HYDROCHEMICAL CHARACTERIZATION OF GROUNDWATER AND SURFACE WATER SUPPORTED BY MULTIVARIATE STATISTICAL ANALYSIS: A CASE STUDY IN THE PO PLAIN (N ITALY)

Marco ROTIROTI [1], Tullia BONOMI [1], Letizia FUMAGALLI [1], Chiara ZANOTTI [1], Sara TAVIANI [1], Gennaro A. STEFANIA [1], Martina PATELLI [1], Veronica NAVA [1], Valentina SOLER [1], Elisa SACCHI [2], Barbara LEONI [1]

[1] Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1-20126 Milano, Italy, marco.rotiroti@unimib.it
[2] Department of Earth and Environmental Sciences, University of Pavia, Via Ferrata 1-27100 Pavia, Italy

Multivariate statistical analysis is a useful method for supporting the interpretation of experimental data, particularly in the case of large datasets. In the present study, cluster analysis (CA) and factor analysis (FA) are used to support the hydrochemical characterization of groundwater and surface water in an area located in the Po Plain (N Italy), highly impacted by human activities related to agriculture.
The study area is located in the Oglio River basin, between the outflow from Lake Iseo and the confluence into Mella River, and covers ~1900 km^2. The northern part of the study area (higher plain) hosts a mono-layer aquifer mainly composed of sands and gravels, whereas the southern part of the area (lower plain) hosts a multi-layer aquifer constituted by a vertical alternation of sands with silty clays; the transition between higher and lower plain is marked by the so called "spring belt".
During a field survey performed in fall 2015, 58 groundwater, 20 river (Oglio River and its main tributaries), 1 Lake Iseo and 7 spring samples were collected for chemical analysis. Physico-chemical parameters, major ions, trace elements and water isotope were measured. The CA was performed on total 86 samples and 18 variables; data were autoscaled. The Ward hierarchical method, based on squared euclidean distance, was used. The FA was done using 82 samples (4 outliers were excluded) using the Kaiser criterion to select significant factors. Results of multivariate statistical analysis were combined with the geomorphological and hydrogeological knowledge of the study area in order to give a hydrogeological explanation of each data cluster.
Results led to the identification of 5 main clusters: (1) higher plain groundwater and springs, characterized by an oxidized *hydrofacies* with higher NO_3, (2) lower plain groundwater, characterized by a reduced *hydrofacies* with higher As, Fe and Mn, (3) Oglio River, (4) Oglio River tributaries and (5) outliers. Within the cluster of higher plain groundwater, three subgroups can be identified: (a) samples with the highest NO_3 and a more enriched isotopic signature attributable to recharge by local precipitation, (b) samples located around the spring belt and characterized by intermediate NO_3 concentrations (average ~50 mg/L) and (c) samples located around the Oglio River and characterized by lower NO_3. Also within the cluster of lower plain groundwater, three subgroups can be identified: (a) samples with more reduced states, (b) samples with earlier reduced states, likely due to some interactions with surface

waters and (c) samples with the highest As concentrations.
The Oglio River cluster can be subdivided into 3 subgroups: (a) the river stretch with losing behavior, (b) the river stretch with draining behavior and (c) groundwater and spring directly fed by Oglio River water.
In conclusion, this work confirms how multivariate statistical analysis can sustain the interpretation of large hydrological datasets in order to support a hydrochemical characterization. The latter will bear the development of the hydrogeological conceptual model of the area, also oriented to groundwater/surface water interactions, that, in turn, will support the numerical flow modeling of the system.

Acknowledgements
This work was supported by Fondazione Cariplo, grant 2014-1282.