Bayesian approach to Extreme Value Theory

How to study unusual weather events?

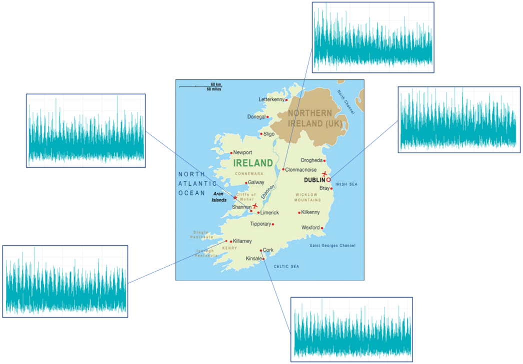
Davide Fabbrucci, Matteo Pierdomenico, Giacomo Randazzo

Politecnico di Milano

07/01/2021

A brief recap of our data

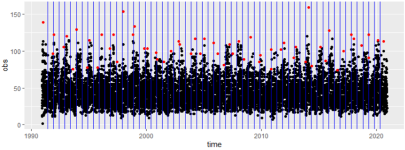
Data: $\{x_{s,t}\}_{s \in S, t \in T}$ where
$x_{s,t} = \textit{intensity of daily highest wind gust(km/h)}$,
$S := \{\textit{5 different counties of Ireland}\}$,
$T := \{\textit{30 years of daily measuraments, about 11'000 observations}\}$.

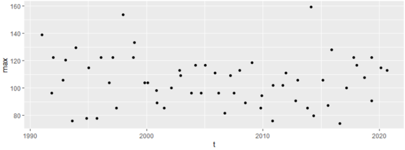


Approaches to select extreme data from time series

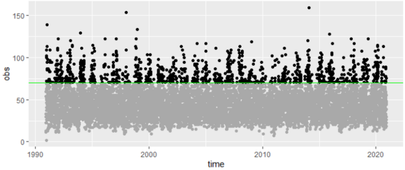
- *Block Maxima approach*:

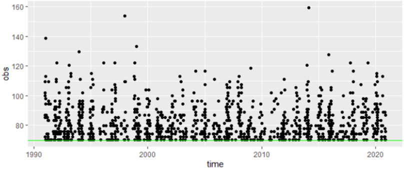
$$x_{\cdot,n}^{block-maxima} = \max_{N\cdot(n-1)<t\leq N\cdot n} x_{\cdot,t},$$ where N is the dimension of the block.





- *Peaks Over Threshold approach*: $x_{\cdot,t}^{POT} = x_{\cdot,t}\ \ s.t\ \ x_{\cdot,t} \geq threshold$





Temporal dependence assumptions

- Short-range dependence:

The plot of the time series against the version at lag 1, for each site, shows correlation between successive observations, so we have to make some assumptions:

> -for *Block Maxima* approach: taking blocks sufficiently large removes short-range dependence for extremes;
>
> -for *P.O.T.* approach: clustering the nearby peaks and select the max of each cluster removes the correletion between possible successive extremes.

- Long-range dependence:

We assumed long-range independence by using the *Leadbetter's D condition*.

A first model: **Block Maxima approach** with STAN

Fixing $s \in S$:

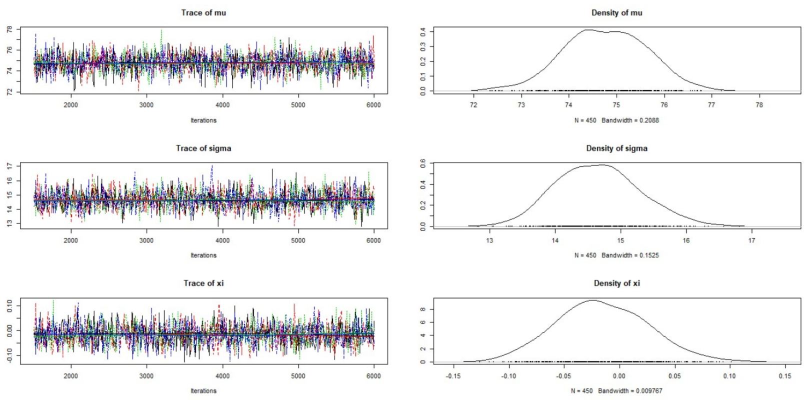
$$x_{s,\cdot}^{block-maxima} | \mu, \sigma, \xi \sim \mathcal{GEV}(\mu, \sigma, \xi),$$
$$\mu \sim \mathcal{N}(0,\ 10000),$$
$$\sigma \sim \Gamma(0.001,\ 0.001),$$
$$\xi \sim \mathcal{N}(0,\ 10),$$

where the CDF of the GEV distribution is

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

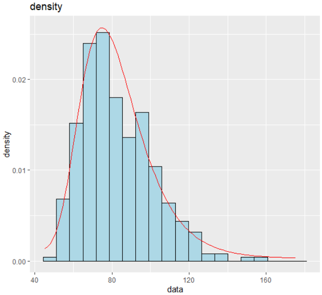
defined on the set $\{x : 1 + \xi(\frac{x-\mu}{\sigma}) > 0\}$ with $\sigma > 0$.

Results for *Clare county*

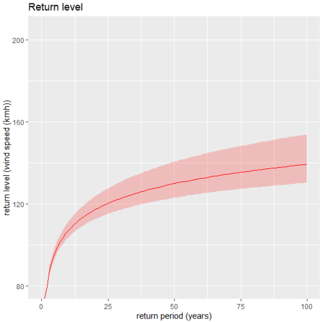


Results for *Clare county*

Parameter	Mean	SD
μ	74.77	0.88
σ	14.65	0.64
ξ	-0.02	0.04



(a) Pointwise Predictive Density



(b) Return Level

Consideration about the priors

Q: Can we extract information from our data?
Fixing a location $s \in S$, we divided our set of daily observations $\{x_{s,t}\}_{t=1}^{T}$ in two parts: $\{x_{s,t}\}_{t=1}^{T^*}$ and $\{x_{s,t}\}_{t=T^*+1}^{T}$, $T^* \approx T/2$.
We then adopted the *Block Maxima approach* model with non-informative priors for the first part of observations $\{x_{s,t}\}_{t=1}^{T^*}$ in order to generate an MCMC sample of each parameter with which we gave informations to the priors of the *Block Maxima approach* model for the second part of data $\{x_{s,t}\}_{t=T^*+1}^{T}$, i.e. :

$$
\begin{gathered}
x_{s,\cdot}^{block-maxima} | \mu, \sigma, \xi \sim GEV(\mu, \sigma, \xi), \\
\mu \sim \Gamma(3593.95,\ 48.47), \\
\sigma \sim \Gamma(240.27,\ 16.78), \\
\xi + \frac{1}{2} \sim \Gamma(68.68,\ 124.98),
\end{gathered}
$$

Results for *Clare county*: $WAIC_{non-informative} = 1561.9$
$WAIC_{informative} = 1596.9$

A first model: **Peaks over thresholds approach** with JAGS

Fixing $s \in S$, u_1 lower bound for the threshold (also used for declustering) and u_2 upper bound for the threshold:

$$x_{s,\cdot}^{POT,u_1} | \sigma, u, \xi \sim \begin{cases} \frac{1}{2} \frac{1}{(u-u_1)} & \text{if } u_1 \leq x < u \\ \frac{1}{2} h(x|\sigma, u, \xi) & \text{if } x \geq u \end{cases},$$

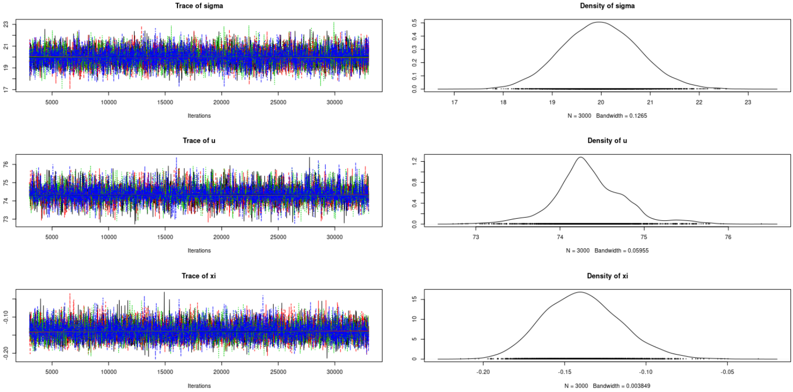
$$\sigma \sim \Gamma(0.001,\ 0.001),$$

$$u \sim \mathcal{N}(u_s^{guess},\ \sigma^2_{u_s^{guess}}), \quad u \in (u_1, u_2),$$

$$\xi \sim \mathcal{N}(0,\ 10), \quad \xi \in \left(-\frac{\sigma}{x_{max} - u}, +\infty\right),$$

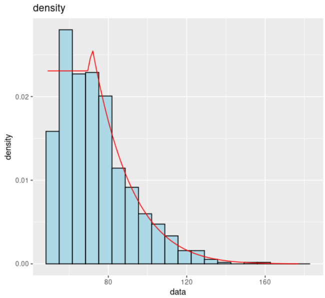
where $h(x|\sigma, u, \xi)$ is the PDF of the GPD distribution.

Results for *Clare county*

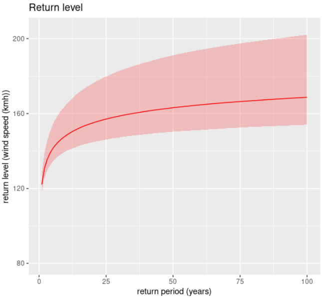


Results for *Clare county*

Parameter	Mean	SD
u	70.70	0.56
σ	18.47	1.22
ξ	-0.10	0.04



(a) Pointwise Predictive Density



(b) Return Level

A random effect model for spatiotemporal data

Problem: We have observations in five different sites and, in each site, extreme wind gust can suffer from seasonal variation. How can we model this variability in our analysis of extremes?

Solution: Fitting a site-seasonal varying *GPD*!

Idea: We partition the annual cycle in 12 "seasons" (months actually) for each site, thus our model will yeld parameters pairs

$$(u_{s,m}, \sigma_{s,m}, \xi_{s,m}), \qquad s = 1, \ldots, 5 \qquad m = 1, \ldots, 12$$

where s and m are indeces for site and season respectively.

Hierarchical Random effects model

- **Assumptions**:

  -the GPD is valid for exceedances over a high threshold for each month at each site,

  -extremes between sites and between months are independent,

  -there is no interaction between monthly and site effects,

  -both spatial effects and monthly effects are exchangeable.

So, the first layer of our model is:

$$x^{POT}_{(s,m),\cdot} | u_{s,m}, \sigma_{s,m}, \xi_{s,m} \sim \mathcal{GPD}(u_{s,m}, \sigma_{s,m}, \xi_{s,m}), \quad \text{for} \quad m = 1, \ldots, 12$$

$$s = 1, \ldots, 5.$$

Hierarchical Random effects model

Building the random effects:

$$log(\sigma_{s,m}) = \epsilon_\sigma^{(s)} + \gamma_\sigma^{(m)}$$

$$log(\xi_{s,m}) = \epsilon_\xi^{(s)} + \gamma_\xi^{(m)}$$

where we take $log(\sigma_{s,m})$ to retain the positivity of the scale parameter $\sigma_{s,m}$.
All random effects are normally and independently distributed:

$$\epsilon_.^{(s)} \sim \mathcal{N}(a_., 1/\xi_.), \quad for \quad s = 1, ..., 5 \quad the\ site\ effect,$$

$$\gamma_.^{(m)} \sim \mathcal{N}(0, 1/\tau_.), \quad for \quad m = 1, ..., 12 \quad the\ monthly\ effect.$$

The final layer is then:

$$a_. \sim \mathcal{N}(0,\ 1000000), \ \ \xi_. \sim \Gamma(0.01,\ 0.01), \ \ \tau_. \sim \Gamma(0.01,\ 0.01).$$

Objectives

Ultimating the implementation and the analysis of the Hierarchical random effect model.

Modelling temporal dependence: as seen before, in our data there is a serial correlation between successive observation in the *POT* approach

-A simple way to handle it is by means **declusterization**.

-A more difficult way, but very interesting, is to **explicitly model** it by means of a first-order Markov model.

References

- L.Fawcett,D. Walshaw. *Modelling Environmental Extremes* Short Course for the 19th Annual Conference of The International Environmetrics Society, The University of British Columbia kanagan, Kelowna, Canada (2008)
- Jacob M., Neves C., Vukadinović Greetham D. (2020) *Extreme Value Theory. In: Forecasting and Assessing Risk of Individual Electricity Peaks* Mathematics of Planet Earth. Springer, Cham.
- Behrens, C.N.; Lopes, H.F. Gamerman, D. (2004). *Bayesian analysis of extreme events with threshold estimation* Statist. Mod., 4, 227–244.
- N.A.M. Amina, M.B. Adam, A.Z. Aris *Bayesian Extreme for modeling high PM10 concentration in Johor* Procedia Environmental Sciences 30 (2015) 309 – 314
- Fawcett, L. and Walshaw, D., *A hierarchical model for extreme wind speeds.* Journal of the Royal Statistical Society: Series C (Applied Statistics), 55: 631-646 (2006).
- Behrens CN, Lopes HF, Gamerman D. *Bayesian analysis of extreme events with threshold estimation.* Statistical Modelling. 2004;4(3):227-244.
- Scarrott, Carl, and Anna MacDonald. *Scarrott, Carl, and Anna MacDonald. A review of extreme value threshold es-timation and uncertainty quantification.* REVSTAT–Statistical Journal 10.1 (2012): 33-60.

Leadbetter's $D(u_n)$ condition

Leadbetter's $D(u_n)$ condition ensures that long–range dependence is sufficiently weak so that it does not affect the asymptotics of an extreme value analysis. This condition is stated more formally in this way:

Definition

A stationary series $x_1, x_2, \ldots$ is said to satisfy the $D(u_n$ condition if, for all $i_1 < \ldots < i_p < j_1 < \ldots < j_q$ with $j_1 - i_p > l$,

$$\Big| Pr\Big\{ x_{i_1} \leq u_n, \ldots, x_{i_p} \leq u_n, x_{j_1} \leq u_n, \ldots, x_{j_q} \leq u_n \Big\} -$$

$$\mathrm{Pr}\Big\{ \mathrm{x}_{i_1} \leq u_n, \ldots, x_{i_p} \leq u_n \Big\} Pr\Big\{ x_{j_1} \leq u_n, \ldots, x_{j_q} \leq u_n \Big\} \Big| \leq \alpha(n, l),$$

where $\alpha(n, l) \to 0$ for some sequence l_n s.t. $l_n/n \to 0$ as $n \to \infty$.

$D(u_n)$ condition holds only for a specific sequence of thresholds u_n that increases with n. For such a sequence, the $D(u_n)$ condition ensures that, for sets of variables that are far enough apart, the difference in probabilities in the definition is sufficiently close to zero to have no effect on the limit laws for extremes.