Predicting Base Flow Index of ungaged catchments using Self Organizing Maps to derive similarity

Youri Rothfuss[1] and Ludovic Oudin[1]
(1) Univ. Paris6, UMR 7619 Sisyphe, youri.Rothfuss@upmc.fr, ludovic.oudin@upmc.fr

1 Introduction

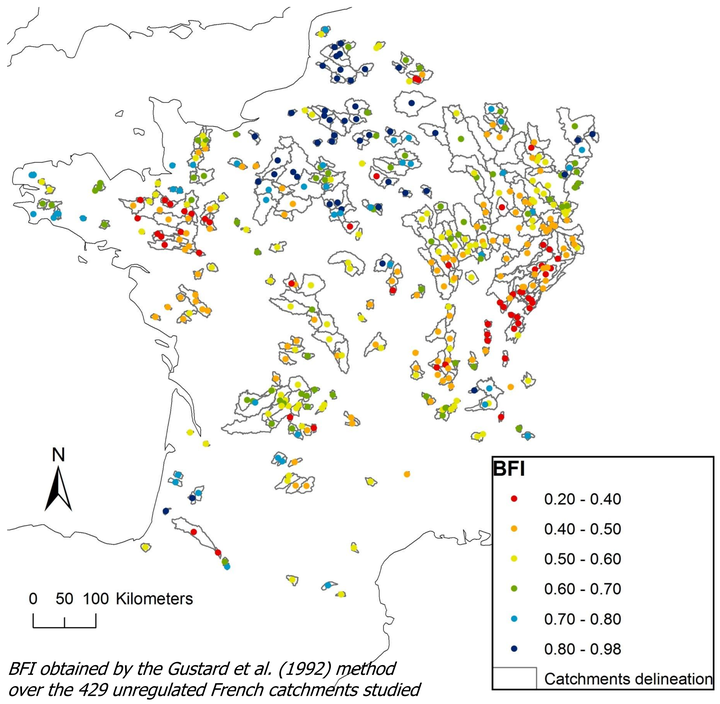
Knowledge on river low-flow characteristics such as base flow is needed for water resources management and quality assessment as low-flow is generally associated with catchment storage. Base Flow Index (BFI, Institute of Hydrology, 1980), the long-term ratio of baseflow to total streamflow, is widely used to provide information on quick- and low-flow components of a catchment. In ungaged sites, where stream flow data is not available, BFI can be predicted either by geostatistical means (i.e. by using the data from surrounding gaged catchments) or by equations, often linear, relating BFI to a number of catchment properties including topography, climate, soil and subsoil materials.

In this present study, information on soil, parent material, geomorphology, and land cover/use of over 429 French catchments were gathered. We used the Self Organizing Maps (SOM, Kohonen, 2001) in a supervised mode for predicting BFI of ungaged catchments

2 Materials

We used a database of **429** French basins located throughout France for which daily streamflow time series over the 1995–2005 period were available. **32** catchment descriptors were considered for our purpose, corresponding to four classes:

- ❖ Geomorphologic properties (e.g. catchment area, mean slope, median altitude, river network density);
- ❖ Climatic properties (e.g. PE, Rainfall, Normalized Rainfall Range);
- ❖ Geologic and lithologic properties (e.g. ESDB Hydrological Class, Parent Material Hydrological Class)
- ❖ Land use / land cover properties (CORINE Land Cover Class)



BFI obtained by the Gustard et al. (1992) method over the 429 unregulated French catchments studied

3 Methods

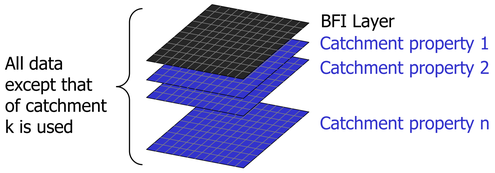
Self Organizing Map (**SOM**) is an artificial neuron network technique that can help identifying structures from multi-dimensional data sets (Kohonen, 2001; Herbst et al., 2009) by means of 2-D maps.

Our methodology follows here the jackknife procedure: each catchment in turn was considered as ungaged and data from all other gaged catchments were used to perform classification thanks to SOM.

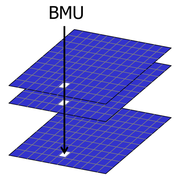
For each catchment k:

1 – Catchment data is removed from the data set

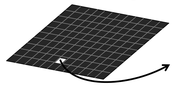
2 – Self Organized Maps (SOM) of BFI along a selection of catchment properties are constructed



3 – Best Matching Unit (BMU) for the pseudo ungaged catchment is found from the SOM n layers



4 – Affect the corresponding BFI value from the BFI layer…



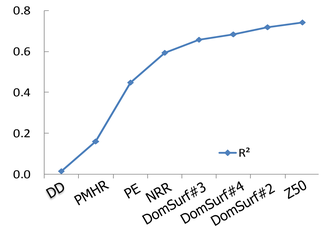
Two settings of the methodology are hereafter optimized: the SOM grid-size and the catchment properties selected to construct the SOM:

- ❖ We tested an increasing number of grid sizes from [5*5] to [30*30]
- ❖ The catchment property that increases best the final BFI prediction is selected until no improvement is observed

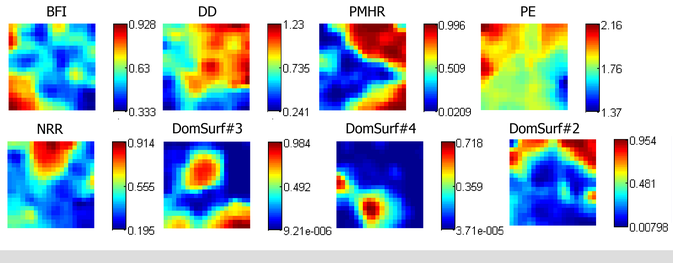
4 Optimal settings and results

Best BFI prediction was obtained when the following catchment descriptors were taken into account, by order of importance:

(1) Catchment Drainage Density (km-1)
(2) Porous Parent Material hydrological type (-)
(3) Mean Annual Potential Evaporation (mm d-1)
(4) Normalized Rainfall range (-)
(5) ESDB Dominant surface texture #3 ("medium fine") (-)
(6) ESDB Dominant surface texture #4 ("fine")
(7) ESDB Dominant surface texture #2 ("medium")
(8) Catchment Median Altitude (m)



Gain of BFI prediction when taking into account selected catchment descriptors

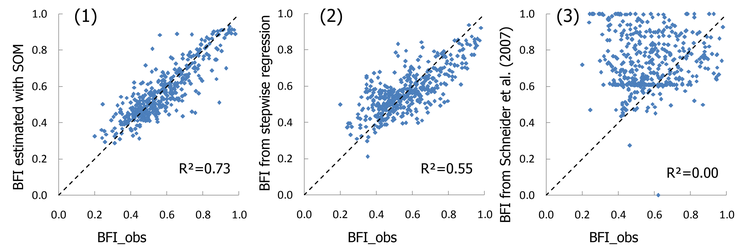


Optimal SOM results: 2-D grids of the 8 most relevant catchment descriptors along BFI

5 Comparison with benchmark methods

The performance of the method (1) is compared to two benchmark methods:

(2) Stepwise multiple linear relationship using the same possible catchment properties as for the SOM approach
(3) The algorithm proposed by Schneider et al. (2007) who used soil (ESDB) data only



BFI estimation son ungaged catchments using the proposed methodology and two other Benchmark methods

6 Discussion and conclusion

- Parameters that were selected belonged essentially to the following classes: climate (PE, NRR), geologic (parent material) and lithologic (ESDB) surface texture) properties. Vegetation information was however not selected for BFI prediction. Five variables enabled reaching nearly optimal results.
- A grid size of [20*20] representing 400 elements (i.e. close the number of studied catchments) gave the best results.
- BFI prediction was improved when using our methodology as compared to classical stepwise multiple regressions. The algorithm proposed by Schneider et al. (2007) was tested as well and showed no correlation between measurements and predictions.
- Studied catchments were mainly located in the north part of France (see map). Observed BFI for catchments from the south of France were found to have highly year-to-year variability values on a ten year period and were therefore removed from in this preliminary study. Longer periods should be used in order to compute the latter BFI values and incorporate them into the analysis.

References

- ❖ Gustard, A., Bullock, A., Dixon, J.M., 1992. Low flow estimation in the United Kingdom. Institute of Hydrology Report 108.
- ❖ Herbst M., Gupta H.V., Casper M.C., 2009. Mapping model behaviour using Self-Organizing Maps, *Hydrol.Earth Syst. Sci.*, 13, 395-409
- ❖ Kohonen T., 2001. Self-Organizing Maps, *3rd ed., Information Sciences*, Springer, Berlin, Heidelberg, New York, 501 pp
- ❖ Laaha G., Blöschl G., 2006. A comparison of low flow regionalisation methods – catchment grouping. *J. Hydrol.*, 323, 193-214
- ❖ Vogel R.M, Kroll C.N., 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics, *Water Resour. Res.*, 28, 2451-2458
- ❖ Schneider M.K., Brunner F., Hollis J.M. , Stamm C., 2007. Towards a hydrological classification of European soils: preliminary test of its predictive power for the base flow index using river discharge data, *Hydrol. Earth Syst. Sci.*, 11, 1501–1513

Acknowledgments

This research is based on a very large hydrometeorological data set, and we would like to acknowledge: Météo France for providing the SAFRAN pluviographic archive; the SCHAPI for providing access to the HYDRO streamflow archive (special thanks go to Laurent Cobos); the European Soil and INRA BDAT data bases, and Jean-louis Rosique who digitized most of the catchment boundaries.