Desalination and Water Treatment
www.deswater.com
doi: 10.1080/19443994.2016.1147377

57 (2016) 22152–22165
October

Taylor & Francis
Taylor & Francis Group

Assessment of heavy metals pollution using multivariate statistical analysis methods in Wadi El Bey (Tunisia)

Taoufik Gasmi[a,b], Imen Khouni[a,*], Ahmed Ghrabi[a]

[a]*Laboratory of Wastewater Treatment, Water Research and Technologies Centre, Technopark Borj Cedria, BP 273-8020 Soliman, Tunisia, Tel. +216 71493196; Fax: +216 71289569; email: taoufikguesmi@gmail.com (T. Gasmi), Tel. +216 79325122; Fax: +216 79325802; emails: imen.khouni@yahoo.fr (I. Khouni), a.ghrabi@yahoo.fr (A. Ghrabi)*
[b]*National Agronomic Institute of Tunisia, University of Carthage, 43 Avenue Charles Nicolle, Mahrajène, 1082 Tunis, Tunisia*

Received 12 August 2015; Accepted 22 January 2016

ABSTRACT

Heavy metal pollution is a major environmental problem worldwide because of the longstanding toxicity and bioaccumulation of metals. The risk is pronounced in the environment adjacent to large industrial complexes and cities historically located along rivers. The aim of this paper was the application of multivariate statistical techniques: cluster analysis (CA), principal component analysis (PCA), and discriminant analysis (DA) to evaluate spatial and temporal variations of the heavy metals fraction of Pb, Cu, Zn, Fe, Ni, Cr, and Al during monitoring of surface water of Wadi El Bey in northern part of Tunisia. Water samples were collected seasonally from 13 sites along the Wadi during 2 years (2012–2013). Results indicated that the concentrations and distribution of heavy metals appear to be largely controlled by natural processes, anthropogenic activities, and climatic conditions. CA classified the sampling sites into three clusters: Lowly Polluted sites "LP" (S2, S3 and S6–S11), Moderately Polluted sites "MP" (S4 and S5), and Highly Polluted Site "HP" (S1) based on similarities of water quality characteristics. PCA applied to the data-sets indicated that the total variance of water quality in the HP, MP, and LP sites were about 94.91, 91.06, and 80.14%, respectively. DA results seem to identify Cu, Fe, Zn, and Al as the most significant parameters for discrimination between seasons; whereas, Pb, Cu, Ni, Cr, and Al were the most significant parameters used to control the spatial variation. Thus, DA allowed reduction in dimensionality of the large data-set, delineating a few indicator parameters responsible for large variations in water quality. This study presents necessity and usefulness of multivariate statistical techniques for evaluation and interpretation of large complex data-sets with a view to get better information about the water quality and design of monitoring network for effective management of water resources.

Keywords: Heavy metals; Multivariate analysis; Tunisia; Wadi El Bey; Water quality

*Corresponding author.

Presented at the Sustainable Water Management Conference on Sustainable Domestic Water Use in Mediterranean Regions (SWMED), 19–21 February 2015, Tunis, Tunisia

1. Introduction

Urban areas and industrial activities as well as mineral exploitation contribute to the introduction of various types of contaminants, affecting the aquatic environment where they are deposited [1]. The pollution of riverine systems with different types of contaminants constitutes a major issue for water resources [2–6]. Heavy metals are one of the most common pollutants which have severely deteriorated the aquatic ecosystems [7,8]. This type of pollution can contaminate the river through many pathways, including discharge of contaminated effluents, terrestrial runoff, agricultural activities, and atmospheric deposition [9]. Unlike organic pollutants in water, heavy metals are not degraded through biological process [10] and then leads to a serious human health risk through the food chain and the loss of biodiversity [11]. This topic has attracted and interested many researchers all over the world [12–16].

Due to its arid to semi-arid climate, Tunisia is facing water scarcity problems, where the estimated available freshwater is only about 450 m^3/citizen/year [17]. There is an urgent need to preserve hydraulic resources, to tackle water penury and to fulfill part of the increased demand. The watershed of Wadi El Bey is one of the most vulnerable areas to pollution, influenced by anthropogenic and natural processes. This watershed (475 km^2) is located in the northeast part of Tunisia and drains Grombalia, Beni Khaled, and Soliman plains [18]. It is located between the Jebal Bouchoucha and Jebal Halloufa in the west, Jebal Abderrahman in the east, Jebal Reba El Ain in the south, and the Gulf of Tunis in the north [19]. The Wadi El Bey river has been very polluted by a variety of contaminant sources such as industrial effluents, domestic wastewater, agriculture drainage, and non point sources. Therefore, prevention of Wadi El Bey water quality requires effective monitoring of several heavy metals parameters.

The multivariate statistical methods are used in environmental studies dealing with measurement and monitoring. These methods have been used to identify important components and sources of the pollution that explain the variations in water quality and influence the water ecosystems [20]. The application of different multivariate statistical techniques such as cluster analysis (CA), principal component analysis (PCA), factor analysis (FA), and discriminant analysis (DA) facilitates the interpretation of complex data matrices to better understand the water quality and the ecological status of studied systems [21].

Few investigators have worked on heavy metal pollution of Wadi El Bey river [18]. However, there is no report pertaining on the use of multivariate statistical methods to evaluate spatial and temporal variations in total heavy metals pollution of Wadi El Bey, hence the need of this study. The aims of this investigation were: (i) to determine the total fraction of each heavy metal (Pb, Cu, Zn, Fe, Ni, Cr, and Al) in the surface water of Wadi El Bey, (ii) to assess the seasonal and spatial distribution of heavy metals pollution and (iii) to identify the sources of heavy metals contamination.

2. Materials and methods

2.1. Monitoring area

The watershed of Wadi El Bey (Fig. 1) is located in the northeast of Tunisia between 36° 35′ 00′′–36° 42′ 00′′ N and 10° 28′ 00′′–10° 33′ 00′′ E [22]. It has an area of 475 km^2 and drains the Grombalia plain. The two major affluents of Wadi El Bey are Wadi Maleh and Wadi Tahouna, both are controlled by dam's hill. Wadi El Bey runs through the agricultural lands; and it is used to discharge the treated wastewater and several industrial effluents (textile, tannery, and food industries).

Thirteen monitoring sites (Fig. 1, Table 1) were selected along the Wadi representing the pollution sources and the discharging industrial effluents.

2.2. Samples collection

Water samples were collected seasonally (twice per season) during two years (2012–2013) from the surface water of Wadi El Bey and/or pollution sources directly discharged into the Wadi and stored into pre-acid-cleaned polyethylene bottles of one liter capacity. Unfiltered samples were immediately acidified with nitric acid (HNO_3) (5 mL ultrapure HNO_3 was added to 1 L of samples) and were conserved in the dark at 4°C until analysis. Before analysis, the samples were acid digested and prepared according to standard methods [23,24].

2.3. Analytical determinations

Seven heavy metals (Pb, Cu, Zn, Fe, Ni, Cr, and Al) were analyzed according to the NF T90–210 using a spectrometer ICP-AES ULTIMA-C brand Jobin–Yvon.

2.4. Data treatment and multivariate statistical methods

Multivariate analysis of the water quality data-set was performed through CA, PCA, and DA [25].

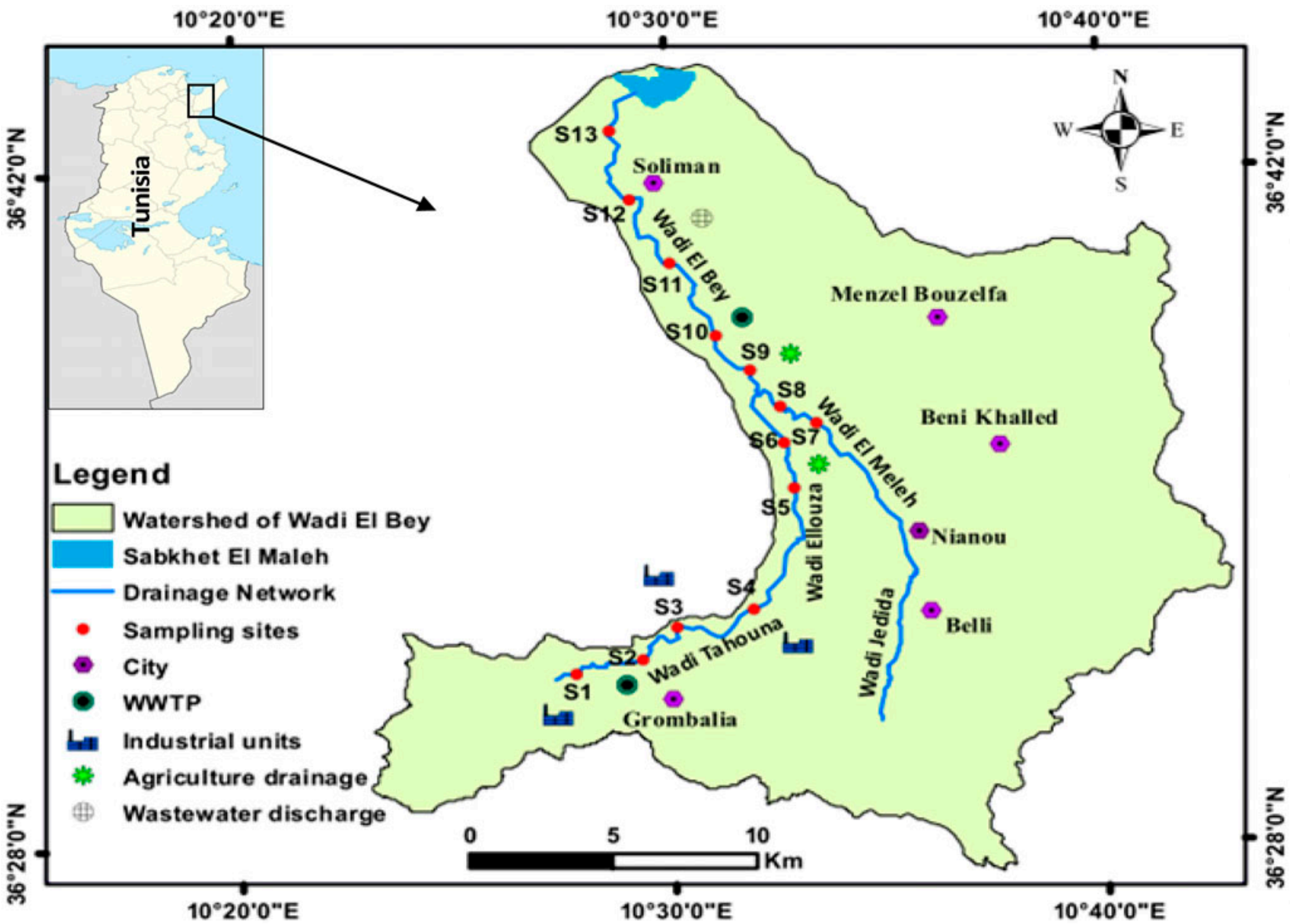

Fig. 1. Map of the watershed of Wadi El Bey and sampling locations (ArcGIS 10.1).

Table 1
Identification of sampling sites along Wadi El Bey river

Sites	Identification of sampling sites	Latitude	Longitude
S1	Discharge of Grombalia industrial area	36°36´37.18´´N	10°29´57.70´´E
S2	Discharge of beer industry (SONOBRA)—Grombalia	36°36´38.52´´N	10°29´59.15´´E
S3	Treated wastewater—WWTP Grombalia.	36°36´39.00´´N	10°30´00.42´´E
S4	Tannery discharges—Grombalia	36°36´40.81´´N	10°30´2.16´´E
S5	Bridge over Wadi El Bey	36°38´24.18´´N	10°30´59.70´´E
S6	Agricultural drainage—Wadi Maleh	36°38´29.96´´N	10°31´6.97´´E
S7	Bridge over Wadi Maleh before Wadi El Bey	36°38´29.55´´N	10°31´5.14´´E
S8	Wadi El Bey Before WWTP discharges—Béni Khaled	36°39´32.52´´N	10°30´37.49´´E
S9	Agricultural drainage—Béni Khaled	36°39´34.56´´N	10°30´36.75´´E
S10	WWTP discharges—Béni Khaled	36°39´35.18´´N	10°30´35.75´´E
S11	Bridge over Wadi El Bey, road Soliman—Grombalia	36°41´3.39´´N	10°29´34.96´´E
S12	Bridge over Wadi El Bey, Soliman city	36°41´42.51´´N	10°28´47.19´´E
S13	Wadi El Bey Before Sebkha	36°43´22.03´´N	10°28´21.94´´E

Mathematical and statistical computations were made using Microsoft Office Excel and XL STAT (2013). The ANOVA test (level of significance $\alpha = 0.05$) was employed to understand the spatial and seasonal variation in the heavy metal concentrations.

2.5. Cluster analysis

CA is an unsupervised pattern recognition technique that uncovers intrinsic structure or underlying behavior of data-set without making a priori assumption about the data, in order to classify the objects of

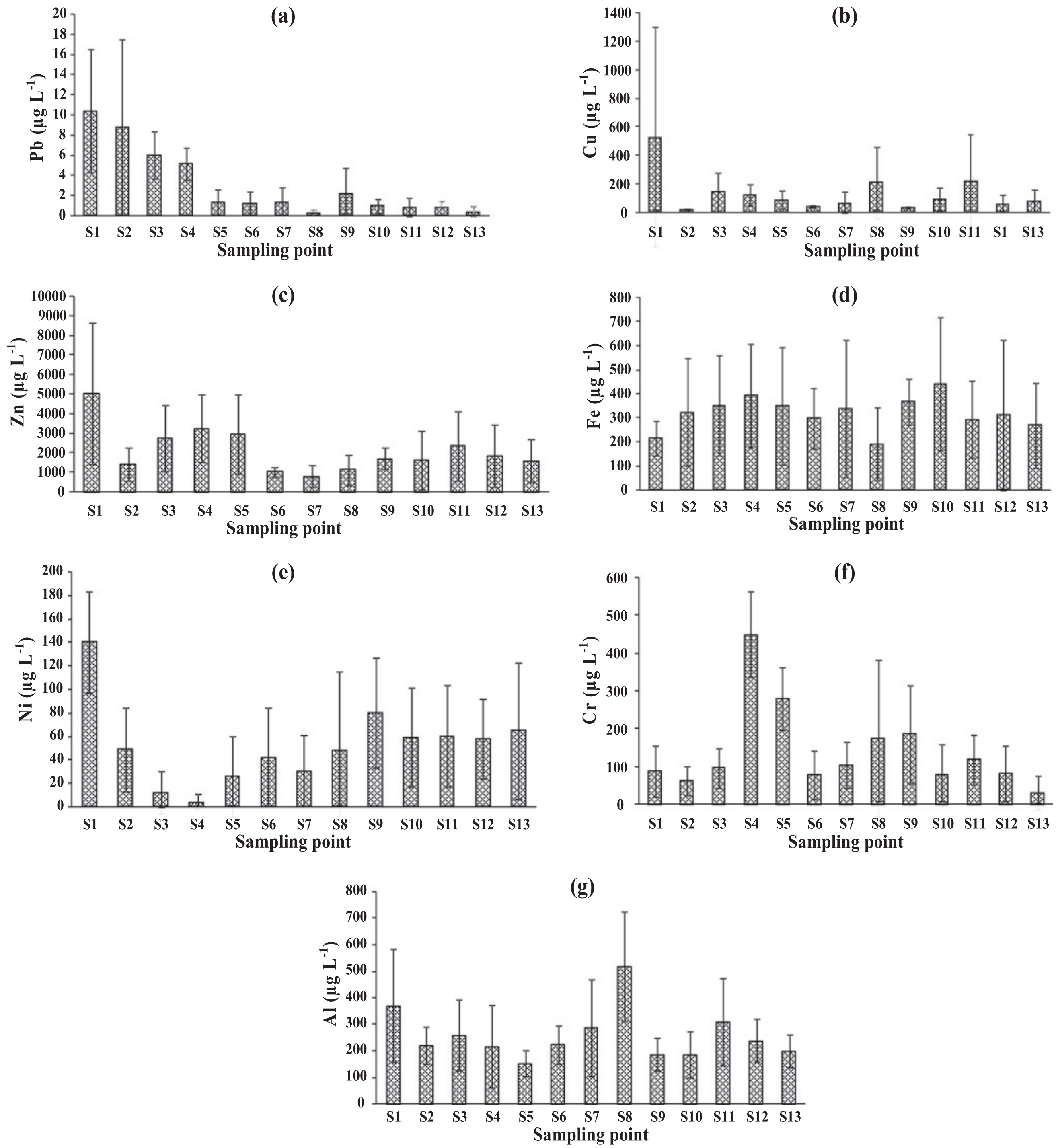

Fig. 2. Spatial variation of total fraction of heavy metals during the period study: Pb (a), Cu (b), Zn (c), Fe (d), Ni (e), Cr (f), and Al (g).

the system into categories or clusters based on their nearness or similarity. Hierarchical agglomerative CA is performed on the normalized data-set by means of the Ward's method, using Euclidean distances as a measure of similarity.

CA is applied to the water quality data-set with view to group the similar sampling sites spread over the stretch and in the resulted dendrogram [26,27].

2.6. Principal component analysis

The goals of PCA are to: (1) extract the most important information from the data; (2) to compress the size of the data-set by keeping only this important information; (3) to simplify the description of the data-set; and (4) to analyze the structure of the observations and the variables. Based on Varimax rotation (raw), these principal components (PCs) are transferred to varifactors (VFs), which provides information on the most meaningful variables that describe a whole data-set while allowing data reduction with minimum loss of original information [28]. Varimax rotation is a change of coordinates used in PCA that maximizes the sum of the variances of the squared loadings. Thus, all the coefficients (squared correlation with factors) will be either large or near zero, with few intermediate values. The goal is to associate each variable to at most one factor. The interpretation of the results of the PCA will be simplified. Then each variable will be associated to one and one only factor, they are split into disjoints sets.

2.7. Discriminant analysis

The DA consists of examination of the dependence of one qualitative (classification) variable from several quantitative variables [29]. DA is used to determine the variables, which discriminate between two or more naturally occurring groups [30]. The DA technique builds up a discriminant function for each group, which operates on raw data and the technique constructs a discriminant function for each group [31,32].

3. Results

All the water samples collected from the different monitoring sites of the Wadi El Bey were analyzed and heavy metals Pb, Cu, Zn, Fe, Ni, Cr, and Al concentrations were determined. Range, mean, and standard deviations are illustrated in Fig. 2. The spatial variation in heavy metals indicated maximum values of Pb (10.37 μg L^{-1}); Cu (527.27 μg L^{-1}), Zn (5,050.5 μg L^{-1}), and Ni (140.5 μg L^{-1}) at site S1 (Discharge of Grombalia industrial area), those of Cr (449.55 μg L^{-1}), Al (518.43 μg L^{-1}) and Fe (439.4 μg L^{-1}) were found at the sites S4 (Tannery discharges—Grombalia), S8 (Wadi El Bey Before WWTP discharges—Béni Khaled), and S10 (WWTP discharges—Béni Khaled), respectively. These findings were higher than those reported in literature on the same river [18]. Zn and Fe were the most abundant elements in the Wadi, whereas Ni and Pb were the less abundant ones. Each element exhibits a wide variation as reflected by the large standard deviation values indicative of seasonal changes [33].

It is known that Zn is the most abundant element found in the earth's crust [34]. In the present investigation, among the seven heavy metals analyzed, Zn was the most abundant all along the Wadi El Bey river with a mean level of 2,131.67 μg L^{-1}. The mean order of heavy metals was Zn (2,131.67 μg L^{-1}), Fe (317.19 μg L^{-1}), Al (260.3 μg L^{-1}), Cr (140.93 μg L^{-1}), Cu (130.96 μg L^{-1}), Ni (51.71 μg L^{-1}), and Pb (3.01 μg L^{-1}). The concentration of Pb (Fig. 2(a)) ranged from 0.26 to

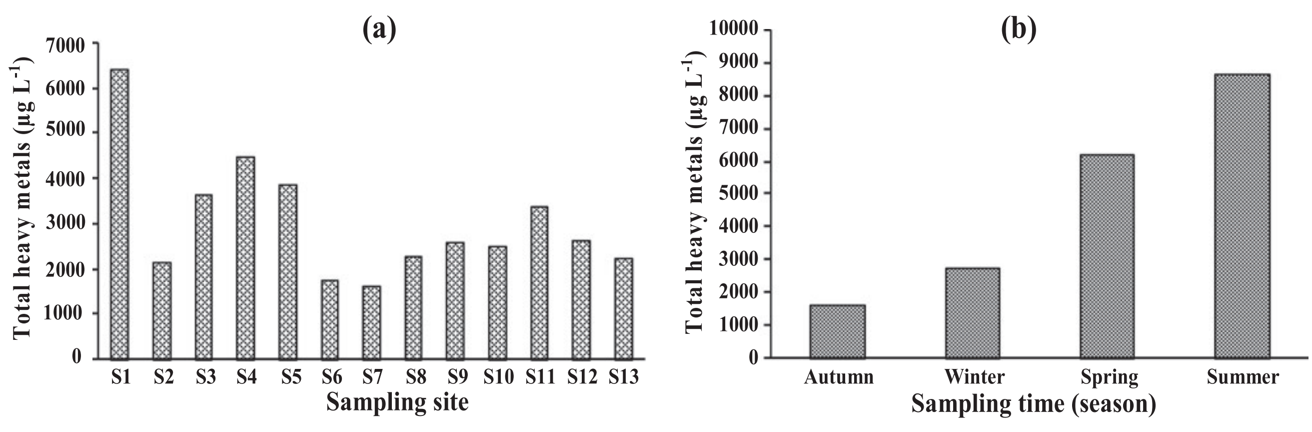

Fig. 3. Sum concentrations of total fraction of heavy metals at the different monitored sites (a) and season (b).

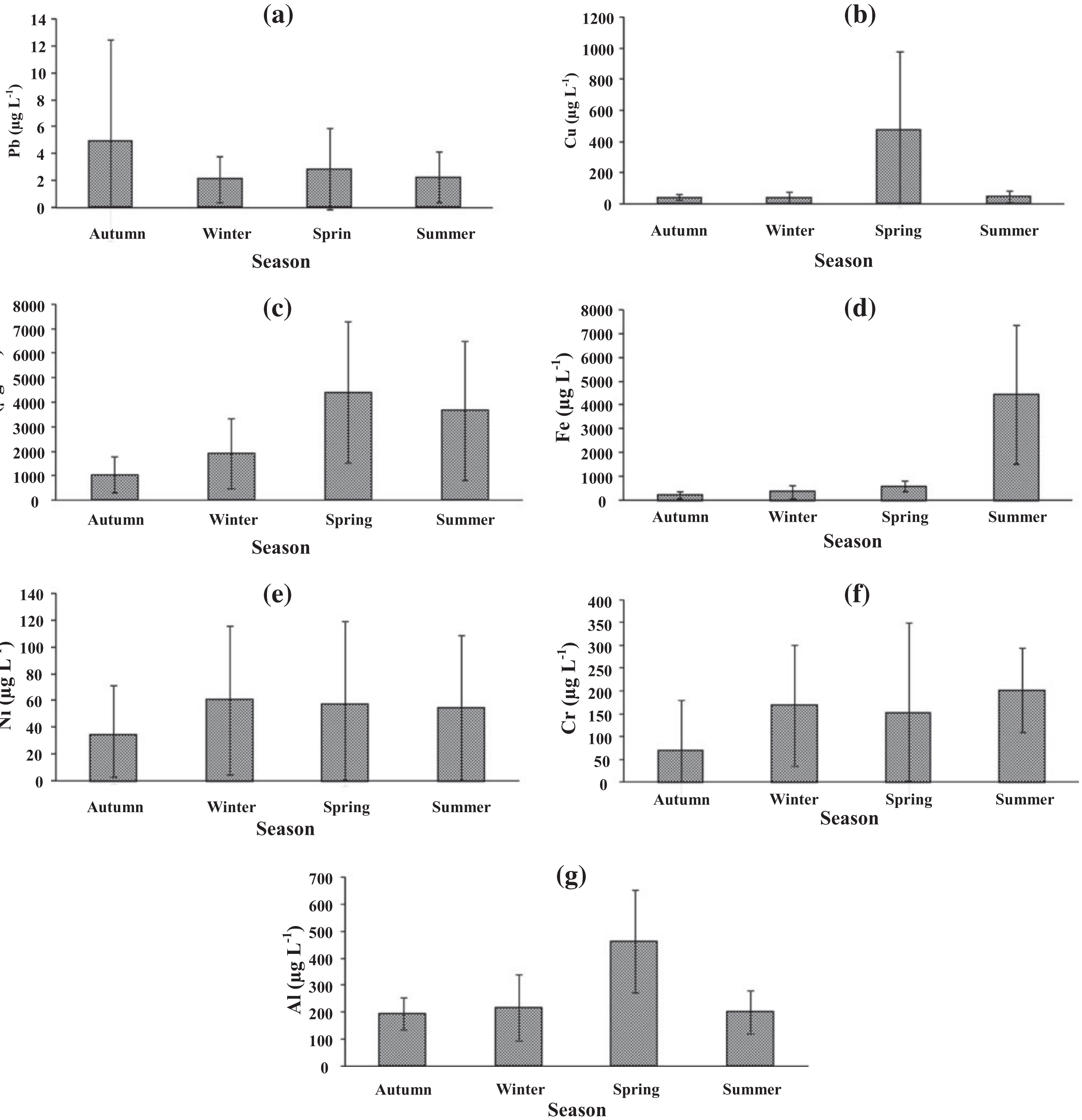

Fig. 4. Seasonal variation of total fraction of heavy metals during the period study: Pb (a), Cu (b), Zn (c), Fe (d), Ni (e), Cr (f), and Al (g).

10.37 µg L^{-1}. The highest Pb value was observed on site S1 (Discharge of Grombalia industrial area) and then decreased gradually. The principal sources of Pb contamination were the Grombalia industrial discharge and the WWTP effluents (mainly textile effluent). Cu content ranged from 21.33 to 527.27 µg L^{-1} (Fig. 2(b)). This clearly indicated that anthropogenic input was the major source of Pb and Cu.

The mean concentration of Zn (Fig. 2(c)) ranged from 809.75 to 5,050.5 µg L^{-1}. The highest value of Zn depicted in site S1 (Discharge of Grombalia industrial area) has indicated that industrial discharge was the principal source of Zn. It was comparatively high at sites S3 (Treated wastewater—WWTP Grombalia), S4 (Tannery discharges—Grombalia) and S5 (Bridge over Wadi El Bey) and was relatively low at the sites S6

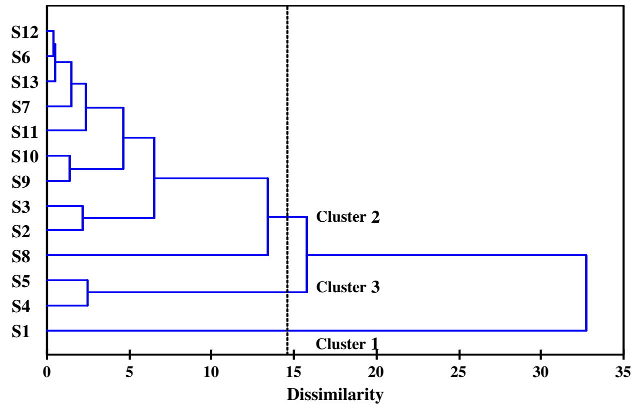

Fig. 5. Dendrogram showing clustering of sampling sites according to Ward's method using squared Euclidean distance.

(Agricultural drainage—Wadi Maleh), S7 (Bridge over Wadi Maleh before Wadi El Bey), and S8 (Wadi El Bey Before WWTP discharges-Béni Khaled). Zn can enter the aquatic environment from a number of sources like industrial discharges, sewage effluents, and terrestrial runoff [35].

Concentration of Fe (Fig. 2(d)) was between 188.75 and 439.2 μg L^{-1}. The distribution of the iron was approximately similar on all the sites. The mean concentration of Ni (Fig. 2(e)) ranged from 3.25 to 140.5 μg L^{-1}. Two highest values of Ni were observed in sites S1 (Discharge of Grombalia industrial area) and S9 (Agricultural drainage—Béni Khaled). The values observed in site S1 (Discharge of Grombalia industrial area) were due to the wide use of this metal and its alloys in the metallurgical, chemical, and food processing industries, especially as catalysts and pigments. Whereas, the values observed in S9 (Agricultural drainage—Béni Khaled) were due to the excessive application of Ni-containing phosphate fertilizers in agricultural activities. In the rest of the sites, Ni concentration was low, because this element is easily accumulated in the phytoplankton or other aquatic plants and it can be deposited in the sediment by many processes such as precipitation, complexation, and adsorption on clay particles [36].

The concentration of Cr (Fig. 2(f)) ranged from 31.85 to 449.55 μg L^{-1}. The highest values were observed at site S4 (Tannery discharges—Grombalia) and were due to the tannery discharge. Indeed, Cr was the most commonly used tanning agent. Nearly 90% of all leather produced were tanned using Cr [37]. The Al values varied from 153.5 to 518.43 μg L^{-1} (Fig. 2(g)). The repartition of Al was homogeneous in the totality, except for the site S1 due to industrial discharge and the site S8 (Wadi El Bey Before WWTP discharges—Béni Khaled) due to the direct deposit of wastes from urban sites adjacent. The abundance of Al along the Wadi was due to metallurgical industry and excess usages of coagulant and flocculent agents to eliminate particulate matter including micro-organisms and soluble organic matter in water treatment plant. These results are comparable to those reported by others studies [38–44].

The relative abundance of heavy metal in the different sampling sites (Fig. 3(a)) was in decreasing order of: S1 > S4 > S5 > S3 > S11 > S12 > S10 > S8 > S13 > S2 > S6 > S7.

The highest total heavy metal concentrations in the Wadi El Bey were recorded at the industrial area of Grombalia (S1) and the tannery discharges (S4). The lowest total heavy metals concentrations were recorded at the sites S6 (Agricultural drainage—Wadi Maleh) and S7 (Bridge over Wadi Maleh before Wadi El Bey), due to dilution by agricultural drainage which indicated that Wadi El Maleh effluent was less polluted (LP) than that of Wadi El Bey. The results of the seasonal variations in total fraction of heavy metal (mean ± SD) showed significant temporal variations in all the heavy metals studied (Fig. 4). In fact, the maximum values of Pb appeared in autumn, Ni in winter, Zn, Cu, and Al in spring, whereas maximum Cr and Fe values appeared in summer.

The variation in total heavy metals concentrations in the surface water of the Wadi (Fig. 3(b)) followed an increasing order from autumn to summer (autumn: 1,615.95 μg L^{-1}; winter: 2,774.39 μg L^{-1}; spring: 6,171.58 μg L^{-1} and summer: 8,604.58 μg L^{-1}). The high level of heavy metals (Cu, Zn, Fe, Cr, and Al) in summer could be attributed to the intense evaporation. In the wet seasons (autumn and winter), dilution by high precipitation was the most plausible explanation for the low heavy metals concentrations, while Pb and Ni have similar concentrations whatever the season. This was due to their mixed sources, to natural contributions and to human activities such as agriculture and industrial activities. It should be noted that this result is consistent with other previous studies [45–47].

4. Discussion

The hierarchical CA, DA, and PCA were applied to evaluate heavy metals pollution status of Wadi El Bey.

4.1. Hierarchical CA and site grouping

The most similar points were grouped forming one cluster and the process was repeated until all points

Table 2
Classification functions for DA of temporal variation in heavy metals levels (total fraction) in the Wadi El Bey

	Standard mode				Forward stepwise mode				Backward stepwise mode			
Parameters	Autumn coefficient	Winter coefficient	Spring coefficient	Summer coefficient	Autumn coefficient	Winter coefficient	Spring coefficient	Summer coefficient	Autumn coefficient	Winter coefficient	Spring coefficient	Summer coefficient
Pb	0.247	0.050	0.024	0.032								
Cu	−0.005	−0.006	−0.001	−0.008					0.001	0.000	0.008	−0.002
Fe	0.007	0.008	0.015	0.004	0.008	0.010	0.016	0.007	0.008	0.009	0.017	0.006
Zn	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001
Ni	0.010	0.022	0.004	0.013								
Cr	0.002	0.007	0.005	0.008								
Al	0.015	0.018	0.026	0.019	0.012	0.013	0.025	0.012				
Constant	−4.434	−5.918	−11.978	−6.204	−3.518	−4.498	−11.715	−4.554	−2.364	−3.103	−7.929	−3.491

Table 3
Classification matrix for DA of temporal variation in total fraction of heavy metals

Monitoring seasons	% Correct	Season assigned by DA			
		Autumn	Winter	Spring	Summer
Standard DA mode					
Autumn	76.92	10	3	0	0
Winter	53.85	2	7	1	3
Spring	76.92	0	3	10	0
Summer	38.46	2	5	1	5
Total	61.54	14	18	12	8
Forward stepwise DA mode					
Autumn	76.92	10	2	0	1
Winter	15.38	6	2	2	3
Spring	69.23	0	4	9	0
Summer	38.46	3	4	1	5
Total	50	19	12	12	9
Backward stepwise DA mode					
Autumn	69.23	9	1	1	2
Winter	23.08	5	3	2	3
Spring	69.23	1	3	9	0
Summer	46.15	3	4	0	6
Total	51.92	18	11	12	11

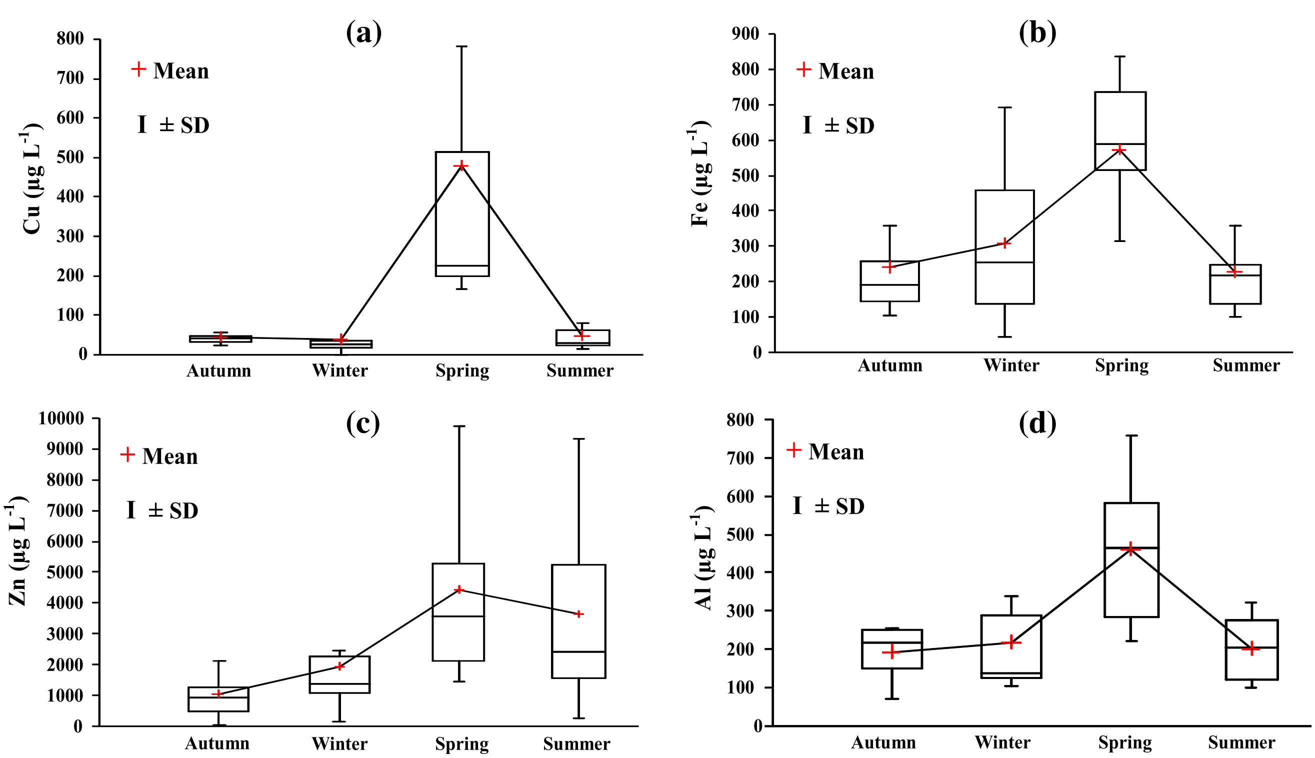

Fig. 6. Box plots for temporal variation in total fraction of heavy metals: Cu (a), Fe (b), Zn (c), and Al (d).

Table 4
Classification functions for DA of spatial variation in heavy metals levels (total fraction) in the Wadi El Bey

	Standard mode			Forward stepwise mode			Backward stepwise mode		
Parameters	HP coefficient	MP coefficient	LP coefficient	HP coefficient	MP coefficient	LP coefficient	HP coefficient	MP coefficient	LP coefficient
Pb	1.296	0.340	0.385	0.844	0.292	0.220	0.816	0.294	0.185
Cu	−0.021	0.002	−0.014				−0.001	0.003	−0.007
Fe	−0.007	0.004	0.003						
Zn	0.001	0.000	0.000						
Ni	0.195	−0.020	0.068	0.091	−0.022	0.024	0.091	−0.026	0.032
Cr	−0.024	0.038	−0.015	−0.002	0.039	0.007	−0.006	0.040	0.000
Al	0.046	−0.006	0.034				0.016	−0.001	0.019
Constant	−25.199	−9.039	−7.211	−13.258	−9.295	−1.518	−15.561	−9.624	−3.424

Table 5
Classification matrix for DA of spatial variation in total fraction of heavy metals levels

		Regions assigned by DA		
Monitoring regions	% Correct	HP	MP	LP
Standard DA mode				
HP	100	17.33	0	0
MP	100	0	17.33	0
LP	90	1.3	0.43	15.6
Total	96.67	18.63	17.76	15.6
Forward stepwise DA mode				
HP	100	4	0	0
MP	75	0	6	2
LP	95	1	1	38
Total	92.31	5	7	40
Backward stepwise DA mode				
HP	50	2	0	2
MP	100	0	8	0
LP	97.5	1	0	39
Total	94.23	3	8	41

belong to a cluster. The results obtained were presented by dendrogram (Fig. 5) [48]. In our case, CA allows to group water samples based on the similarity of their chemical composition. The dendrogram using CA displayed important information about the grouping of the 13 monitored sites into 3 different clusters that were identified as cluster 1, cluster 2, and cluster 3. Cluster 1, represented by S1 (Discharge of Grombalia industrial area) was characterized with very high levels of Pb, Cu, Zn, and Ni concentrations. It corresponds to high-polluted site (HP). This site is situated at upstream of the Wadi El Bey and it receives pollution from industrial discharge. The cluster 2, grouping S2, S3, and S6–S11, was characterized by low metals concentrations and corresponds to the LP sites situated at the middle and the downstream of the Wadi. These sites received diffuse pollution from agricultural drainage, domestic wastewater, and runoff from villages. The cluster 3 (S4 and S5) was characterized essentially by high concentration of Cr and it corresponded to relatively moderate pollution sites (MP), which received pollution from tannery discharges and runoff from Grombalia. CA provides a useful classification of the monitoring sites and can be used to design an optimal spatial monitoring network with lower costs [38,49].

4.2. DA: temporal and spatial variation of heavy metals

Temporal variations were further evaluated through DA ($p < 0.01$). The classification of functions (DFs) and matrices (CMs) for DA of temporal concentrations of heavy metals in the Wadi El Bey obtained from the standard modes are shown in Tables 2 and 3. According to the standard DA mode using all discriminant variables, the corresponding CMs assigned 61.54% of the cases correctly. However, in forward stepwise and backward modes, DA gave CMs with 50 and 51.92% correct assignations, respectively, with little difference between seasons. Consequently, the temporal DA results seem to identify Cu, Fe, Zn, and Al as the most significant parameters for discrimination between seasons. As identified by DA, box plots of the selected parameters (Fig. 6) showed seasonal trends. In the wet seasons, heavy metal concentrations decrease by dilution. In dry season, essentially in spring, the phenomenon of evaporation increased the metal levels. However, in summer when temperature is near 35°C, evaporation was maximal and the flow was very low. In these conditions, concentrations of suspended solids and heavy metals in the reservoir

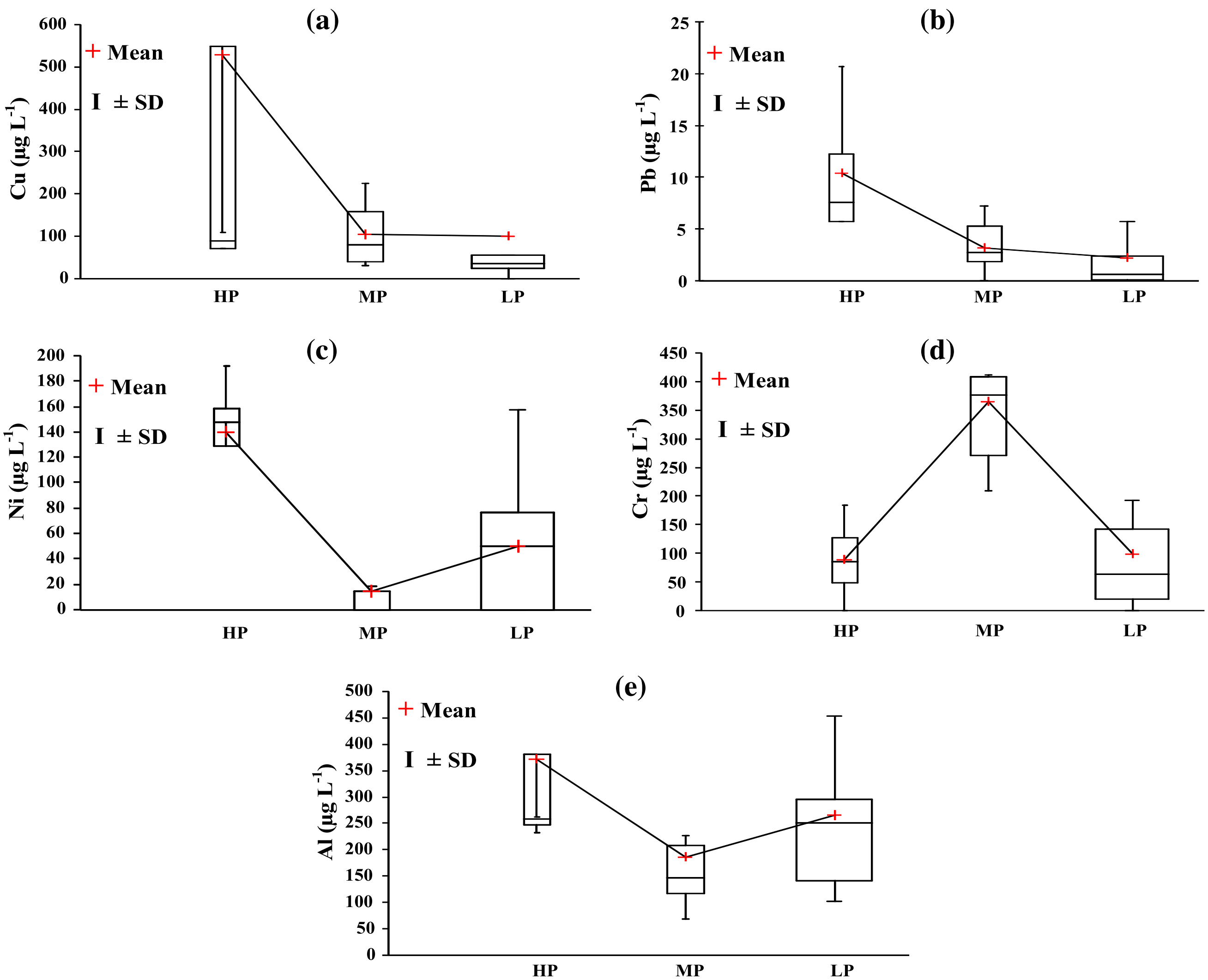

Fig. 7. Box plots for spatial variation in total fraction of heavy metals: Cu (a), Pb (b), Ni (c), Cr (d), and Al (e).

water will be decreased significantly due to sediment deposition [50,51]. It was clear with iron and copper, which were compatible with the metals order of sedimentation establishes by Varol and Şen [43].

After grouping into three classes (HP, MP, and LP), the classification of functions (DFs) and matrices (CMs) are illustrated in Tables 4 and 5. As the temporal DA, the spatial standard DA mode including seven parameters gives CMs with 96.67% cases correctly. The forward stepwise mode assigned 92.31% cases correctly, whereas the backward stepwise mode showed Cu, Pb, Ni, Cr, and Al with CMs corresponding to 94.23%.

To evaluate heavy metals associated with spatial variations in Wadi El Bey, the Box and whisker plots of discriminating parameters identified by backward stepwise mode were constructed (Fig. 7). Cu, Pb, Ni, and Al have the highest values in the HP sites due to the discharges of Grombalia industrial area (S1). Cr is the highest in the MP sites, which was due to the discharge of the tannery industry. In LP sites, values of the five metals were the lowest; this was due to dilution and the nonpoint sources like agriculture activities and urban runoff.

4.3. Data structure determination and source identification

PCA allows identifying the different groups of metals that correlate and may be considered as having a similar behavior and common origin [52]. According to Table 6, two PC at 94.91% explained the information treated for HP, three PC at 91.06% for MP and four PC at 80.14% for LP. For all monitored sites classified under HP group, the first component (VF1),

Table 6
Loading of experimental variables (7) on significant PCs for HP sites, MP sites, and LP sites data-sets

Variables	VF1	VF2	VF3	VF4
HP sites (two significant principal components)				
Pb	0.414	**0.539**		
Cu	**0.628**	0.370		
Fe	**0.974**	0.000		
Zn	**0.794**	0.001		
Ni	**0.947**	0.021		
Cr	0.161	**0.794**		
Al	**0.648**	0.352		
Eigenvalue	4.567	2.078		
% Total variance	65.23	29.679		
Cumulative % variance	65.23	94.916		
MP sites (three significant principal components)				
Pb	0.000	**0.641**	0.239	
Cu	**0.980**	0.000	0.003	
Fe	**0.907**	0.021	0.014	
Zn	0.246	0.229	**0.452**	
Ni	0.162	**0.575**	0.035	
Cr	**0.703**	0.205	0.011	
Al	**0.654**	0.119	0.178	
Eigenvalue	3.653	1.790	0.931	
% Total variance	52.18	25.574	13.306	
Cumulative % variance	52.18	77.762	91.069	
LP sites (four significant principal components)				
Pb	0.004	0.050	**0.873**	0.029
Cu	**0.753**	0.067	0.003	0.072
Fe	0.197	0.023	0.050	**0.649**
Zn	**0.512**	0.071	0.053	0.048
Ni	0.048	**0.566**	0.017	0.124
Cr	0.020	**0.663**	0.001	0.023
Al	**0.633**	0.043	0.016	0.002
Eigenvalue	2.167	1.483	1.014	0.947
% Total variance	30.95	21.18	14.48	13.52
Cumulative % variance	30.95	52.13	66.62	80.14

Note: Bold values represent correlation with significance.

explained 65.23% of total variance with a strong positive correlation with metals Fe, Ni, and a little less with Cu, Zn, and Al. These factors mainly represent the contribution of different sources of pollution like industrial effluent, domestic wastewater, and agricultural drainage. The second component (VF2) showed 29.68% of the total variance that was associated with Pb and Cr. This component was also attributed to the industrial discharge, in particular the tannery industry.

For MP group of sites, the first component (VF1) showed 52.18% of total variance, and it has a strong positive correlation with Cu and Fe and less with Cr and Al. VF2, explaining 25.57% of the total variance, has a positive correlation with Pb and Ni. For LP group, the first component explaining 30.95% was correlated positively with Cu, Zn, and Al (industrial discharges).

Ni and Cr were correlated to the second component with 21.18% of total variance and linked to tannery effluent; Pb was correlated to the third component and Fe to the fourth with 14.48 and 13.52% of total variance, respectively. They were characterized by wastewater treatment plants, domestic wastewater, and agricultural drainage effluents.

5. Conclusion

In the present study, total fraction concentrations of seven heavy metals in surface water of Wadi El Bey were studied. To investigate spatial and temporal variations in water quality among the Wadi, different multivariate statistical analysis (CA, PCA, and DA) were used. Hierarchical CA grouped 13 sampling sites into 3 clusters of similar water quality characteristics: lowly polluted sites "LP" (S2, S3 and S6-S11), moderately polluted sites "MP" (S4 and S5), and highly polluted Site "HP" (S1). Based on obtained information, it will be possible to design a future, optimal sampling strategy, which could reduce the number of sampling stations and associated costs. To evaluate heavy metals associated with spatial and temporal variations in Wadi El Bey, the Box and whisker plots of discriminating parameters were constructed. Although the PCA did not result in a significant data reduction, it helped extract and identify the factors/sources responsible for variations in river water quality at three different sampling sites. PCA was performed on the normalized data-sets separately for the three different clusters (HP, MP, and LP). It was applied to compare the compositional patterns between the analyzed water samples and to identify the factors that influence each one. These factors represent the contribution of different pollution sources such as the industrial effluent, domestic wastewater, and agricultural drainage. DA gave the best results both spatially and temporally. For three sampling clusters on the Wadi, DA yielded an important data reduction, as it used only four heavy metals (Cu, Fe, Zn, and Al) affording about 50% correct assassinations in temporal variation, and five heavy metals (Cu, Pb, Ni, Cr and Al) affording more than 90% correct assassinations in spatial variation. Therefore, DA allowed a reduction in the dimensionality of the large data-set, delineating a few indicator parameters responsible of the large variations in water quality. Hence, this study illustrated the usefulness of multivariate statistical techniques for analysis and interpretation of complex data-sets, and in water quality assessment, identification of pollution sources/factors, and understanding temporal/spatial variations in water quality for effective river water quality management. Otherwise, the present study suggested that the Wadi El Bey should be given priority for effective waste management in order to sustain the ecological integrity.

Acknowledgements

This work is funded and conducted under the frame of objectives contact between Water Research and Technologies Centre (CERTE) and Ministry of Higher Education, Scientific Research.

Abbreviations

CA	—	cluster Analysis
DA	—	discriminant analysis
FA	—	factor analysis
HCA	—	hierarchical cluster analysis
HP	—	highly polluted
LP	—	lowly polluted
MP	—	moderately polluted
PCA	—	principal component analysis
VF	—	varimax factors
WWTP	—	wastewater treatment plant

References

[1] E.K. Atibu, N. Devarajan, F. Thevenon, P.M. Mwanamoki, J.B. Tshibanda, P.T. Mpiana, K. Prabakar, J.I. Mubedi, W. Wildi, J. Poté, Concentration of metals in surface water and sediment of Luilu and Musonoie Rivers, Kolwezi-Katanga, Democratic Republic of Congo, Appl. Geochem. 39 (2013) 26–32.

[2] S.T.Y. Tong, W. Chen, Modeling the relationship between land use and surface water quality, J. Environ. Manage. 66 (2002) 377–393.

[3] M. Matthies, J. Berlekamp, S. Lautenbach, N. Graf, S. Reimer, System analysis of water quality management for the Elbe river basin, Environ. Model. Software 21 (2006) 1309–1318.

[4] M.M.V. Landeghem, M.D. Meyer, S.B. Cox, B. Sharma, R. Patiño, Spatial and temporal patterns of surface water quality and ichthyotoxicity in urban and rural river basins in Texas, Water Res. 46 (2012) 6638–6651.

[5] X. Zhai, J. Xia, Y. Zhang, Water quality variation in the highly disturbed Huai River Basin, China from 1994 to 2005 by multi-statistical analyses, Sci. Total Environ. 496 (2014) 594–606.

[6] V.C. Viswanathan, J. Molson, M. Schirmer, Does river restoration affect diurnal and seasonal changes to surface water quality? A study along the Thur River, Switzerland, Sci. Total Environ. 532 (2015) 91–102.

[7] Z. Ali, R.N. Malik, A. Qadir, Heavy metals distribution and risk assessment in soils affected by tannery effluents, Chem. Ecol. 29 (2013) 676–692.

[8] S. Nazeer, M.Z. Hashmi, R.F. Malik, Heavy metals distribution, risk assessment and water quality characterization by water quality index of the River Soan, Pakistan, Ecol. Indic. 43 (2014) 262–270.

[9] Y. Huang, W. Zhu, M. Le, X. Lu, Temporal and spatial variations of heavy metals in urban riverine sediment: An example of Shenzhen River, Pearl River Delta, China, Quat. Int. 282 (2012) 145–151.

[10] A.K. Hegazy, H.F. Kabiel, M. Fawzy, Duckweed as heavy metal accumulator and pollution indicator in industrial wastewater ponds, Desalin. Water Treat. 12 (2009) 400–406.

[11] S. Li, Q. Zhang, Spatial characterization of dissolved trace elements and heavy metals in the upper Han River (China) using multivariate statistical techniques, J. Hazard. Mater. 176 (2010) 579–588.

[12] C. Neal, M. Neal, L. Hill, H. Wickham, The water quality of the River Thame in the Thames Basin of south/south-eastern England, Sci. Total Environ. 360 (2006) 254–271.

[13] R. Cidu, R. Biddau, Transport of trace elements under different seasonal conditions: Effects on the quality of river water in a Mediterranean area, Appl. Geochem. 22 (2007) 2777–2794.

[14] A.K. Krishna, M. Satyanarayanan, P.K. Govil, Assessment of heavy metal pollution in water using multivariate statistical techniques in an industrial area: A case study from Patancheru, Medak District, Andhra Pradesh, India, J. Hazard. Mater. 167 (2009) 366–373.

[15] S. Li, Q. Zhang, Response of dissolved trace metals to land use/land cover and their source apportionment using a receptor model in a subtropic river, China, J. Hazard. Mater. 190 (2011) 205–213.

[16] Y. Yang, Z. He, Y. Wang, J. Fan, Z. Liang, P.J. Stoffella, Dissolved organic matter in relation to nutrients (N and P) and heavy metals in surface runoff water as affected by temporal variation and land uses—A case study from Indian River Area, south Florida, USA, Agric. Water Manage. 118 (2013) 38–49.

[17] M.H. Louati, R. Khanfir, A. Alouini, M. El Echi, L. Frigui, A. Marzouk, Guide pratique de gestion de la sécheresse en Tunisie (Practical Guide of drought Management in Tunisia), Rapport interne, Ministère de l'agriculture, 2000.

[18] S. Khadhar, A. Mlayah, A. Chekirben, A. Charef, M. Methammam, S. Nouha, Z. Khemais, Vecteur de la pollution métallique du bassin versant de l'Oued El Bey vers le Golfe de Tunis (Tunisie) Vector of metal pollution of the watershed of Wadi El Bey towards the Gulf of Tunis (Tunisia), Hydrol. Sci. J. 10 (2013) 835–487.

[19] M. Ennabli, Etude Hydrogéologique des Aquifères du Nord-est de la Tunisie, Pour une Gestion Intégrée des Ressources en eau (Hydrogeology Study of Northeastern Aquifers of Tunisia, for an Integrated Management of Water Resources), Thèse de Doctorat. Université. Nice, France, 1980.

[20] E.M. Andrade, H.A.Q. Palácio, I.H. Souza, R.A. Leão, M.J. Guerreiro, Land use effects in groundwater composition of an alluvial aquifer (Trussu River, Brazil) by multivariate techniques, Environ. Res. 106 (2008) 170–177.

[21] M. Varol, B. Gökot, A. Bekleyen, B. Şen, Spatial and temporal variations in surface water quality of the dam reservoirs in the Tigris River basin, Turkey, Catena 92 (2012) 11–21.

[22] DGRE, Direction Générale des Ressources en eau: Annuaire Piézométrique de la Tunisie (General Department of Water Resources: Piezometric directory of Tunisia), Tunis, 1990.
[23] USEPA, 1992, Method 3005A: Acid Digestion of Waters for Total Recoverable or Dissolved Metals for Analysis by FLAA or ICP Spectroscopy. Method 3005A, Revision 1, Washington, DC. Available from: <http://www.epa.gov/osw/hazard/testmethods/sw846/pdfs/3005a.pdf>.
[24] APH Association, AWW Association, Standard Methods for the Examination of Water and Wastewater, eighteenth ed., APHA, AWWA, American Water Works Association and Water Pollution Control Federation (WPCF), Washington, DC, 1990.
[25] F. Al-Badaii, M.S. Othman, Heavy metals and water quality assessment using multivariate statistical techniques and water quality index of the Semenyih, Peninsular Malaysia, Iranica J. Energy Environ. 5 (2014) 132–145.
[26] C. Hagedorn, S.L. Robinson, J.R. Filtz, S.M. Grubbs, T.A. Angier, J.R. Reneau, Determining sources of fecal pollution in a rural Virginia Watershed with antibiotic resistance patterns in fecal streptococci, Appl. Environ. Microbiol. 65 (1999) 5522–5531.
[27] V. Librando, Chemometric evaluation of surface water quality at regional level, Fresenius, J. Anal. Chem. 339 (2000) 613–619.
[28] B. Helena, R. Pardo, M. Vega, E. Barrado, J.M. Fernandez, L. Fernandez, Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis, Water Res. 34 (2000) 807–816.
[29] K. Kočišová, M. Mišanková, Discriminant analysis as a tool for forecasting company's financial health, Procedia—Social Behav. Sci. 110 (2014) 1148–1157.
[30] K.P. Singh, A. Malik, D. Mohan, S. Sinha, Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—A case study, Water Res. 38 (2004) 3980–3992.
[31] R.A. Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis, third ed., 642, Prentice–Hall International, Englewood Cliffs, New Jersey, USA, 1992.
[32] W.D. Alberto, D.M. Del Pilar, A.M. Valeria, P.S. Fabiana, H.A. Cecilia, B.M. de Los Angeles, Pattern Recognition Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality. A Case Study, Water Res. 35 (2001) 2881–2894.
[33] L. Siyue, Z. Quanfa, Spatial characterization of dissolved trace elements and heavy metals in the upper Han (China) using multivariate statistical techniques, J. Hazard. Mater. 176 (2010) 579–588.
[34] P.E. John De Zuane, Chemical Parameters Inorganic in Quality Standards and Constraints, Von Nostrand and Reinhold, New York, 1990, pp. 47–151.
[35] S.G. John, J.G. Park, Z. Zhang, E.A. Boyle, The isotopic composition of some common forms of anthropogenic zinc, Chem. Geol. 245 (2007) 61–69.
[36] M. Cempel, G. Nikel, Nickel: A review of its sources and environmental toxicology, Pol. J. Environ. Stud. 15 (2006) 375–382.
[37] M.I.C. Monteiro, I.C.S. Fraga, A.V. Yallouz, N.M.W. De Oliveira, S.H. Ribeiro, Determination of total chromium traces in tannery effluents by electrothermal atomic absorption spectrometry, flame atomic absorption spectrometry and UV/visible spectrophotometric methods, Talanta 58 (2002) 1–6.
[38] H. Pekey, D. Karakaş, M.B. Lu, Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses, Mar. Pollut. Bull. 49 (2004) 809–818.
[39] D.K. Sinha, N. Kumar, Monitoring of trace metals in Gagan water at Moradabad, Indian J. Environ. Prot. 26 (2006) 516–520.
[40] S. Li, Z. Xu, X. Cheng, Q. Zhang, Dissolved trace elements and heavy metals in the Danjiangkou Reservoir, China, Environ. Geol. 55 (2008) 977–983.
[41] M. Frankowski, M. Sojka, A. Ziola-Frankowska, M. Siepak, S. Murat-Blazejewska, Distribution of heavy metals in the Mala Welna system (western Poland), Oceanol. Hydrobiol. Stud. 38 (2009) 51–61.
[42] K.M. Mohiuddin, H.M. Zakir, K. Otomo, S. Sharmin, N. Shikazono, Geochemical distribution of trace metal pollutants in water and sediments of downstream of an urban river, Int. J. Environ. Sci. Technol. 7 (2010) 17–28.
[43] M. Varol, B. Şen, Assessment of nutrient and heavy metal contamination in surface water and sediments of the upper Tigris River, Turkey, Catena 92 (2012) 1–10.
[44] M.E. Goher, A.M. Hassan, I.A. Abdel-Moniem, A.H. Fahmy, S.M. El-sayed, Evaluation of surface water quality and heavy metal indices of Ismailia Canal, Nile River, Egypt, Egypt. J. Aquat. Res. 40 (2014) 225–233.
[45] M. Vega, R. Pardo, E. Barrado, L. Debán, Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis, Water Res. 32 (1998) 3581–3592.
[46] S. Li, Q. Zhang, Risk assessment and seasonal variations of dissolved trace elements and heavy metals in the Upper Han, China, J. Hazard. Mater. 181 (2010) 1051–1059.
[47] G. Soma, K.S. Abhay, Risk assessment, statistical source identification and seasonal fluctuation of dissolved metals in the Subarnarekha, India, J. Hazard. Mater. 265 (2014) 305–314.
[48] A. Irpino, R. Verde, A new Wasserstein based distance for the hierarchical clustering of histogram symbolic data, in: V. Batagelj, H.H. Bock, A. Ferligoj, A. Žiberna (Eds.), Data Science and Classification, Springer-Verlag, Berlin-Heidelberg, 2006, pp. 185–192.
[49] V. Simeonov, J.A. Stratis, C. Samara, G. Zachariadis, D. Voutsa, A. Anthemidis, Assessment of the surface water quality in Northern Greece, Water Res. 37 (2003) 4119–4124.
[50] H. Karadede-Akin, E. Ünlü, Heavy metal concentrations in water, sediment, fish and some benthic organisms from Tigris, Turkey, Environ. Monit. Assess. 131 (2007) 323–337.
[51] D. Zhu, Environmental issues of Three Gorges Project, in: E. Broch, D.K. Lysne, N. Flatabo, E. Helland-Hansen (Eds.), Proceedings of the 3rd International Conference on Hydropower, Balkema Publishers, Rotterdam, 1997, pp. 141–149.
[52] M. Tahri, F. Benyaïch, M. Bounakhla, E. Bilal, J.J. Gruffat, J. Moutte, D. Garcia, Multivariate analysis of heavy metal contents in soils, sediments and water in the region of Meknes (central Morocco), Environ. Monit. Assess. 102 (2005) 405–417.