Geophysical Research Abstracts
Vol. 16, EGU2014-7321, 2014
EGU General Assembly 2014

Reconstructing missing information on precipitation datasets: impact of tails on adopted statistical distributions.

Daniele Pedretti and Roger Daniel Beckie

Earth, Ocean and Atmospheric Sciences, University of British Columbia, 2207 Main Mall, Vancouver, BC V6T1Z4 , Canada (dpedretti@eos.ubc.ca)

Missing data in hydrological time-series databases are ubiquitous in practical applications, yet it is of fundamental importance to make educated decisions in problems involving exhaustive time-series knowledge. This includes precipitation datasets, since recording or human failures can produce gaps in these time series. For some applications, directly involving the ratio between precipitation and some other quantity, lack of complete information can result in poor understanding of basic physical and chemical dynamics involving precipitated water. For instance, the ratio between precipitation (recharge) and outflow rates at a discharge point of an aquifer (e.g. rivers, pumping wells, lysimeters) can be used to obtain aquifer parameters and thus to constrain model-based predictions. We tested a suite of methodologies to reconstruct missing information in rainfall datasets. The goal was to obtain a suitable and versatile method to reduce the errors given by the lack of data in specific time windows. Our analyses included both a classical chronologically-pairing approach between rainfall stations and a probability-based approached, which accounted for the probability of exceedence of rain depths measured at two or multiple stations. Our analyses proved that it is not clear a priori which method delivers the best methodology. Rather, this selection should be based considering the specific statistical properties of the rainfall dataset. In this presentation, our emphasis is to discuss the effects of a few typical parametric distributions used to model the behavior of rainfall. Specifically, we analyzed the role of distributional "tails", which have an important control on the occurrence of extreme rainfall events. The latter strongly affect several hydrological applications, including recharge-discharge relationships. The heavy-tailed distributions we considered were parametric Log-Normal, Generalized Pareto, Generalized Extreme and Gamma distributions. The methods were first tested on synthetic examples, to have a complete control of the impact of several variables such as minimum amount of data required to obtain reliable statistical distributions from the selected parametric functions. Then, we applied the methodology to precipitation datasets collected in the Vancouver area and on a mining site in Peru.