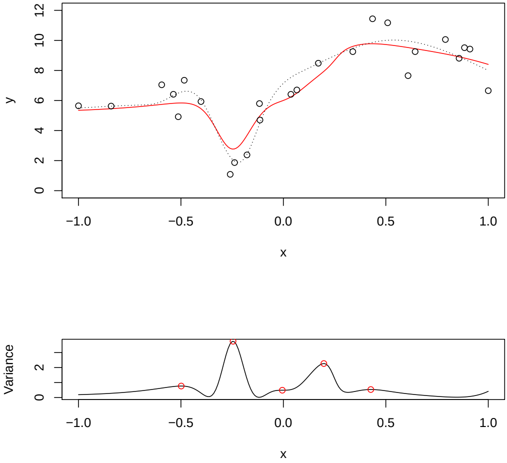


Figure 10: Step 3.

1D Illustration

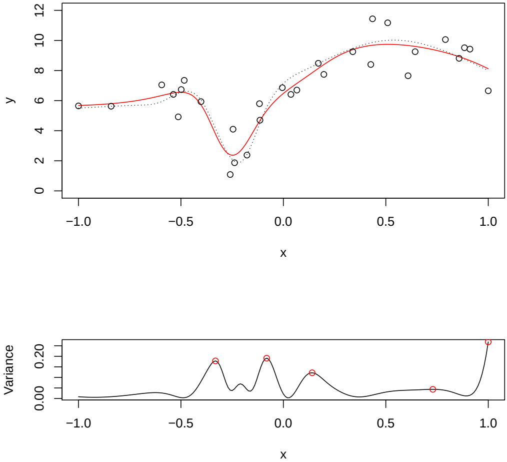


Figure 10: Step 4.

Problems that arise with meteorological stations

In the previous example, the aim was to add stations. It is not possible in our weather stations case since it is impossible to get information where there is no stations.
⇒ We focus on the problem of removing stations

Problems that arise with meteorological stations

In the previous example, the aim was to add stations. It is not possible in our weather stations case since it is impossible to get information where there is no stations.
$\Rightarrow$ We focus on the problem of removing stations
We split the network into three parts :

Problems that arise with meteorological stations

In the previous example, the aim was to add stations. It is not possible in our weather stations case since it is impossible to get information where there is no stations.
⇒ We focus on the problem of removing stations
We split the network into three parts :

- A set $\mathcal{X}_1$ of stations that cannot be removed (airports and such likes).

Problems that arise with meteorological stations

In the previous example, the aim was to add stations. It is not possible in our weather stations case since it is impossible to get information where there is no stations.
$\Rightarrow$ We focus on the problem of removing stations
We split the network into three parts :

- A set $\mathcal{X}_1$ of stations that cannot be removed (airports and such likes).
- A set $\mathcal{X}_2$ of stations used to test the quality of interpolation.

Problems that arise with meteorological stations

In the previous example, the aim was to add stations. It is not possible in our weather stations case since it is impossible to get information where there is no stations.
$\Rightarrow$ We focus on the problem of removing stations
We split the network into three parts :

- A set $\mathcal{X}_1$ of stations that cannot be removed (airports and such likes).
- A set $\mathcal{X}_2$ of stations used to test the quality of interpolation.
- A set $\mathcal{X}_3$ of stations that we can delete.

Transformation of the l’algorithme

The idea is to compute the disagreement function only at the stations of $\mathcal{X}_3$ in order to see the most informative ones.

Transformation of the l'algorithme

The idea is to compute the disagreement function only at the stations of $\mathcal{X}_3$ in order to see the most informative ones.
We apply the QBC the following way :

Protocol

Transformation of the l'algorithme

The idea is to compute the disagreement function only at the stations of $\mathcal{X}_3$ in order to see the most informative ones.
We apply the QBC the following way :

Protocol

- We use the set $\mathcal{X}_1$ as our initial training sample in the algorithm

Transformation of the l'algorithme

The idea is to compute the disagreement function only at the stations of $\mathcal{X}_3$ in order to see the most informative ones.
We apply the QBC the following way :

Protocol

- We use the set $\mathcal{X}_1$ as our initial training sample in the algorithm
- We create a committee of 100 experts (i.e. 100 neural networks)

Transformation of the l'algorithme

The idea is to compute the disagreement function only at the stations of $\mathcal{X}_3$ in order to see the most informative ones.
We apply the QBC the following way :

Protocol

- We use the set $\mathcal{X}_1$ as our initial training sample in the algorithm
- We create a committee of 100 experts (i.e. 100 neural networks)
- We choose a disagreement function

Transformation of the l'algorithme

The idea is to compute the disagreement function only at the stations of $\mathcal{X}_3$ in order to see the most informative ones.
We apply the QBC the following way :

Protocol

- We use the set $\mathcal{X}_1$ as our initial training sample in the algorithm
- We create a committee of 100 experts (i.e. 100 neural networks)
- We choose a disagreement function
- At each step, we add to the training sample $\mathcal{X}_1$ the stations of $\mathcal{X}_3$ for which the cost function is maximal.

Disagreement function

We choose to consider the three following cost functions :

- The variance between the experts modelling $\sigma \Rightarrow d_\sigma$

We will compare the performances of the different cost functions

Disagreement function

We choose to consider the three following cost functions :

- The variance between the experts modelling $\sigma \Rightarrow d_\sigma$
- The variance between the experts modelling $\xi \Rightarrow d_\xi$

We will compare the performances of the different cost functions

Disagreement function

We choose to consider the three following cost functions :

- The variance between the experts modelling $\sigma \Rightarrow d_\sigma$
- The variance between the experts modelling $\xi \Rightarrow d_\xi$
- The variance between the 95% quantiles of the GP distribution associated to the parameters interpolated by the experts $\Rightarrow d_p$

We will compare the performances of the different cost functions

Correlation between the estimators

Problems linked to the estimation of the two parameters σ and ξ :

1. There is a negative correlation between the two estimators.
2. The shape parameter ξ is harder to estimate than the scale parameter σ.

We simulated 100 samples from a GPD with $(\sigma, \xi) = (25, 0.5)$ and applied our estimator on each sample.

Illustration of the correlation

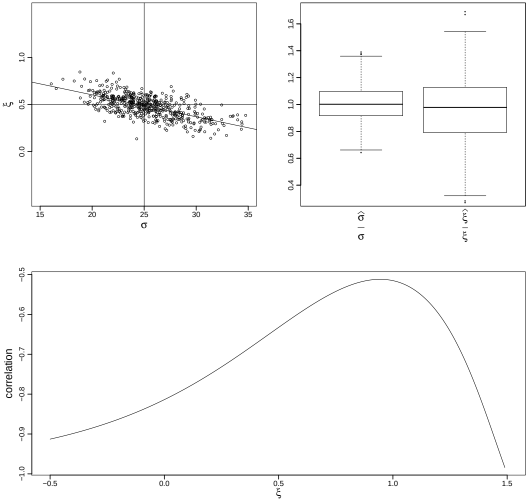


Figure 11: Correlation between the estimators of the parameters of the GPD.

Final protocol

We applied this resampling technique for each station

Final protocol

We applied this resampling technique for each station
⇒ 100 estimations by station

Final protocol

We applied this resampling technique for each station
⇒ 100 estimations by station

For each set of estimations we applied the QBC algorithm

Final protocol

We applied this resampling technique for each station
⇒ 100 estimations by station

For each set of estimations we applied the QBC algorithm

We compared the frequency of dismissal of each station by the algorithm over the 100 replications

1. Extreme value theory

2. Kullback-Leibler divergence and extreme values
   - Problem
   - Divergence
   - Simulations
   - Real data

3. Query by Committee and optimal design
   - Motivation
   - Behaviour of the extremes
   - Query by Committee
   - Simulations
   - Real data

4. Perspectives

Studied cases

Functions:

- $\sigma(x) = 30$
- $\xi(x) = \frac{1}{2}\left(1.3 - \exp\left(-16x^2\right)\right) - 0.1$

Studied cases

Functions:

- $\sigma(x) = 30$
- $\xi(x) = \frac{1}{2}\left(1.3 - \exp\left(-16x^2\right)\right) - 0.1$

Parameters :

- 55 points
- 20 in $\mathcal{X}_1$
- 15 in $\mathcal{X}_2$
- 20 in $\mathcal{X}_3$

Studied cases

Functions:

- $\sigma(x) = 30$
- $\xi(x) = \frac{1}{2}\left(1.3 - \exp\left(-16x^2\right)\right) - 0.1$

Parameters :

- 55 points
- 20 in $\mathcal{X}_1$
- 15 in $\mathcal{X}_2$
- 20 in $\mathcal{X}_3$

We add 3 stations to $\mathcal{X}_1$ at each iteration and do 5 iterations $\Rightarrow$ 5 stations suppressed

Results

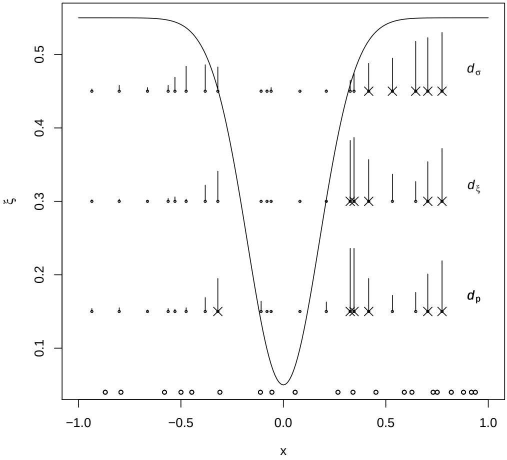


Figure 12: Case 1.

1. Extreme value theory

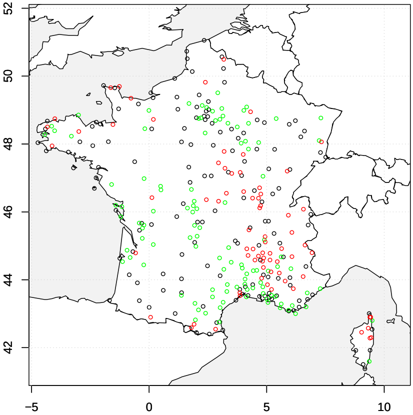
2. Kullback-Leibler divergence and extreme values
   - Problem
   - Divergence
   - Simulations
   - Real data

3. Query by Committee and optimal design
   - Motivation
   - Behaviour of the extremes
   - Query by Committee
   - Simulations
   - Real data

4. Perspectives

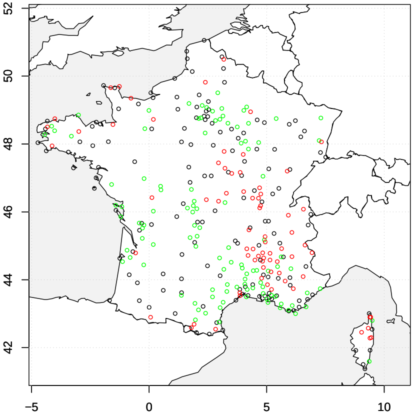
Repartition

- 147 stations in $\mathcal{X}_1$
- 110 stations in $\mathcal{X}_2$
- 74 stations in $\mathcal{X}_3$



Repartition

- 147 stations in $\mathcal{X}_1$
- 110 stations in $\mathcal{X}_2$
- 74 stations in $\mathcal{X}_3$



We add 15 stations to $\mathcal{X}_1$ at each iteration and do 4 iterations.

$\Rightarrow$ we dismiss 14 stations

Results

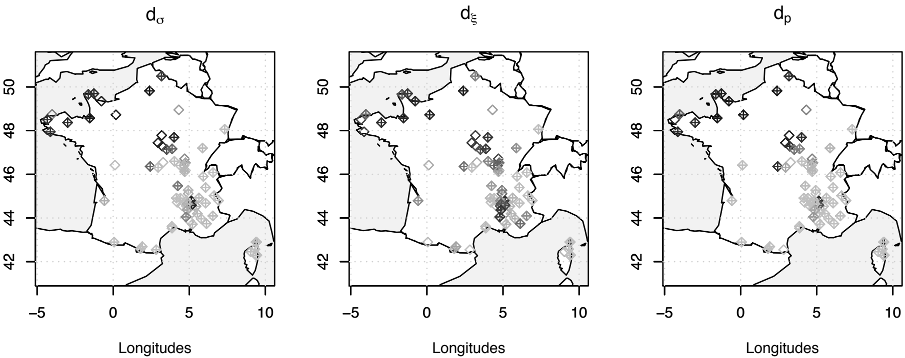


Figure 13: Frequency of suppression of each station for each cost function.

Results

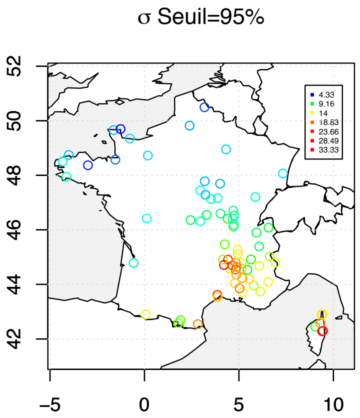


Figure 13: Initial estimations of the parameters at the stations of $\mathcal{X}_3$.

Part 1

- Asymptotic normality
- Direction of the evolution of the behaviour of the extremes
- Work with more stations (spatial structure ?)

Part 2

- Develop a more solid theory to justify our approach
- Apply our method with other quantities (wind gusts, temperatures...)
- Adapt the method to the construction of stations