

# Analysis of the environmental liabilities generated by past activities in uranium mining exploitation in the Province of Córdoba, Argentina

Laura Folguera<sup>1,2</sup>  · Jorge Magallanes<sup>1,2</sup> · Daniel Cicerone<sup>1,2</sup>

Received: 28 April 2015 / Accepted: 19 October 2015 / Published online: 24 February 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** Worldwide, old disposal sites of uranium mine and mill wastes are the objects of environmental restoration programs or have already been remediated. This is the case of Los Gigantes in Córdoba, Argentina, where uranium was extracted and processed in the 1980s; a local source of pollution to watercourses was generated as a consequence of disposal of solid and liquid wastes. The present study aims at describing the physicochemical characteristics of the surface watercourses that run across the complex, finding a grouping structure of the sampling sites according to their degree of pollution, and defining the variables that are significant to that grouping. The problem is addressed with both traditional and robust statistics techniques; additional chemometrical tools are also applied. It was found that streams close to the tailings and the pond exhibit lower pH and higher concentration of anions and metals compared to upstream watercourses. As the distance downstream to these pollution sources increases, all physicochemical parameters recover gradually, reaching levels close to background and complying with provincial and national regulations, proving that pollution is locally constrained.

**Keywords** Uranium mining · Water pollution · Chemometrics · Robust statistics

## Introduction

Uranium mining remains controversial, principally because of environmental and health impacts associated with the early years of development of this industry. During the largely unregulated early phase several environmental impacts took place worldwide. Consequently many old disposal sites of uranium mine and mill tailings are now the objects of environmental restoration programs or have already been remediated. With over 430 nuclear reactors operational worldwide at the end of 2013, more than 70 under construction and many more under consideration, the provision of fuel for these long-lived facilities will be essential for the uninterrupted generation of significant amounts of base load electricity for decades to come. As a result, enhancing awareness of leading practices in uranium mining is increasingly important (International Atomic Energy Agency (IAEA) 2008; Nuclear Energy Agency (NEA) 2014).

Uranium mining and milling operations have been performed in Argentina since the 1950s. These activities took place in the past under a much less restrictive framework than current legislation. As a result different volumes of tailings and waste ponds exist that have been proved to be a source of pollution (Uranium Mining Environmental Restoration Project (PRAMU) 2005) and thus require to be adequately addressed. The country is currently running a “Uranium Mining Environmental Restoration Project” (PRAMU) to remediate legacies from former uranium mining and processing activities at eight different sites. This project, the first of its kind in the region, has so far enabled progress in civil works in the first site addressed, Malargüe, and financed baseline studies and radiological studies in two other sites: El Chichón and Los Gigantes (The World Bank 2014). This paper evaluates the situation

---

✉ Laura Folguera  
laura.folguera@gmail.com

<sup>1</sup> Comisión Nacional de Energía Atómica, Centro Atómico Constituyentes, Buenos Aires, Argentina

<sup>2</sup> Instituto de Investigación e Ingeniería Ambiental (3iA), Universidad Nacional de San Martín, Buenos Aires, Argentina

of the former uranium production site Los Gigantes, Córdoba, Argentina.

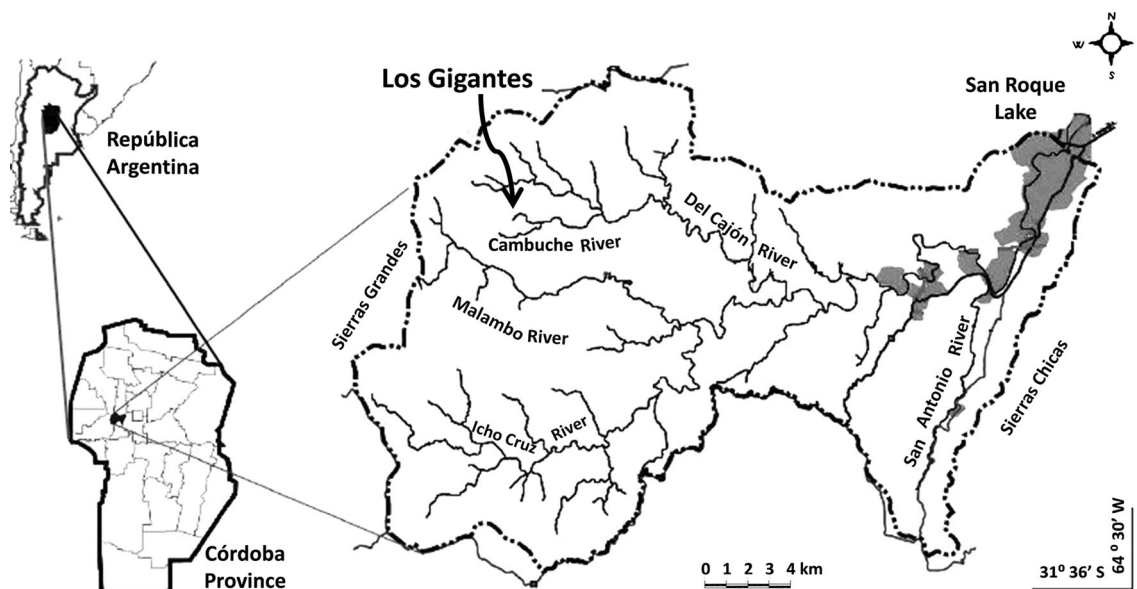
Los Gigantes uranium mining site is located in the largely unpopulated mountainous granite Sierra Grande area, about 30 km southwest from the city of Villa Carlos Paz, Córdoba Province, Argentina (Fig. 1). In the area there have been recognized Paleozoic granitic rocks and detrital Quaternary sediments. In general, granites contain potassium feldspar, quartz, sodium plagioclase, biotite and muscovite and have a porphyritic texture which makes them easily strippable by weathering agents as well as structural factors, facilitating the entry and circulation of water. The presence of pegmatitic tabular intrusions with varying thicknesses and positions along rifts is frequent, especially the sub-horizontal or low angle rifts. This granite is the host rock for uranium mineralization (autunite and metaautunite). Rifts affecting these materials are the main ways for the circulation of water in them. There are many springs of meager flow that originate from these structures; most of these springs water emerges from horizontal to subhorizontal position fractures. The Quaternary in the area is mainly represented by a layer of loose granular material (regolith), which covers the granite bedrock. It has an important horizontal extension but is of little vertical development, with variable thickness from a few centimeters to 1.5–2.0 m. The sediment is predominantly sandy loam. Finally, fluvial deposits partially fill the main watercourses of the area, with thicknesses below the meter. It is generally sandy sediment. The ore is of very low grade, yielding only 300 g of uranium per ton. The open pit mine had an extension of 300 × 500 m. The exploitation

was operated by a private company, under a contract with the Argentinean Comisión Nacional de Energía Atómica (CNEA), from 1982 to 1989.

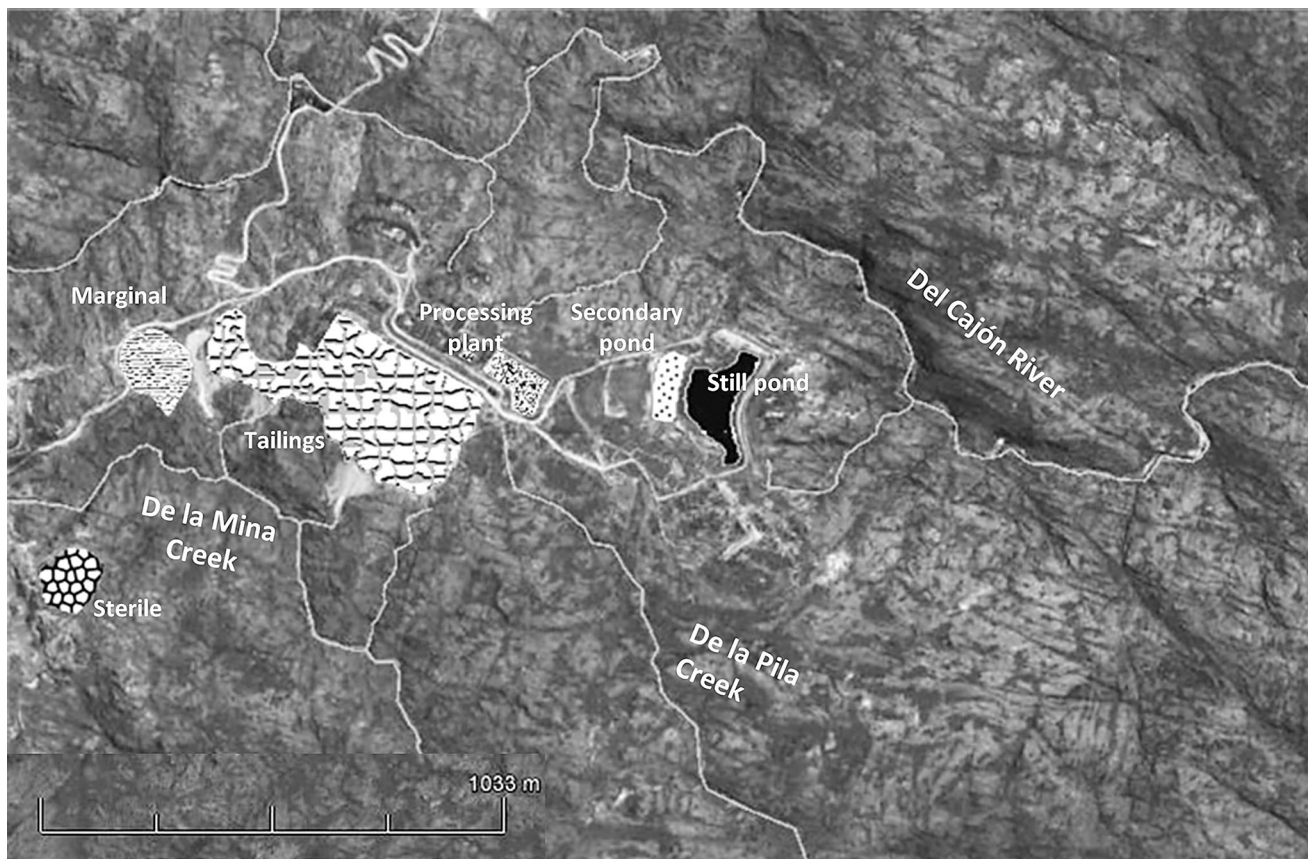
Mined ore was heap leached on-site; this process implied that the ore that was removed from the open-pit was crushed in a large mill, ground to sand consistency and mounded above grade on a prepared pad with an impermeable asphaltic membrane in its bottom. The ore was then leached with sulfuric acid so the uranium was mobilized:  $\text{UO}_3 + 2\text{H}^+ \rightleftharpoons \text{UO}_2^{2+} + \text{H}_2\text{O}$ . The uranyl ion in solution forms stable complexes with the sulphate ion. The uranium in this pregnant leach liquor was recovered by ion exchange technique and then recovered from the eluate by precipitation with ammonium hydroxide:  $2\text{UO}_2\text{SO}_4 + 6\text{NH}_4\text{OH} \rightleftharpoons (\text{NH}_4)_2\text{U}_2\text{O}_7 + 2(\text{NH}_4)_2\text{SO}_4 + 3\text{H}_2\text{O}$ . The diuranate was finally dewatered and dried to yield  $\text{U}_3\text{O}_8$  (yellowcake). The effluent and leached residue were neutralized by the addition of  $\text{Ca}(\text{OH})_2$  and discharged as a slurry into the waste retention pond.

When the mine had effectively ceased operation, 206 tons of uranium had been produced as ammonium diuranate. As a result of mineral extraction and processing, 3.5 million tons of solid wastes were generated and disposed of in separate piles: 2.5 million tons of tailings, 600,000 tons of low-grade (marginal) ore and 1 million tons of sterile overburden. Furthermore, 105 m<sup>3</sup> of slurry along with 105 m<sup>3</sup> of liquid effluents were collected and stored in a still pond (PRAMU 2005) (Fig. 2).

The major concern in mill tailings from a geochemical standpoint is the generation of acid pore waters. In low-pH waters, desorption of radionuclides and hazardous elements



**Fig. 1** Location of the ex-mining industrial complex Los Gigantes in Córdoba, Argentina. The gray surface on the right corresponds to populated areas



**Fig. 2** Waste facilities in Los Gigantes mining area

occurs, increasing their concentration. Seepage of such acid waters, in particular from unlined mill tailings, may affect the surrounding environment, including soils, surface water, and groundwater. In acid water leaching from tailings, dissolved uranium and daughter products can reach levels that are a direct threat to nearby ecosystems (Abdelouas 2006).

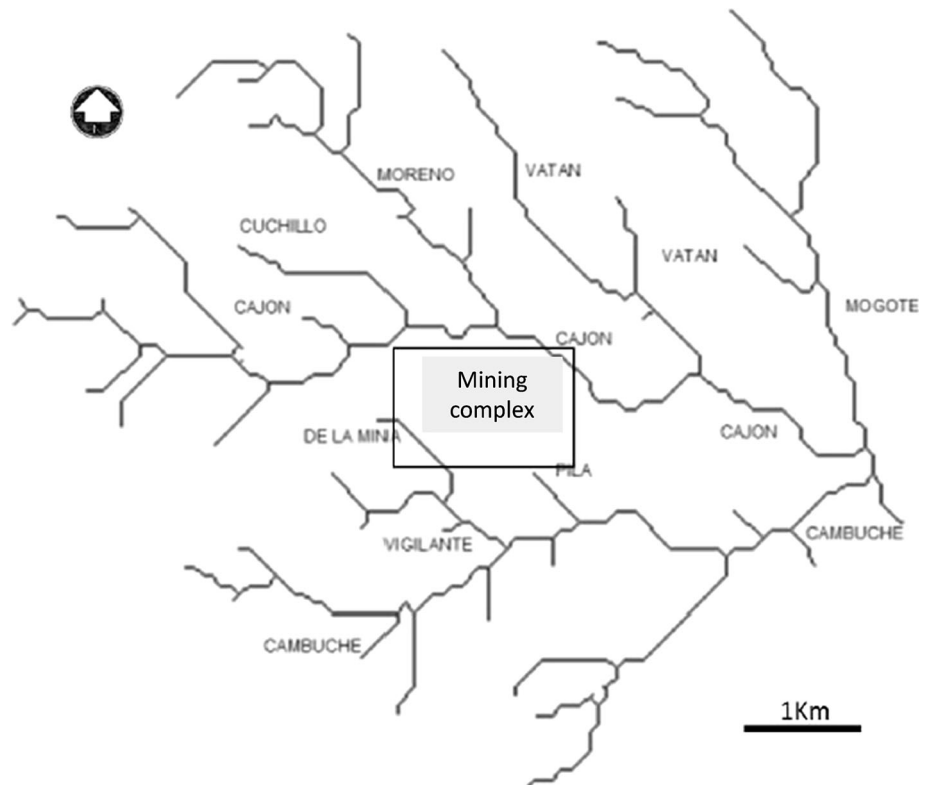
The watercourses that run across the industrial complex are headstreams of the San Antonio River Basin that flows down into the San Roque Lake. This lake is the major water source which provides drinking water to Villa Carlos Paz and Córdoba cities, the former located 30 km northeast from the industrial complex (Fig. 1). On the basis of previous hydrological research (Instituto Nacional del Agua (INA) 2001; División Química del Agua y del Suelo (DQAS) 2012) it is possible to determine that the streams that run through the mining solid or liquid wastes contribute with a flow rate of only  $11.5 \text{ l s}^{-1}$ . Considering that the principal collector of the basin, the San Antonio River, has a mean discharge of  $1057 \text{ l s}^{-1}$  before it flows into the lake, this input is less than 1 %.

At a regional scale, the zone coincides with the fracture parallel to the mountain range, Los Gigantes, a groundwater recharge and discharge area, where water outflows as

springs. The specific land where the industrial complex was built is known as La Mesada; it is part of the Achala batholith, a formation of fractured granitic rocks which holds an aquifer that is defined as poor according to its low permeability:  $10^{-2}$  to  $10^{-4} \text{ m}^3 \text{ day}^{-1}$  (Instituto Nacional del Agua (INA) 2001). The drainage system is related to the structure of the underlying rock (Fig. 3). The region has a high drainage density and an angular drainage pattern, governed by fractured and diaclased granites and pegmatites. Steep natural slopes occur along with a shallow vegetation cover. The occurrence of severe convective storms and orographic effect result in a marked tendency to sudden rising and flooding. Rainfall makes up for the largest runoff volume during the summer time. During the dry months, the drainage system is supplied from underground flows. The River Cajón subbasin has a length of 0.13 km; in addition the river's frequency is  $2.28 \text{ rivers km}^{-2}$ , which altogether indicate short catchment response time.

Rainwater falls on the mining wastes before infiltrating into the granitic rocks, thus altering, locally, its chemical characteristics. As a result, acid mining drainage, characterized by low pH and high concentration of cations and anions is found in the most impacted sites. Ions come from both natural occurring minerals that mobilize in contact

**Fig. 3** El Cajón subbasin runoff structure. At a regional scale water flows due east



with the acidic water and chemical products used in the industrial process such as sulphuric acid and calcium hydroxide. Previous monitoring data showed that the contamination sources are basically presently contained, with background values of chemical constituents (uranium, radon, heavy metals and other ions) downstream from the confluence of the Cambuche and Cajón Rivers (Fig. 2). The main concern is with its potential impact on spreading contamination downstream by eroded waste rock or polluted run-off from the site. The primary goal of the restoration works will be to keep the long-term leaching and off-site tailings transport to a minimum (Argentina 2003).

An environmental baseline and radiological study has recently been completed as the first step of the corresponding restoration plan. This study included periodic sampling of sediments, surface waters and groundwater for the physicochemical characterization and assessment of water quality, along with the development of a bioindicator based on benthonic diatoms communities. As a result, a large data set that imposes computational treatment has been generated. A first challenge arises from the fact that measurements are multivariate in nature and variables generally do not follow statistical normal distributions. In fact, in environmental chemistry it is very frequent to find lognormal distributions, which are particularly common for trace elements in natural waters, and also no statistical

modeled distributions at all. Furthermore, when normal distributions occur it is almost the rule that the assumed normal distribution models (e.g. a linear regression model with normal errors) describe fairly well most observations, but fail to account for some observations that follow no pattern at all. Such atypical data are called outliers, and even a single outlier can have a large distorting influence on a classical statistical method under the assumption of normality or linearity, leading to results with unacceptable low statistical efficiency (Maronna et al. 2006). Before using classical statistics methods, the data should be transformed and tested for normality. If this cannot be provided by data transformation algorithms, non-classical statistics methods, e.g. robust statistics have to be applied (Einax et al. 1997).

The robust approach to statistical modeling and data analysis aims at deriving methods that produce reliable parameter estimates and associated tests and confidence intervals, not only when the data follow a given distribution exactly, but also when this happens only approximately, and even when there is no statistical distribution at all. Robust methods fit the bulk of the data correctly: if the data contain no outliers the robust method gives approximately the same results as the classical method, while if a small proportion of outliers are present the robust method gives approximately the same results as the classical method applied only to the bulk data. As a consequence of



better fitting the bulk of the data, robust methods provide a very reliable method for detecting outliers, even in high-dimensional multivariate situations (Maronna et al. 2006).

The existence of outliers in environmental databases is a very important topic since they can have a considerable influence on the results of the analysis. There exist rigid definitions of outliers that consider them as “aberrant values that might be obtained through a wrong operation, going from a wrong setting of an instrument to transcription errors” (Massart et al. 1997), for which they have to be detected and removed. However, the presence of extreme values (provided their correct measurement) can also represent phenomena of interest. In fact, in polluted-site studies, it is common to find that the data contain some extremely high values that should not be eliminated (Berthold and Hand 2003; British Columbia Ministry of Environment (BCME) 2001).

This work presents the results of five surface water sampling campaigns developed for the Site Characterization Project “Study of the technical feasibility for the design and installation of a permeable reactive barrier” (Cicerone et al. 2009). The aim of the data analysis is to study the degree of pollution of watercourses that run across the mining complex. Sampling sites are classified according to their degree of pollution, a grouping structure of these sites according to their physicochemical characteristics is found, variables that are significant to that grouping are defined. The problem is addressed with both traditional and robust statistics techniques; additional chemometrical tools are also applied.

## Methodology

### Sampling and analytical methods

Samples were collected from 27 sites distributed throughout Los Gigantes mining complex (Fig. 5 in “Results”) during five sampling campaigns: (1) April 2010, (2) August 2010, (3) November 2010, (4) March 2011, and (5) November 2011. Sampling sites 3, 4, 5, 10, 13, 14 and 15 locate upstream from the exploitation; sites 16 and 29 are springs flowing from the still pond area; sites 31, 32 and 34 locate on springs that emerge from the tailings area, samples taken approximately 200 m (sites 31 and 32) and 400 m (site 34) from the piles; sites 62 and 63 locate on a creek, 600 and 1300 m downstream from the tailings, respectively; sites 1 and 28 are situated along the Del Cajón River within the complex; sites 8, 11, 33 and 73 are situated along the Cambuche River within the complex; sites 43, 46, 47, 75, 76, RIC and RSA locate along different water courses 1 km and further from the exploitation. Sample collection was performed according to EPA

Method 0012.1 and Method 0008.7. Temperature, pH, electrical conductivity (EC), dissolved oxygen concentration (DO) and turbidity were measured in the field with a multiparameter Sension 156 HACH®. ICP optical spectrometry was used for the determination of Mn, Be, Cu, Zn, U, As, B, Pb (EPA Method 200.7); Ca, Mg and Fe were determined by means of atomic absorption spectrometry (EPA Method 242.1), chloride and sulfate were measured with ionic chromatography (EPA Method 300.0).

### Database pretreatment

The database that arises from surface waters characterization consists of 138 objects (samples) and 17 physicochemical measured parameters. Initially, there were concentration levels below the detection limits; these were replaced by random numbers that ranged between zero and the corresponding detection limit. On the other hand the database contained a total of 45 missing data, distributed among 13 variables of 29 samples. Incomplete data matrices are a problem that is repeatedly encountered in environmental research (Junninen et al. 2004; Mwale et al. 2012). Missing values were filled up by using an artificial neural network of the class Self Organized Map (Folguera et al. 2015; García-Reiriz et al. 2011; Magallanes et al. 2011; Zupan et al. 1997).

Those variables with lognormal distribution were  $\log_{10}$  transformed. Finally, every parameter distribution was scaled from 0 to 1 using minmax transformation (Zupan and Gasteiger 1999):  $e = \frac{x - \min}{\max - \min}$ , where min and max are the minimum and maximum values of the variable, respectively,  $x$  is a value within the range of the variable and  $e$  is the scaled value, which results non dimensional.

### Data analysis

First of all, with the objective of summarizing the water quality data in a single number that expresses the level of pollution of the sampling sites, the Canadian Council of Ministers Water Quality Index (WQI) (Canadian Council of Ministers of the Environment (CCME) 2001) was calculated for each sample. This index compares the measured data with specified quality objectives, which in our case were the guidelines for the protection of aquatic life defined in the national legislation (Argentinean Legislation 1993). According to the index value obtained, water quality is ranked from excellent to poor (index 100 to 0, respectively). Additionally, the relationships among environmental variables were assessed by means of their correlation matrix.

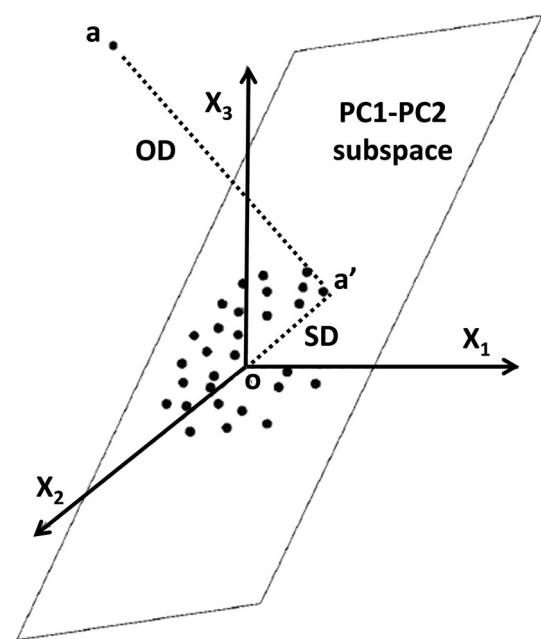
On the other hand, the multivariate structure of the data was studied by means of principal component analysis

(PCA). This is a widespread chemometrical method which tries to explain the covariance structure of data by a small number of weighed linear combinations of variables, called components (Berthold and Hand 2003; Massart et al. 1997). Because PCA is concerned with data reduction, it is widely used for the analysis of high-dimensional data which are usually present in environmental analysis. The first component corresponds to the direction in which the projected observations have the largest variance. The second component corresponds to an orthogonal direction to the first one that again maximizes the variance of the data points projected on it. Continuing in this way, the method produces a number of principal components equal to the number of variables, which tie in with the eigenvalues and eigenvectors decomposition of the empirical covariance matrix. The very first PC's explain the largest possible variation in the data and therefore those PC's account for most of the information. Unfortunately, the classical variance (which is being maximized) and the classical covariance matrix (which is being decomposed) are very sensitive to anomalous observations. The first components are often attracted towards outlying points, and may not capture the variation of the regular observations. Therefore, data reduction based on classical PCA (CPCA) becomes unreliable if outliers are present in the data (Hubert et al. 2003). The goal of robust PCA (RPCA) methods is to obtain principal components that are not heavily influenced by outliers. The method proposed by Verboven and Hubert (2005) starts from an  $n \times p$  data matrix  $X$ , where  $n$  stands for the number of objects and  $p$  for the number of variables. Robust estimates of the center  $\mu$  and the scatter matrix  $\Sigma$  of  $X$  are obtained by the minimum covariance determinant (MCD) estimator. The MCD method looks for the  $h (> n/2)$  (out of  $n$ ) observations whose classical covariance matrix has the lowest possible determinant, that is the subset of observations with the minimum spatial volume, which implies to leave out the possible outliers. The raw MCD estimate of location is then the average of these  $h$  points, whereas the raw MCD estimate of scatter is their covariance matrix.

The physicochemical database was analyzed using both CPCA and RPCA, with a twofold goal: firstly, to find linear combinations of the original variables that contain most of the information, displaying the objects structure in the new PC subspace (score graphic) and analyzing variables contributions (loadings) to that structure; secondly, to flag extreme values by use of a diagnostic plot. The latter displays on the horizontal axis the score distance within the PC1–PC2 plane,  $SD = \sqrt{t_i^T L^{-1} t_i}$ , where  $L$  is the eigenvalues' matrix and  $t_i$  is the vector' scores of the object  $i$ . On the vertical axis, the orthogonal distance to the PC1–PC2 plane  $OD = \|x_i - \hat{\mu}_i - P_{p,k} t_i\|$  where  $x_i$  is the vector of

object  $i$ ,  $\hat{\mu}_i$  is the average vector of the data matrix and  $P_{p,k}$  is the loading matrix (Fig. 4). The algorithm defines two respective cut-off values beyond which objects are considered anomalous (Verboven and Hubert 2005).

To better visualize the structure of the data that emerges from their physicochemical characteristics, samples were clustered on the bases of their PCA scores, both classical and robust. The methods used for this purpose were  $k$ -means and silhouette plotting (Berthold and Hand 2003; Massart et al. 1997).  $K$ -means is a non-hierarchical clustering algorithm that treats each observation (sample) as an object having a neighboring location in the multidimensional space, in our case defined by the principal components. Objects are classified into  $k$ -clusters obtained iteratively by use of an optimization technique that minimizes within-cluster Euclidean distance (the distance of each point to the centroid of the cluster it was assigned to) or maximizes between-cluster distance (Massart et al. 1997). The number of clusters,  $k$ , is defined by the operator. Next, the silhouette plot can be used to visualize the resulting clusters and to assess how well-separated they are. This graphic displays a measure of how close each point in one cluster is to points in the neighboring clusters, which is called the silhouette value. This measures ranges from 1, indicating points that are very distant from neighboring clusters, through 0, indicating points that are barely distinguishable from one cluster or another, to  $-1$ , indicating points that are probably assigned to the wrong



**Fig. 4** Projection of point 'a' from the space defined by the original variables  $X_1$ ,  $X_2$  and  $X_3$  on a two-dimensional PC1–PC2 subspace.  $a$ – $a'$  is the orthogonal distance (OD);  $o$ – $a'$  is the score distance (SD)

cluster or that belong to more than one cluster (Kaufman and Rousseeuw 1990).

### Results

As it was expected, surface waters physicochemical characteristics varied greatly depending on the sampling sites positions relative to the pollution sources (Fig. 5; Table 1). The unaffected surface waters (upstream from the mine) were characterized by circumneutral pH and low conductivity (less than  $60 \mu\text{S cm}^{-1}$ ), meaning that the cations and anions background levels are globally low. Samples collected at highly impacted sites, springs flowing from the still pond area (sites 16 and 29) and the tailings area (sites 31, 32 and 34), showed low pH values, high conductivity, high concentration of sulphate and significantly enrichment in Ca, Mg, Fe, and trace elements. In some cases concentrations of Mn, Be, Cu, Zn and U appear above the guidelines defined by the national legislation for the protection of aquatic life—Mn  $100 \mu\text{g l}^{-1}$ , Be  $0.05 \mu\text{g l}^{-1}$ , Cu  $2 \mu\text{g l}^{-1}$ , Zn  $30 \mu\text{g l}^{-1}$ , U  $20 \mu\text{g l}^{-1}$  (Argentinean Legislation 1993). As the distance to the pollution sources increases downstream (sites 28, 1, 8, 33, 11, 73, situated on the rivers downstream from the springs confluences) pH rises, while EC as well as concentrations of sulphate, chloride and metals decrease (probably due to dilution effects, precipitation of metals as oxides and atmospheric carbon dioxide equilibrium). Concentrations levels close to the background values and within the national guidelines are found in the sites sampled right out of the exploitation, that is, site 43 in the case of the Del Cajón River (1000 m from the exploitation), and site 46 in the case of the Cambuche River (4000 m from the exploitation).

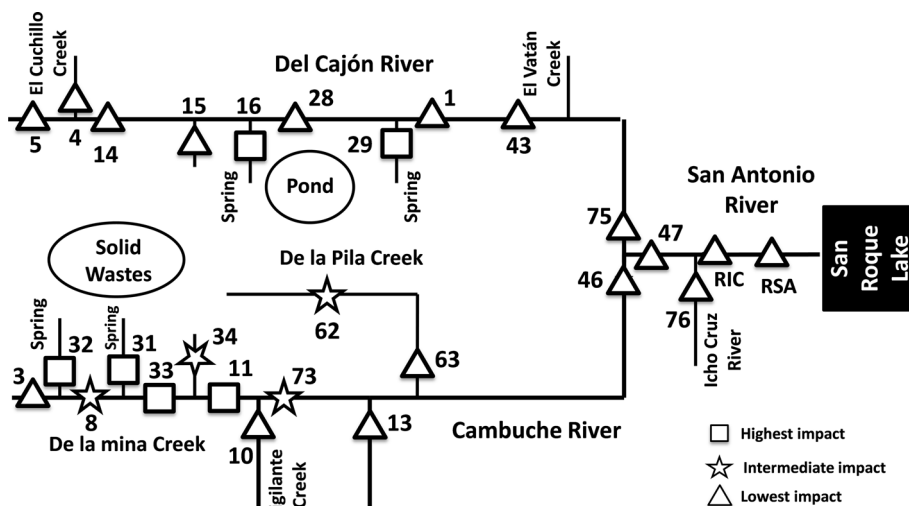
Table 2 shows the correlation matrix between physicochemical parameters. In general, cations are correlated positively among themselves and negatively with pH, due to the fact that acidic conditions solubilize minerals such as potassium feldspate (peritic microcline), sodium plagioclase, muscovite, oxides (quartz, zircon, rutile) and salts (apatite) nonspecifically. Ca and Mg exhibit the highest correlation (0.91), both cations have a large correlation coefficient with sulphate as well (0.72 and 0.70, respectively) which can be explained in terms of the high contents of Ca and Mg sulphate in the dam and the piles. These compounds arose from the neutralization of sulphuric acid with calcium hydroxides (with significant contents of magnesium).

In the case of CPCA the capture of variance for the first three components is 59, 13 and 7 %, respectively (accumulated variance 79 %). For RPCA the capture of variance is slightly higher: 62, 16 and 6 % (accumulated variance 84 %).

The RPCA analysis reveals the existence of four groups of samples (Fig. 6): group 1 consists of samples that correspond to the most impacted sites, average WQI = 58.1; group 2 comprises samples from sites with an intermediate level of impact, average WQI = 80.5; group 3 corresponds to samples from the least impacted sites, average WQI = 98.4; group 4 includes most samples of the forth campaign, except for those that were collected at the most impacted sites thus assigned to group 1. Table 3 shows the loadings of the first three PC's: cations (except As) and anions concentrations have positive loadings to PC1, while pH and DO have negative loadings.

Since mining impact implies an increase in the concentrations of ions in water bodies as well as a decrease in pH it can be concluded that sites separate along PC1

**Fig. 5** Water sampling sites in Los Gigantes. The grouping that result from PCA analysis, according to their level of impact, is mapped here



**Table 1** Physicochemical parameters of the surface water samples

Site	pH	EC ( $\mu\text{S cm}^{-1}$ )	DO (%)	Turbidity NTU	Ca (ppm)	Mg (ppm)	Cl (ppm)	Sulphate (ppm)	
1	7.1 ± 0.2	76 ± 23	78 ± 6	0.6 ± 0.2	6.1 ± 3.3	1.0 ± 0.3	1.0 ± 0.7	15.9 ± 5.1	
3	6.3 ± 0.3	60 ± 13	52 ± 13	2.1 ± 0.3	4.7 ± 2.1	0.8 ± 0.3	0.4 ± 0.2	2.7 ± 4.7	
4	7.5 ± 0.6	34 ± 9	82 ± 10	0.3 ± 0.1	1.3 ± 0.8	0.2 ± 0.2	0.4 ± 0.4	2.7 ± 2.5	
5	7.5 ± 0.4	37 ± 10	78 ± 14	0.5 ± 0.2	2.1 ± 1.1	0.4 ± 0.1	0.8 ± 0.8	2.7 ± 2.7	
8	6.1 ± 0.3	296 ± 52	76 ± 5	1.3 ± 0.5	31.3 ± 7.5	3.7 ± 0.9	0.8 ± 0.5	111.0 ± 12.5	
10	7.1 ± 0.8	48 ± 7	85 ± 4	0.4 ± 0.2	3.0 ± 1.3	0.8 ± 0.5	0.6 ± 0.6	2.7 ± 4.0	
11	3.7 ± 0.5	433 ± 172	78 ± 5	0.7 ± 0.4	35.3 ± 14.1	5.9 ± 2.8	0.9 ± 0.7	295.8 ± 309.5	
13	7.2 ± 0.3	55 ± 24	83 ± 5	0.4 ± 0.2	4.1 ± 4.3	0.7 ± 0.4	0.6 ± 0.5	6.1 ± 7.8	
14	7.3 ± 0.6	37 ± 9	79 ± 7	0.4 ± 0.1	1.0 ± 1.6	0.3 ± 0.1	0.9 ± 0.7	1.1 ± 0.3	
15	7.8 ± 0.0	40 ± 0	92 ± 0	0.7 ± 0.0	2.9 ± 0.0	0.4 ± 0.0	0.8 ± 0.0	6.0 ± 0.0	
16	4.7 ± 1.0	932 ± 349	81 ± 6	0.4 ± 0.1	109.0 ± 86.3	15.2 ± 5.1	4.2 ± 3.0	454.5 ± 190.2	
28	7.7 ± 0.5	58 ± 19	86 ± 9	0.4 ± 0.1	4.7 ± 2.6	0.7 ± 0.2	0.8 ± 0.6	5.5 ± 4.0	
29	3.7 ± 0.2	1686 ± 448	70 ± 8	0.4 ± 0.1	250.0 ± 89.4	39.5 ± 11.5	11.4 ± 3.7	855.3 ± 200.5	
31	3.3 ± 0.1	2527 ± 725	62 ± 18	1.0 ± 1.0	204.0 ± 60.9	48.3 ± 19.3	3.8 ± 2.8	1055.3 ± 527.7	
32	4.3 ± 0.2	812 ± 125	76 ± 3	0.3 ± 0.2	114.8 ± 30.8	11.4 ± 3.2	1.5 ± 0.8	387.8 ± 97.7	
33	3.6 ± 0.7	644 ± 254	75 ± 6	0.3 ± 0.1	42.8 ± 16.2	23.1 ± 38.1	1.0 ± 0.7	239.5 ± 83.0	
34	4.9 ± 0.2	279 ± 75	69 ± 9	0.8 ± 0.6	17.0 ± 10.7	1.9 ± 1.0	0.5 ± 0.5	113.5 ± 34.1	
43	7.2 ± 0.5	88 ± 21	83 ± 2	0.5 ± 0.1	6.9 ± 4.4	1.1 ± 0.4	1.2 ± 0.2	20.0 ± 5.4	
46	6.7 ± 0.5	64 ± 16	78 ± 3	0.4 ± 0.2	4.9 ± 2.6	0.7 ± 0.2	0.7 ± 0.6	7.1 ± 5.1	
47	6.9 ± 0.9	73 ± 29	84 ± 4	0.4 ± 0.1	6.5 ± 4.1	1.7 ± 1.1	0.9 ± 0.7	13.1 ± 6.7	
62	6.2 ± 0.0	390 ± 107	71 ± 14	1.3 ± 0.7	51.2 ± 38.9	4.5 ± 3.1	1.5 ± 1.1	156.0 ± 46.5	
63	7.5 ± 0.8	127 ± 27	94 ± 7	0.4 ± 0.1	13.1 ± 5.9	1.6 ± 0.5	0.5 ± 0.3	37.6 ± 12.7	
73	6.3 ± 0.5	132 ± 22	81 ± 3	3.1 ± 0.3	10.2 ± 6.3	1.7 ± 1.0	0.7 ± 0.6	40.2 ± 8.3	
75	6.8 ± 0.3	66 ± 32	78 ± 4	0.4 ± 0.1	5.7 ± 2.7	1.0 ± 0.1	1.1 ± 0.4	12.0 ± 9.1	
76	6.6 ± 0.6	55 ± 14	86 ± 5	1.0 ± 0.5	5.1 ± 2.9	1.9 ± 2.1	0.9 ± 0.7	3.2 ± 4.5	
RIC	7.4 ± 0.5	90 ± 44	87 ± 18	0.9 ± 0.2	6.2 ± 4.2	1.4 ± 1.1	2.4 ± 1.7	4.3 ± 2.7	
RSA	7.5 ± 1.1	77 ± 66	93 ± 22	2.6 ± 3.0	8.7 ± 6.4	2.2 ± 1.6	2.7 ± 1.7	4.2 ± 2.8	
11b	6.1 ± 0.5	29 ± 6	69 ± 7	0.5 ± 0.1	9.5 ± 17.5	1.4 ± 2.3	0.8 ± 1.2	1.5 ± 0.8	
Fe (ppb)	Mn (ppb)	Be (ppb)	Cu (ppb)	Zn (ppb)	U (ppb)	As (ppb)	B (ppb)	Pb (ppb)	WQI
33 ± 38	62 ± 10	0.8 ± 0.1	0.5 ± 0.2	2.4 ± 0.6	0.7 ± 0.0	0.7 ± 0.0	1.91 ± 0.32	0.02 ± 0.01	100
169 ± 116	142 ± 176	0.1 ± 0.1	0.6 ± 0.2	0.9 ± 0.1	3.8 ± 1.0	0.4 ± 0.0	8.99 ± 3.45	1.23 ± 1.73	94
19 ± 16	6 ± 1	0.1 ± 0.0	0.3 ± 0.1	0.4 ± 0.2	0.6 ± 0.2	0.5 ± 0.1	7.29 ± 4.39	0.23 ± 0.21	99
23 ± 17	7 ± 5	0.1 ± 0.0	0.2 ± 0.1	0.2 ± 0.1	0.8 ± 0.2	0.5 ± 0.3	5.13 ± 5.39	0.04 ± 0.01	98
81 ± 142	243 ± 220	3.3 ± 0.5	0.3 ± 0.1	19.3 ± 7.4	4.0 ± 2.8	0.3 ± 0.1	6.25 ± 4.27	0.33 ± 0.21	86
10 ± 12	6 ± 3	0.1 ± 0.0	0.2 ± 0.0	0.6 ± 0.2	0.8 ± 0.3	0.4 ± 0.1	7.50 ± 3.89	0.17 ± 0.14	100
23 ± 20	1110 ± 551	11.3 ± 3.6	1.8 ± 1.0	49.3 ± 19.1	21.5 ± 7.8	0.6 ± 0.2	2.24 ± 1.27	0.90 ± 0.10	65
47 ± 69	8 ± 5	0.2 ± 0.2	0.4 ± 0.0	0.4 ± 0.4	0.4 ± 0.2	0.4 ± 0.1	4.89 ± 3.13	0.11 ± 0.09	100
29.0 ± 21.5	10.0 ± 11.4	0.09 ± 0.01	0.40 ± 0.35	1.0 ± 1.4	0.8 ± 0.3	0.6 ± 0.2	8.3 ± 1.5	0.20 ± 0.32	98
42 ± 0	8 ± 0	0.0 ± 0.0	0.6 ± 0.0	1.0 ± 0.0	0.2 ± 0.0	0.6 ± 0.0	4.31 ± 0.00	0.12 ± 0.00	97
38 ± 14	1790 ± 764	12.8 ± 0.4	2.0 ± 1.4	53.0 ± 18.4	2.9 ± 3.0	0.5 ± 0.0	10.50 ± 0.71	0.05 ± 0.00	71
24 ± 27	14 ± 13	0.2 ± 0.1	0.2 ± 0.1	0.4 ± 0.4	0.7 ± 0.1	0.6 ± 0.2	1.88 ± 1.76	0.29 ± 0.22	92
67 ± 48	4387 ± 1461	36.0 ± 6.7	2.4 ± 0.5	61.0 ± 3.4	7.8 ± 1.7	3.2 ± 5.1	5.53 ± 4.19	0.95 ± 0.05	50
170 ± 160	11,549 ± 13,013	87.8 ± 11.1	13.5 ± 1.3	421.0 ± 29.2	128.5 ± 47.8	0.8 ± 0.2	6.12 ± 5.28	2.99 ± 0.01	38
34 ± 31	1621 ± 1204	16.5 ± 3.9	1.0 ± 0.8	83.3 ± 11.8	35.8 ± 12.5	0.4 ± 0.1	6.42 ± 6.13	0.41 ± 0.02	61
40 ± 43	1606 ± 1581	17.8 ± 5.3	2.8 ± 1.0	83.3 ± 22.2	33.3 ± 11.2	0.5 ± 0.1	7.38 ± 4.14	1.50 ± 0.50	64
22 ± 19	489 ± 381	5.4 ± 3.8	1.9 ± 2.1	23.5 ± 16.6	5.0 ± 6.0	0.3 ± 0.1	5.03 ± 3.37	0.10 ± 0.10	73
19.0 ± 10.4	41.8 ± 22.4	0.60 ± 0.37	0.40 ± 0.17	2.0 ± 1.8	0.6 ± 0.1	0.4 ± 0.2	4.3 ± 3.6	0.10 ± 0.08	90



**Table 1** continued

Fe (ppb)	Mn (ppb)	Be (ppb)	Cu (ppb)	Zn (ppb)	U (ppb)	As (ppb)	B (ppb)	Pb (ppb)	WQI
32 ± 24	5 ± 5	0.2 ± 0.1	0.5 ± 0.2	1.1 ± 1.3	0.9 ± 0.8	0.3 ± 0.3	5.05 ± 1.47	0.12 ± 0.08	91
24 ± 18	7 ± 2	0.3 ± 0.1	0.4 ± 0.1	0.6 ± 0.3	1.0 ± 0.0	0.4 ± 0.1	5.13 ± 3.04	0.41 ± 0.51	92
124 ± 93	93 ± 136	0.5 ± 0.2	0.3 ± 0.1	2.4 ± 2.4	1.3 ± 0.6	0.5 ± 0.0	8.41 ± 3.40	0.09 ± 0.02	91
27 ± 25	6 ± 3	0.1 ± 0.1	0.5 ± 0.2	1.0 ± 0.7	1.0 ± 0.0	0.6 ± 0.2	0.96 ± 0.92	0.53 ± 0.46	93
14 ± 16	283 ± 198	1.1 ± 0.6	0.4 ± 0.1	9.8 ± 3.6	2.4 ± 0.5	0.3 ± 0.1	7.54 ± 3.98	0.46 ± 0.33	86
10.0 ± 11.6	10.0 ± 5.7	0.30 ± 0.15	0.30 ± 0.30	1.0 ± 0.7	0.4 ± 0.2	0.6 ± 0.2	11.0 ± 7.2	0.10 ± 0.19	90
21 ± 24	11 ± 6	0.1 ± 0.1	0.6 ± 0.3	0.8 ± 0.2	0.6 ± 0.3	0.8 ± 0.2	10.55 ± 1.22	0.30 ± 0.30	99
14 ± 22	20 ± 23	0.1 ± 0.0	0.7 ± 0.1	0.5 ± 0.3	0.8 ± 0.2	0.8 ± 0.1	10.53 ± 3.37	0.15 ± 0.14	98
17 ± 24	11 ± 7	0.1 ± 0.0	0.7 ± 0.1	0.5 ± 0.1	0.9 ± 0.2	0.9 ± 0.0	15.50 ± 2.38	0.43 ± 0.31	99
44.0 ± 10.4	6.8 ± 4.2	0.04 ± 0.03	0.30 ± 0.01	1.9 ± 1.5	0.2 ± 0.2	0.6 ± 0.2	3.4 ± 2.3	0.55 ± 0.59	89

For every site the median and MADN calculated on the basis of the five campaigns are shown

according to their degree of pollution. Thus, sites with a minimum impact show low scores on PC1, while highly positive scores along PC1 are indicative of mining pollution. Group 4 appears as a whole different group, separated from group 3 along PC2. The former group comprises samples taking during the fourth campaign, which was conducted during a period of unusually high rainfall. The exceptional weather conditions might account for the differences among the variables that make the greatest contributions to PC2, since these unusual rainfalls might have occasioned a decrease in some compounds concentration levels as a result of dilution as well as a more significant increase in other ions due to mobilization from mining wastes.

As regards CPCA, samples scores lead to the similar clustering as RPCA. However, there exists a difference in the parameters loadings on the PC's (Table 3). PC1 and PC2 are mostly characterized by the same variables in both CPCA and RPCA, but their order is different. In CPCA the loading for pH on PC1 is in fourth place, while as a result of RPCA this parameter holds the seventh place. Arsenic and turbidity make a similar contribution to PC2 in CPCA, while in the robust version As loading is one order larger compared to that of turbidity. Cations loadings order for the first two principal components differs as well. In the case of PC3 parameters make their contributions in the same order for both PCA methods.

The partitioning technique applied to PCA scores (Fig. 7) confirms PCA results: sites are grouped in four clusters: silhouette 1 comprises samples from sites with an intermediate level of impact (group 2 in Fig. 6), silhouette 2 includes most samples of the fourth campaign, except for those collected at the most impacted sites (group 4 in Fig. 6), silhouette 3 corresponds to the most impacted sites

(group 1 in Fig. 6) and silhouette 4 comprises the least impacted sites (group 3 in Fig. 6). Both CPCA and RPCA throw identical results as regards sites assignment to each silhouette although there exists a slight difference in the sites silhouette values, these being higher in the robust version.

The diagnostic plot obtained from RPCA (Fig. 8) spots as outliers two types of samples: first, those taken at highly impacted sites (mean WQI = 52); a second group of samples consisting on samples with WQI = 100 and samples with extreme values in one or two parameters but below the guidelines in the rest. Table 4 shows the average values of the physicochemical parameters measured in the bulk of the data (data without outliers), and, separately, in the two groups of outliers. The latter locate clearly far away from the majority of the data, i.e. their physicochemical parameters means are different. If the corresponding classical statistics were calculated including them, the results would be inaccurate. This is the reason why the classical method (Fig. 9) is less efficient in the detection of extreme values: the number of sites labeled as such is lower than those detected with the robust technique.

In Fig. 5 the groups of sites obtained with PCA are mapped on the mining area. This image reveals the relationship between the location of the sites relative to the waste facilities and their level of mining impact. The group of the most impacted sites comprises the area close to the piles of solid wastes and the still ponds; the group of sites with an intermediate impact locate either further from the contamination sources or enjoy the dilution effect of clean incoming streams; the group of least impacted sites correspond to two different zones: the area upstream from the piles and ponds, and the area downstream of these where the rivers recover naturally.

**Table 2** Correlation matrix of the physicochemical variables

	pH	CE	DO	Turbidity	Ca	Mg	Cl	Sulphate	Fe	Mn	Be	Cu	Zn	U	As	B	Pb
pH	1.00	-0.75	0.21	0.10	-0.55	-0.62	-0.22	-0.60	-0.22	-0.56	-0.72	-0.56	-0.78	-0.74	-0.03	0.03	-0.32
CE		1.00	-0.11	-0.06	0.77	0.77	0.44	0.79	0.16	0.63	0.77	0.50	0.74	0.70	-0.05	-0.06	0.23
DO			1.00	-0.06	-0.10	-0.11	-0.10	-0.03	-0.15	-0.07	-0.06	-0.06	-0.09	-0.05	0.16	-0.01	-0.05
Turbidity				1.00	0.00	-0.05	0.01	-0.13	0.06	-0.05	-0.10	0.12	0.02	-0.04	-0.04	0.06	-0.12
Ca					1.00	0.91	0.57	0.72	0.24	0.71	0.69	0.43	0.65	0.58	-0.24	-0.02	0.06
Mg						1.00	0.53	0.70	0.23	0.71	0.69	0.46	0.66	0.60	-0.18	0.04	0.18
Cl							1.00	0.45	0.04	0.39	0.29	0.31	0.30	0.21	-0.43	0.15	-0.36
Sulphate								1.00	0.05	0.59	0.72	0.43	0.68	0.59	-0.11	-0.06	0.10
Fe									1.00	0.39	0.15	0.30	0.29	0.26	0.15	0.07	0.29
Mn										1.00	0.66	0.43	0.65	0.63	-0.07	0.10	0.16
Be											1.00	0.56	0.86	0.78	-0.04	-0.02	0.23
Cu												1.00	0.75	0.60	0.06	0.12	0.18
Zn													1.00	0.78	-0.01	0.09	0.24
U														1.00	0.04	0.07	0.31
As															1.00	-0.04	0.60
B																1.00	-0.04
Pb																	1.00

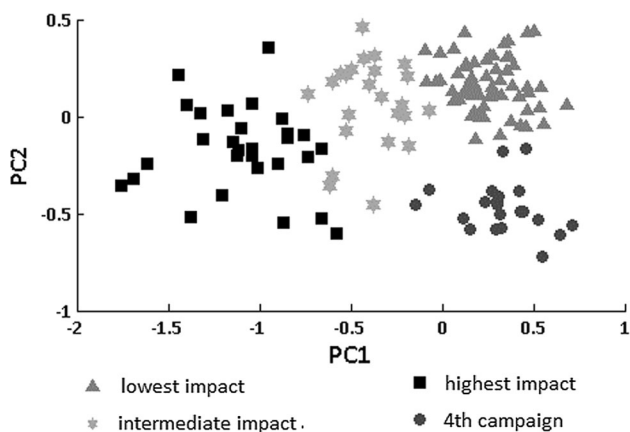


Fig. 6 Robust PCA scores of water samples

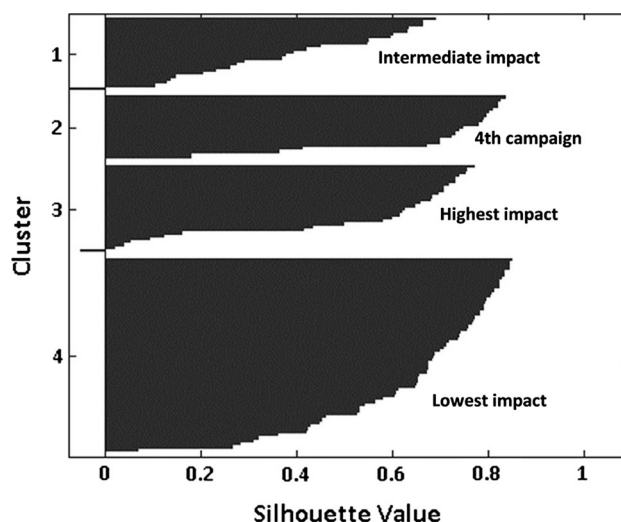


Fig. 7 Silhouette plot. Each silhouette corresponds to a group of sites with a specific impact level

**Conclusions**

Restoration of abandoned uranium mines in Argentina has been taking place since 2005. Environmental monitoring plans imply handling large multivariate, non-normal distributed, outliers carrying data sets. This poses a computational challenge that must be addressed by use of proper chemometrical methods.

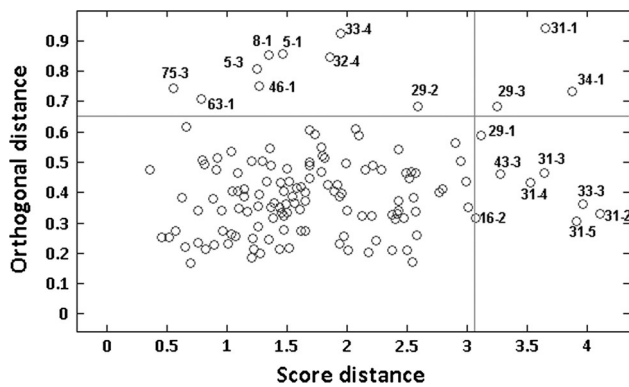
The analysis of surface water samples from the mining and milling complex Los Gigantes shows that there still exist pollution sources that need to be correctly addressed. Surface streams close to the tailings and the pond exhibit lower pH and higher concentration of anions and metals compared to upstream watercourses. All physicochemical

parameters recover gradually downstream from the complex, reaching levels close to background and complying with provincial and national regulations at the sampling sites located outside of the mining complex proving thus that pollution is locally constrained.

CPCA grouped objects in four clusters according to their level of mining pollution. The same result was obtained with RPCA, and confirmed by *k*-means clustering and silhouette plotting. In both CPCA and RPCA the first PC, separated samples according to their general level of

**Table 3** Parameters loadings to the first three PC's in RPCA and CPCA

	RPCA loadings			CPCA loadings			
	PC1	PC2	PC3	PC1	PC2	PC3	
Be	0.381	0.053	-0.222	CE	0.375	-0.039	-0.209
CE	0.369	-0.043	-0.111	Be	0.374	0.036	-0.227
Mn	0.352	0.021	0.411	Mn	0.338	-0.056	0.436
Mg	0.339	-0.068	0.043	pH	-0.330	-0.167	0.222
Ca	0.324	-0.156	0.116	Zn	0.328	0.079	-0.053
Zn	0.317	0.093	-0.123	Mg	0.325	-0.148	0.120
pH	-0.289	-0.179	0.327	Ca	0.300	-0.218	0.135
Sulphate	0.264	-0.128	-0.099	Sulphate	0.252	-0.112	-0.200
U	0.238	0.108	-0.065	U	0.250	0.118	-0.058
Cu	0.166	0.065	0.036	Cu	0.174	0.066	0.064
Chloride	0.160	-0.363	0.214	Chloride	0.123	-0.374	0.113
Fe	0.093	0.271	0.725	Pb	0.102	0.693	0.012
As	-0.037	0.404	0.054	Fe	0.098	0.221	0.687
Pb	0.035	0.722	-0.069	DO	-0.018	0.003	-0.036
DO	-0.023	-0.008	0.013	Turbidity	-0.018	-0.061	0.221
Turbidity	0.005	-0.046	0.142	As	-0.017	0.432	-0.009
B	0.000	-0.025	0.131	B	0.006	-0.032	0.220

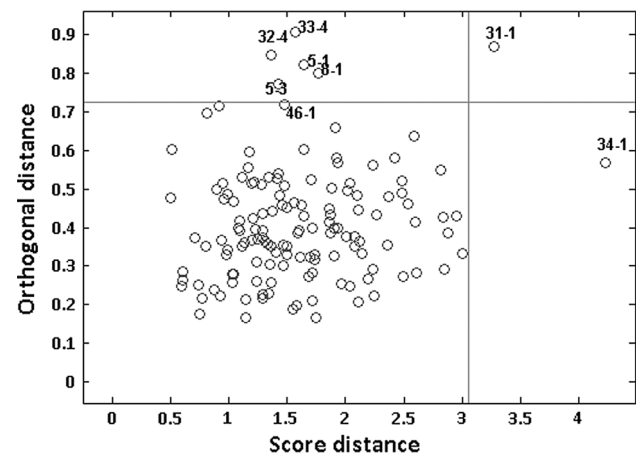


**Fig. 8** Robust diagnostic plot. Objects beyond the *horizontal* and/or *vertical* lines are considered *outliers*. *Numbers before hyphens* indicate sampling sites, *numbers after hyphens* indicate sampling campaign

**Table 4** Means of physicochemical parameters of the bulk of the data, all the outliers, outliers of most impacted sites (A) and outliers of least impacted sites (B)

	Bulk	Outliers	Outliers A	Outliers B
pH	6.32	5.02	3.83	7.05
CE	219.84	1050.32	1612.21	87.09
DO	78.56	76.46	73.28	82.19
Turbidity	0.78	2.29	0.53	4.91
Ca	22.92	107.80	163.91	11.61
Mg	3.96	26.33	40.82	1.49
Chloride	1.55	4.46	6.31	1.29
Sulphate	89.04	532.08	833.67	15.09
Fe	35.65	151.76	178.75	97.78
Mn	524.58	7550.93	8937.39	149.62
Be	3.71	30.70	47.04	0.76
Cu	0.95	7.91	7.75	8.14
Zn	18.60	144.78	229.31	11.94
U	5.32	75.93	119.29	1.59
As	1.42	1.28	1.72	0.51
B	8.17	8.07	8.81	6.59
Pb	2.25	2.14	3.19	0.34

pollution. Along the second PC it was possible to identify a complete sampling campaign performed under exceptional weather conditions. The variables loadings order and magnitudes differed if they were calculated with CPCA or RPCA. Mapping the four groups allowed seeing the spatial distribution of zones with a different degree of mining impact. Samples with extreme values were detected by means of a diagnostic plot in the classical and robust versions; in the robust plot a larger number of outliers were spotted. These results show that the classical definition of outliers, that considers them as aberrant values that have to



**Fig. 9** Classic diagnostic plot. Objects beyond the *horizontal* and/or *vertical* lines are considered *outliers*. *Numbers before hyphens* indicate sampling sites, *numbers after hyphens* indicate sampling campaign

be removed, might not be applicable to environmental problems because they can actually imply an important piece of information.

So far it has been shown that in certain cases CPCA and RPCA are likely to throw dissimilar results, which in the case of environmental management might lead to different decision making. Assessing which method is more efficient for the purposes of the study becomes then necessary. In our case, identification of parameters that are able to separate samples according to their levels of impact is of high importance for the monitoring plan; so is the recognition of pristine and extremely polluted sites. It has been pointed out that the presence of extreme values in a population attracts statistic parameters that are calculated upon normality assumption towards those outliers, thus leading to unreliable results. Since ROBPCA does not include extreme values in the variance and covariance estimation, their distorting influence is averted, thus this method is preferred over CPCA.

After an exploratory campaign, robust statistics proved to be a reliable tool for the detection of sampling sites of interest for the monitoring plan; highly polluted sites were spotted as well as pristine sites. Identification of the former is crucial to environmental management since it allows planning the resources assignment where they are needed, including the level or goals to be achieved by the restitution/remediation/restoration actions. Recognition of clean sites might be useful for the definition of quality targets in a restoration process. Finally, robust techniques are suitable for finding the most contributing physicochemical parameters, which are also a key step for directing reasonable efforts in the monitoring plan and for identification of potential sources of contamination.

**Acknowledgments** This study was made possible with the financial support provided by the Comisión Nacional de Energía Atómica through the project PP-SNA\_GQ\_GQA-DSA-01. The authors would like to thank the coworkers of the División Química del Agua y del Suelo and División Química Analítica of the Centro Atómico Constituyentes.

## References

- Abdelouas A (2006) Uranium mill tailings: geochemistry, mineralogy, and environmental impact. *Elements* 2(6):335–341
- Argentina (2003) Argentina Mining and Milling Uranium Wastes Sites Restoration (PRAMU) environmental assessment. Joint Convention on the Safety of Spent Fuel Management and on the safety of radioactive waste management. National report
- Argentinean Legislation (1993) Regulation decree 831/93 of Hazardous Wastes National Law 24.051. In: Official state bulletin 27.630. Buenos Aires
- Berthold MR, Hand DJ (2003) *Intelligent data analysis: an introduction*, 2nd edn. Springer, Berlin
- British Columbia Ministry of Environment (BCME) (2001) Outliers, a Guide for data analysts and interpreters on how to evaluate unexpected high values. Contaminated sites statistical applications guidance document 12-8 (Internet). [http://www.env.gov.bc.ca/epd/remediation/guidance/technical/pdf/12/gd08\\_all.pdf](http://www.env.gov.bc.ca/epd/remediation/guidance/technical/pdf/12/gd08_all.pdf). Accessed 2 Nov 2014
- Canadian Council of Ministers of the Environment (CCME) (2001) CCME Water Quality Index 1.0, user's manual. In: Canadian Council of Ministers of the Environment (eds) Canadian environmental quality guidelines for the protection of aquatic life, Winnipeg
- Cicerone D, Gómez del Río J, Sánchez Proaño P, Grande Cobián D (2009) Technical feasibility study for the design and installation of a permeable reactive barrier. Comisión Nacional de Energía Atómica internal report: PP-SNAGQ\_DAS-01
- División Química del Agua y del Suelo (DQAS) (2012) Comisión Nacional de Energía Atómica internal report IN-GQ\_DQAS-00
- Einax JW, Zwanziger HW and Geiß S (1997) *Chemometrics in environmental analysis*. VCH Verlagsgesellschaft mbH Weinheim
- Folguera L, Zupan J, Cicerone DS, Magallanes JF (2015) Self-organizing maps for imputation of missing data in incomplete data matrices. *Chemometr Intell Lab* 143:146–151
- García-Reiriz A, Magallanes JF, Zupan J, Líberman S (2011) Artificial neural networks to evaluate the boron concentration decreasing profile in blood-BPA samples of BNCT patients. *Appl Radiat Isotopes* 69:1793–1795
- Hubert M, Rousseeuw PJ, Branen KV (2003) ROBPCA: a new approach to robust principal component analysis. *Technometrics* 47(1):64–79
- Instituto Nacional del Agua (INA) (2001) Investigación Hidrogeológica en el Área Los Gigantes. Subsecretaría de Recursos Hídricos de la Nación, Ministerio de Panificación Federal, Inversión Pública y Servicios de la República Argentina. Technical report IT-201
- International Atomic Energy Agency (IAEA) (2008) Estimation of global inventories of radioactive waste and other radioactive materials. Technical document IAEA-TECDOC-1591
- Junninen H, Niska H, Tuppurainen K, Ruuskanen J, Kolehmainen M (2004) Methods for imputation of missing values in air quality data sets. *Atmos Environ* 38:2895–2907
- Kaufman L, Rousseeuw PJ (1990) *Finding groups in data: an introduction to cluster analysis*. Wiley, New York
- Magallanes JF, García Reiriz AG, Líberman S, Zupan J (2011) Kohonen classification applying 'missing variables' criterion to evaluate the pboronophenylalanine human-body-concentration decreasing profile of boron neutron capture therapy patients. *J Chemometrics* 25(6):340–347
- Maronna RA, Douglas DM, Yohai VJ (2006) *Robust statistics, theory and methods*. Wiley, England
- Massart DL, Vandeginste BGM, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J (1997) *Handbook of chemometrics and qualimetrics*. Elsevier, Amsterdam
- Mwale F, Adeloye A, Rustum R (2012) Infilling of missing rainfall and streamflow data in the Shire river basin, Malawi. A self organizing map approach. *Phys Chem Earth* 50–52:34–43
- Nuclear Energy Agency (NEA) (2014) Managing environmental and health impacts of uranium mining. Organization for Economic Cooperation and Development. NEA 7062, Paris
- The World Bank (2014) Restructuring paper on a proposed project restructuring of Argentina mining environmental restoration project (P110462). Report 79771-AR
- Uranium Mining Environmental Restoration Project (PRAMU) (2005) Evaluación Ambiental, document Marco. Comisión Nacional de Energía Atómica. Internal report INF-04-010
- Verboven S, Hubert M (2005) LIBRA: a MATLAB library for robust analysis. *Chemometr Intell Lab* 75:127–136
- Zupan J, Gasteiger J (1999) *Neural networks in chemistry and drug design*. Wiley-VCH, Weinheim
- Zupan J, Novič M, Ruisánchez I (1997) Kohonen and counterpropagation artificial neural networks in analytical chemistry. *Chemometr Intell Lab* 38:1–2