



Outlier detection for groundwater data in France

Larissa VALMY (1,2), Chantal DE FOUCQUET (1), and Bernard BOURGINE (2)

(1) Geosciences center, Mines ParisTech, Fontainebleau-France (larissa.valmy@mines-paristech.fr), (2) BRGM, Orléans-France

Quality and quantity water in France are increasingly observed since the 70s. Moreover, in 2000, the EU Water Framework Directive established a framework for community action in the water policy field for the protection of inland surface waters (rivers and lakes), transitional waters (estuaries), coastal waters and groundwater. It will ensure that all aquatic ecosystems and, with regard to their water needs, terrestrial ecosystems and wetlands meet 'good status' by 2015. The Directive requires Member States to establish river basin districts and for each of these a river basin management plan. In France, monitoring programs for the water status were implemented in each basin since 2007. The data collected through these programs feed into an information system which contributes to check the compliance of water environmental legislation implementation, assess the status of water guide management actions (programs of measures) and evaluate their effectiveness, and inform the public.

Our work consists in study quality and quantity groundwater data for some basins in France. We propose a specific mathematical approach in order to detect outliers and study trends in time series. In statistic, an outlier is an observation that lies outside the overall pattern of a distribution. Usually, the presence of an outlier indicates some sort of problem, thus, it is important to detect it in order to know the cause. In fact, techniques for temporal data analysis have been developed for several decades in parallel with geostatistical methods. However compared to standard statistical methods, geostatistical analysis allows incomplete or irregular time series analysis. Otherwise, tests carried out by the BRGM showed the potential contribution of geostatistical methods for characterization of environmental data time series. Our approach is to exploit this potential through the development of specific algorithms, tests and validation of methods. We will introduce and explain our method and approach by considering the Loire Bretagne basin case.