

This article was downloaded by:

On: 17 January 2011

Access details: *Access Details: Free Access*

Publisher *Taylor & Francis*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Environmental Analytical Chemistry

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713640455>

Chemometrical Interpretation of Lake Waters after their Chemical Analysis by Using Aas, Flame Photometry and Titrimetric Techniques

Şerife Tokalioğlu^a; Şenol Kartal^a

^a Department of Chemistry, Faculty of Arts and Sciences, Erciyes University, Kayseri, Turkey

Online publication date: 17 September 2010

To cite this Article Tokalioğlu, Şerife and Kartal, Şenol(2002) 'Chemometrical Interpretation of Lake Waters after their Chemical Analysis by Using Aas, Flame Photometry and Titrimetric Techniques', *International Journal of Environmental Analytical Chemistry*, 82: 5, 291 – 305

To link to this Article: DOI: 10.1080/03067310290018802

URL: <http://dx.doi.org/10.1080/03067310290018802>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

CHEMOMETRICAL INTERPRETATION OF LAKE WATERS AFTER THEIR CHEMICAL ANALYSIS BY USING AAS, FLAME PHOTOMETRY AND TITRIMETRIC TECHNIQUES

ŞERİFE TOKALIOĞLU and ŞENOL KARTAL*

*Department of Chemistry, Erciyes University, Faculty of Arts and Sciences,
TR-38039, Kayseri, Turkey*

(Received 1 August 2001; In final form 5 February 2002)

This article reports the distribution of trace metals and major ions in lake waters of Sultansazlığı, a reedfield near Kayseri, Turkey. The determination of elements such as Cr, Ni, Cu, Cd and Pb in lake water samples was performed by AAS after a preconcentration step using a column packed with Amberlite XAD-16 resin. Both flame- and graphite furnace-atomic absorption spectrometry was used to determine these metals in lake water samples. The concentrations of the metals bound to humic substances, and free metal ions were determined after their sorption-elution on the resin, separately. The column method optimized with sodium tetraborate reagent was used in determining the free metal ions. In the determinations of Ca^{2+} and Mg^{2+} , K^{+} , and SO_4^{2-} , Cl^{-} and total hardness, flame atomic absorption spectrometry, flame photometry, and titrimetry were used, respectively. In order to evaluate the analytical data by multivariate statistical techniques which enable feature reduction and grouping of the pollutant sources in lake waters from their chemical composition, principal component analysis (PCA), cluster analysis (CA) and correlation analysis were used. As a consequence of multivariate statistical evaluation, main anthropogenic sources like traffic, industry and agricultural processes were drawn to be responsible from the pollution in the environment investigated.

Keywords: Principal component analysis; Cluster analysis; Correlation analysis; Lake water; Preconcentration

INTRODUCTION

Both natural weathering and human activities are sources of metals, which are transported to lake waters through small stream and atmosphere. However, the relative impact of natural and anthropogenic sources on the geochemistry of lake waters and sediments is not always clear, especially in the area in which the local mineral contain high natural metal concentrations. Therefore, for a better interpretation of the pollution process in the surface waters (e.g., lake or river water), it is crucial to be able to distinguish between natural and human-related metal enrichments in the lake waters [1]. For this purpose, one of the most widely used techniques is principal component

*Corresponding author. Fax: +90-352-4374933. E-mail: kartal@erciyes.edu.tr

analysis (PCA) which is an important chemometric tool and seeks to establish combinations of variables capable of describing the principal data tendencies observed. PCA relies upon an eigenvector decomposition of the covariance and correlation matrix. Statistical and principal components analysis have been used to determine possible pollution sources relating to some major and trace ions in freshwater lake systems. A multivariate statistical analysis of the collected concentration data by the principal component method allows to identify the sources and typology of pollution, and to discriminate between the quality of waters at the different sampling sites. On the other hand, cluster analysis, a multivariate tool for classifying objects or variables, is a useful technique used for the same purpose. In addition, correlation analysis is another utilizing tool to investigate the binary and the multivariate relationships between variables and samples. All these techniques have received numerous applications in environmental sciences, recently [1–3].

Various applications of chemometrics techniques in the field of environmental studies for identifying pollution sources and characterization of samples were presented and discussed in literature. In application, principal component analysis, cluster analysis and to some extent the correlation analysis for lake waters [1,4,5], river waters [6,7], seawaters [3,8–11], rainwaters [12,13], sediments [2,14–17], soils [18], atmospheric aerosols [19–21], and historical marbles [22], have been used to establish the model to predict the origin of pollutants and for classifying both samples and pollutants.

One of the methods, often used, for orthogonal factor rotation is called varimax rotation. This is based on the assumption that the interpretability of factor j can be measured by the variance of the square of its factor loadings, i.e., the variance of $a_{1j}^2, a_{2j}^2, \dots, a_{pj}^2$. If this variance is large then the a_{ij}^2 values tend to be either close to zero or close to unity. Varimax rotation therefore maximizes the sum of these variances for all the factors. Kaiser first suggested this approach [23]. Later, he modified it slightly by normalizing the factor loadings before maximizing the variances of their squares, then this appears to give improved results. Varimax rotation can therefore be carried out with or without Kaiser normalization. Numerous other methods of orthogonal rotation have been proposed. However, varimax is recommended as the standard approach. In our example, z -scores for standardizing of the original data and Kaiser normalization for orthogonal factor rotation were applied as running the software (SPSS) [23].

In this study, the concentrations of variables in the lake waters of Sultansazlığı, a watershed, a wetland and a closed even area surrounded by high mountains, were measured. The area around the basin of the Sultansazlığı, in Kayseri, Turkey, is a typical case of mixed influence from natural and anthropogenic metal enrichments. It is in the triangle of Develi-Yahyalı-Yeşilhisar and a A-class watery districts which was included in the International Ramsar Agreement. Sultansazlığı is very rich in both fauna and flora and is mainly covered with reeds. Also, it contains both salty and fresh water ecosystems and is on the migration route of the birds and home for about 300 bird species.

In the collection of analytical data for performing the multivariate and/or bivariate statistical analyses, mainly three techniques were used: atomic absorption spectrometry for Ni, Cu, Pb, Ca, Mg, Cd and Cr, flame photometry for K, and tritrimetric methods for SO_4^{2-} , Cl^- and total hardness (TH: as CaCO_3). In general, especially flame- and graphite furnace atomic absorption spectrometric techniques require a preconcentration/separation procedure before the determination step due to low levels of numerous

metals in aquatic or other environmental media [3]. All parameters used in statistical analyses performed for this study were determined in the samples filtered through 0.45 μm membrane filter papers, and total concentrations of the variables (for 38 samples) were treated for the statistical evaluation of the data obtained. The metal levels of the lake waters (Cd, Cr, Cu, Ni and Pb) and of sediments (Co, Cu, Cr, Fe, Mn, Ni, Pb and Zn) in the investigated area were studied in detail in our previous investigations using speciation procedures [24,25]. A comparative study among correlation analysis, PCA and cluster analysis using the data obtained by chemical analysis was performed for this investigation.

EXPERIMENTAL

Apparatus and Chemicals

A Hitachi atomic absorption spectrometer (Model Z-8000) with an air-acetylene burner and a Zeeman background corrector was used to determine the metal ions in water samples. Whilst the determination of Cu, Pb and Ni metals found in the water samples was performed by FAAS, using an injection method [26], the measurements of Cr and Cd, present in low levels in the samples, carried out using a graphite tube atomizer (Model 180-704, Hitachi) because of their low levels after the preconcentration step. The operating conditions suggested by the manufacturer in the instrument data processor were followed to determine the metals. Injection volumes were 100 μL for FAAS and 10 μL for GFAAS. In determinations of K ion, a Jenway PFP7 model flame photometer was employed. All pH measurements were performed with a Nel digital pH meter and a combined glass electrode.

Analytical reagents grade chemicals and doubly distilled water were used to prepare the solutions. The stock solutions, 1000 $\mu\text{g mL}^{-1}$ in 1% HNO_3 , were made from the analytical grade nitrate salt (Merck, Darmstadt, Germany) of each element studied.

Site and Sampling

Sultansazlığı (the Reedfield of Sultan) is one of the most important marshy areas of Turkey, near Kayseri, in the view of bioornitological point. It is located in Central Anatolia. It is 70 km away, in the south of Kayseri. Being a marshland is crucial especially for breeding-resting and hibernationplace for bird species. During autumn and winter about 600 000–1 500 000, excluding water birds especially duckbirds, gather there among them the teal with 50 000–70 000 individuals form the greatest population of Europe and Asia. The average annual precipitation is about 394 mm and renewal of waters depends upon a little occasional influent streams.

The lakes in the area have different surfaces, totally the fresh and salty areas is 17 200 hectares (3650 hectares of Lake Yay, 1900 hectares of Kepir Reedfield, 3 300 hectares of Sultan Reedfield, and 8350 hectares of grassland). They are shallow (mean depth: 2–3 m, maximum 6–8 m) and temperate lakes. It is surrounded by high mountains on the four sides and completely even and sloped 2%. In the North of it, the highest volcano of the region, Mount Erciyes (3917 m), in the East the mountain of Develi, Akpınar, Çiçekliyurt (2047 m, 2057 m), in the South the Taurus Mountains, Aladağ (3373 m), Elmalı (2235 m), and in the West, Kartalkaya (1958 m), İncildağı (1759 m) exists.

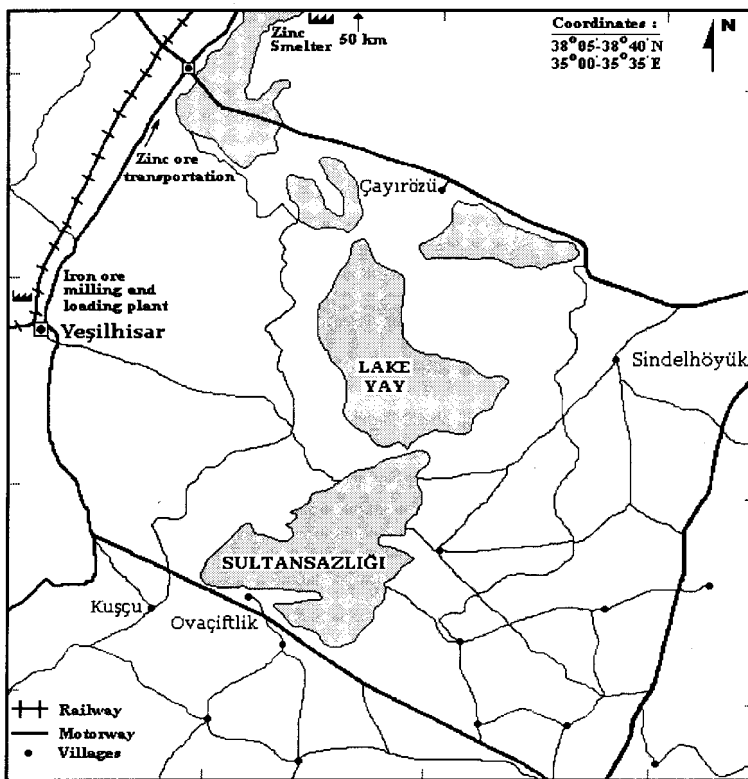


FIGURE 1 Map of the sampling sites.

As probable sources of metal contamination influencing the area, a zinc smelter and an industrialized region locate in the North at about 50 km far from Sultansazlığı and an iron ore milling and loading plant at Yeşilhisar County is approximately 10 km away from the wetland. Transporting of the zinc and iron ores along the motorway and railway which pass through from neighbours of the wetland, is also a crucial pollution sink on the basin. Particulate matter and dusts are transported into the basin by wind [26]. Map of the sampling site is shown in Fig. 1.

Totally 38 water samples were taken from different freshwater lakes of Sultansazlığı on October 21, 1995 and on March 30, 1996. At these dates, seven and thirty-one (as total thirty-eight) water samples were collected, respectively. The samples were kept in the refrigerator at 4°C before the analysis [27]. The pH values of the water samples were approximately 7.7.

Speciation Procedure

Speciation procedure have been widely used for the determination of heavy metals in different environmental media such as soil, sediment, seawater, and terrestrial surface water (river, lake etc.). In general, before determining heavy toxic metals in natural waters, an enrichment and separation stage is needed for reliable, precise and accurate measurements. For this reason, to determine the metals in lake waters, a solid-phase

extraction procedure including a glass column and an adsorbent resin was employed. In the determination of Ni, Cu, Pb, Cd and Cr metals bound to organic substances and the free ions, a two-step sequential solid phase extraction procedure was employed [23].

Column preparation procedure and characteristic properties of Amberlite XAD-16 resin and the speciation procedure were given elsewhere in details [23,26]. The speciation procedure is briefly given below.

Separation of Metal-humic Complexes from Free Metal Ions

First, the water samples were filtered through a 0.45 μm Nuclepore membrane filter (47 mm diameter) to separate the metals bound to the organic substances and free metal ions. A 250 mL aliquot of the filtrate was passed through the column to adsorb the metals bound to humic substances (i.e., humic and fulvic acids). The effluent was reserved in order to determine the free metal ions. The metals sorbed on the resin were eluted with 1 mol L⁻¹ HCl in acetone. The eluate was evaporated near to dryness and the residue was dissolved in 0.7 mL of 2 mol L⁻¹ nitric acid and measurements were performed by AAS.

Free Metal Ions

To determine the free metal ions in the effluent, sodium tetraborate reagent was added to the effluent to adjust pH 9.2. This solution was passed through the column filled with Amberlite XAD-16 resin. After this step, the procedure mentioned above was applied to the eluate for determining the metals.

Statistical Evaluation of the Data

Principal Component Analysis (PCA) is a powerful tool for performing exploratory data analysis, such as a detective seeking to uncover hidden relationships in the data table. When it begins, it is not known what the relationships one is looking for will be like. Instead, one has to wait until it uncovers hidden relationships. PCA, using an eigenvector rotation, can be able to decompose the data matrix into a series of loadings and scores. In chemistry, this approach can be used by diagonalizing the correlation or covariance matrix. Typically, the largest eigenvalues (based on the size of the eigenvalues) are the most important. Algebraically, principal components are particular linear combination of the variables X_1, X_2, \dots, X_p . Geometrically, these linear combinations represent the selection of a new coordinate system obtained by rotating the original system with X_1, X_2, \dots, X_p as the coordinate axes. The new axes represent the directions with maximum variability and provide a simpler and more parsimonious description of the covariance structure. PCA may also be obtained for the standardized variables. For this purpose, z-scores described just below as standardized variables were used to evaluate the data with the multivariate statistical analysis techniques in this study [28,29].

Cluster analysis (CA) To carry out a clustering, one of the decisions to be taken is the selection of the clustering algorithm. Another is the selection of the similarity measure, and the last one is to choose a possible transformation of the variables. Euclidean distance and correlation-based measures are the two main types of the distance, of

which the former is used most often. The Euclidean distance between two objects (or samples) i and i' is explained as follows:

$$D_{ii'} = \sum_{j=1}^m \sqrt{(x_{ij} - x_{i'j})^2}$$

where x_{ij} and $x_{i'j}$ are the values of variable j for the objects i and i' and m is the number of variables. Also, another variant of the Euclidean distance called squared Euclidean distance, which is the sum of the squared distances over all variables, is used frequently as a distance measure for clustering [30]. An alternative criterion is the choice of the number of factors could be based on a cluster analysis and also it aims at the determination of relationships between the samples without using prior information about these relationships. In cluster analysis, certain samples are assumed to fall into predetermined groups or classes, which are used to produce class models. The most common method of cluster analysis is called hierarchical clustering. There are several variants for hierarchical cluster analysis like between-groups linkage, within-groups linkage, nearest neighbor, furthest neighbor, centroid clustering, median clustering, and Ward's method. Among these methods the Ward's method gives more meaningful results. Therefore, only the Ward's method combined with Euclidean distance for clustering of the samples and squared Euclidean distance for clustering of the variables, respectively [1,22], will be considered here.

In order to avoid problems caused by scale differences in data set, it is generally required to standardize or transform the data, because variables with large values contribute more to the distances measures than those with small values. Therefore, z -scores (also called standard scores) for quantitative variables were used in this work before applying multivariate statistical techniques. These scores are transformations of the data values to standard deviation units and indicate the relative position of each value within its distribution. For each variable that is specify, SPSS subtracts the variable's sample mean, \bar{x} , from each value, x_i , and then divides the difference by the sample standard deviation, s .

$$z_i = \frac{x_i - \bar{x}}{s}$$

The mean z -score and the standard deviation are always 0 and 1, respectively. When variables are measured in different units, a z -score transformation places the variables on a common scale, making it easier to compare their values. The distribution of scores has the same shape as the original distribution, so they are not remedy for problem data. Because z -scores (standardized values) are most useful for data that follow a normal distribution (or at least a symmetric distribution), population is analyzed in log units [31].

Correlation analysis In the presence of one independent and one dependent variable, the classical measure of linear relationship is given by the correlation coefficient. The correlation coefficient has often been considered as another measure of similarity between the variables. The closer the correlation coefficient is to unity the more similar the two samples are.

For the statistical interpretation of the analytical data obtained from the analysis of lake water samples, a package program of SPSS V9.05 was used for PCA, CA and correlation analyses.

RESULTS AND DISCUSSION

Analysis Results of Lake Water Samples

The concentration of trace elements and major ions, total hardness values and also their basic statistics for the 38 lake water samples from Sultansazlıği are given in Table I.

TABLE I Concentration of the variables and basic statistics of the input data set for the lake waters

Sample No.	Variables										
	^a Cu ²⁺	^b Pb ²⁺	^a Ni ²⁺	^a Cd ²⁺	^a Cr ³⁺	^b Ca ²⁺	^b Mg ²⁺	^b K ⁺	^b Cl ⁻	^b SO ₄ ²⁻	^{b,c} TH
1	1.83	2.73	10.4	0.06	0.74	29.5	67.5	26.0	329	144	961
2	1.96	2.99	4.72	0.13	0.62	42.8	100	33.0	344	196	667
3	1.68	2.68	20.6	0.07	0.65	52.1	130	24.0	474	234	761
4	1.75	2.29	4.12	0.07	0.49	47.7	125	27.0	576	229	746
5	2.45	1.89	5.21	0.06	0.86	50.6	125	36.5	591	243	821
6	2.84	2.86	10.7	0.16	0.57	60.2	148	28.0	496	234	765
7	2.09	2.57	9.30	0.09	0.65	34.3	97.5	19.5	285	191	512
8	1.16	3.65	20.6	0.04	0.09	63.6	134	14.6	344	196	648
9	0.62	7.42	15.7	0.13	0.38	85.4	155	19.8	394	205	693
10	0.77	0.81	12.6	0.09	0.29	70.2	130	16.0	416	196	684
11	0.58	1.06	10.7	0.04	0.32	79.2	145	18.6	424	154	672
12	0.65	1.62	19.0	0.16	0.36	67.4	125	16.2	442	186	693
13	0.84	1.37	24.1	0.20	0.39	60.2	110	13.8	445	187	697
14	1.11	3.28	19.9	0.12	0.86	56.3	89.9	13.6	412	186	721
15	1.40	1.06	20.9	0.16	1.32	66.2	110	13.2	291	157	612
16	1.53	0.89	14.2	0.06	0.31	61.0	95.0	11.2	231	119	546
17	1.01	1.00	10.4	0.08	0.36	68.0	105	12.2	269	187	572
18	1.76	0.64	14.6	0.16	0.29	63.8	140	13.0	284	191	624
19	1.31	1.71	10.6	0.12	0.39	69.6	115	14.0	299	210	612
20	0.90	1.63	8.33	0.07	0.35	64.4	100	12.8	284	172	592
21	1.01	1.86	10.6	0.03	0.32	68.4	110	13.0	261	237	560
22	1.12	1.34	14.3	0.08	0.31	58.6	80.0	10.0	170	208	484
23	1.20	1.52	13.3	0.03	0.87	59.0	70.0	8.4	94.6	156	396
24	1.29	1.59	16.3	0.30	0.39	68.2	70.0	8.6	125	126	423
25	2.57	1.90	7.58	0.11	1.17	61.8	70.0	7.0	111	165	412
26	1.77	1.56	18.3	0.04	1.13	59.0	60.0	9.6	104	165	408
27	1.06	1.63	15.7	0.04	1.21	55.0	55.0	8.6	81.7	141	372
28	1.63	1.26	26.5	0.05	0.14	58.4	65.0	8.45	187	144	323
29	0.82	1.44	11.1	0.02	1.58	60.6	65.0	8.0	72.3	156	356
30	2.82	3.28	14.9	0.04	0.14	57.2	58.0	6.8	54.9	130	338
31	0.71	1.95	28.4	0.16	0.13	31.8	44.0	4.7	74.4	99.8	385
32	3.56	3.15	24.4	0.09	0.30	55.9	64.0	7.7	78.1	140	304
33	1.76	1.34	19.1	0.03	0.27	56.4	61.0	6.2	47.6	130	308
34	1.81	1.93	39.1	0.03	0.34	53.8	45.0	5.8	47.6	130	309
35	1.79	1.48	16.0	0.08	0.97	41.1	44.0	4.1	83.2	163	207
36	1.48	4.33	22.8	0.03	0.09	29.6	39.0	5.0	67.2	163	491
37	1.33	2.08	36.6	0.03	0.96	43.0	38.0	6.2	23.5	97.9	290
38	1.62	2.86	30.6	0.16	0.98	30.6	33.0	5.2	45.4	163	491
Mean	1.52	2.12	16.64	0.09	0.57	56.3	89.9	13.6	246	172	538
^d S.D.	±0.68	±1.23	±8.22	±0.06	±0.38	±13.3	±35.5	±8.1	±167	±38	±178
S.D./mean	0.45	0.58	0.49	0.67	0.67	0.24	0.39	0.60	0.68	0.22	0.33

^aµg/L.^bmg/L.^cTotal hardness (as CaCO₃).^dStandard deviation.

TABLE II Percentage concentration based on sum of levels of metals bound to organic substances and free metals and their standard deviation ($n=38$)

Form of the metal	Metals (concentration in %)				
	Cu	Pb	Ni	Cd	Cr
Bound to organic matter	49.9	53.1	57.9	38.2	55.0
Free metal ions	50.1	46.9	42.1	61.8	45.0
Standard deviation	15.2	16.3	13.8	19.4	15.8
Ratio of bound- to free metal	0.99	1.13	1.37	0.62	1.23

The heavy metal speciation has a substantial interest in water analysis since their geochemical behaviour and biological effects depend greatly on their physico-chemical forms. Although the metal contents obtained from solid phase extraction steps give informative explanation about their bioavailability and mobility, the total metal concentrations give more information on metal pollution and/or accumulation in lake waters or any different environment. Therefore, the total concentrations of metal and the other variables were taken into consideration in the evaluation of the analytical results by PCA, CA and correlation analysis. The results given in Table II contain the percentage concentrations of physico-chemical forms of the metals Cr, Ni, Cu, Cd, and Pb in the lake waters in which the total metal concentrations (filtered through $0.45\ \mu\text{m}$ membrane filter) were based upon the summation of free or hydrated metal ions and bound to organic substances (e.g., humic and fulvic acids). As can be seen from Table II, the ratios of the metals bound to organic substances in their soluble forms are in the range of 38–58%, while the ratios of the free metal forms are in the range of 42–62%, and the free form of cadmium element is more abundant than the others. In other words, the bound forms of Ni, Cr and Pb to organic substances in aquatic medium are more effective than those of their free forms. This is a crucial feature in the determination of metal in natural waters, soils and sediments by using instrumental techniques due to their bioavailability and mobility. For example, if one wants to analyze the metals by using ion selective electrode and polarographic techniques, one can only collect an information about free metal ions.

Statistical Evaluation of Analytical Results

Correlation Analysis Table III shows the linear correlation coefficients for all the variables using the data collected from the sampling sites of the region. The correlation coefficients lower than ± 0.10 are not illustrated in the table. The values with boldface attest highly significant correlation between corresponding variables (according to the criterion of $r=0.40$, $p=0.01$ and $n=38$). The correlation coefficient of about 0.40 and greater would indicate a statistically significant relationship between variables [32]. Ni exhibits moderate, but negative correlations with Mg, K, Cl^- , SO_4^{2-} , and total hardness. The variables such as Mg, K, Cl^- , SO_4^{2-} , and total hardness have strongly correlated to each other ($0.64 < r < 0.89$) since these variables are the main components of natural waters. Ca is only moderately correlated with Mg ($r=0.60$).

PCA Results The PCA is a powerful technique for displaying the relationships between objects using scores plots and the impact of variables on each principal components using loadings plots. For this purpose, it was taken advantage of a multivariate

TABLE III Correlation matrix between the concentrations of variables in lake water samples

<i>Variable</i>	<i>Cu</i>	<i>Pb</i>	<i>Ni</i>	<i>Cd</i>	<i>Cr</i>	<i>Ca</i>	<i>Mg</i>	<i>K</i>	<i>Cl⁻</i>	<i>SO₄²⁻</i>	<i>TH^b</i>
<i>Cu</i>	1.00 ^a										
<i>Pb</i>	0.12	1.00									
<i>Ni</i>	- ^b		1.00								
<i>Cd</i>				1.00							
<i>Cr</i>		-0.14	-0.10		1.00						
<i>Ca</i>	-0.32		-0.24	0.11	-0.15	1.00					
<i>Mg</i>	-0.18		-0.50	0.22	-0.24	0.60	1.00				
<i>K</i>	0.15	0.19	-0.57	0.10			0.69	1.00			
<i>Cl⁻</i>	-0.12	0.10	-0.48	0.22	-0.15	0.28	0.87	0.85	1.00		
<i>SO₄²⁻</i>		0.18	-0.53			0.23	0.71	0.67	0.74	1.00	
<i>TH^c</i>	-0.15	0.18	-0.45	0.23	-0.10	0.13	0.74	0.82	0.89	0.64	1.00

^aCorrelation coefficients greater than 0.40 are meaningful at 99% confidence level for 38 samples [31].

^bLess than ± 0.10 not indicated.

^cTH as CaCO₃.

TABLE IV The results of PCA analysis for lake water samples from Sultansazlıđı (total variance explained), $n = 38$

<i>Component</i>	<i>Initial eigenvalues</i>			<i>Extraction sums of squared loadings</i>			<i>Rotation sums of squared loadings</i>		
	<i>Total</i>	<i>% of variance</i>	<i>Cumulative %</i>	<i>Total</i>	<i>% of variance</i>	<i>Cumulative %</i>	<i>Total</i>	<i>% of variance</i>	<i>Cumulative %</i>
1	4.622	42.022	42.022	4.622	42.022	42.022	4.438	40.348	40.348
2	1.558	14.165	56.187	1.558	14.165	56.187	1.614	14.673	55.021
3	1.252	11.386	67.573	1.252	11.386	67.573	1.381	12.551	67.573
4	0.966	8.779	76.352						
5	0.830	7.545	83.896						
6	0.737	6.698	90.594						
7	0.447	4.064	94.658						
8	0.349	3.177	97.835						
9	0.118	1.075	98.910						
10	0.069	0.629	99.539						
11	0.051	0.461	100.000						

statistical package program, SPSS, for all the computations. A 38×11 data matrix was used for the multivariate analysis. Table IV reflects the eigenvalues and the variance percentages (accounted for cumulative) corresponding to the principal components (PCs) using varimax rotation with Kaiser normalization for the lake water samples. As can be seen in Table IV, only the eigenvalues of over 1 are taken into consideration and accounted for 67.6% of the total data variance. As a result of this, the number of PCs are reduced to three.

The PC1 has normally the highest factor loadings and the biggest contribution to the total variance of the data system (accounted for variance percentage of 40.3). The second and the third PCs provide a contribution of 14.7 and 12.6% to the total variance, respectively. Table V reflects the varimax rotated factor loadings for each of PCs and communalities for each of variables from the lake waters of Sultansazlıđı. Cd has very low communality (0.234) and therefore it has a weak factor loadings (0.444) on the PC3. The PC1 has very high factor loadings for variables: Mg, K, Cl⁻, SO₄²⁻, TH and Ni (but negative and may be related to fuel oil combustion). This PC may be attributed to crustal elements or soil originated. The PC2 has been

TABLE V Varimax rotated PC loadings for lake water samples

<i>Variable</i>	<i>Communalities</i> (h^2)	<i>PC1^a</i>	<i>PC2</i>	<i>PC3</i>
Cu	0.528	c	0.711	-0.141
Pb	0.579	0.143	0.477	0.575
Ni	0.659	- 0.709	-	0.384
Cd	0.234	0.124	-0.146	0.444
Cr	0.531	-	0.154	- 0.713
Ca	0.675	0.238	- 0.779	0.107
Mg	0.917	0.833	-0.397	0.256
K	0.889	0.921	0.201	-
Cl ⁻	0.906	0.918	-0.144	0.207
SO ₄ ²⁻	0.708	0.839	-	-
TH ^b	0.808	0.867	-	0.235
% of total variance	0.676	40.35	14.67	12.55
Possible sources		Soil and combustion	Mixed (soil, agricultural, traffic and industry)	Traffic and industry

^aRotation method: Varimax with Kaiser normalization and rotation converged in 4 iterations.

^bTotal hardness as CaCO₃.

^cPC loading less than ± 0.10 are not indicated.

correlated with Cu and Ca variables and to some extent with Pb (0.477). Pb has more efficient loading for PC3 than PC2. The PC2 seems to be a mixed component; that is, soil, agricultural, traffic and industrial originated. Copper sulphate is used with agricultural purposes for combating harmful microorganisms such as pathogens and bacteria in the areas of grapevines. The PC3, in addition to Pb, has the biggest factor loadings for Cd and Cr. This principal component is easily attributed to industrial and traffic originated sources, because the area is mainly influenced from motorway (far from approximately 5 km) and a zinc smelter (emits mainly Pb, Cd and Zn) which is about 50 km far from the region investigated. Relationships between samples and the impact of variables on each principal components can be visualized by utilizing scores plots and loadings plots, respectively. This includes plotting a graph of the scores of the samples against those for a different component; for example, PC1 versus PC2, and Fig. 2 illustrates this scores plot for the 38 samples. As can be seen from Fig. 2, there is a clear separation between the samples to some extent. This plot makes better separation among the samples: 1 and 2; 4, 5 and 6; 10, 11 and 12; 31, 34 and 37; and 32, 36 and 38 from the same origin, while the rest of the samples are grouped around the geometric origin of the principal components. For better understanding of the association among the eleven considered variables, it can be plotted the varimax rotated factor loadings on a PC1–PC2 axes plane (Fig. 3); this figure includes 55% of the total variance. It is evident from this plot that Mg, Ni, Cr, TH (as CaCO₃), SO₄²⁻ and Cl⁻ are strongly correlated with the PC1, whereas Pb, Ca and Cu show a correlation with the PC2; Cd is weakly correlated to PC1 and PC2, but moderately with PC3. In addition, Cd has the lowest total variance (with a communality of 0.234) when compared with those of the other variables. Therefore, this indicates that the exploratory information about the Cd is very poor acquired from this data set. In Fig. 4, factor loadings are given for eleven variables. Factor scores (Fig. 3) and factor loadings (Fig. 4) seem in a good agreement to each other.

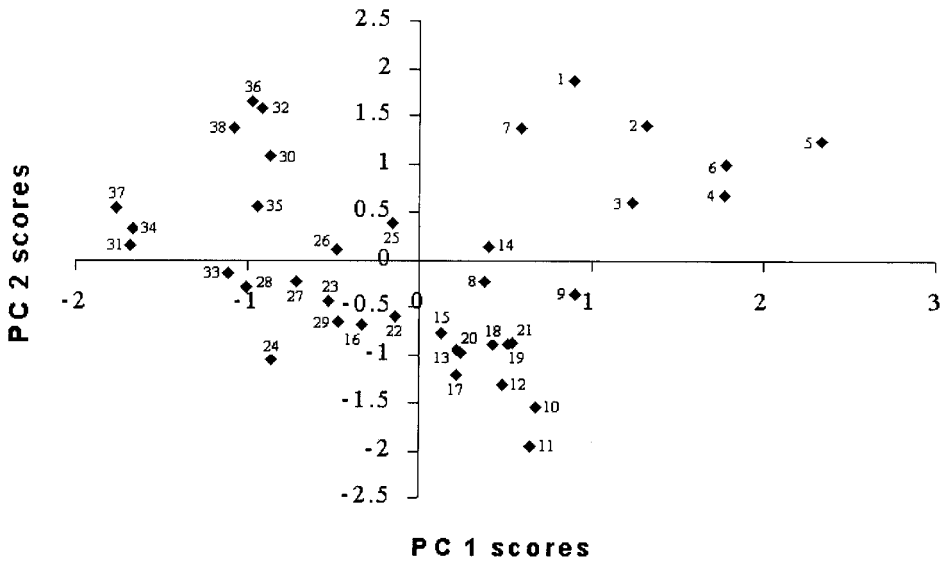


FIGURE 2 Scores plot of the 38 samples (for 11 variables). In this figure, the PC1 is represented by the horizontal axis and the PC2 by vertical axis.

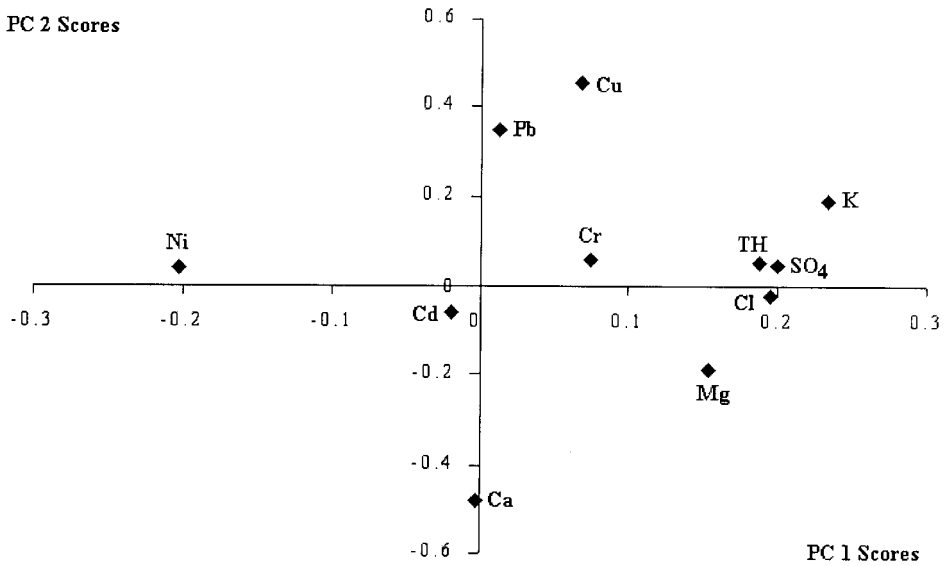


FIGURE 3 Scores plot of the 11 variables (for 38 samples). The PC1 is represented by the horizontal axis and the PC2 by vertical axis.

Hierarchical Cluster Analysis

Finally, a hierarchical clustering (the Ward's method) of the standardized data set was performed and the results are shown in the dendrograms. Figures 5 and 6, for samples

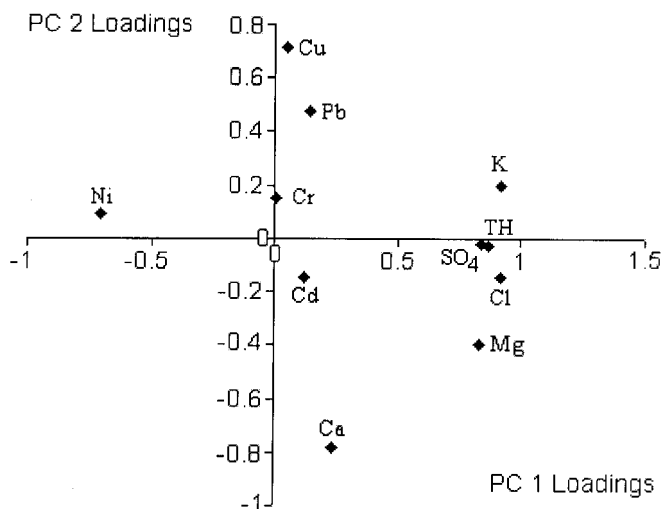


FIGURE 4 Loadings plot of the 11 variables (for 38 samples). The PC1 is represented by the horizontal axis and the PC2 by vertical axis.

(similarities obtained by Euclidean distance) and variables (similarities obtained by squared Euclidean distance), respectively. The classification of the input data by the use of multivariate statistical approaches makes it possible to suggest some hypotheses for the explanation of the possible interactions in the water system [8]. Before cluster analysis, variables were standardized by means of z -scores; then Euclidean distances or squared Euclidean distances were calculated. In Fig. 5, all clusters would be interpreted to be divided into two bigger subgroups: the first group contains the samples between 23 and 38, excluding 28, and the second contains the samples between 2 and 22, excluding 1. Factor scores results in Fig. 2 are in agreement with the clustering results for samples (Fig. 5). The cluster analysis results for variables are given in Fig. 6. This dendrogram has two bigger subgroups: the first groups contains the variables Cu, Pb, Ni, Cr and the second contains the rest of the variables. This results may indicate that the first group is relevant to anthropogenic sources while the second is related to natural sources.

CONCLUSION

The determined pollutants or other variables within a study provide a basis for the comparison with situations previously investigated or to be performed in future. We can conclude that the quality of the lake waters in the point of toxic heavy metals is acceptable and comparable to the analogous areas. This is, especially, important for the bio-ornitological point of view of the area. When a whole evaluation of the statistical results was done, it is clearly seemed that some water parameters; heavy metals such as Pb, Cu, and to some extent Ni and Cr, are related to anthropogenic origin whilst the others investigated mainly naturally originated. Cd showed low factor loading which means that is explained variance(%) is low and it is clustered with major water components, especially with Ca. This behaviour may

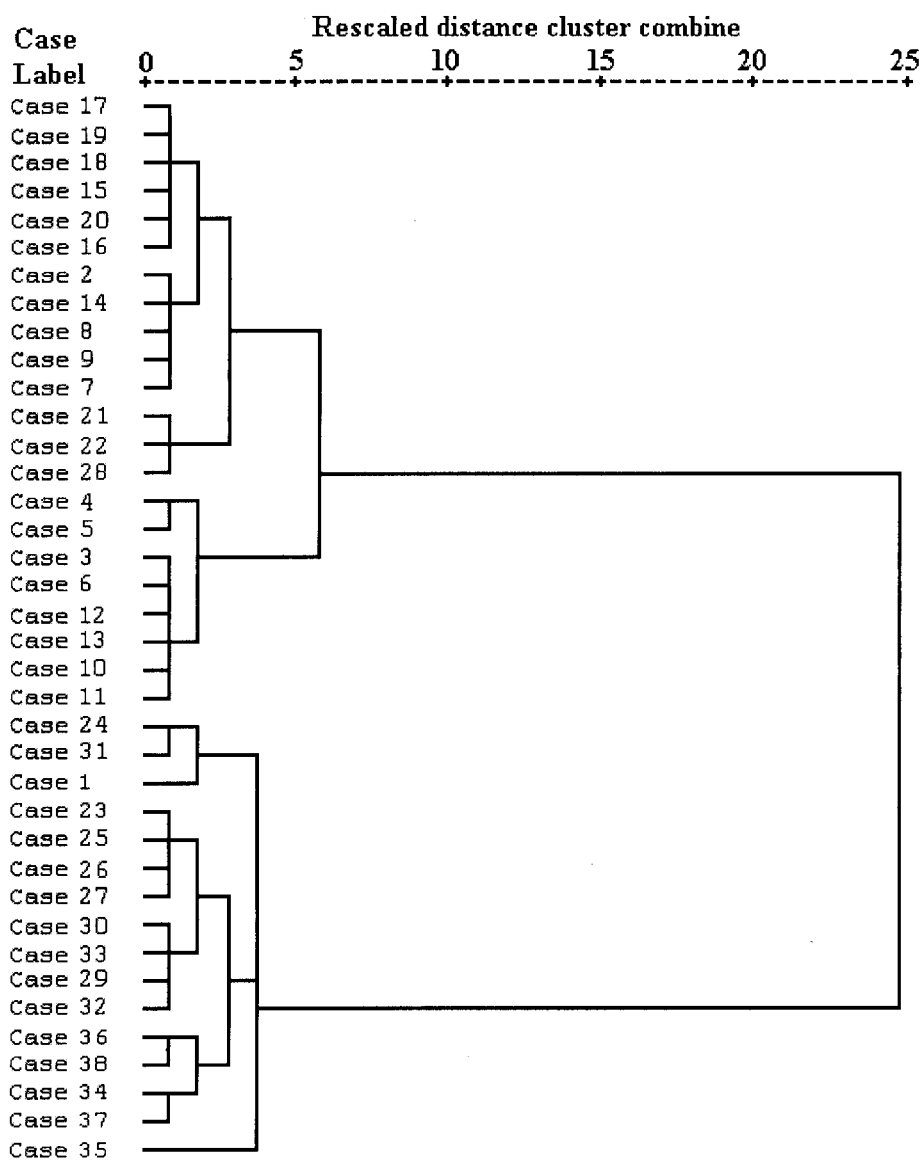
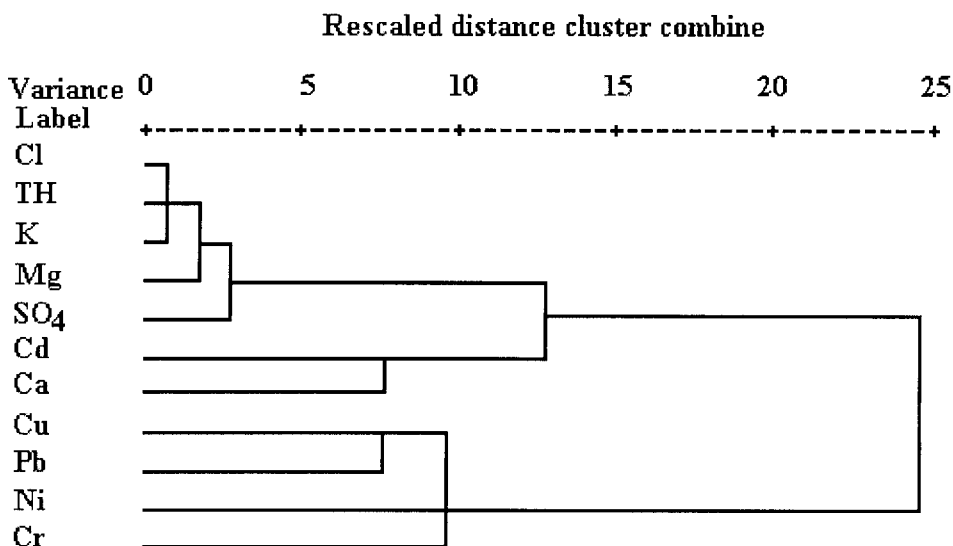


FIGURE 5 Dendrogram resulting from the Ward's method of hierarchical cluster analysis of the 38 samples (for the 11 variables). Similarities have been calculated from Euclidean distance.

be probably related to its natural origin. At the same time, Cd had very low concentrations in the lake waters studied.

The multivariate statistical approaches seem to be very useful for the interpretation of data sets from environmental studies. The results of principal components and cluster analyses demonstrate that they provide information with respect to reality in fields of scientific research.



Dendrogram by using Ward's Method with squared Euclidean distance

FIGURE 6 Dendrogram resulting from the Ward's method of hierarchical cluster analysis for the 11 variables (in 38 samples). Similarities have been calculated from Euclidean distance.

References

- [1] E. Reisenhofer, A. Picciotto and D. Li, *Anal. Chim. Acta*, **306**, 99–106 (1995).
- [2] M.O. Angelidis and M. Aloupi, *Mar. Pollut. Bull.*, **40**, 77–82 (2000).
- [3] M.M. Morales, P. Marti, A. Llopis, L. Compos and S. Sagrado, *Anal. Chim. Acta*, **394**, 109–117 (1999).
- [4] R. Leardi, C. Armanino, S. Lanteri and L. Alberrotanza, *J. Chemometr.*, **14**, 187–195 (2000).
- [5] E. Mentasti, O. Abollino, M. Aceto and C. Sarzanini, *J. Environ. Anal. Chem.*, **71**, 245–255 (1998).
- [6] I.L. K uchler, N. Miekeley and B.R. Forsberg, *J. Braz. Chem. Soc.*, **11**, 286–292 (2000).
- [7] N. Mazlum, A.  zer and S. Mazlum, *Turk. J. Eng. Environ. Sci.*, **23**, 19–26 (1999).
- [8] V. Simeonov, G. Andreev and M. Karadjov, *Fresenius J. Anal. Chem.*, **345**, 744–747 (1993).
- [9] E. Reisenhofer, G. Adami and A. Favretto, *Fresenius J. Anal. Chem.*, **354**, 729–734 (1996).
- [10] P. Rivaro, M. Grotti, C. Ianni and E. Magi, *Mar. Pollut. Bull.*, **36**, 880–886 (1998).
- [11] P. Barbieri, G. Adami, A. Favretto and E. Reisenhofer, *Fresenius J. Anal. Chem.*, **361**, 349–352 (1998).
- [12] S.L. Tsakovski and V.D. Simeonov, *Toxicol. Environ. Chem.*, **65**, 203–216 (1998).
- [13] T. Ozeki, K. Koide, N. Ogawa, T. Adzuhata, M. Kajikawa and T. Kimoto, *Anal. Sci.*, **13**, 169–176 (1997).
- [14] V. Simeonov, S. Tsakovski and D.L. Massart, *Toxicol. Environ. Chem.*, **72**, 81–92 (1999).
- [15] M.C. Kennicutt II, P.N. Boothe, T.L. Wade, S.T. Sweet, R. Rezak, F.J. Kelly, J.M. Brooks, B.J. Presley and D.A. Wiesenberg, *Can. J. Fish. Aquat.*, **53**, 2554–2566 (1996).
- [16] P. Heininger, J. Pelzer, R. Henrion and G. Henrion, *Fresenius J. Anal. Chem.*, **360**, 344–347 (1998).
- [17] V. Simeonov, D.L. Massart, G. Andreev and S. Tsakovski, *Chemosphere*, **41**, 1411–1417 (2000).
- [18] J.W. Einax and U. Soldt, *Fresenius J. Anal. Chem.*, **361**, 10–14 (1998).
- [19] N. Ferrer and J.J. Perez, *Atmos. Environ.*, **24B**, 181–184 (1990).
- [20] Ş. Kartal, M. Dođan, C.M. Rojas and R.E. van Grieken, *Sci. Total Environ.*, **133**, 83–97 (1993).
- [21] Z. Jeran, R. Jacimovic, F. Batic, B. Smodis and H.Th. Wolterbeek, *Fresenius J. Anal. Chem.*, **354**, 681–687 (1996).
- [22] H. Ak ay, C. Demir, U.  zer and Ş. G cer, *Chem. Anal. (Warsaw)*, **44**, 577–589 (1999).
- [23] B.F.J. Manly, *Multivariate Statistical Methods: A Primer*, Chapman and Hall, London, pp. 75–76 (1986).
- [24] Ş. Tokaliođlu, Ş. Kartal and L. El i, *Anal. Sci.*, **16**, 1169–1174 (2000).
- [25] Ş. Tokaliođlu, Ş. Kartal and L. El i, *Anal. Chim. Acta*, **413**, 33–40 (2000).

- [26] L. Elçi, Ş. Kartal, A. Ülgen, M. Doğan and P. Doğan, *Turk. J. Chem.*, **14**, 294–302 (1990).
- [27] Ş. Tokaloğlu, Ph. D. Thesis, University of Erciyes, Kayseri (1997).
- [28] L.K.J. Lajunen, *Spectrochemical Analysis by Atomic Absorption and Emission*, p. 214. Royal Society of Chemistry, Cambridge (1992).
- [29] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, 2nd Edn., pp. 340–348. Prentice Hall, Inc., New Jersey (1988).
- [30] <http://ull.chemistry.uakron.edu/chemometrics/pca/>
- [31] B. Massart, Q. Guo, F. Questier, D.L. Massart, C. Boucon, S. de Jong and B.G.M. Vandeginste, *Trends Anal. Chem.*, **20**, 35–41 (2001).
- [32] SPSS Base 10.0 Applications Guide (SPSS Inc., Chicago, 1999), p. 32.
- [33] T. Yamane, *Statistics: An Introductory Analysis*, 3rd Edn., Harper and Row Publishers, New York, pp. 1092–1093 (1973).