Risk Analytics 2024 – Practical 2

Frédéric Aviolat and Juraj Bodik

Winter semester 2024-2025, HEC, UNIL

Precipitation in Lausanne

Lausanne faces significant challenges from extreme weather events. The city's vulnerability to heavy rainfall, snowstorms, and occasional heatwaves underscores the critical need for robust risk analytics and proactive mitigation strategies. Extreme weather not only impacts daily life but also poses threats to infrastructure, public safety, and environmental stability. Many buildings and infrastructure are built to sustain only certain amount of rain. For example, it is estimated that a storm delivering 150 mm of precipitation in a single day can potentially overwhelm the city's drainage systems and water infrastructure. In this practical, we will analyze precipitation data in Lausanne and estimate the probability of a storm exceeding 150 mm per day.

Figure 1: Extreme weather in Lausanne

Part 1: Block maxima approach

The dataset provided includes daily precipitation measurements in Lausanne from 1930 to 2014. Your goal is to model and analyze the precipitation extremes. In this part, we are focusing on block maxima approach.

(a) Read in the data. Draw an histogram of the daily precipitation values. Which distribution would best fit the data ?

(b) Extract the yearly maximum values. Draw their histogram. Which distribution would best fit the data ?

(c) Fit a linear model to the yearly maximum precipitation values and predict the values for the next 10 years. Provide confidence intervals for your predictions and plot it. Do you think that this a reasonable approach?

(d) Fit a GEV with constant parameters to the historical yearly max values. We recommend using `fevd` function in `extRemes` library or `gev.fit` function in `ismev` library. Fit a second GEV model with time varying location parameter. Compare the two models using AIC or BIC. Which one do you recommend using?

(e) Draw diagnostic plots of your GEV fit (for example, using `gev.diag` function). Is it a good fit?

(f) Using the model chosen in the previous parts, predict the 10-year return level. Draw your predictions of the 10-year return levels together with your data.

(g) Broadly speaking, each year, there is a chance of 1/10 that the observed value is above the 10-year return level. Comment on the results for both the linear model prediction (from c) and the GEV approach (from f). How many historical values were above this 10-year return level? Answer the same question with 20, 50 and 85-year return level.

(h) Using the fitted model, compute the return period of 100 mm of precipitation.

(i) Using the fitted model, compute the probability that there will be a day in the next year when the precipitation exceeds 150 mm.

Part 2: Peaks-over-threshold approach

In this part, we will use the Peaks-Over-Threshold (POT) approach to analyze extreme precipitation events in Lausanne.

(a) Display a time series plot of the daily precipitation across the data range.

(b) We want to model the high precipitation levels using the POT approach. First step is choosing a threshold. Draw Mean Residual Life Plot (for example using `mrlplot` in `POT` library) for the full range of your data. Choose a reasonable threshold. In the plot from part a) highlight the data that exceeds this threshold.

(c) Fit a GPD for the data exceeding the threshold and draw a diagnostic plot. Is it a reasonable fit? (Hint: if not, you may reconsider the choice of the threshold)

(d) Using the fitted model, compute the 10-year, 20-year, 50-year and 85-year return levels.

(e) Using the fitted model, compute the return period of 100 mm of precipitation.

(f) Using the fitted model, compute the probability that there will be a day in the next year when the precipitation exceeds 150 mm.

(g) Compare the results with the block maxima method. Explain the drawbacks and advantages of using the POT approach compared to the block maxima method. Which method do you prefer?

Part 3: Clustering and Seasonal Variations

In this part, we will handle extreme events occurring in clusters during specific seasons.

(a) Upload the Geneva temperature data. Plot the data. Subset the data for the summer months (June to September).

(b) Compute the extremal index of the subsetted series with appropriatelly chosen threshold (for example, you can use `extremalindex` function in `extRemes` package). Do the extremes occur in clusters? What is the probability that if the temperature today is extreme (above the chosen threshold) then tomorrow will be also extreme?

(c) Decluster the data using a suitable threshold. Plot the resulting declustered data. (Hint: you may want to use `decluster` function in the `extRemes` package.)

(d) Fit a Generalized Pareto Distribution (GPD) to the data, both raw and declustered. Compare the models and compute 10-year return level.