Simulation of extreme functionals in meteoceanic data

Nathan Gorse*[1], Olivier Roustant[2], Jérémy Rohmer[3], and Déborah Idier[3]

[1]Institut de Mathématiques de Toulouse – Institut National des Sciences Appliquées (INSA) - Toulouse – France

[2]Institut National des Sciences Appliquées de Toulouse (INSA) – Institut National des Sciences Appliquées - Toulouse – France

[3]Bureau de Recherches Géologiques et Minières – Bureau de Recherches Géologiques et Minières (BRGM) – France

Abstract

The role of meteoceanic conditions is crucial in coastal flooding. Our objective is to simulate new observations that have the same behaviour as the observed conditions but are extrapolated towards high values. The motivation is to enhance the modelling of coastal flooding by a design of experiments (Rohmer et al. 2022) where the inputs of such numerical models are time series. We work specifically on the site of Gâvres in French Brittany and use a database based on the paper of (Idier et al. 2020). This small town being located in a macro-tidal area, we focus on meteoceanic conditions occurring (+/-)3h around the high tide (with a fixed time step of 10 minutes). Thus, we are interested in the evolution of forcing conditions over tidal cycles and our database consists of time series of length 37 at different events. We use the notation XtM to describe the value obtained at time t for the Mth tidal cycle.
Our observations do not meet the standard assumptions of independence and regular variations. The first step of our method thus consists in a "whitening" pre-processing of the data. We focus on detrended winter time series X˜tM. Then, we impose a minimal duration between each event and introduce an autoregressive model with residuals ϵtM to account for the temporal dependence between tidal cycles while preserving the dependence within each one of them. In a second step, we show how to combine the approach of (Opitz, Allard, and Mariethoz 2021) and (Clémen¸con, Huet, and Sabourin 2024) to simulate new extreme residuals. This involves a marginal transformation T to meet the assumption of regular variations. Finally, we apply the reverse transformations and show how to simulate extreme time series, verifying (T(ϵM))> u where is the L2 norm. This step depends on an initial time series X˜M–, that can be chosen to tune the desired level of extremes.

We apply our method to the surge data and generate simulations of extreme time series. Since we use autoregressive models and we aim to obtain realistic extreme simulations, we use the level of (ϵsim) to draw X˜M–. We assess the quality of our method by checking that simulations and extreme observations share common behaviours. First of all, they should have shape similarities. As the simulations extrapolate the L2 norm of the observations, we apply PCA on the normalized data X˜M/(X˜M). The simulated time series generally behave the same way as the extremes observations as their respective coordinates are close.

*Speaker

Then, we apply two-sample classification tests (Lopez-Paz and Oquab 2016; Watson et al. 2023), to assess whether a classifier can distinguish between the observed and simulated time series. Given we have n = 259 extreme time series in the database and we simulate nsim = 2, 000 extreme time series, we estimate a confidence interval for the rate of correct prediction. Since the value 50% often lies within these intervals for several classifiers, our simulations are consistent with the observations.

Keywords: Extreme values, Pareto processes, Regular variations, Autoregressive models