Exploring the stochastic uncertainty of Weather Generators' extreme estimates in different practical available information scenarios.

Carles Beneyto[a], José Ángel Aranda[a] and Félix Francés[a]

[a] Research Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, Camino de Vera s/n, E-46022 Valencia, Spain

Stochastic Weather Generators (WG) have been extensively used in recent years for hydrologic modeling and climate downscaling, among others. Synthetic simulations are based on the main statistics of the observed weather, thus, relying upon the length and spatial distribution of the input data series. In most cases, and especially in arid/semiarid regions, these are scarce, which makes it difficult for WGs to obtain reliable quantile (X_T) estimates, particularly those associated with low-frequency events. Different approaches have been recently adopted to modeling high precipitation amount, mainly related with its distribution being heavy-tailed or the integration of additional information in the parametrization of the WG. The present study aims to quantify the uncertainty of the higher precipitation X_T estimates generated by a WG for different practical available information scenarios.

A one-rain gauge synthetic case study was implemented. The WG used for the experiment was GWEX, which includes a three-parameter (σ, κ, and ξ) cumulative distribution function (E-GPD) to model precipitation amounts, being the shape parameter ξ the one directly governing the upper tail of the distribution function. Four information scenarios were considered:

0. No additional information and ξ is fixed to its default value
1. No additional information and ξ is estimated from the sample
2. There exists a "perfect" regional study of maximum daily precipitation (RSMDP) where the distribution function is different from the E-GPD. Here, ξ is estimated with different estimated X_T from the population
3. There exists a "perfect" RSMDP where the values are fitted to an E-GPD and ξ is set to the population value.

A synthetic "base" population was created from a 66-years sample recorded in the Spanish Mediterranean coast (i.e. semiarid climate). All parameters defining this population were estimated and a ξ value of 0.11 was assigned for all months of the year for the sake of simplicity.

The methodology consists of a Monte Carlo simulation with packages of 50 x 60-year samples, estimating the parameters with GWEX for each and calculating the simulated X_T. The Relative Root Mean Square Error (*RRMSE*), Relative Bias (*RB*) and the Coefficient of Variation (*CV*) were calculated and analyzed for each package for both the simulated X_T and for the ξ parameter.

For each sample, all WG parameters related with the precipitation occurrence, amount and its temporal correlation were firstly fitted from the sample records for then, according to the information scenarios, modifying the shape parameter ξ.

Additionally, a sensitive analysis for the sample length and for the population ξ value were conducted.

Results showed that, as expected, integrating a RSMDP reduces considerably the uncertainty of the simulated highest X_T. It was also demonstrated that knowing the X_T from the RSMDP is more informative (i.e. less uncertainty) than knowing the ξ value. Furthermore, the lower the comparison X_T is used, the less informative it is. From the sensitivity analysis, it was concluded that longer sample sizes (up to current available lengths) do not significantly improve extreme X_T estimates, and the more extreme de climate is (i.e. higher ξ), the higher the uncertainty of the highest X_T are.

Keywords: weather generator; uncertainty; regional study of maximum daily precipitation (RSMDP); E-GPD.