Geophysical Research Abstracts
Vol. 20, EGU2018-2220, 2018
EGU General Assembly 2018
© Author(s) 2017. CC Attribution 4.0 license.

Classification of flood hydrographs with unsupervised learning in the spectral space

Henning Oppel, Benjamin Mewes, and Andreas Schumann

Ruhr-University Bochum, Hydrology, Water Resources Management and Environmental Engineering, Bochum, Germany (henning.oppel@rub.de)

Floods are hydrological events, which are caused by different meteorological- and antecedent catchment conditions. Hence, the dominant processes creating different flood waves (and its characteristics) within a data set may differ. This heterogeneity of flood origins makes explorative analysis of large data sets very complex. Important features and information might vanish in mist of the data. Such complex data sets create a higher amount of uncertainty for the development and calibration of process-based hydrological models and reduce their performance. Therefore, it is important to structure big data sets. In case of data sets consisting of flood hydrographs, the most straightforward way is to classify them into flood types.
Here we are facing some problems: How many event types exists in the data set? Which parts of flood waves define their similarity? To cope with these problems, we propose the use of an unsupervised learning strategy to perform the classification of hydrographs with flood signatures derived from their spectral decomposition. Unsupervised learning gives us the opportunity to search the data for dominant patterns rather than starting with a pre-defined number of event types. For this purpose we apply a combined cluster technique, merging the benefits from hierarchical- and k-means clustering with an unsupervised learning technique. The dominant processes causing a flood event create the underlying dominant shape, or oscillation, of an hydrograph. Processes (or catchment properties) of minor importance add a variety of details to the observed hydrograph. The superimposition of these details with the relevant process makes the identification of similarities so challenging. Here we propose the use of flow signatures drawn from a spectral decomposition by means of the Fourier Transformation. These signatures incorporate only the underlying relevant oscillation.
Our combined approach of flow signature extraction from spectral decomposition and unsupervised learning has been applied to different catchments in mid- and southern Germany. The defined clusters were discussed for relevance and validity. Statistical tests indicate that the defined clusters provide homogeneous populations necessary for reliable flood statistics. A comparison of flood types defined for different data sets, drawn from different catchments, showed that these results are independent from catchment scale and heterogeneity of the data set.