Environ Earth Sci (2016) 75:335
DOI 10.1007/s12665-015-5142-8

ORIGINAL ARTICLE

Joint interpretation of the hydrochemistry of two neighbouring basins by N-way multivariate methods

Rosario Soledad Barranquero[1,4] · Marcelo Raúl Varni[2] · Rafael Pardo[3] · Marisol Vega[3] · María Emilia Zabala[2,4] · Víctor Alejandro Ruiz de Galarreta[1]

Received: 20 July 2015 / Accepted: 5 October 2015 / Published online: 11 February 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract The objective of this work is to compare the chemical composition and the spatial and temporal variabilities of groundwater in two basins, the Langueyú and Del Azul creeks basins, located in the Pampean plain, Buenos Aires province (Argentina). The Pampean plain is the most productive region in Argentina, agriculture and livestock being the main economic activities. Groundwater is the principal water resource in the region, with a strong and growing demand for human supply and for agriculture and industrial activities. Several sampling campaigns were carried out on shallow wells of the two studied basins along a period of 3 years (2010–2013) to identify seasonal variations. Electrical conductivity, pH, bicarbonate, chloride, sulphate, nitrate, calcium, magnesium, sodium and potassium were determined following standard methods. For hydrochemical interpretation, descriptive statistical analyses, matrix augmentation principal component analysis, MA-PCA, and multidimensional principal component analysis, N-PCA (Parafac and Tucker3 models), were applied to the hydrochemical datasets from both basins. Three main hydrochemical processes have been identified in both basins: saline enrichment in the groundwater flow direction caused by dissolution of carbonates; exchange of calcium and magnesium by sodium in the same direction, and located areas of nitrate pollution. The paper shows that N-PCA is a good tool to deepen in the understanding of the hydrochemical features of groundwater from two neighbour basins, simplifying the analysis of large amounts of data, as well as establishing relations between the compared basins. Therefore the work is considered an interesting contribution to the study of groundwater resources with a regional scope. This knowledge is essential in basins with high socio-economic interests it causes a direct impact on resources management.

Keywords Groundwater · Pampean plain · N-way principal component analysis · Multivariate statistics · Hydrochemistry · Spatial and temporal variabilities · Regional scope

Introduction

The scientific research related to water resources has grown without major impact upon decisions regarding their sustainable use and management (Millennium Ecosystem Assessment 2005), mainly due to a lack of regional action lines for a sustainable management. For the development of these lines, isolated studies of basins in a region are insufficient and a comparison between neighbouring basins is necessary for a regional scope.

In many areas of Argentina, including the Pampean plain (Fig. 1), groundwater resources are severely exploited and used for economic activities and human consumption. In this sense, knowledge about the

---

✉ Rosario Soledad Barranquero
rosariobarranquero@yahoo.com.ar;
rosario0405@gmail.com

[1] Centro de Investigaciones y Estudios Ambientales, Facultad de Ciencias Humanas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN), Paraje Arroyo Seco s/n, Campus Universitario Tandil, 7000 Tandil, Argentina

[2] Instituto de Hidrología de Llanuras 'Dr Eduardo J. Usunoff', UNICEN, República de Italia 780, B7300 Azul, Argentina

[3] Departamento de Química Analítica, Facultad de Ciencias, Universidad de Valladolid, Campus Miguel Delibes, 47011 Valladolid, España

[4] Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina

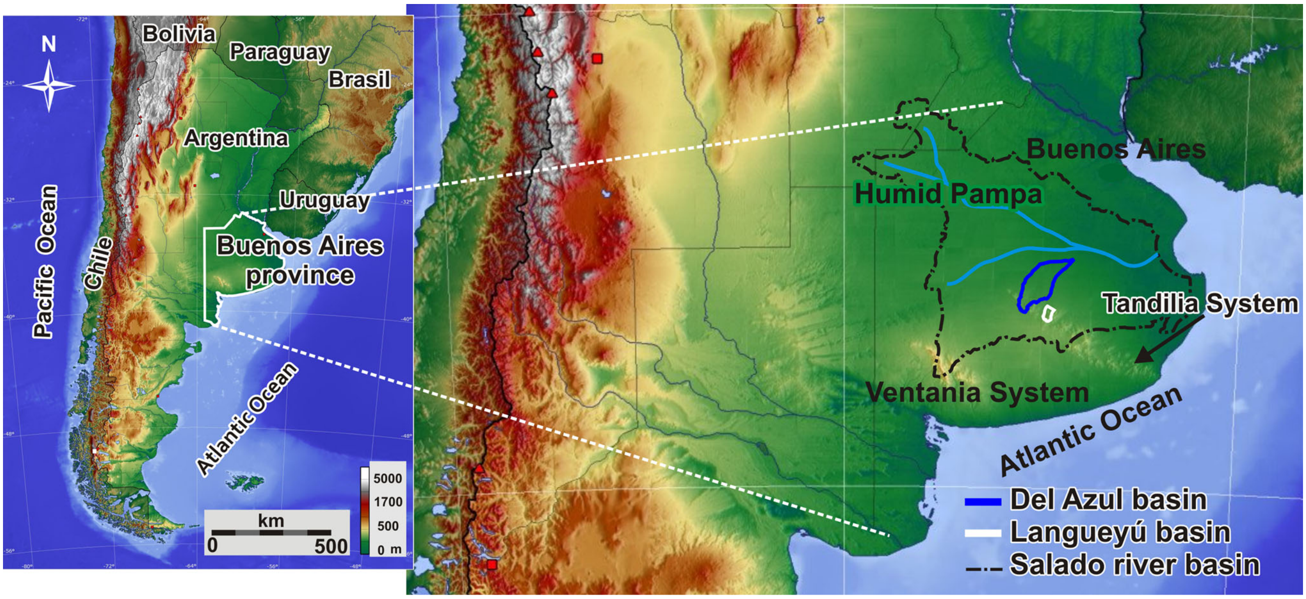

Fig. 1 Location of Argentina Humid Pampa, Tandilia System and representation of Langueyú and Del Azul creek basins

hydrochemical features and spatial and temporal variabilities of groundwater is essential for an adequate water management.

The Pampean plain is the main area of agricultural production and the strongest economic region of Argentina. In this area emerges the Tandilia System, an orographic belt located in the central part of Buenos Aires province (Fig. 1). The Tandilia System has been divided into subsystems designated by the name of the closest city or district, without correspondence to hydrogeological characteristics. Studies at different spatial scales carried out by several authors (Hernández et al. 2002; Ruiz de Galarreta 2006; Barranquero 2009; Varni 2013), have identified common features between basins with headwaters in the Northern slope of the Tandilia System. Therefore, it is important the hydrochemical study of these basins with a regional scope.

Groundwaters of two neighbouring basins, Langueyú and Del Azul creek basins (Fig. 1), have been analyzed and compared. Both creeks are tributaries of the Salado River basin and were selected for this study as they are approximately 80 km away, and a similar degree of knowledge is available about the groundwater behavior and uses in both areas.

In environmental and hydrological studies, the large number of samples and variables usually involved produce large data matrices, thus hindering the interpretation of the results. Multivariate statistical analysis, especially classical two-way principal component analysis, PCA, is frequently used to reduce the dimensionality of the dataset and to facilitate the visualization and interpretation of the underlying information by uncovering relationships between samples and variables (Vega et al. 1998; Simeonov et al. 2003; Singh et al. 2004). To assess the spatial and temporal variability of groundwater hydrochemistry, some studies have applied multivariate statistical techniques such as cluster analysis, PCA or discriminant analysis (Mohammadi 2009; Nosrati and Van Den Eeckhaut 2012; Xu et al. 2012; Singaraja et al. 2013; Jayalakshmi et al. 2014; Pazand and Pazand 2014; Thomas et al. 2015). In other cases, a combination of analytical techniques with simulation modeling has been used (Dai et al. 2012, 2014).

Because the limitations of classical (two-way) PCA to interpret multidimensional datasets in which, besides sampling points and parameters analyzed, a third dimension such as variations with time, must be considered, more recently multi-mode PCA methods (N-PCA) have been developed. These methods have made possible including the totality of information in a multi-dimensional structure-sampling sites, parameters and time (Faber et al. 2003; Smilde et al. 2004).

N-PCA methods have been used to interpret datasets in different fields such as spectroscopy (Bro 1997) or physical and environmental studies (Pere-Trepat et al. 2007; Gussani et al. 2008; Pardo et al. 2008; Reis et al. 2010). Only a few number of studies have applied N-PCA or other simulation methods to analyze multidimensional groundwater data (Dai and Samper, 2006; Singh et al. 2007; Galego-Fernandes et al. 2008; Ramesh-Kumar and Riyazuddin 2008; Bacon et al. 2014; Barranquero et al. 2014). The last-named work has been done in the study area, but restricted to the Langueyú creek basin. In addition, none of the detailed background works have applied the N-PCA methods to compare two basins.

This work proposes the evaluation of hydrochemical features and spatial and temporal variabilities of groundwater from the Langueyú and Del Azul creek basins. The most important contributions are to simplify the analysis of the results using N-PCA methods, and to show the comparability of two basins with different monitoring networks. This new methodological approach to the investigation of groundwater resources can complement classical hydrochemical methods.

The comparison of two basins representative of a region with high socio-economic interest represents a contribution to the regional scope, necessary to produce an impact on water resources management.

Study area

Langueyú and Del Azul creek basins are located in Tandil and Azul districts, with 123,871 and 65,280 inhabitants, respectively (INDEC 2010). As the main cities of both districts are settled within the basin, there is a strong demand of groundwater for human supply and for a variety of socio-economic activities, mainly agriculture.

Both basins have a subhumid-humid mesothermal climatic regime, with little or no water deficiency (C2B’2r), according to Thornthwaite and Mather (1957). The long-term balances (1900–2000 in Tandil, 1901–2013 in Azul) showed that excesses occur in the most of the year with a total of 144 mm in Tandil and 177 mm in Azul. Deficits occur only occasionally in summer months.

During the years monitored in this work (2010, 2011 and 2012), the occurrence of water excesses and deficits were much more irregular than in the long-term balance. Although amounts were different in each basin, water availability patterns were coincident the only noticeable difference being that Langueyú creek basin in the 2010–2012 period was slightly more humid than Del Azul, with more months with positive differences in comparison to the long-term balance.

From the geomorphological point of view, both basins have three sectors (Varni and Usunoff 1999; Ruiz de Galarreta and Banda Noriega 2005; Zárate et al. 2010):

- *Hills* with a strong structural control of the hydrological dynamics that design integrated and dendritic drainage. Slopes have average values of 6 % in Langueyú creek basin and of 5 % in Del Azul (Varni et al. 2013).
- *Foothills* a well-defined and distributary design drainage network being the flow divergent according to morphology. Slopes are smoother than in hills with values of 0.3 % for both basins.
- *Plains* drainage is poorly defined and integrated, with narrow channels and temporary courses that are often lost in slight depressions. Water movement has a predominance of vertical component over horizontal flows. Slopes have average values of 0.1 % in Langueyú and between 0.05 and 0.1 % in Del Azul (Varni et al. 2013).

Langueyú and Del Azul basins stretch over an area of approximately 600 and 6000 km^2, respectively, excluding the area of the regional groundwater discharge occurring in the Salado river, located Northeast. The difference in extension makes flow paths in Del Azul basin longer, i.e., elongated groundwater residence time, determining a greater representation of plains sector in this basin. Another difference is that Tandil city is located in a hilly area, unlike Azul city that is located in a transitional zone between foothills and plain. Even considering these differences, both areas have many common features that make them comparable, as location of its headwaters in the North slope of the Tandilia System, and direction of surface runoff and groundwater flow towards the Northeast with discharge in the Salado river.

The formation in which the headwaters are settled (Tandilia) consists of a crystalline basement of Precambrian age covered by Postpampean and Pampean Cenozoic sediments (Teruggi and Kilmurray 1975).

The crystalline basement is mostly made up of a variety of rocks of granitic aspect.

The Cenozoic sedimentary cover is mainly composed by Pampean Formation that forms a continuous mantle across the Pampean plain. This formation is composed by loessial sediments of silty composition, with large proportion of volcanic glass and subordinate fractions of sand and clay. Within these sediments, calcretes with varying thickness and distribution are common. Postpampean Formation is represented in smaller proportions in the sedimentary cover composed by more superficial deposits usually located along the river valleys.

In both basins, groundwater resources are located on two hydrogeological environments with different behavior in terms of groundwater circulation and admission.

In the hills sector, bedrock is in surface or very shallow and is fissured, with secondary porosity and permeability. Here the sedimentary cover is thin, and therefore the aquifers have low performance of exploitation.

To the Northeast, in the direction of groundwater flow, the basement deepens, thus lowering the limit of the aquifer in sectors of foothills and plains. Also, the thickness of the sedimentary cover increases. This cover corresponds to porous medium, which is a multi-unit phreatic aquifer, since it has discontinuities in depth, although these are only of local character (Sala et al. 1981, 1987).

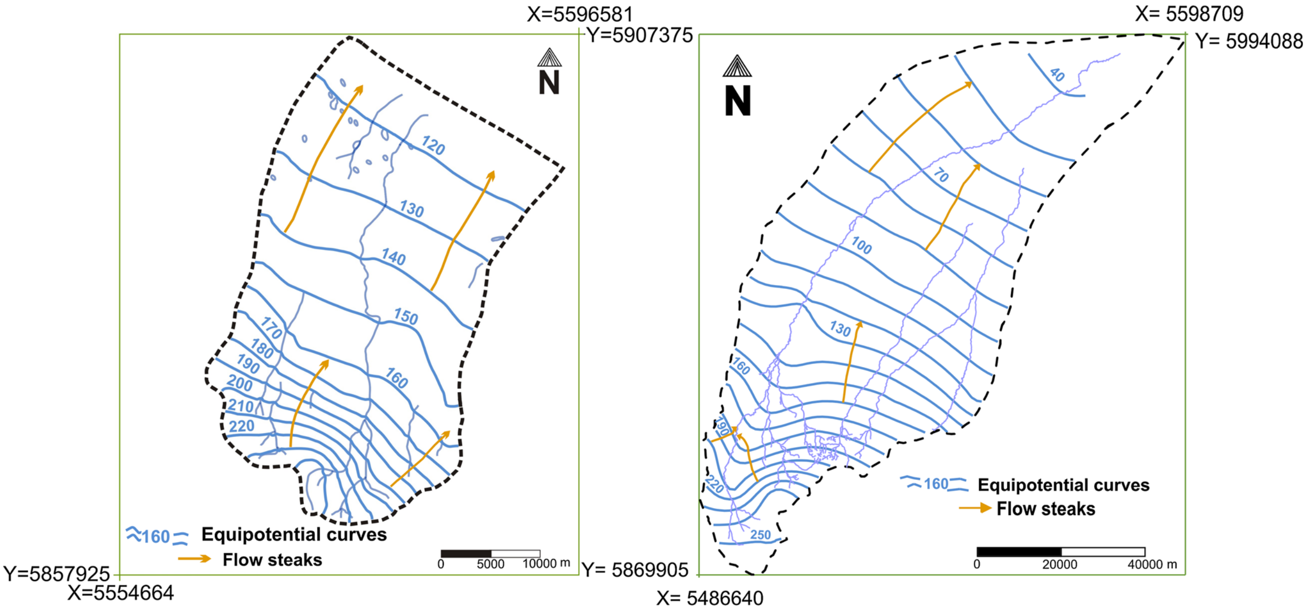

Fig. 2 Equipotential curves map for October 2010 in Langueyú creek basin (*left*) and for August 2010 in Del Azul creek basin (*right*)

Saturated thickness is larger in Del Azul creek basin than in Langueyú because the basement deepens up to 800 m in Del Azul (Zabala 2013) and only to 220 m in Langueyú (Barranquero 2014). It is important to note that according to regional studies, deepening of basement in both basins has similar characteristics (Yrigoyen 1975); the difference is likely due to the distinct length from headwaters to the Northern limit considered for each basin in this study: 143 km in Del Azul basin and only 39 km in Langueyú area. In fact, the basement depth at 39 km from the headwaters in Del Azul is 250 m.

Groundwaters flow has direction from Southwest to Northeast in both basins (Fig. 2). In addition, the water table is generally near the surface (between 2 and 5 m) except in the hills sector, where it deepens.

Materials and methods

Sampling and analytical procedures

Figure 3 shows groundwater sampling network grids for the two studied areas. These monitoring networks were defined in different times by different research centers, but were considered adequate for comparison because they were planned with similar objectives. In both basins groundwater monitoring network is representative of the upper section of the aquifer.

In Del Azul basin, 24 wells were sampled along 5 sampling campaigns (August 2010, June 2011, January 2012, January 2013 and August 2013). These wells belong to the Instituto de Hidrología de Llanuras "Dr Eduardo Usunoff" (IHLLA) monitoring network. Depths of wells vary from 5 to 10 m and the samples were taken by pumping after purge. Electrical conductivity (E.C.), pH and water temperature were measured on-site through continuous flow cells with a multiparametric probe. Alkalinity was also measured on-site by titration. The rest of parameters (chloride, sulphate, nitrate, calcium, magnesium, sodium and potassium) were determined at IHLLA laboratory using standard methods of analysis (APHA 2005).

In Langueyú basin, 21 wells were sampled along 6 sampling campaigns (October 2010, February 2011, June 2011, October 2011, February 2012 and June 2012). In this case, the monitoring network consisted mostly of shallow private domestic wells. Samples were taken as above, measuring on-site E.C., pH and water temperature with a multiparametric probe and alkalinity by titration method. The same chemical parameters than in Del Azul creek basin were analyzed, following standard methods (APHA 2005), at Laboratorio de Análisis Fisicoquímicos y Minerales (LAByM) of Facultad de Ciencias Veterinarias of UNICEN and at IHLLA laboratory.

Despite the disparity in the number and dates of sampling campaigns, monitoring can be considered representative of hydrological variations as it allowed to identify the greatest excesses and deficits peaks in the water balance for both basins.

Data pre-processing was performed considering that, for multivariate statistical analysis, all campaigns should have the same number of samples. Also, only samples with ionic

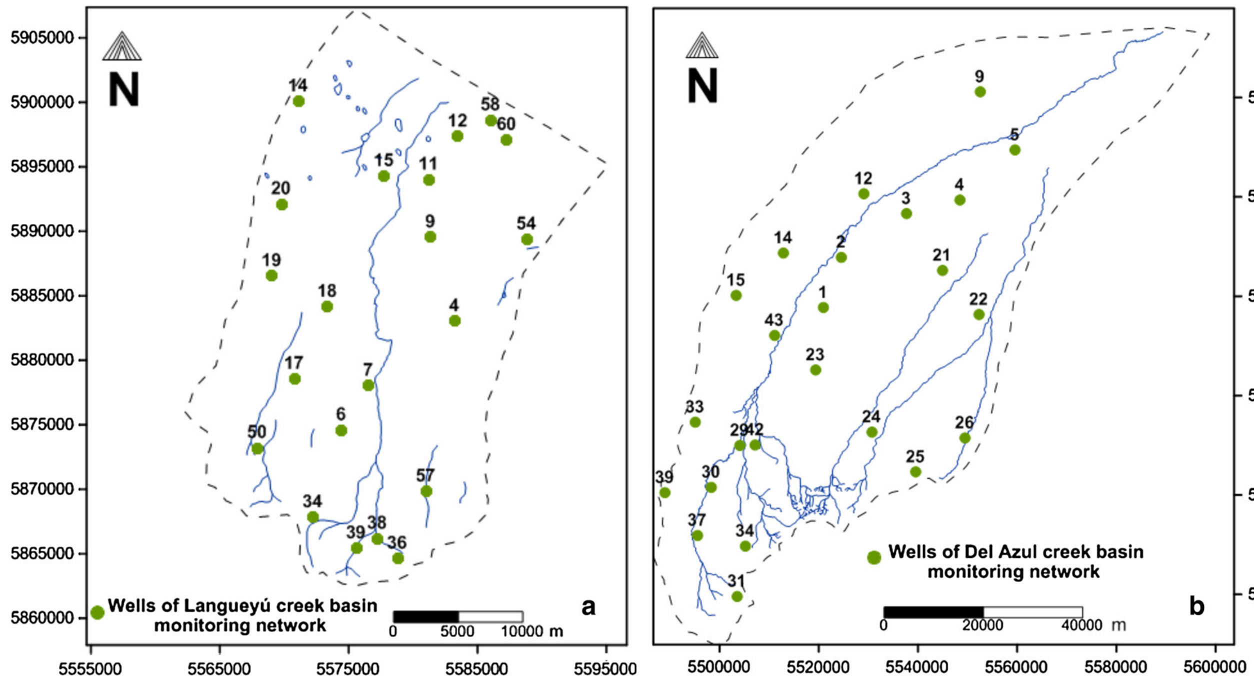

Fig. 3 Groundwater monitoring networks of Langueyú (**a**) and Del Azul (**b**) creek basins

balance error lower than 10 % were considered. The physico-chemical parameters determined and used for statistical analysis were: E.C., pH, bicarbonate, chloride, sulphate, nitrate, calcium, magnesium, sodium and potassium. The dataset of each basin was z-standardized by column to compensate the differences in magnitude and scale of the different hydrochemical parameters.

Statistical analysis

For each basin, the experimental dataset can be arrayed into a tridimensional matrix $\underline{\mathbf{X}}$, of dimensions ($n_{\text{sites}} \times n_{\text{parameters}} \times n_{\text{campaigns}}$). The dataset of the Langueyú basin contained 1260 data, 21 sites × 10 parameters × 6 campaigns, whereas in the Del Azul matrix, dimensions were 24 sites × 10 parameters × 5 campaigns, giving a total of 1200 data.

The information contained in both datasets can be visualized by standard PCA, after unfolding the tridimensional matrix $\underline{\mathbf{X}}$ along one of the dimensions or modes (in our case campaign mode) to give a bidimensional augmented matrix $\mathbf{X}^{\mathbf{aug}}$, with dimensions (($n_{\text{sites}} \times n_{\text{campaigns}}$) $\times\, n_{\text{parameters}}$). The procedure, called Matrix Augmentation PCA (MA-PCA) allows visualizing the relationships among the physicochemical parameters, by finding a small number of significant underlying factors. However, the spatio-temporal information contained in the other two modes (sites and campaigns) appears mixed and is not easily ascertained (Barranquero et al. 2014), therefore limiting the usefulness of the model and making difficult its interpretation.

On the contrary, N-PCA methods (Barranquero et al. 2014) are applied on the original N-dimensional $\underline{\mathbf{X}}$ matrices, allowing the separation of the information present in each dimension or mode and making easier to understand the groundwater conceptual model.

In this paper, the most common N-PCA algorithms, PARAFAC and Tucker3, have been applied. PARAFAC builds up models with the same number of significant factors in each dimension (Bro 1997) according to:

$$x_{ijk} = \sum_{f=1}^{NF} a_{if} b_{jf} c_{kf} + e_{ijk} \tag{1}$$

On the other hand, Tucker3 algorithm is more flexible since it enables different number of factors in each mode, and yields simpler models (Henrion 1994):

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk} \tag{2}$$

Both models originate three matrixes A, B and C, carrying the information present, respectively, in sampling sites, physico-chemical parameters and sampling campaigns. Tucker3 has an additional tridimensional G matrix that accounts for the importance of the factor interaction.

MINITAB 13.0 (Minitab Inc.) and MATLAB R2013B (TheMathWorks Inc.) software were used for statistical calculations. PARAFAC and Tucker3 modelling was carried out with the N-way toolbox for MATLAB (Andersson and Bro 2000).

Results and discussion

Descriptive statistical analysis

Langueyú creek basin

Table 1 shows descriptive statistics for the dataset. It is important to highlight that standard deviation is greater than half of mean concentration for sulphate, nitrate and potassium, thus indicating high dispersion (high variability) of the values of these parameters.

Table 1 Descriptive statistics of hydrochemical variables measured in Langueyú creek basin ($n = 126$)

Parameter[a]	Median	Mean	SD	Minimum	Maximum
E.C.	801	828	147	530	1035
pH	7.6	7.5	0.2	7.0	8.3
Bicarbonate	476	489	107	268	830
Chloride	34	34	15	3	106
Sulphate	13	17	15	3	106
Nitrate	26	29	17	4	84
Calcium	41	44	21	12	116
Magnesium	21	23	8	8	43
Sodium	118	116	57	17	253
Potassium	11	15	14	1	67

[a] All parameters are expressed in mg l^{-1} except E.C., in µS cm^{-1}, and pH, in pH units

The high dispersion and the presence of outlying values can be visualized in the box-and-whisker plots shown in Fig. 4. The interquartile ranges (IQR), defined by the extremes of the box, of most distributions are approximately symmetrical, but sulphate, potassium, calcium or nitrate have some values lying far above 1.5 IQR (the limits of the whiskers). Whiskers are approximately symmetric in the majority of parameters; E.C., bicarbonate, nitrate and sodium show longer whiskers towards high values in Del Azul basin. The same applies, although not so markedly, for nitrate and calcium in Langueyú groundwaters.

The Kolmogorov–Smirnoff normality test confirmed that only nitrate and sulphate follow non-normal distributions.

The groundwater samples show low salt contents with an average electrical conductivity of 801 µS cm^{-1}, i.e., 561 mg l^{-1} of total dissolved salts, TDS, using a conversion factor of 0.7, extracted from previous regional studies (Barranquero et al. 2013). The concentration of dissolved salts increases to the Northeast, i.e., in the flow direction. In the hills sector, E.C. ranges varied from 530 to 730 µS cm^{-1}, i.e., from 371 to 511 mg l^{-1} TDS. Towards the Northeast, E.C. increases up to values between 740 and 1035 µS cm^{-1}.

Bicarbonate is the most abundant anion in groundwaters from Langueyú basin, with contents varying from 268 to 830 mg l^{-1}. It increases slightly along the flow direction,

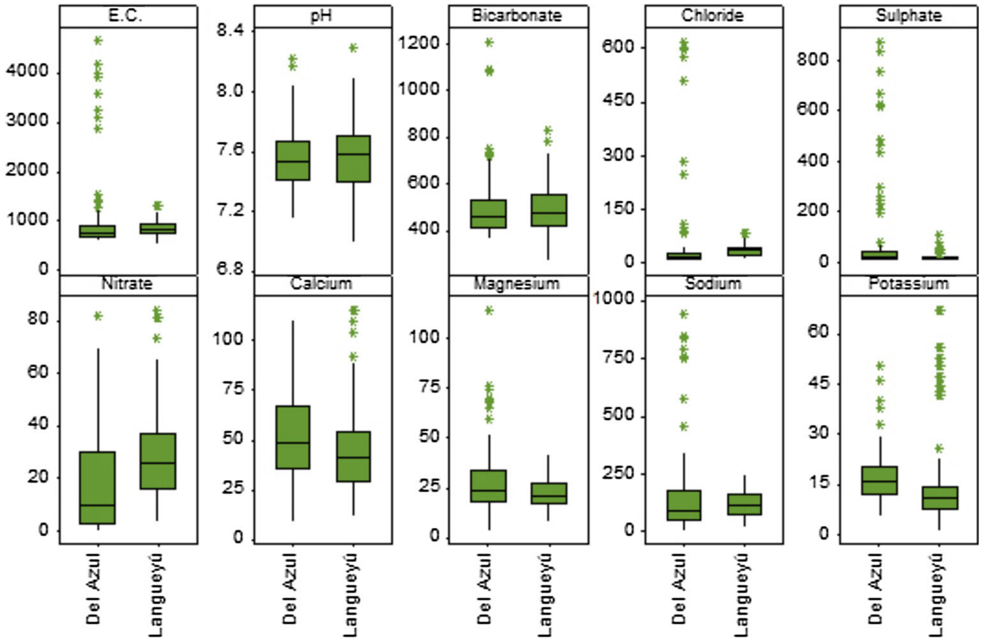

Fig. 4 Box-and-whisker plots of hydrochemical parameters measured in the studied basins (variable units as in Tables 1, 2)

being samples 14 and 15 those having the maximum values, above 600 mg l^{-1}.

Water hardness has allowed the classification of most samples as "slightly hard" (up to 100 mg l^{-1} $CaCO_3$) and eight of them as "moderately hard" according to Custodio and Llamas (1983).

The base exchange index (b.e.i.) has been calculated using Eq. (3) (Pulido Bosch 2007), where concentrations are expressed in meq l^{-1}. Negative values were obtained for all samples, indicating the occurrence of ionic exchange.

$$\text{b.e.i.} = \frac{Cl^- - (Na^+ + K^+)}{Cl^-} \qquad (3)$$

The groundwater samples can be classified as calcium/magnesium bicarbonate type in the hills sector, mostly sodium bicarbonated type in the plains sector and a mix of both types in the foothills sector. The corresponding Piper diagrams can be seen in Barranquero et al. (2014).

Del Azul creek basin

Statistics for the dataset of Del Azul basin are shown in Table 2. The values of the standard deviation and the box-and-whisker plots show a high dispersion for most variables, thus pointing to non-normal distributions and the occurrence of several outlying values, above the 1.5 IQR value.

The groundwater samples from this basin show higher salt contents than those from Langueyú, with an average value for E.C. of 1017 µS cm^{-1}, i.e., 712 mg l^{-1} TDS. It should be noted that this basin has a longer flow path than the Langueyú basin, and therefore groundwater has longer contact with sediments and older waters. The TDS concentration presents a minimum of 398 mg l^{-1} in sampling site 23, located in the South-center of the basin, and a maximum of 3255 mg l^{-1} in well 5, located at the Northeast.

Table 2 Descriptive statistics of variables measured in Del Azul creek basin ($n = 120$)

Parameter[a]	Median	Mean	SD	Minimum	Maximum
E.C.	740	1017	813	586	4650
pH	7.5	7.5	0.2	7.1	8.2
Bicarbonate	459	503	150	361	1211
Chloride	15	53	130	4	620
Sulphate	16	78	174	4	876
Nitrate	10	16	17	1	82
Calcium	49	53	24	9	111
Magnesium	24	28	17	4	114
Sodium	93	153	182	4	947
Potassium	16	17	7	5	51

[a] All parameters are expressed in mg l^{-1} except E.C., in µS cm^{-1}, and pH, in pH units

According to Custodio and Llamas (1983), samples 5 and 12 could be classified as brackish (more than 2000 mg l^{-1} TDS). All remaining samples have TDS values between 500 and 1400 mg l^{-1}, which are classified as sweet.

Although bicarbonate is again the predominant anion in all samples in this basin, samples 5, 9 and especially 12, present also high chloride and sulphate concentrations.

As regards to the cations, sodium is predominant in 71 % of samples, and calcium in the remaining 19 %, this last group of samples being mainly located in the headwaters area.

Calcium and magnesium concentrations decreased along the flow direction (towards the Northeast), while sodium and potassium increased. Potassium shows values between 10 and 30 mg l^{-1} in most samples, with a mean value of 18 mg l^{-1} and low standard deviation, thus behaving as a quasi-conservative ion.

The water samples can be classified as calcium/magnesium bicarbonate type in the South and sodium bicarbonated type in the center and North of the basin. The b.e.i. is negative in all samples of the basin indicating that ionic exchange is happening.

MA-PCA

Langueyú creek basin

As indicated above (3.2), MA-PCA is carried out by unfolding the original tridimensional $\underline{\mathbf{X}}$ matrix into a bidimensional deployed $\mathbf{X}^{\mathbf{aug}}$ matrix of dimensions 126×10. The variables were auto-scaled along the column dimension (variables) to zero mean and unit variance. Three significant principal components (PC's) with eigenvalues greater than unity were retained, explaining 65.4 % of total variance. First component, PC1, accounts for 31.4 % of variance, and is mainly associated with E.C., chloride, sulfate, sodium and potassium. PC2, explaining 21.0 % of variance, is positively related to calcium, magnesium and negatively to pH. PC3 accounts for 13.0 % of variance and is mainly associated with bicarbonate and nitrate.

Figure 5 (left) shows a summary of MA-PCA results for Langueyú creek basin. The loadings plot of the two first PC's (Fig. 5a left) explains 52.4 % of information. Two groups of variables can be identified along PC1 axis. The first includes E.C., bicarbonate, chloride, sulfate, sodium, and potassium with positive PC1 values while the second group consists of calcium and magnesium with negative PC1 values. pH is located opposite to calcium and magnesium in the PC's space, whereas nitrate appears in an

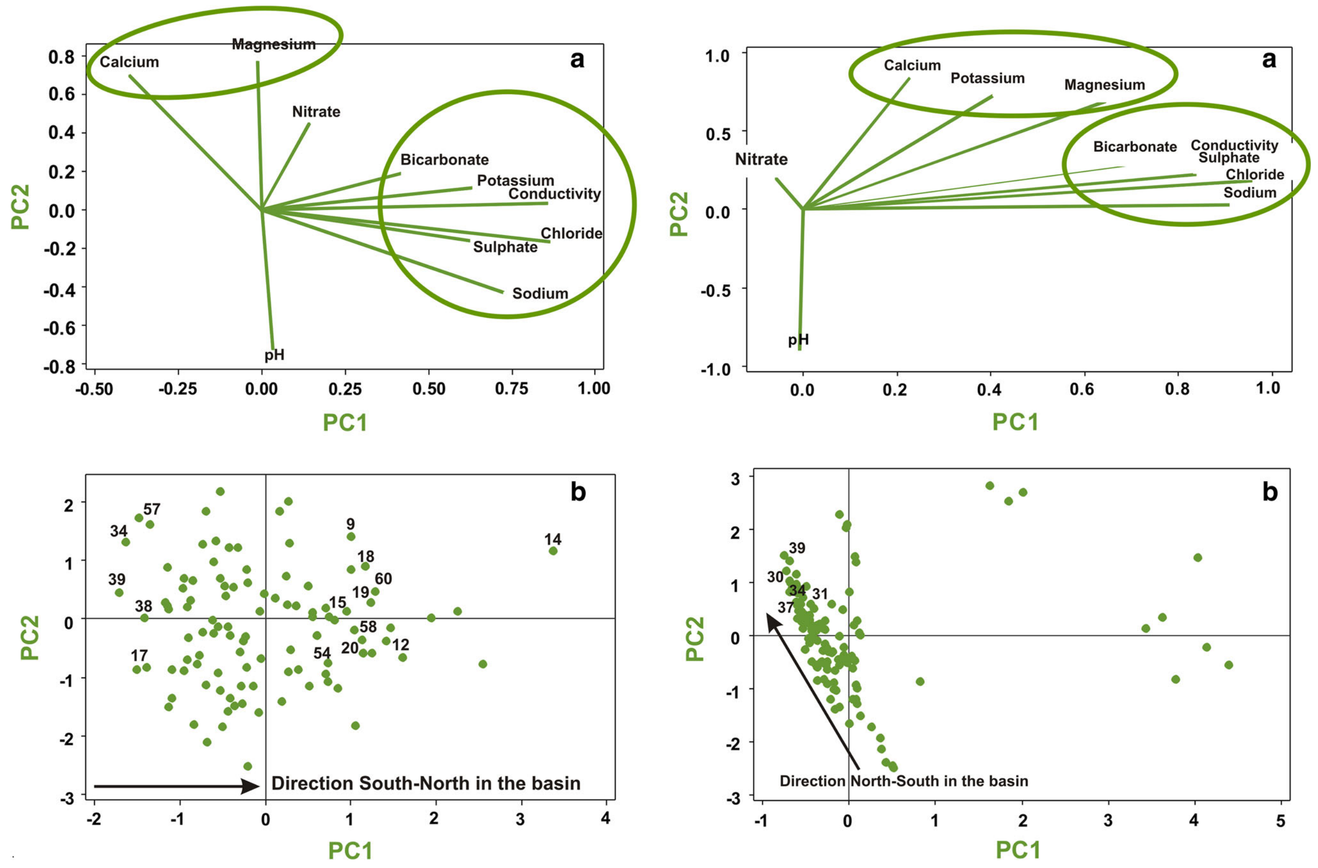

Fig. 5 MA-PCA results: **a** loadings; **b** scores of Langueyú (*left*) and Del Azul (*right*) creek basins

intermediate position, consistent with its higher loading in PC3.

Figure 5b (left) shows the scores of the samples in the first two PC's, whose main characteristic is the tendency of sampling sites to be ordered along the PC1 axis, according to flow direction, i.e., the South-North axis (see Fig. 2). Both, the grouping of variables and the ranking of samples along groundwater flow direction, seem to be in agreement with a conceptual model of the groundwater system based on salinity (E.C.) that increases, mainly, in the flow direction and ionic exchange of calcium and magnesium per sodium and (in smaller proportion) potassium, occurring also in this direction. On the other hand, nitrate is not associated with the natural hydrochemical evolution indicated above, and may be caused by anthropogenic processes.

Del Azul creek basin

In this case, the resulting unfolded $\mathbf{X}^{\mathbf{aug}}$ matrix has dimensions (120 × 10) and was also auto-scaled before processing. MA-PCA allowed finding three significant PC's in terms of the explanation of the dataset variance (89.5 %).

PC1 (48.7 % of variance) is associated with E.C., bicarbonate, chloride, sulphate and sodium. PC2 (27.9 %) is related to calcium, magnesium, potassium and pH (again with opposite sign), whereas PC3 (12.9 %) is related to nitrate.

Figure 5 (right) shows a graphical summary. The loadings plot of the first two components (Fig. 5a right), is very similar to that found for Langueyú (Fig. 5a left), except for potassium that now is associated to calcium and magnesium. pH has also a similar location in the PC's space, showing again a differential behavior regarding the rest of parameters. These variable groupings are again consistent with a hydrochemical model mainly based on ionic exchange of calcium and magnesium by sodium.

Figure 5b (right) shows the scores of samples on the first two components. Although at a first glance the distribution pattern of groundwater samples seems very different from that of Langueyú basin shown in Fig. 5b (left), there is also a marked relationship between the spatial location of sampling sites along the South-North axis. In this case, the alignment along PC2 direction is linked with ion-exchange processes, being the Southern samples, richer in calcium and magnesium, located in the upper zone of the main grouping. Nitrate shows again a different behavior likely due to its anthropogenic origin.

N-PCA

The application of MA-PCA to the $\mathbf{X}^{\mathbf{aug}}$ matrix allows extracting useful insights about the hydrochemical modeling, through the interpretation of the loadings and scores of the more significant PC's. However, and because of the unfolding procedure, spatial (sampling sites) and temporal (sampling campaign) information present in the original tridimensional $\underline{\mathbf{X}}$ matrix, appear confounded in the scores (Pardo et al. 2008; Cid et al. 2011). On the contrary, N-PCA methods allow disaggregating the spatial–temporal information.

Firstly, PARAFAC modeling was applied to the two hydrochemical datasets. In both cases, the most consistent models had two factors in each dimension: spatial, parameters and temporal. The main characteristics of these two-factors, i.e., [2 2 2], models, were: (1) conservation of the relationships amongst the variables previously found with MA-PCA, thus confirming the validity of the extracted conclusions and (2) the close similarity between the patterns of the two factors for the temporal dimension. Tucker3 model can circumvent this inconvenient since its algorithm allows to test models having different number of factors in each dimension, such as the [2 2 1] model.

Langueyú creek basin

Tucker3 [2 2 1] model was found to be mathematically consistent (Leardi et al. 2000) with a superdiagonal $\underline{\mathbf{G}}$ matrix with $g_{111} = -19.8$ and $g_{221} = 11.3$.

This model accounts for a 54.9 % of variance (a similar amount to MA-PCA). Figure 6 shows the loadings (in N-PCA the term scores is not used) for modes A, B and C, i.e., sites, parameters and campaigns. The existence of two main hydrochemical features is again confirmed: the ionic exchange of calcium and magnesium by sodium and potassium, can be seen in the opposite loadings than these parameters present in B1 and B2. On the other side, nitrate is present in both B factors, with opposite loadings on each of them, thus confirming that is not related with the natural ion-exchange processes.

To explain jointly the information of sampling locations and parameters, the signs of the terms of the superdiagonal G matrix g_{111} and g_{221} must be taken into account. In the case of [1 1 1] interaction, since g_{111} is negative and all C1 loadings are positive, sites with positive A1 loadings will be correlated with parameters having negative B1 loadings. Therefore, samples 6, 34, 39, 50 and 57 (in the Southern zone of basin) have higher levels of nitrate, calcium and magnesium, whereas sites 12, 14 and 15 present higher levels of the rest of parameters.

Del Azul creek basin

Also in this case, a Tucker [2 2 1] model, explaining 74.7 % of variance, was mathematically consistent as shows the superdiagonal structure of $\underline{\mathbf{G}}$ matrix with $g_{111} = 25.8$ and $g_{221} = -14.9$. It must be noted that this matrix has the signs of g_{111} and g_{221} opposed to those found for Langueyú creek basin. The loadings of modes A (sampling points), B (chemical variables) and C (sampling campaigns) are represented in Fig. 7.

For [1 1 1] factor, since g_{111} is positive and C1 is negative, sampling sites with positive A1 loadings will be correlated with parameters with negative B1 values. This factor therefore contains information related to both ionic exchange and saline enrichment processes. Nitrate only appears now in factor [2 2 1], correlated with calcium, magnesium and potassium. Since both g_{221} and C1 loadings are negative, sampling sites with positive A2 loadings (9, 29, 30, 33, 34, 37, 39 and 42), mainly located in the Southern sector of studied area, present, in general, higher concentrations of nitrate, calcium, magnesium and potassium (positive B2 loadings), whereas the rest of samples are correlated with the parameters with negative B2 loadings. In addition, B2 loadings show calcium, magnesium and potassium with opposite signs to sodium, so this factor is in agreement with ionic exchange processes, and linked to nitrate appearance mainly in the South of the basin.

The main conclusion to be drawn for both basins, after application of MA-PCA and N-PCA, is the existence of three principal hydrochemical processes: salinity (E.C.) tends to increase in the groundwater flow direction; ionic exchange also increases from South to North; nitrate occurs as polluting agent, but with different source and patterns in each basin. Ionic exchange and the occurrence of nitrate show a slight temporal variability.

Temporal analysis

For the analysis of the temporal variability we firstly assessed which hydrochemical parameters may be affected by changes in the water table levels. Nitrate was excluded from this analysis because of its anthropogenic origin. Calcium was also excluded because its concentration showed no significant seasonal variation (Zabala 2013; Barranquero 2014). Magnesium was not taken into account either as it is closely related to calcium. Therefore, there have been considered the variables associated with PC1 on MA-PCA (Fig. 5), except bicarbonate and potassium in Langueyú and bicarbonate in Del Azul, because these variables have smaller loadings in PC1. The remaining parameters (E.C., chloride, sulphate and sodium) are related to A1 negative loadings, sites 12, 14 and 15 in Langueyú, and 4, 5, 9 and 12 in Del Azul. All these sites are

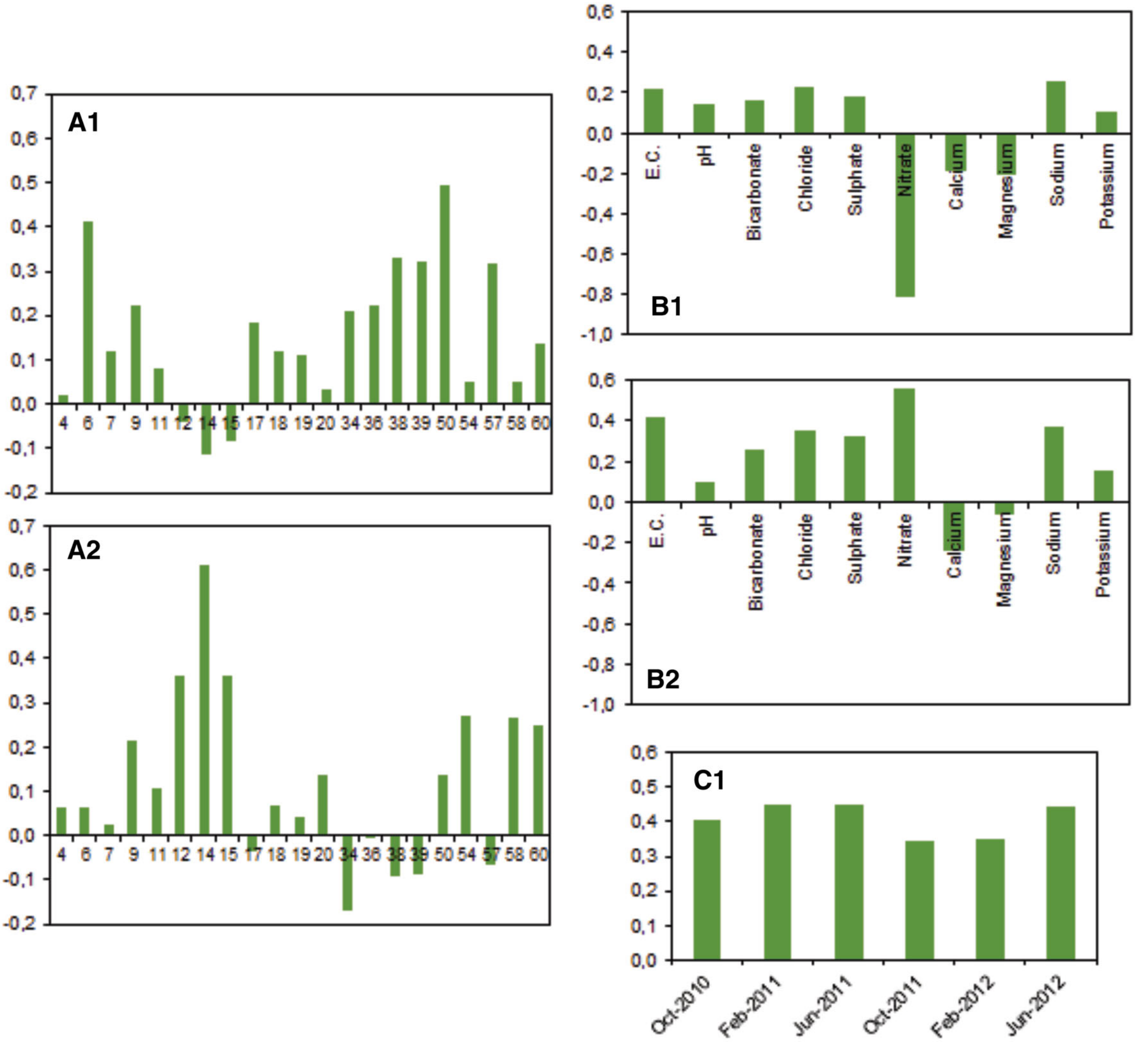

Fig. 6 Loadings plot of two factors of Tucker3 model in **A**, **B** and **C** modes for Langueyú creek basin

located in the Northeast area of their respective basin that is where waters have had a major circulation in the flow direction.

Secondly, the groundwater excesses and deficits in the campaign period, related with water table fluctuations between campaigns at sites identified above, were evaluated and compared with variation observed in factor C1 (product of C1 value per g_{111}). From this analysis, in the Langueyú basin it was found that factor C1 decreases when important excesses in water balance occur in the months previous to sampling, thus showing a dilution effect (Fig. 8).

In Del Azul basin was not possible to establish such relationships, because the water table variations at sites 4, 5, 9 and 12 show a uniform behavior only in January 2013 campaign. Varni (2013) identified, to the Northeast of this basin, exchange of groundwater with temporal water bodies located in the small depressions that are abundant in this sector. Probably because this flows, more frequent sampling campaigns are needed to explain seasonal variability.

Hydrogeological interpretation

The results of MA-PCA and N-way PCA are consistent with the hydrogeological conceptual model. Both statistical techniques reflect the main hydrochemical processes: saline enrichment in the groundwater flow direction by dissolution mainly of carbonates; exchange of calcium and magnesium by sodium in the same direction, and localized nitrate pollution.

The ionic exchange process is clearly displayed in both multivariate models (MA-PCA and Tucker3), even more strongly in Tucker3 model, with significant loadings on both 1 and 2 components. The process is showed more clearly in Del Azul basin, in which the samples are joined according this process in the scores plot shown in Fig. 5b

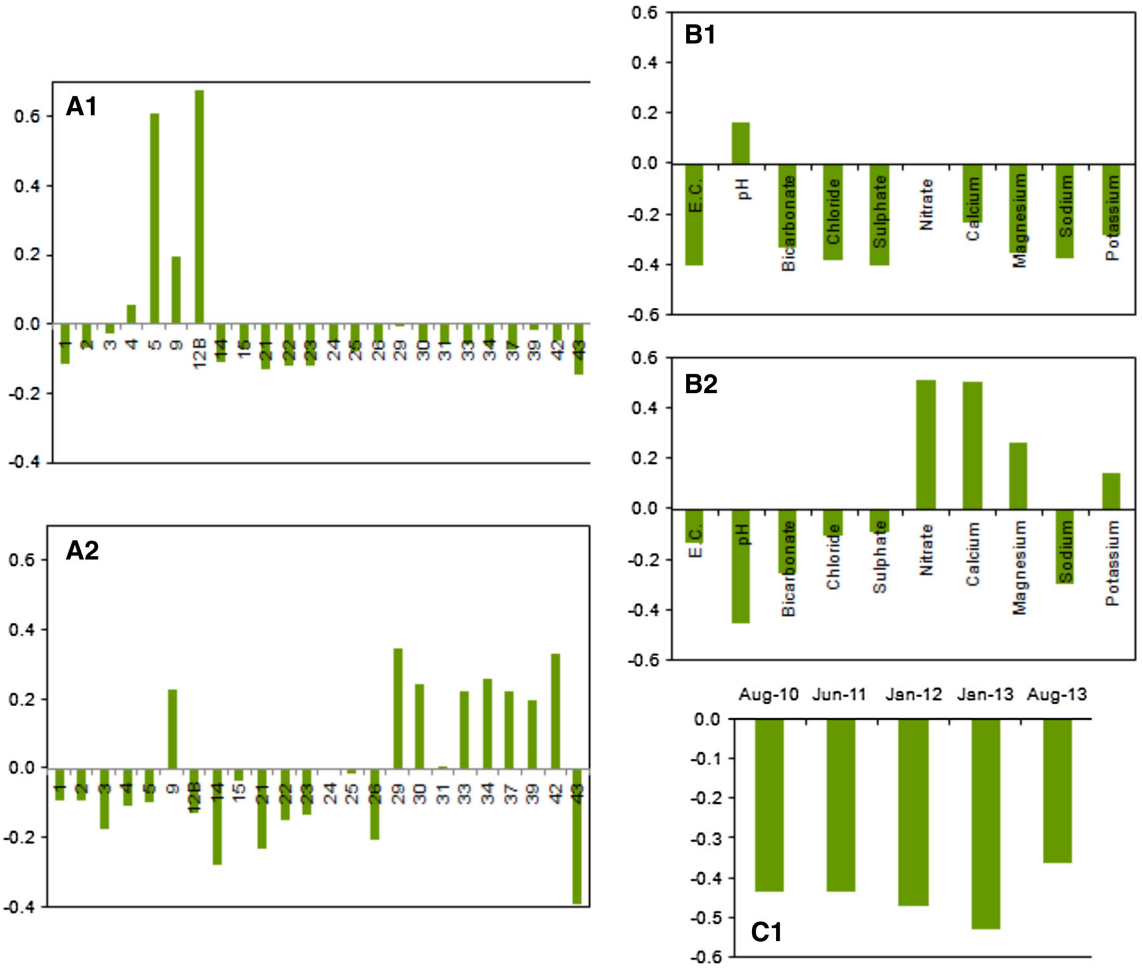

Fig. 7 Loadings plot of two factors of Tucker3 model in **A**, **B** and **C** modes for Del Azul creek basin

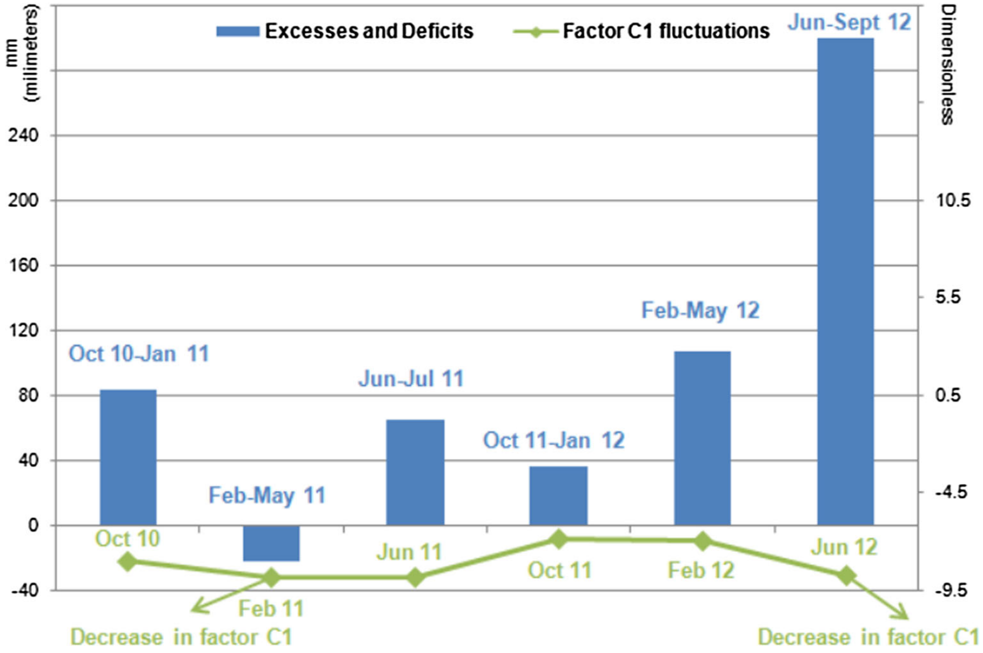

Fig. 8 Groundwater excesses and deficits and factor C1 fluctuations for Langueyú creek basin in the studied period

right. In the Langueyú basin the clustering is defined by saline enrichment (Fig. 5b left).

The different behavior of nitrate can be seen most clearly in Langueyú and is reflected in the PC3 component of MA-PCA and in B1 factor of Tucker3 model. In Del Azul basin this process is identified in PC3 of MA-PCA and factor B2 of Tucker3 model, being possible only in the latter case the association of the variable with specific samples located in the Southern sector of the basin, where the main human activity is intensive agriculture, except sample 9 that shows the highest nitrate concentrations (>35 mg l^{-1} when the average for August 2010 is 16 mg l^{-1}). Samples also differ in their bicarbonated calcium or magnesium composition, so that the ionic exchange process would be mixed in this association.

C1 loadings carry information on temporary or seasonal variability. This mode is nearly constant. However, although of low magnitude, this mode shows a correlation with some hydrochemical variables. As demonstrated in the previous section, C1 variation in Langueyú creek basin seems to be linked to hydrochemical dilution processes that occur under certain climatic conditions and for certain variables. It has been clarified through the application of Tucker3.

As noted in section, "MA-PCA" was not able to show clearly temporal hydrochemical variation, although the main hydrochemical processes were identified. Therefore N-PCA, that allowed the temporal discrimination, was applied.

Conclusions

Fundamental hydrochemical processes that explain most of data variability are common to the two investigated basins: (a) increased salinity towards the Northeast, i.e., in the groundwater flow direction, associated mainly to variables E.C., chloride, sulphate and sodium; (b) calcium-magnesium exchanged with sodium ion also in the groundwater flow direction; and, (c) localized occurrence of nitrate of anthropogenic origin and its differential behavior regarding the rest of hydrochemical variables. In Langueyú creek basin, nitrate pollution was likely associated to in situ disposal of household wastewater or loose animals around. In Del Azul creek basin, nitrate occurrence is mainly due to the use of nitrogen fertilizers in intensive agriculture. This means that, despite some slight differences, both basins show approximately the same hydrochemical behavior.

The importance of multivariate statistical support was demonstrated by the strength of the results of different statistical methods and by the consistency with the hydrogeological conceptual model. In particular N-PCA can be used to compare the study areas even when the available information does not correspond to the same sampling periods in both cases. An improved interpretation of the three dimensional nature of the hydrochemical datasets was obtained after implementation of the Tucker3 algorithm of N-way principal component analysis. The approach is validated by the amount of variance explained. As PARAFAC, Tucker3 enables robust multidimensional analyses that lead to a simpler and clearer interpretation. That is, the method ensures obtaining a simple mathematical model because it only considers full multiplicative interactions among the different dimensions. The solution obtained is unique, in contrast to classic PCA where rotational freedom exists, thus allowing the results rotation without reducing the quality of the modeling.

Background studies have explained the main hydrochemical processes in Pampean and Postpampean sediments of the Buenos Aires province (Miretzky et al. 2001; Quiroz-Londoño et al. 2008; Carol et al. 2009; Zabala et al. 2015). The contribution of this paper is to compare these processes in two basins by applying statistical methods.

Multivariate statistical analysis by N-PCA methods has demonstrated to be a useful tool to corroborate the hydrogeological conceptual model in both basins. It has also enabled the deeper comprehension of hydrochemical characteristics, as well as establishing relations between the two study areas by comparison. Therefore the work can be considered an interesting contribution to groundwater research with a regional scope. The hydrochemical characterization in the basins is a tool for a more rational management of the groundwater resources in the region. The regional scope is not only of interest from the environmental point of view, but also especially to the groundwater resources assessment in basins with different sampling designs, which may improve the efficiency of groundwater monitoring networks.

Acknowledgments Rosario Soledad Barranquero and María Emilia Zabala participate in this project thanks to the support of CONICET fellowship.

References

Andersson CA, Bro R (2000) The N-way toolbox for MATLAB. Chemom Intell Lab Syst 52:1–4

APHA (American Public Health Association) (2005) Standard methods for the examination of water & wastewater. APHA, Washington, DC

Bacon DH, Dai Z, Zheng L (2014) Geochemical impacts of carbon dioxide, brine, trace metal and organic leakage into an unconfined, oxidizing limestone aquifer. Energy Proc 63:4684–4707

Barranquero R (2009) Análisis hidrogeológico y evaluación de cargas contaminantes en la cuenca del arroyo Langueyú. Msc Thesis, Universidad Nacional de La Pampa

Barranquero R (2014) Análisis y evaluación del sistema hidrogeológico ambiental en la cuenca del arroyo Langueyú, Tandil, Buenos Aires. PhD Thesis, Universidad Nacional de Córdoba

Barranquero R, Varni M, Ruiz de Galarreta A, Quiroga M, Landa R (2013) Assessment of the hydrochemical characteristics of the Langueyú creek basin applying multivariate statistical analysis. J Argent Chem Soc 100:9–24

Barranquero R, Pardo R, Varni M, Ruiz de Galarreta A, Vega M (2014) Modelling of the groundwater hydrological behaviour of the Langueyú creek basin by using N-way multivariate methods. Hydrol Process 28:4743–4755

Bro R (1997) PARAFAC. Tutorial and applications. Chemom Intell Lab Syst 38:149–201

Carol E, Kruse E, Mas-Pla J (2009) Hydrochemical and isotopical evidence of ground water salinization processes on the coastal plain of Samborombón Bay, Argentina. J Hydrol 365:335–345

Cid FD, Antón RI, Pardo R, Vega M, Caviedes-Vidal E (2011) Modelling spatial and temporal variations in the water quality of an artificial water reservoir in the semiarid Midwest of Argentina. Anal Chim Acta 705:243–252

Custodio E, Llamas M (1983) Hidrología Subterránea. Barcelona, España

Dai Z, Samper J (2006) Inverse modeling of water flow and multicomponent reactive transport in coastal aquifer systems. J Hydrol 327:447–461

Dai Z, Wolfsberg A, Reimus P, Deng H, Kwicklis E, Ding M, Ware D, Ye M (2012) Identification of sorption processes and parameters for radionuclide transport in fractured rock. J Hydrol 414–415:516–526

Dai Z, Keating E, Bacon D, Viswanathan H, Stauffer P, Jordan A, Pawar R (2014) Probabilistic evaluation of shallow groundwater resources at a hypothetical carbon sequestration site. Sci Rep 4:4006

Faber N, Bro R, Hopke PK (2003) Recent developments in candecomp/parafac algorithms: a critical review. Chemom Intell Lab Syst 65:119–137

Galego-Fernandes P, Carreira P, Oliveira Da Silva M (2008) Anthropogenic sources of contamination recognition—Sines coastal aquifer (SW Portugal). J Geochem Explor 98:1–14

Giussani B, Monticelli D, Gambillara R, Pozzi A, Dossi C (2008) Three-way principal component analysis of chemical data from Lake Como watershed. Microchem J 88:160–166

Henrion R (1994) N-way principal component analysis theory, algorithms and applications. Chemom Intell Lab Syst 25:1–23

Hernández M, Giaconi LM, González N (2002) Línea de base ambiental para las aguas subterráneas y superficiales en el área minera de Tandilia, Buenos Aires, Argentina. In: Bocanegra E, Martínez D, Massone H (eds) Groundwater and human development. A A Balkema Publishers, Mar del Plata, pp 336–343

INDEC (2010) http://www.censo2010.indec.gov.ar/resultadosdefinitivos.asp. Accessed 1 Aug 2014

Jayalakshmi B, Ramachandramoorthy T, Paulraj A (2014) Statistical interpretation on seasonal variations of groundwater quality in Ramanathapuram coastal tract, Tamil Nadu, India. Environ Earth Sci 72:1271–1278

Leardi R, Armanino C, Lanteri S, Alberotanza L (2000) Three-mode principal component analysis of monitoring data from Venice lagoon. J Chemometr 14:187–195

Millenium Ecosystem Assessment (2005) Ecosistemas y Bienestar Humano: Informe de Síntesis. World Resources Institute, Washington, DC

Miretzky P, Conzonno V, Fernández Cirelli A (2001) Geochemical mechanism controlling pampasic ponds hydrochemistry Salado River drainage basin, Argentina. RBRH 6(4):29–39

Mohammadi Z (2009) Assessing hydrochemical evolution of groundwater in limestone terrain via principal component analysis. Environ Earth Sci 59:429–439

Nosrati K, Van Den Eeckhaut M (2012) Assessment of groundwater quality using multivariate statistical techniques in Hashtgerd Plain, Iran. Environ Earth Sci 65:331–344

Pardo R, Vega DL, Cazurro C, Carretero C (2008) Modelling of chemical fractionation patterns of metals in soils by two-way and three-way principal component analysis. Anal Chim Acta 606:26–36

Pazand K, Pazand K (2014) Hydrogeochemical investigation using multivariate analytical methods in Esfadan basin, eastern Iran. Environ Earth Sci 72:483–489

Pere-Trepat E, Ginebreda A, Tauler R (2007) Comparison of different multiway methods for the analysis of geographical metal distributions in fish, sediments and river waters in Catalonia. Chemom Intell Lab Syst 88:69–83

Pulido Bosch A (2007) Nociones de hidrogeología para ambientólogos. Universidad de Almería, Almería

Quiroz-Londoño OM, Martínez DE, Dapeña C, Massone H (2008) Hydrogeochemistry and isotope analyses used to determine groundwater recharge and flow in low-gradient catchments of the province of Buenos Aires, Argentina. Hydrogeol J 16(6):1113–1127

Ramesh-Kumar RA, Riyazuddin P (2008) Application of chemometric techniques in the assessment of groundwater pollution in a suburban area of Chennai city, India. Curr Sci India 94(8):25

Reis MM, Reis MG, Luz CC, Bastos WR, Yamashita M (2010) Aplicação do modeloTucker-3 para a análise da biodegradação de diesel. Qui Nova 33:1464–1470

Ruiz de Galarreta A (2006) Geohidrología y balance hidrológico de la zona no saturada en la cuenca superior del arroyo Tandileofú, provincia de Buenos Aires. PhD Thesis, Universidad Nacional de La Plata

Ruiz de Galarreta A, Banda Noriega R (2005) Geohidrología y evaluación de nitratos del Partido de Tandil, Buenos Aires, Argentina. In: Actas del IV Congreso Argentino de Hidrogeología y II Seminario Hispano-Latinoamericano sobre temas actuales de la Hidrología Subterránea, Córdoba, pp 99–108

Sala JM, Rojo A, Martín E, Ulibarrena J (1981) Investigaciones geohidrológicas de la cuenca de Tandil y su zona de influencia. Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, La Plata

Sala JM, Kruse E, Aguglino R (1987) Investigación hidrológica de la Cuenca del Arroyo Azul, Provincia de Buenos Aires. 37, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, La Plata

Simeonov V, Stratis JA, Samara C, Zachariadis G, Voutsa D, Anthemidis A, Sofoniou M, Kouimtzis TH (2003) Assessment of the surface water quality in northern Greece. Water Res 37:4119–4124

Singaraja C, Chidambaram S, Prasanna MV, Thivya C, Thilagavathi R (2013) Statistical analysis of the hydrogeochemical evolution of groundwater in hard rock coastal aquifers of Thoothukudi district in Tamil Nadu. Environ Earth Sci, India. doi:10.1007/s12665-013-2453-5

Singh KP, Malik A, Moham D, 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 Res 38:3980–3992

Singh KP, Malik A, Sinha S, Mohan D, Singh VK (2007) Exploring groundwater hydrochemistry of alluvial aquifers using multi-way modelling. Anal Chim Acta 596:171–182

Smilde A, Bro R, Geladi P (2004) Multi-way analysis with applications in the chemical sciences. Wiley, England

Teruggi M, Kilmurray J (1975) Tandilia. Relatorio Geología Provincia de Buenos Aires. VI Congreso Geológico Argentino, Buenos Aires, pp 55–77

Thomas J, Sabu J, Thrivikramji KP (2015) Discriminant analysis for characterization of hydrochemistry of two mountain river basins of contrasting climates in the southern Western Ghats, India. Environ Monit Assess 187:365. doi:10.1007/s10661-015-4589-0

Thornthwaite CW, Mather JR (1957) Instructions and tables for computing potential evapotranspiration and the water balance. 10, Laboratory of Climatology, New Jersey

Varni M (2013) Application of several methodologies to estimate groundwater recharge in the Pampeano aquifer, Argentina. Water Technol Sci 4(3):63–85

Varni M, Usunoff E (1999) Simulation of regional-scale groundwater flow in the Azul River basin, Buenos Aires Province, Argentina. Hydrogeol J 7(2):180–187

Varni M, Comas R, Weinzettel P, Dietrich S (2013) Application of the water table fluctuation method to characterize groundwater recharge in the Pampa plain. Hydrol Sci J, Argentina. doi:10.1080/02626667.2013.833663

Vega M, Pardo R, Barrado E, Debán L (1998) Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res 32:3581–3592

Xu J, Chen Y, Li W, Zhang L, Hong Y, Bi X, Yang Y (2012) Statistical analysis of groundwater chemistry of the Tarim River lower reaches, Northwest China. Environ Earth Sci 65:1807–1820

Yrigoyen M (1975) Geología del Subsuelo y Plataforma Continental. In: Relatorio Geología Provincia de Buenos Aires. VI Congreso Geológico Argentino. CONI S.A.C.I.F.I, Bahía Blanca, pp 139–168

Zabala ME (2013) El origen de la composición química del acuífero freático en la cuenca del arroyo del Azul. PhD Thesis, Universidad Nacional de Córdoba

Zabala ME, Manzano M, Vives L (2015) The origin of groundwater composition in the Pampeano Aquifer underlying the Del Azul Creek basin, Argentina. Sci Total Environ 518–519:168–188

Zárate M, Mehl A (2010) Geología y geomorfología de la cuenca del arroyo del Azul, provincia de Buenos Aires, Argentina. In: I Congreso Internacional de Hidrología de Llanuras. Azul, Buenos Aires, pp 65–78