Journal of the Argentine Chemical Society

ASSESSMENT OF THE HYDROCHEMICAL CHARACTERISTICS OF THE LANGUEYÚ CREEK BASIN APPLYING MULTIVARIATE STATISTICAL ANALYSIS

Rosario S. Barranquero[1,♥], Marcelo Varni[2], Alejandro Ruiz de Galarreta[3], Miguel Quiroga[4] and Roberto Landa[5]

[1]*Centro de Investigaciones y Estudios Ambientales (CINEA)-UNICEN. 399 Gral. Pinto St. C.P. B7000GHG. Tandil, Buenos Aires, Argentina.*

[2]*Instituto de Hidrología de Llanuras "Dr. Eduardo J. Usunoff" (UNICEN-CIC-Municipalidad de Azul). 780 República de Italia St. (B7300), Azul, Buenos Aires, Argentina.*

[3]*CIC-CINEA-UNICEN.*

[4]*Laboratorio de Análisis Bioquímicos y Minerales (LAByM)-FCV-UNICEN.*

[5]*LAByM-FCV-UNICEN.*

Received October 7, 2011. In final form August 22, 2013.

Abstract

This work aimed to analyze a hydrochemical dataset of Langueyú creek basin (Buenos Aires, Argentina) applying two multivariate statistical analysis methodologies: cluster analysis and main component analysis. In the basin is settled the Tandil city, which relies on groundwater for different purposes (drinking, industry, agriculture, etc.). This is why it is considered essential to a thorough understanding of hydrochemical characteristics. Phreatic levels were measured and groundwater

♥ Corresponding author. E-mail: *rosariobarranquero@yahoo.com.ar* Tel: +54-249-154495613

samples were taken with a quaterly frecuency in a monitoring net of thirty wells. Physicochemical parameters, conductivity, pH, major ions and nitrate, were determined. The results of cluster and main component analysis showed that it is possible to separate two groups of samples: One group is composed of samples located in the recharge zone and the proximal foothill (elevation 150 m.a.s.l); and a second group is composed of samples that are located to the NE. It was concluded that two processes are determinants to define hydrochemical characteristics of the study area: a) the salts acquisition during the way through the porous medium, which begins with the dissolution of carbonates, due to the acidity acquired by the reaction with CO_2 during its infiltration through the unsaturated zone, and b) calcium-sodium exchange combined with the retention of potassium in clays. Multivariate statistical analysis has proved to be a useful tool for interpreting hydrochemical results and the validation of the hydrogeological conceptual model.

Key words: groundwater, multivariate statistical analysis, Tandil, Langueyú creek basin

Resumen

Este trabajo tuvo como objetivo analizar un conjunto de datos hidroquímicos de la cuenca del arroyo Langueyú (Buenos Aires, Argentina) aplicando dos metodologías de análisis estadístico multivariado: análisis de conglomerados y de componentes principales. En la cuenca se asienta la ciudad de Tandil, que depende del agua subterránea para diferentes fines (consumo humano, industria, agricultura, etc.). Es por esto que se considera indispensable un conocimiento profundo de sus características hidroquímicas. Se midieron los niveles freáticos y se tomaron muestras de agua con una frecuencia cuatrimestral en una red de monitoreo de treinta pozos. Se determinaron los parámetros físico-químicos, conductividad, pH, iones mayoritarios y nitrato. Los resultados de los análisis de conglomerados y componentes principales mostraron que es posible separar dos grupos de muestras: Un grupo se compone de las ubicadas en la zona de recarga y piedemonte proximal (elevación 150 msnm); y otro grupo se compone de las muestras ubicadas hacia el NE. Se concluyó que los dos procesos determinantes para definir las características hidroquímicas son: a) la adquisición de sales en el recorrido por el medio poroso, que comienza con la disolución de carbonatos, debido a la acidez adquirida por la reacción con CO_2 durante su infiltración a través de la zona no saturada, y b) el intercambio iónico calcio-sodio, combinado con la retención de potasio en las arcillas. El análisis estadístico multivariado ha demostrado ser una herramienta útil para interpretar resultados hidroquímicos y para la validación del modelo hidrogeológico conceptual.

Palabras clave: agua subterránea, análisis estadístico multivariado, Tandil, cuenca del arroyo Langueyú

INTRODUCTION

Multivariate statistical analysis is a tool of immense usefulness for the knowledge of the hydrochemical characteristics of groundwater, as it allows discerning variables and relevant processes from a dataset [1].

For the current work, the general objective has been set in studying in depth the hydrochemical knowledge of groundwater in the Langueyú creek basin, applying two multivariate statistical analysis methodologies: cluster analysis and main component analysis.

The Langueyú creek basin has the particularity of displaying a surface runoff that is poorly defined and scattered in the biggest part of its area. The runoff is only concentrated in well-defined courses in its headwaters because they are controlled by hills. This surface dynamics clearly corresponds to the groundwater. Thus, it is only in the hilly zone where a precise delimitation of the underground basin can be established, in clear accordance with the water divide determined by the hills, which is controlled by outcrops. Downstream, as it advances to the plain zone, lateral limits are not so clearly visible according to the surface and can vary because of natural phenomena and/or human activity. In principle, the lateral limits have been perpendicularly calculated to the topographic level curves and they have been verified after obtaining the ground waters' equipotentials. The North edge of the basin has been established as the arbitrary political delimitation between the Tandil and the Ayacucho counties so as to enclose the study area, although the real limit is found more to the North, in the depressed zone of the Salado river, that is to say that there is a water flow (superficial and underground) that exits through this limit. Considering that, for this paper, it is more appropriate to refer to the study universe as the study area, rather than as the basin. This criterion will be applied from now on.

The study area is located to the SE-centre of the Buenos Aires Province, within the Tandilia hilly system. It is approximately 600 km^2 big, from the source of the North hillside of the hills of Tandil up to the limit with the county carrying the same name and the Ayacucho one (Figure 1). In the study area, it is located a city called Tandil which has a 50 km^2 surface and a population of 123 343 inhabitants according to the data provided by the 2010 INDEC census [2]. Taking into account the groundwater is the only possible water supply source to the population for different purposes (human consumption, industry, agriculture, etc.) thorough knowledge of its hydrochemical characteristics is considered highly important.

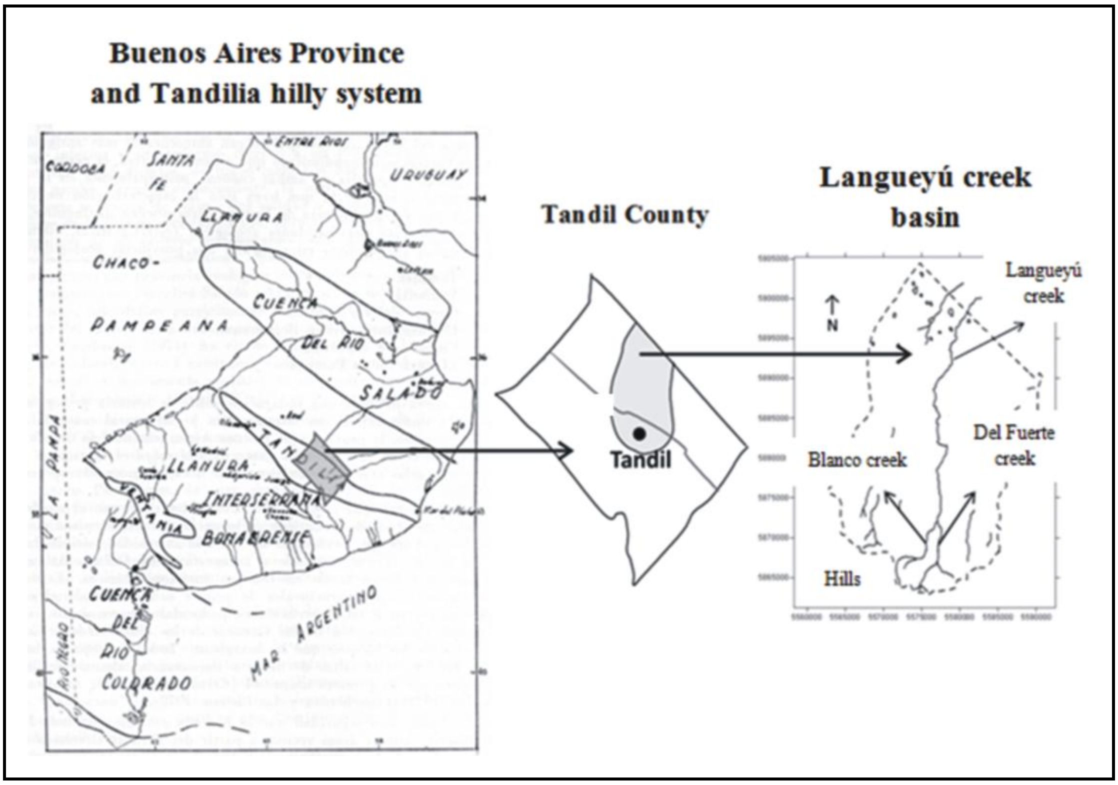

Figure 1. Location of the study area in the Tandil County and in the Tandilia hilly system.

In spite of the importance of the groundwater in the area, there are no hydrogeological study records to this spatial scale, although some papers in other basins o else in reduced areas of the basin of the Langueyú creek have contributed to the knowledge of the general characteristics of the abovementioned area.

Between 1983 and 1990, Ruiz de Galarreta has conducted hydrogeological investigations in a systematical way in the upper basin of the Tandileofú creek [3], a basin adjoining the Langueyú basin, with productions of papers to a semidetail and detail scale. The assessment of these data has resulted in the generation of a hydrological conceptual model of the Tandileofú basin [4]. The obtaining of this model has been of great usefulness to understand the functioning of the hydrological system in the basin of the Langueyú creek, because it displays common characteristics to it by the start of both in the hill field. Between the years 2002 and 2004, Ruiz de Galarreta has conducted the study "Assessment of future exploitation zones of the water underground resource to supply drinkable water to the Tandil city" (*"Evaluación de futuras zonas de explotación del recurso hídrico subterráneo para aprovisionamiento de agua potable a la ciudad de Tandil"*) (Agreement with the Tandil town hall), in the framework of which the hydrological and nitrate characterization has been performed in the NW sector of the El Chato hill [5]. It was also evaluated, specifically in the Tandil Industrial Park (TIP), the pollutant charges of the industrial activity and the potential affecting degree to the groundwater resource taking into account the vulnerability of the water system [6].

In the framework of the cooperation agreements of UNICEN with the Obras Sanitarias de Tandil (OST) *(Tandil Public Health Department)*, public body which provides drinkable water services, and with the Comisión de Lucha contra Plagas Agrícolas de Tandil *(Struggle Commission against agricultural plagues of Tandil)*, extensive papers were carried out evaluating nitrate levels in waters coming from OST exploitation wells and the underground hydrochemistry to a nitrate and frequently used agrochemicals concentration reconnaissance scale in the Tandil County [7]. Within the OST cooperation agreement, during the years 2003 and 2004 the status of drinkable water supply for Tandil was also evaluated, particularly with respect to its chemical quality. This study has provided not only data related to the presence of nitrates in the exploitation boreholes, but also data referred to the sustainability degree in the handling of the resource [8,9].

Climate, soils and geomorphology

According to monthly temperature and precipitation data of 101 years [7] corresponding to the Tandil Station of the National Weather Forecast, the climate of the city according to Thornthwaite and Mather's [10] weather classification has been typified as sub humid-humid, mesothermal, with a fairly significant water deficit and fresh summers. However, if data belonging to the last 30 years is analyzed, the weather could be typified as humid, being invariable the thermal efficiency characters and the summer concentration index.

Using the Thornthwaite and Mather's [10] water balance for the 1990-2000 period (Table 1), an annual precipitation mean value of 838 mm is obtained, real and potential evapotranspiration are 694 and 712 mm, respectively, with a fairly significant deficit value (18 mm) corresponding to January, February, and March. Water excess is 144 mm and it is distributed during May, June, July, August, September, October, and November [7].

Table 1. Water balance for the 1990-2000 period [7].

	J	F	M	A	M	J	J	A	S	O	N	D	Total
P (mm)	86	78	99	68	67	49	41	44	61	83	83	79	838
ETP	124	95	83	51	29	17	18	22	33	55	78	107	712
P-ETP	-38	-17	16	17	38	32	23	22	28	28	5	-28	
P a	-66	-83										-28	
Storage	96	85	101	118	150	150	150	150	150	150	150	124	
Stor. Dif.	-28	-11										-26	
Deficit	10	6										2	18
Excesses	-	-	-	-	6	32	23	22	28	28	5	-	144
ETR	114	89	83	51	29	17	18	22	33	55	78	105	694

From a taxonomical point of view, according to the Soil Survey Staff [11], the soils of the study area are represented by a 90% of mollisols. In detail, it can be indicated that in the hills, there are mostly lytic hapludolls, in the foothills arguidols stand out and to the NE of the study area, the presence of thinner materials allow the development of typical natracuols.

The Tandil County, according to its geomorphologic characteristics, has been divided into de following zones: hills, foothills, and plains [7]. Through different investigations that have been taking place in the study area since 2007, these zones (Figure 2) have been delimited and characterized and it has been pointed out that its division would be appropriate for a first close up to the description of the conceptual model of the groundwater system.

The hills zone displays sectors with steep relief with emerging rocks and other sectors of wavy relief with a crystalline bed covered by a thin layer of loess. This determines a marked structural control of the hydrological dynamics. The waters that are concentrated in small sub-basins generate an integrated superficial drainage with dendritic design and with a maximum gradient which fades out to the NE, that is, in the direction of the underground flow.

The crystalline bed in the hills zone plays the role of a fast recharge source of precipitation water. Meanwhile, the bed in its depth acts as an impermeable base, in its practical aspect, of the porous means, especially when it has reached great depth and the thickness of the sedimentary package is important.

The foothill zone has softer slopes than the hills and has some isolated depressions that interrupt the regularity of the relief. Within the study area, it is extended from the slope break next to the hill zone to, approximately, the topographic curve of 150 m.a.s.l (Figure 2). The superficial drainage net is well defined and it has a distributive design. The drainage is divergent in accordance with the foothill regional morphology. In this environment, the Langueyú flow is found in solitude, without receiving additional contributions.

The plain zone is extended downstream the topographic curve of 150 m.a.s.l and it continues to the Salado river depression. As it has previously been pointed out, the study area has not covered the whole of the zone, because of which the regional underground discharge would not be included. In this environment, the slopes are highly soft and there are superficial materials that were originally aeolian and later suffered from the transport and redepositation by means of laminar flow. These characteristics make the superficial drainage poorly defined and integrated, with narrow courses and, in sectors, temporary flows that are lost in soft depressions.

The phreatic surface evidence the low water gradients present in this zone and they portray the flow lines with almost parallel directions to the Langueyú creek. Nonetheless, some local underground discharges occur, which make up the basic flow of the creek and assure its permanent regimen (Figure 3).

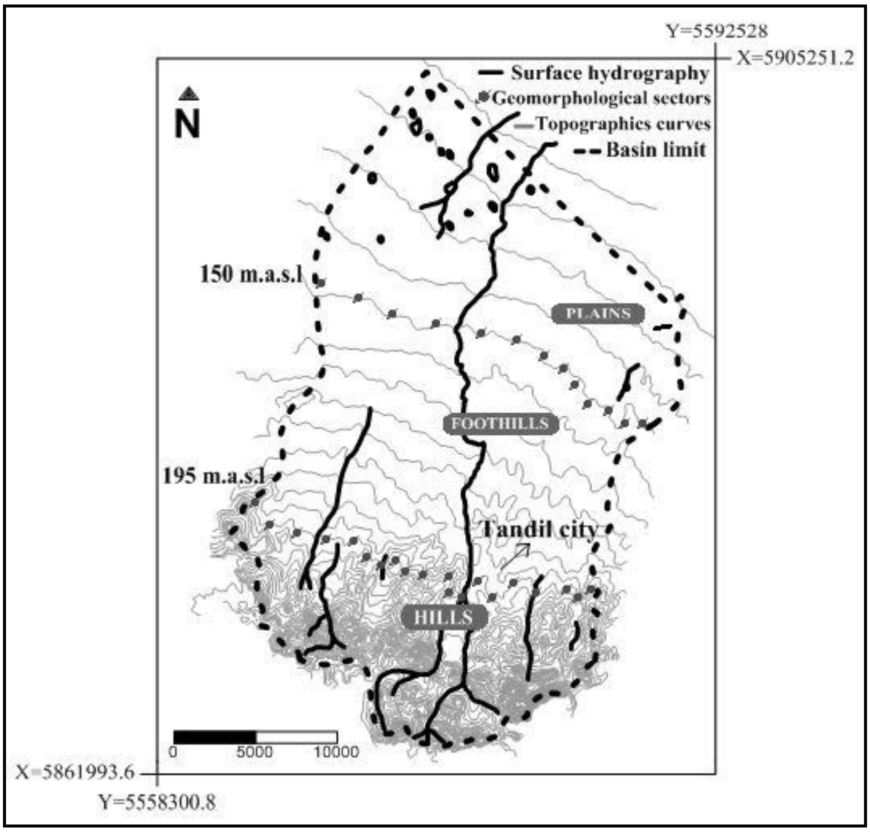

Figure 2. Geomorphologic zones.

Geology and Hydrolitology

From the stratigraphic point of view, the Tandilia hilly system has a crystalline bed base of Precambrian age, constituted by a typical association of igneous-metamorphic rocks called Buenos Aires Complex [12] and sedimentary rocks of young Paleozoic age or Precambrian age which originally covered the entire hills zone. This material has been denuded by the erosion after the elevation of the blocks in the Cenozoic period, leaving in some sectors the crystalline bed uncovered. In the study area, the bed is only superficial in the hills zone and, superimposed to it, there is the Cenozoic sedimentary cover mainly composed of Pampean Sediments, and to a lesser extent, Postpampean ones, which are originated in the loess and are massive with calcrete presence (caliche). From the mineralogical point of view, this Pampean loess is characterized by an abundance of plagioclases (20 to 60%), relatively little quartz (20 to 30%) and a considerable percentage of volcanic glass (15 to 30 %) [13, 14].

As regards the hydrolitology, the crystalline bed on the one hand and the Pampean and Postpampean Sediments on the other can be differentiated as two hydrogeologic units that present different behavior towards admission and circulation of groundwater, given their constitution, texture and structure [7].

The crystalline bed is composed of rocks which are primarily aquifuge with different fracture degrees which consequently make up a poor aquifer. The cracked environment in this crystalline bed presents secondary porosity and permeability. The caudal distributed in private wells located in these types of rocks do not generally go beyond 1 m^3 h^{-1}.

This crystalline bed, emerging in the hills, is intensified towards the N, where it is over 200 meters deep, representing the impermeable base of the aquifer.

Meanwhile, the Pampean and Postpampean Sediments are the porous means in which the set of exploitation wells of OST is located for the drinkable water supply of the Tandil city. The performance of these wells is variable and according to the transmissivity of the aquifer, the performance can be higher than 100 $m^3\ h^{-1}$.

In spite of the vertical heterogeneity, the deepest wells carried out in the study area done by OST have presented phreatic levels comparable to the ones found in superficial wells, because of which it can be pointed out the presence of a single aquifer. The saturated thickness within the hilly environment is very variable, according to the thickness of the sedimentary cover on the cracked rock blocks constituting the crystalline bed. In the extra hill environment, the thickness of the sedimentary package has been verified between 55 and 90 meters, considering the lithological profiles of the OST wells.

Summing up, the groundwater of the study area is presented in a single multiunit aquifer characterized by some vertical heterogeneity. In general terms, the dynamics of its upper part is representative of what happens in all its thickness in most of the part of the study area, except for the wells located towards the North limit that, because of being superficial, can be affected by a direct evaporation process located and restricted to the upper part of the aquifer.

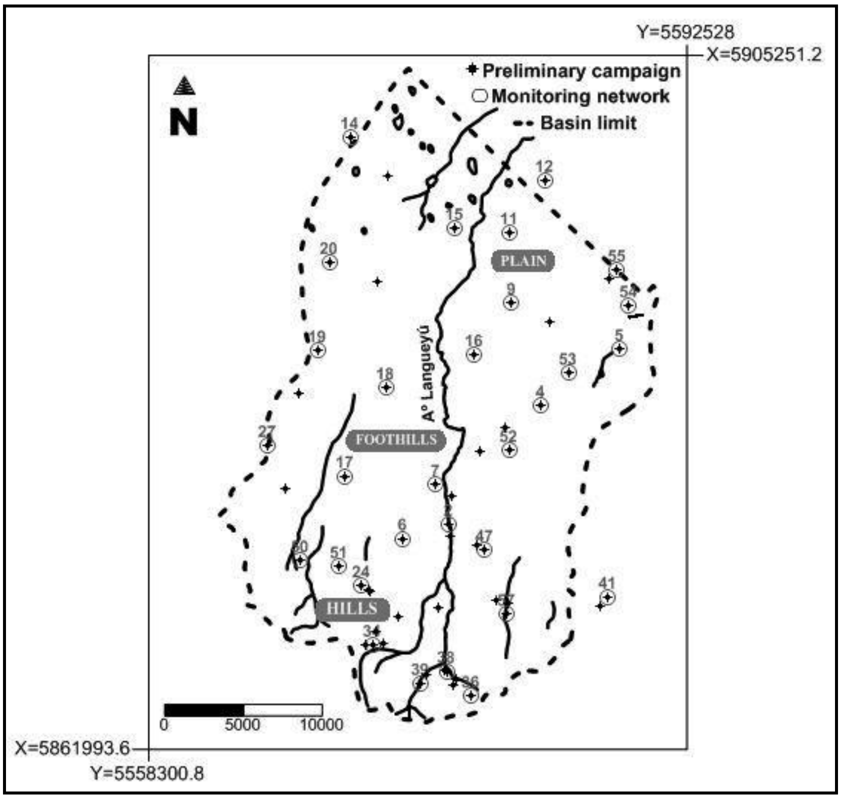

Figure 3. Phreatic surface map corresponding to February 2008. Samples taken in the preliminary campaign and selected for the monitoring net.

Hydrodynamics and hydrochemistry

According to Ruiz de Galarreta et al. [15], the depth of the water table varies between 15 m in some wells to the S of the study area and 2 m towards the NE. as it can be seen in Figure 3, the phreatic levels display a NE underground flow, in accordance with the superficial morphological characteristics, though with a smaller gradient.

The aquifer recharge is regionally native by precipitations, mainly locating in the hills. Meanwhile, the regional discharge is produced towards the depressed zone of the Salado river and, locally, in the perennial streams and main affluent of the Langueyú creek. In the OST wells exploitation zone, this relationship can be inverted.

The obtained specific electrical conductivity values allow classifying the groundwater as freshwater [16]. As regards hardness, most of the samples are “moderately hard”. In all the cases, the values are within the permitted values of the Código Alimentario Argentino *(Argentinean Food Code)* for drinkable water (400 mg/l). All the samples of the study area have been typified as sodium bicarbonated.

METHODOLOGY

The hydrochemical data used for multivariate statistical analysis come from a monitoring net created in 2007 for the characterization of the hydrochemical system of the basin, from a set of existing wells. The determining of the monitoring net was based on the general objective of the investigation [17], considering that if the goal is obtaining a regional diagnosis of a relatively homogenous area, an approximately regular distribution of the wells must be prioritized, which covers all the study area and makes the obtaining of representative samples possible to know what happens to a regional level. Figure 3 displays the location of the identified wells during the preliminary campaign and the selected ones for the monitoring net. The measuring was seasonal: samples were taken during June and October of 2007 and February and June of 2008, thus completing an annual cycle.

The wells that are part of the monitoring net have different depths depending where they are geographically located, but in all the cases, they are superficial wells that go into no more than 2 or 3 meters under the depth up to which the water table is. The wells have manual extraction devices (hand pump), eolian ones (mills) and electrical ones (centrifuge submersible pump, pumpers) which were turned on some minutes before taking the sample. These were bottled in 1 lt. PET (Polyethylene Terephthalate) containers which, differently from other plastics such as Polyethylene, it does not allow gas diffusion [17]. The containers were rinsed with the water which was about to be collected and were filled in, without leaving airlock, taking into account that the alkalinity would not be determined in the field. The samples were transported towards the laboratory in coolers and the alkalinity determination was conducted by colorimetric titration (SM 2320B) before the first 24 hours.

The electrical conductivity and the pH were measured in the field, considering the susceptibility that present with respect to the changes of the environmental conditions [18]. The electrical conductivity was measured with an Orion 105 Aplus conductimeter, with automatic correction by temperature. The pH was determined with a pH meter with a 0.01 resolution and was calibrated with buffer solutions for 7 and 10 pH. The remaining determinations, conducted by the Laboratorio de Análisis Bioquímicos y Minerales *(Laboratory of Biochemical and Mineral Analysis)* of the Facultad de Ciencias Veterinarias *(Veterinary Science Faculty)* (UNICEN), were: chlorides, sulfates, nitrates, calcium, magnesium, potassium, and sodium. The determinations were conducted utilizing normalized methods [19].

Multivariate statistical analysis

To conduct multivariate statistical analysis, the clustering and factor analyses were employed, using the software Statistica 5.5 [20]. Although the number of samples that were used is not high, from the statistical point of view, the application of the multivariate analysis is valid, nonetheless, given the fact that it is intended to determine the relationships between the analyzed parameters and their relative contribution to the explanation of the hydrogeochemical variation in the area. In no case will be intended to draw statistical inferences on the population, because for the latter, the sampling should have been at random, unbiased and with a number of observations much higher than the hereby used.

Clustering Analysis

The clustering analysis consists of forming sample groups that are associated according to their similarity. This imprecise definition is due to the great variety of existing techniques to fulfill such clustering objective. In this paper, it has been decided to use the agglomerative hierarchical analysis technique, which groups things or objects in groups which get bigger and bigger until there is only one group [21].

The first problem to be solved is how to decide on the similarity of two samples, that is, how do decide if sample A is more similar to B or to C? To do such thing, a similarity coefficient is utilized. Numerous similarity coefficients have been formulated, which can be grouped up in distance, correlation, and association coefficients.

To understand the distance coefficients, let us suppose that there are two samples, P1 and P2, over which three characteristics have been measured, X, Y, and Z. These two samples can be illustrated in the tridimensional space generated by the variables X, Y, and Z. This can be seen in Figure 4, where it is also shown the distance between P1 and P2, which is easily calculated by means of algebraic calculations. It is supposed that the samples with less distance between them are the most similar ones. Of course, this reasoning can be widened from a tridimensional space to an n-dimensional space (n variables) [22].

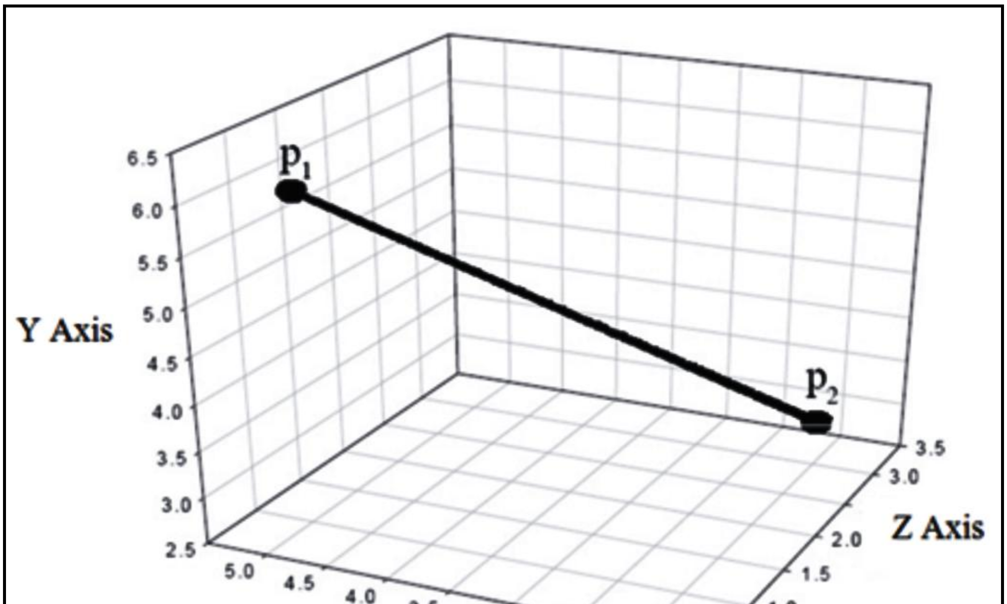

Figure 4. Distance between two points in a tridimensional Euclidean space.

In the case of the correlation coefficients, if they are considered vectors that are initiated from the origin of the n-dimensional variable system and their ends are located in the coordinates of each sample, the angles among these vectors can be measured. It is considered that the narrower the angles, the more similarity they have.

The association coefficients are used for data of the double state type (presence-absence). In this paper, distance coefficients will be used.

The similarity coefficients among different samples can be ordered in a matrix: this matrix is insufficient to express the relationships among the different samples; it only displays the similarities between pairs. There is a great variety of techniques of similarity matrix analysis to allow distinguishing the relationships among all the samples. Within these techniques, we are particularly interested in the clustering analysis, or cluster, which is the one that will be used in this paper.

There are numerous clustering analysis techniques which are classified according to:

- Techniques that form exclusive groups versus techniques that form non-exclusive groups;
- Techniques that form hierarchical groups versus techniques that form non-hierarchical groups;
- Divisive techniques versus agglomerative techniques, and
- Sequential techniques versus simultaneous techniques.

In this case, we will use a clustering technique which is exclusive, hierarchical, agglomerative, and sequential.

The clustering of two samples is easy; we simply associate the two samples with the highest similarity coefficient. Now, once that the first group is formed, the remaining samples can associate to another sample or another group. The similarity coefficient between the group and the sample is defined from the simple, complete and average linkage concepts. In the simple linkage, it is adopted the similarity between the group and the sample as the similarity of the group sample which is the most similar to the most extreme one. In the case of the complete linkage, the similarity of the sample with lower similarity coefficient is taken into account and in the average linkage, the average of the coefficients. In this paper, we will use the complete linkage.

Analysis of factors

The analysis of factors is a statistical method used to describe the variability among certain observed variables in relation to a lower amount of hypothetical variables, called factors [23, 24]. That is, it is possible, for example, that the variations among ten observed variables can be explained to a great extent by two or three variables which are not directly measurable (the factors). When replacing the original variables with the factors, part of the variance to be explained is lost, but the model becomes simpler. The observed variables can be modeled as lineal combinations of the factors plus an error term. The analysis of these linear combinations can be used to explain the factors as new variables with some sense with respect to the processes that are being explained.

The commonest way of factor analysis is the analysis of principal components. In it, we seek a linear combination of observed variables so as to obtain the greatest variance of them (Principal Component 1). Then, a second linear combination is investigated, orthogonal to the first one, which explains the highest amount of remaining variance, and so on. Thus, a set of uncorrelated or orthogonal independent factors is obtained in a geometric sense.

To evaluate the clustering of the samples, the Q factor analysis is used in this case, which analyzes the variance of the samples, not of the variables. That is, it works with a transposed data matrix. In this analysis, the main components will be extreme pure samples, whose combination produces original samples. In a historical work, Klovan [25] uses this analysis to distinguish the deposition environments from the distributions of sediments in the Barataria Bay beaches, distinguishing three predominant types of energy in the deposition of these sediments: wave, ocean currents, and gravity energy.

RESULTS AND DISCUSSION

Cluster analysis

For the analyzed data set, as shown in Figure 5, if it is decided to cut into the distance 20, two groups of samples are defined.

One group is composed of samples located in the recharge zone and the proximal foothill (elevation 150 m.a.s.l) of the study area (Group I), except for sample 11, which is located to the NE (Figure 6). A second group is composed of samples that are located to the NE (Group 2), mainly on the right bank of the Langueyú creek, except for sample 23, which joins this group even when being upstream.

The main difference in the hydrochemical characteristics between these two groups is the electrical conductivity. It has conductivity values between 664 and 825 μS cm^{-1} for Group I and between 825 and 1116 μS cm^{-1} for Group II. A major important difference is also detected in the bicarbonate values, which range between 7.6 and 10.6 milliequivalents per liter (meq l^{-1}) for the former and 5.5 to 8 for the latter.

From a special analysis of sample 23, it can be inferred that its link to the Group II lies, first, in a much higher conductivity value (1145 μS cm^{-1}) than the samples which are located nearby. It also differs from them because of its chloride content, which is 1.91 meq l^{-1}, where the average value of close by wells is only around 0.68 meq l^{-1}. The sample also presents a 137.7 mg l^{-1} nitrate value, which allows inferring that it is affected by specific pollution of an anthropogenic nature.

In the case of sample 11, it is evidenced that its association with Group I is due to its low sulphate content. This constitutes 3.4% of the total anion content, being 4.2% the average content of Group I. While virtually all the samples that form Group 2 have sulphate values which are greater than 5%.

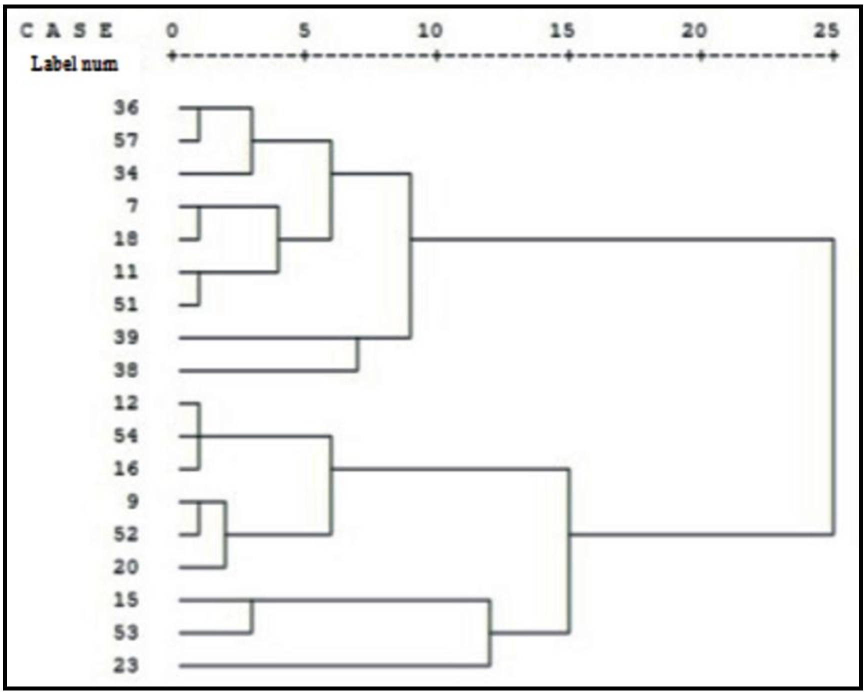

Figure 5. Phenogram for the set of analyzed samples with complete linkage and Euclidean distance.

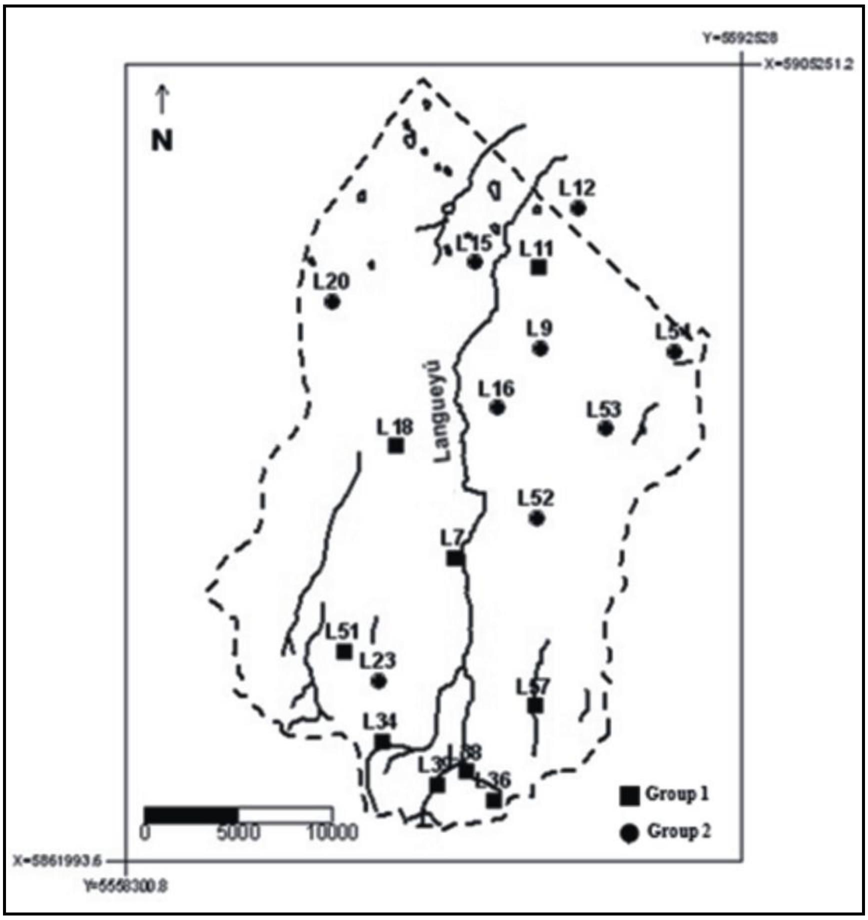

Figure 6. Location of the set of analyzed samples.

Principal Components Analysis

First, the principal components analysis is applied to the transposed data matrix (variables in rows and samples in columns). Figure 7 shows the samples in the plane Principal Component 1 (PC1) - Principal Component 2 (PC2), which would be, as stated above, two extreme samples. In this case, PC1 represents a relatively saline (older) sample to be located to the N, while PC2 represents young water recently recharged. PC2 is associated with the 23 and 39 wells first, while 38 and 36 are associated with lower values. In the case of PC1, a more strongly associated group could be determined by samples 9, 12, 16, 20 and 54.

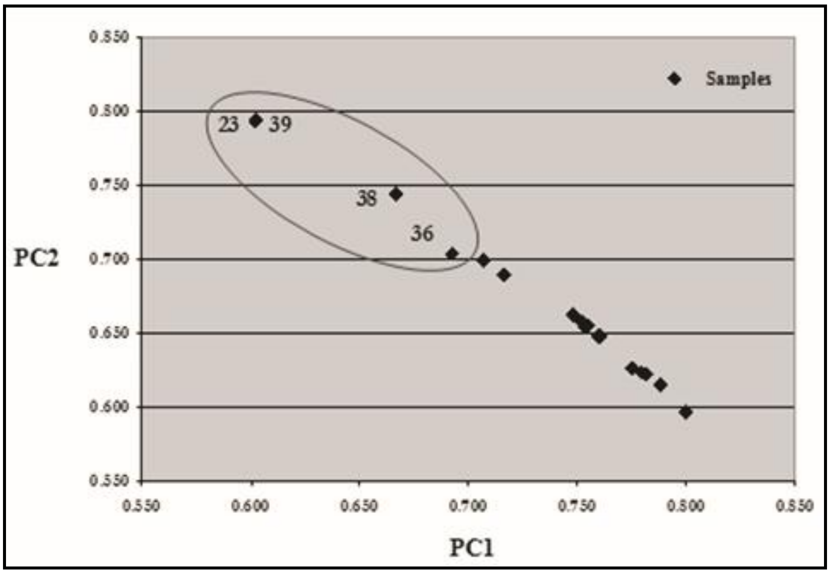

Figure 7. Principal components chart showing samples associated with PC2 (marked with an ellipse).

In order to understand in a more complete way the processes through which the samples acquire and change their composition, a principal component analysis has been applied to the physicochemical variables analyzed for the set of samples of October, 2007. Table 2 shows the results of the analysis. In this case, as in the analysis applied to the samples, Varimax rotation is used. The solution is limited to three main components, as the fourth explains a very small variance percentage. The variance explained by the three components represents 79.3% of the total variance.

Table 2. Physicochemical variables that explain the various factors obtained for the dataset.

October 07	pH	EC	HCO_3	Cl	NO_3	SO_4	Ca	Na	Mg	K
Factor 1	-0.098	**0.939**	**0.715**	**0.830**	0.098	**0.796**	0.034	**0.550**	0.091	**0.627**
Factor 2	0.158	0.076	-0.229	0.055	**0.851**	0.102	**0.986**	**-0.793**	**0.843**	-0.344
Factor 3	**0.817**	0.076	0.488	-0.217	-0.311	-0.204	0.063	-0.018	0.377	0.047

In the first factor, the electrical conductivity appears with a correlation which is slightly higher than 0.9. In the same vein, chlorides, sulfates, bicarbonates, potassium and sodium are associated, all of them with positive correlation. All these variables are strongly responsible for the salinity content in natural conditions, in general, of the water in the study area and its surrounding basins. The close relationship conductivity / bicarbonates is based on the chemical attack in the infiltration area, facilitated by the presence of carbon dioxide, to carbonates such as calcrete, enabling the high content of bicarbonates of young waters. But also, as this PC is very broad and it contains all the variables that contribute to the increase in total salts, it makes the other anions (chlorides and sulfates) also present a positive association with it. This means that the anions vary in the same way; they all grow in the direction of the flow, as groundwater acquires them in its path. PC1, then, could be called "increase of salinity."

PC2 is represented by calcium and magnesium that have the same behavior in a spatial sense across the study area; both decrease in the direction of the flow, in accordance with the increase of the $Na^{+}/Ca^{++}+Mg^{++}$ and the Na^{+}/Ca^{++} relationships, which would point out the existence of the ionic exchange process. The appearance of sodium in factor 2 showing a negative association reinforces this hypothesis. PC2 would therefore represent the ionic exchange process.

The presence of ionic exchange is also proved by applying the imbalance index (i.d.d.) proposed by Catalán Lafuente [26].

$$i.d.d. = r\frac{Cl^{-} - \left(Na^{+} + K^{+}\right)}{SO_4^{+} + HCO_3^{-} + NO_3^{-}} \qquad (2)$$

The i.d.d. is negative throughout the study area, ranging from -0.5 to -1.65, which indicates that ionic exchange is indeed occurring. In a spatial sense, it corresponds to an imbalance index less negative (higher) as we move towards the NE.

In PC2, nitrates are also associated. They have elevated values in some samples by anthropogenic impacts. The association is positive with respect to calcium and magnesium because the samples with higher levels of nitrates are found in the recharge zone. That is, in the direction of the flow, the three ions decrease but in different ways. In the case of nitrates, the decline is not gradual as it happens with calcium and magnesium, because the contamination is specific.

The pH appears as a variable that defines PC3.

To complete the analysis, was performed a scatter graph in which it is displayed on the horizontal axis the electrical conductivity in mS cm^{-1} and the chloride content (in meq l^{-1}) on the vertical axis (Figure 8). As reference, the values of a seawater sample are placed in the graph. Directed at these features, samples 15 and 53 have a water table depth of 1.5 and 1.9 m respectively. It could be assumed, in principle, that they would be undergoing the process of direct evaporation from that level, which would explain their high chlorides proportion. However, to prove this hypothesis, more specific analysis should be conducted in the plain area. Sample 23, located in the hilly area, is also in this sector of the graph but not because it is affected by the same process, but because it presents specific pollution as detailed above.

At the bottom of the graph, sample 36 is placed, with a conductivity characteristic of a low-water route (0.782 mS cm^{-1}) and low chloride content, so it could be assumed as water from local recharge.

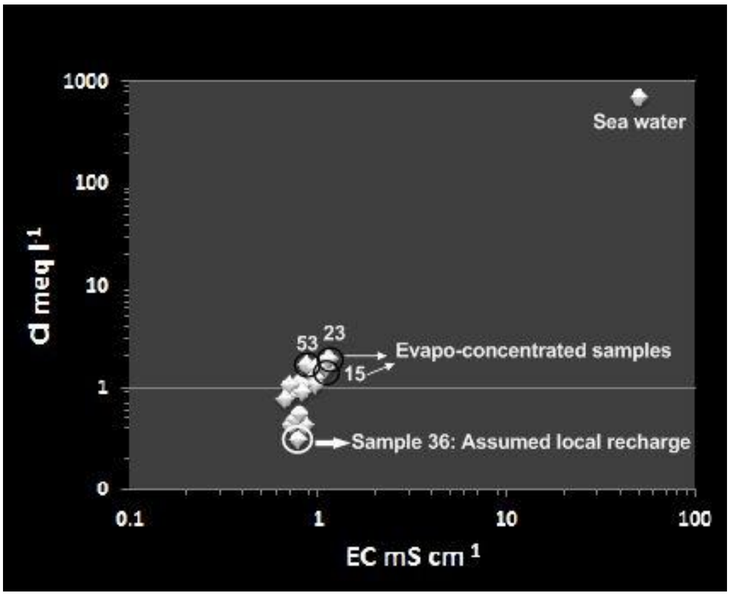

Figure 8. EC vs. Cl graph for samples of the campaign in October, 2007.

CONCLUSIONS

It is concluded that the hydrochemical behavior in the study area is differentiated in different geomorphological areas. In the hilly area, rapid water flow determines a low degree of salts acquisition (the medium there is a combination of cracks in compact rocks and a porous medium), whereas in the foothills, there is a marked increase in salinity, stressed by water flow only in the porous medium, and thus with longer contact with sediments. Meanwhile, in the plain area, although there has been an expected increase in salts content given the longer route, the level of sulfates and chlorides show that the regional discharge occurs downstream, out the limit imposed for the study. In this area, it was also observed that, in the upper aquifer, solutes can concentrate because of direct evaporation, as the water table is very shallow. It should be noted that this solute concentration is recorded because of the shallowness that, in some cases, the sampled wells have; that is, the chemical results in these cases would represent what is happening at the top of the aquifer and not the hydrochemical regional evolution.

The processes identified as determinants of the hydrochemical characteristics of the study area are: a) the salts acquisition during the way through the porous medium, which begins with the dissolution of carbonates, due to the acidity acquired by the reaction with CO_2 during its infiltration through the unsaturated zone, and b) calcium-sodium exchange combined with the retention of potassium in clays.

In accordance with these processes, the samples are grouped by their similarity in their conductivity values, as an indirect measure of the total content of dissolved salts and the proportions of different cations.

Multivariate statistical analysis has proved to be a useful tool for interpreting hydrochemical results and the validation of the hydrogeological conceptual model, especially considering that there is no history of such studies in the study area. It should also be noticed that, for the moment being, there is a lack of economic and infrastructure resources to carry out investigations involving the analysis of a larger amount of samples and/or variables.

REFERENCES

[1] J. F. Hair, R. Anderson, R. Tatham y W. C. Black, *Análisis Multivariante*, 5ta ed. Prentice Hall Iberia, Madrid, 1999. Cáp. 3.

[2] www.indec.gov.ar (Censo Nacional de Población y Vivienda 2010-Instituto Nacional de Estadística y Censos).

[3] Hernández, M. y Ruiz de Galarreta, A., *Comportamiento hidrodinámico del acuífero freático en el sector superior de la Cuenca del arroyo Tandileofú, Pcia de Bs. As.*, en: *I Jornadas Geológicas Bonaerenses*, Tandil, 1985, pp. 391-408.

[4] Ruiz de Galarreta, A., *Geohidrología y Balance Hidrológico en Zona no Saturada en la cuenca alta del arroyo Tandileofú, Provincia de Buenos Aires.* Tesis doctoral. Inédita. 2006.

[5] Ruiz de Galarreta, A., *Caracterización Hidrológica y de Nitratos en el sector NO del cerro el Chato, Tandil*. Informe final. Convenio CINEA. Universidad Nacional del Centro-Municipalidad de Tandil. Inédito. 2004.

[6] Ruiz de Galarreta, A., Banda Noriega, R. y Granato, *Riesgo de contaminación hídrica en el Parque Industrial de Tandil, Provincia de Buenos Aires*, en: *Seminario vulnerabilidad de acuíferos*, Lima, 2004.

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

[8] Barranquero, R., Ruiz de Galarreta, A., Banda Noriega, R., *Evaluación de nitratos en los pozos de explotación en la ciudad de Tandil, Buenos Aires, Argentina*, en: Actas resúmenes *VIII Congreso de ALHSUD*, Asunción, 2006, p. 29. Trabajo en Cd.

[9] R. Barranquero, A. Ruiz de Galarreta, y R. Banda Noriega. Cuadernos del CURIHAM-FCEIA (UNR). **2006**, *12*, 65-75.

[10] C. W. Thornthwaite y J. R. Mather, *Instructions and tables for computing potential evapotranspiration and the water balance*, Laboratory of Climatology, Publication núm. 10. Centertown, N. J. 1957.

[11] Soil Survey Staff, *Soil Taxonomy*, Second Edition NRCS-USDA, Agric. Handbook 436, US., Washington DC, 1999.

[12] H.G. Marchese y E. Di Paola. *Revista de la Asociación Geológica Argentina*. **1975**, 30 (1).

[13] M.E. Teruggi, *Journal of Sedimentary Petrology*, **1957**. *27*, 322-332.

[14] J.M. Sayago, M.M. Collantes, A. Karlson y J. Sanabria. *Quaternary International*, **2001**, *76/77*, 247-25.

[15] Ruiz de Galarreta A., Varni M., Banda Noriega, R. y Barranquero, R., *Caracterización geohidrológica preliminar en la cuenca del arroyo Langueyú, Partido de Tandil, Buenos Aires*, en: Actas *V Congreso Argentino de Hidrogeología*, Asociación Internacional de Hidrogeólogos, Entre Ríos, 2007.

[16] E. Custodio y M. Llamas, *Hidrología Subterránea*. Tomos I y II, ed. Omega. Barcelona, 1983.

[17] G. Castany, *Prospección y explotación de las aguas subterráneas*, ed. Omega, Barcelona, 1975.

[18] C.A.J. Appelo y D. Postma, *Geochemistry, groundwater and pollution*, ed. A. A. Balkema, 1992.

[19] APHA, *Standard Methods for the Examination of Water & Wastewater*, 21st ed., APHA, ISBN: 0875530478 / ISBN-13: 9780875530475, 2005.

[20] Statsoft, Inc. *Statistica for Windows [Computer program manual]*. Tulsa, OK: StatSoft, Inc., 2300 East 14th Street, Tulsa, WEB: http://www.statsoft.com. 1995.

[21] J. V. Crisci y M. F. López Armengol, *Regional Program for Scientific and Technological Advice*. Series of biology, Monograph 26, 1983.

[22] C. A. Núñez Colin, J. E. Rodríguez Pérez, R. Nieto Ángel y A. F. Barrientos Priego. *Revista Chapingo. Serie horticultura*. **2004**, V10 N2.

[23] H.H. Harman, *Modern factor analysis*, 3rd ed., The University of Chicago Press, Chicago, 1976.

[24] K.G. Joreskog, J.E. Klovan and R.A. Reyment. *Geological factor analysis*, Elsevier, New York, 1976.

[25] J.E. Klovan. *Journal of Sedimentary Research*. **1966**, *36*; No. 1.

[26] J. Catalán Lafuente, *Química del agua*, ed. Blume, Barcelona, 1969.