Geophysical Research Abstracts
Vol. 17, EGU2015-7437-1, 2015
EGU General Assembly 2015

Comparison of Threshold Detection Methods for the Generalized Pareto Distribution (GPD): Application to the NOAA-NCDC Daily Rainfall Dataset

Roberto Deidda (1), Antonis Mamalakis (2), and Andreas Langousis (2)

(1) Dipartimento di Ingegneria Civile, Ambientale e Architettura, University of Cagliari, Italy (rdeidda@unica.it), (2) Department of Civil Engineering, University of Patras, Greece (andlag@alum.mit.edu)

One of the most crucial issues in statistical hydrology is the estimation of extreme rainfall from data. To that extent, based on asymptotic arguments from Extreme Excess (EE) theory, several studies have focused on developing new, or improving existing methods to fit a Generalized Pareto Distribution (GPD) model to rainfall excesses above a properly selected threshold u. The latter is generally determined using various approaches that can be grouped into three basic classes: a) non-parametric methods that locate the changing point between extreme and non-extreme regions of the data, b) graphical methods where one studies the dependence of the GPD parameters (or related metrics) to the threshold level u, and c) Goodness of Fit (GoF) metrics that, for a certain level of significance, locate the lowest threshold u that a GPD model is applicable.

In this work, we review representative methods for GPD threshold detection, discuss fundamental differences in their theoretical bases, and apply them to daily rainfall records from the NOAA-NCDC open-access database (http://www.ncdc.noaa.gov/oa/climate/ghcn-daily/). We find that non-parametric methods that locate the changing point between extreme and non-extreme regions of the data are generally not reliable, while graphical methods and GoF metrics that rely on limiting arguments for the upper distribution tail lead to unrealistically high thresholds u. The latter is expected, since one checks the validity of the limiting arguments rather than the applicability of a GPD distribution model. Better performance is demonstrated by graphical methods and GoF metrics that rely on GPD properties. Finally, we discuss the effects of data quantization (common in hydrologic applications) on the estimated thresholds.

Acknowledgments:
The research project is implemented within the framework of the Action «Supporting Postdoctoral Researchers» of the Operational Program "Education and Lifelong Learning" (Action's Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.