



Assessment of natural and anthropogenic impacts in groundwater, utilizing multivariate statistical analysis and inverse distance weighted interpolation modeling: the case of a Scopia basin (Central Greece)

Nikos Charizopoulos¹ · Eleni Zagana² · Aris Psilovikos³

Received: 16 October 2017 / Accepted: 14 May 2018 / Published online: 18 May 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In Scopia basin, central Greece, a hydrochemical investigation was completed. Groundwater samples from 41 sites were used to assess the natural and anthropogenic impacts in groundwater, utilizing the principal component analysis (PCA) involved with the inverse distance weighted (IDW) interpolation modeling and hierarchical cluster analysis (HCA). Best fit model to explain the spatial distribution of both hydrochemical parameters and PCA was chosen by optimizing the IDW interpolator's parameters. Precision of the model was picked based on less root-mean-squared prediction error (RMSPE) amongst predicted and actual values measured at the same locations. Groundwater exhibit Ca–Mg–HCO₃ as the dominant hydrochemical type and their greater part are mixed waters with non-dominant ion. Interpolation models demonstrate high estimations of nitrates in zones with agricultural activities and high estimations of nickel and chromium in regions with the strong presence of ultrabasic rocks. Dominant part of the groundwater samples surpasses in many cases the European Community (EC) drinking water permissible limits. Thus, they are unsuitable for human consumption. PCA illustrated four factors, which clarified 80.62% of the aggregate variance of data and HCA classified two statistically significant clusters of sampling sites. Results show natural procedures ascribed to the weathering of the minerals contained in the ultrabasic rocks and anthropogenic influences related to the use of fertilizers and wastewater leak. In light of FAO standards and Richards's classification, the groundwaters are reasonable for irrigation purposes, featuring waters with low sodium hazard and moderate salinity hazard.

Keywords Groundwater hydrochemistry · Principal component analysis · Hierarchical cluster analysis · Inverse distance weighting · Scopia basin

Introduction

The groundwater bodies are associated with the earth's solid phase (soils, rocks), and they change subjectively through natural procedures (McArthur et al. 2004; Bourette et al. 2009; Jiang et al. 2009; Mills et al. 2011) and different anthropogenic impacts. Agrarian practices, for example, fertilizers, pesticides, animal wastes (Ju et al. 2006; Jackson

et al. 2008; Hansen et al. 2011; Magesh et al. 2012) and residential development (wastes and refuse) (Mull et al. 1992; Stollenwerk 1996; Eiswirth et al. 2000; Wakida and Lerner 2005) play a critical part in water resources quality degradation. However, for the sustainable management of water resources, there is a requirement for understanding the water quality deterioration from diffusive or nonpoint contamination sources, as well as the background chemistry related to the natural geochemical forms. These factors may not distinguish the chemical composition of groundwater alone (Kyoung-Ho et al. 2014). Thus, in recent years, different methods have been created for the proficient administration and prediction of groundwater quality. Such methods are Multivariate Statistical Analysis (MSA) and GIS. The MSA with the techniques of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) is a quantitative and autonomous approach of groundwater classification permitting the grouping of groundwater samples and the

✉ Nikos Charizopoulos
nchariz@gmail.com

¹ Lab. of Mineralogy-Geology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece

² Department of Geology, Laboratory of Hydrogeology, University of Patras, Rio, 26110 Patras, Greece

³ Department of Ichthyology and Aquatic Environment, University of Thessaly, Odos Fitokou, 384 46 Volos, Greece

correlation between chemical parameters and groundwater samples (Cloutier et al. 2008). The PCA can uncover a straightforward basic structure inside a multivariate dataset by diminishing the dimensions of the original variables into new principal components (Jolliffe 2002) and constitute a standout amongst the most vital statistical techniques for the elucidation of groundwater chemistry (Dunteman 1989). Clustering is an unsupervised technique of data grouping, utilizing a given measure of similarity. In a hydrogeochemical study, a cluster analysis serves the purpose of isolating a group of representative clusters (also known as water type or a hydrogeochemical facies) that reflects the processes generating the natural variation found in an hydrogeochemical parameter (Nguyen et al. 2015). Furthermore, the depiction of Geochemical data in a GIS platform shown as an integrated tool to investigate factors which control the hydrogeological processes. The Inverse Distance Weighting method (IDW) of ArcGIS Geostatistical Analyst application is a commonly used interpolation technique, which is used to obtain the spatial distribution of groundwater quality parameters (Asadi et al. 2007; Arif et al. 2014). The IDW in correlation with different techniques, most particular, kriging is more straightforward to programming and does not require pre-modeling or subjective expectations in choosing a semi-variogram model (Henley 1981; Tomczak 1998). In the IDW, the optimal interpolator can be achieved by optimizing a set of parameters; the number of the neighbors to be included, the least numbers of the neighbors, the shape and section type, the major/minor axis of shape type and the optimal power value. PCA, HCA, and IDW consist of a scientifically robust way of quantifying data, and they have been applied successfully in the combination or individually on a hydrochemical groundwater researcher. For instance, Olmez et al. (1994) found that PCA is a valuable tool in helping to differentiate among several possible sources of groundwater pollution, each having similar gross Geochemical characteristics. Demirel and Güler (2006) applied PCA, HCA and Geochemical modeling systems to decide the principal factors and mechanisms controlling the chemistry of groundwaters in the Mediterranean coastal aquifer, Mersin–Erdemli basin (Turkey). Monjerezi et al. (2011) utilized a coordinated utilization of HCA and PCA and found that the chemical character of groundwater in the lower Shire River valley in Malawi is extremely changeable, with localized areas of predominantly brackish water. They applied the IDW technique to model the spatial distribution of the hydrogeochemical factors. Yidana (2010) performed PCA and HCA to classify groundwater samples spatially and determine the probable sources of variation in groundwater salinity in southeastern Ghana. Gong et al. (2014) compare the accuracy of IDW and Kriging interpolations to estimate groundwater arsenic concentration in Texas USA. They found that the correlation coefficient between the measured

and estimated arsenic levels was greater with IDW than Kriging Gaussian, Kriging spherical or Co-Kriging interpolations when analyzing data from wells in the entire Texas. Ghosh and Kanchan (2014) coupled geochemical analysis of various parameters with statistical analysis (PCA and HCA) to distinguish the pollution zones of groundwater in the central alluvial tract of the Bengal plain of India. They applied the IDW for spatial interpolation and mapping for both Geochemical parameters and statistical analysis outcomes.

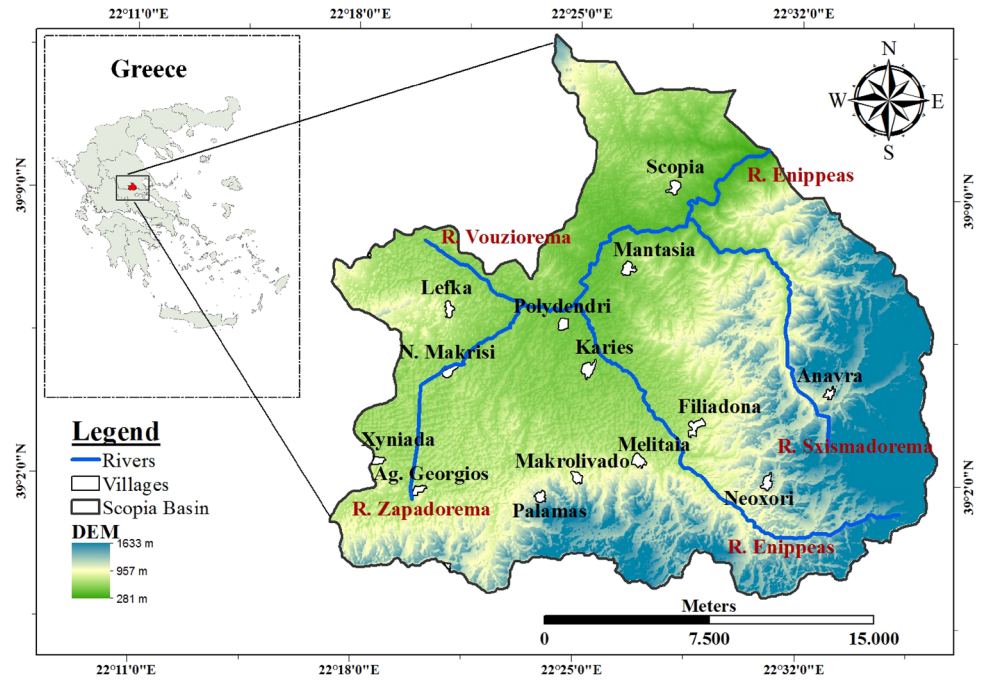
Scopia basin is the study region of the present work. It is arranged in the Domokos plateau in the Central Greece and undergoes the impacts of urbanization, seriously irrigated farming activity and the broad utilization of fertilizers. Furthermore, a broad ophiolites mass constitutes the basin's bedrock influencing groundwaters naturally. The hydrological status of Scopia basin has been recognized by the specialized reports of IGME (2010) and the Ministry of Environment and Energy (Ypen 2012) under the Directive (2000/60) of the European Council (EC 1998). However, none of these specialized reports, neither past analysis, have utilized PCA in blend with an optimized set of IDW parameters such as power value, shape type (circle/ellipse) and section type (divided circle/ellipse into one sector, four sectors, four sectors with 45° offset and eight sectors). Accordingly, this study introduces a new approach to IDW interpolation modeling for both groundwater hydrochemical parameters and PCA factors spatial dispersion. The method applied in the Scopia basin so as to: (a) uncover natural procedures and anthropogenic activities representing groundwater quality and their spatial distribution, (b) decide the propriety of groundwater for human utilization in the study zone, with the comparison of the concentrations of the main and trace elements in groundwater with the values established by the EC. In addition, (c) distinguish the suitability of groundwater for irrigation purposes as per FAO irrigation water guidelines and Richards diagram.

Study site

Site description

Scopia basin is arranged between latitudes 39°12'N and 39°02'N and longitudes 22°17'E and 22°36'E. It is the hydrological basin of the upper part of Enippeas River, which originates from the Othrys Mount, and runs through the southeast plain of Thessaly and streams into the Pinios River in central Greece (Fig. 1). Geographically, the basin consists mainly of a flat central and northern part, the high central peaks of Othrys Mount to the south and east and small territory hills to the north and west. It covers a region of 438.79 km² and has a perimeter of 116.06 km indicative of a large-sized basin. The lowest and highest points are 281 m and 1633 m, respectively, while the mean altitude

Fig. 1 Geomorphological setting map of Scopia basin



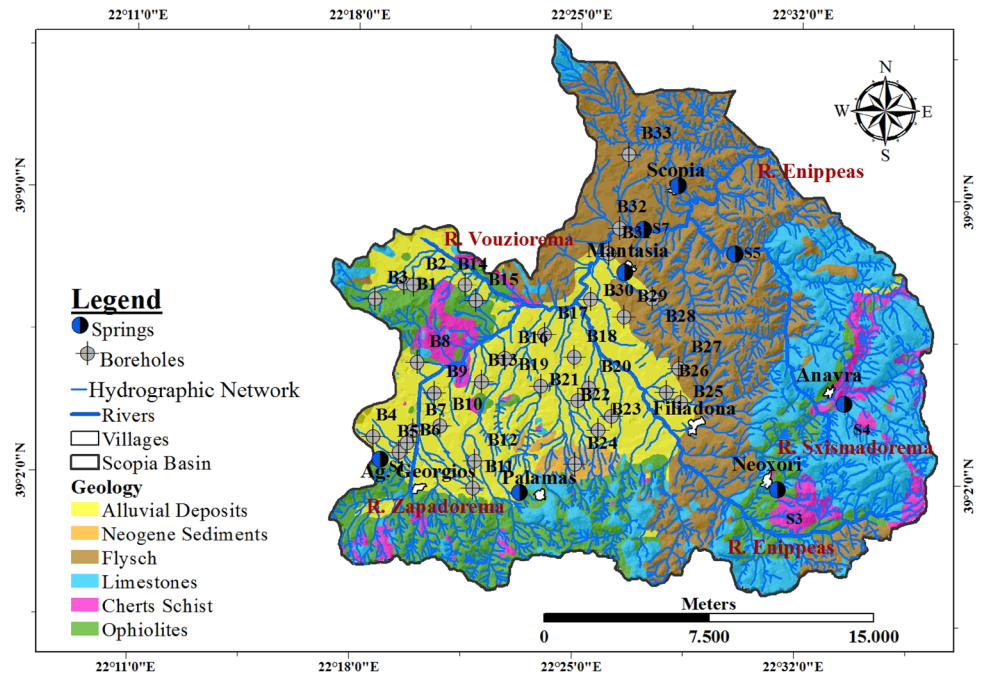
is 629 m, classifying the basin site as a semi-mountainous (Fig. 1) (Charizopoulos and Psilovikos 2015). The climate is portrayed from June to August as super dry, as per Lang's drought index (Trewartha and Horn 1980), and the irrigation period lasts from the past 10 days of April until the end of September. The yearly precipitation is 697.9 mm, and the mean yearly temperature is 14.8 °C. In winter, temperatures fall beneath zero and in summer transcends 40 °C. The aggregate yearly runoff is $49.7 \times 10^6 \text{ m}^3$ for the period October 2009–March 2011 (Charizopoulos 2013).

Geological and hydrogeological settings of the area

The study area is a part of the Pelagonian Geotectonic Zone of Eastern Greece, which is mainly characterized by the ophiolite series and the Schist–chert formation (Xypoliass et al. 2010). It presents a complex Geological structure, because of the intense tectonic activity. The oldest Geological formation in the study area dates back to the Triassic–Cretaceous age and includes ophiolites (Peridotites and Diabases), which also comprises part of the basin margin and the bedrock of the plain of the basin (Mountrakis 1985; Katsikatsos et al. 1986; Karmis 2010). The lithology of the study area consists of (a) alluvial deposits with mainly clays and conglomerates and a thickness up to 200 m. They are found in the central plain of the basin. (b) Neogene sediments with marls, clays, gravels, conglomerate and Marly limestone found in the south part of the basin. (c) Upper Cretaceous–Paleocene Flysch. It outcrops at the north and east parts of the basin (Fig. 2) and consists mainly of conglomerates, Marly carbonate sediments,

and clay sandstones. (d) Upper Cretaceous transgressive carbonate sediments mainly at the south part of the basin (Fig. 2) with a thickness of 200–300 m. (e) Triassic–Cretaceous Schist–chert formation and ophiolites (Karmis 2010). The ophiolites include Dunites with Chromites, Peridotites, Diabases, Gabbros and pillow lavas with a high concentration of alkali compounds (Na_2O 3.4–5.4% and K_2O 0.1–1.7%). Serpentinities are also found in the area as a metamorphic product of ultrabasic rocks (Feri re 1982). They occur in the south and northwest parts of the basin (Fig. 2). Groundwater in the study area is represented by Karst springs, spring from the fractured formations and a significant unconfined aquifer. In the west part of the area, close to the villages Neochori and Anavra, Karst springs discharge the karstified carbonate sediments. During the wet period, the discharge rate of the springs exceeds 50 l/s and a portion of the karst water flows towards the Enippeas River (Charizopoulos 2013). Some of the karst springs fall dry during the dry period. At the southeast and northwest parts of the basin, groundwater circulates through the fracture zone of the ophiolites and the flysch formations. Springs with low discharge rates (1–2 l/s) emerge. A highly productive shallow unconfined aquifer is developed in the alluvial deposits at the central plain of the basin. The recharge of the aquifer takes place by direct infiltration of precipitation, the lateral feed of the karst water of carbonate rocks and the lateral feed on the Enippeas River. A large number of boreholes with depths from 65 to 125 m and discharge rates from 30 to 80 m^3/h exploit the unconfined aquifer for drinking and mainly for

Fig. 2 Geology setting map of Scopia basin



irrigation purposes. The absolute groundwater level ranges about 404–484 m (IGME 2010).

Methodology

Data collection and chemical analysis

An aggregate of 41 groundwater samples was gathered in April 2010, since, in the wet period every one of the springs is functioning, while in the dry period most of the springs fall dry. Groundwater samples were gathered from 8 spring's outlets which discharge from ophiolites (S1), karst aquifer (S2–S5) and flysch (S6–S8) and 33 boreholes (B1–B33) in the plain of the basin (Fig. 2). The sites of groundwater tests were recorded utilizing a Garmin Dakota 20 GPS (Fig. 2). The samples were gathered in two polyethylene bottles (100 and 1000 ml volume) and were stored in a frozen cooler amid the fieldwork. The initial segment of 100 ml was filtered through 0.45- μm pore size Millipore filters and acidified to pH about 2 with 65% ultra pure HNO_3 . It was utilized for the determination of heavy metals concentration (Fe, Mn, Cu, Cr, Ni, Pb, Cd, Co and Zn). The second non-acidified part (1 l) was held to determine major cations and anion analysis (CO_2 , Ca^{2+} , Mg^{2+} , Na^+ , K^+ , HCO_3^- , Cl^- , SO_4^{2-} , NO_3^- , NH_4^+ , PO_4^{3-} , F^- , Br^- , I_2) (Lloyd and Heathcote 1985; APHA 1998; Appelo and Postma 2005). All the water samples were stored in the icebox in the laboratory at a temperature of 4 °C and were analyzed inside 3 days in

the wake of sampling. Total Hardness, Temporal Hardness, CO_2 and also Cl^- were determined using titration kits (Total Hardness: ManVer Buret Titration Method 8226, Temporal Hardness: Buret Titration Method, HCl 0.1N Titrant, Methyl red indicator, CO_2 : Titration Method, 0.0227 N Sodium Hydroxide Standard Solution Titrant, Phenolphthalein indicator and Cl^- : Buret Titration Method AgNO_3 0.1 N Titrant and K_2CrO_4 indicator). SO_4^{2-} , NO_3^- , NH_4^+ , PO_4^{3-} , F^- , Br^- and I^- were determined by spectrophotometry (HACH DR/3000) utilizing the appropriate HACH kits (SO_4^{2-} Sulfaver 4 Method 8051, NO_3^- Cadmium Reduction Method 8039, NH_4^+ Nessler Method 8038, PO_4^{3-} Phosver 3 Method 8048, F^- SPADNS Method 8029, Br: DPD Method 8016 and I_2 : DPD Method 8031). The elements Ca^{2+} , Mg^{2+} , Fe, Mn, Cu, Cr, Ni, Pb, Cd, Co, and Zn were determined by an atomic absorption spectroscopy (GBC/908AA), with a detection limit for heavy metals the value of 0.001 mg L^{-1} , which is the lower furthest limit of detection. The Na^+ and K^+ were determined utilizing a Flame photometer (INTECH/420). Every one of the analysis was led in the lab of the Institute of Mineralogy and Geology of the Agricultural University of Athens. The charge-balance error for major ionic species because of the conceivable deviation of the analytical procedure was under 5%, and it was statistically acceptable (Freeze and Cherry 1979; Reed and Mariner 1991). Since the purpose of this study is the assessment of natural and anthropogenic impacts in groundwater based on major and trace elements, farther C isotope analyses in groundwater samples for an extensive approach of C in groundwater were not performed.

Multivariate statistical analysis

In the present study, the MSA performed with statistical packet SPSS 19, with the methods PCA and HCA.

Principal component analysis

The PCA method establishes participation of individual chemicals in influence several factors, which commonly affect hydrochemistry (Vega et al. 1998). PCA is used to study all the existing variations (standard, unique and error) producing components to “extract” the largest percentage of variance with the minimum number of factors. The steps of PCA process are:

1. Construction of the correlation matrix, which shows what variables have high correlation and will count in factor extraction. Two indicators are used to check the effect of the PCA implementation on the data (a) the Kaiser–Meyer–Olkin index (KMO) which is an indicator of data adequacy (> 0.50) and checks if the original variables can be factorized efficiently (Norusis 2011) and (b) Bartlett’s test of sphericity which compares the correlation matrix with a matrix of zero correlations (identity matrix); it checks if there is a certain redundancy between the variables that can be summarized with a few numbers of factors ($p < 0.05$) (Cattell 1978).
2. Factor extraction, which considers just factors that have eigenvalues greater than “1” (Davis 2002).
3. Factors rotation with the Varimax method (Kaiser 1958) to achieve a more significant distribution of the weights of the different variables on the components (Davis 2002).
4. Calculation of factor scores, which indicates the contribution of each factor at every site (Voudouris 2009).

Hierarchical cluster analysis

The CA is an effective tool that reveals the fundamental structure or underlying conduct of a dataset without making a priori suppositions about the data to classify the objects of the system into categories or clusters in light of their nearness or similarity. In CA, the distance between samples is utilized as a measurement of similarity (Otto 1998; Vega et al. 1998). The outcome is a dendrogram which gives a visual portrayal of the clustering procedure by showing a figure of the groups and their proximity with a significant reduction in dimensionality of primary data (Kruskal and Landwehr 1983). There are two major categories of CA: hierarchical and non-hierarchical. The hierarchical CA (HCA) is the most widely recognized approach in which clusters are formed sequentially, starting with the most identical pair of objects and forming higher clusters step by step

and used in this work. The squared Euclidean distance was used as a similarity or dissimilarity measurement, whereas Ward’s linkage method was used to link clusters (Ward 1963). Ward’s method is capable of minimizing the distorting effect or sum of squared distances of centroids from two hypothetical groups generated at each stage (Lambrakis et al. 2004). The combination of these methods has been observed as the best in optimal results in HCA (Lambrakis et al. 2004; Mencio´ and Mas-Pla 2008; Lin et al. 2012). The observed water quality data, x_{ji} were standardized by z scale transformation as given below:

$$z = \frac{x_{ji} - \bar{x}_j}{s_j}, \quad (1)$$

where x_{ji} value of the j th water quality parameter measured at i th site \bar{x}_j mean (spatial) value of the j th parameter, and s_j standard deviation (spatial) of the j th parameter (Machiwal and Jha 2015).

GIS interpolation modeling

The depiction of the spatial distribution of hydrochemical parameters as well as factor scores was achieved with the Inverse Distance Weighted (IDW) interpolation modeling. The attribute of this method is that nearby locations are more likely to have similar values and the linear interpolator weights the interpolated data $\hat{z}(x_0)$, at unsampling location x_0 , as follows (Fortin et al. 2005; Mantzafleri et al. 2009):

$$\hat{z}(x_0) = \sum_{j=1}^m w_j \times z(x_j). \quad (2)$$

In Eq. (2), the $z(x_j)$ is the value of the water quality parameter z at the sampling location j , m is the number of neighboring sampling sites, w_j are the weights according to the distance between the unsampling location x_0 and the sampling locations x_j such that $\sum_{j=1}^m w_j = 1$. The formula of IDW method is finally obtained as follows (Fortin et al. 2005; Zisou and Psilovikos 2012).

$$\hat{z}(x_0) = \frac{\sum_{j=1}^m z(x_j) \times d_{ij}^{-k}}{\sum_{j=1}^m d_{ij}^{-k}}, \quad (3)$$

where power parameter k is the distance influence coefficient, d_{ij} is the distances between the unsampling location i (x_0) and the sampling locations j (x_j). Weights in IDW can be raised to the power of k (i.e., Linear, squared, cubed, etc.), so as the distance increases, the weights decrease rapidly. The power of k can also be optimized. The optimization process calculates several models and chooses the power value that leads to the model with the minimum Root Mean Square

Prediction Error (RMSPE). IDW assumes that the simulated surface is being driven by the local variation which can be captured through the neighborhood. Changing neighborhood options (shape, sectors and the number of neighbors) may lead to a better model. The shape confines how far and where to look for the measured values, which will be utilized in the prediction. If the neighborhood is divided into sectors, the maximum and minimum constraints will be applied to each sector (Esri 2012). The interpolation accuracy of the method was measured by computing the Mean Prediction Error (MPE) and the RMSPE for the data. The MPE of the interpolated values was calculated with the following Eq. (4):

$$\text{MPE} = \frac{1}{n} \sum_{i=1}^n (\hat{x}_i - x_i). \quad (4)$$

RMSPE is the square root of the average squared distance of a data point from the fitted line calculated with the following Eq. (5):

$$\text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{x}_i - x_i)^2}, \quad (5)$$

where x_i and \hat{x}_i are the measured and estimated values, respectively, of the i th data points, and n is the total number of data points (Gong et al. 2014; Wilford et al. 2016). The optimal value is determined by minimizing RMSPE, which is finally a summary statistic quantifying the error of the predicted surface (Esri 2012). In this study, with the aim of selecting the best-fitting, interpolation model based on the minimum RMSPE value (Eq. 5), weight raised to optimize power, circle/ellipse shapes were used and also four different neighborhood sectors (a, one sector; b, four sectors; c, four sectors with 45° offset and d, eight sectors). Although all the IDW parameters may be optimized, the following assumptions were made to simplify the method: (1) the neighborhood type was standard, (2) the neighboring points considered in the process, were minimum 10 and maximum 15, (3) the anisotropy angle was set in 0° and (4) the minor semi-axes were set equal to three times the major semi-axes in ellipse shape type. The anisotropy factor (the ratio of the major to the minor semi-axes lengths) was 1 for circle and 0.36 for ellipse shape type. All the above counts were acknowledged by applying the Geostatistical Analyst of the Arc 10.0 GIS® software (Philip and Watson 1982; Watson and Philip 1985).

Results and discussion

Hydrochemical data and IDW interpolation modeling

Table 1 presents the Univariate statistics summary ($n=41$) for 24 water quality parameter values, hardness, total (TH), TDS, Eh, pH, electric conductivity (EC) and water temperature ($T^{\circ}\text{C}$).

Table 2 summarizes the results of the accuracy assessment of IDW interpolation method for various hydrochemical parameters. The goodness-of-fit criteria suggested that the best-fitting models, to understand the spatial distribution of the hydrochemical parameters, are the ellipse shape type and neighborhood deviation of four sectors for TDS and NO_3^- , ellipse/eight sectors for SO_4^{2-} and Cd and ellipse/one sector for Ni. Furthermore, it is observed that the best-fitting, interpolation model represented by circle shape type and neighborhood deviation of one sector for Mg^{2+} and Cl^- and circle/eight sectors for Cr.

Physicochemical parameters

The pH ranges from 6.9 to 8.2. The spring presents lower values (6.9–7.4), while the boreholes have a basic character with pH from 7.1 to 8.2 due to the hydrolysis of magnesium-bearing minerals, such as Olivines of ophiolitic origin, contained in the alluvial deposits. Eh with values from 200 to 630 mV shows that the oxidizing conditions prevail in the study area, and this is enhanced by the presence of nitrates and sulfates. EC ranges from 283 to 1138 $\mu\text{S cm}^{-1}$. Springs present EC values from 283 to 1138 $\mu\text{S cm}^{-1}$, while boreholes from 416 to 947 $\mu\text{S cm}^{-1}$. The highest EC values in S6, S8 springs and B31 and B33 boreholes found at the northeast and southwest parts of the basin, are related to the farming activities and the presence of septic tanks. TDS varies from 350.5 to 1035.0 mg L^{-1} . In spring, the TDS values range from 350.5 to 1035 mg L^{-1} , while in boreholes from 456.7 to 965.1 mg L^{-1} . The lowest TDS values are measured at the southeast part of the basin in karst springs. It is obviously associated with the high flow rates within karstified carbonate formations and the smaller residence time therein. The best-fit interpolation model for TDS as it was presented by ellipse shape type and neighborhood deviation of four sectors (Table 2), shows high values in the north and southwest areas of the study site associated with urban wastewater and cultivation methods (Fig. 3a). The values of Total Hardness ranges from 11.8 to 36.9 °dH with 78% of the samples with values above 18 °dH; thus the majority of the samples is characterized as hard waters.

Table 1 Chemical analyses and descriptive statistics of Scopia basin groundwater

	Springs (<i>n</i> :8)						Boreholes (<i>n</i> :33)					MAV ^a
	Units	Min	Max	Mean	SD	Median	Min	Max	Mean	SD	Median	
<i>T</i>	°C	11.4	16.1	14.4	1.6	14.6	14.5	18.5	15.9	0.8	15.8	25
EC	μS cm ⁻¹	283.0	1,138.0	629.4	274.3	556.0	416.0	947.0	643.4	146.5	618.0	–
TDS	mg L ⁻¹	350.5	1,035.0	660.3	214.4	636.5	456.7	965.1	653.3	125.1	641.8	350.5
pH		6.9	7.4	7.2	0.2	7.3	7.11	8.2	7.6	0.2	7.6	6.5–9.5
Eh	mV	200.1	630.0	349.6	148.6	289.8	225.0	332.7	262.0	28.0	246.3	–
O ₂	mg L ⁻¹	3.1	6.9	5.3	1.7	6.4	3.9	8.4	6.58	0.84	6.5	–
CO ₂	mg L ⁻¹	60.0	230.0	134.5	63.9	107.5	45.0	155.0	99.0	28.6	103.0	–
Tot. hardness ^b	°dH	11.8	36.9	22.5	8.0	20.5	13.3	29.0	21.6	4.1	21.6	–
Ca ²⁺	mg L ⁻¹	75.2	215.2	126.8	47.76	105.6	8.8	155.2	97.1	35.18	100.0	100
Mg ²⁺	mg L ⁻¹	5.5	46.6	20.5	13.2	17.3	12.9	91.4	34.7	18.8	27.62	50
Na ⁺	mg L ⁻¹	2.5	36.3	14.5	10.8	11.9	0.94	61.5	25.6	14.3	21.0	200
K ⁺	mg L ⁻¹	0.39	11.4	3.5	3.79	1.56	ND	2.0	0.4	0.52	0.2	12
NH ₄ ⁺	mg L ⁻¹	ND	0.53	0.08	0.18	0.01	ND	0.45	0.09	0.12	0.04	0.5
HCO ₃ ⁻	mg L ⁻¹	244.0	488.0	388.9	75.8	390.4	280.6	549.0	384.9	54.05	372.1	–
Cl ⁻	mg L ⁻¹	10.6	67.4	27.0	18.4	23.0	7.1	95.7	29.5	21.7	18.1	250
SO ₄ ²⁻	mg L ⁻¹	2.5	131.4	38.8	48.0	11.6	0.5	91.2	36.7	28.4	31.7	250
NO ₃ ⁻	mg L ⁻¹	6.2	90.2	38.9	34.8	25.1	11.9	134.6	43.6	25.1	41.4	50
PO ₄ ³⁻	mg L ⁻¹	0.08	0.38	0.24	0.11	0.26	0.1	0.75	0.43	0.16	0.44	≈6.69
Br ⁻	mg L ⁻¹	0.02	0.64	0.12	0.21	0.03	ND	0.26	0.05	0.06	0.03	0.01
I ⁻	mg L ⁻¹	ND	1.01	0.17	0.34	0.04	ND	0.4	0.08	0.1	0.04	–
F ⁻	mg L ⁻¹	ND	0.71	0.35	0.23	0.37	ND	0.65	0.25	0.17	0.21	1.5
Fe	mg L ⁻¹	0.106	0.221	0.138	0.037	0.122	0.075	0.231	0.153	0.044	0.144	0.2
Mn	mg L ⁻¹	<0.001	0.008	0.003	0.002	0.002	<0.001	0.046	0.010	0.011	0.005	0.05
Ni	mg L ⁻¹	0.012	0.211	0.127	0.064	0.128	0.015	0.289	0.141	0.068	0.143	0.02
Cr	mg L ⁻¹	<0.001	0.062	0.022	0.026	0.007	<0.001	0.121	0.040	0.034	0.042	0.05
Cd	mg L ⁻¹	<0.001	<0.001	–	–	–	<0.001	0.040	0.006	0.008	0.001	0.005
Zn	mg L ⁻¹	<0.001	0.004	0.001	0.001	0.001	<0.001	0.318	0.045	0.078	0.007	0.10
Pb	mg L ⁻¹	<0.001	<0.001	–	–	–	<0.001	0.125	0.005	0.022	0.001	0.01
Co	mg L ⁻¹	<0.001	<0.001	–	–	–	<0.001	<0.001	–	–	–	–
Cu	mg L ⁻¹	<0.001	<0.001	–	–	–	<0.001	0.024	0.002	0.004	0.001	2.0

^aMAV max acceptable value given by the Dir 98/83/EC

^bTot. hardness total hardness

Main elements

Among cations, Ca²⁺ is the most abundant with concentration from 8.8 to 215.2 mg L⁻¹. In springs water, the concentration ranges from 75.2 to 215.2 mg L⁻¹, while in boreholes from 8.8 to 155.2 mg L⁻¹. The high values are attributed to the dissolution of carbonate minerals and are measured in karst springs. Mg²⁺ is the second most abundant cation with concentration from 5.5 to 91.4 mg L⁻¹. The interpolation model shows that the highest concentrations are observed in boreholes, which are fed from the ophiolites (B1 91.4 mg L⁻¹, B2 68.4 mg L⁻¹, B8 74.8 mg L⁻¹). The lowest is found in karst springs and springs from flysch formation (S4 5.5 mg L⁻¹, S5 10.0 mg L⁻¹) (Fig. 3b). Na⁺ concentration ranges from 2.5 to 36.3 mg L⁻¹ in the springs and

from 0.94 to 61.5 mg L⁻¹ in boreholes. This is related to the dissolution of sodium-rich pillow lavas and anthropogenic activities such as the use of sodium-rich dung as fertilizer, which is a very common agricultural practice in the study area. K⁺ concentration is very low from not detected (ND) to 11.4 mg L⁻¹. Among anions HCO₃⁻ is the most abundant with concentration from 244 to 488 mg L⁻¹ in springs and 280.6–549 mg L⁻¹ in boreholes (Table 1), because of the dissolution of the carbonate minerals. Cl⁻ concentration ranges from 7.1 to 95.7 mg L⁻¹, while the 80% of the samples show values below 40 mg L⁻¹. The best-fit interpolation model (Table 2) shows high estimations of Cl⁻ in the North and Southwest margins of the basin. They are associated with boreholes B9 (67.4 mg L⁻¹), B11 (95.7 mg L⁻¹), B17 (74.5 mg L⁻¹) and spring S8 (67.4 mg L⁻¹) in

Table 2 Values of the goodness-of-fit criteria to select the best-fit interpolation model of hydrochemical parameters

Variable	Sector type ^A	Shape type					
		Circle			Ellipse		
		Optimized power value	Goodness-of-fit criteria		Optimized power value	Goodness-of-fit criteria	
		<i>k</i>	MPE	RMSPE	<i>k</i>	MPE	RMSPE
TDS	a	1.00	0.32	131.79	1.71	-10.98	129.14
	b	1.07	-4.64	130.22	2.07	-13.06	128.46
	c	1.10	-2.24	132.79	1.80	-10.76	128.99
	d	1.36	-3.97	133.09	2.21	-12.79	129.13
Mg ²⁺	a	1.09	-0.55	11.77	1.51	0.39	12.96
	b	1.53	-0.33	12.14	1.70	0.71	13.03
	c	1.55	0.03	12.15	1.74	0.65	13.12
	d	1.71	0.09	12.35	1.76	0.74	13.11
Cl ⁻	a	1.00	-0.17	20.18	1.10	-0.60	21.01
	b	1.22	-0.89	20.45	1.26	-1.28	20.90
	c	1.21	-0.51	20.41	1.18	-0.87	20.85
	d	1.32	-0.75	20.46	1.26	-1.27	20.81
SO ₄ ²⁻	a	1.95	-1.03	30.69	2.35	-2.99	28.94
	b	1.98	-1.52	30.16	2.37	-3.24	28.70
	c	2.00	-1.32	30.26	2.32	-3.00	28.78
	d	2.05	-1.42	30.20	2.40	-3.17	28.74
NO ₃ ⁻	a	1.50	0.89	26.10	1.34	-0.53	25.32
	b	1.30	0.32	25.27	1.50	-1.23	25.18
	c	1.44	0.42	25.47	1.55	-1.00	25.37
	d	1.54	0.23	25.39	1.52	-1.20	25.19
Ni	a	1.00	0.001	0.070	1.00	0.003	0.067
	b	1.00	0.002	0.070	1.00	0.002	0.068
	c	1.00	0.002	0.070	1.00	0.002	0.068
	d	1.00	0.002	0.070	1.00	0.002	0.068
Cr	a	1.63	-0.000	0.032	1.00	0.000	0.033
	b	1.77	0.000	0.032	1.00	0.000	0.033
	c	1.71	0.000	0.032	1.00	0.000	0.033
	d	1.70	0.000	0.031	1.00	0.000	0.033
Cd	a	1.00	0.000	0.007	1.00	0.000	0.007
	b	1.00	0.000	0.007	1.07	0.000	0.007
	c	1.00	0.000	0.007	1.14	0.000	0.007
	d	1.11	0.000	0.007	1.12	0.000	0.006

^ASector type a, one sector; b, four sectors; c, four sectors with 45° offset; d, eight sectors

k: ≥1

MPE: best value is close to 0, RMSPE: best value is the lowest

Scopia village and are related to agrarian and municipal wastes (Fig. 3c). SO₄²⁻ concentration varies between 2.5 and 131.4 mg L⁻¹ in the springs and from 0.5 to 91.2 mg L⁻¹ in boreholes (Table 1). The best-fitting model for the spatial distribution of sulfates (Table 2) indicates high concentrations in North of the study area. It is attributed to the oxidation of pyrites of ophiolitic origin but also to the use of sulfur-ammonium fertilizers in the cultivated plain of the study range (Fig. 3d). Regarding NO₃⁻ is the second

most abundant anion among anions and ranges from 6.2 to 90.2 mg L⁻¹ in spring waters as well as from 11.9 to 134.6 mg L⁻¹ in boreholes. NO₃⁻ concentration in 29% of the water samples surpasses the European drinking water threshold limit (EC 1998). The best IDW predicted interpolator model as expressed by ellipse shape type and neighborhood deviation of four sectors (Table 2), present high values of NO₃⁻ in specific areas where groundwaters suffered by intense cultivation and related to the application of nitrogen

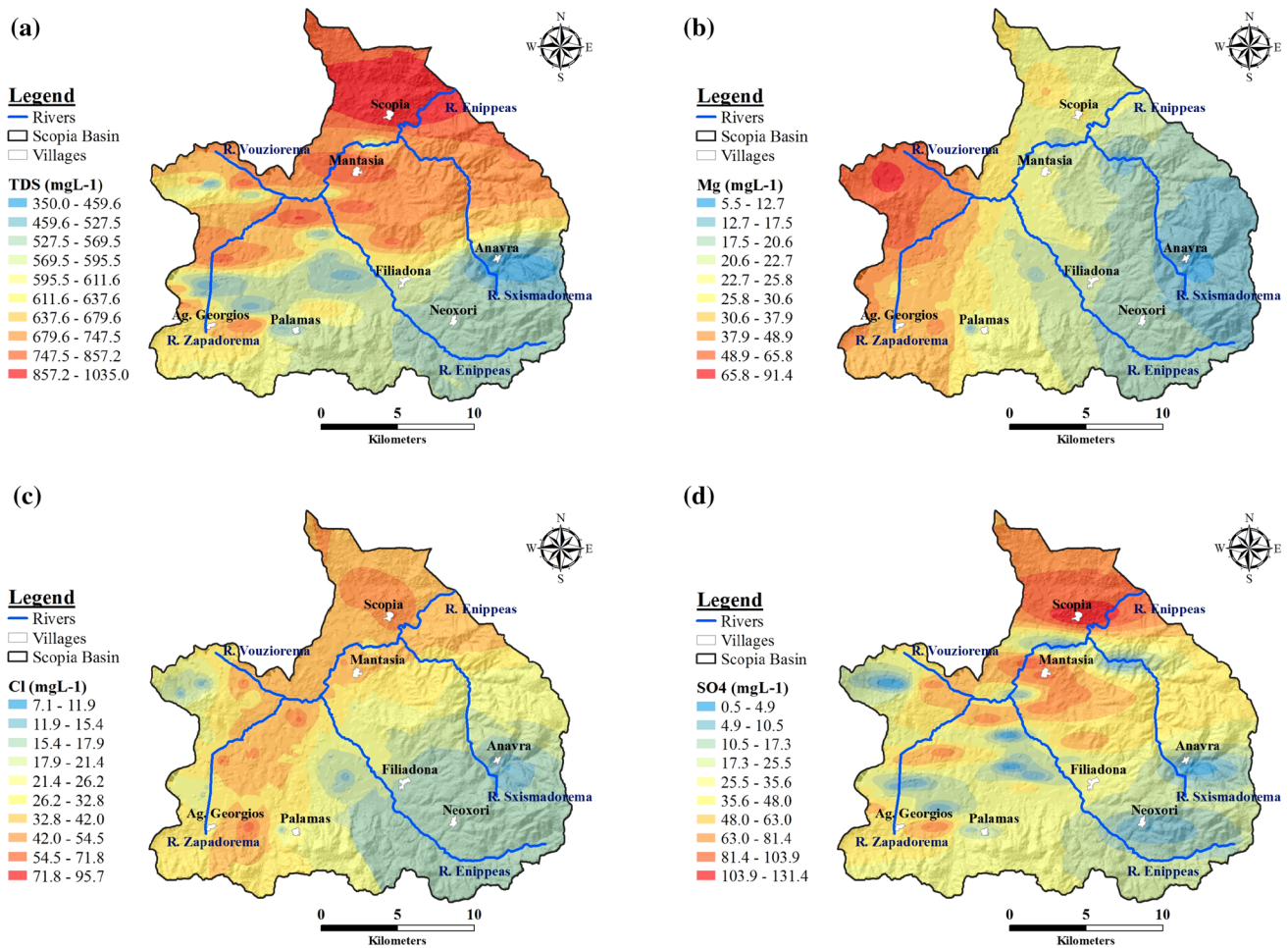


Fig. 3 a TDS spatial distribution, b Mg^{2+} spatial distribution, c Cl^{-} spatial distribution, d SO_4^{2-} spatial distribution

fertilizers (Fig. 4a). NH_4^+ values range from ND to 0.53 mg L^{-1} (Table 1). The highest concentrations are reported in residential and cultivated regions.

Trace elements

Regarding trace elements, Ni is the most abundant with concentration from 0.012 to 0.289 mg L^{-1} . In 93% of the groundwater samples, Ni surpasses the European drinking water limit of 0.02 mg L^{-1} . The spatial distribution guide of Ni in light of the best IDW interpolation model with ellipse shape type and neighborhood deviation of one sector (Table 2) is delineated in Fig. 4b. The high content of Ni is credited to the weathering of nickel-bearing minerals of the ultrabasic rocks found in the study area. Total iron is the second most abundant trace element with concentration from 0.075 to 0.231 mg L^{-1} . In 21% of the groundwater samples, total iron surpasses the drinking water threshold point of 0.2 mg L^{-1} (EC 1998). The presence of iron in the

study region is related to the presence of iron oxides in the clay sediments of ophiolitic origin.

The concentration of Cr_{tot} ranges from BDL to 0.121 mg L^{-1} , and 34% of the groundwater samples exceed 0.05 mg L^{-1} , which is the European drinking water threshold limit. The goodness-of-fit criteria recommended that the best-fitting model for the spatial distribution of Cr_{tot} is represented by circle shape type and neighborhood deviation of eight sectors (Table 2). The presence of the element is caused by the weathering of the minerals contained in the ophiolites, which cover a major piece of the study range (Fig. 4). Externally, the northwest margin of the area an abandoned open mine of chromite is found, which collects surface water of the area. Mn content ranges from below detection limit (BDL) to 0.046 mg L^{-1} and only in one borehole the concentration surpasses the 0.05 mg L^{-1} , of the European drinking water thresholds limit. Mn has a geogenic origin and is caused by the weathering of the minerals contained in the ophiolites. Cd values range from BDL to 0.040 mg L^{-1} , while in 24% of the groundwater samples, the concentration

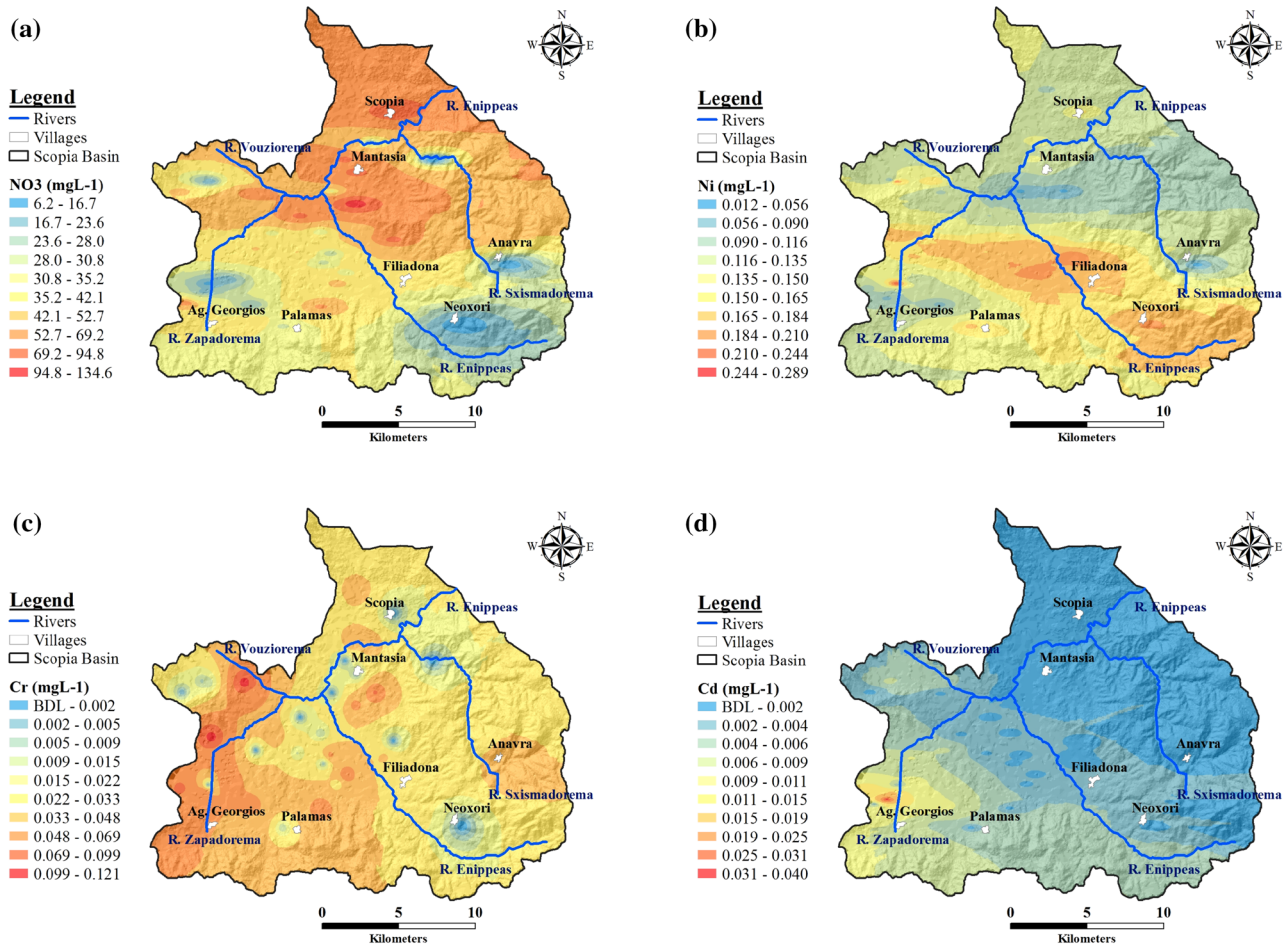


Fig. 4 a NO₃⁻ spatial distribution, b Ni spatial distribution, c total Cr spatial distribution, d Cd spatial distribution (BDL: Below detection limit)

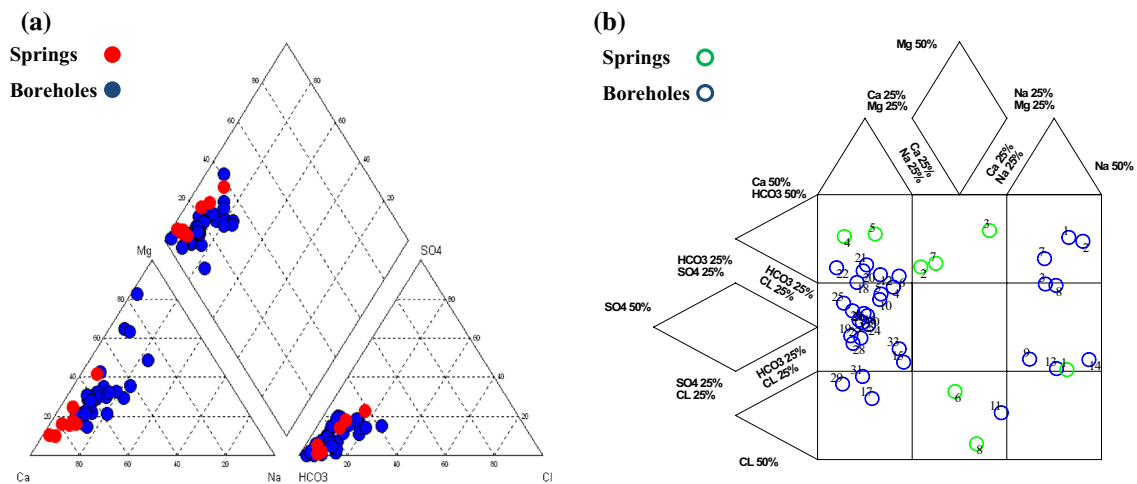


Fig. 5 a Classification of groundwater in Piper diagram, b classification of groundwater in the extended Durov diagram

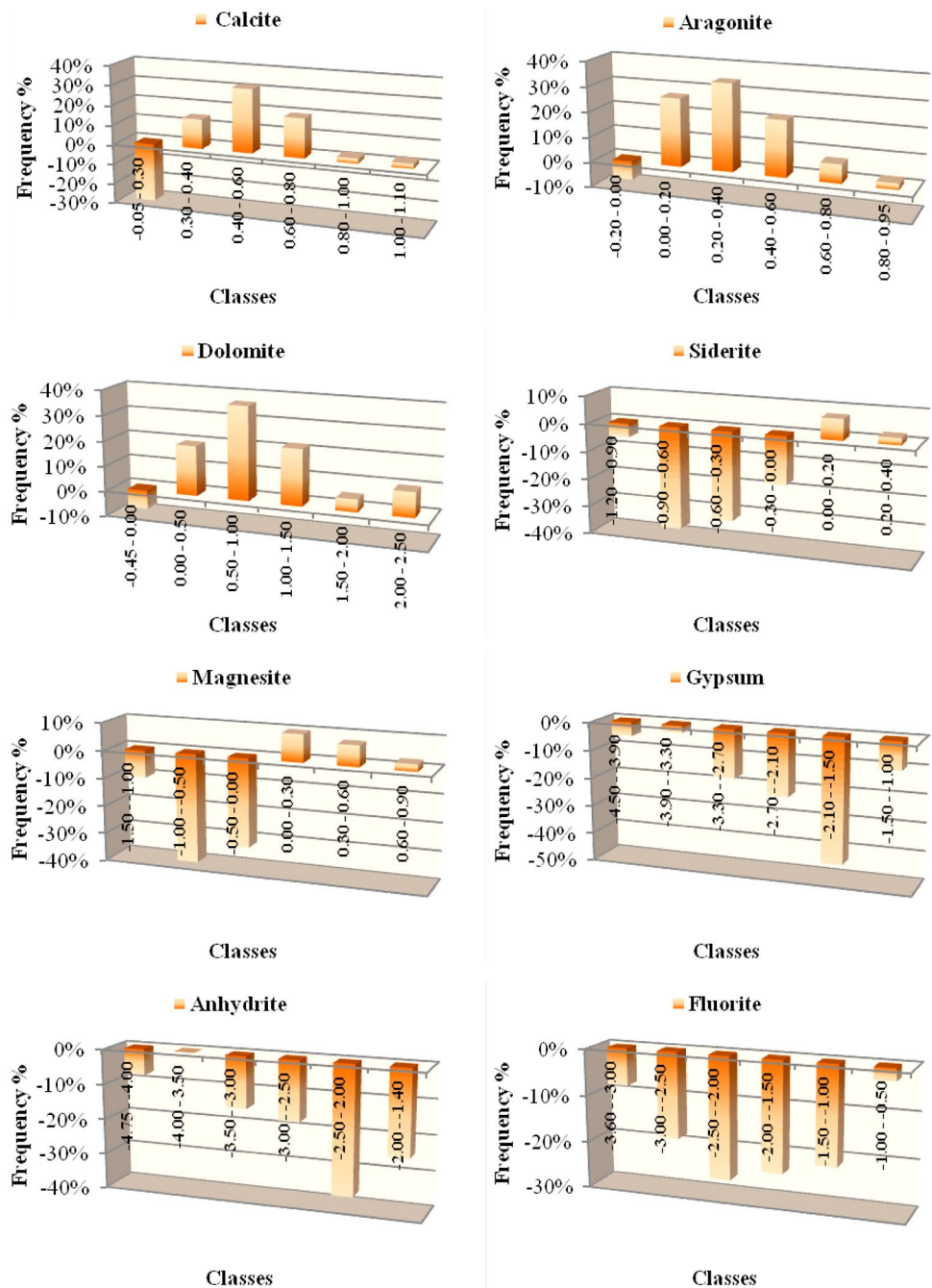
is recorded above the drinking water limit of EC, which is 0.005 mg L^{-1} . The best-fit interpolation model presents the high values of the Cd in the Southwest part of the basin. It can be credited both to the presence of the component in the minerals contained in ophiolitic rocks and the use of phosphate fertilizers in the cultivated part of the area (Fig. 4d). As indicated by Alloway (1995) and Kabata-Pendias and Pendias (2001), Cd is a fundamental trace component in phosphate fertilizers. The Zn values range from BDL to 0.318 mg L^{-1} , exceeding in four boreholes the European drinking water limit of 0.10 mg L^{-1} . Zn is contained in

pesticides, which are utilized in the cultivated plain in the area. The Cu and Pb concentrations range in BDL values of all water samples. Only borehole B1 present value of 0.024 mg L^{-1} for Cu and 0.125 mg L^{-1} for Pb which is above the 0.01 mg L^{-1} EC value.

Groundwater classification

The Piper outline (Fig. 5a) uncovers two principal water types in the study region. The greater percentage of the

Fig. 6 Saturation index diagrams of several minerals for groundwater samples



samples (90%) belongs to the first group with dominant hydrochemical type Ca–Mg–HCO₃ (Fig. 4a). The second group represents 10% of the water samples with a Ca–Mg–Na–HCO₃ water type. In the extended Durov diagram (Durov 1948), is obtained that the most of the samples (55%) belong to the fields fourth and sixth (Fig. 5b).

The next dominant fields, with a rate of 31%, are the first, second and third. Most of these samples are displayed in the first field (Ca–HCO₃) and considered as fresh waters (Appelo and Postma 2005), younger than groundwaters of the other types. The rest of the samples concerning cation exchange waters with hydrochemical processes determined by the phases of Mg–HCO₃ and Na–HCO₃. The seventh and eight fields are represented by the 14% of the samples, concerning waters which the reverse cation exchange phenomenon is in full advance. In this category, the samples with high pollution load belong, having mainly anthropogenic origin, while the hydrochemical processes determined by the phases of Ca–Cl and Mg–Cl.

Origin of groundwater elements

The PHREEQC software (Parkhurst and Appelo 2013) was utilized for the determination of saturation indices of specific minerals. The saturation index (SI), shows if a solution is in equilibrium, undersaturated or supersaturated with regards to a solid phase (Merkel and Planer-Friedrich 2008). The SI is expressed by the ratio $SI = \log(IAP/K)$. The IAP and K are the ion activity product and the mineral equilibrium constant at a given temperature, respectively. A negative value indicates undersaturation (possible mineral solution) and a positive value indicates supersaturation (possible mineral precipitation). If the SI is equal to zero, it reflects the solubility equilibrium with respect to the mineral phase of the water, a phenomenon which rarely achieved in nature (Appelo and Postma 2005). In the Scopia basin, every one of the samples is undersaturated with respect to evaporate mineral phases gypsum (CaSO₄·6H₂O) ($-4.49 < SI < -1.18$), anhydrite (CaSO₄) ($-4.74 < SI < -1.43$) and fluorite (CaF₂) ($-3.53 < SI < -0.81$). Undersaturated with

respect in siderite (FeCO₃) ($-1.16 < SI < +0.36$) and magnesite (MgCO₃) ($-1.46 < SI < +0.81$) are water samples in 88% and 90%, respectively (Fig. 6). The phenomenon is suggesting that these phases are minor in the host rocks of Scopia basin. The majority of the samples is saturated in calcite (CaCO₃) ($-0.05 < SI < +1.1$) in 98% percentage and in aragonite (CaCO₃) ($-0.02 < SI < +0.95$) and dolomite (CaMg(CO₃)₂) ($-0.43 < SI < +2.22$) in 93% (Fig. 6). The oversaturation of calcite, aragonite, and dolomite indicates that these minerals play a significant role in controlling the groundwater chemistry in the mixing zone. Calcite is an essential mineral which tends to dissolve or precipitate quite rapidly in natural waters (Alley 1993). Moreover, waters in the Meteoric environment, as long as they are still in contact with carbonates, maybe several times oversaturated with respect to calcite (James and Choquette 2013). In Fig. 6, the plots of the obtained SI and their classes are presented. Negative SI values represent the undersaturated conditions, while the positive SI values the oversaturated conditions.

Table 4 Factor loadings of the varimax rotated 4-factor model

Variable	Factor				Communality
	I	II	III	IV	
Ca	0.61	-0.55	-0.36	-0.31	0.90
Mg	0.20	0.90	0.20	0.12	0.91
Na	0.78	-0.07	0.23	0.16	0.69
Cl	0.80	0.04	-0.03	0.15	0.66
SO ₄ ²⁻	0.93	-0.08	-0.07	-0.17	0.90
NO ₃ ⁻	0.71	-0.12	-0.20	-0.36	0.69
NH ₄ ⁺	0.79	0.21	0.24	0.16	0.75
Cr	0.05	0.01	0.94	-0.12	0.90
Pb	-0.17	0.81	-0.27	-0.27	0.84
Ni	0.05	-0.05	-0.12	0.91	0.84
Total variance (%)	36.70	18.46	13.00	12.46	80.62

Significant loads >0.75, Moderate loads >0.65

Table 3 Correlation matrix of chemical data of Scopia basin groundwater's

	Ca	Mg	Na	Cl	SO ₄ ²⁻	NO ₃ ⁻	NH ₄ ⁺	Cr	Pb	Ni
Ca	1									
Mg	-0.50	1								
Na	0.34	0.09	1							
Cl	0.44	0.18	0.63	1						
SO ₄ ²⁻	0.66	0.10	0.65	0.65	1					
NO ₃ ⁻	0.64	0.03	0.36	0.31	0.78	1				
NH ₄ ⁺	0.26	0.38	0.56	0.59	0.62	0.39	1			
Cr	-0.22	0.19	0.19	-0.03	0.02	-0.04	0.21	1		
Pb	-0.31	0.51	-0.21	-0.14	-0.18	-0.10	-0.10	-0.14	1	
Ni	-0.14	0.05	0.09	0.05	-0.05	-0.09	0.10	-0.11	-0.20	1

High correlation >0.75, Moderate correlation >0.65

Multivariate statistical analysis

Principal component analysis

Table 3 presents the chemical data correlation matrix of Scopia basin groundwaters. A high correlation (> 0.75) is seen between SO₄²⁻ with NO₃⁻ (0.78) and moderate correlation between Ca with SO₄²⁻ (0.66) credited to the utilization of fertilizers. Furthermore, moderate correlation (0.65) between Na with SO₄²⁻ and Cl with SO₄²⁻ has been observed, associated with wastewater leaks. Utilization of PCA extracted four factors that clarify the 80.62% of the aggregate variance. The technique showed trustworthy as the Keiser–Meyer–Olkin index, and Bartlett’s test of sphericity were 0.62 (> 0.50) and 0.00 (< 0.05) respectively, while communalities were above (> 0.5) which implies the satisfactory efficiency of the 4-factor model (Table 4). Table 5 represents the results of the assessment of the accuracy of the IDW interpolation method for PCA factors.

Factor I Accounts 36.7% of the variance of the information matrix and contains noteworthy loads of SO₄²⁻ and Cl⁻ as well as moderate loads of NH₄⁺, Na⁺, NO₃⁻ and Ca²⁺. It portrays the anthropogenic effect on the groundwater quality. The high correlation between SO₄²⁻ Cl⁻ and Na⁺ is related to the leakage of agricultural and municipal

wastes (Sikora et al. 1976). The correlation between Ca²⁺ and NO₃⁻ (0.65) is connected with the use of fertilizers as NH₄NO₃·CaCO₃ (22% N and 33% NCaCO₃) (Tisdale and Nelson 1975), which is very common in cultivated regions of the study area. Figure 7a depicts the spatial distribution of factor I as represented by ellipse shape type and neighborhood deviation of four sectors (Table 5). The high values of the factor prevail in residential and cultivated areas, due to the absence of drainage system and the intensive usage of fertilizers respectively.

Factor II Accounts 18.4% of the variance of the information matrix and contains critical loads of Mg and Pb. It portrays the geogenic effect on the groundwater quality and is related to the weathering of the magnesium-bearing minerals that are contained in the ophiolites. The goodness-of-fit criteria suggested that the best-fitting model for the spatial distribution of factor II expressed by circle shape type and neighborhood deviation of one sector (Table 5). Figure 7c demonstrates that the positive values of the factor appear in the northwest part of the region, where ophiolites prevail.

Factor III Records 13% of the fluctuation of the information matrix with a noteworthy load of Cr_{tot} and portrays the impact of the weathering of chromites that are contained in the ophiolites. The best-fit model to understand the spatial distribution of factor III is the circular shape type with four

Table 5 Values of the goodness-of-fit criteria to select the best-fit interpolation model of PCA factors

Variable	Sector type ^A	Shape type					
		Circle			Ellipse		
		Optimized power value	Goodness-of-fit criteria		Optimized power value	Goodness-of-fit criteria	
		k	MPE	RMSPE	k	MPE	RMSPE
Factor I	a	1.66	-0.04	0.92	2.17	-0.09	0.89
	b	1.81	-0.06	0.91	2.28	-0.11	0.88
	c	1.82	-0.05	0.92	2.19	-0.10	0.89
	d	1.92	-0.06	0.92	2.33	-0.10	0.89
Factor II	a	1.00	-0.03	0.77	1.05	0.02	0.82
	b	1.15	-0.02	0.78	1.19	0.04	0.81
	c	1.16	0.00	0.78	1.20	0.03	0.82
	d	1.29	0.01	0.79	1.23	0.04	0.82
Factor III	a	1.42	-0.02	0.93	1.00	-0.00	0.97
	b	1.54	-0.02	0.92	1.27	-0.00	0.98
	c	1.55	-0.02	0.93	1.50	-0.01	0.99
	d	1.65	-0.02	0.93	1.37	-0.00	0.98
Factor IV	a	1.00	-0.00	1.07	1.00	0.03	1.01
	b	1.00	0.01	1.06	1.00	0.02	1.04
	c	1.00	0.01	1.05	1.00	0.02	1.04
	d	1.00	0.02	1.04	1.00	0.02	1.03

^ASector type a, one sector; b, four sectors; c, four sectors with 45° offset; d, eight sectors

k: ≥1

MPE: best value is close to 0, RMSPE: best value is the lowest

sector's type (Table 5). The high values of the factor occur around the old chromite mine and in regions where aquifers are fed by ultramafic rocks. In addition to the parts of the alluvial deposits, where the clastic material of ophiolitic origin prevails in the lithology, factor III shows high values (Fig. 7c).

Factor IV Shows a significant load of Ni and accounts 12.46% of the variance of the total matrix. The best-fitting, interpolation model for the spatial distribution of the factor is an ellipse shape type and a neighborhood deviation of one sector (Table 5). The factor describes the weathering of the nickel-rich minerals of the ophiolites (Fig. 7d).

Hierarchical cluster analysis

HCA was applied to groundwater sampling sites, according to their hydrogeochemical parameters, to be matched at each sampling site, the number of the cluster. Based on the dendrogram, the 41 sampling sites categorized into two statistically significant clusters (Fig. 8a). The first cluster contains 31 sampling sites and is connected with natural influences

in groundwaters. It includes aquifers with a lateral feed from the consolidated carbonate and ultrabasic rocks. The second cluster comprises ten sampling sites with significant anthropogenic pollution loads. In this cluster sampling, sites with high values of EC ($684\text{--}1138\ \mu\text{S cm}^{-1}$) are included. This fact is connected with the use of fertilizers and wastewater leak. The spatial distribution of all groundwater sampling sites is characterized by the two clusters over the geologic formations of the study area. These are shown in Fig. 8b. It is obtained that the sites of cluster II are lying in villages and in areas of intensive agriculture activity, where alluvial and flysch geological formations appear. In these regions, the significant loads of factor I are also found (Fig. 7a).

Water suitability

Chemical analyses in the present study have shown that the mean total hardness values are going from > 21.6 dH. These values characterize groundwaters as hard waters. The concentrations of NO_3^- , total Fe, Ni, Cr, Cd and Zn in 29, 21, 93, 34, 24 and 10% of the samples, respectively, surpasses

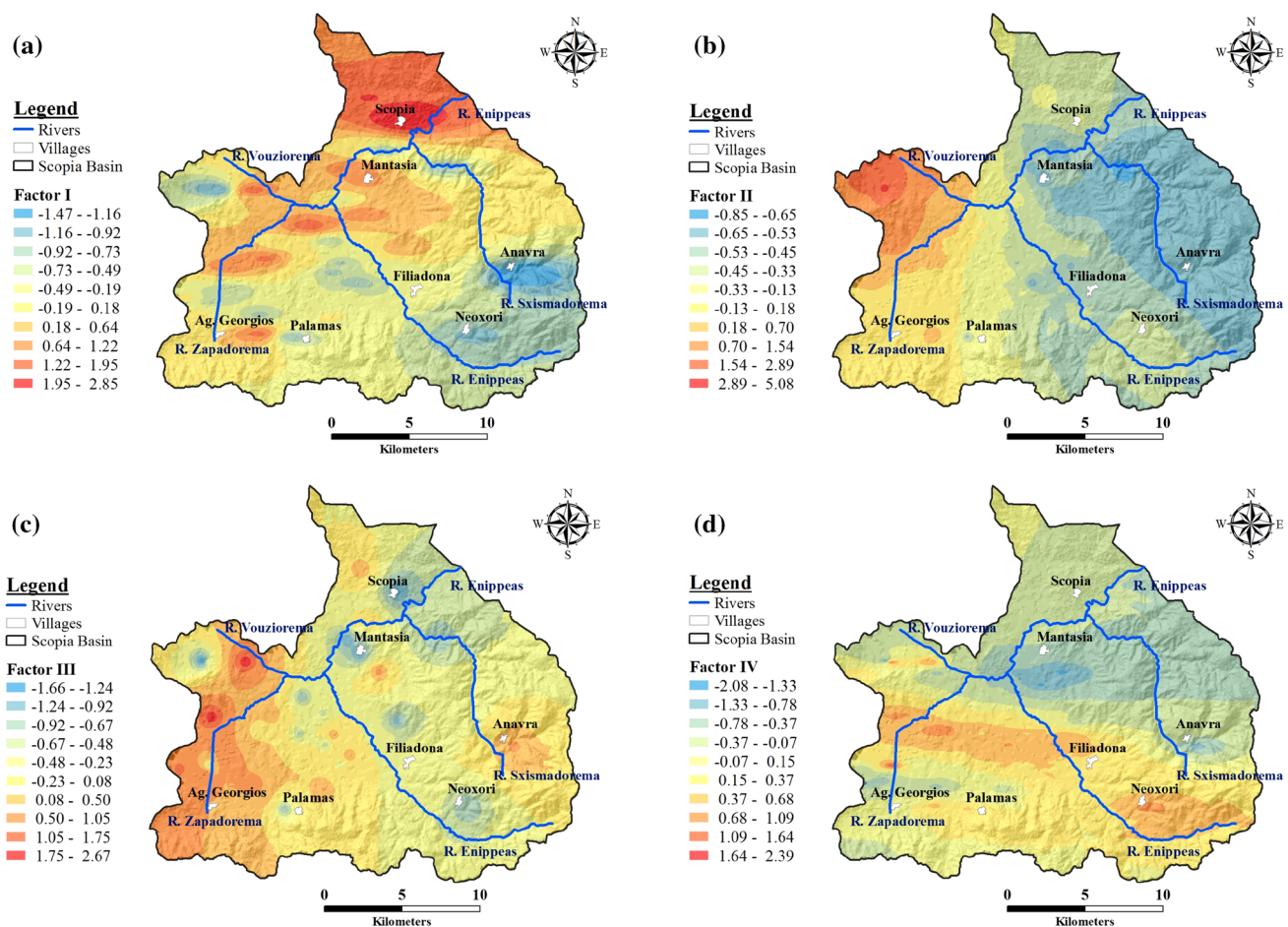
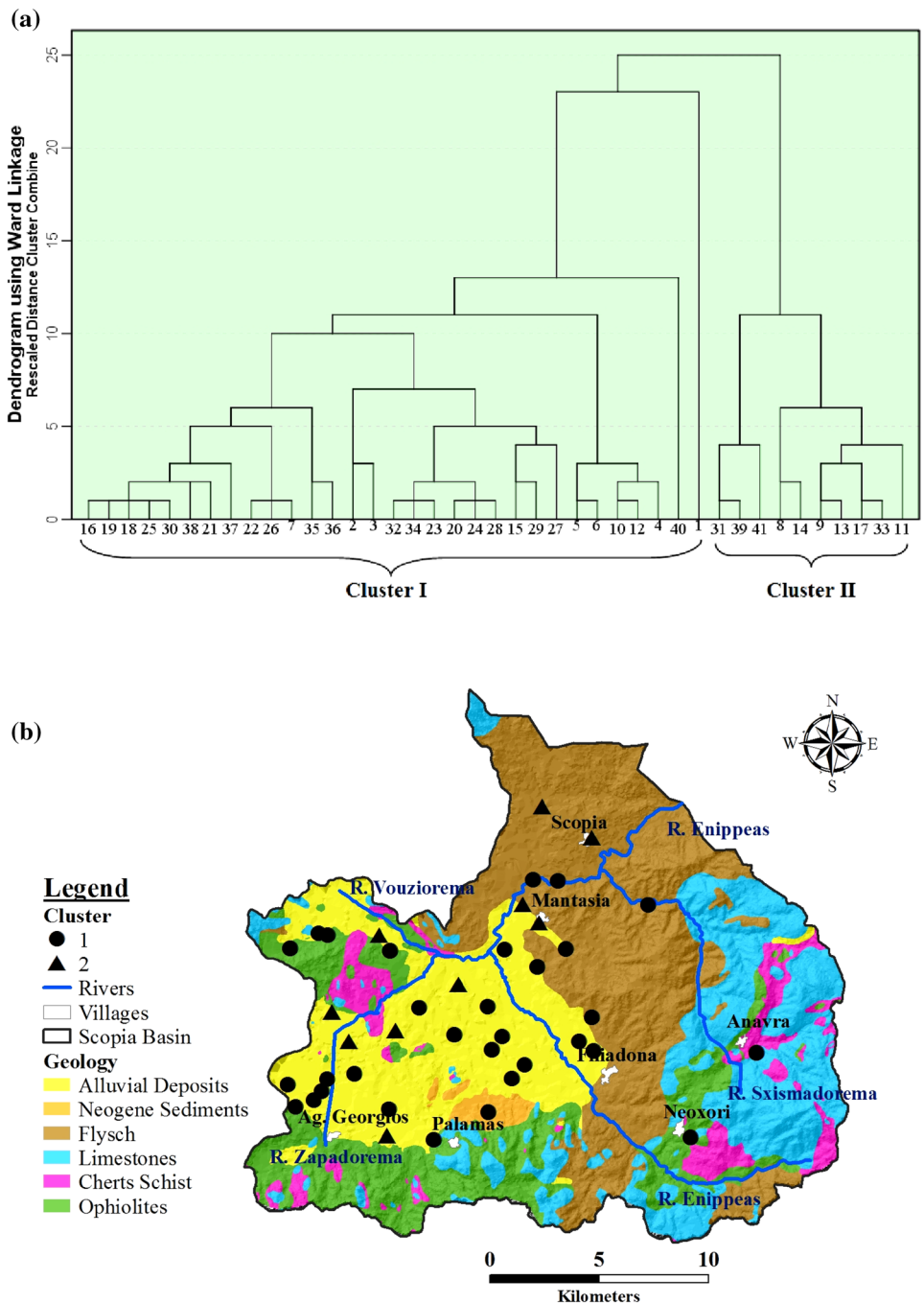


Fig. 7 Spatial distribution of factor analysis scores: **a** Factor I, **b** Factor II, **c** Factor III, **d** Factor IV

the drinking water limit of E.C. The boreholes B1, B2, B3, B24 and springs S2, S3, and S4 are utilized for water supply, required in the Scopia region, while borehole B20 is utilized for both irrigation purposes and human utilization (Charizopoulos 2013). The comparison with the EC drinking water standards demonstrated an excess of NO_3^- for B24, Mg for B1, B2, and B3, Fe for B2 and B3, Cr for B20 and B24, Ni and Pb for B1. Thus, just groundwaters from the carbonate rocks at the southern part of the region are appropriate for human utilization (Fig. 2). The suitability of groundwater

for irrigation purposes estimated with the sodium adsorption ratio (SAR) and Richards diagram (Richards 1954). SAR values range from 0.02 to 1.27. These values in combination with EC values demonstrate that the degree of restriction on the use of irrigation water ranges from none to slight or moderate, in accordance with FAO irrigation water standards (Ayers and Westcot 1994). As indicated by Richards outline, waters classified to fields C2.S1 and C3.S1 (Fig. 9a). In the field, C2.S1 is arranged most of the examples (76%), featuring waters with low sodium hazard and moderate salinity

Fig. 8 a Clusters dendrogram showing two clusters of natural and anthropogenic influences in groundwater in the study area, b Geographical location of 41 groundwater sampling sites according to their corresponding cluster over the geologic formations of the study area



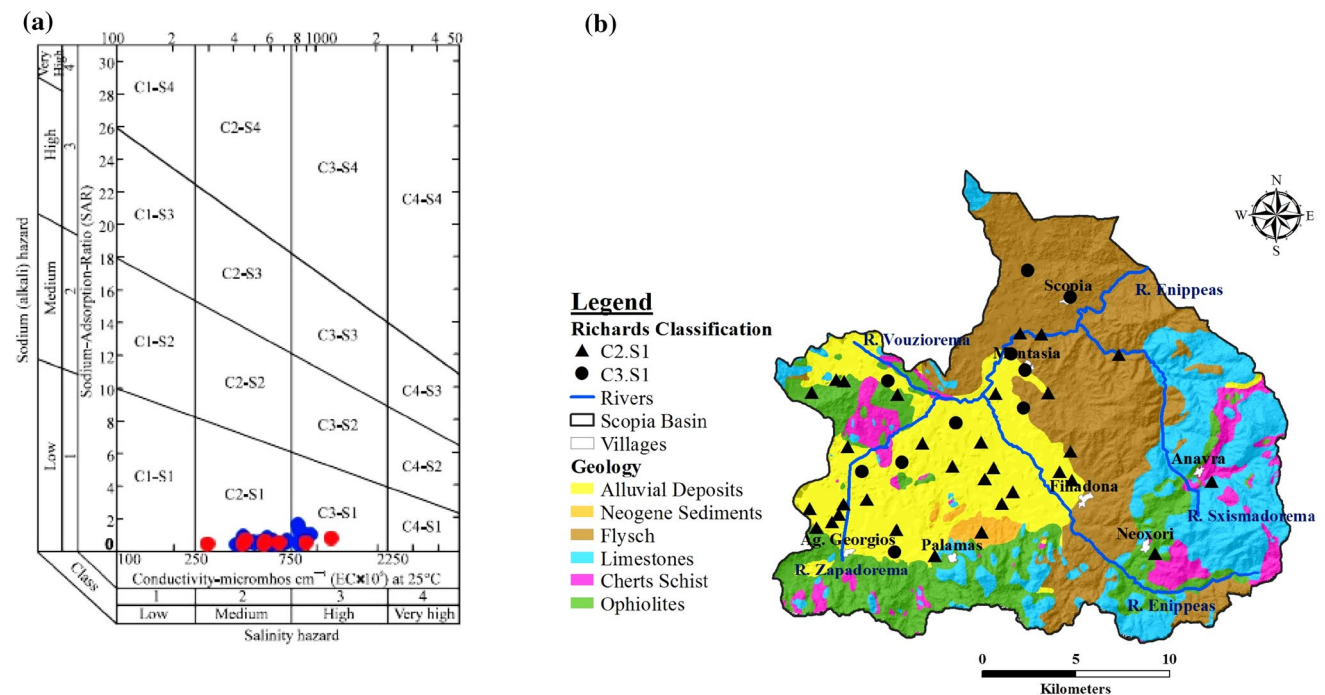


Fig. 9 **a** Diagram of irrigation water hazard classification, for samples of springs (red) and boreholes (blue) based on total salts (EC) and SAR (Richards 1954), **b** Irrigation water suitability map based on Richards classification

hazard. Water that falls in the medium salinity hazard class (C2) can be used in most cases without any special practices for salinity control (Sappa et al. 2014). In the category C3.S1 classified 24% of the samples, concerning waters with low sodium hazard and high salinity hazard. These waters are unsuitable for the cultivation of plants sensitive to salinity and the irrigation of soils with limited drainage (Fig. 9b).

Conclusions

In the present study, a new approach to IDW interpolation modeling of groundwaters has been proposed and HCA was also applied to assess the natural and anthropogenic impacts in Scopia basin groundwaters in Central Greece. The IDW parameters optimized to achieve the best-fit model of both hydrochemical parameters and PCA factors spatial distribution. The results predicted by the IDW interpolator compared with the actual values measured at the same locations based on less root-mean-squared prediction error. The goodness-of-fit criteria suggested that the finest IDW models, to understand the spatial distribution of the hydrochemical parameters, are defined by the ellipse shape type and neighborhood deviation of four sectors for TDS and NO_3^- , ellipse/eight sectors for SO_4^{2-} and Cd and ellipse/one sector for Ni. The best interpolation model for TDS demonstrates high values in the north and the southwest areas of the study site

related to urban wastewater and cultivation methods. The best-fitting model for the spatial distribution of SO_4^{2-} demonstrates high concentrations in North of the study area. It is ascribed to the oxidation of pyrites of ophiolitic origin but also to the use of sulfur–ammonium fertilizers in the cultivated plain in the study area. The best IDW predicted interpolator model presents high values of NO_3^- in specific areas where groundwaters suffered by intense cultivation and related to the application of nitrogen fertilizers. The most accurate interpolation model presents the high values of the Cd in the Southwest part of the basin. It can be credited both to the existence of the component, in the minerals contained in ophiolitic rocks and also in the use of fertilizers in the cultivated part of the basin. Furthermore, it is observed that the best-fitting, interpolation model represented by circle shape type and neighborhood deviation of one sector for Mg^{2+} and Cl^- . The model demonstrates high estimations of Cl^- in the North and Southwest margins of the basin. The goodness-of-fit criteria suggested that the most accurate model for the spatial distribution of Cr_{tot} is represented by circle shape type and neighborhood deviation of eight sectors. The presence of the element is caused by the weathering of the minerals contained in the ophiolites, which cover a big part of the study area. The best-fit models to understand the spatial distribution of the PCA factors determined by ellipse/four sectors, circle/one sector, circle/four sectors, and ellipse/one sector for factor I, factor II, factor III, and factor

IV, respectively. HCA presents two statistically significant clusters. The outcomes demonstrate that geogenic processes ascribed to the weathering of the minerals contained in the ultrabasic rocks and anthropogenic influences related to the utilization of fertilizers and wastewater leak impact the groundwater chemistry and quality. Groundwaters are unsuitable for water supply needs and suitable for irrigation purposes, presenting in their majority low sodium hazard and moderate salinity hazard.

References

- Alley WM (1993) Regional ground-water-quality. Van Nostrand Reinhold, New York
- Alloway BJ (1995) Heavy metals in soil, Second edn. Wiley, New York
- APHA (American Public Health Association) (1998) Standard Methods for the Examination of Water and Wastewater, method 3030, 20th edn. American Water Works Association, Water Environment Federation, Washington
- Appelo CAJ, Postma D (2005) Geochemistry, groundwater, and pollution, second edn. Balkema, Rotterdam
- Arif M, Hussain J, Hussain I, Kumar S, Bhati G (2014) GIS based inverse distance weighting spatial interpolation technique for fluoride occurrence in ground water. *Open Access Libr J* 1:1–6
- Asadi SS, Vuppala PM, Reddy A (2007) Remote sensing and GIS techniques for evaluation of groundwater quality in municipal corporation of Hyderabad (Zone-V), India. *Int J Environ Recour* 4(1):45–52
- Ayers RS, Westcot DW (1994) Water quality for agriculture. FAO irrigation and drainage water paper. <http://www.fao.org/DOCRP/003/T0234e/T0234E00.htm>. Accessed 1989
- Bourette C, Bertolo R, Almodovar M, Hirata R (2009) Natural occurrence of hexavalent chromium in a sedimentary aquifer in Uraina, State of Sao Paulo, Brazil. *Ann Braz Acad Sci* 81:227–242
- Cattell RB (1978) The scientific use of factor analysis in behavioral and life sciences. Plenum Press, New York
- Charizopoulos N (2013) Investigation on mechanisms of quantitative and qualitative deterioration of water and soil resources in Domokos catchment by natural and anthropogenic processes. PhD dissertation, Agriculture University of Athens. <http://dspace.aau.gr/xmlui/handle/10329/5812>. Accessed 10 Feb 2015
- Charizopoulos N, Psilovikos A (2015) Geomorphological analysis of Scopia catchment (Central Greece), using DEM data and GIS. *Fresen Environ Bull* 24(11):3973–3983
- Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J Hydrol* 353(3–4):294–313
- Davis JC (2002) Statistics and data analysis in geology. Wiley, Singapore, pp 526–540
- Demirel Z, Güler C (2006) Hydrogeochemical evolution of groundwater in a Mediterranean coastal aquifer, Mersin-Erdemlibasin (Turkey). *Enviro Geol* 49:477–487
- Dunteman GH (1989) Principal component analysis. Sage, Thousand Oaks
- Durov SA (1948) Natural waters and graphic representation of their composition. *Dokl Akad Nauk SSSR* 59:87–90
- EC (1998) Council Directive 98/83/EC, Directive of the European Parliament on the quality of water intended for human consumption. The European Parliament and the Council of the European Union, Off J L 330
- Eiswirth M, Hotzl H, Burn LS (2000) Development scenarios for sustainable urban water systems. In: Sililo et al (eds) *Groundwater: past achievements and future challenges*. Balkema, Rotterdam, pp 917–922
- Esri (2012) An Overview of the Interpolation Toolset. <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/>. Accessed 28 Sept 2012
- Ferriere J (1982) Paleogeographies et tectoniques superposees dans les Hellenides internes: les massifs de l'Othrys et du Pelion (Grece continentale). These Univ des Sciences et Techniques de Lille, S.G.N. 8, vol I
- Fortin M, Randall M, Dale T (2005) Spatial analysis: a guide for ecologists. Cambridge University Press, Cambridge
- Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs
- Ghosh T, Kanchan R (2014) Geoenvironmental appraisal of groundwater quality in Bengal alluvial tract, India: a geochemical and statistical approach. *Environ Earth Sci* 72:2475–2488
- Gong G, Mattevada S, O'Bryant Sid (2014) Comparison of the accuracy of kriging and IDW interpolations in estimating groundwater arsenic concentrations in Texas. *Environ Res* 130:59–69
- Hansen B, Laerke T, Dalgaard T, Erlandsen M (2011) Trend reversal of nitrate in Danish groundwater—a reflection of agricultural practices and nitrogen surpluses since 1950. *Environ Sci Technol* 45:228–234
- Henley S (1981) Nonparametric geostatistics. Applied Science Publishers Ltd, London
- IGME (Institute of Geology and Mineral Exploration) (2010) Recording groundwater data of East Central Greece. Third Community Support Framework (2000-06), Report series, p 156
- Jackson BM, Browne CA, Butler AP, Peach D, Wade AJ, Wheeler HS (2008) Nitrate transport in Chalk catchments: monitoring, modeling and policy implications. *Environ Sci Policy* 11(2):125–135
- James J, Choquette P (2013) Diagenesis 9. Limestones— The Meteoric Diagenetic Environment. In: Scholle PA, James NP, Read JF (eds) *Carbonate sedimentology and petrology*. American Geophysical Union, Washington, DC, pp 45–78
- Jiang Y, Wu Y, Groves C, Yuan D, Kambesis P (2009) Natural and anthropogenic factors affecting the groundwater quality in the Nandong karst underground river system in Yunan, China. *J Cont Hydr* 109(1–4):49–61
- Jolliffe IT (2002) Principal component analysis, Springer series in statistics. Springer, New York, p 487
- Ju TX, Kou CL, Zhang SF, Christie P (2006) Nitrogen balance and groundwater nitrate contamination: comparison among three intensive cropping systems on the North China Plain. *Environ Pollut* 143(1):117–125
- Kabata-Pendias A, Pendias H (2001) Trace elements in soils and plants, third edn. CRC Press, New York, p 413
- Kaiser H (1958) The varimax criterion for analytic rotation in factor analysis. *Psychom* 23(3):187–200
- Karmis P (2010) Electromagnetic geophysical survey in the area of drained Lake Xyniada. IGME, Report series, p 50
- Katsikatsos G, Migiros G, Triantaphyllis M, Mettos A (1986) Geological structure of internal Hellenides (E. Thessaly, SW Macedonia, Euboea, Attica, Northern Cyclades islands and Lesvos). IGME Geol Geophys Res Special Issue 191–212
- Kruskal JB, Landwehr JM (1983) Icicle plots: better displays for hierarchical clustering. *Am Stat* 37(2):162–168
- Kyoung-Ho K, Seong-Taek Y, Seong-Sook P, Yongsung J, Tae-Seung K (2014) Model-based clustering of hydrochemical data to demarcate natural versus human impacts on bedrock groundwater quality in rural areas, South Korea. *J Hydrol* 519:626–636
- Lambrakis N, Antonakos A, Panagopoulos G (2004) The use of multicomponent statistical analysis in hydrogeological environmental research. *Water Res* 38:1862–1872

- Lin CY, Abdullah MH, Praveena SM, Yahaya AHB, Musta B (2012) Delineation of temporal variability and governing factors influencing the spatial variability of shallow groundwater chemistry in a tropical sedimentary island. *J Hydrol* 432–433:26–42
- Lloyd J, Heathcote J (1985) Natural inorganic hydrochemistry in relation to groundwater. Clarendon Press, Oxford
- Machiwal D, Jha MK (2015) Identifying sources of groundwater contamination in a hard-rock aquifer system using multivariate statistical analyses and GIS-based geostatistical modeling techniques. *J Hydrol* 4(A):80–110
- Magesh NS, Krishnakumar S, Chandrasekar N, Soundranayagam JP (2012) Groundwater quality assessment using WQI and GIS techniques, Dindigul district, Tamil Nadu, India. *Arab J Geosci*. <https://doi.org/10.1007/s12517-012-0673-8>
- Mantzafleri N, Psilovikos AR, Mplanta A (2009) Water quality monitoring and modeling in Lake Kastoria, using GIS. Assessment and management of pollution sources. *Water Resour Manage* 23(15):3221–3254
- McArthur JM, Banerjee DM, Hudson-Edwards KA, Mishra R, Purohit R, Ravenscroft P et al (2004) Natural organic matter in sedimentary basins and its relation to arsenic in anoxic ground water: the example of West Bengal and its worldwide implications. *Appl Geochem* 19:1255–1293
- Mencio A, Mas-Pla J (2008) Assessment by multivariate analysis of groundwater–surface water interactions in urbanized Mediterranean streams. *J Hydrol* 352:355–366
- Merkel B, Planer-Friedrich B (2008) Groundwater geochemistry: a practical guide to modeling of natural and contaminated aquatic systems, second edn. Springer, Berlin
- Mills CT, Morrison JM, Goldhaber MB, Ellefsen KJ (2011) Chromium(VI) generation in vadose zone soils and alluvial sediments of the southwestern Sacramento Valley, CA: a potential source of geogenic Cr(VI) to groundwater. *Appl Geochem* 26:1488–1501
- Monjerezi M, Vogt R, Aagaard P, Saka J (2011) Hydro-geochemical processes in an area with saline groundwater in lower Shire River valley, Malawi: an integrated application of hierarchical cluster and principal component analyses. *Appl Geochem* 26:1399–1413
- Mountrakis D (1985) Geology of Greece. University Studio Press, Thessaloniki, p 205
- Mull R, Harig F, Pielke M, Clark L, Hartwell D, Beraud JF, Harris RC (1992) Groundwater management in the urban area of Hanover, Germany. *J Inst Water Environ Manage* 6:199–206
- Nguyen TT, Kawamura A, Tong TN, Nakagawa N, Amaguchi H, Gilbuena RJr (2015) Clustering spatio–seasonal hydrogeochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam. *J Hydrol* 522:661–673
- Norusis M (2011) IBM SPSS statistics 19 guide to data analysis. Pearson Publishers, London, p 672
- Olmez I, Beal J, Villaume J (1994) A new approach to understanding multiple-source groundwater contamination: factor Analysis and chemical mass balance. *Water Res* 28(5):1095–1101
- Otto M (1998) Multivariate methods. In: Kellner R, Mermet JM, Otto M, Widmer HM (eds) Analytical chemistry. Wiley-VCH, Weinheim, p 916
- Parkhurst DL, Appelo CAJ (2013) Description of input and examples for PHREEQC version 3—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. United States Geological Survey. <http://pubs.usgs.gov/tm/06/a43/>. Accessed 23 Jan 2013
- Philip GM, Watson DF (1982) A precise method for determining contoured surfaces. *Austral Petr Expl Assoc J* 22:205–212
- Reed M, Marine R (1991) Quality control of chemical and isotopic analyses of geothermal water samples. In: Proceedings of the 16th Workshop on Geothermal Reservoir Engineering. Stanford University, California, SGP-TR-134, pp 9–13
- Richards LA (1954) Diagnosis and improvement of saline and alkali soils. USDA Agricultural Handbook No 60. US Department of Agriculture, Washington DC, p 160
- Sappa G, Ergul S, Ferranti F (2014) Water quality assessment of carbonate aquifers in southern Latium region, Central Italy: a case study for irrigation and drinking purposes. *Appl Water Sci* 4:115–128
- Sikora LG, Bent MG, Corey RB, Keeney DR (1976) Septic nitrogen and phosphorus removal test system. *Groundwater* 14(5):309–314
- Stollenwerk KG (1996) Simulation of phosphate transport in sewage-contaminated groundwater Cape Cod Massachusetts. *Appl Geochem* 11(1–2):17–324
- Tisdal S, Nelsson W (1975) Soil fertility and fertilizers. Macmillan Publishing Co Inc, New York, p 240
- Tomczak M (1998) Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW)—cross-validation/Jackknife approach. *J Geogr Inf Decis Anal* 2(2):18–30
- Trewartha GT, Horn LH (1980) An introduction to climate, 5th edn. McGraw-Hill Inc, New York
- Vega M, Pardo R, Barrado E, Deban L (1998) Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res* 32:3581–3592
- Voudouris K (2009) Hydrogeology of Environment. Tziola press, Thessaloniki, 460
- Wakida F, Lerner D (2005) Non-agricultural sources of groundwater nitrate: a review and case study. *Water Res* 39:3–16
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236–244
- Watson DF, Philip GM (1985) A Refinement of Inverse Distance Weighted interpolation. *Geoprocessing* 2:315–327
- Wilford J, Caritat P, Bui E (2016) Predictive geochemical mapping using environmental correlation. *Appl Geochem* 66:275–288
- Xypolias P, Spanos D, Chatzaras V, Kokkalas S, Koukouvelas I (2010) Vorticity of flow in ductile thrust zones: examples from the Attico–Cycladic Massif (Internal Hellenides, Greece). *Geol Soc Lond Spec Publ* 335:687–714
- Yidana SM (2010) Groundwater classification using multivariate statistical methods: Southern Ghana. *J Afr Earth Sci* 57:455–469
- Ypen (Ministry of Environment and Energy) (2012) Management plans and river basin water districts Thessaly, Epirus and western mainland Greece under Directive 2006/118/EC. Thessaly Water District: 10th Report series, p 234
- Zisou C, Psilovikos A (2012) Monitoring and Modeling of the Physicochemical Parameters Geographic Distribution in Lake Karla using GIS. In: 2nd Hellenic Conference of “EYE–EEDYP”, pp 1250–1261