Assessment of seasonal variations of chemical characteristics in surface water using multivariate statistical methods

A. Zare Garizi; *V. Sheikh; A. Sadoddin

[1]*Department of Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran*

ABSTRACT: Water pollution has become a growing threat to human society and natural ecosystems in the recent decades. Assessment of seasonal changes in water quality is important for evaluating temporal variations of river pollution. In this study, seasonal variations of chemical characteristics of surface water for the Chehelchay watershed in northeast of Iran was investigated. Various multivariate statistical techniques, including multivariate analysis of variance, discriminant analysis, principal component analysis and factor analysis were applied to analyze river water quality data set containing 12 parameters recorded during 13 years within 1995-2008. The results showed that river water quality has significant seasonal changes. Discriminant analysis identified most important parameters contributing to seasonal variations of river water quality. The analysis rendered a dramatic data reduction using only five parameters: electrical conductivity, chloride, bicarbonate, sulfate and hardness, which correctly assigned 70.2 % of the observations to their respective seasonal groups. Principal component analysis / factor analysis assisted to recognize the factors or origins responsible for seasonal water quality variations. It was determined that in each season more than 80 % of the total variance is explained by three latent factors standing for salinity, weathering-related processes and alkalinity, respectively. Generally, the analysis of water quality data revealed that the Chehelchay River water chemistry is strongly affected by rock water interaction, hydrologic processes and anthropogenic activities. This study demonstrates the usefulness of multivariate statistical approaches for analysis and interpretation of water quality data, identification of pollution sources and understanding of temporal variations in water quality for effective river water quality management.

Keywords: *Discriminant analysis; Factor analysis; Multivariate analysis of variance; Principal component analysis; Water quality*

INTRODUCTION

Rivers and their catchments are highly important parts of the natural heritage. Rivers have been utilized by mankind for thousands of years to the extent that few of them are now in their natural condition (Ngoye and Machiwa, 2004). Human and ecological use of in-stream water requires to be considered for both the quantity and the quality of water (Chang, 2008; Masamba and Mazvimavi, 2008). Water quality is important to assess the health of a watershed and to make necessary management decisions to control current and future pollution of receiving water bodies (Khadam and Kaluarachchi, 2006; Behbahaninia et al., 2009). The information on water quality and pollution sources is important for implementation of sustainable water-use management strategies (Sarkar et al., 2007; Zhou et al., 2007; Nouri et al., 2008a; Bu, et al., 2009; 2010; Soner Kara and Onut, 2010).

The chemical composition of surface water and its properties in a particular region is basically governed by natural processes such as precipitation rate, weathering processes, soil erosion and anthropogenic effects which is point or non-point sources of pollution such as urban, industrial and agricultural activities (Jarvie et al., 1998; Giridharan et al., 2009; Nouri et al., 2009; 2011). Seasonal variations in precipitation, surface run-off, ground water flow, interception and abstraction strongly affect river discharge and consequently the concentrations of chemical compositions of river

✉ *Corresponding Author Email: v.sheikh@yahoo.com*
Tel.:+98171 224 5965; Fax: +98171 224 5886

water (Vega *et al*., 1998; Khadka and Khanal, 2008; Mtethiwa *et al*., 2008; Juang *et al.,* 2009; Pejman *et al*., 2009). Pollutants entering a river system normally result from many transport pathways including storm water runoff, discharge from ditches and creeks, vadose zone leaching, groundwater seepage and atmospheric deposition (Ouyang *et al*., 2006; Nouri *et al.,* 2008b; Jha *et al.,* 2010). These pathways are also seasonal-dependent. Therefore, seasonal changes in surface water quality must be considered when establishing a water quality management program (Ouyang *et al*., 2006).

To provide a reliable estimation of surface water quality, monitoring programs attempt to include frequent water sampling at numerous sites and perform full analysis of water for a large number of physicochemical parameters. Therefore, these programs usually yield large data matrices (Boyacioglu *et al*., 2004). Interpretation of water quality data matrices for extracting useful information and effective management of surface water can be approached through the use of multivariate statistical methods, which can be called environmetrics. Environmetrics, also known as chemometrics, can be considered as a branch of environmental analytical chemistry that uses multivariate statistical modeling approaches such as regression, Discriminant analysis (DA), Cluster analysis (CA) Principal component analysis (PCA) and Factor analysis (FA) (Simeonov *et al*., 2000; Juahir *et al*., 2010). Environmetrics is deemed to be one of the best approaches to reveal the information which is concealed in the quality variables observed in water quality monitoring networks and avoid misinterpretation of large complex monitoring data (Mazlum *et al*., 1999; Simeonov *et al*., 2002).

Multivariate statistical techniques have been applied to characterize and evaluate freshwater quality; they are useful in verifying temporal and spatial variations caused by natural and anthropogenic factors linked to seasonality (Helena *et al*., 2000; Singh *et al*., 2005; Shrestha and Kazama 2007). Multivariate analysis of variance (MANOVA) determines if there are any significant differences between several groups of multivariate data. DA determines the variables that discriminate between groups by constructing a discriminant function (DF) (Singh *et al*., 2005). FA attempts to explain the correlations between the observations in terms of

the underlying factors, which are not directly observable (Yu *et al*., 2003). It helps in identifying the possible factors / sources that influence water quality (Reghunath *et al*., 2002; Panda *et al*., 2006; Bu *et al*., 2009a).

In this study, the Chehelchey River in Golestan province in Iran was chosen for water quality assessment. The Chehelchay River is one of the largest tributaries of the Gorganrood River. Currently, the watershed of the Chehelchay River is the water source for about 14000 inhabitants in rural areas, 30 villages, within the watershed as well as urban areas and agricultural fields in the downstream of the watershed. In near future, the Narmab-Chehelchay dam is going to be constructed at the downstream of the Chehelchay River by the local water authorities. This dam is planned to provide water supply, particularly potable water of Minoodasht and other adjacent cities. This is while the Chehelchay watershed has long been subjected to land-use change from forest to croplands and suffers from severe soil erosion. Thus, there is a growing concern about the potential effects of extensive land-use change and resultant changes in hydrological conditions on the availability and quality of the water from this watershed. Despite its significance, there is a lack of knowledge and understanding regarding the water quality of the Chehelchay River and its variation through the seasons. Therefore, this study attempts to facilitate the understanding of system behavior with respect to water quality issues and management within the Chahelchay watershed in the Golestan province, Iran. This study was carried out over the period August 2009 through August 2010. The objectives of this study are; 1) To examine the seasonal variations of river water quality parameters recorded during 13 years monitoring during 1995–2008; 2) To determine the important discriminant variables; and 3) To identify factors and sources influencing the chemistry of the river water. The overall aim of the present study is to provide useful information for water resources management at the watershed scale.

MATERIALS AND METHODS
Study area
The Chehelchay River, one of the largest tributaries of the Gorganrood River, is located in Golestan province in the northeastern part of Iran. The river

has a drainage area of about 25000 hectares which lies between longitude of 55°23´ and 55°38´ E and latitude of 36°59´ and 37°13´ N. The location of Chehelcay watershed is shown in Fig. 1. The watershed has an altitude range from 190 to 2570 m and a mean slope of 45.82 %. The underlying geology is mainly limestone formations, alluvial deposits near the streams and quaternary sedimentary formations,loess, in the lowlands of the study area. The climate of the region is mediterranean, which is warm and dry in summers, and temperate and rainy in winters. Average annual precipitation is about 750 mm, which mostly falls in winter and spring. Hence, the discharge into the river is relatively high during winter and spring, but low in summer and fall. Land use in the watershed is predominantly forest (55 %) and croplands (45 %).

Data

The data set of the Chehelchay water quality monitoring station was obtained from the Golestan Province Regional Water Agency that takes samples of surface water every month and analyzes the samples using standard methods. The selected parameters for the assessment of surface water quality characteristics were river discharge (Q), Electrical conductivity (EC), Total dissolved solids (TDS), pH, Total hardness (TH), sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), magnesium (Mg^{2+}), chloride (Cl^-), sulfate (SO_4^{2-}) and bicarbonate (HCO_3^-). The water quality data available for the analysis in this study belongs to the observed period from 1995 to 2008. The main step before data analysis was the assessment of the quality of the data which was accomplished by calculating the ion balance error using Eq. 1 (Hounslow, 1995):

$$Error\ of\ ion\ balance = \frac{\sum Cations - \sum Anions}{\sum Cations + \sum Anions} \qquad (1)$$

An error of up to ±5 % is tolerable, while the accuracy of the measurements with a calculated error outside this range is questionable (Hounslow, 1995). Approximately all of the water quality data for this study were within this range, so the data quality was acceptable for further analysis.

To meet the purpose of this study, raw data were segregated into four seasonal groups of Spring (from March 21st. to June 21st.), Summer (from June 22nd. to September 22nd.), Fall (from September 23rd. to December 21st.) and Winter (from December 22nd. to March 20th.). Descriptive statistics of water quality parameters according to seasons are given in Table 1.

Statistical analysis

Seasonal variations of the water quality parameters were first evaluated through season-parameter correlation matrix, using the Spearman non-parametric correlation coefficient (Spearman R). To do this, a specific integer number was assigned to each season (winter = 1, spring = 2, summer = 3 and fall = 4). Then Spearman correlation was established between all of the water quality parameters and the ordinal variables (Wunderlin *et al*., 2001; Shrestha and Kazama, 2007). Multivariate statistical techniques including MANOVA, DA, PCA, and FA were then applied for the explanation of seasonal variations and interpretation of river water quality data set.

All mathematical and statistical computations and graphical displays were made using R statistical package version 2.11.0.

MANOVA and DA

MANOVA performs an ANOVA style analysis on several dependent variables simultaneously. The objective of MANOVA is to investigate whether or not the combination of dependent variables varies with respect to the independent variables. The dependent variables are the predictors and the independent variables are the groups. The matrix of total variances and covariances; and the matrix of pooled within-group variances and covariances are compared via multivariate F tests in order to determine if there are any significant differences between groups, with regard to all variables (French *et al*., 2010). The second aaspect is prediction or allocation of observations to groups, in which Classification functions (CFs) are used to assign an observation to one of the groups (Rencher, 2002; Najafpour *et al*., 2008). The DA technique operates on original data and the method constructs a DF for each group, as in the Eq. 2 (Wunderlin *et al*., 2001; Shrestha and Kazama, 2007):

$$f(G_i) = k_i + \sum_{j=1}^{n} w_{ij} \cdot p_{ij} \qquad (2)$$

Where *i* is the number of groups (*G*), k_i is the constant inherent to each group, *n* is the number
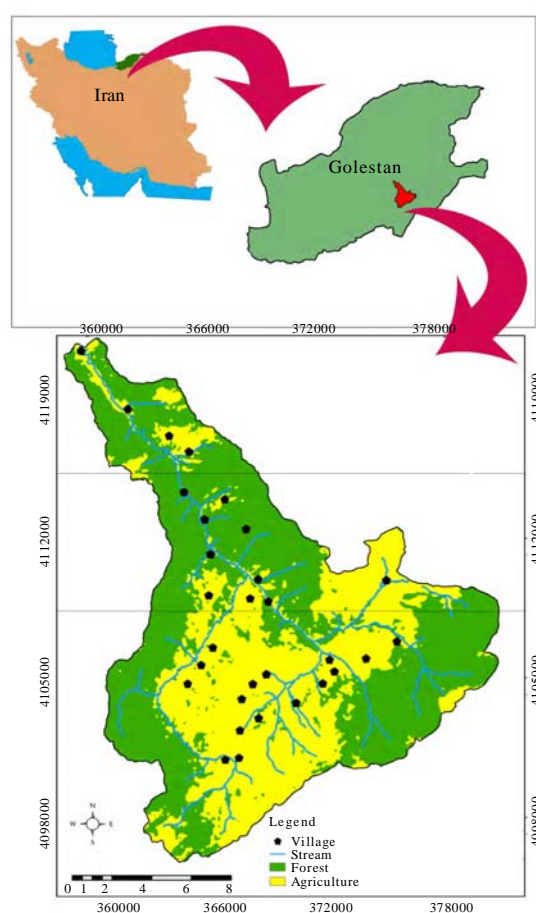
Fig. 1: Location of the Chehelchay watershed

of parameters used to classify a set of data into a given group and w_j is the weight coefficient, assigned by DA to a given selected parameters (p_j). At the beginning, DA extracts a first function that will separate the groups as much as possible. It then chooses a second function that is both uncorrelated with the first function and provides as much further separation as possible. The procedure continues adding functions in this way until reaching the maximum number of functions as determined by the number of predictors and categories in the dependent variable (Wuensch, 2007). In DA, the seasons were considered as temporal grouping variables. These grouping variables were used in the analysis as dependent variables, while the water quality parameters were considered independent variables. DA was performed to determine the most significant water quality parameters associated with the differences between the seasons. The classification

table, also called confusion matrix, was used to assess the performance of DA.

Principal component analysis and factor analysis

PCA is a statistical method used to determine components that are linear combinations of the original variables. In PCA, a set of p correlated variables is transformed to a smaller set of uncorrelated hypothetical constructs called principal components (PCs). For this purpose, a covariance matrix or correlation matrix is used. In PCA, the first PC, is the linear combination of the variables with maximal variance and represents the largest variability of the original data set. The second component is the linear combination with the next largest variability that is orthogonal to the first components and so on (Deb *et al.*, 2008). Eq. 3 introduces the basic concept of PCA:

584

Table 1: Descriptive statistics of the selected water quality parameters in the Chehelchay River

		Q (m³/s)	EC (µs/cm)	TDS (mg/L)	pH	Na⁺ (meq/L)	K⁺ (meq/L)	Ca²⁺ (meq/L)	Mg²⁺ (meq/L)	Cl⁻ (meq/L)	SO₄²⁻ (meq/L)	HCO₃⁻ (meq/L)	Hardness (mg/L)
Spring	Min.	0.30	278	182	6.30	0.10	0.02	1.30	1.00	0.30	0.10	2.25	139.2
	Max.	26.88	810	529	8.40	3.30	0.10	5.80	2.80	2.70	4.88	5.10	368.7
	Mean	3.94	497.6	324.4	7.40	1.20	0.05	2.58	1.70	1.14	0.94	3.49	212.8
	S.D.*	4.08	112.38	73.14	0.52	0.60	0.02	0.72	0.41	0.51	0.59	0.60	37.42
Summer	Min.	0.08	408	265	6.40	0.12	0.02	1.15	1.20	0.30	0.10	1.90	138.9
	Max.	3.52	1005	675	8.40	4.10	0.12	4.20	3.50	3.90	2.90	5.00	330
	Mean	0.82	639	413.4	7.51	2.28	0.05	2.52	2.03	2.22	1.32	3.38	226.1
	S.D.	0.57	128.46	83.11	0.48	0.74	0.02	0.73	0.47	0.73	0.49	0.70	38.11
Fall	Min.	0.19	368	230	6.50	0.11	0.02	1.50	0.59	0.30	0.30	2.80	149
	Max.	17.00	1390	909	8.50	6.14	0.15	4.20	4.60	6.40	4.22	5.10	411.9
	Mean	1.23	717	465.4	7.42	2.53	0.06	2.92	2.24	2.52	1.52	3.71	256.6
	S.D.	2.28	139.47	92.55	0.44	0.97	0.02	0.58	0.61	0.90	0.58	0.50	34.86
Winter	Min.	0.35	320	214	6.20	0.30	0.00	1.60	0.90	0.40	0.10	1.30	159
	Max.	20.00	1310	890	8.30	6.60	0.10	4.25	4.30	5.50	3.75	4.60	362.1
	Mean	2.30	611.2	398.9	7.41	1.77	0.05	2.95	1.92	1.72	1.10	3.85	242.1
	S.D.*	2.88	149.66	100.78	0.47	1.07	0.02	0.64	0.63	0.89	0.55	0.57	34.69

* S.D.: Standard Deviation

$$Z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + a_{i3}x_{3j} + ... + a_{im}x_{mj} \quad (3)$$

Where Z is the component score, a is the component loading, x is the measured value of a variable, i is the component number, j is the sample number and m is the total number of variables.

FA follows PCA. The main purpose of FA is to reduce the contribution of less significant variables to simplify even more of the data structure coming from PCA. This can be achieved by rotating the axis defined by PCA, according to well-established rules and constructing new variables, also called varifactors (Shrestha and Kazama, 2007).

In mathematical terms, PCA and FA involve the following five major steps: 1) Start by coding the variables x_1, x_2... x_p to have zero means and unit variances. Standardization of the measurements ensures that they all have equal weight in the analysis; 2) Calculate the covariance matrix; 3) Find the eigenvalues and corresponding eigenvectors; 4) Discard any components that only account for a small proportion of the variations in the dataset; and 5) Develop the factor loading matrix and perform a rotation on the factor loading matrix to infer the principal parameters (Ouyang *et al.*, 2006).

PCA and FA were performed on the water quality data for different seasons to discover the structure in the relationships between water quality parameters, in order to identify the important parameters which affect the chemistry of surface water in each season and to investigate the possible sources of different pollutants.

RESULTS AND DISCUSSION
Seasonal variations of river water quality

Evaluation of seasonal variations in river water quality parameters through season-parameter correlation matrix showed that except pH, all other parameters are significantly ($p < 0.01$) correlated with season. Among these, chloride, river discharge, sodium and sulfate exhibited highest correlation as Spearman's $R = 0.55, -0.54, 0.51$ and 0.44, respectively. Seasonality in various water quality parameters can be explained in terms of seasonal variations in climatologic and hydrologic characteristics associated with the wet (cold or temperate) and dry (warm) seasons.

Looking for more evidences on seasonality, the data for water quality parameters were further analyzed by MANOVA and DA. The classical Wilks' Lambda statistic was used as a significance test for the equality of the group means in MANOVA (Nath and Pavur, 1985). According to the results (Wilks' lambda = 0.3141, p-value = 0.00) four seasons are significantly different with regard to all water quality parameters. DA was carried out to determine the most significant parameters associated with the differences between seasons. Three discriminant DFs were found, explaining 100 % of the total variance between seasons (Table 2). The first

function explained 71.4 % of the total variance between seasons; the second function explained 22.1%, while the third one explained only 6.5 % of the total variance between the seasons. The relative contribution of each parameter is reported in Table 3.

As it can be seen in Table 3, EC, Cl⁻ and Na⁺ have large absolute values in DF-1. These parameters have strong contribution in discriminating four seasons and account for most of the expected seasonal variations in water quality. The second group of parameters which have less contribution in explaining the seasonal variations are TDS, HCO_3^- and TH in DF-1 and DF-2 and SO_4^{2-} and Ca^{2+} in DF-3.

The classification matrix (Table 4) shows that 79.2 % of the cases are correctly assigned to their respective groups, indicating that the performance of DA has been quite good. After standard DA, backward stepwise variable selection technique using DA helped in finding the most important parameters. In this method, a model of discrimination is built including all

variables and then variables are removed step-by-step, beginning with the less significant until no significant changes in removal criteria are obtained (Wunderlin *et al.*, 2001; Shrestha and Kazama, 2007). The results suggest that EC, Cl⁻, HCO_3^-, SO_4^{2-} and TH are the most significant parameters. Stepwise, DA gave 70.2 % correct assignations using only the five discriminant parameters. Box and Whisker plots of the selected parameters showing seasonal trends are given in Fig. 2. The combined use of DA and Box and Whisker plots simplifies the interpretation of the results.

As it can be seen in Fig. 2, higher values of EC, Cl⁻ and SO_4^{2-} are observed in summer and fall and lower values in spring and winter. These parameters are reactive components that are partially of anthropogenic origin (Khazheeva *et al.*, 2007). EC qualitatively reflects the status of inorganic pollution and is a measure of TDS in waters (McCutcheon *et al.*, 1993). In the Chehelchay River, conductance values follow an inverse pattern with the river discharge, which shows the dilution effect. Average river discharge is higher in spring and winter as compared to summer and fall (Table 1). High EC values are observed as a result of decreased rainfall /discharge and increased agricultural land-use and built-up intensity in the river catchment. Of the dissolved solids, Cl⁻ is the easiest to account for because of its conservative nature and limited numbers of potential sources. The behavior of Cl⁻ is described as conservative because its circulation through the hydrological cycle is determined by physical rather than chemical processes (Grasby *et al.*, 1997). High Cl⁻ concentrations in summer and fall indicate the hydrologic effect and anthropogenic pollution with NaCl, as observed in comparison with the Na⁺ concentrations (Table 1). Dissolved SO_4^{2-} can be derived from the dissolution of SO_4 minerals; oxidation of pyrite and other forms of reduced S; oxidation of organic sulfides in natural soil processes; and anthropogenic inputs, i.e. fertilizers (Hem, 1985; Grasby *et al.*, 1997). S-based fertilizers are not commonly used in the Chehelchay watershed and plausible origin of SO_4^{2-} in the river water is weathering of the sedimentary rocks. Seasonal variations of dissolved SO_4^{2-} could be attributed to changes in the intensity of weathering process.

TH and HCO_3^- concentration of the Chehelchay River water is high in Fall and Winter and low in Spring and Summer (Fig. 2c and e). These are

Table 2: Eigen-values for the DFs

Function	Eigen value	Percent of variance	Cumulative variance (%)
1	1.106	71.4	71.4
2	0.300	22.1	93.5
3	0.077	6.5	100.0

Table 3: DF coefficients for water quality parameters

Parameter	DF-1	DF-2	DF-3
Discharge	-0.09	-0.02	0.75
TDS	-0.73	1.12	0.46
EC	1.85	-0.64	0.43
pH	0.16	0.02	0.15
HCO_3^-	-0.69	0.62	0.13
Cl^-	1.06	-0.23	1.37
SO_4^{2-}	-0.28	0.35	1.28
Ca^{2+}	0.16	0.17	-0.96
Mg^{2+}	-0.01	-0.11	-0.74
Na^+	-1.02	-0.55	-2.46
K^+	0.28	0.50	0.19
Hardness	-0.26	1.14	0.23

Table 4: Classification matrix for DA

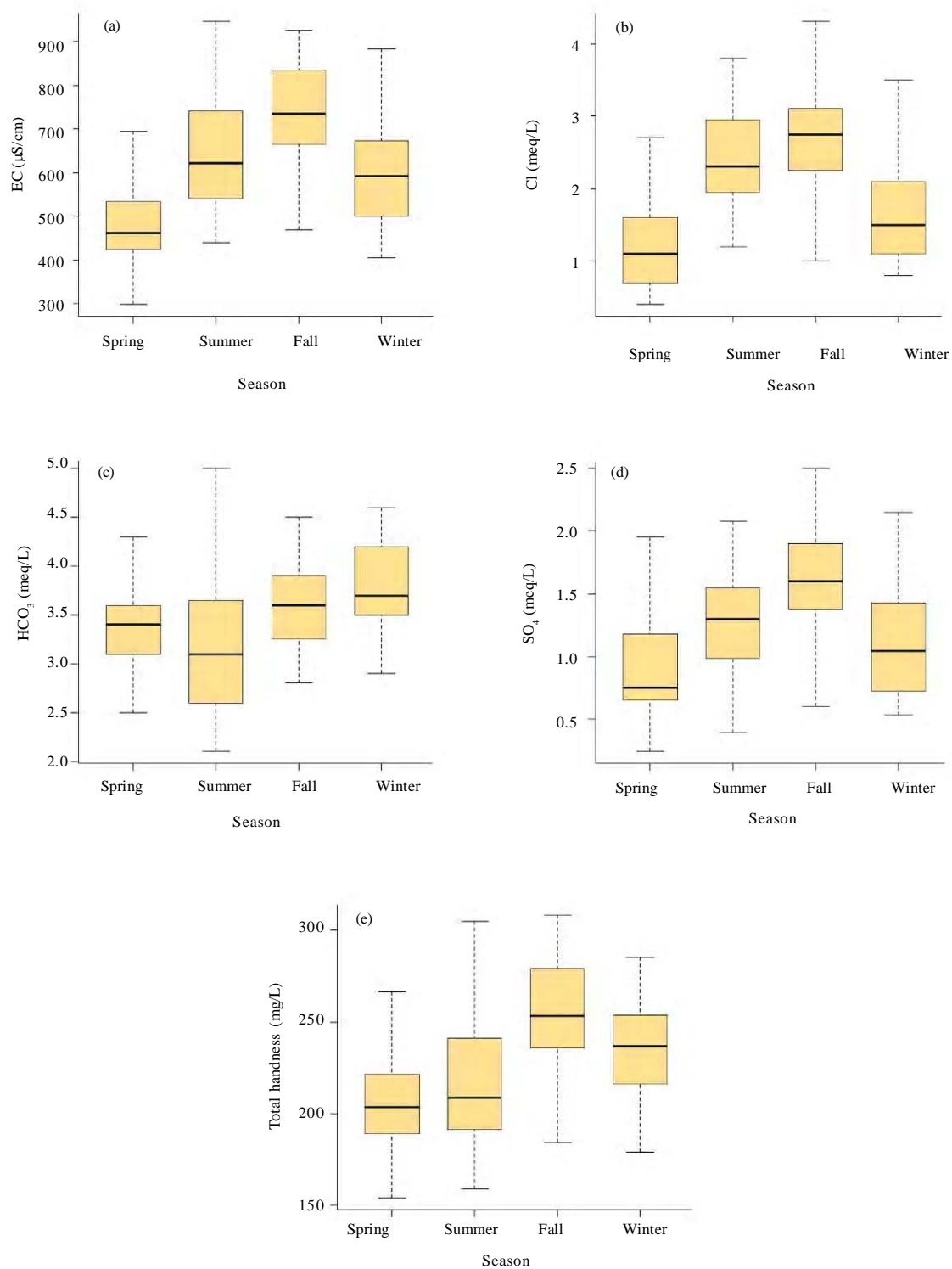| Season	Seasons assigned by DA				Correct (%)
	Spring	Summer	Fall	Winter	
Spring	30	1	1	4	83.3
Summer	1	28	4	3	77.8
Fall	0	2	29	5	80.6
Winter	6	1	2	27	75
Total	37	32	36	39	79.2

Fig. 2: Seasonal variations of: a) EC, b) Cl^-, c) HCO_3^-, d) SO_4^{2-} and e) TH in surface water quality of the Chehelchay watershed

nonreactive components that are determined by natural hydrogeochemical processes (Khazheeva *et al.*, 2007). TH is dependent on the presence of Ca^{2+} and Mg^{2+} in water. Similar pattern in seasonal variation of TH and HCO_3^- is possibly due to dissolution of minerals such as calcite and dolomite through HCO_3^- weathering process. This process is represented by the following simplified reaction (Cai *et al.*, 2007):

$$CaCO_3 + CO_2 + H_2O <-> Ca^{2+} + 2HCO_3^- \qquad (4)$$
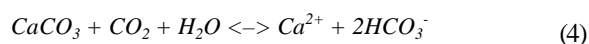
The reaction of HCO_3^- minerals with dissolved carbon dioxide (carbonic acid) releases HCO_3^- and Ca^{2+} in the river water. There is an inverse relationship between temperature and dissolved CO_2 (Langmuir, 1997), thus the intensity of carbonate weathering and concentration of HCO_3^- and TH is higher in cold seasons, i.e. Fall and Winter and lower in warm seasons.

PCA and FA of water quality data

PCA was performed on the normalized data sets containing 12 variables, separately for four different seasons, to compare the compositional pattern between water quality parameters and to identify the factors influencing them. PCA of the four data sets yielded three PCs with eigenvalues >1, explaining more than 80 % of the total variance in respective water quality data sets. An eigenvalue gives a measure of the significance of the PC; thus the PCs with the highest eigenvalues are the most significant. Eigenvalues of 1.0 or greater are considered significant (Shrestha and Kazama, 2007). For each season three varifactors (VFs) were obtained through the FA performed on the PCs. Results of FA including factor loadings, eigenvalues and total and cumulative variance values are presented in Table 5. Fig. 3 represents the results obtained for the first and
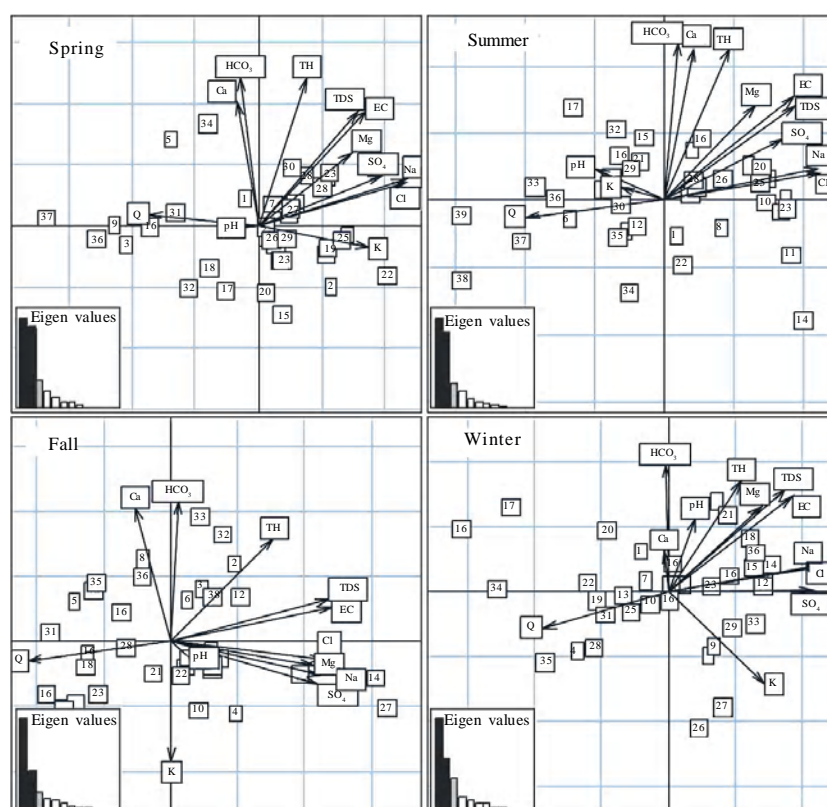
Fig. 3: Scatter diagrams of FA for four seasons; vectors represent loadings of each water quality parameter in VF1 (horizontal axis) and and VF2 (vertical axis); numbers are observations

Table 5: Factor loading matrix and explained variance of water quality parameters in four seasons

Spring				Summer			
Parameter	VF1	VF2	VF3	Parameter	VF1	VF2	VF3
Discharge	-0.668	0.076	0.181	Discharge	-0.800	-0.098	0.219
TDS	0.613	0.723	0.038	TDS	0.765	0.567	-0.078
EC	0.657	0.725	0.040	EC	0.759	0.622	-0.029
pH	-0.083	-0.012	0.852	pH	-0.398	0.192	0.584
HCO_3^-	-0.111	0.938	0.190	HCO_3^-	0.086	0.943	-0.055
Cl^-	0.903	0.271	0.068	Cl^-	0.904	0.168	0.079
SO_4^{2-}	0.754	0.319	-0.308	SO_4^{2-}	0.689	0.365	0.258
Ca^{2+}	-0.123	0.780	-0.515	Ca^{2+}	0.173	0.908	-0.095
Mg^{2+}	0.571	0.465	0.461	Mg^{2+}	0.534	0.569	0.245
Na^+	0.898	0.292	0.079	Na^+	0.894	0.194	0.011
K^+	0.676	-0.138	0.118	K^+	-0.251	0.078	0.876
Hardness	0.298	0.933	-0.081	Hardness	0.376	0.910	0.050
Eigenvalue	4.348	3.908	1.401	Eigenvalue	4.577	3.827	1.310
Total variance %	36.236	32.569	11.677	Total variance %	38.144	31.889	10.914
Cumulative %	36.236	68.806	80.482	Cumulative %	38.144	70.033	80.947
Fall				Winter			
Parameter	VF1	VF2	VF3	Parameter	VF1	VF2	VF3
Discharge	-0.838	-0.116	-0.101	Discharge	-0.774	-0.232	0.222
TDS	0.931	0.256	0.066	TDS	0.721	0.647	0.097
EC	0.952	0.211	0.073	EC	0.767	0.609	0.097
pH	0.119	-0.100	0.885	pH	0.169	0.459	0.002
HCO_3^-	0.045	0.853	-0.118	HCO_3^-	-0.019	0.811	0.441
Cl^-	0.906	-0.115	0.224	Cl^-	0.933	0.174	-0.030
SO_4^{2-}	0.869	-0.256	-0.143	SO_4^{2-}	0.892	0.014	0.112
Ca^{2+}	-0.206	0.816	-0.219	Ca^{2+}	-0.025	0.261	0.953
Mg^{2+}	0.848	-0.142	-0.128	Mg^{2+}	0.573	0.545	-0.516
Na^+	0.891	-0.209	0.295	Na^+	0.938	0.159	-0.052
K^+	0.006	-0.728	-0.177	K^+	0.594	-0.587	0.052
Hardness	0.610	0.624	-0.319	Hardness	0.450	0.705	0.482
Eigenvalue	5.996	2.590	1.172	Eigenvalue	5.166	2.975	1.687
Total variance %	49.969	21.582	9.765	Total variance %	43.049	24.795	14.061
Cumulative %	49.969	71.551	81.316	Cumulative %	43.049	67.844	81.905

second varifactors in different seasons. The factor loadings express the correlation between the original variables and the newly formed varifactors. The VF loadings can be used to determine the relative importance of a variable as compared to other variables in a factor and do not reflect the importance of the factor itself (Ouyang *et al.*, 2006). Classification of factor loading is 'strong', 'moderate' and 'weak', corresponding to absolute loading values of > 0.75, 0.75-0.50 and 0.50-0.30, respectively (Liu *et al.,* 2003).

Close inspection of eigenvalues, loadings and significant water quality parameters in each season reveals that beside geochemical aspects, seasonal regime of the Chehelchay River water quality is controlled by four important hydrologic/anthropogenic processes:

1) Dilution, during springtime, when river discharge reaches its maximum value; 2) Concentration, due to evapotranspiration and the base-flow effect,

during the dry seasons, i.e. Summer and Fall; 3) Impact of non-point sources of pollution especially agricultural activities, that are dominant during the warm seasons and; 4) Flushing of soil-salts, due to floods occurring in Fall / Winter. Having the mentioned processes functioning over the watershed, relevant management policies and actions need to be implemented. To reduce the concentrations of chemical compositions of river water during the Summer, water extraction from the river for irrigation of rice paddies on floodplains of the main river should be regulated.

Although the chemical fertilizers and pesticides effects have not been dealt with in this study, a large amount of chemical fertilizers and pesticides is applied on these paddies, which is finally released into the river ecosystem and deteriorates the river water quality. Moreover, plowing the steep slopes for dry farming throughout the watershed and application of chemical fertilizers and pesticides on these farms can

cause non-point source pollution and flushing of soil-salts during the Fall and Winter, when the soil surface is rarely covered with vegetation or plant residues. It seems appropriate to apply the best management practices over the entire watershed to curb some of the environmental issues raised above. Of the best ways to mitigate the downgrading trend of river water quality, establishment of riparian vegetation is recommended.

Obviously, a carefully planned sampling procedure is required to fill the current gap in available data in terms of sampling points distribution with regards to the geology and land use as well as including pathogens and fertilizers usage for the study area.

CONCLUSION

In this study, various multivariate statistical methods were utilized to evaluate seasonal variations in surface water quality of the Chehelchay watershed. The significance of the seasonal variations of water quality was evaluated by MANOVA. The results showed that the seasonal variations of the four seasons are significantly different. DA provided significant DFs and recognized the most significant parameters. It yielded a dramatic data reduction, as it used only five parameters (EC, Cl$^-$, HCO$_3^-$, SO$_4^{2-}$ and TH), leading to 70.2 % correct assignations of observations to respective season. The study shows that DA is the best data reduction and pattern recognition method for the study area in comparison to other techniques, that allows reduction in the dimensionality of the large data set, delineating a few indicator parameters responsible for large variations in water quality. PCA and FA assisted in extracting and recognizing factors or sources of origin responsible for water chemistry in four seasons of the year. PCA / FA identified three latent factors that explained more than 80 % of the total variance of 12 parameters, standing for salinity, carbonate weathering and alkalinity, respectively. Contribution of hydrologic processes and anthropogenic activities was also demonstrated by FA. The analytical results indicate that the surface water of the Chehelchay watershed is a unique example for the compound impact of weathering, hydrologic and anthropogenic processes. The chemical composition of surface water in each season is strongly influenced by rock-water interactions; dissolution and dilution; and anthropogenic inputs.

This study demonstrates the usefulness of multivariate statistical techniques for analysis and interpretation of complex data sets, and also for identification of pollution sources and better understanding of seasonal variations in water quality for effective river water quality management.

ACKNOWLEDGEMENTS

REFERENCES

Behbahaninia, A.; Mirbagheri, S. A.; Khorasani, N.; Nouri, J.; Javid, A. H., (2009). Heavy metal contamination of municipal effluent in soil and plants. J. Food, Agri. Environ., 7 (3-4), 851-856 (**6 pages**).

Boyacioglu, H.; Gunduz, O., (2004). Application of factor analysis in the assessment of surface water quality in Buyuk Menderes River Basin. In EWRA symposium on water resources management risks and challenges for the 21st Century 2. Izmir, Turkey, September 2-4.

Bu, H.; Tan, X.; Li, S.; Zhang, Q., (2009). Water quality assessment of the Jinshui River (China) using multivariate statistical techniques. Environ. Earth Sci., 60 (8),1631–1639 (**9 pages**).

Bu, H.; Tan, X.; Li, S.; Zhang, Q., (2010). Temporal and spatial variations of water quality in the Jinshui River of the South Qinling Mts., China. Ecotox. Environ. Safe., 73 (5), 907–913 (**7 pages**).

Cai, W. J.; Guo, X. H.; Chen, C. T. A.; Dai, M. H.; Zhang, L .J.; Zhai, W. D.; Lohrenz, S. E.; Yin, K., (2007). A comparative overview of weathering intensity and HCO$_3^-$ flux in the world's major rivers with emphasis on the Changjiang, Huanghe, Zhujiang (Pearl) and Mississippi Rivers. Cont. Shel. Res., 28 (12), 1538-1549 (**12 pages**).

Chang, H., (2008). Spatial analysis of water quality trends in the Han River basin, South Korea. Water Res., 42 (13), 3285-3304 (**20 pages**).

Deb, D.; Deshpande, V. N.; Das, K. C., (2008). Assessment of water quality around surface coal mines using principal component analysis and fuzzy reasoning techniques. Mine. Water Environ., 27 (3), 183–193 (**11 pages**).

French, A.; Macedo, M.; Poulsen, J.; Waterson, T.; Yu, A., (2010). Multivariate analysis of variance (MANOVA). Avaialabel at: http://userwww.sfsu.edu/~efc/classes/biol710/manova/MANOVAnewest.pdf.

Giridharan, L.; Venugopal, T.; Jayaprakash, M., (2009). Assessment of water quality using chemometric tools: A case study of River Cooum, South India. Arch. Environ. Con. Tox., 56 (4), 654–669 (**16 pages**).

Grasby, S. E.; Hutcheon, I.; Krouse, H. R., (1997). Application of the stable isotope composition of SO$_4^{2-}$ to tracing anomalous TDS in Nose Creek, southern Alberta, Canada. Appl. Geochem., 12 (5), 567-575 (**8 pages**).

Helena, B.; Pardo, R.; Vega, M.; Barrado, E.; Fernandez, J. M.; Fernandez, L., (2000). Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal component analysis. Water Res., 34 (3), 807-816 (**10 pages**).

Hem, J. D., (1985). Study and interpretation of the chemical characteristics of natural water. United States Geological Survey Water Supply Paper # WSP 2254.

Hounslow, A. W., (1995). Water quality data: Analysis and interpretation. CRC Press LLC, Lewis publishers.

Jarvie, H. P.; Whitton, B. A.; Neal, C., (1998). Nitrogen and phosphorus in east coast British rivers: Speciation, sources, and biological significance. Sci. Total Environ., 210-211, 79–109 (**31 pages**).

Jha, R. B.; Waidbacher, H.; Sharma, S.; Straif, M., (2010). Study of agricultural impacts through fish base variables in different rivers. Int. J. Environ. Sci. Tech., 7 (3), 609-615 (**7 pages**).

Juahir, H.; Zain, S. M.; Yusoff, M. K.; Hanidza, T. I. T.; Armi, A. S. M.; Toriman, M. E.; Mokhtar , M., (2010). Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. Environ. Monit. Assess., 173 (1-4), 625–641 (**17 pages**).

Juang, D. F.; Lee, C. H., Hsueh, S. C., (2009). Chlorinated volatile organic compounds found near the water surface of heavily polluted rivers. Int. J. Environ. Sci. Tech., 6 (4), 545-556 (**12 pages**).

Khadam, I. M.; Kaluarachchi, J. J., (2006). Water quality modeling under hydrologic variability and parameter uncertainty using erosion-scaled export coefficients. J. Hydrol., 330 (1-2), 354-367 (**14 pages**).

Khadka, R. B.; Khanal, A. B., (2008). Environmental management plan (EMP) for Melamchi water supply project, Nepal. Environ. Monit. Assess., 146 (1-3), 225-234 (**10 pages**).

Khazheeva, Z. I.; Tulokhonov, A. K.; Dashibalova, L. T., (2007). Seasonal and spatial dynamics of TDS and major ions in the Selenga River. Water Res., 34 (4), 444–449 (**6 pages**).

Langmuir, D., (1997). Aqueous environmental geochemistry, Prentice-Hall Inc.

Liu, C. W.; Lin, K. H.; Kuo, Y. M., (2003). Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Sci. Total Environ., 313 (1-3), 77-89 (**13 pages**).

Masamba, W. R. L.; Mazvimavi, D., (2008). Impact on water quality of land uses along Thamalakane-Boteti River: An outlet of the Okavango Delta. Phys. Chem. Earth, 33 (8-13), 687-694 (**8 pages**).

Mazlum, N.; Ozer, A.; Mazlum, S., (1999). Interpretation of water quality data by principal components analysis. Turk. J. Eng. Environ. Sci., 23, 19-26 (**8 pages**).

McCutcheon, S. C.; Martin, J.; Barnwel, T. O., (1993). Water Quality. Maidment, D. R. (Eds.), Handbook of Hydrology. McGraw-Hill Inc., New York.

Mtethiwa, A. H.; Munyenyembe, A.; Jere, W.; Nyali, E., (2008). Efficiency of oxidation ponds in wastewater treatment. Int. J. Environ. Res., 2 (2), 149-152 (**4 pages**).

Najafpour, S.; Alkarkhi, A. F. M.; Kadir, M. O. A.; Najafpour, G. D., (2008). Evaluation of spatial and temporal variation in river water quality. Int. J. Environ. Res., 2 (4), 349-358 (**10 pages**).

Nath, R.; Pavur, R., (1985). A new statistic in the one way multivariate analysis of variance. Comput. Stat. Data An., 2 (4), 297–315 (**19 pages**).

Ngoye, E.; Machiwa, J. F., (2004). The influence of land-use patterns in the Ruvu river watershed on water quality in the river system. Phys. Chem. Earth, 29 (15-18), 1161-1166 (**6 pages**).

Nouri, J.; Danehkar, A.; Sharifipour, R. (2008b). Evaluation of ecotourism potential in the northern coastline of the Persian Gulf. Environ. Geo. 55 (3), 681-686 (**6 pages**).

Nouri, J.; Karbassi, A. R.; Mirkia, S., (2008a). Environmental management of coastal regions in the Caspian Sea. Int. J. Environ. Sci. Tech., 5 (1), 43-52 (**10 pages**).

Nouri, J.; Khorasani, N.; Lorestani, B.; Karami, M.; Hassani, A.H.; Yousefi, N. (2009). Accumulation of heavy metals in soil and uptake by plant species with phytoremediation potential. Environ. Earth Sci., 59 (2), 315-323 (**9 pages**).

Nouri, J.; Lorestani, B.; Yousefi, N.; Khorasani, N.; Hasani, A. H.; Seif, S.; Cheraghi, M., (2011). Phytoremediation potential of native plants grown in the vicinity of Ahangaran lead–zinc mine (Hamedan, Iran). Environ. Earth Sci., 62 (3), 639-644 (**7 pages**).

Ouyang, Y.; Nkedi-Kizza, P.; Wu, Q. T.; Shinde, D.; Huang, C. H., (2006). Assessment of seasonal variations in surface water quality. Water Res., 40 (20), 3800-3810 (**11 pages**).

Panda, U. C.; Sundaray, S. K.; Rath, P.; Nayak, B. B.; Bhatta, D., (2006). Application of factor and cluster analysis for characterization of river and estuarine water systems-a case study: Mahanadi River (India). J. Hydrol., 331 (3-4), 434–445 (**12 pages**).

Pejman, A. H.; Nabi Bidhendi, G. R.; Karbassi, A. R.; Mehrdadi, N.; Esmaeili Bidhendi, M., (2009). Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques. Int. J. Environ. Sci. Tech., 6 (3), 467-476 (**10 pages**).

Reghunath, R.; Murthy, T. R. S.; Raghavan, B. R., (2002). The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. Water Res., 36 (10), 2437–2442 (**6 pages**).

Rencher, A. C., (2002). Methods of Multivariate Analysis, Second Edition. John Wiley and Sons Inc.

Sarkar, S. K.; Saha, M.; Takada, H.; Bhattacharya, A.; Mishra, P.; Bhattacharya, B., (2007). Water quality management in the lower stretch of the river Ganges, east coast of India: an approach through environmental education. J. Clean Prod., 15 (16), 1559–1567 (**9 pages**).

Shrestha, S.; Kazama, F., (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environ. Modell. Softw., 22 (4), 464-475 (**12 pages**).

Simeonov, V.; Einax, J. W.; Stanimirova, I.; Kraft, J., (2002). Envirometric modeling and interpretation of river water monitoring data. Anal. Bioanal. Chem., 374 (5), 898-905 (**8 pages**).

Simeonov, V.; Stefanov, S.; Tsakovski, S., (2000). Environmetrical treatment of water quality survey data from Yantra River, Bulgaria. Mikrochim. Acta, 134 (1-2), 15–21 (**7 pages**).

Singh, K. P.; Malik, A.; Sinha, S., (2005). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: A case study. Anal. Chim. Acta, 538 (1-2), 355-374 (**20 pages**).

Soner Kara, S.; Onut, S., (2010). A stochastic optimization approach for paper recycling reverse logistics network design under uncertainty. Int. J. Environ. Sci. Tech., 7 (4), 717-730 (**14 pages**).

Vega, M.; Pardo, R.; Barrado, E.; Deban, L., (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res., 32 (12), 3581-3592 (**12 pages**).

Wuensch, K. L., (2007). Multivariate analysis with SPSS. East Carolina University. Available at: http://core.ecu.edu/psyc/wuenschk/SPSS/SPSS-MV.htm.

Wunderlin, D. A.; Diaz, M. P.; Ame, M. V.; Pesce, S. F.; Hued, A. C.; Bistoni, M. A., (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). Water Res., 35 (12), 2881-2894 (**14 pages**).

Yu, S.; Shang, J.; Zhao, J.; Guo, H., (2003). Factor analysis and dynamics of water quality of the Songhua River Northeast China. Water Air Soil Pollut., 144 (1-4), 159–169 (**11 pages**).

Zhou, F.; Huang, G. H.; Guo, H. C.; Zhang, W.; Hao, Z. J., (2007). Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong, Water Res., 41 (15), 3429–3439 (**11 pages**).

AUTHOR (S) BIOSKETCHES

Zare Garizi, A., M.Sc. student, Department of Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran. Email: *arash.zare86@gmail.com*

Sheikh, V., Ph.D., Assistant Professor, Department of Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran. Email: *v.sheikh@yahoo.com*

Sadoddin, A., Ph.D., Assistant Professor, Department of Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran. Email: *amir.sadoddin@gmail.com*