Available online at www.sciencedirect.com
ScienceDirect
Chemometrics and Intelligent Laboratory Systems 86 (2007) 86 – 94
Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

Characterization of Turkish bottled waters using pattern recognition methods

Cüneyt Güler *

Mersin Üniversitesi, Çiftlikköy Kampüsü, Jeoloji Mühendisliği Bölümü, 33343 Mersin, Turkey

Received 22 August 2006; accepted 23 August 2006
Available online 29 September 2006

Abstract

In this study, 130 (out of total 229 certified) Turkish bottled water brands were characterized by using multivariate pattern recognition methods. The database used in this study was compiled from the production licenses issued to bottlers by the Turkish Ministry of Health, which is the regulating body of the bottled water industry since 1997. The production licenses provided information on up to 34 physico-chemical parameters and were a valuable information source for this study. The relationships among eight selected major ion chemistry variables (calcium, magnesium, sodium, potassium, chloride, sulfate, bicarbonate, and fluoride) were examined by principal components analysis and hierarchical cluster analysis. In general, the concentrations of most water quality constituents reported were much higher in the natural mineral water brands compared to the natural spring water and drinking water brands of bottled water. This study has also shown that Turkish bottled waters have a diverse character reflected by their chemical compositions, which in turn is a result of country's peculiar geologic framework.

Keywords: Bottled water; Mineral water; Pattern recognition; Principal components analysis; Hierarchical cluster analysis; Turkey

1. Introduction

Turkey is geographically located on the "Alpine–Himalayan orogenic belt" where continuing tectonic activities not only resulted in the formation of country's peculiar geologic framework but also resulted in the formation of numerous cold and hot water springs. Ground water and water from springs are the sources used for bottling water in the country. However, at present, only 20% of the existing springs are utilized by the Turkish bottled water industry. Growing population and the population shift from rural to urban areas have increased the consumption of bottled water world wide [1]. Nowadays, most people living in urban environments prefer bottled water because it is associated with naturalness [2], objection to unpleasant tastes and odors from municipal water supplies [3], and is often regarded as safer and healthier than tap water [4]. There is also a common belief that natural (spring) waters have beneficial medicinal and therapeutic effects [5]. Archeological findings suggest that utilization of spring waters for therapeutic purposes goes as far back as to Phrygians, who lived in Kütahya and Eskisehir region in Turkey between 1200 and 700 BC. The Anatolia (a.k.a. Asia Minor), being strategically located on the ancient east–west trade route called the Silk Road, was the cradle of very important civilizations in the history. Many springs in the Anatolia had been utilized for drinking and/or therapeutic purposes during the longstanding reigns of Hittite (1900–1193 BC), Persian (546–333 BC), Roman (27 BC–476 AD), Byzantine (330–1453 AD), Seljuk (1055–1243 AD), and Ottoman (1299–1922 AD) Empires and they are still in use today.

Today, there are 229 domestic brands of bottled water (excluding fruit flavored bottled waters) recognized by the Turkish Ministry of Health, which is the regulating body of the bottled water industry since 1997. The Ministry of Health distinguishes between several types of water according to their sources (the point of emergence) and physical/chemical properties, each with its own definition. In Turkey, commercialization of surface waters (e.g. lake, reservoir, and river waters) is prohibited, in other words, surface waters cannot be bottled and sold for human consumption. Turkish bottled waters are divided into several classes as follows: (1) natural spring water; (2) drinking water; (3) processed drinking water; and (4) natural mineral water. Natural spring, drinking, and processed drinking water types are still (uncarbonated) waters; whereas natural mineral waters are sparkling (carbonated) waters and they are bottled from sources carbonated naturally. Still water is the most consumed soft drink

* Tel.: +90 324 361 0001x7314; fax: +90 324 361 0032.
E-mail address: cguler@mersin.edu.tr.

0169-7439/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2006.08.009

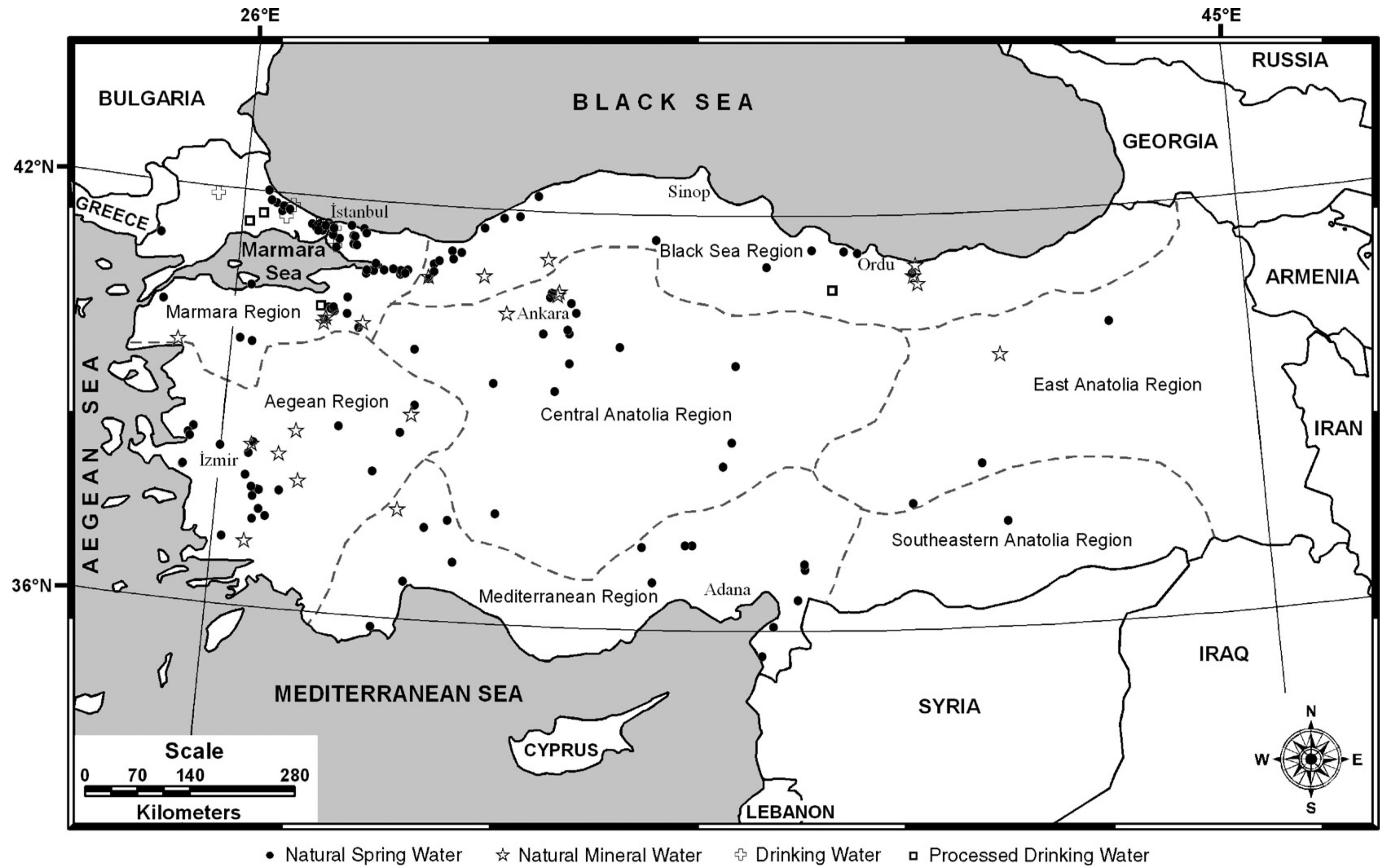

Fig. 1. Location of springs and wells utilized for bottling various type of waters in Turkey.

with meals in Turkey, and it became a substitute for tap water especially in urban households. However, natural mineral water is generally perceived by Turkish people as a product having a medicinal value due to its high dissolved mineral content.

According to Turkish legislation [6] natural spring water must be derived from an underground formation with favorable geologic conditions, from which water flows naturally and constantly to the surface of the earth. Drinking water must be derived from an underground formation with favorable geologic conditions, from which water flows naturally to the surface of the earth or extracted by a method (generally drilling a borehole) approved by the Ministry of Health [6]. Natural spring water and drinking water types may be subjected to disinfection processes such as filtration, ozonation, and UV (ultraviolet) irradiation. Processed drinking water is extracted from an unpolluted aquifer by a method approved by the Ministry of Health and this type of water may be subject to disinfection (e.g. with ozone and/or UV irradiation), filtration (e.g. with activated carbon), purification (e.g. with reverse osmosis) or other suitable processes and its physical/chemical properties may be changed by removal and/or addition of minerals [6]. Natural mineral water must be derived from an underground formation with favorable geologic conditions, contains at least 1000 mg/L dissolved minerals and/or trace elements (water may be hot or cold at source). This type of water also naturally contains carbon dioxide (CO_2) gas and radioactive elements and flows naturally to the surface of the earth or extracted by an approved method. Natural mineral waters can only be disinfected by using UV irradiation and must have clinically proven health benefit effects determined by the Ministry of Health [6].

The Turkish bottled water market has grown quickly since the 1990s. Beginning with the period 1992–1994, especially major cities started to experience a severe drinking water problem, which have dramatically increased bottled water consumption in Turkey. The reasons for the problem were threefold: (1) insufficient water supply from city water distribution networks; (2) problems in the taste, purity, color, and odor of tap water; and (3) erratic power supply. In order to solve the problem, so-called “water stations” were opened where people were able to buy water in bulk. This presented a very big opportunity for companies which sell bottled water and resulted in a sharp increase in sales. Nevertheless, the health and hygiene conditions in the water stations were not up to the necessary standards, and sales from water stations were therefore banned in October 1997 by a government decree to regulate and control sales of bottled water. As a result, consumers responded and started to buy well-known brands, therefore intensifying competition. Due to the public distrust to municipal tap water for aforementioned reasons, in Turkey, bottled water has become a US$500 million industry with a per capita consumption of approximately 78 l [7].

The bottled water industry in Turkey is very dynamic and competitive market, with key players focused on natural spring water, which translated into an industry with regional brands tied to local spring water sources. They have the infrastructure to benefit from cost-effective nationwide distribution with strategically located production plants (Fig. 1). Smaller companies are mostly century-old and family/municipality-owned, whereas larger firms are under control of major multinational soft drink companies. Currently, of the 229 registered bottlers in Turkey, about 52% are located in Marmara region, which also accounts for almost half of bottled water consumption (Fig. 1). The second leading region is Aegean, which is less industrialized than Marmara region, but is a main tourist resort/attraction. Tourism places severe seasonal pressures on water resources at the local level especially in this part of the country. For this reason, Aegean region is also important for water producers (Fig. 1).

The purpose of this paper is to investigate the physico-chemical characteristics of domestic brands of bottled waters sold in Turkey and to classify them by utilizing parameters reported on their government issued production licenses. In the present study, 130 domestic brands of bottled water were characterized using multivariate pattern recognition methods including principal components analysis (PCA) and hierarchical cluster analysis (HCA).

2. Materials and methods

2.1. Bottled water database

All 229 registered bottled water manufacturers in Turkey were contacted by telephone and/or electronic mail to obtain a copy of their production licenses, which reports on up to 34 physico-chemical parameters measured in bottled water. Where available, production licenses were also obtained from the company websites found on the Internet. In total, information about 189 brands was obtained, which represent 82.5% of the domestic brands of bottled water currently sold in Turkish market. Information about 40 (out of 229) brands were not available by no means, which are all small-scale local bottled water producers with a total share in the market less than 2%. A further evaluation of the database revealed that production licenses of 59 brands (out of 189) have not reported on some of the major ion parameters that were utilized in this study. As a result, those 59 brands were excluded from the multivariate analysis but included for the descriptive statistical calculations (i.e. mean, min., max., etc.). To keep the brand names anonymous they were numbered from 1 through 130 and this convention used throughout the text and in the multivariate analyses. The data were arranged in rows for brands and in columns for measured physico-chemical parameters. Bottled water dataset can be downloaded from the ftp site via following link: ftp://ftp.clarkson.edu/pub/hopkepk/Chemdata/Turkish_water_data. Of the 34 physico-chemical variables (consisting of major ions, minor ions, trace elements, and physical parameters) in the compiled database, eight variables (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-}, $SO_4^{=}$, $HCO_3^{=}$, and F^{-}) occur most often and were thus utilized by the multivariate pattern recognition methods. Analysis of chemical and physical properties of these bottled waters were carried out by official laboratories that have been certified by the Ministry of Health and accuracy and precision of the laboratory results were not questioned in this study. However, as an independent check on the quality of the chemical analyses in the database they were tested for charge balance error [8]:

$$\%\ \text{Charge Balance Error} = \frac{\sum z \cdot m_c - \sum z \cdot m_a}{\sum z \cdot m_c + \sum z \cdot m_a} \times 100 \qquad (1)$$

In Eq. (1), z is the absolute value of the ionic valence, m_c the molality of cationic species and m_a the molality of the anionic species. Calculated charge balance errors are less than ±10% for all the samples in the database, which is an acceptable error for the purpose of this study.

2.2. Pattern recognition methods

In this study, Turkish bottled water brands were characterized by means of univariate and multivariate analysis techniques using the parameters reported on their government issued production licenses. Pattern recognition methods can provide a powerful tool for analyzing a large amount of water chemistry data [9]. These methods can be used to test water quality data and determine if samples can be grouped into distinct populations that may be significant from a statistical point of view. In this study, principal components analysis (PCA) and hierarchical cluster analysis (HCA) were used to classify the bottled water brands into distinct groups based on their (chemical) similarity. The Statistica® Release 5.5 [10] commercial software package was utilized for the multivariate analysis techniques performed. In addition to the multivariate methods, descriptive statistics and correlation were also analyzed.

The data used in this study were universally skewed positively; the data contained small numbers of high values. This result was expected, since most naturally occurring element distributions follow this pattern [11]. The data were log-transformed (base 10) so that they more nearly corresponded to normally distributed data. Then, all the eight variables were standardized by calculating their standard scores (z-scores) as follows:

$$z_i = \frac{x_i - \mu}{s} \quad (2)$$

In Eq. (2), the z_i is the standard score of the sample i, x_i the value of sample i, μ the mean, and s the standard deviation. Besides normalizing and reducing outliers, these transformations also tend to homogenize the variance of the distribution [12]. Standardization scales the log-transformed data to a range of approximately −3 to +3 standard deviations, centered about a mean of zero and gives a standard deviation of unity. In this way, each variable has equal weight in the statistical analyses. Without scaling, the results are influenced most strongly by the variable with the greatest magnitude [13].

2.2.1. Correlation analysis

Correlation analysis was applied to describe the degree of relation between two water chemistry parameters. Correlation is the ratio of the covariance of two variables to the product of their standard deviations [14]. The resulting correlation coefficient is a unitless number that ranges between −1.0 and +1.0. A value of −1.0 represents a perfect inverse relationship between the two variables, whereas +1.0 occurs when the two variables react in exactly the same way as their values change. A correlation coefficient of zero suggests that the two variables are independent of each other. In this study, "Pearson r correlation" was used and the results of the correlation analysis are presented in the following sections.

2.2.2. Principal components analysis

As a pattern recognition method, PCA reduces a large number of variables (measured parameters in water samples) to a small number of variables (principal components or PCs) [15]. More concisely, PCA linearly combines two or more correlated variables into one variable. This approach has been used to extract related variables and infer the processes that control water chemistry [16]. The PCA-defined new variables can then be displayed in a scatter diagram, presenting the individual samples (bottled water brands) as points in a lower-dimensional (generally 2-D or 3-D) space. PCA is useful for data reduction and to assess the continuity/overlap of clusters or clustering/similarities in the data. Varimax normalized rotation was applied to the PCs in order to reduce the contribution of variables with minor significance. This rotation is called Varimax because the goal is to maximize the variance (variability) of the "new" variable, while minimizing the variance around the new variable [17]. The number of PCs extracted (to explain the underlying data structure) is defined by using the "Kaiser criterion" [18] where only the PCs with eigenvalues greater than unity are retained. In other words, unless a PC extracts at least as much information as the equivalent of one original variable, it is dropped [17].

2.2.3. Hierarchical cluster analysis

HCA was used to determine if the bottled water brands can be grouped into statistically distinct groups (clusters). The assumptions of cluster analysis techniques include homoscedasticity (equal variance) and normal distribution of the variables [19]. Equal weighing of all variables requires the standardization (z-scores) of the data. Comparisons based on multiple parameters from different samples are made and the samples grouped according to their "similarity" to each other. The classification of samples according to their parameters is termed Q-mode classification. In this study, Q-mode HCA was used to classify the samples into distinct groups. This approach is commonly applied to water chemistry data in order to define groups of samples that have similar chemical and physical characteristics. The linkage rule used here is Ward's method [20]. Linkage rules iteratively link nearby points (similar samples) by using the similarity (distance) matrix. The initial cluster is formed by linkage of the two samples with the greatest similarity. Ward's method is distinct from all other methods because it uses an analysis of variance (ANOVA) approach to evaluate the distances between clusters and forms smaller distinct clusters than those formed by other methods [17]. The Euclidian distance was selected as the similarity measurement, which is straight line distance between two points in c-dimensional space defined by c number of variables. Classification results of HCA are generally presented in a graphical form called "dendogram".

3. Results and discussion

3.1. Physico-chemical characteristics of bottled waters

The mean, standard deviation (±S.D.), minimum (Min.), and maximum (Max.) values of the 33 physico-chemical parameters of 189 bottled water brands were determined and presented in Table 1.

Table 1
Descriptive statistics for the water quality parameters for each of the water classes

Parameter	Still (uncarbonated) waters									Sparkling (carbonated) waters		
	Natural spring water (149 brands)			Drinking water (12 brands)			Processed drinking water (5 brands)			Natural mineral water (23 brands)		
	n [a]	Mean±S.D.	Min.–Max.	*n* [a]	Mean±S.D.	Min.–Max.	*n* [a]	Mean±S.D.	Min.–Max.	*n* [a]	Mean±S.D.	Min.–Max.
Calcium (mg/l)	122	16.82±15.09	0.48–97.1	11	11.65±5.94	6–25	5	2.73±3.09	0.8–8.16	23	179.78±119.2	46.1–420.8
Magnesium (mg/l)	123	3.68±3.41	0.012–17.4	11	3.69±2.29	0.9–8.5	5	1.66±1.91	0.07–4.86	23	92.26±108.9	9.24–447.3
Sodium (mg/l)	110	6.94±6.89	0.04–30	11	22.94±14.43	7.3–50	–	–	–	19	390.29±493.2	1.61–1769.9
Potassium (mg/l)	115	1.81±2	0.005–10	11	4.03±3.16	1.3–11.2	–	–	–	17	11.05±22.25	1–77
Chloride (mg/l)	145	12.21±8.18	0.15–40	12	32.12±18.96	10.7–73.8	5	10.59±6.09	2.94–19	20	119.58±182.3	7.3–759.7
Sulfate (mg/l)	149	9.25±7.96	1–35.85	12	13.63±9.32	2.4–31	5	8.87±5.64	2.3–15.2	20	105.06±155.9	1–702.4
Bicarbonate (mg/l)	103	68.92±59.04	1–291.65	10	57.7±55.91	7.65–204.2	–	–	–	20	1725.8±979.4	251.3–4026
Fluoride (mg/l)	147	0.16±0.16	0.006–0.9	12	0.26±0.3	0–1	4	0.21±0.33	0.05–0.7	18	0.58±0.49	0.03–1.52
Aluminum (μg/l)	71	26.5±45.71	2–380	7	23.43±19.1	3–50	–	–	–	8	329.56±877.1	4–2500
Antimony (μg/l)	43	3.37±1.95	0.5–5	6	5±0	5	–	–	–	3	3.67±2.31	1–5
Arsenic (μg/l)	50	7.5±3.58	1–10	6	7±4.65	1–10	–	–	–	5	8.8±2.68	4–10
Barium (μg/l)	–	–	–	–	–	–	–	–	–	6	183.83±399.9	16–1000
Boron (μg/l)	60	136.5±203.2	4–1000	8	107.5±84.64	10–300	–	–	–	8	4306.3±7062	160–21380
Cadmium (μg/l)	43	3.39±1.52	0.05–5	6	3.67±1.03	3–5	–	–	–	5	2.48±1.16	0.4–3
Chromium(μg/l)	49	16.98±8.35	0.8–50	7	16.86±16.99	2–50	–	–	–	5	15.4±7.8	2–20
Copper (μg/l)	50	16.56±9.91	2–60	7	18.57±3.78	10–20	–	–	–	6	15.5±7.31	3–20
Cyanide (μg/l)	45	3.05±3.06	0.1–10	6	4.67±4.13	2–10	–	–	–	5	1.8±0.45	1–2
Iron (μg/l)	87	60.56±327.8	4–3080	8	18.75±3.54	10–20	–	–	–	19	803.1±1422.3	20–4500
Lead (μg/l)	42	8.54±2.83	1–10	6	10±0	10	–	–	–	6	8.13±3.21	2.2–10
Manganese (μg/l)	55	16.18±8.43	2–62	6	18.33±4.08	10–20	–	–	–	8	232.6±283.2	20–900
Mercury (μg/l)	38	0.85±0.34	0.05–1	6	0.92±0.2	0.5–1	–	–	–	4	1±0	1
Nickel (μg/l)	56	15.73±6.77	2–20	6	30.33±34.88	2–100	–	–	–	7	13.71±8.2	2–20
Selenium (μg/l)	39	8.63±2.64	1–10	6	9.17±2.04	5–10	–	–	–	4	20±20	10–50
Zinc (μg/l)	68	36.24±33.62	3.5–170	9	24.44±10.14	20–50	–	–	–	8	197.3±336.2	6–980
Spring Discharge (l/s)	128	16.8±141.6	0.041–1600	10	4.4±3.3	1.2–12.6	1	1.4±0	1.4	18	2.4±2.8	0.4–10
pH (Standard Units)	147	7.15±0.64	5.6–8.2	12	6.93±0.74	6.04–8.2	5	6.78±0.81	5.75–8.02	19	6.36±0.6	5.3–7.9
Ammonium (mg/l)	120	N.D.	N.D.	11	N.D.	N.D.	5	N.D.	N.D.	7	0.22±0.16	0.01–0.44
Nitrite (mg/l)	120	N.D.	N.D.	11	N.D.	N.D.	5	N.D.	N.D.	6	0.02±0.02	0.006–0.05
Nitrate (mg/l)	117	4.5±4.32	0.05–19.4	11	1.84±2.13	1–7.92	5	1.77±1.72	1–4.84	18	6.43±6.51	0.1–20
Phenolics (μg/l)	3	267±461.6	0.01–800	11	N.D.	N.D.	5	N.D.	N.D.	1	0.0001±0	0.0001
PAHs (μg/l)	36	6.24±7.04	0.00008–19.6	11	N.D.	N.D.	–	–	–	3	8.55±1.15	7.78–9.87
α-Activity (pCi/l)	98	0.73±0.66	0.054–6.14	11	0.67±0.31	0.22–1.21	–	–	–	9	2.54±3.21	0.11–10.65
β-Activity (pCi/l)	98	1.89±1.75	0.054–10.1	11	1.53±0.64	0.54–2.73	–	–	–	9	6.45±5.53	0.54–16.38

Parameters were reported on the production licenses of 189 Turkish bottled water brands.
N.D. = Not detected.
[a] Number of samples.

All elements in bottled water brands from Turkey presented a wide spread in composition. For most elements the difference between the lowest and the highest elemental concentration was one to three orders of magnitude. These large variations can be attributed to the different geological origins of the waters. The physico-chemical composition of natural waters is controlled by many factors that include chemistry of atmospheric precipitation, mineralogy of the rocks encountered along the underground flow path, residence time of the groundwater in the aquifer, climate and topography of the area [9]. These factors combine to create diverse water types that change spatially and temporally. Of all the water types studied, natural mineral water showed the highest mean concentration for the chemical constituents reported on the production licenses (Table 1). A comparison of the mean values of the elements in the natural spring and natural mineral waters showed that the concentrations of Ca, Cl, SO_4, Al, Fe, and Mn are approximately 10 times higher in the natural mineral waters compared to natural spring waters (Table 1). Whereas, concentrations of Mg, Na, and B were 25 to 60 times higher in the natural mineral waters compared to natural spring waters. This difference in mean elemental concentrations is expected since natural mineral waters are mostly produced from areas close to geothermal regions with deep ground water circulation patterns and recent tectonic/volcanic activity. However, only small differences were observed in the mean elemental concentrations of As, Cd, Cr, Cu, Ni, Pb, and Sb. Turkish bottled waters also contain several α- and β-emitting isotopes in widely varying concentrations and elevated concentrations were mostly restricted to natural mineral waters. Mean gross α- and β-activities of the natural mineral waters were almost four times higher than all other water types (Table 1). The pH values of the natural mineral waters are generally lower (mean pH = 6.36) than any other water type because they naturally contain high levels of carbon dioxide (CO_2) gas. The minimum pH values and maximum concentrations of the major elements were generally observed for the natural mineral water samples. The minimum pH value of 5.3 was observed for the brands 120 and 122. The maximum concentrations of the major elements (in mg/L) were 420.8 for calcium, 447.3 for magnesium, 1769.9 for sodium, 77 for potassium, 759.7 for chloride, 702.4 for sulfate, 4026 for bicarbonate, and 1.52 for fluoride (Table 1).

To get an insight into natural diversity in the composition of mineral waters the types were identified for all bottled water brands (Table 2). The type of water is defined by all ionic constituents that contribute at least 20% milliequivalent to the total anionic or cationic composition of water, where total equivalents of cations and anions were recorded as 100%. In order to present the composition of water, cations or anions are listed in the decreasing order, starting with the most prevailing one.

Of the total 48 different water types identified, the most frequently observed water type is Ca–Mg–HCO_3, which characterizes 22 natural spring water brands (Table 2). However, there are no two drinking water brands that were classified in the same water type. As Table 2 shows a total of 26 bottled water brands were classified in unique water types ranging from Ca–Mg–Cl–HCO_3 to Na–Mg–Ca–Cl–HCO_3. The type of water does not give detailed information about the composition of a water sample. Waters of the same type can still have a significantly different composition concerning the major constituents. In order to perform a comparison between different bottled water types, main components (sodium, potassium, calcium, magnesium, chloride, sulfate, and bicarbonate) of the 130 bottled waters are plotted on the Piper diagram [21]. The diagram displays the relative concentrations of the major cations and anions on two separate trilinear plots, together with a central diamond plot where the points from the two trilinear plots are projected (Fig. 2). As Fig. 2 shows the data are broadly distributed rather than forming distinct clusters. Employing the water classification scheme of Back and Hanshaw [22], the samples are classified into a variety of water types. This diagram provides little information that allows us to discriminate between separate groups (types or clusters) of samples. Therefore, multivariate pattern recognition methods are utilized to sort bottled water brands into groups or clusters.

3.2. Multivariate analysis

Associations among variables can be demonstrated statistically by correlation analysis. In correlation analysis, correlation coefficients are calculated for all possible pairs of variables. The Pearson r correlation with a pre-established significance level of 5% (0.05) showed that high significant correlation coefficients exist between Na and Cl (r=0.92), Na and HCO_3 (r=0.85), Na and K (r=0.78), Cl and K (r=0.78), Mg and SO_4 (r=0.74), and between Cl and HCO_3 (r=0.72). However, no significant correlations (at 5% level) were observed for parameters pairs including K and Ca (r=0.13), K and Mg (r=0.11), Cl and Mg (r=0.14), Cl and SO_4 (r=0.07), Na and SO_4 (r=0.15), and K and SO_4 (r=0.09). High correlations between the variable pairs are also desired for the application of pattern recognition methods used in this study, especially for PCA. Bottled water composition was studied by means of both PCA and HCA.

PCA technique was used to reduce the number of dimensions present in the data matrix (reducing 8 variables to 2 PCs in this

Table 2
Forty-eight different water types (or hydrochemical facies) of 130 Turkish bottled water brands

No.	Water type	Count			No.	Water Type	Count		
		NSW	NMW	DW			NSW	NMW	DW
1	Ca–HCO_3	12	4	–	26	Ca–Na–SO_4–HCO_3–Cl	1	–	–
2	Ca–HCO_3–Cl	5	–	–	27	Mg–Ca–HCO_3	1	1	–
3	Ca–HCO_3–SO_4	3	–	–	28	Mg–Ca–HCO_3–Cl	1	–	–
4	Ca–Mg–Cl	4	–	–	29	Mg–Ca–Na–HCO_3	2	–	–
5	Ca–Mg–Cl–HCO_3	1	–	–	30	Mg–HCO_3	–	1	–
6	Ca–Mg–Cl–SO_4	1	–	–	31	Mg–HCO_3–SO_4	–	1	–
7	Ca–Mg–HCO_3	22	–	–	32	Na–Ca–Cl	6	–	1
8	Ca–Mg–HCO_3–Cl	1	–	–	33	Na–Ca–Cl–HCO_3	4	–	1
9	Ca–Mg–HCO_3–SO_4	2	1	–	34	Na–Ca–Cl–HCO_3–SO_4	1	–	–
10	Ca–Mg–HCO_3–SO_4–Cl	1	–	–	35	Na–Ca–Cl–SO_4	1	–	–
11	Ca–Mg–Na–Cl–SO_4	2	–	–	36	Na–Ca–HCO_3	1	1	–
12	Ca–Mg–Na–HCO_3	–	1	–	37	Na–Ca–HCO_3–Cl	2	–	–
13	Ca–Mg–Na–HCO_3–Cl	2	1	–	38	Na–Ca–HCO_3–Cl–SO_4	–	–	1
14	Ca–Mg–SO_4–Cl	1	–	–	39	Na–Ca–HCO_3–SO_4	1	–	–
15	Ca–Na–Cl–HCO_3	2	–	1	40	Na–Ca–Mg–Cl–SO_4–HCO_3	–	–	1
16	Ca–Na–HCO_3	4	–	–	41	Na–Ca–Mg–HCO_3	–	2	–
17	Ca–Na–HCO_3–Cl	4	–	1	42	Na–Ca–Mg–HCO_3–Cl	2	–	–
18	Ca–Na–HCO_3–Cl–SO_4	1	–	–	43	Na–Ca–Mg–HCO_3–Cl–SO_4	1	–	–
19	Ca–Na–HCO_3–SO_4	1	–	–	44	Na–Cl	–	–	1
20	Ca–Na–Mg–Cl	1	–	–	45	Na–Cl–HCO_3	1	–	–
21	Ca–Na–Mg–Cl–HCO_3	1	–	–	46	Na–HCO_3	–	1	–
22	Ca–Na–Mg–HCO_3	2	1	–	47	Na–HCO_3–Cl	1	2	1
23	Ca–Na–Mg–HCO_3–Cl	2	–	1	48	Na–Mg–Ca–Cl–HCO_3	–	–	1
24	Ca–Na–Mg–HCO_3–SO_4–Cl	1	–	–					
25	Ca–Na–SO_4–Cl–HCO_3	1	–	–					

NSW=Natural spring water, NMW=Natural mineral water, DW=Drinking water.

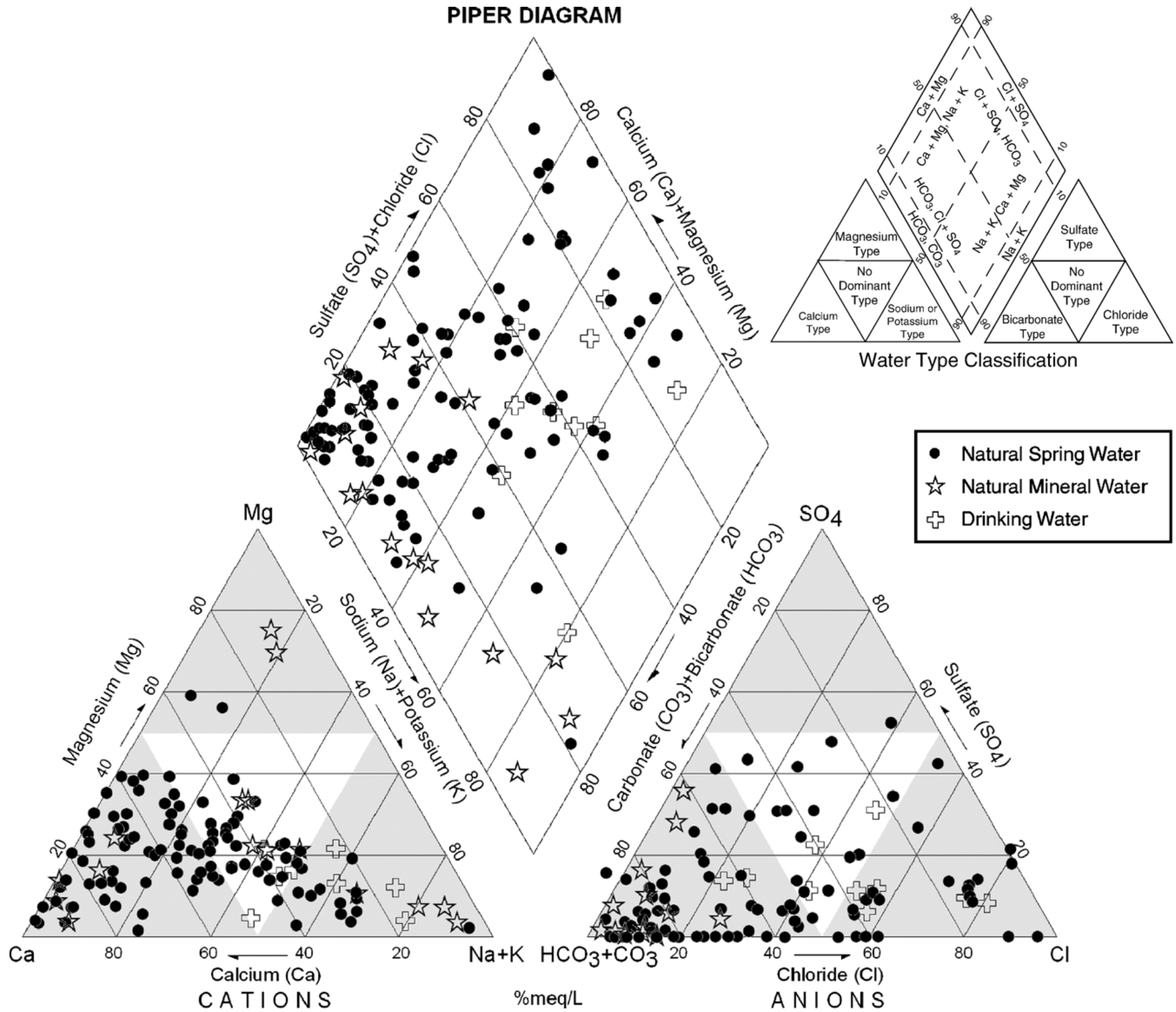

Fig. 2. Piper diagram of the 130 Turkish bottled waters.

study), to select the most discriminating parameters, and to investigate the overall variation of data. Rotation of principal components was carried out using the Varimax normalized method and only factors with eigenvalues greater than one were taken into consideration (Kaiser criterion). Varimax normalized procedure for eigenvector rotation resulted in two principal components (PC1 and PC2), which explained 66% of the total variance. Fig. 3 shows the results of the PCA analysis of the 130 bottled water brands. The first principal component (PC1) contains 49.8% of the total variance and the second component (PC2) represents 16.2% of the total variance. PC1 is concerned mainly with Ca, Mg, and HCO_3 (with loadings 0.92, 0.87 and 0.91, respectively), whereas PC2 was related to Na, Cl, and K (with loadings 0.78, 0.79 and 0.75, respectively). Thus, PC1 may be considered as an index of water hardness, whereas PC2 appears to represent the water saltiness. As Fig. 3 shows, the 17 natural mineral waters with high-mineral content were grouped on the upper side of the plot (stars), whereas a lack of clear grouping was observed between natural spring waters and drinking waters. Representation of the bottled waters using PCA does not seem to be particularly useful as the first two components accounted for only 66% of the total variance. Although there appears to be reasonable statistical discrimination, there is no objective means to clearly distinguish boundaries between some of the groups.

HCA was used for searching the natural grouping among bottled waters from different sources. The bottled water brands were classified according to their major ion composition using hierarchical cluster analysis. The data were standardized (*z*-scores) and the Euclidean distance was used as similarity measurement. The Ward's method was used as amalgamation rule to obtain the hierarchical associations. The data matrix used for the multivariate analysis was composed of a 130×8 data matrix (130 samples and 8 variables). The result of the HCA is presented as a dendogram (Fig. 4). The resulting dendogram was interpreted to have classified the 130 brands in four major groups based on a similarity of the eight study parameters. However, greater or fewer groups could be defined by moving the horizontal line of asterisks (phenon line) up or down. The first group is composed of 17 brands, all of which were natural mineral water brands. The

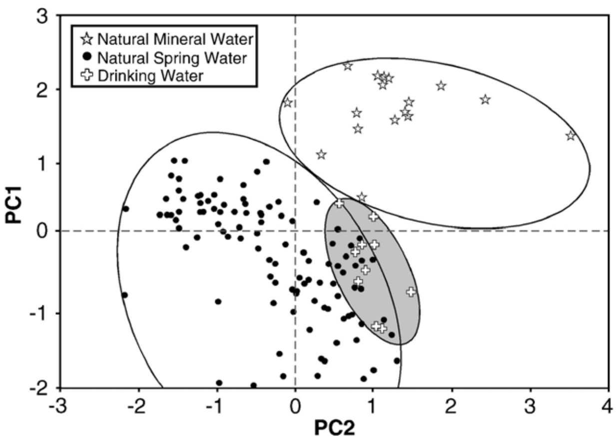

Fig. 3. Score plot of PC1 vs. PC2 illustrating the grouping of bottled water brands in 2-D PCA space.

second group comprised 15 natural spring water brands. The third group comprised 40 brands, 5 of drinking water brands, and 35 of natural spring water brands. The fourth group is composed of 58 brands, 5 of drinking water brands and 53 of natural spring water brands. In Fig. 4, the clusters in the dendogram are ordered from right to left in a near monotonic increase of total dissolved solids (TDS) content, with few exceptions.

The first group (natural mineral waters) presented the highest values of the analyzed parameters (TDS = 431.77–6774.75), while the fourth group presented the lowest ones (TDS = 12.77–206.59). It is worth noting that according to the Turkish legislation [6] natural mineral waters must contain at least 1000 mg/L dissolved minerals and/or trace elements. However,

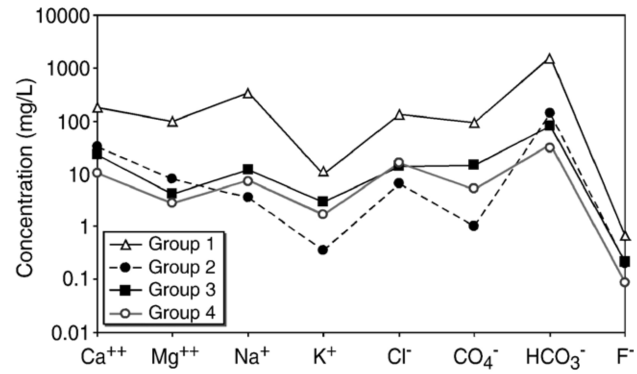

Fig. 5. Plot of the mean parameter values for the four principal bottled water groups determined by using hierarchical cluster analysis (HCA).

brands 122, 128 and 126 have TDS values (in mg/L) of 431.77, 486.72, and 983.23, respectively and therefore these brands do not qualify as natural mineral waters according to legislation. The second group (TDS = 97.25–313.88) and third group (TDS = 43.17–433) presented intermediate TDS concentrations between groups 1 and 4. The means of the parameters in each group reveal some trend between the major groups (Fig. 5): Group 1 shows a higher mineralization (mean TDS = 2490.29) than group 2 (mean TDS = 191), group 3 (mean TDS = 152), and group 4 (mean TDS = 79). The difference is due to higher Ca, Mg, Na, K, Cl, SO_4, HCO_3, and F contents in group 1 samples (Fig. 5). Group 2 samples have significantly higher Ca, Mg, HCO_3, and F contents than group 3 and 4 samples. Meanwhile, group 3 samples are distinguished from group 2 and 4 samples by their higher Na, K, and SO_4, contents. Finally, group 4 samples display a higher Cl content than groups 2 and 3 samples.

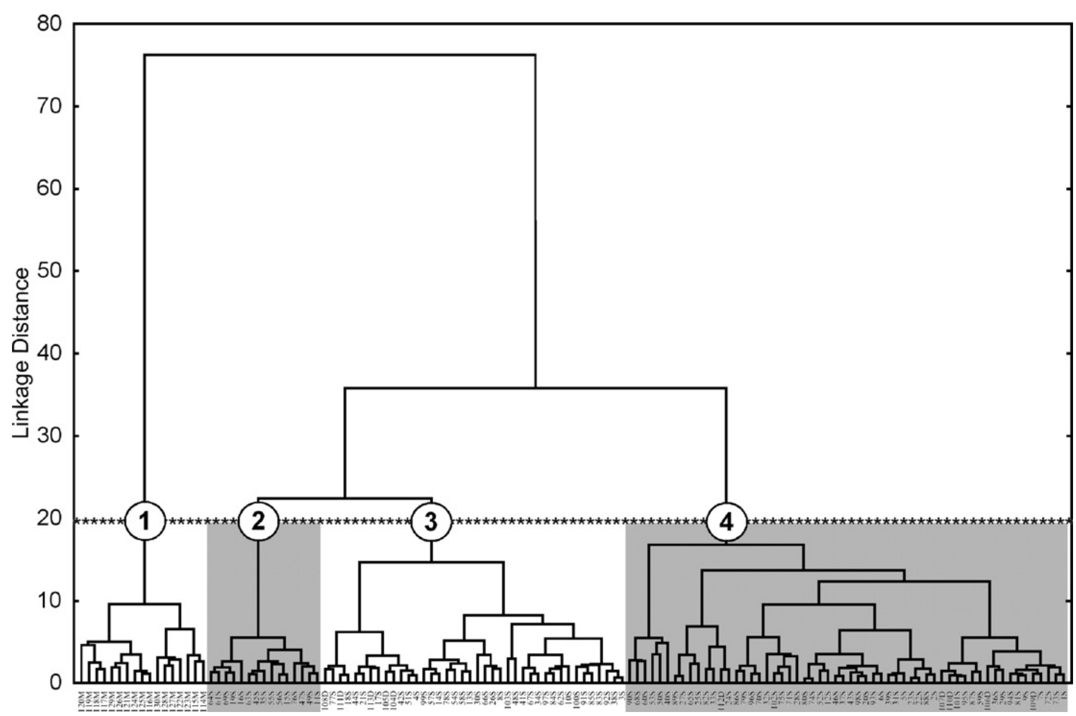

Fig. 4. Hierarchical dendogram from the HCA for the 130 bottled water brands. Line of asterisks defines “phenon line”, which is chosen by analyst to select number of groups or subgroups.

4. Conclusion

Water is essential for life and it is a freely tradable economic good in many forms and packaging materials. In the present manuscript we intend to highlight the main characteristics of the natural mineral waters, natural spring waters, and drinking waters from Turkey, all of which were admitted by the Turkish legislation as "bottled drinking waters" intended for human consumption. In this study, 189 Turkish bottled water brands comprising different water types were characterized by descriptive statistical measures (mean, minimum, and maximum) and 130 of them with complete major ion chemistry data were subjected to pattern recognition methods (HCA and PCA) to reveal the natural grouping among the brands. Pattern recognition techniques applied to the bottled water data set provided information on composition and characterized the waters from various sources. The HCA rendered four significant water groups on the basis of similarities/dissimilarities in their chemical characteristics. PCA provided information on the overall composition of the samples and continuity/overlap of the clusters. According to results of this study there is a substantial variation in the chemical composition of bottled waters studied. These variations in the chemical compositions of the bottled water brands are mostly caused by natural environment (geologic formations, climate, topography, etc.) and they are further modified by the treatment technologies (disinfection and purification) employed by bottled water manufacturers during the production.

Acknowledgements

The author gratefully acknowledges bottled water companies for providing the production licenses.

References

[1] S.P. Williams, Put the lid on bottled water, Newsweek 137 (May 21 2001) 61.
[2] B. Saad, F.W. Pok, A.N.A. Sujari, M.I. Saleh, Analysis of anions and cations in drinking water samples by capillary ion analysis, Food Chemistry 61 (1998) 249–254.
[3] L.M. Tamagnini, R.D. González, Bacteriological stability and growth kinetics of *Pseudomonas aeruginosa* in bottled water, Journal of Applied Microbiology 83 (1997) 91–94.
[4] A.B. Armas, J.P. Sutherland, A survey of the microbiological quality of bottled water sold in the UK and changes occurring during storage, International Journal of Food Microbiology 48 (1999) 59–65.
[5] D.W. Warburton, K.I. Dodds, R. Burke, M.A. Johnston, P.J. Laffey, A review of the microbiological quality of bottled water sold in Canada between 1981 and 1989, Canadian Journal of Microbiology 38 (1992) 12–19.
[6] Resmi Gazete, İçilebilir nitelikteki suların istihsali, ambalajlanması, satışı ve denetlenmesi hakkında yönetmelik, Resmi Gazete, vol. 23144, 18 October 1997 (in Turkish).
[7] P. Çelik, İlk sucumuz Hamidiye'ydi, Milliyet Gazetesi (01 September 2003) (Link accessed March 2004: http://www.milliyet.com.tr/2003/09/01/business/bus06.html#) (in Turkish).
[8] R.A. Freeze, J.A. Cherry, Groundwater, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1979.
[9] C. Güler, G. Thyne, J.E. McCray, A.K. Turner, Evaluation of graphical and multivariate statistical methods for classification of water chemistry data, Hydrogeology Journal 10 (2002) 455–474.
[10] StatSoft, Inc., STATISTICA for Windows Computer program manual Tulsa, OK, 2000.
[11] A.T. Miesch, Geochemical Survey of Missouri—Methods of Sampling, Laboratory Analysis and Statistical Reduction of Data, USGS Professional Paper, vol. 954-A, 1976.
[12] R.J. Rummel, Applied Factor Analysis, Northwestern University Press, Evanston, 1970.
[13] A.G. Judd, The use of cluster analysis in the derivation of geotechnical classifications, Bulletin of the Association of Engineering Geologists 17 (1980) 193–211.
[14] J.C. Davis, Statistics and Data Analysis in Geology, second edition, John Wiley and Sons, Inc., New York, 1986.
[15] G. Qian, G. Gabor, R.P. Gupta, Principal components selection by the criterion of the minimum mean difference of complexity, Journal of Multivariate Analysis 49 (1994) 55–75.
[16] B. Helena, R. Pardo, M. Vega, E. Barrado, J.M. Fernandez, L. Fernandez, Temporal evolution of ground water composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis, Water Research 34 (2000) 807–816.
[17] StatSoft, Inc., Electronic Statistics Textbook, Tulsa, OK, 1997.
[18] H.F. Kaiser, The application of electronic computers to factor analysis, Educational and Psychological Measurement 20 (1960) 141–151.
[19] G.A. Alther, A simplified statistical sequence applied to routine water quality analysis: a case history, Ground Water 17 (1979) 556–561.
[20] J.H. Ward, Hierarchical grouping to optimize an objective function, Journal of American Statistics Association 69 (1963) 236–244.
[21] A.M. Piper, A graphic procedure in the geochemical interpretation of water-analyses, Transactions of the American Geophysical Union 25 (1944) 914–923.
[22] W. Back, B.B. Hanshaw, Chemical geohydrology, Advances in Hydroscience 2 (1965) 49–109.