



## **Data-driven catchment classification: combining linear and non-linear approaches to improve hydrological predictions in ungauged catchments**

M. Di Prinzio, A. Castellarin, and E. Toth

University of Bologna, DICAM, Bologna, Italy (attilio.castellarin@unibo.it, +39 051 2093140)

A promising approach to catchment classification makes use of unsupervised neural networks (Self Organising Maps, SOM's), which organise input data through non-linear techniques depending on the intrinsic similarity of the data themselves. Our study considers  $\sim 300$  Italian catchments scattered nationwide, for which several descriptors of the streamflow regime and geomorphoclimatic characteristics are available. We compare a reference classification, identified by using indices of the streamflow regime as input to SOM, with four alternative classifications, which were identified on the basis of catchment descriptors that can be derived for ungauged basins. One alternative classification adopts the available catchment descriptors as input to SOM, the remaining classifications are identified by applying SOM to sets of derived variables obtained by applying Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA) to the available catchment descriptors. The comparison is performed relative to a PUB problem, that is for predicting several streamflow indices in ungauged basins. We perform an extensive cross-validation to quantify nationwide the accuracy of predictions of mean annual runoff, mean annual flood, and flood quantiles associated with given exceedance probabilities. Results of the study indicate that performing PCA and, in particular, CCA on the available set of catchment descriptors before applying SOM significantly improves the effectiveness of SOM classifications by reducing the uncertainty of hydrological predictions in ungauged sites.