Geophysical Research Abstracts
Vol. 13, EGU2011-3596, 2011
EGU General Assembly 2011

Automating Peak-over-Threshold with the Information Matrix test

Sophie Fukutome (1), Mark Liniger (1), and Maria Süveges (2)

(1) Federal Office of Meteorology and Climatology MeteoSwiss, Climate Services, Zuerich, Switzerland (mark.liniger@meteoswiss.ch), (2) ISDC Data Centre for Astrophysics, Geneva Observatory, Geneva, Switzerland

MeteoSwiss provides return level estimates of meteorological parameters to the public (e.g. warning services, media, ..), governmental (hydrology, security,..) and the private sector (e.g. engineers, energy, insurances,...). It is confronted with a wide range of parameters in varying aggregations and with a large number of stations. In addition, due to the complex topography, parameters such as precipitation display great spatial heterogeneity. Thus, an automatic procedure for the statistical analysis of extremes and the corresponding uncertainties is essential.

The purpose of the present study is to automate the peak over threshold method for a large number of meteorological time series. The main goal is to develop an objective choice for the threshold and run parameter for subsequent declustering. The data used for this study is daily precipitation from the MeteoSwiss observational network covering all of Switzerland, including low-land and mountainous stations both north and south of the Alps; some stations have data for more than 100 years.

The information matrix test is a goodness-of-fit test developed to identify combinations of threshold and run parameter for which the model proposed for the estimation of the extremal index, i.e. the inverse mean cluster size, must be rejected. It is used to choose visually an appropriate combination of threshold and run parameter for subsequent extreme value analysis with Peak-over-Threshold. In our particular application, we try to mimic the visual choice quantitatively by selecting the threshold/run parameter combination from a set of points for which the model is not rejected. The selected combination is then used to decluster the time series, and the Generalized Pareto distribution (GPD) is fitted to the resulting data set.

Results show that the threshold/run parameter combination chosen is often, but not always, similar to a subjective expert choice. Comparison with the classical threshold diagnostics indicates that the choice is within the acceptable range of values. The method is applied to automatically produce return level estimates at all stations of the observational network. Return level maps of daily and 5-day precipitation, for instance, yield realistic climatologies that reflect regional characteristics of the prevailing climate. The difference of the results and their uncertainties to more conventional methods such as block maxima is presented.