Bayesian Approach to Extreme Value Theory:
How to study unusual weather events?

Davide Fabbrucci, Matteo Pierdomenico, Giacomo Randazzo

Abstract

Keywords:

---

1. Introduction

1.1. Extreme Value Theory

Due to their rarity, events like epidemies, high atmospheric pollutions or natural disasters have long captured the public's attention and often they are shrouded in mystery.

A particular branch of statistics, Extreme Value Theory (*EVT*), offers insights to their inherent scarcity and stark magnitude.

Moreover, in the few last decades, the *EVT* literature has grown considerably, with applied interest in engineering, oceanography, environment and economics, among others, and that is because it allows us extrapolate informations beyond the range of available data.

Unlike other statistical approaches, which are more concerned with the 'center' of probability distributions, Extreme Value Theory focuses on the tails of the distributions, i.e. either on maxima or minimal values, the extreme deviations from the median.

1.2. Approaches for practical extreme value analysis

Given an ordered sample of a given random variable (or Given a set of continuous observations?), there exist two approaches to select the extremes for the analysis.

The first approach is called *Block Maxima* approach and it relies in dividing the dataset into blocks sufficiently large and then, picking the maximum over each sample, deriving the block maxima series.

The second approach is called *Peak Over Threshold (POT)* and it consists in

picking a threshold and extracting, from a continuous record, all the observation exceeding that threshold.

1.3. Extreme daily high gust in Ireland (or The Dataset)

In this paper we use a Bayesian approach to develop models for daily highest wind gusts over a region of central and southern Ireland. The data adopted for the analysis consists of daily maximum wind gusts recorded by the Irish National Meteorological Service at 5 locations (Fig 1), namely we considered data recorded by 5 different station, each of them situated in a different county (*Clare, Clarke, Dublin, Kerry* and *Westmeath*). For all the sites, data are recorded over a period of 30 years, from December 1st, 1990 to November 30th, 2020, constituting 54790 observations at each site.
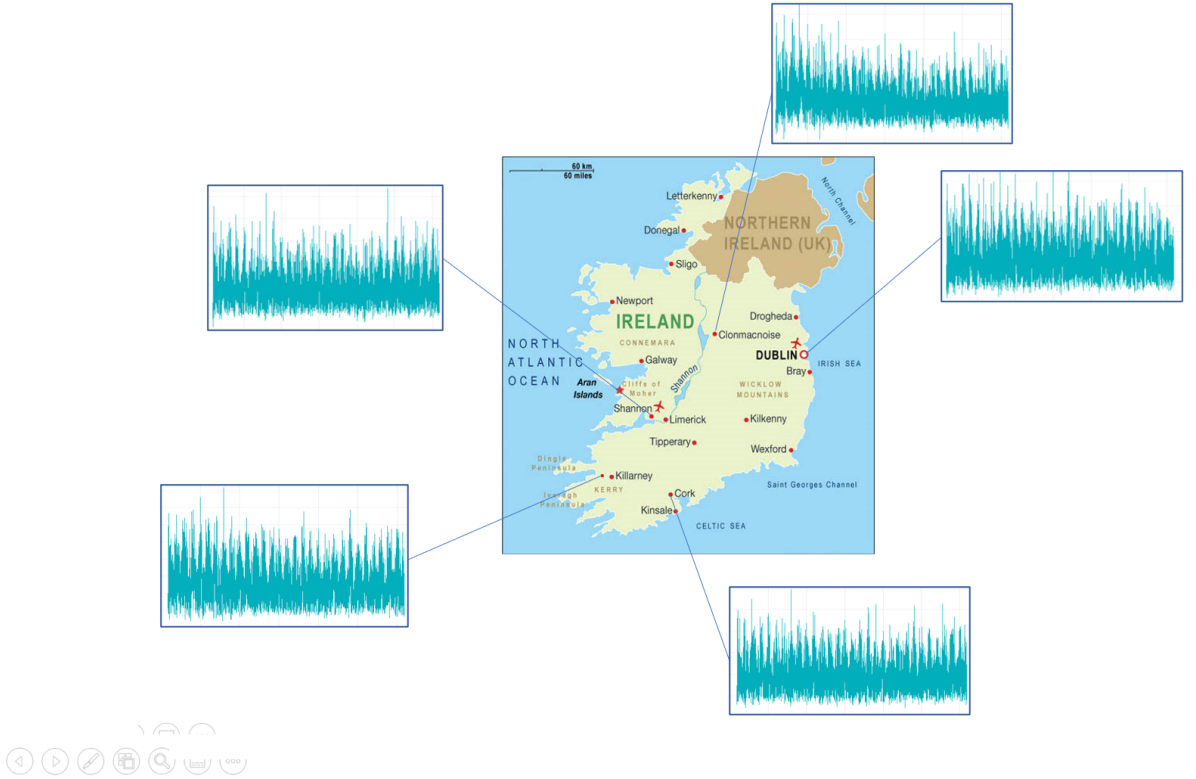

Figure 1: Map of the wind speed stations considered with the relative 30 years daily time series of maximum daily wind gust.

In Fig. 2 is illustrated an exploratory analysis of data from two contrasting sites, Kerry and Westmeath. The first is one of the most mountainous regions of Ireland, it is bounded on the west by the Atlantic Ocean and it is more often exposed to intense atmospheric phenomena; the second is a county in the center of Ireland, the territory is mostly constituted by hills and lakes and it has not boundaries with the sea. The effects of these differences are clearly seen in Fig 2, where a time series plot, a plot of the time series against the lagged series and an histogram of the observations is represented for both the counties.
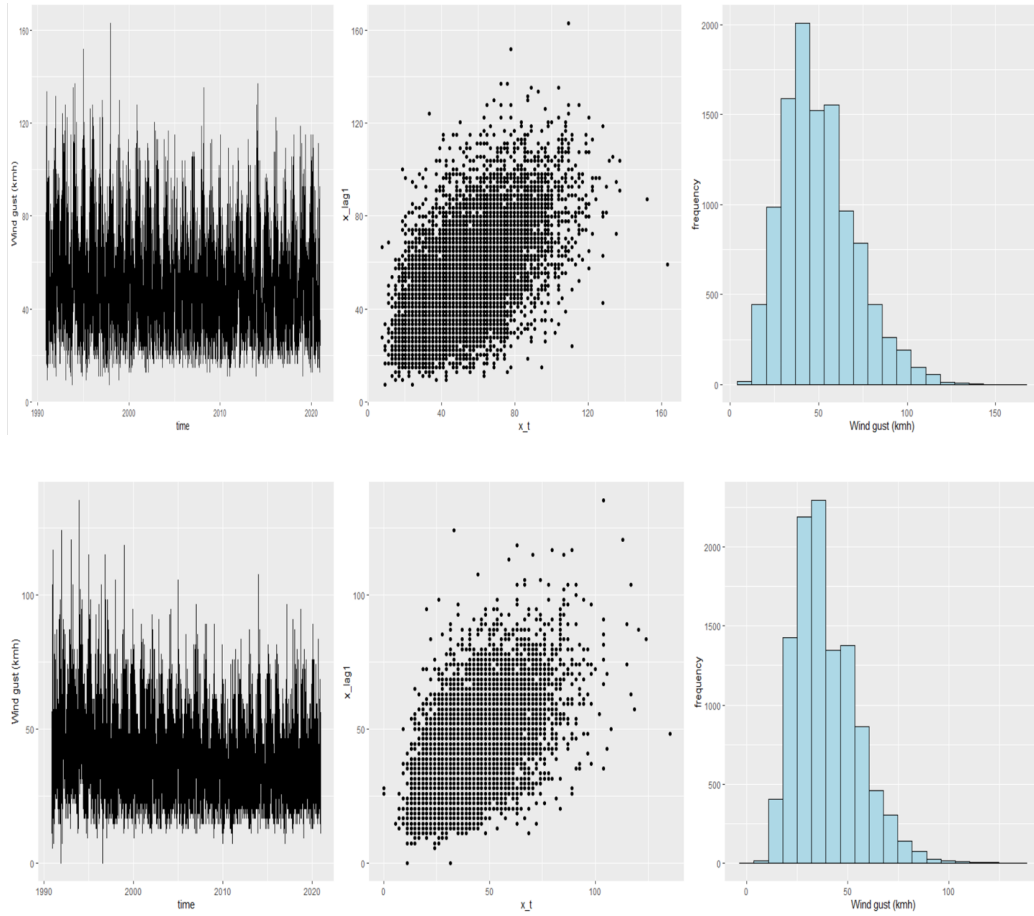

Figure 2: Time series plot, plot of the time series against the lagged series and histograms of daily maximum wind gusts in Kerry (above) and Westmeath (below).

3

1.4. The models

In the first part of this article we will focus on two simple models to analyze the EVT for daily highest wind gusts in just one single site, using respectively Block Maxima Approach and POT.

In the second part we will focus on a more complicated model which is a hierarchical random effects model that incorporates random effects for the sites, the monthly (seasonal?) variation and the serial dependence that it is ineherent in the time series of daily maximum speeds obtained at each site. As we have seen before, wind gusts can be very different in magnitude from one site to another, but it is also possible to find differences in intensity if we consider data observed in winter w.r.t data observed in summer for every single site. The aim of our hierarchical model is to use the complex structure inherent in the data to improve the "simplicistic" inferential procedure adopted in the first models which is presented in section two and trhee.

As highlighted in [7], Bayesian approach is a really good means of exploiting data on extreme events. The methods adopted to address our problems of inference are Markov Chain Monte Carlo methods, implemented by means of softwares such as STAN,JAGS and NIMBLE.

2. Block Maxima Approach Model

3. POT Model

4. Hierarchical random effects model

5. Conclusions

References

[1] 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)

[2] Jacob M., Neves C., Vukadinović Greetham D. *Extreme Value Theory. In: Forecasting and Assessing Risk of Individual Electricity Peaks* Mathematics of Planet Earth. Springer (2020), Cham.

[3] Behrens, C.N., Lopes H.F., Gamerman D. *Bayesian analysis of extreme events with threshold estimation* Statist. Mod. (2004), 4, 227–244.

[4] 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

[5] 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).

[6] 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.

[7] Coles, Stuart G., and Jonathan A. Tawn. *A Bayesian Analysis of Extreme Rainfall Data.* Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 45, no. 4, (1996), pp. 463–478.