Extreme Value Theory for observations

Théophile Caby, Davide Faranda, Sandro Vaienti, Pascal Yiou
Rencontre distribution des événements extrêmes application á la climatologie

28 novembre 2019

Motivations

- Consider a dynamical system (M, T, μ) and a smooth observable $f : M \to \mathbb{R}^k$.
- We want to describe the statistics of recurrence of f near a particular value f_0 when computed along a trajectory of the system.
- Methods based on recurrence properties of the system are used in climate to compute local dimensions and inverse persistence. They necessitate to work with trajectories of the original system, while physicists have often only access a lower dimensional representation of it through measurements.
- Can we still recover informations on the *real* underlying attractor using observational data ?

Previous results

- Rousseau and Saussol (2010) showed that for a large class of chaotic systems, the recurrence rate for the observations defined as

$$R_f(z) = \lim_{r \to 0} \frac{\log \inf\{k \in \mathbb{N}^* : f(T^k z) \in B(f(z), r)\}}{\log r}$$

  and the local dimension of the image measure at the point f_0 are equal.
- They were able to compute this quantity explicitly for absolutely continuous measures and C^∞ observables.
- We place ourselves in an EVT context, which allows a full asymptotic description of the statistics of recurrence, but also of the statistics of the number of visits of the observable near f_0.

The formal approach

- We compute along a trajectory of the system the function

$$\phi_f(x) = -\log |f(x) - f_0|.$$

- ϕ_f takes high values whenever $f(x)$ is close to f_0 : by studying the extreme values taken by ϕ_f along a trajectory, one can have access to some local properties of the image of the system near f_0.
- For example, one can consider the variable

$$M_n(x) = \max\{\phi_f(x), \cdots, \phi_f(T^{n-1}(x))\}.$$

- $\mu(M_n < u_n)$ gives the the probability that the observable has not entered the ball $B(f_0, e^{-u_n})$ after n iterations of the system (its hitting time statistics).
- Computing $\mu(M_n < s)$ is a standard problem of EVT.

Definitions

- *Image measure* of a set $A \subset f(M)$:

$$f_*\mu(A) = \mu(f^{-1}(A)).$$

- *Local dimension* of $f_*\mu$ at f_0 :

$$d_f(f_0) = \lim_{r\to 0} \frac{\log f_*\mu(B(f_0, r))}{\log r}.$$

Main result

- It is very easy to see that $\mu(\phi_f > u_n) \underset{n\to\infty}{\sim} e^{-u_n d_f(f_0)}$.
- Applying the machinery of EVT, we find that for highly chaotic systems,

$$\mu(M_n < u_n) \underset{n\to\infty}{\sim} 1 - \exp(-\theta_f \exp(d_f(f_0)u_n - \log n)).$$

- The extremal index θ_f may be different from 1, in case of clustering of the extreme events.

Values of θ_f and $d_f(f_0)$?

- When applying EVT based methods used in climate, we actually compute some local quantities from observational data : θ_f and $d_f(f_0)$.
- We want to know if these quantities can provide informations on the underlying attractor.
- More generally, we want to know the value of these quantities since they allow to predict the hitting time statistics of the observable in the neighborhood of f_0.

The extremal index θ_f

- Let us consider the ball $B_r := B(f_0, r)$.
- One can show with the spectral approach that if the limits defining the following quantities exist :

$$q_k := \lim_{r \to 0} \frac{f_*\mu(B_r \cap T^{-1}\overline{B}_r \cap \ldots \cap T^{-k}\overline{B}_r \cap T^{-k-1}B_r)}{f_*\mu(B_r)},$$

for $k \geq 1$ and

$$q_0 := \lim_{r \to 0} \frac{f_*\mu(B_r \cap T^{-1}B_r)}{f_*\mu(B_r)},$$

the EI is given by

$$\theta_f = 1 - \sum_{k=0}^{\infty} q_k.$$

The extremal index θ_f

- In the standard EVT problem ($f = Id$), $q_k = 0$ only when z is (k+1)-periodic.
- When f is continuous, we expect that if z is k-periodic, the quantity q_{k-1} is not 0, so $\theta \neq 1$.
- There is another kind of clustering that can appear, when there are points $z' \in X$ such that $f(z) = f(z')$ and $f(T^k z) = z'$, for $k \in \mathbb{N}$. In this situation, the value of q_k could be positive and $\theta_f < 1$.
- For generic observables, this is very unlikely to happen, and we expect that in physical application, $\theta = 1$ when z is not periodic.

Numerical computation of θ_f

- Because several values of q_k may be different from 1, we introduced an estimate of θ that computes each value of q_k up to an order m :

$$\hat{q}_k = \frac{\sum_{i=0}^{N-2-m} \mathbf{1}(\phi(T^i x_0) \geq u \cap \max_{l=1,\ldots,m} \phi(T^{l+i} x_0) < u \cap \phi(T^{i+k+1} x_0) \geq}{\sum_{i=0}^{N-1} \mathbf{1}(\phi(T^i x_0) \geq u)/N}$$

- We then compute $\hat{\theta}_m = 1 - \sum_{i=0}^{m} \hat{q}_k$.
- In the standard problem, the values of q_k can carry some information on the periodicity of the target point. This estimate allows to characterize periodic patterns of the dynamics

An interesting example

- In the standard problem, there are some (partially stochastic) systems for which the EI is well defined and none of the q_k in the spectral formula is 0.
- Example : take two maps $f_0 = 2x - \mathrm{mod} 1$ and $f_1 = 2x + b - \mathrm{mod} 1$, $0 < b < 1$ and apply successively either f_0 or f_1 with probablity $1/2$. Then for the target point 0,

$$q_k = 1/4^{k+1}.$$

$$\theta = \sum_{k=0}^{\infty} q_k = 2/3.$$

Values of $d_f(f_0)$ for Lebesgue-a.c. measures

- Rousseau Saussol (2010) : If μ is a.c. with respect to Lebesgue on $\mathbb{R}^l$ and f is $C^\infty(\mathbb{R}^l, \mathbb{R}^k)$, then $d_f(f(z))$ is integer valued and is equal to the rank of $Df(z)$ μ-a.e..
- For most observables of physical interest, it is equal to $\min(l, k)$ almost everywhere.
- We give an example where $d_f(f(z))$ is not integer : consider in S^1 the map $2x \mod 1$ ($\mu = Leb$) and the observable $f(x) = x^a$. In this case,

$$d_f(f(0)) = 1/a.$$

Values of $d_f(f_0)$ for fractal measures

- Many physical systems have (multi)fractal natural measures. It is believed to be the case in climate (we proposed recently that the wide distribution of local dimensions found in EVT methods orginates from the multifractal structure of the atmospheric attractor).
- The simplest case is when f a diffeormorphism : the local dimensions are then preserved.
- Observables used in physics are not diffeomorphisms... It is more appropriate to see them as lower dimensional projections of the original system.

Hunt-Kaloshin theorem (1998)

- Hunt-Kaloshin (1998) : For a *prevalent* set of C^1 observables $f : \mathbb{R}^l \to \mathbb{R}^k$ and $\mu - a.e.\ z$,

$$d_f(f(z)) = \min(k, d(z)).$$

- The notion of prevalence is analogue to the one of almost everywhere for infinite dimensional spaces : when considering a generic observable, the theorem should hold.
- We propose some illustrations and discuss some important implications of this result.

A toy model of fractal measure : the baker's map

- The baker's map is defined in the unit square by

$$x_{n+1} = \begin{cases} \lambda_a x_n, \;\; y_n < \alpha, \\ (1-\lambda_b) + \lambda_b x_n, y_n > \alpha \end{cases}$$

$$y_{n+1} = \begin{cases} \dfrac{y_n}{\alpha}, y_n < \alpha, \\ \dfrac{y_n - \alpha}{1-\alpha}, y_n > \alpha, \end{cases}$$

  where $\alpha \in (0, 1/2]$, $\lambda_a + \lambda_b \leq 1$.
- The invariant measure has a multifractal structure. The value of its information dimension D_1 can be computed explicitly from the different parameters.

Structure of the image measure

- For a given observable f, we want to know the structure of the image measure around a point $f(z)$, $z \in M$.
- In the observable has values in $\mathbb{R}$, we take a point in the ball $(f(z) - r, f(z) + r)$ and look wether it has a preimage in M. If it is the case for all point of this ball, we expect that the image measure is 'full' around the point z.
- This would give a local dimension equal to 1 for the image measure.
- Geometrically, this problem is the same as looking whether the level curves $\{x \in M, f(x) = f(z) \pm \varepsilon\}$ intersect the attractor for all small values of ε.

The real observable $f(x, y) = \frac{x+y}{2}$

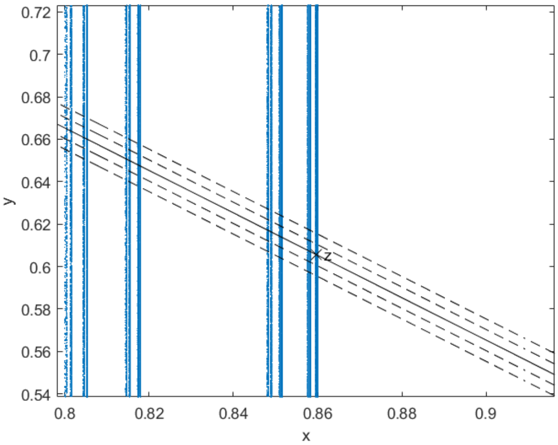

Figure – The baker's attractor is in blue, the level curves in dotted lines

For a normal multivariate gaussian observable centered at z

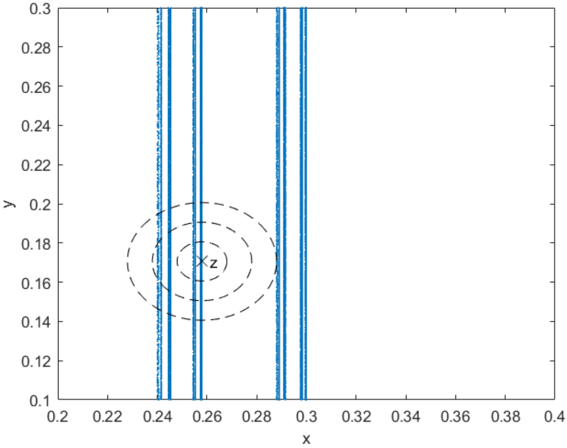

Figure – The baker's attractor is in blue, the level curves in dotted lines

Numerical computations of $d_{f_*\mu}$ and θ_f

	$\alpha = 1/5$	$\alpha = 1/4$	$\alpha = 1/3$
$f_0 = 0.1$	1.00 ± 0.01	1.00 ± 0.02	1.00 ± 0.01
$f_0 = 0.3$	1.00 ± 0.01	1.00 ± 0.01	1.00 ± 0.01
$f_0 = 0.8$	1.00 ± 0.01	1.00 ± 0.01	1.00 ± 0.02

Table – Values of $d_{f_*\mu}(f_0)$ computed for the mean value observable, for different values of α and f_0.

	$\alpha = 1/5$	$\alpha = 1/4$	$\alpha = 1/3$
$f_0 = 0.1$	1 ± 0	1 ± 0	1 ± 0
$f_0 = 0.3$	1 ± 0	1 ± 0	1 ± 0
$f_0 = 0.8$	1 ± 0	1 ± 0	1 ± 0

Table – Values of θ_f computed for the mean value observable, for different values of α and f_0. The error is the standard deviation of the results

Is it possible to find a C^1 observable which not in this prevalent set ?

- We expect to happen when the measure has a very lacunary, completely disconnected structure in which the level curves can squeeze through.
- This is the case of a product of Cantor sets. We take the gaussian observable centered at a point z.

Peculiar case

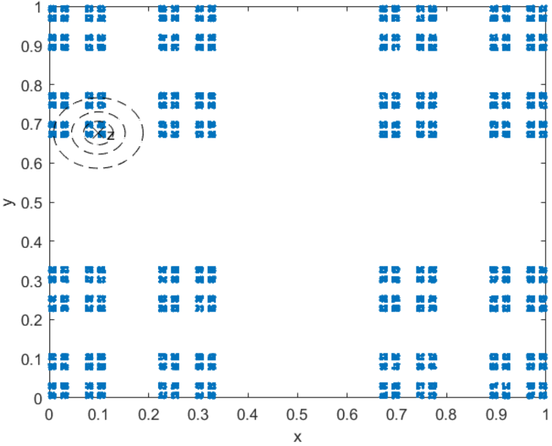

Figure – The product of Cantor is in blue, the level curves for the gaussian observable is in dotted lines.

Numerical results

z	$(0.994, 0.0029)$	$(0.6679, 0.9914)$	$(0.0861, 0.2565)$
$d_{f_*\mu}(f(z))$	0.61 ± 0.002	0.60 ± 0.002	0.62 ± 0.002

Table – Values of $d_{f_*\mu}(f(z))$ computed for the gaussian observable g, for different points z. The error is the standard deviation of the results

- We find values close to the local dimensions of the Cantor set (≈ 0.63).
- In this case, the image measure has a Cantor-like structure.
- For other generic observables, the image measure is lacunary, but not well organized. We see integer local dimensions.

Application to climate data

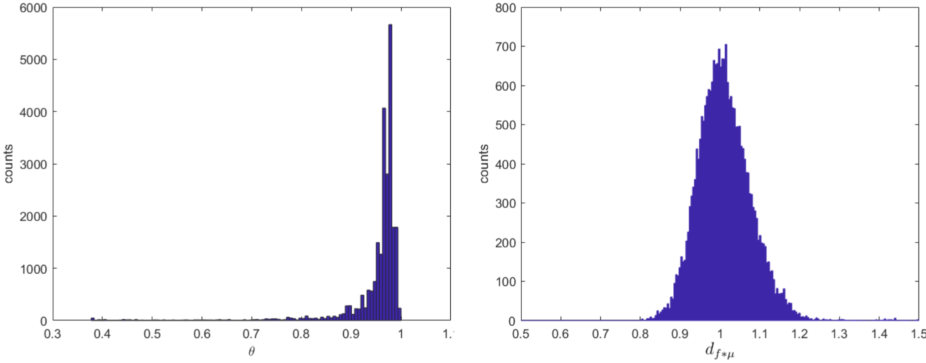

Figure – Distributions for the SLP of the EI (left) and of the $d_{f_*\mu}(f(z))$ (right) found for different points z of the attractor. We took f equal to the average value of the pressure field.

Discussions

- We might think that by taking an observable $f : \mathbb{R}^l \to \mathbb{R}^k$, with $k << l$, $d_{f_*\mu} = k$, and all information on the fine structure of the attractor is lost.
- Hunt-Kaloshin theorem says that it is enough for k to be larger than D_1, the information dimension of the underlying system to have that $d_{f_*\mu} = \min(k, D_1) = D_1$ almost everywhere.
- By measuring enough observations, we are able to recover some information on the fine structure of the real underlying system !

Distribution of local dimensions for SLP data

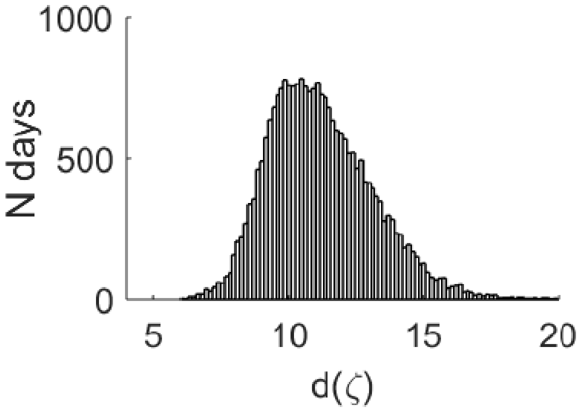

Figure – Distribution of local dimensions for SLP data (Faranda et al. 2016). The local dimensions are much lower than $k = 1060$, (on average around 12.8) suggesting that the information of the underlying attractor is ≈ 12.8.

The statistics of number of visits of the observable near a particular value

- We are interested in the law of the number of visits in the ball $B_n = B(f_0, e^{-u_n})$ up to a suitably rescaled time :

$$N_n(t) = \sum_{l=1}^{\lfloor \frac{t}{\mu(B_n)} \rfloor} \mathbf{1}_{B_n}(T^l(x)).$$

- EVT predicts that for chaotic enough systems, the law of $N_n(t)$ is given asymptotically by a compound Poisson distribution ν :

$$\mu\left(N_n(t) = k\right) \underset{n\to\infty}{\rightarrow} \nu(k)$$

Compound-Poisson distributions

- A random variable W is compound Poisson distributed if there are i.i.d. integer valued random variables $\pi_i \geq 1$ and an independent Poisson disribution distributed P such that $W = \sum_{i=1}^{P} \pi_i$.
- In the present case, π_i represents the size of a cluster and P the number of clusters appearing in the interval of time considered.

- For the standard observable ($f = Id$), the results are well known :
  - if z is non periodic, ν is a Poisson distribution of parameter t
  - if z is periodic, $N_n(t)$ follows for large n a Polya-Aeppli law of parameters θ and t :

$$\mu(N_n(t) = k) \underset{n\to\infty}{\to} e^{-\theta t} \sum_{j=1}^{k} (1-\theta)^{k-j} \theta^j \frac{(\theta t)^j}{j!} \binom{k-1}{j-1}.$$

- What happens when $f \neq Id$?

The statistics of number of visits of the observable

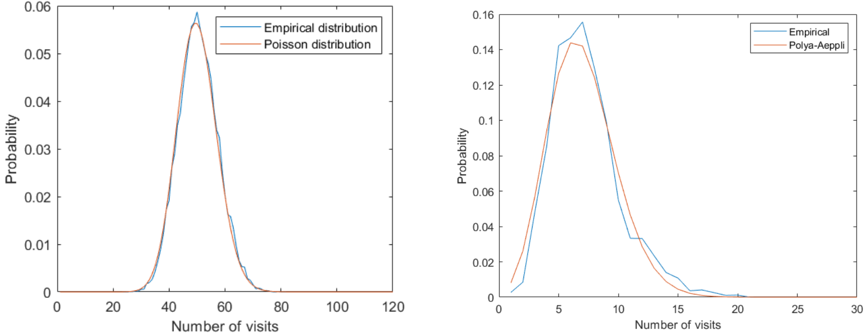

Figure – Comparison between the empirical distributions of the number of visits of the observable "mean value" in a small ball centered in $f(z)$ and different compound Poisson distributions for the Baker map (left) and the climate data (right).

Conclusions

- The hitting time statistics of observations are described by a Gumbel law whose scale parameter is related to the local dimension of the image measure.
- This allows to recover some information on the fine structure of the underlying attractor using EVT based methods, whenever the dimensionalty of the observations is large enough.
- The statistics of the number of visits of the observable near a point of interest is described asymptotically by a compound-Poisson distribution, and a pure Poisson distribution when $\theta = 1$.

Thank you !