Available online at www.sciencedirect.com

SCIENCE DIRECT®

Analytica Chimica Acta 515 (2004) 143–149

ANALYTICA CHIMICA ACTA

www.elsevier.com/locate/aca

Using chemometric tools to assess anthropogenic effects in river water A case study: Guadalquivir River (Spain)

Carolina Mendiguchía, Carlos Moreno*, Mª Dolores Galindo-Riaño, Manuel García-Vargas

Department of Analytical Chemistry, Faculty of Marine and Environmental Sciences, University of Cádiz, Puerto Real, Cádiz 11510, Spain

Received 29 August 2003; received in revised form 20 November 2003; accepted 30 January 2004

Available online 21 April 2004

Abstract

The Guadalquivir River is the major water source of the south of Spain. In this work, we have selected an area, about 115 km long, from Alcalá del Río to the mouth of the river, to study the effects caused by human activities on water quality. The area under study includes several sensible points as the town of Seville, the Guadiamar River (where an important mine spill took place on April 1998) and the Doñana National Park.

We measured several physico-chemical variables in the 26 sampling stations located along the river, and in three different campaigns from 2001 to 2002. With the results we built the data matrix, which was analysed by factor analysis/principal components analysis (FA/PCA) and cluster analysis (CA). This analysis allowed the identification of four different zones in the river, with different water quality. The first zone (zone 1A) comprised from Alcalá del Río to Seville. The second zone (zone 1B) was the city of Seville, and as a consequence, presented higher concentrations of several variables such as nitrite, ammonium or manganese. The third zone (zone 2) included from Seville to the Guadiamar River. In this area, agriculture is the main activity, and then, higher concentrations of suspended solids and phosphate were measured. In terms of water quality, this zone was partially similar to zone 1A, and partially similar to the fourth zone, starting in the Guadiamar River and finishing in the mouth of the Guadalquivir River. The water in this last zone (zone 3) is mainly estuarine water. Thus, its quality is influenced by seawater input, and also by the inputs from the Guadiamar River (coming from a mining area), and presented higher copper concentration.

Three principal components were extracted, explaining the 79.1% of the data variance. PC1 (46.9% variance) was mainly associated with nitrite, ammonium and manganese. PC2 (22.5% variance) was mainly associated with suspended solids and phosphates. PC3 (9.7% variance) was mainly correlated to nitrate and copper concentration.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Guadalquivir River; Principal components analysis; Cluster analysis; Estuarine water; Anthropogenic effects

1. Introduction

The Guadalquivir River flows in a south-westerly direction, almost across the whole of the Spanish region of Andalusia, being, together with its tributaries, the main water source of this region with more than 7 millions inhabitants. As a consequence of such a high population, the river receives many inputs, from both natural and anthropogenic origin, that may cause deterioration of water quality. Besides, more than 700,000 has of its basin are devoted to agriculture, with very high production of rice, olives or fruits, with the corresponding environmental effects in the river [1]. This is especially significant in the river's lower basin, between Seville and the mouth, where the most agriculturally productive part of Andalusia is located.

Despite its importance, no global study has been performed on the environmental effects caused by human activities in the river. Some authors have described local effects caused in the river [2] or groundwater [3,4] by nutrients and pesticides used in agriculture. Other studies have been focused on the concentration of heavy metals in the Gulf of Cádiz, where Guadalquivir River joins Atlantic Ocean [5]. These studies have increased in number after the mine spill occurred in 1998 in Aznalcóllar [6,7].

In the present work, we have done a complete study of the water quality of Guadalquivir River, from Seville to the mouth. Besides, we have used the power of chemometrics to establish the effects caused by the different human activities performed in the margins of the river.

---

* Corresponding author. Tel.: +34-956-016433; fax: +34-956-016040.
E-mail address: carlos.moreno@uca.es (C. Moreno).

0003-2670/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2004.01.058

2. Materials and methods

2.1. Study site

The study was done in the navigable part of the Guadalquivir River (SW Spain), from Alcalá del Río to the mouth of the river (Fig. 1). At the investigated section, with a length of about 115 km, the river flows in direction Northeast-Southwest. Along its course, about 15 km downstream the starting point of the study, the river passes through the town of Seville, with a population of more than 700,000 inhabitants. It is the main urban and industrial centre of the zone, although both industrial and domestic wastewaters are treated before discharging into the river. Although some small villages (with their corresponding effects) may be found between Seville and the estuary (beginning in sampling station G14), agriculture and related industries are the main activity carried out in this zone. Finally, in the estuarine zone the Guadalquivir River participates in the Doñana National Park, giving an added ecological value to this area.

Up to 26 sampling sites were selected to cover the study (see Fig. 1). To select the location of the sampling stations we used two different criteria. In the first part of the river (stations G1–G13) we used a criterion based on geographical considerations and in the activities developed in the riversides. The location of these stations was fixed for all the sampling campaigns. On the other hand, in the estuarine zone (stations G14–G26), we used a criterion based on water salinity, which varies with tide situation, to cover the entire estuary. For this reason, the location of stations was variable in this zone.

2.2. Sampling

A total of 59 water samples were collected along the river during three sampling expeditions, in October 2001, March 2002, and October 2002. Table 1 indicates the stations sampled in each campaign.

Samples were collected using LDPE bottles, which were pre-cleaned with 6N HCl, rinsed with 10% HNO_3, and stored in polyethylene bags until required. In each sampling station, three different samples were taken in duplicate, to analyse suspended solids, nutrients (nitrate, nitrite, ammonium, and phosphate), and metals (copper and manganese), respectively. Except of those used for suspended solids, all of them were filtered on-line through 0.45 μm pore-size acid-washed polypropylene Calyx Capsule (Osmonics, USA).

Samples for the analysis of suspended solids were processed within 24 h of collection, while samples for nutrients and metals analysis were frozen and acidified with nitric acid, respectively, until analysis.

2.3. Analytical procedures

An electrochemical portable device (Sension 156, Hach Co., USA) was used for in situ measurements of T, pH and conductivity/salinity. Suspended solids (SS) were gravimetrically quantified by using 0.45 μm nylon filters (Osmonics, USA). To analyse nutrients, different spectrophotometric

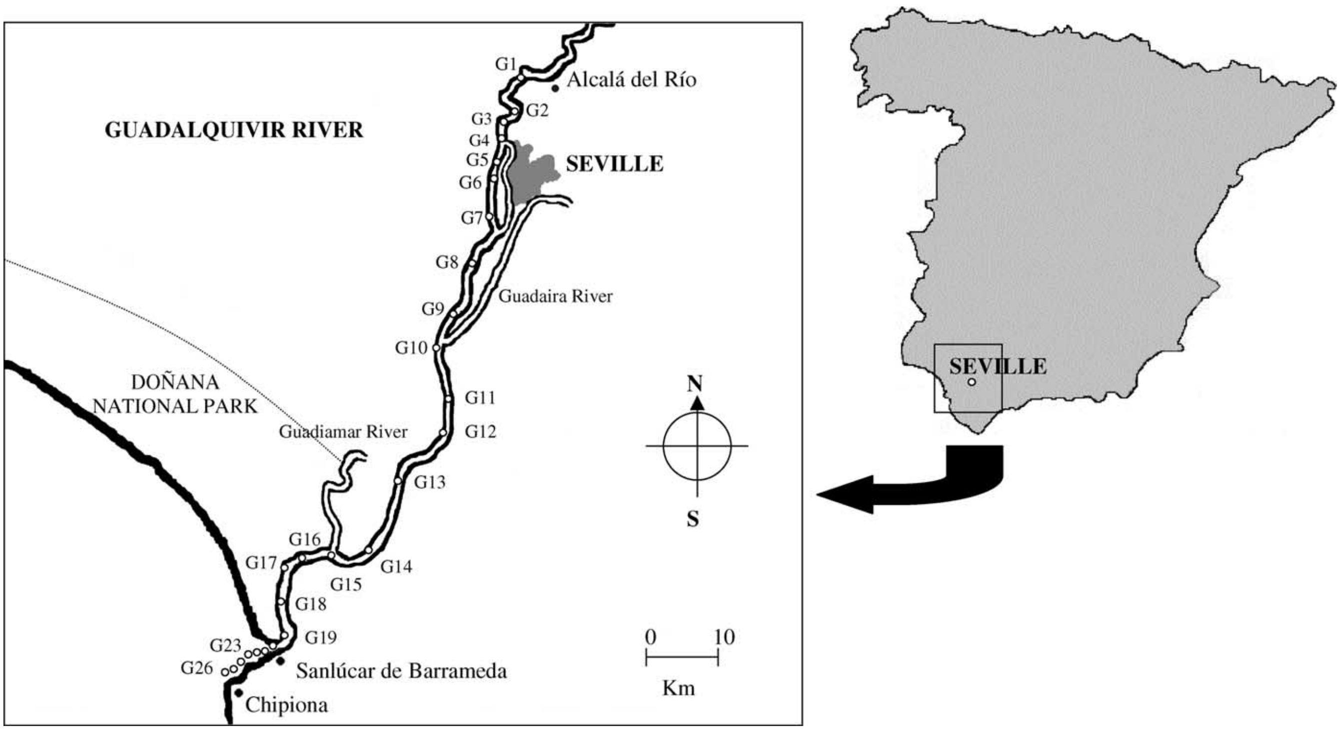

Fig. 1. Location of Guadalquivir River and water sampling stations.

Table 1
Sampling stations studied during each sampling campaign in the Guadalquivir River

Sampling campaign[a]	Sampling station[b]																									
	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26
I	X	X	X	X	X	X	X	X	X	X	X	X	X	–	X	–	X	–	X	–	X	–	–	X	X	–
II	X	X	X	X	X	X	X	X	X	X	X	X	X	–	–	–	X	–	–	–	X	X	–	–	X	X
III	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	–	X	–	X	–	X	X	X	X	–

[a] I: October 2001 (19 samples collected), II: April 2002 (18 samples collected), III: October 2002 (22 samples collected).
[b] X: Sample collected, –: sample not collected.

Table 2
Methodology applied to the quantification of nutrients

Nutrient	Method
$N\text{-}NH_4^+$	Spectrophotometry (indophenol blue)
$N\text{-}NO_2^-$	Spectrophotometry (sulphanilamide/*N*-(1-naphthyl)-ethylenediamine dihydrochloride)
$N\text{-}NO_3^-$	Spectrophotometry (reduction in a Cd column and nitrite method)
$P\text{-}PO_4^{3-}$	Spectrophotometry (ammonium molybdate/ascorbic acid)

standard methods were applied (see Table 2) [8–11]. Copper and manganese concentrations were determined by the standard addition method in a flameless atomic absorption spectrometer (Model SolaarM, Thermo Electron, USA).

2.4. Statistical procedures

In principal components analysis (PCA), eigenanalysis of the experimental data was performed to extract principal components (PCs), using two selection criteria: the scree plot test and corrected average eigenvalue. For hierarchical cluster analysis (CA), the squared Euclidean distance between normalised data was used to measure similarity between samples. Both average linkage between groups and Ward's method were applied to standardised data and the results obtained were represented in a dendrogram [12].

Statistical treatments were performed by using both Statistica for Windows 5.1 (StatSoft, Inc., USA) and SPSS for Windows 11.0.1 (SPSS Inc., USA) software packages.

3. Results and discussion

3.1. Spatial variations

The results obtained for the spatial variation of the studied variables are shown in Figs. 2 and 3 in the form of box-whiskers plots. The results for each station include all the data collected for the three sampling expeditions, since the absence of significant temporal variation was confirmed by analysis of variance (ANOVA). Salinity and temperature of each sample were also measured, but they were not plotted since they did not provide any relevant information to this study. For the same reason they were not included in the data treatments reported below.

As observed, there are several variables (Mn^{2+}, $N\text{-}NO_2^-$ and $N\text{-}NH_4^+$) that present a maximum between sampling stations G5 and G8, where the town of Seville was located. Similar behaviour was observed for $N\text{-}NO_3^-$, but it also presents very high concentrations in the initial part of the study site (sampling stations G1–G3). In the case of $P\text{-}PO_4^{3-}$ and SS, maximum concentrations were measured in the central part of the river, between stations G9 and G15, where agriculture is the main activity carried out.

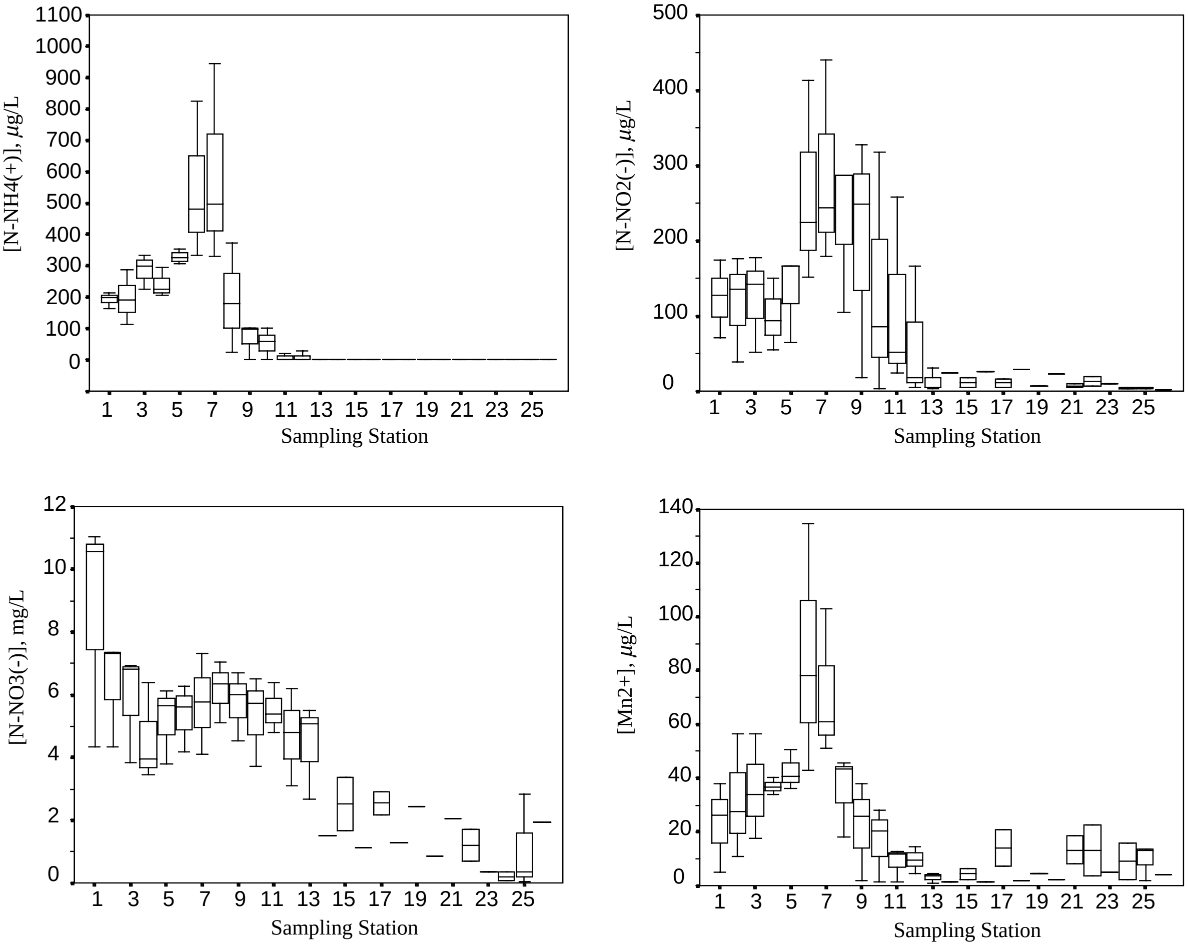

Fig. 2. Spatial variations of ammonium, nitrite, nitrate, and manganese in the sampling stations of Guadalquivir River.

Particular behaviours were observed for pH and Cu^{2+}. The former exhibited an increase starting about sampling stations G13–G14, which was mainly caused by the influence of the seawater input in the estuarine zone. Maximum concentrations of copper were measured between sampling stations G14–G16, were Guadiamar River joins Guadalquivir River. Guadiamar River comes from the area of Riotinto, were several mines are located, and its water typically contents high concentration of several heavy metal such as copper.

3.2. Principal components analysis

To select the number of components to be extracted we used both Jolliffe's method of corrected average eigenvalue [13] and Cattel's Scree plot test (see Fig. 4) [14]. In both cases we retained three PCs and, as a consequence, 79.1% of the variance of the original data may be explained.

In Fig. 5 we have represented both the scores of samples and loadings of variables, corresponding to PC1 and PC2. Each sample is identified by the name of the corresponding sampling station plus a superscript (I, II or III) indicating the sampling campaign. As can be observed, samples sites are classified in three different groups. First group corresponds to the zone with higher urban influence (zone 1). This zone can be divided in two sub-zones. On the one hand, zone 1A comprises from the town to Alcalá del Río to Seville (sampling stations G1–G5) and its water quality is controlled by the concentrations of nitrogen-related nutrients and manganese. On the other hand, zone 1B includes the two sampling stations located closest to Seville (G6 and G7) and present very high concentrations of the same chemical species.

The sampling stations included in zone 2 are those located from Seville to the beginning of the estuary (G8–G13). In this area, the margins of the river are used mainly for agriculture, and this fact probably conditions that water quality is mainly controlled by concentration of phosphates and suspended solids.

The zone 3 includes all the sites sampled in the estuary of the river, which begins in the mouth of Guadiamar River (G14) that, as mentioned before, has typically high copper concentrations. Thus, the water characteristics of this zone are mainly controlled by copper concentration and by pH, as a consequence of the decreasing acidity caused by seawater input.

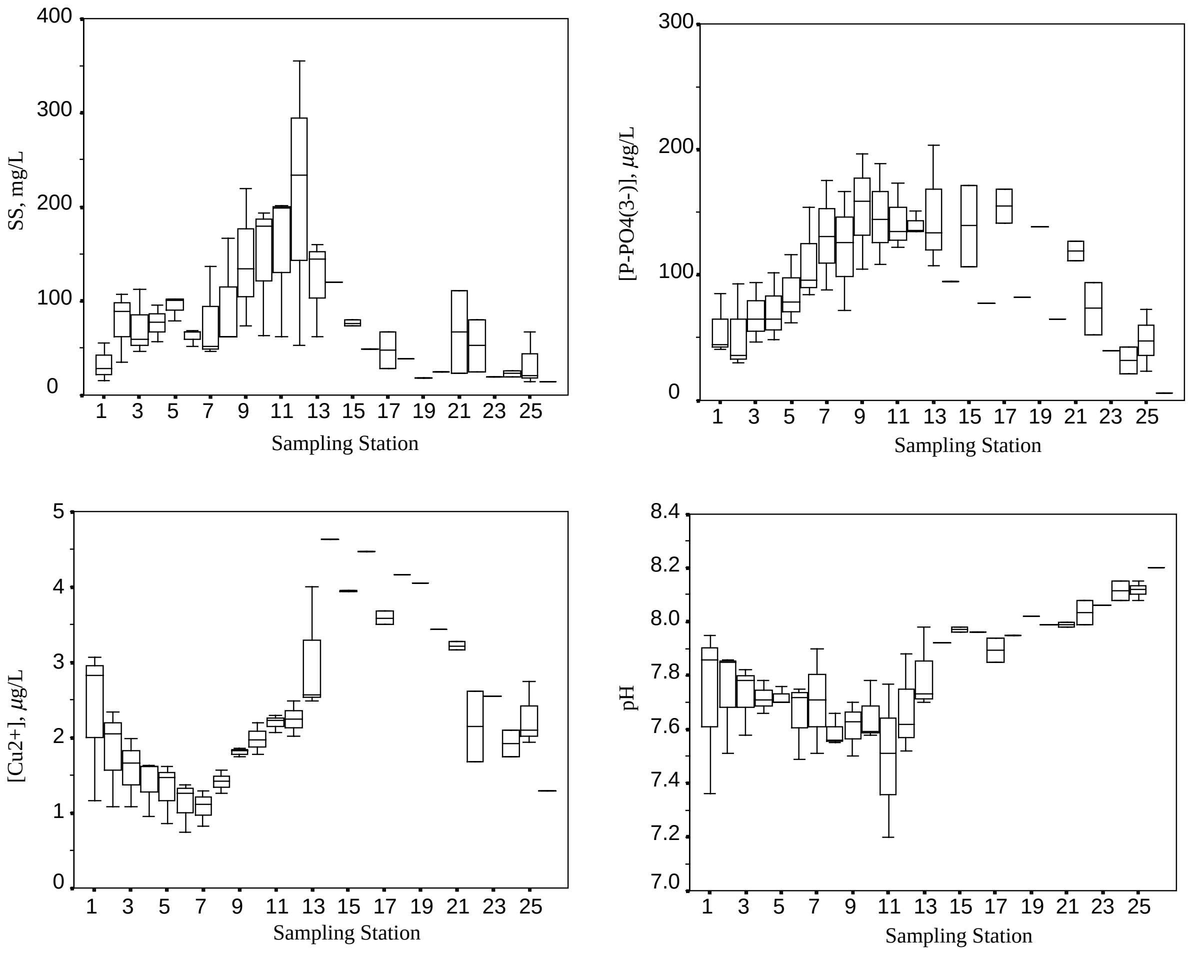

Fig. 3. Spatial variations of suspended solids, phosphates, copper, and pH in the sampling stations of Guadalquivir River.

Note that the borders of the different zones are not clearly marked and then, some small alterations may appear, probably as a consequence of tidal influences, e.g. in the first sample campaign, an upstream shift was observed in the first sample campaign, and then, samples collected in stations G6–G10 presented characteristics similar to those found in zone 1A. Similar effect may be observed in the separation of zones 2 and 3, which may be established between stations G13 and G14. Thus, while G13 was located in zone 2 in campaigns I and II, it was in zone 3 in campaign III.

These results may be complemented by examination of the loadings of the three retained PCs (see Table 3). PC1 (46.9% of the variance) is highly and positively contributed by the variables related to urban activities (NO_2^-, Mn^{2+}, NH_4^+ and NO_3^-), and negatively by Cu^{2+} and pH. PC2 (22.5% of the variance) is mainly contributed by the variables affected by agriculture activities (PO_4^{3-} and SS), while PC3, that explains only 9.7% of the variance, is participated mainly by NO_3^- and other variables like Mn^{2+}, PO_4^{3-} and Cu^{2+}.

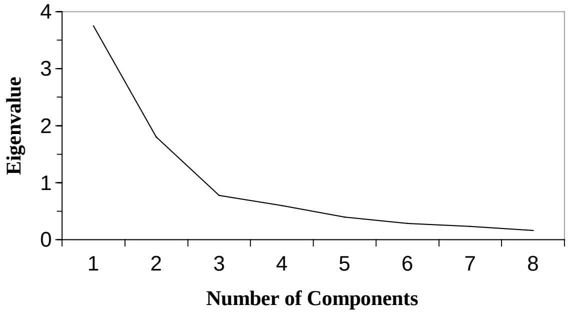

Fig. 4. Scree plot of the eigenvalues of principal components.

Table 3
Loadings of experimental variables on the first three principal components

	PC1	PC2	PC3
Cu^{2+}	−0.724	−0.317	0.359
Mn^{2+}	0.810	0.219	0.401
SS	0.190	−0.800	−0.296
NO_3^-	0.738	−0.137	−0.470
NO_2^-	0.866	0.160	0.133
NH_4^+	0.781	0.302	0.126
PO_4^{3-}	0.233	−0.825	0.379
pH	−0.757	0.444	−0.048
Eigenvalue	3.749	1.801	0.778
Variance (%)	46.9	22.5	9.7

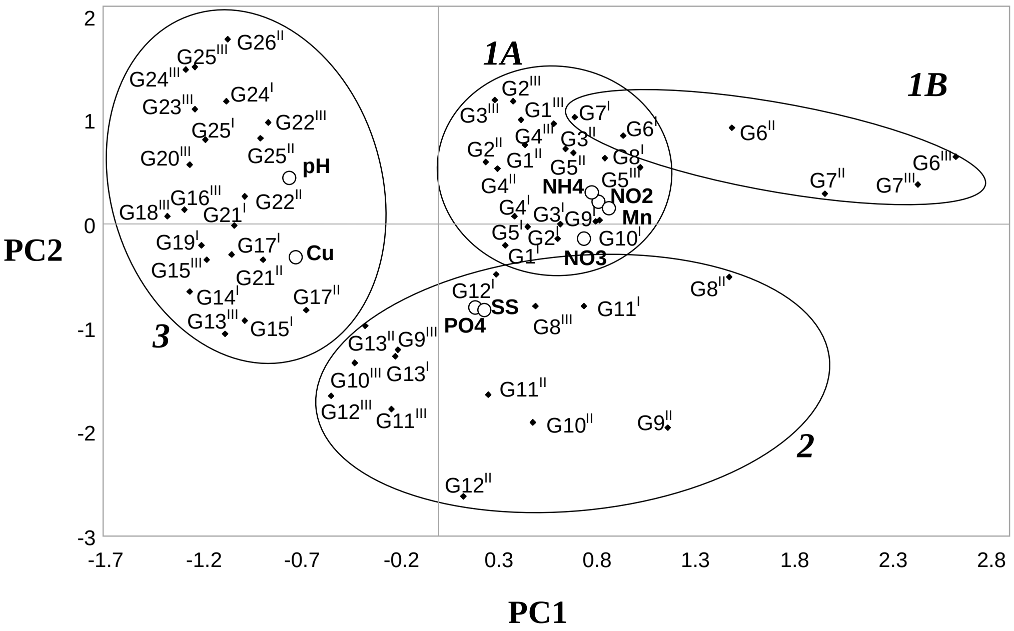

Fig. 5. Scores of river water samples (●) and loadings of variables (○) on the plane defined by the first two principal components.

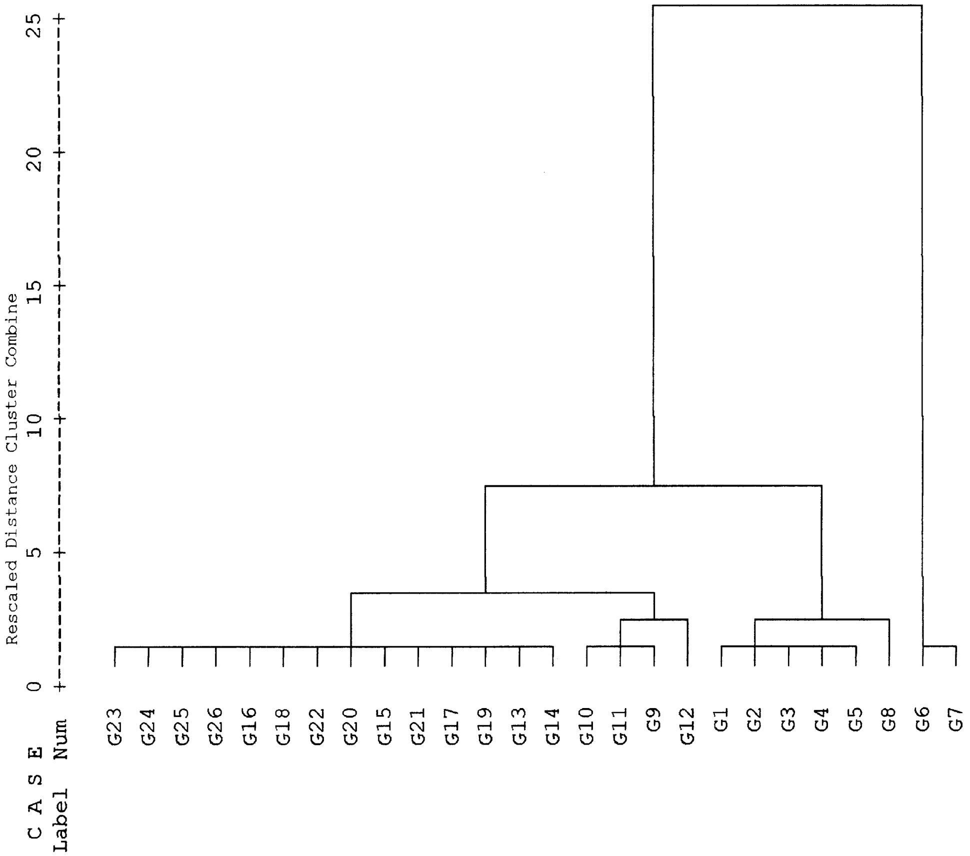

Fig. 6. Dendrogram of the 26 sampling stations studied in the Guadalquivir River.

Table 4
Loadings of experimental variables on the first three rotated principal components

	Varifactor 1	Varifactor 2	Varifactor 3
Cu^{2+}	−0.495	0.199	−0.685
Mn^{2+}	0.920	0.083	0.106
SS	−0.263	0.736	0.390
$NO_3{}^-$	0.300	0.210	0.806
$NO_2{}^-$	0.808	0.094	0.363
$NH_4{}^+$	0.784	−0.063	0.314
$PO_4{}^{3-}$	0.108	0.920	−0.145
pH	−0.479	−0.617	−0.403
Eigenvalue	2.761	1.870	1.696
Variance (%)	34.5	23.4	21.2

A varimax rotation allows a better and more explicit assignment of experimental variables to PCs. As seen is Table 4, varifactor 1 is mainly contributed by Mn^{2+}, $NO_2{}^-$, and $NH_4{}^+$, and then, may be related to urban pollution. Varifactor 2 is mainly contributed by $PO_4{}^{3-}$ and SS and then, may be related to the effects caused in the river by agriculture activities. Finally, varifactor 3 is participated by variables originated from different sources, as $NO_3{}^-$ or Cu^{2+}.

3.3. Cluster analysis

Hierarchical cluster analysis was used to group sampling stations based on the similarity of the water chemical composition. Both average linkage between groups and Ward's method were applied to normalised data using squared Euclidean distance to quantify similarity between stations. The classification patterns obtained by both methods were very similar. Fig. 6 shows the dendrogram obtained by average linkage between groups method. Three very well differentiated clusters can be observed. The one placed on the right groups sampling stations G6 and G7, corresponding to the town of Seville, which are clearly differentiated from the rest or the river. The second cluster contents the stations located in the vicinity of Seville, and slightly differentiates the stations locates upstream (G1–G5) and downstream (G8) Seville. These two clusters correspond to the zones of the river denoted as 1B and 1A, respectively.

The third cluster (left in Fig. 6) includes most of the samples taken between Seville and the mouth, and also differentiates in different sub-clusters the estuary (G13–G26, zone 3) and the sampling stations of zone 2 located close to estuarine area (G9–G12).

4. Conclusions

The water quality of Guadalquivir River has been established from Seville to the mouth. PCA and CA reinforced the results obtained by each other, confirming the existence of three zones clearly separated (zones 1A, 1B and 3), and a transition zone (zone 2), with water characteristics between those obtained in zones 1A and 3. The human activities mainly affecting the river are urban and industrial wastewaters, especially in zones 1A and 1B, and agriculture, especially in zone 2.

Acknowledgements

This study has been financed by the Spanish Ministerio de Educación y Cultura (REN 2000-1101). The authors gratefully acknowledge the collaboration of the Consejería de Medio Ambiente (Junta de Andalucía) in sampling campaigns.

References

[1] Atlas Agrario y Pesquero de Andalucía, Consejería de Agricultura y Pesca, Junta de Andalucía, Sevilla, 1992.
[2] M. Espigares, C. Coca, M. Fernandezcrehuet, O. Moreno, A. Bueno, R. Galvez, Environ. Toxicol. Water Qual. 12 (1997) 249.
[3] J.E. Fernandez, J.M. Murillo, F. Moreno, F. Cabrera, E. Fernandezboy, Commun. Soil Sci. Plant 29 (1998) 2829.
[4] M. Villalba, F.J.B. Delarosa, M.T. Rodríguez, J.C.J. S'anchez, Water Res. 29 (1995) 1361.
[5] C. Braungardt, E.P. Achterberg, M. Nimmo, Anal. Chim. Acta 377 (1998) 205.
[6] E.P. Achterberg, C. Braungardt, N.H. Morley, F. Elbatz-Poulichet, M. Leblanc, Water Res. 33 (1999) 3387.
[7] E. Garcia-Luque, J.M. Forja, T.A. Delvalls, A. Gomez-Parra, Environ. Monit. Assess. 83 (2003) 71.
[8] K. Grasshoff, M. Ehrhardt, K. Kremling, Methods of Seawater Analysis, 2nd ed., Verlag Chemie, Weinheim, 1983.
[9] J. Rodier, Análisis de las Aguas (Spanish Edition), Omega, Barcelona, 1990.
[10] A. Tovar, C. Moreno, M.P. Mánuel-Vez, M. García-Vargas, Quim. Anal. 20 (2001) 37.
[11] A. Tovar, C. Moreno, M.P. Mánuel-Vez, M. García-Vargas, Anal. Chim. Acta 469 (2002) 235.
[12] G. Ramis Ramos, M.C. García Alvárez-Coque, Quimiometría (in Spanish), Síntesis Madrid, 2001.
[13] I.T. Jolliffe, J. R. Stat. Soc. C. Appl. 21 (1972) 160.
[14] R.B. Cattel, Multivar. Behav. Res. 1 (1966) 245.