Discrimination of different sub-basins on Tajo River based on water influence factor

R. Bermudez (1,2), A. Marcano Cerdeño (2), A.M. Tarquis (2,3), and D. Andina (2)

(1) MOPU (rafael.bermudez@upm.es), (2) Group for Automation and Soft Computing, ETSI Telecomunicación GASP/UPM (amarcano@gc.ssr.upm.es andina@upm.es), (3) ETSIA-CEIGRAM, Universidad Politecnica de Madrid, Spain (anamaria.tarquis@upm.es)

The SOM algorithm and numerical taxonomy has been applied to classify the Tajo hydrographic basin (Spain) till Portugal border. A total of 52 stations have collected data. 15 variables of water have been used in this study: flow (m3/sec), pH, electrical conductivity, concentration of several cations and anions (mgr/l) such as Cl-, SO4=, alkalinity (CO_3H-, CO_3=, OH-), Ca2+, Mg2+, Na+ and K+.

The analysis proceeds in two different lines. First, the Kohonen algorithm is trained using as input the corresponding mean each of the variables for each station. This will allow us to establish different clusters characterized in terms of the variables and other auxiliary variables before that referred to us the help we understand and explain the construction of the groups. Second, a new Kohonen network will be trained using as inputs averages of the latent variables or factors obtained after carrying out factor analysis on all available data.

Different groups were obtained by applying a Euclidean distance between stations (distance classification) and a Euclidean distance between each station and the estimated centre of gravity between them (centroid classification), varying the number of parameters and with or without standardized variable.

The partitions are performed in the entire network. Each group is identified by a number and a colour assigned to the network. The partitions the network evolved to get the best choice. Final partition is associated with a dendrogram. The averages for the variables associated with each cluster and total cluster are shown in a table.

To compare the classification of a log-log relationship is established between the number of groups created and distances, to select the best. We found that the centroid is more appropriate classification after a logic way more natural constraints that the minimum distance between stations. Standardized variable does not improve the classification, except when the centre of gravity method is applied. It taking into account the ions and the sum of them as variables, we obtain the best classification. The stations are grouped according to the electrical conductivity (EC), total anions (ta), total cations (ct) and the ratio of the ions (Na / Ca y Mg / Ca). For a given classification and comparison of different groups created some variation in ion concentration and the ratio of the ions are observed. However, the variation in each ion is different between groups as appropriate. For the latter group, regardless of classification, the increase of all ions is general.

Comparing the dendrograms, and the groups that originated the Tajo river basin can be sub dived into five distinct sub-basins by the main influence on the water:
1. With a greater influence ombrogenic (rainfed).
2. The ombrogenic and soil influence (rain and groundwater fed).
3. With pedogenic influence.
4. With lithogenic influence (geological foundation).
5. With greater and lithogenic ombrogenic additional influence.