Multivariate statistics to understand the geochemical processes induced by groundwater pollution. Multi-scale applying.

Julie Gesels, Claire Joniau, Jordi Batlle Aguilar, Alain Dassargues and Serge Brouyère

Hydrogeochemical data set contains many sampling points and many parameters. Usual data treatments (classical diagrams, cartography, univariate statistics) are not suitable to many parameters.

Different hydrogeochemical approaches (classical diagrams, spatial distribution maps, geochemical equations and multivariate statistics) are combined to obtain a global understanding of the hydrogeochemical processes at regional and at local scale. The combined approaches are applied to 2 examples in the Meuse Valley near Liège: the alluvial aquifer of the Meuse River (regional scale) and a contaminated site (local scale).

Many sources of contamination contribute to the deterioration of groundwater quality in a sector of the alluvial aquifer of the Meuse River: anthropogenic sources related to the urban and industrial area and natural sources, in relation to underground mine drainage. We performed a hydrogeochemical investigation using different combined approaches to determine the most contaminated sectors of the alluvial aquifer and the origin of the inorganic contaminations. Groundwater of the alluvial aquifer can be essentially characterized as calcium sulfate or calcium bicarbonate. If we look at the spatial distribution of groundwater types, we can clearly observe that some areas are highlighted. To investigate the possible occurrence of acid mine drainage, we look at the main chemical equations involved. Gypsum and calcite saturation indexes indicate changes in groundwater composition in relation with acid mine drainage. To investigate a step further the dataset, we have applied self-organizing maps, which is a non-linear multivariate statistical technique. In short, the technique allows highlighting relations between parameters based on the similarity of patterns obtained in the resulting rectangles. The non-linear multivariate statistical technique allows also separating the data set into different groups, in this case 4 groups. With the combined approaches, we obtained location of the most contaminated sectors of the alluvial aquifer and first evidence on the natural or anthropogenic contamination of strongly impacted sector.

The second example of the using of combined approaches is a brownfield in the area of Liège. Soil and groundwater are highly polluted by organic (BTEX and PAHs) and inorganic (metals, iron and sulphate) compounds. Many hydrogeochemical data are available and their spatial distributions on groundwater are complex. To investigate the possible occurrence of natural attenuation of benzene, we look at the main chemical equations involved. A non-linear multivariate statistical method is applied (Self Organizing Maps) and divides the data set into 4 clusters. The 4 clusters of piezometers are located on a map and combined with the detailed characterization of the hydraulic conductivity (K_H) and the characterization of groundwater – river interactions. Multivariate statistics allow a better understanding of the geochemical processes and their location.