Statistical Analysis of Ground Water Table Distribution in Alessandria Province (Piedmont – Italy)

E. Robotti[1], E. Marengo[1], M.C. Gennaro[1], A. Maiocchi[2], G. Pavese[3], A. Indaco[3]

[1]University of Eastern Piedmont, Dept. Environmental & Life Sciences, Alessandria, Italy
[2]Bracco SPA, Via Folli 14, Milan, Italy
[3]ARPA, Via Don Gasparolo 1, Alessandria, Italy

Abstract

This communication is focussed on the presentation of the results of two campaigns of samples collection for the characterisation of the ground water table distribution in the North-West area of Alessandria province (Piedmont – Italy). 44 and 66 wells were explored during the first and second campaigns respectively: in the first case 29 chemical and physical-chemical features were measured, while the content of inorganic ions and the elemental content were determined during the second campaign. The use of Principal Component Analysis together with refined mapping tools provided a clear understanding of the dynamic aspects of the subterranean waters of the investigated area. Kohonen Artificial Neural Networks were used as well to evaluate the wells similarity.

Introduction

The first dataset consisted of 44 samples coming from the investigated artesian wells in Fraschetta zone, in Alessandria district (Piemonte, North-West Italy), including 9 villages in a total area of about 50 km^2. As the main activity in the zone is agriculture, the wells were primarily used for irrigation purposes and most of them are in private houses courtyards, equipped with electrical pumping systems. The sampling was performed in all the cases by collecting the water from a tap at depths of the wells always greater than 4 meters. The sampling took place during the spring and summer of 2001. 29 chemical and chemical-physical variables were measured on each sample. The second dataset set is constituted by 66 samples coming from the investigated artesian wells in Fraschetta, collected in the spring and summer of 2005. The collection procedure was the same as in the campaign of 2001. In this case, 56 variables were determined including: pH, conductibility, 6 inorganic anions and 50 elements determined by ICP-MS.

Statistical calculations, Principal Component Analysis and Cluster Analysis were performed by Statistica 7.5 (StatSoft Inc., USA), Unscrumbler 9.5 (Camo, Norway) and Excel 2000 (Microsoft Corporate, USA); Kohonen Artificial Neural Networks were performed by a self-developed software (Visual Basic, Microsof USA).

Results

- First campaign

Principal component analysis was applied on the final matrix of 44 observations and 29 variables (autoscaled). The first 5 PCs explain the 53.14 % of the total variance contained in the dataset. The results obtained show a very low correlation between the variables, being the total variance highly partitioned over many components.

The loadings and the scores plots for the first two PCs allowed to point out some peculiar relationships between observations and variables:

1) PC_1 mainly discriminates well waters having higher conductivities (higher concentration levels of dissolved salts) from those at lower salts content enriched of free oxygen.

2) PC_2 discriminates well waters having an extra content of metal ions such as Na, Ca, Fe and Pb, from those having higher concentrations of P_2O_5 and aluminium sulphate.

Mapping techniques were then used to efficiently analyse the relationships between geographic position and samples characteristics. This was performed by using kriging mapping technique [1] representing each individual PC on a thermographic scale (figure 1). The maps present a similar hot spot in correspondence with the same zone, localised in the area occupied by wells number 48, 50, 51, 55, 56. A further geological analysis showed that this is indeed a very important zone, with a deep change of the characteristics of the ground table water composition, due to a thermal pool. With respect to this focal point PC_1 and PC_2 exhibit a similar behaviour, with a differentiation of the water composition along the North-South direction.

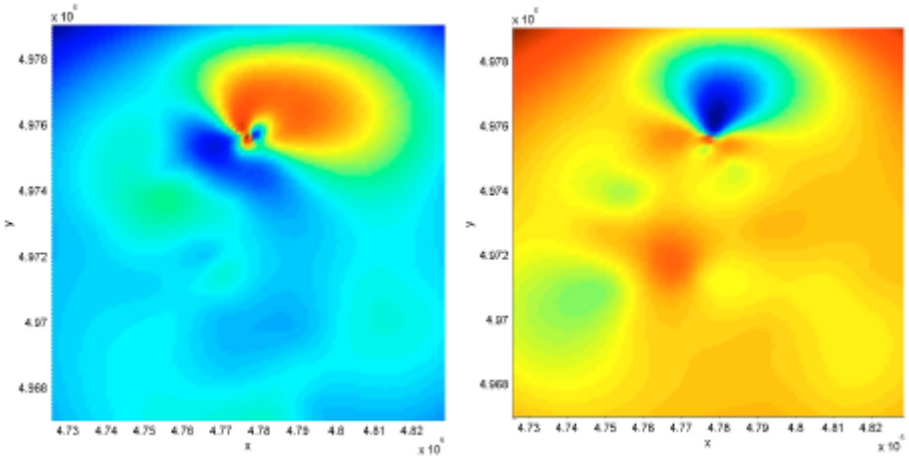

Figure 1: Thermographic maps of PC_1 and PC_2

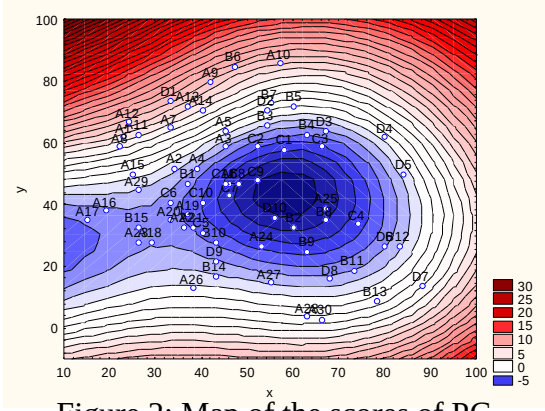

Figure 2: Map of the scores of PC_1

- **Second campaign**

The second dataset consisted of a matrix of 66 observations and 56 variables (autoscaled before PCA). The first 5 PCs explain about 45% of the total variance. Again, the results show a very low correlation between the variables. Mapping techniques were exploited to clarify the distribution of the elements and the variables in the region investigated, according to the PCs identified as significant; interesting behaviours were pointed out by this analysis, showing the presence of hot spots in the region investigated. An example is given in figure 2, were the scores of PC_1 are represented on the original map. The analysis was further developed through the use of the Kohonen ANNs, able to group the objects according to their similarity. The overall analysis allowed to point out interesting sources of information: one due to Cl, Mn, Zn and Al (possible presence of a source of anthropic pollution); a second one due to Cd and Zn, indicating another probable source of anthropic pollution; a third one due to nitrates (agricultural pollution).

References

1) D.L. Massart, B.G. Vandeginste, et al., "Chemometrics: a textbook", Elsevier: Amsterdam, 1988, ISBN: 0-444-42660-4

2) J.C. Davis, "Statistics and data analysis in geology", Wiley; 1986, ISBN: 0471172758

3) D.L. Massart, L. Kaufman, Hierarchical clustering methods., In: Elving PJ, Winefordner JD (eds) The interpretation of analytical chemical data by the use of cluster analysis. Wiley, New York, 1983, ISBN: 0471078611

4) J. Zupan, J. Gasteiger, "Neural network for chemists: an introduction", John Wiley & Sons, New York, USA, 1993, ISBN: 1560817933