Univariate Stochastic Modeling of Rainfall Extremes in Montpellier

Anne Bernard*[1,2], Nicolas Meyer[1,2], and Gwladys Toulemonde[1,2]

[1]Institut Montpelliérain Alexander Grothendieck – Centre National de la Recherche Scientifique, Université de Montpellier – France

[2]Inria Montpellier – Inria LEMON – France

Abstract

Flood risk is particularly high in urban areas due to soil impermeability, which prevents water absorption. Flooding can occur after periods of intense rainfall or during prolonged episodes of moderate rain. A precise understanding of such events is therefore necessary to control this risk. This is especially true in Montpellier, where heavy precipitation events frequently result in urban flooding.

In this context, the modeling of rainfall patterns plays a crucial role. Over time, statistical approaches such as the use of Exponential, Gamma, or Poisson-type distributions have been widely employed to model daily or hourly precipitation data. These methods offer flexibility in fitting a range of rainfall characteristics. While it performs well across the full range of values, it fails to accurately capture extreme events. Following extreme value theory, large values can be modeled by a Generalized Pareto Distribution (GPD), which corresponds to the limit of threshold exceedances when the threshold goes to infinity. However, in a statistical setting, this approach faces the choice of an appropriate threshold, which is a challenging issue. An extended version of the GPD (EGPD) has therefore been proposed to model the entire range of the rainfall intensities without choosing any threshold. This model is however unable to capture dry periods since it relies on a continuous distribution.

In this work, our goal is to model univariate rainfall at a high temporal resolution and by taking into account dry periods, moderate, and heavy rainfall. The data we use comes a network of 17 rain gauges that has been deployed across the Verdanson watershed in Montpellier by the Urban Observatory. This mesuring system has generated a large, high-frequency data collected at one-minute intervals since 2019. In order to properly model this high resolution, we use a discrete model based on the EGPD, called the Discrete Extended Generalized Pareto Distribution (DEGPD). While it provides some accurate results when fitting it to the data, this distribution still fails to properly capture the high frequency of zeros in our dataset, even when a zero inflation method is applied. To address this issue, we develop a hurdle model, which separates the modeling of zero rainfall from that of positive values. Among the two common strategies - truncation and shifting - we choose the shifted approach. The results of the fitting procedures on our dataset support the use of the previously proposed distribution in modeling this type of data. However, incorporating temporal dependence remains a challenge, as the current distribution is not easily extendable to time series.

Keywords: Extreme value theory, rainfall modeling, DEGPD, hurdle model, univariate modeling

*Speaker