Multivariate statistical analysis for estimating surface water quality in reservoirs

Matias Bonansea*

Consejo Nacional de Investigaciones Científicas
y Técnicas (CONICET),
Departamento de Estudios Básicos y Agropecuarios,
Facultad de Agronomía y Veterinaria (FAV),
Universidad Nacional de Río Cuarto (UNRC),
Ruta Nacional 36 Km 601, 5800 Río Cuarto, Córdoba, Argentina
Email: mbonansea@ayv.unrc.edu.ar
*Corresponding author

Raquel Bazán

Departamento de Ingeniería Química y Aplicada,
Facultad de Ciencias Exactas Físicas y Naturales,
Universidad Nacional de Córdoba (UNC),
Instituto Superior de Estudios Ambientales (ISEA-UNC), Argentina
Email: rbazan84@gmail.com

Susana Ferrero

Departamento Matemática,
Facultad de Ciencias Exactas,
Físico-Químicas y Naturales (FCEFQyN),
Universidad Nacional de Río Cuarto (UNRC), Argentina
Email: sferrero@exa.unrc.edu.ar

Claudia Rodríguez and Claudia Ledesma

Departamento de Estudios Básicos y Agropecuarios,
Facultad de Agronomía y Veterinaria (FAV),
Universidad Nacional de Río Cuarto (UNRC), Argentina
Email: m_c_rodriguez@yahoo.com
Email: cledesma@ayv.unrc.edu.ar

Lucio Pinotti

Consejo Nacional de Investigaciones Científicas
y Técnicas (CONICET),
Departamento de Geología,
Físico-Químicas y Naturales (FCEFQyN),
Universidad Nacional de Río Cuarto (UNRC), Argentina
Email: lpinotti@exa.unrc.edu.ar

Abstract: Regular water quality monitoring programs are an important aspect of water management. Different multivariate statistical techniques were applied for interpretation and evaluation of the data matrix obtained during a six-year monitoring program (2006 to 2011) in the principal reservoirs of the central region of Argentina. Eleven sampling sites located in two reservoirs were surveyed each climatic season for 18 parameters. Cluster analysis grouped the sampling sites into three clusters and classified the different climatic seasons into two clusters based on their similarities. Principal component analysis/factor analysis showed the existence of five significant varifactors (VF) which account for 79.3% of the variance, related to soluble salts, nutrients, physico-chemical parameters, and non-common source. Source contribution was calculated using multiple regression of sample mass concentration on the absolute VF scores. This study demonstrates the usefulness of multivariate statistical techniques helping managers to get better information about surface water systems.

Biographical notes: Matias Bonansea is an Assistant Investigator of the National Scientific and Technical Research Council (CONICET-Argentina) and is a Professor at the Department of Basic Studies and Agriculture at the National University of Río Cuarto. He obtained his PhD in Biological Sciences in University of Río Cuarto (Argentina).

Raquel Bazán is a Professor in the Department of Chemical Engineering and Applied at the National University of Córdoba. She obtained her Master's degree in Environmental Engineering in the National University of Córdoba.

Susana Ferrero is a Professor in the Mathematics Department at the National University of Río Cuarto. She obtained her Master's degree in Statistics in the National University of Río Cuarto.

Claudia Rodríguez is a Professor in the Department of Basic Studies and Agriculture at the National University of Río Cuarto. She obtained her PhD in Biological Sciences in University of Córdoba.

Claudia Ledesma is a Professor in the Department of Basic Studies and Agriculture at the National University of Río Cuarto. She obtained her Master's degree in Environmental Engineering in the National University of Río Cuarto.

Lucio Pinotti is an Investigator of the National Scientific and Technical Research Council and is a Professor in the Department of Geology at the National University of Río Cuarto. He obtained his PhD in Geological Sciences in University of Río Cuarto.

1 Introduction

Surface water quality in a region depends not only on natural factors such as precipitation inputs, erosion, weathering of crustal materials, but also on human activities which can alter these systems, either by polluting them or by modifying them to such an extent that

the use of surface water system become restricted (Banihabib et al., 2015; Su et al., 2015). Thus, public officials, land owners, and the general public express concern over perceived deterioration of water quality and seek to determine the magnitude of the impact that public and private investments and regulatory actions are having on the attainment of water quality goals in order to decide about investing in further actions (Hirsch et al., 2015).

Regular water quality monitoring programs are an important aspect of water resource management (Megalovasilis, 2014; Yan et al., 2015). Due to spatial and temporal variations in water quality, such programs need to include a large number of physicochemical parameters taken at different times and from many sites involving huge financial inputs and resulting in a large and complex data matrix which is often difficult to interpret towards drawing meaningful conclusions. Thus, there is a need to optimise the monitoring networks, reducing the number of water quality parameters, sampling sites and monitoring periods to representative ones without losing useful information (Singh et al., 2004). To achieve this goal, some data-driven approaches, such as the projection pursuit technique and neural networks, have been applied to water quality assessment (Wang et al., 2013). However, compared with these approaches, multivariate statistical techniques can be used to analyse large water quality data sets providing a representative and reliable estimation of surface water quality. These techniques can be used to achieve great efficiency of data compression from the original data and to interpret natural associations between samples or variables, highlighting the information which is not available at first glance (Varol et al., 2012). Furthermore, the multivariate treatment of environmental could be successfully used to identify possible pollution sources that influence water systems and offers a valuable technique for reliable management of water resources as well as rapid solutions on pollution problems (Wang et al., 2012).

The objective of this study was to evaluate the efficiency of different pattern recognition methods for the analysis of water quality of the two major dam reservoirs in Calamuchita valley, located in the central region of Argentina. This procedure could be used by water management authorities and decision-makers to generate a standardised monitoring program for regional-scale water quality assessment, allowing an easier comparison of water quality from different reservoirs at a regional scale. Different multivariate statistical techniques [cluster analysis (CA) and principal component analysis/factor analysis (PCA/FA)] were applied to the data matrix obtained during a six years (2006–2011) monitoring program in the principal reservoirs of Calamuchita valley to evaluate data and draw conclusions about the similarities and dissimilarities existing between sampling sites and seasons, as well as to identify water quality variables responsible for spatial and temporal variations in water quality. The influence of pollution sources on reservoirs water quality was ascertained. Thus, we illustrated the usefulness of the multivariate statistical analysis to improve the understanding of the surface water system.

2 Methodology

2.1 Study area

Calamuchita valley, located in the central region of Argentina, is characterised by a mountainous system called Sierras Pampeanas by the west and Sierras Chicas by the east, which encompass approximately a 500 km long and 150 km wide area. The climate of the region is temperate subtropical, air temperature range between –8.0°C and 40.0°C with a mean value of 16.5°C. The rainy period is spring-summer (more than 200 mm per month, September-March) and the annual mean rain is 743 mm (Mariazzi et al., 1992). This region presents different moderately eutrophic reservoirs greater than 0.3 km^2, which were built between 1930 to 1950 for multiple purposes, such as water supply, power generation, flood control, irrigation, tourism and recreational activities (Bonansea et al., 2015a). As part of a monitoring program several physical, chemical and biological properties were surveyed in the most important reservoirs of this area (Figure 1).

Figure 1 Principal reservoirs of Calamuchita valley and position of sampling sites in Los Molinos and Río Tercero reservoir (see online version for colours)
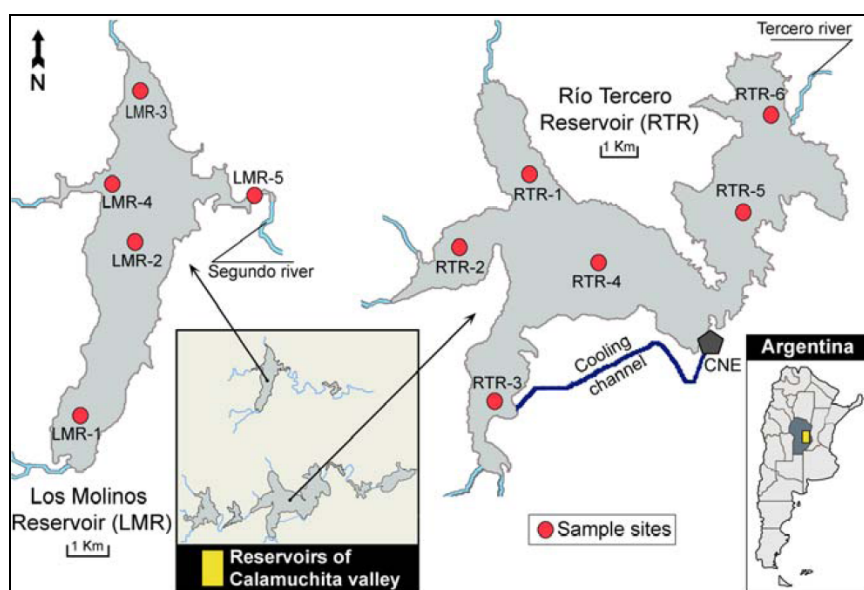
Río Tercero reservoir (RTR: 32.210844 S; 64.473056 W) is the largest artificial reservoir in Córdoba province. This multiple purposes reservoir, which is used to supply drinking water to three cities located on its shores, has an area of 46 km^2, a maximum volume of 733 hm^3 and a maximum and mean depth of 46.5 and 12.2 m, respectively (Bonansea

et al., 2015b). RTR is divided in two basins by a strait. Western basin has three branches where rivers flow, while eastern basin presents the only effluent called Tercero River (Mariazzi et al., 1992). In 1986 a nuclear power plant (CNE: 600 MWa) was installed. Water for cooling the nuclear reactor is taken from the middle section of the reservoir and is returned to the western basin by a 5 km long open-sky channel (Bonansea et al., 2015b). Los Molinos reservoir (LMR: 31.818656 S, 64.531586 W), is used to supply drinking water to the 30% of Cordoba city (with 1.4 million inhabitants). This reservoir has an area of 21.1 km^2, a maximum volume of 399 hm^3 and a maximum and mean depth of 53.0 and 16.3 m, respectively (Bazán et al., 2005). LMR has three branches where rivers flow and its effluent is called Segundo River.

2.2   Sampling and analytical procedures

Water sampling was conducted once during each climatic season (defined in the southern hemisphere as winter: July–September, spring: October–December, summer: January–March, and fall: April–June) from 2006 to 2011. Eleven sampling sites were selected in order to reasonably represent the water quality of the reservoirs (Figure 1). Five sample sites were located in LMR and six sites were located in RTR. Coordinates of sample sites were recorded using a GPS device. Water samples were collected at 20 cm depth. Preservation and transportation of water samples were performed according to standard methods (APHA-AWWA-WEF, 2000). In total, samples were analysed for 18 parameters. The selected water quality parameters included water surface temperature (WST), pH, dissolved oxygen (DO) and electrical conductivity (EC) which were measured in-situ using portable electronic instruments. Secchi disk transparency (SDT) was measured using a standard Secchi disk with alternating black and white quadrants. In laboratory, chlorophyll-a concentration (Chl-a), total phosphorus (TP), total nitrogen (TN), nitrate nitrogen (NO$_3$-N), sodium (Na), potassium (K), calcium (Ca), magnesium (Mg), chloride (Cl), sulphate (SO$_4$), fluoride (F), total hardness (T-Hard) and total alkalinity (T-Alk) were determined according to standard analytical methods and protocols (APHA-AWWA-WEF, 2000; WHO, 2006).

2.3   Statistical analysis

One way multivariate analysis of variance (MANOVA) and one way analysis of variance (ANOVA) were performed as a first approach to compare the significant spatial and temporal differences of water quality variables (p < 0.05; least significance difference, LSD). Relationships among variables were tested using Pearson's coefficient with statistical significance set at p < 0.01 and 0.05. Multivariate statistical analysis can help to simplify and organise large data set to provide meaningful insight (Chen et al., 2007). Multivariate analysis of the water quality data set was performed using CA, PCA/FA techniques. These multivariate statistical techniques were applied on experimental data previously standardised using z-scale transformation in order to avoid misclassifications arising from the different orders of magnitude of both numerical value and variance of the parameters analysed (Wunderlin et al., 2001; Singh et al., 2004). The standardisation procedure, which is commonly used in multivariate statistical analysis, was used to eliminate the influence of different units of measurements and rendered the data dimensionless (Varol et al., 2012).

2.3.1 Cluster analysis

CA is an unsupervised pattern recognition technique whose primary purpose is to assemble objects into categories or clusters based on their independent variables or characteristics. This technique classifies objects, so that each object is similar to the others in the cluster with respect to a predetermined selection criterion. The resulting clusters should then exhibit high internal (within-cluster) homogeneity and high external (between clusters) heterogeneity (Shrestha and Kazama, 2007). Hierarchical clustering is the most common approach in which clusters are formed sequentially, by starting with the most similar pair of objects and forming higher clusters step-by-step. The result can be displayed as a dendrogram (tree diagram) which provides a visual summary of the clustering process, presenting a picture of the groups and their proximity, with a dramatic reduction in dimensionality of the original data (Wunderlin et al., 2001). The Euclidean distance usually gives the similarity between two samples and a distance can be represented by the difference between analytical values from the samples (Otto, 1998). In this study, hierarchical CA was performed using the Ward's method and Euclidean distance, which has proven to be a very efficient measure of similarity (Kotti et al., 2005). This method uses an analysis of variance approach to evaluate the distances between clusters, attempting to minimise the sum of squares of any two clusters that can be formed at each step (Wang et al., 2013). CA was applied to the water quality data set to group the similar sampling sites (spatial variability) and seasonal (temporal) similarity among the samples, resulting in spatial and temporal dendrograms. The linkage distance is reported as Dlink/Dmax, which represents the quotient between the linkage distances for a particular case divided by the maximal distance, multiplied by 100, as a way to standardise the linkage distance represented on the y-axis (Varol et al., 2012). Finally, ANOVA was performed ($p < 0.05$) to analyse the significant differences between the clusters obtained by the spatial CA. As the temporal CA generated only two clusters, differences between clusters were tested for statistical differences using a student T-test ($p < 0.05$).

2.3.2 Principal component analysis/factor analysis

Kaiser-Meyer-Olkin (KMO) and Bartlett's sphericity test were performed previously to examine the suitability of the data set for PCA/FA technique. KMO is a measure of sampling adequacy that indicates the proportion of variance that is common, i.e., variance that may be caused by underlying factors (Shrestha and Kazama, 2007). In this study the KMO was 0.78 (close to 1), indicating that PCA/FA may be useful. Bartlett's test of sphericity indicates whether a correlation matrix is an identity matrix, which would indicate that variables are unrelated (Varol et al., 2012). The significance level of 0.00 in this study (less than 0.05) indicated that there were significant relationships among the variables. The principal component analysis (PCA) extracts the eigenvalues and eigenvectors from the covariance matrix of original variables (measured parameters). An eigenvalue gives a measure of the significance of the factor, while an eigenvector is a list of coefficients (loadings or weightings) by which we multiply the original correlated variables to obtain new uncorrelated (orthogonal) variables, called principal components (PCs), which are weighted linear combinations of the original variables. A PC is the product of the original data and an eigenvector; the result of projecting the data on to a new axis is a new variable. There are as many PCs as original variables, however the first

PC loading explains the most variance and each subsequent PC explains progressively less. As a result, a small number of factors usually account for approximately the same amount of information as the much larger set of the original observations do (Kennedy et al., 2015). Thus, PC provides information on the most meaningful parameters, which describe the whole data set affording data reduction with minimal loss of information (Helena et al., 2000).

Factor analysis (FA) attempts to extract a lower dimensional linear structure from the data set. It further reduces the contribution of less significant variables obtained from PCA and the new group of variables known as varifactors (VFs) is extracted through rotating the axis defined by PCA. It should be noted that a PC is a linear combination of observable water quality variables, while a VF can include unobservable, hypothetical, 'latent' variables (Wunderlin et al., 2001). PCA/FA was performed on correlation matrix of rearranged data (all observations for each group of sites), so that it explains the structure of the underlying data set. The correlation coefficient matrix measures how well the variance of each constituent can be explained by relationship with each of the others (Singh et al., 2004). PCA of the normalised variables (water quality data set) was performed to extract significant PCs and to further reduce the contribution of variables with minor significance; these PCs were subjected to a varimax rotation (raw) generating VFs. According to Helena et al. (2000), by this method loadings and VF are obtained in which original variables participate more clearly.

2.3.3   Source apportionment

The source apportionment is an important environmetric approach aiming to the estimation of contribution of identified sources to the concentration of each parameter (Simeonov et al., 2003). After the determination of the number and identity of possible sources affecting surface waters by using PCA/FA, source contribution was calculated using multiple regression of sample mass concentration on the absolute VF scores. A detailed description of the modelling approach can be found in Thurston and Spengler (1985). This technique makes it possible to apportion the component mass among various source components obtained by PCA/FA (five components in our case). The PCA/FA assumes the total concentration of each element is made up of the sum of elemental concentrations from each identified component. The approach calculates the weight of source in the total sum using multiple regressions.

3   Results and discussion

3.1   Physico-chemical characteristics of reservoirs water

The basic statistics calculated for the studied reservoirs of Calamuchita valley are summarised in Table 1. This table reveals that the concentrations of all analysed parameters were below the prescribed maximum limits by World Health Organization guidelines for drinking water (WHO, 2006).

Table 1 Mean, range and standard deviation of different water quality parameters at different sampling sites

Parameter		Sampling sites										
		LMR					RTR					
		1	2	3	4	5	1	2	3	4	5	6
WST (°C)	M ± Sd	18.4 ± 5.8	17.9 ± 5.4	18.2 ± 5.4	17.6 ± 5.6	18.2 ± 4.5	19.7 ± 5.5	19.4 ± 5.3	23.8 ± 5.7	18.8 ± 5.2	20.3 ± 5.4	20.8 ± 5.7
	R	9.9 – 25.0	9.4 – 23.3	11.4 – 23.9	9.5 – 23.4	10.7 – 22.5	13.0 – 26.5	13.0 – 26.0	16.0 – 31.0	13.0 – 26.0	14.0 – 28.0	13.0 – 29.0
pH	M ± Sd	8.1 ± 0.9	7.5 ± 0.5	8.2 ± 1.8	7.8 ± 1.2	7.3 ± 0.5	7.8 ± 0.6	7.7 ± 0.7	7.7 ± 0.4	7.5 ± 0.4	7.9 ± 0.8	7.6 ± 0.5
	R	6.9 – 9.8	67 – 8.1	6.5 – 12.3	6.5 – 10.9	6.5 – 8.0	7.1 – 9.2	6.6 – 9.1	6.8 – 8.4	7.0 – 8.3	7.0 – 9.9	6.8 – 8.3
DO (mg/L)	M ± Sd	9.3 ± 1.5	8.6 ± 2.0	8.8 ± 0.9	9.1 ± 2.1	6.7 ± 1.6	8.3 ± 1.8	7.9 ± 1.6	8.1 ± 1.0	8.3 ± 1.6	8.3 ± 1.6	8.2 ± 1.5
	R	7.5 – 12.9	6.5 – 12.2	7.5 – 10.3	6.7 – 12.5	4.9 – 9.4	5.0 – 11.1	5.0 – 9.5	6.7 – 10.3	4.7 – 10.0	5.0 – 10.3	5.5 – 10.4
EC (µS/cm)	M ± Sd	88 ± 23	86 ± 20	88 ± 21	81 ± 21	88 ± 21	153 ± 23	161 ± 18	164 ± 23	164 ± 18	166 ± 19	186 ± 54
	R	58 – 130	57 – 117	61 – 119	58 – 114	57 – 118	127 – 192	138 – 202	128 – 200	143 – 198	142 – 192	101 – 301
SDT (m)	M ± Sd	1.1 ± 0.5	1.7 ± 0.5	1.6 ± 0.4	1.5 ± 0.4	2.2 ± 1.2	2.2 ± 0.5	3.7 ± 1.3	2.1 ± 1.0	3.6 ± 0.8	2.4 ± 0.5	1.5 ± 0.6
	R	0.4 – 1.9	1.0 – 2.7	1.1 – 2.3	1.0 – 2.3	0.7 – 4.5	1.4 – 3.0	1.9 – 6.1	0.7 – 4.0	2.4 – 5.3	1.5 – 3.0	0.9 – 2.8
Chl-a (mg/L)	M ± Sd	0.2 ± 0.1	0.2 ± 0.1	0.1 ± 0.0	0.2 ± 0.1	0.1 ± 0.1	0.2 ± 0.2	0.2 ± 0.2	0.2 ± 0.4	0.1 ± 0.2	0.1 ± 0.1	0.1 ± 0.1
	R	0.1 – 0.4	0.1 – 0.5	0.1 – 0.2	0.1 – 0.4	0.1 – 0.2	0.1 – 0.7	0.1 – 0.7	0.0 – 0.3	0.1 – 0.6	0.1 – 0.3	0.0 – 0.2
TP (mg/L)	M ± Sd	0.3 ± 0.1	0.4 ± 0.2	0.3 ± 0.1	0.4 ± 0.2	0.5 ± 0.3	0.2 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1
	R	0.1 – 0.5	0.2 – 0.8	0.1 – 0.4	0.2 – 0.5	0.2 – 0.7	0.1 – 0.2	0.2 – 0.5	0.2 – 0.5	0.2 – 0.5	0.2 – 0.5	0.2 – 0.4
TN (µg/L)	M ± Sd	0.7 ± 0.3	0.7 ± 0.3	0.7 ± 0.4	0.6 ± 0.2	0.8 ± 0.3	1.3 ± 0.2	1.9 ± 0.8	1.3 ± 0.7	1.0 ± 0.5	1.6 ± 0.7	1.4 ± 0.6
	R	0.5 – 1.3	0.5 – 1.3	0.5 – 1.3	0.5 – 0.9	0.5 – 1.3	1.2 – 1.5	1.2 – 3.4	0.9 – 2.9	0.5 – 1.6	0.7 – 2.3	0.6 – 2.3
NO3-N (mg/L)	M ± Sd	0.1 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.1 ± 0.1	0.2 ± 0.1	1.0 ± 0.5	1.2 ± 0.7	0.8 ± 0.5	1.0 ± 0.6	1.1 ± 0.6	1.1 ± 0.8
	R	0.1 – 0.3	0.1 – 0.3	0.1 – 0.3	0.1 – 0.2	0.1 – 0.3	0.7 – 1.4	0.1 – 2.1	0.1 – 1.2	0.1 – 1.9	0.1 – 1.9	0.2 – 2.7

Notes: M: mean; Sd: standard deviation; R: range.

Table 1   Mean, range and standard deviation of different water quality parameters at different sampling sites (continued)

Parameter		LMR					RTR					
		1	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*	*6*
Na (mg/L)	M ± Sd	8.0 ± 2.1	7.2 ± 1.3	7.2 ± 1.2	7.0 ± 1.0	7.2 ± 1.7	12.9 ± 3.8	12.3 ± 2.8	12.5 ± 2.3	12.1 ± 2.8	13.3 ± 3.8	16.0 ± 7.6
	R	6.1 − 10.2	5.3 − 9.9	6.3 − 8.0	5.3 − 8.3	5.0 − 10.8	9.1 − 17.7	9.1 − 19.2	9.1 − 16.7	8.1 − 18.2	10.1 − 20.2	5.9 − 30.3
K (mg/L)	M ± Sd	1.5 ± 0.1	1.6 ± 0.2	1.5 ± 0.1	1.6 ± 0.1	1.6 ± 0.2	3.1 ± 0.5	2.6 ± 0.5	2.6 ± 0.4	2.6 ± 0.5	2.6 ± 0.5	2.7 ± 0.6
	R	1.4 − 1.6	1.4 − 1.9	1.4 − 1.6	1.4 − 1.7	1.4 − 1.8	2.6 − 3.7	2.1 − 4.0	2.2 − 3.5	2.1 − 3.7	2.1 − 3.7	1.8 − 3.8
Ca (mg/L)	M ± Sd	9.0 ± 0.5	9.2 ± 1.6	9.1 ± 0.2	9.7 ± 1.5	9.2 ± 1.4	17.1 ± 2.6	15.8 ± 2.2	15.9 ± 2.6	15.6 ± 1.9	16.0 ± 1.5	24.0 ± 10.1
	R	8.6 − 9.5	7.1 − 11.8	8.9 − 9.2	7.8 − 11.7	7.3 − 11.4	13.6 − 20.0	13.6 − 20.0	12.0 − 20.0	12.8 − 18.4	14.4 − 17.6	3.5 − 18.0
Mg (mg/L)	M ± Sd	2.0 ± 0.1	2.1 ± 0.1	2.0 ± 0.1	2.1 ± 0.2	2.1 ± 0.1	4.3 ± 1.3	4.3 ± 0.8	3.9 ± 1.2	4.3 ± 0.8	4.1 ± 1.3	4.5 ± 1.3
	R	1.8 − 2.2	2.0 − 2.3	1.9 − 2.1	2.0 − 2.4	2.0 − 2.4	3.4 − 6.3	3.4 − 5.9	2.4 − 6.8	2.4 − 5.4	2.4 − 6.8	2.9 − 7.3
Cl (mg/L)	M ± Sd	2.1 ± 0.6	1.9 ± 0.7	1.8 ± 0.4	2.0 ± 0.5	2.0 ± 0.6	10.0 ± 3.7	7.9 ± 2.8	8.3 ± 3.9	8.0 ± 2.9	7.3 ± 3.5	8.0 ± 2.6
	R	1.6 − 2.8	0.4 − 2.7	1.5 − 2.0	1.3 − 2.7	1.2 − 3.2	5.7 − 14.3	2.9 − 11.4	2.9 − 14.3	2.9 − 11.4	2.9 − 14.3	2.9 − 11.4
SO4 (mg/L)	M ± Sd	5.0 ± 1.0	4.6 ± 1.5	5.0 ± 1.4	5.1 ± 1.6	4.8 ± 1.5	10.3 ± 7.6	14.2 ± 5.7	15.9 ± 5.0	16.5 ± 6.5	16.3 ± 3.7	19.1 ± 7.1
	R	4.0 − 6.0	3.0 − 8.0	4.0 − 6.0	3.0 − 8.0	3.0 − 8.0	3.5 − 21.2	6.3 − 24.8	7.1 − 23.2	2.8 − 24.4	11.0 − 20.8	3.5 − 28.3
F (mg/L)	M ± Sd	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.2	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.4 ± 0.3	0.5 ± 0.2
	R	0. − 0.4	0.2 − 0.5	0.2 − 0.4	0.2 − 0.4	0.2 − 0.4	0.2 − 0.4	0.1 − 0.4	0.1 − 0.4	0.1 − 0.4	0.2 − 1.0	0.2 − 1.0
T-Hard (mg/L)	M ± Sd	30 ± 2	31 ± 4	31 ± 1	33 ± 4	31 ± 4	61 ± 9	57 ± 6	565 ± 9	56 ± 7	57 ± 8	62 ± 11
	R	29 − 32	26 − 38	31 − 32	28 − 38	26 − 37	50 − 70	50 − 65	44 − 72	44 − 64	46 − 70	40 − 80
T-Alk (mg/L)	M ± Sd	43 ± 4	43 ± 5	42 ± 6	43 ± 5	42 ± 5	69 ± 9	65 ± 9	64 ± 10	65 ± 8	64 ± 10	65 ± 16
	R	38 − 46	37 − 50	38 − 46	38 − 51	37 − 50	60 − 80	50 − 85	50 − 80	50 − 80	50 − 85	40 − 90

Notes: M: mean; Sd: standard deviation; R: range.

Table 2   Pearson correlation matrix of the 18 variables determined

	WST	pH	DO	EC	SDT	Chl-a	TP	TN	NO3-N	Na	K	Ca	Mg	Cl	SO4	F	T-Hard	T-Alk
WST	1																	
pH	0.01	1																
DO	*–0.26*	0.08	1															
EC	*0.23*	–0.02	–0.16	1														
SDT	–0.05	–0.09	–0.15	*0.19*	1													
Chl-a	0.08	0.10	0.08	0.04	–0.16	1												
TP	0.18	–0.16	–0.22	*–0.21*	–0.19	0.11	1											
TN	*0.31*	–0.06	–0.19	**0.66**	0.25	–0.06	–0.13	1										
NO3-N	*0.31*	–0.06	–0.01	**0.64**	0.28	–0.22	–0.15	0.48	1									
Na	0.25	0.13	–0.06	**0.87**	–0.03	–0.13	–0.21	0.46	*0.59*	1								
K	0.26	0.02	–0.21	**0.86**	0.08	–0.02	*–0.39*	**0.55**	*0.58*	**0.78**	1							
Ca	0.14	0.08	–0.22	**0.90**	0.09	–0.02	*–0.39*	**0.61**	*0.59*	**0.73**	**0.87**	1						
Mg	0.19	*0.21*	–0.13	**0.81**	0.18	0.01	*–0.41*	**0.52**	*0.50*	**0.73**	**0.75**	**0.75**	1					
Cl	0.22	0.15	–0.16	**0.73**	0.17	0.03	*–0.48*	0.42	0.46	**0.68**	**0.85**	**0.73**	**0.85**	1				
SO4	*0.44*	–0.03	0.09	**0.75**	0.16	–0.07	–0.24	0.44	**0.69**	**0.71**	**0.62**	**0.64**	**0.64**	**0.57**	1			
F	–0.11	0.06	0.05	0.16	–0.19	0.08	–0.12	–0.04	–0.12	0.27	0.09	0.17	0.11	0.06	0.06	1		
T-Hard	0.16	0.14	–0.19	**0.90**	0.14	0.01	*–0.43*	**0.61**	**0.60**	**0.74**	**0.88**	**0.97**	**0.87**	**0.83**	**0.67**	0.13	1	
T-Alk	0.25	0.09	–0.12	**0.90**	0.10	–0.06	*–0.37*	**0.61**	**0.52**	**0.82**	**0.88**	**0.86**	**0.78**	**0.78**	**0.67**	0.07	**0.87**	1

Note: Italic and bold indicate correlation is significant at the 0.05 level, and values higher than 0.50, respectively.

MANOVA indicated significant temporal and spatial variations ($p < 0.05$). ANOVA indicated significant temporal variations ($p < 0.05$) for most water quality parameters except for SDT, Chl-a, TN, NO_3-N and F. WST and TP displayed higher values in summer, while higher values of pH and DO occurred in winter. Significant spatial variations ($p < 0.05$) were observed in almost all water quality variables except for pH, DO, Chl-a and F. Higher concentration of PT were observed in LMR, while other variables showed higher values in RTR. Table 2 shows the Pearson correlation matrix of the 18 analysed variables. Only those correlation values higher than 0.5 were considered and highlighted in this table. Inspection of correlation matrix and analysis of variance showed that EC, Na, K, Ca, Mg, Cl, SO_4, T-Hard and T-Alk, which displayed significant spatial and temporal variations ($p < 0.05$) with lower values in fall and higher concentrations in spring, were closely correlated ($r = 0.50$ to 0.97). We consider that most of these associations are responsible for water mineralisation, being directly related to hydrochemical characteristics of the region.

3.2   Spatial and temporal similarities

CA was applied on reservoirs water quality data, to detect similarity for grouping of sampling sites and climatic seasons. Dendrograms of spatial and temporal variability obtained by the Ward's method are shown in Figure 2. The spatial CA, generated a dendrogram grouping 11 sampling sites into five clusters at (Dlink/Dmax) $\times$ 100 < 28 [Figure 2(a)]. The spatial CA was useful to separate between analysed reservoirs. Thus, cluster 1 corresponded to sites 5 and 6 of RTR. Cluster 2 consisted of one site (RTR-3). Cluster 3 corresponded to sites 1, 2 and 4 of RTR. While cluster 4 corresponded to site 5 of LMR, and cluster 5 consisted of sites 1 to 4 of LMR. Further, the cluster classification varied with significance level and in a very convincing way because sites in these clusters had similar characteristics features and natural backgrounds that were affected by similar sources. Each cluster represented geographical location of sampling sites located at each reservoir. Thus, cluster 2 and 3 could be considered to a relatively high pollution region because were directly influenced by river influxes of RTR as shown in Figure 1, receiving polluted effluents from different sources such as stream inputs, agricultural runoff and soil erosion of the RTR watershed. The results from one-way ANOVA used to compare difference between clusters showed that only pH and Chl-a did not showed significant variations ($p > 0.05$). With the exception of SDT and TP, most variables were significantly higher in these clusters ($p < 0.05$). Further, WST was significantly higher in cluster 2 than in the rest of the groups. This cluster is also influenced by the effluent of CNE which, as it was mentioned, raises the temperature of water. Similar results were found by Policht-Latawiec et al. (2016) studing the effect of a cooling water discharge from a power station on the water quality of a river in Poland. Cluster 5, which is related with river loads of LMR, could be considered as moderate pollution region because watersheds of these rivers present lower anthropic activities and higher areas of natural vegetation than RTR. However, in last years these watersheds have reported an increase in agricultural activities and a decrease in native forest. Hence it is possible that in the near future this cluster will be more related to clusters of high pollution regions. Cluster 1 and 4 correspond to relatively low pollution regions. These clusters, situated in the west region of RTR and LMR respectively, are located far from major point and non-point pollution sources such as river loads. This was confirmed by results of ANOVA, which showed that clusters 1 and 4 presented the higher values of SDT. Thus, the spatial

variations of water quality in the principal reservoirs of Calamuchita valley demonstrated that water quality is better in the eastern areas than in the western portions of the reservoirs related with river loads.

Figure 2 Dendrograms showing hierarchical clustering of (a) sampling sites and (b) climatic seasons using the Ward's method and Euclidean distance

(a)   (b)
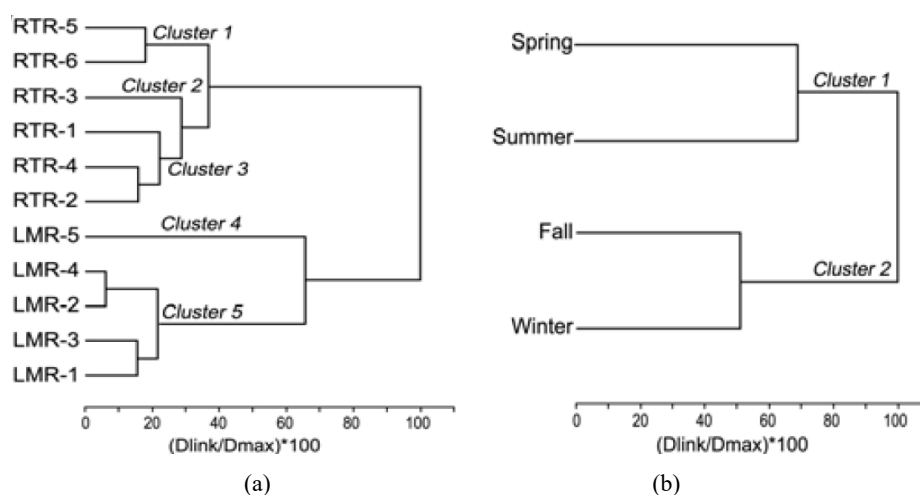
The temporal dendrogram obtained by CA [Figure 2(b)] generated two clusters at (Dlink/Dmax) × 100 < 55. Cluster 1 corresponded to the wet seasons including spring and summer. Cluster 2 corresponded to the dry season (fall and winter). Although this was an expected result, CA was able to corroborate this assumption and suggests that sampling during only two seasons (wet and dry) in a year may suffice for assessment of temporal variations in water quality of the reservoirs. The results from Student's T test showed that WST and SO_4 were significant higher in the rainy season ($p < 0.05$), whereas DO, TN and T-Alk were higher in the dry season.

It is evident that the CA technique is useful in offering reliable classification of surface waters in the study area and helpful for rapid assessment of water quality because only one site of each spatial and temporal cluster may serve as a good spatial assessment of water quality in the whole region. Further, we confirm that it is possible to design an optimal future spatial and temporal monitoring strategy in an optimal manner with fewer sampling sites and lower associated costs without missing much information in reservoirs of Calamuchita valley.

3.3 Data structure determination and source identification

PCA/FA analysis was applied to the normalised data sat (18 variables) to compare the compositional patterns between the analysed water samples and to identify the factors that influence each one. There are several criteria to identify the number of VFs to be retained, in order to understand the underlying data structure (Helena et al., 2000). In the present study, factors with eigenvalues greater than 1.0 were taking into account. Thus, PCA/FA renders five VFs with eigenvalues > 1.0 accounting for 79.3% of total variance in the water quality data set (Table 3).

Table 3   Loadings of experimental variables on the first five VFs for the data set

Parameters	VF1	VF2	VF3	VF4	VF5
WST	0.13	0.05	**0.88**	0.00	0.07
pH	0.10	0.26	–0.08	**0.75**	0.20
DO	–0.13	–0.17	–0.17	**0.86**	–0.20
EC	**0.95**	–0.09	–0.03	0.01	0.02
SDT	0.26	–0.22	–0.28	*0.56*	–0.25
Chl-a	0.14	**0.82**	0.15	0.15	–0.09
TP	.0.46	**0.64**	0.44	–0.20	–0.04
TN	0.44	*0.53*	0.18	–0.14	0.02
NO3-N	0.37	*0.54*	0.35	0.05	–0.19
Na	**0.79**	–0.25	0.23	0.22	–0.01
K	**0.91**	0.01	–0.16	–0.19	0.02
Ca	**0.95**	–0.03	–0.08	–0.11	0.03
Mg	**0.92**	0.15	0.00	–0.02	–0.08
Cl	**0.80**	0.32	–0.28	–0.20	0.05
SO_4	**0.76**	–0.17	0.37	0.26	–0.27
F	0.01	–0.05	0.04	0.08	0.33
T-Hard	**0.98**	0.02	–0.06	–0.08	0.00
T-Alk	**0.89**	0.06	–0.14	0.04	0.00
Eigenvalue	8.320	1.890	1.640	1.380	1.040
% total variance	46.207	10.501	9.102	7.680	5.768
Cumulative % variance	46.207	56.709	65.810	73.490	79.258

Bold and italic values indicate strong and moderate loadings, respectively. According to Liu et al. (2003), the factor loadings were classified as 'strong', 'moderate' and 'weak' corresponding to absolute loading values of > 0.75, 0.75 – 0.50 and 0.50 – 0.30, respectively. Thus, the first VF (VF1) was the most important with 46.2% of the total variance. VF1 showed nine parameters with strong positive loadings > 0.75, being correlated with EC, Na, K, Ca, Mg, Cl, SO_4, T-Hard, and T-Alk. This VF may be interpreted as the 'mineral' factor of the reservoirs, because is related to natural sources of the ionic groups of salts in the watersheds from inflows, soil weathering and runoff. These results agree with those obtained with the Pearson correlation test (Table 2).

The second VF (VF2), which was responsible for 10.5% of the total variance, has strong positive loadings on Chl-a and TP, and moderate positive loadings on TN, and NO_3-N. This VF was named as the 'nutrient' source of the variability, and may be interpreted as representing influences from stream inputs, point sources such as municipal and industrial effluents, and non-point sources such as agricultural runoff, erosion in the watersheds and atmospheric deposition. Farming and cattle rising are common activities in the studied watersheds. Fertilisers and manure, which can contribute to high levels of soil nitrogen and phosphorus, can be incorporated into the reservoirs by streams and

runoff after rainfall, increasing eutrophication processes which often lead to a reduction in the supply of ecosystems services (Saaremäe et al., 2014). Thus, the priority is to develop advanced techniques for decreasing nutrient sources of pollution in this region. VF2 also showed the effect of nutrients on Chl-a, which is one of the most important parameters since it is used to estimate the primary productivity of water ecosystems (Dall'Olmo et al., 2005). Subsequent nutrient abatement strategies implemented by environmental agencies or water authorities could be largely successful resulting in a reduction of algal biomass and better reservoirs water quality. VF3, accounting for 9.1% of the total variance, showed only WST with strong significant loading, so it could be interpreted as the 'temperature' factor. The contribution of WST, which showed a very characteristic annual cycle with highest values in summer and lowest in winter, could also be related with the constant input of heat from the cooling channel of CNE nuclear power plant in RTR. According to Bonansea et al. (2015b), the nuclear power plant generates a thermal plume from the cooling channel, which dissipates to the rest of RTR. Although Mariazzi et al. (1992) suggest that the constant influx of heat has not generated an increase in WST over time, as mean WST in RTR remains stable since CNE was installed, the effect of the nuclear power plant in WST should be analysed in detail in future studies. Additional 7.7% of the total variance was explained in VF4, which was interpreted as the 'physico-chemical' factor. This VF had strong contribution on pH and DO, and moderate contribution on SDT. Finally, VF5 explaining lower variance (5.7%) did not show significant loading for any parameter, suggesting that this VF may indicate a 'non-common origin' or source.

Many potential pollution sources have been identified by using PCA/FA in reservoirs of Calamuchita valley. Thus, the results from PCA/FA suggested that most of the variations in water quality are explained by a set of soluble salts (natural) included in VF1, a substantial contribution comes from anthropogenic pollution sources, while minor contributions arise from physico-chemical and non-common sources. In this study PCA/FA did not result in much data reduction, as we need 16 parameters (88.9% of the 18 parameters) to explain 79.3% of the data variance. However, this technique served as a means to identify those parameters that had the greatest contribution to variation in water quality and suggested possible sets of pollution sources in the region.

3.4 Source apportionment

Results of the contribution of the possible sources in each source type obtained by multiple regression of sample mass concentration on the absolute VF scores is presented in Table 4. As evident from the higher determination coefficients (R^2) and the lower absolute root mean square errors (RMSE), the multiple regression exhibited good adequacy between the measured and predicted values for most parameters. A significant biased part of the total concentration is present for almost all parameters. Results showed that natural sources of the ionic groups of salts in Calamuchita valley from inflows, soil weathering and runoff are the main contribution to the mineral factor. Point (municipal and industrial effluents) and non-point sources (agricultural runoff, erosion, and atmospheric deposition) are the main contributors to the nutrient parameter, whereas the physico-chemical factors appeared to be the main source of most elements.

Table 4   Source contribution to surface waters in Calamuchita valley

Parameter	Intercept	Source types					R^2	RMSE
		1 mineral	2 nutrient	3 temperature	4 physico-chemical	5 non-common source		
WST	18.42	–	–	4.84	–	–	0.80	1.56
pH	7.65	–	0.20	–	0.42	–	0.43	0.28
DO	8.33	–	–0.34	–0.33	1.70	–0.39	0.85	0.42
EC	116.23	37.27	–	–	–	–	0.91	6.28
SDT	2.36	0.31	–0.26	–0.34	–0.67	–0.29	0.58	0.34
Chl-a	0.01	–	0.01	–	–	–	0.75	0.01
TP	0.04	–0.01	0.05	0.01	0.01	–	0.67	0.01
TN	1.03	0.46	0.02	–	–	–	0.59	0.16
NO3-N	0.47	0.34	0.32	0.21	–	–0.11	0.79	0.12
Na	9.02	1.60	–0.51	0.47	0.45	–	0.80	0.49
K	1.98	0.45	–	–0.08	–0.10	–	0.89	0.08
Ca	12.70	4.10	–	–	–0.46	–	0.93	0.58
Mg	2.84	1.00	0.17	–	–	–	0.87	0.10
Cl	3.90	2.26	0.92	–0.81	–0.55	–	0.87	0.35
SO_4	8.53	4.59	–1.02	2.27	1.57	–1.62	0.88	0.65
F	0.33	–	–	–	–	0.12	0.87	0.01
T-Hard	43.43	14.37	–	–0.88	–1.23	–	0.97	1.30
T-Alk	51.07	9.68	–	–	–	–	0.82	1.83

4   Conclusions

Monitoring is the core of all water quality management practices since decisions to be made are based on available information related to water quality processes. The physical, chemical and biological quality of reservoir water is critically important, because they are linked to every aspect of human wellbeing and sustainable development. In this study, different multivariate statistical techniques were successfully used to assess temporal and spatial variation in surface water quality of the most important reservoirs of Calamuchita valley and to identify the main contaminants and their sources.

According to similar water quality characteristics, hierarchical CA grouped the 11 sampling sites into five clusters of relative high, moderate and low pollution, and classified climatic seasons in two clusters (wet and dry season). Based on information obtained from this study, it is possible to design an optimal future monitoring strategy, resulting in a reduced number of sampling sites, monitoring frequency and lower associated costs. Although PCA/FA did not result in much data reduction, it helped to identify the factors or sources responsible for water quality variations. VFs obtained from PCA/FA indicated that the parameters responsible for water quality variations are mainly related to soluble salts (natural), organic pollution and nutrients (point and non-point), physico-chemical parameters, and non-common source. Further, source contribution was

calculated using multiple regression of sample mass concentration on the absolute VF scores. Thus, the used multivariate statistical techniques served as an excellent exploratory tool in the analysis and interpretation of complex data set on water quality, in identifying pollutant sources, and in understanding their temporal and spatial variations for effective reservoirs water management. This study could be considered for future planning and management of reservoirs. The methodology presented here could be very useful not only to the local authorities for the pollution control/management, but also for all aquatic systems, either natural lakes or artificial dams, at a regional scale.

References

APHA-AWWA-WEF (2000) *Standard Methods for the Examination of Water and Wastewater*, 18th ed., APHA-AWWA-WEF, Washington, DC.

Banihabib, M., Arabi, A. and Salha, A. (2015) 'A dynamic artificial neural network for assessment of land-use change impact on warning lead-time of flood', *International Journal of Hydrology Science and Technology*, Vol. 5, No. 2, pp.163–178.

Bazán, R., Corral, M., Pagot, M., Rodríguez, A., Rodríguez, N., Larrosa, N., Cossavella, A., del Olmo, S., Bonfanti, E. and Busso, F. (2005) 'Teledetección y modelado numérico para el análisis de la calidad de agua del embalse Los Molinos', *Ingeniería hidráulica en México*, Vol. 20, No. 2, pp.121–135.

Bonansea, M., Bazán, R., Ledesma, C., Rodriguez, C. and Pinotti, L. (2015a) 'Monitoring of regional lake water clarity using Landsat imagery', *Hydrology Research*, Vol. 46, No. 5, pp.661–670.

Bonansea, M., Rodriguez, M., Pinotti, L. and Ferrero, S. (2015b) 'Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina)', *Remote Sensing of Environment*, Vol. 158, No. 1, pp.28–41.

Chen, K., Jiao, J., Huang, J. and Huang, R. (2007) 'Multivariate statistical evaluation of trace elements in groundwater in a coastal area in Shenzhen, China', *Environmental Pollution*, Vol. 147, No. 3, pp.771–780.

Dall'Olmo, G., Gitelson, A., Rundquist, D., Leavitt, B., Barrow, T. and Holz, J. (2005) 'Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands', *Remote Sensing of Environment*, Vol. 96, No. 2, pp.176–187.

Helena, B., Pardo, R., Vega, M., Fernandez, J. and Fernandez, L. (2000) 'Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis', *Water Research*, Vol. 34, No. 3, pp.807–816.

Hirsch, R., Archfield, S. and De Cicco, L. (2015) 'A bootstrap method for estimating uncertainty of water quality trends', *Environmental Modelling & Software*, Vol. 73, No. 1, pp.148–166.

Kennedy, D., Dukic, V. and Dwyer, G. (2015) 'Combining principal component analysis with parameter line-searches to improve the efficacy of Metropolis-Hastings MCMC', *Environmental and Ecological Statistics*, Vol. 22, No. 2, pp.247–274.

Kotti, M., Vlessidis, A., Thanasoulias, N. and Evmiridis, N. (2005) 'Assessment of river water quality in Northwestern Greece', *Water Resources Management*, Vol. 19, No. 1, pp.77–94.

Liu, C., Lin, K. and Kuo, Y. (2003) 'Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan', *Science of the Total Environment*, Vol. 313, No. 1, pp.77–89.

Mariazzi, A., Donadelli, J., Arenas, P., Di Siervi, M. and Bonetto, C. (1992) 'Impact of a nuclear power plant on water quality of Embalse del Río Tercero Reservoir (Córdoba, Argentina)', *Hidrobiologia*, Vol. 246, No. 2, pp.129–140.

Megalovasilis, P. (2014) 'A new hydrological balance and water management issues in Zakynthos Island, Greece', *International Journal of Hydrology Science and Technology*, Vol. 4, No. 2, pp.139–154.

Otto, M. (1998) 'Multivariate methods', in Kellner, R., Mermet, J.M., Otto, M. and Widmer, H.M. (Eds.): *Analytical Chemistry*, Wiley-VCH, Weinheim, Germany.

Policht-Latawiec, A., Kanownik, W. and Jurek, A. (2016) 'The effect of cooling water discharge from the power station on the quality of the skawinka river water', *Carpathian Journal of Earth and Environmental Sciences*, Vol. 11, No. 2, pp.427–435.

Saaremäe, E., Liira, M., Poolakese, M. and Tamm, T. (2014) 'Removing phosphorus with Ca-Fe oxide granules – a possible wetlands filter material', *Hydrology Research*, Vol. 45, No. 3, pp.368–378.

Shrestha, S. and Kazama, F. (2007) 'Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji River basin, Japan', *Environmental Modelling & Software*, Vol. 22, No. 4, pp.464–475.

Simeonov, V., Stratis, J., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M. and Kouimtzis, T. (2003) 'Assessment of the surface water quality in Northern Greece', *Water Research*, Vol. 37, No. 17, pp.4119–4124.

Singh, K., Malik, A., Mohan, D. and Sinha, S. (2004) 'Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – a case study', *Water Research*, Vol. 38, No. 18, pp.3980–3992.

Su, Z., Lin, C., Ma, R., Luo, J. and Liang, Q. (2015) 'Effect of land use change on lake water quality in different buffer zones', *Applied Ecology and Environmental Research*, Vol. 13, No. 2, pp.489–503.

Thurston, G. and Spengler, J. (1985) 'A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston', *Atmospheric Environment*, Vol. 19, No. 1, pp.9–25.

Varol, M., Gökot, B., Bekleyen, A. and Sen, B. (2012) 'Spatial and temporal variations in surface water quality of the dam reservoirs in the Tigris River basin, Turkey', *Catena*, Vol. 92, No. 1, pp.11–21.

Wang, M., Markert, B., Chen, W., Peng, C. and Ouyang, Z. (2012) 'Identification of heavy metal pollutants using multivariate analysis and effects of land uses on their accumulation in urban soils in Beijing, China', *Environmental Monitoring and Assessment*, Vol. 184, No. 10, pp.5889–5897.

Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L. and Li, B. (2013) 'Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China', *Journal of Hydro-environment Research*, Vol. 7, No. 1, pp.30–40.

WHO (2006) *Guidelines for Drinking Water Quality*, 1st Addendum to 3rd ed., Vol. 1, World Health Organization, Geneva, Switzerland.

Wunderlin, D., Diaz, M., Ame, M., Pesce, S., Hued, A. and Bistoni, M. (2001) 'Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia River basin (Cordoba, Argentina)', *Water Research*, Vol. 35, No. 12, pp.2881–2894.

Yan, F., Liu, L., Li, Y., Zhang, Y., Chen, M. and Xing, X. (2015) 'A dynamic water quality index model based on functional data analysis', *Ecological Indicators*, Vol. 57, pp.249–258.