# Multivariate Statistical Analyses of Groundwater Hydrochemical Data of Tirnavos Sub-basin (Central Greece)

**Vrouhakis I.[1,2,*], Pisinaras V.[2], Panagopoulos A.[2], Stamatis G.[1]**

[1]Agricultural University of Athens, Department of Natural Resources & Agricultural Engineering, Sector of Geological Sciences, Mineralogy and Geology Laboratory, Iera Odos 75, 11855 Athens

[2]Hellenic Agricultural Organisation Demeter, Soil & Water Resources Institute, Sindos 57400, Thessaloniki

*corresponding author: e-mail: i.vrouhakis@swri.gr

**Abstract**

Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) were applied in the groundwater hydrochemical parameters dataset of the Tirnavos alluvial sub basin-central Greece. PCA results suggest occurrence of four principal components both for dry and wet periods. Results of HCA application enabled to divide groundwater samples into two major groups with a large disparity in the number of samples for both periods. Implementation of these methods and spatial distribution of their corresponding results assist revealing key hydrodynamic evolution patterns and hydrochemical dependencies to anthropogenic and geogenic factors.

**Keywords:** hydrochemistry, principal components, cluster analysis, geographical distribution

## 1. Introduction

The hydrogeochemical characterization may be achieved by variable approaches, among which is the multivariate statistical techniques (e.g.Voudouris et al. 2000). Principal Components Analysis and Hierarchical Cluster Analysis have been widely used in environmental sciences and hydrochemical research (e.g. Melloul & Collin 1992, Morell 1996, Pisinaras 2008). The aim of PCA technique is to establish the relationship between variables and define the contribution of each of them, or each combination, to the structure of hydrochemical data. Hierarchical cluster technique divides groundwater samples into clusters based on their squared Euclidean distance, with the samples of the same cluster characterized by high homogeneity. The levels of the similarity at which observations are merged, are used to construct dendrograms (Lokhande et al. 2008, Tziritis et al. 2016).

## 2. Study Area

Tirnavos sub-basin forms the north-easternmost part of the eastern Thessaly plain in central Greece. It is filled by Quaternary alluvial deposits that are bounded along the SW part of the basin by Neogene marls and sandy-clay deposits. At the western margins, karstified marbles of middle-upper Cretaceous crop out. The crystalline bedrock of the Pelagonian unit composed by Mica-schists of Upper Paleozoic and Gneisses of Paleozoic age form the northern boundary of the sub-basin (Figure 1). Pinios and Titarisios Rivers flow across the sub-basin which hydrologically is part of the Pinios River Basin. The Quaternary deposits host an unconfined aquifer near the talus cones of Titarisios at NW, which towards the central parts of the basin converts to a system of a phreatic and a deeper confined aquifer, separated by a sequence of clay layers (Panagopoulos, 1996). Also, the marbles of the west margin of the basin host a karstic aquifer of great potential.

## 3. Sampling and Analysis

153 groundwater samples were collected during 4 sampling periods (2 dry and 2 wet periods between September 2016 and April 2018) from the boreholes of the study area. Totally 25 parameters were determined including major ions, trace elements and physicochemical properties, but only 15 of them (Table 1) were used in multivariate statistical analyses because the rest didn't show any significant fluctuations or didn't present a complete set of measurements. In order to produce more representative results, the calculated average values of the 2 dry and 2 wet periods were considered in the analyses.

**Table 1.** Parameters used in multivariate statistical analyses

| | |
|---|---|
| *Major Ions:* | $K^+$, $Na^+$, $Ca^{2+}$, $Mg^{2+}$, $Cl^-$, $HCO_3^-$, $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ |
| *Trace elements:* | Fe, Mn, Cu |
| *Physicochemical properties:* | pH, EC, $Hardness_{CaCO_3}$ |

## 4. Results

The results of the PCA application for both dry and wet periods are shown in Table 2, where the occurrence of four principal components for each period may be identified. The first two components for both periods are related with groundwater salinization, nitrate pollution and recharge surges. The third component is probably related with the different redox conditions of the aquifers because of the medium to high corellation of the Mn, Fe and $NH_4$ in both periods. The last component is correlated with the Cu ions contained in superphosphate fertilizers as impurity.

| DRY PERIOD | Component | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| pH | 0.275 | ***-0.763*** | -0.006 | 0.107 |
| EC | ***0.924*** | 0.305 | 0.196 | -0.026 |
| K | 0.068 | ***0.656*** | 0.097 | 0.315 |
| Na | ***0.947*** | -0.215 | 0.107 | -0.014 |
| Ca | 0.039 | ***0.901*** | 0.040 | 0.175 |
| Mg | ***0.635*** | 0.438 | 0.336 | -0.190 |
| Hardness | 0.376 | ***0.861*** | 0.213 | 0.016 |
| Cl | ***0.954*** | 0.138 | -0.050 | -0.059 |
| HCO3 | 0.172 | ***0.606*** | ***0.601*** | -0.023 |
| SO4 | ***0.959*** | -0.107 | 0.028 | 0.029 |
| NO3 | 0.054 | ***0.816*** | -0.346 | -0.077 |
| NH4 | -0.073 | 0.133 | ***0.917*** | -0.060 |
| Cu | -0.101 | 0.097 | -0.063 | ***0.921*** |
| Fe | 0.284 | 0.074 | ***0.736*** | 0.253 |
| Mn | 0.076 | -0.243 | ***0.673*** | -0.209 |

| WET PERIOD | Component | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| pH | ***-0.777*** | -0.111 | 0.052 | 0.143 |
| EC | ***0.637*** | ***0.723*** | 0.048 | 0.190 |
| K | ***0.629*** | 0.050 | -0.080 | -0.026 |
| Na | -0.224 | ***0.727*** | 0.354 | 0.450 |
| Ca | ***0.894*** | 0.044 | -0.296 | -0.082 |
| Mg | 0.509 | ***0.755*** | 0.164 | -0.009 |
| Hardness | ***0.873*** | 0.400 | -0.127 | -0.062 |
| Cl | 0.455 | ***0.774*** | -0.095 | 0.053 |
| HCO3 | ***0.783*** | -0.142 | 0.365 | 0.311 |
| SO4 | -0.054 | ***0.946*** | -0.030 | 0.047 |
| NO3 | ***0.788*** | 0.184 | -0.186 | -0.084 |
| NH4 | -0.161 | 0.208 | ***0.633*** | 0.596 |
| Cu | 0.034 | -0.103 | 0.207 | ***-0.666*** |
| Fe | -0.089 | -0.047 | ***0.810*** | -0.311 |
| Mn | -0.128 | 0.090 | ***0.840*** | 0.016 |

**Table 2.** Results of PCA application

The application of HCA divides groundwater samples into two major groups with a large disparity in the number of samples for both periods (Figure 2). In both cases, the smaller group spatially identifies regions of high anthropogenic pollution with elevated concentrations of specific parameters, and more specifically $NO_3^-$. The larger group was further divided into two subgroups with a specific spatial distribution within the sub-basin. In the wet period, the first subgroup (1.1) includes 31 boreholes and extends in the central and western parts of the basin; the second subgroup (1.2), includes only 4 boreholes (three of which are located out of the main study area) and extends in a small area to the eastern part of the basin. The first subgroup (1.1) also in the dry period extends in the central and western parts of the basin and includes 21 boreholes. Whereas the second subgroup (1.2) includes 12 boreholes and mostly extends on a larger area at the central and eastern part of the basin. .

This separation into groups resulting from the application of the HCA is following the general hydrochemical classification of the area according to which, in the western part of the basin and towards its center, fresh water dominates, while progressively towards the eastern extent of the system the hydrochemical type of water results from mixing and / or ion exchange processes. Group 2 zone, on the other hand, is attributed to the observed shallow groundwater levels especially along the western and northwestern margins of the basin, the local geological setting and the intensified agriculture, which altogether justify the documented elevated $NO_3^-$ concentrations; this is especially, so at the southern part of the study area, where extensive cotton crops occur that usually receive large quantities of fertilizers often employing low efficiency irrigation systems, in conjunction to low recharge rates of the aquifer through lateral crossflows from the surrounding formations.

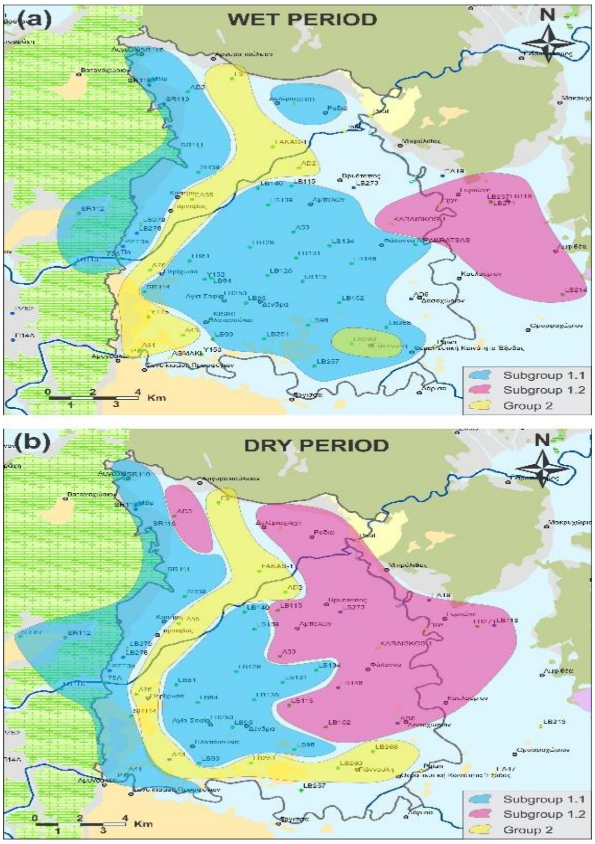


**Figure 1.** Study area with spatial distribution of HCA groups for wet (a) and dry (b) periods


**References**

Logeshkumaran A., Magesh N. S., Godson P. S. and Chandrasekar N. (2015), Hydro-geochemistry and application of water quality index (WQI) for groundwater quality assessment, Anna Nagar, part of Chennai City, Tamil Nadu, India, *Applied Water Science*, **5(4)**, 335-343.

Melloul A. and Collin M. (1992), The 'principal components' statistical method as a complementary approach to geochemical methods in water quality factor identification; application to the Coastal Plain aquifer of Israel, *Journal of Hydrology*, **140(1-4)**, 49-73.

Morell I., Giménez E. and Esteller M. V. (1996), Application of principal components analysis to the study of salinization on the Castellon Plain (Spain), *Science of the Total Environment*, **177(1-3)**, 161-171.

Panagopoulos A. (1996), A methodology for groundwater resources management of a typical alluvial aquifer system in Greece, School of Earth Sciences, Ph.D. Thesis, *University of Birmingham*, 46p.

Tziritis E., Skordas K. and Kelepertsis A. (2016), The use of hydrogeochemical analyses and multivariate statistics for the characterization of groundwater resources in a complex aquifer system. A case study in Amyros River basin, Thessaly, Central Greece, *Env. Earth Sciences*, **75(4)**, 339.

Voudouris K., Panagopoulos A. and Koumantakis J. (2000), Multivariate statistical analysis in the assessment of hydrochemistry of the Northern Korinthia prefecture alluvial aquifer system (Peloponnese, Greece), *Natural Resources Research*, **9(2)**, 135-146.

Pisinaras V. (2008), Development of a methodology framework for the integrated management of complexed groundwater systems Ph.D. Thesis, *Democritus University of Thrace*, 98p.