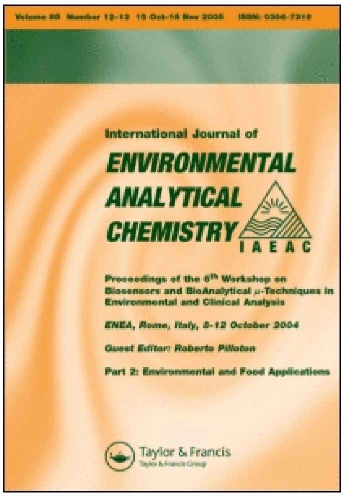

International Journal of Environmental Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

Factor Analysis for Assigning Sources of Groundwater Pollution

G. Rauret[a]; M. T. Galceran[a]; R. Rubio[a]; F. X. Rius[b]; M. S. Larrechi[b]
[a] Department of Analytical Chemistry, University of Barcelona, Barcelona, Spain [b] Faculty of Chemistry of Tarragona, University of Barcelona, Tarragona, Spain

To cite this Article Rauret, G. , Galceran, M. T. , Rubio, R. , Rius, F. X. and Larrechi, M. S.(1990) 'Factor Analysis for Assigning Sources of Groundwater Pollution', International Journal of Environmental Analytical Chemistry, 38: 3, 389 — 397

To link to this Article: DOI: 10.1080/03067319008026943
URL: http://dx.doi.org/10.1080/03067319008026943

Intern. J. Environ. Anal. Chem., Vol. 38, pp. 389–397
Reprints available directly from the publisher
Photocopying permitted by license only

FACTOR ANALYSIS FOR ASSIGNING SOURCES OF GROUNDWATER POLLUTION*

G. RAURET, M. T. GALCERAN and R. RUBIO

Department of Analytical Chemistry, University of Barcelona, Barcelona, Spain

F. X. RIUS and M. S. LARRECHI

Faculty of Chemistry of Tarragona, University of Barcelona, Tarragona, Spain

(*Received 6 July 1989*)

Factor Analysis (FA) is applied to tentatively establish either the sources of pollution in the groundwater in a delimited zone in Catalonia which has been polluted by permeation with non-controlled wastewaters or the groundwater evolution with time in this area. Three types of chemical analysis were considered: general parameters indicating the chemical quality of water, inorganic micropollutants and organic micropollutants. These analyses were carried out with forty-five samples collected in three different periods, and eleven parameters were considered. From the final solutions of the FA method applied several factors were retained for the different sampling periods and these factors were related to different sources of pollution. FA has proved to be more useful when punctual pollution incidents takes place rather than when groundwater evolution with time is studied.

KEY WORDS: Factor analysis, sources, groundwater pollutants, cyanide, organic pollutants, gas chromatography.

INTRODUCTION

At the beginning of 1987 an episode of groundwater pollution was observed at the municipality of Les Franqueses del Vallès (Catalonia, Spain). This pollution caused problems to the inhabitants of the area, specially farmers, since the zone is mainly agricultural and the groundwater is used to irrigate the fields. In a previous study[1] carried out at the Analytical Chemistry Department of the University of Barcelona it has been concluded that a cause–effect relationship between some organic micropollutants found in the groundwater samples analysed and some (non-controlled) wastewater from a factory located in this zone exists.

Once the zone of study has been delimited and the sampling points have been established in cooperation with the Environmental Service of Diputació de Barcelona, the sampling was performed in three different periods, in order to acquire knowledge on the evolution of the pollution with time. Three types of

---

*Presented 5 September 1988 at the 18th International IAEAC Symposium on Environm. Analytical Chemistry, Barcelona, Spain.

chemical analyses were carried out involving the determination of: 1st, general parameters indicating the chemical quality of water, 2nd, inorganic micropollutants and 3rd, organic micropollutants. From the results found, eleven features or compounds were chosen as representatives of the chemical groundwater pollution. Chemical criteria and the fact that these parameters are in a measurable concentration in a large part of the analysed samples have been taken into account to select these parameters.

Finally chemometrics has been applied to the results obtained, in order to find the possible sources of groundwater pollution. Recently, pattern recognition methods have been used to classify several types of surface waters,[2,3] to study the influence of the sampling procedure on the features of the water samples,[4] to separate the areas characterized by different types of water quality or to determine the variation in the data.[5,6] Factor analysis techniques have been widely applied to describe air pollution in highly industrialized regions,[7,8] however, no references have been found relating to receptor modeling in aqueous environmental chemistry. In this paper classical Factor Analysis, FA,[9,10] has been applied in order to tentatively establish not only the sources of pollution but also the groundwater evolution with time in the studied area.

Factor analysis methods are well suited to describe these types of problems. In fact, the aims of FA are threefold:[9,10] (i) to identify the true dimensionality of the problem, which in the present study would mean to recognize the number of possible contamination sources, (ii) to measure the interrelationship between variables in a multivariate space (the correlation coefficient can only show a linear association between pairs of variables, therefore it is not well adapted to multidimensional data), and (iii) to supply a quantitative measurement of the importance of each of these actual effects, or pollution sources, on the observed parameter values.

EXPERIMENTAL

Forty-five samples corresponding to nineteen wells from Les Franqueses del Vallès (Figure 1) were analyzed. Samples were collected in three periods, namely during March, July and November of 1987. Samples, which were collected manually or by pumping depending on the well characteristics, and sampling dates are indicated in Table 1.

Eleven parameters, some of them general characteristic parameters of water quality as conductivity or organic matter, expressed as permanganate consumption, were considered. Among the inorganic compounds cyanide ion was chosen and among the organic pollutants the total amount of basic plus neutral compounds, and acid compounds were included. The concentration of some organic acids, phenols and other unusual substances indicating pollution in this particular case, were also considered. All these parameters were selected since they were present in almost all the studied samples in variable amounts, which could allow the distinction of different possible sources of pollution. Other variables such as chloride, pH, nitrogen (ammonia), MBAS (detergents expressed as

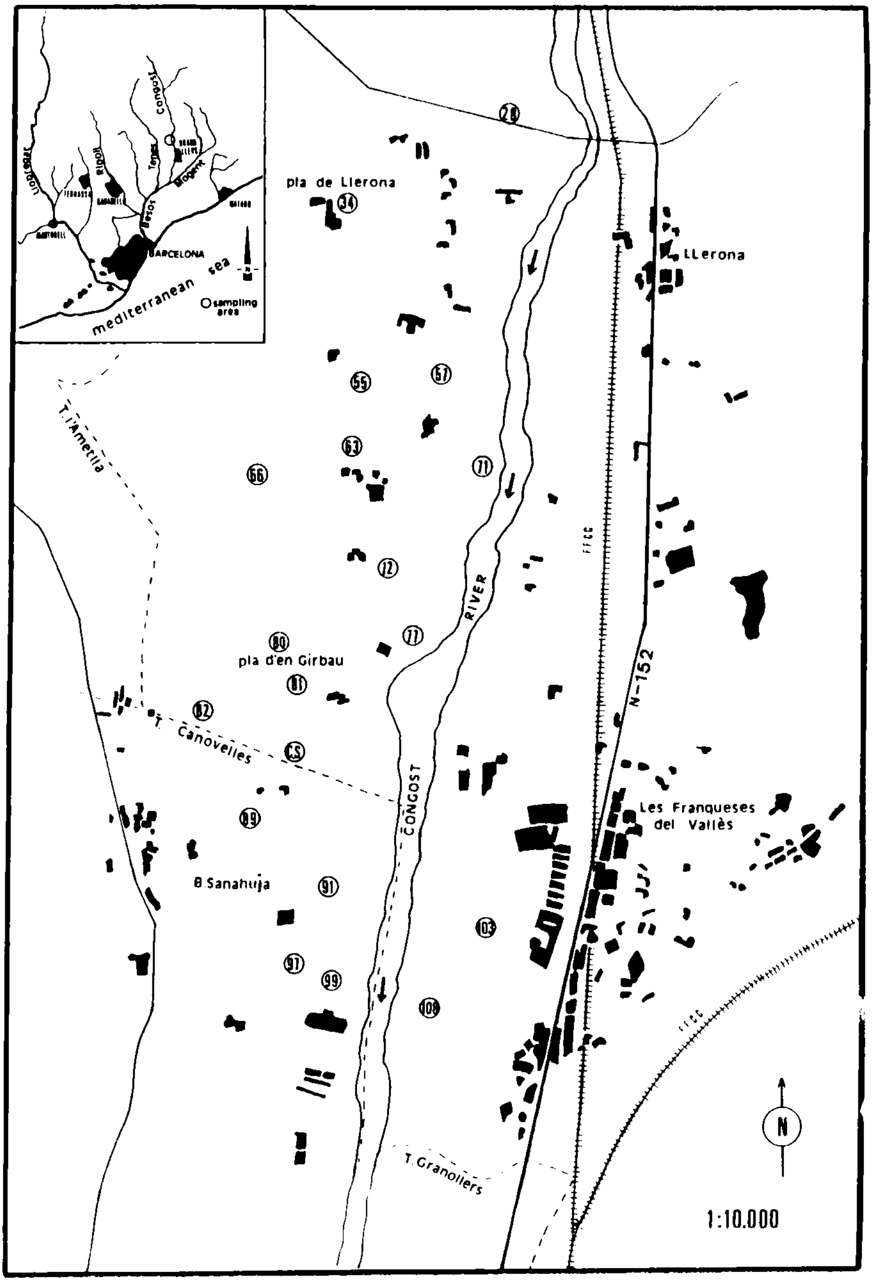

Figure 1 Location of sampling points in the studied area.

Table 1 Sampling and sampling data

Sample	*Date*	*Sample*	*Date*	*Sample*	*Date*
1-28[a]	16-3-1987	–	–	–	–
1-34	16-3-1987	2-34	6-7-1987	3-34	10-11-1987
1-55	16-3-1987	2-55	6-7-1987	3-55	10-11-1987
1-57	16-3-1987	2-57	6-7-1987	3-57	10-11-1987
–	–	2-63	1-7-1987	3-63	11-11-1987
1-66	16-3-1987	2-66	6-7-1987	3-66	11-11-1987
1-71	17-3-1987	2-71	6-7-1987	3-71	11-11-1987
1-72	16-3-1987	2-72	1-7-1987	3-72	10-11-1987
1-77	16-3-1987	3-77	1-7-1987	3-77	10-11-1987
–	–	2-80	6-7-1987	–	–
1-81	16-3-1987	2-81	6-7-1987	3-81	12-11-1987
1-82	16-3-1987	–	–	–	–
1-89	16-3-1987	2-89	1-7-1987	3-89	12-11-1987
1-91	16-3-1987	2-91	1-7-1987	3-91	11-11-1987
1-97	16-3-1987	2-97	1-7-1987	3-97	11-11-1987
–	–	2-99	1-7-1987	3-99	11-11-1987
1-103	16-3-1987	–	–	3-103	12-11-1987
–	–	2-108	1-7-1987	3-108	12-11-1987
1-CS	16-3-1987	–	–	–	–

[a] First figure indicates the period of sampling. Second figure indicates the well number.

methylene blue active substances), non filtrable residue, hexavalent chromium and other toxic heavy metals, iron and manganese as well as some organic compounds identified by gas chromatography/mass spectrometry were not considered as relevant parameters because of their own chemical nature or due to their low concentrations, below the detection limits. Therefore, although they have been analyzed, they have not been taken into account in the chemometric analysis.

The general parameters of water quality and the inorganic compounds were determined according to standard analytical methods.[11] The organic extractables were separated in two fractions, the first one obtained at pH 8 (basic plus neutral and weak acid compounds) and the second one obtained at pH 1 (acidic compounds). The two fractions were analysed by gas chromatography with FID detection and by gas chromatography/mass spectrometry. The results obtained are indicated in Table 2. For computation it is necessary to have a value for each parameter, we have used the minimum detectable amount, indicated as < in Table 2 for cyanide, benzocyano compounds, phenols and morpholine. Further details of the procedure and the results obtained for all measured parameters are indicated elsewhere.[1]

Calculations were performed using the well known pattern recognition package SPSS-X.[12] The program was run on an IBM 3083 of the University of Barcelona.

RESULTS AND DISCUSSION

Correlation of Variables

Correlation matrices obtained for the three sets of measurements reveal a different

Table 2 Analytical results for the groundwater samples

Sample	*Cond* $\mu mhos\ cm^{-1}$	MnO_4^- $mgL^{-1}O_2$	*B+N.Comp.* mgL^{-1}	*A.Comp.* μmgL^{-1}	CN^- μgL^{-1}	*ϕ-CN* %	*Phen* %	*Mi.Ac.* %	*Pa.Ac.* %	*Mor.* %	*St.Ac.* %
1-28	2617	7.07	0.026	0.046	<1	0.1	0.3	4.0	16.0	7.2	5.3
1-34	2197	3.40	0.070	0.030	<1	<0.06	<0.06	1.2	3.5	19.3	0.9
1-55	2187	8.48	0.018	0.080	3.1	<0.06	1.1	4.0	15.5	6.5	4.3
1-57	2292	3.52	0.017	0.061	<1	0.34	5.5	2.1	7.4	2.9	6.5
1-66	1900	2.44	0.023	0.120	<1	<0.06	0.6	2.2	10.2	3.5	3.8
1-71	1108	16.40	0.154	0.201	<1	4.2	1.9	1.9	3.4	<0.06	1.7
1-72	4250	62.30	0.225	0.175	147	16.5	2.2	3.0	2.2	1.4	0.5
1-77	5348	16.00	0.014	0.095	<1	0.06	0.6	3.9	15.7	1.7	2.6
1-81	2139	51.90	0.287	0.086	148	12.3	2.0	2.1	1.0	0.4	0.2
1-82	1337	1.88	0.008	0.087	<1	0.4	1.7	3.4	26.1	0.9	5.4
1-89	778	0.28	0.009	0.145	<1	<0.06	1.2	2.6	15.1	3.4	1.1
1-91	2841	11.90	0.058	0.044	19.2	3.9	10.0	0.7	3.9	0.2	1.0
1-97	2006	2.68	0.038	0.330	<1	1.3	1.3	3.1	11.9	0.5	1.3
1-103	1261	0.72	0.019	0.042	<1	1.2	1.4	4.9	20.5	0.6	2.7
1-CS	2282	28.00	0.051	0.053	14.0	0.2	2.7	2.9	4.3	0.9	20.6
2-34	3060	6.52	0.008	0.044	<1	0.6	0.2	8.6	24.0	0.9	11.7
2-55	3110	7.36	0.012	0.128	<1	0.6	0.3	9.1	23.7	3.1	7.4
2-57	4030	8.83	0.014	0.040	<1	2.8	0.3	5.9	20.6	2.6	8.3
2-63	3240	18.95	0.461	0.051	49	2.6	6.0	4.3	3.0	7.9	7.5
2-66	1650	2.91	0.013	0.093	<1	0.6	0.3	12.4	30.8	0.8	8.0
2-71	1016	14.34	0.015	0.139	<1	<0.06	1.5	0.1	9.2	0.8	1.5
2-72	7000	14.81	0.024	0.019	10	0.1	5.8	7.2	8.1	14.5	10.0
2-77	5580	18.07	0.009	0.026	<1	2.4	0.2	7.0	21.6	1.0	14.9
2-80	1780	3.41	0.016	0.047	<1	3.3	0.4	6.4	17.8	1.5	5.1
2-81	2930	8.50	0.126	0.050	30	0.8	9.8	7.4	5.9	14.3	7.2
2-89	1017	2.53	0.009	0.030	<1	1.7	0.3	7.7	30.2	1.3	7.7
2-91	3000	9.00	0.026	0.050	<1	10.2	8.7	2.2	10.2	10.8	1.7
2-97	2820	7.15	0.011	0.033	<1	0.9	0.6	6.8	21.2	3.8	8.7
2-99	4210	12.23	0.009	0.040	<1	0.6	2.3	6.9	17.6	3.9	3.1
2-108	3200	13.68	0.080	0.090	<1	7.6	4.0	4.4	3.8	12.7	8.4
3-34	2440	25.22	0.003	0.050	<1	<0.06	<0.06	2.7	11.9	1.2	0.2

Table 2 (continued)

Sample	*Cond* $\mu mhos\,cm^{-1}$	MnO_4^- $mgL^{-1}O_2$	*B + N. Comp.* mgL^{-1}	*A. Comp.* μmgL^{-1}	CN^- μgL^{-1}	*ϕ-CN* %	*Phen* %	*Mi. Ac.* %	*Pa. Ac.* %	*Mor.* %	*St. Ac.* %
3-55	2370	9.18	0.007	0.087	4	<0.06	0.9	0.7	7.5	0.8	2.7
3-57	2290	15.61	0.007	0.106	<1	<0.06	0.3	0.6	15.1	0.6	2.6
3-63	2710	18.95	0.073	0.050	<1	<0.06	3.5	8.5	0.6	8.5	7.5
3-66	2010	20.39	0.007	0.085	<1	<0.06	<0.06	4.2	4.8	2.5	2.1
3-71	685	28.83	0.024	0.111	<1	<0.06	0.5	2.8	6.4	0.5	3.3
3-72	3380	43.98	0.021	0.123	<1	<0.06	0.2	1.1	12.8	11.8	5.9
3-77	5810	24.19	0.004	0.156	<1	<0.06	6.9	1.6	8.1	0.4	0.2
3-81	2570	23.85	0.030	0.110	<1	<0.06	2.8	2.2	6.9	20.0	13.3
3-89	1620	23.78	0.019	0.144	1	<0.06	0.1	3.8	6.3	1.2	14.2
3-91	2800	25.83	0.064	0.091	<1	<0.06	0.4	5.5	5.8	8.0	2.7
3-97	2400	28.37	0.011	0.101	8	<0.06	0.3	12.7	7.1	9.7	0.6
3-99	3553	27.69	0.018	0.098	4	<0.06	1.1	5.3	7.2	3.2	0.5
3-103	1360	26.47	0.013	0.096	<1	<0.06	0.06	2.2	3.3	2.9	1.1
3-108	2590	28.18	0.024	0.050	<1	<0.06	0.3	2.1	9.5	15.6	6.4

Cond: conductivity; MnO_4^-: permanganate consumption expressed as O_2; B + N. Comp.: organic basic plus neutral compounds; A. Comp.: organic acid compounds; CN^-: cyanides; ϕ-CN: isocyanotoluene (tentative); Phen: 2,4 bisdimethylethylphenol; Mi. Ac: miristic acid; Pa. Ac: palmitic acid; Nor.: 4 (methylphenyl) sulfonylmorpholine; St. Ac.: stearic acid.

Table 3 Varimax rotation of principal factors

Specie	*Set 1*				*Set 2*				*Set 3*				
	a_1	a_2	a_3	a_4	a_1	a_2	a_3	a_4	a_1	a_2	a_3	a_4	a_5
Cond.	0.36	−0.01	0.06	0.04	0.17	−0.02	*0.89*	0.34	*0.82*	−0.21	0.13	0.01	0.07
MnO_4^-	*1.16*	−0.02	0.13	0.11	0.09	0.34	0.25	*0.94*	0.04	−0.17	0.53	−0.02	0.09
B + N Comp.	0.60	−0.23	0.07	0.63	0.17	*0.75*	−0.03	0.23	0.11	0.65	0.36	−0.06	−0.31
A Comp.	0.26	−0.40	*0.75*	−0.01	−0.09	−0.01	−0.57	0.07	0.23	−0.13	0.09	−0.05	*0.93*
CN^-	*0.77*	−0.17	0.11	0.59	0.12	*1.15*	0.08	0.06	−0.08	−0.09	−0.04	*0.73*	0.05
ϕ-CN	*0.69*	−0.34	0.25	*0.79*	0.53	−0.11	−0.12	0.13	−0.22	0.26	0.44	0.25	0.15
Phen.	0.01	−0.40	*0.73*	−0.07	*0.87*	0.43	0.07	0.04	*0.97*	0.32	−0.19	−0.10	0.19
Mi. Ac.	0.15	*1.05*	−0.16	−0.28	−0.49	0.04	0.42	−0.62	0.03	0.36	0.18	*0.88*	−0.16
Pa. Ac.	−0.57	*0.73*	−0.08	−0.06	−0.67	−0.43	0.08	−0.53	0.01	*0.76*	0.13	−0.23	0.06
Mor.	−0.11	−0.29	*−0.67*	−0.02	*0.81*	0.30	0.26	0.09	0.02	0.27	*0.84*	0.01	−0.09
St. Ac.	−0.01	0.07	0.08	−0.42	−0.34	0.06	*0.72*	0.06	−0.19	0.48	0.34	−0.32	0.27
% of total variance	43.2	19.1	13.8	9.4	39.9	21.2	13.4	10.2	24.0	18.3	15.8	13.8	9.9

behaviour in the association between pairs of variables in each set. The measurements carried out in March are more correlated than those carried out in July and November. In March, the basic plus neutral fraction is highly correlated to cyanides ($r = 0.89$), organic matter content ($r = 0.87$) and benzocyanocompounds ($r = 0.90$), the cyanides are highly correlated to organic matter content ($r = 0.93$) and benzocyano compounds ($r = 0.95$) and the organic matter content is related to benzocyano compounds ($r = 0.90$). In July these correlations disappeared and only the basic plus neutral fraction is highly correlated to another variable: cyanides ($r = 0.92$). In November not any pair of variables shows a strong relationship.

It is well known that high correlation coefficients do not necessarily imply causation between the variables, however, it might well be that the strong correlation found in March between basic plus neutral fraction, cyanides and benzocyano compounds indicate a common source for these pollutants.

Selection of the Method

Three different methods of extracting the initial factors, all of them available in subroutine FACTOR of SPSS-X, have been used. In all cases, the main diagonal of the correlation matrix has been replaced by some estimates of communality, obtaining in this way the principal factors, as a different choice to the principal components (PCA) computed when the main diagonal of the correlation matrix is not altered. The invariance of the resulting coefficient matrix obtained adds confidence in the overall solution of the problem. Several rotation procedures, two orthogonal as VARIMAX and QUARTIMAX and one oblique: OBLIQUE has been applied to each of the factoring methods. Although the final solutions do not significantly differ from one method to another, the VARIMAX-rotated factors have been preferred since their results are the easiest to interpret.

Number of Factors to be Retained

The eigenvectors having eigenvalues higher than 1.0 have been considered as significant factors. The VARIMAX-rotated loading matrix for the three data sets as well as the percentage of variance retained by each factor, are given in Table 3. Four factors were extracted for March and July data and five factors were retained for November data.

Interpretation of Principal Component Method

For the first set of samples (March, 1987), when the pollution problem appeared, the first component (with the largest eigenvalue) was mainly related to the organic matter content (expressed as permanganate consumption), cyanide ion and benzocyano, all of them highly correlated as it has been previously pointed out, the second factor was associated with fatty acids, the third with phenols and acidic organic compounds and the fourth with benzocyano compounds. These factors could be related with two different sources of pollution, the most differentiating factor can be attributed to the factory that, having uncontrolled effluents, had been related to the episodic pollution in the previous work.[1] The second factor has been assigned to a contribution from a fat factory which is near the studied zone. The third and fourth factors are difficult to be correlated to any single source.

A decrease in the importance of factor 1 together with a change in the coefficients associated to the differentiating compounds of this factor has been observed for the samples corresponding to the second and third sampling periods. For instance, the most differentiating compounds are phenols and morpholine derivatives in set 2, while phenols and conductivity take this role in set 3. Furthermore in these cases the compounds are not as correlated as in the first sampling period.

According to these facts Factor Analysis has provided useful information in the case that a punctual pollution incident happens, but the overall picture changes considerably when this source disappears. In the present case, where the pollutants are quickly distributed across the aquifer, the non-equilibrium conditions increase the difficulties in the assignation of pollution sources.

Acknowledgements

The authors thank the Diputació de Barcelona, specially to Mr L. Alonso for their cooperation in sampling location and for the economical support.

References

1. G. Rauret, M. T. Galceran and R. Rubio, Rapport delivered to Diputació de Barcelona (1987). Unpublished results.
2. J. H. M. Bartels, T. A. H. M. Janse and F. W. Pijpers, *Anal. Chim. Acta* **177,** 35 (1985).
3. Ch. Lee and S. Kim, *Bull. Korean Chem. Soc.* **9,** 5 (1988).
4. G. Rauret, R. Rubio, F. X. Rius and M. S. Larrechi, *Intern. J. Environ. Anal. Chem.* **12,** 255 (1988).
5. Zoltanné Varsanyi, *Hidrologiai Közlöny* **66,** EVF.6.Szam (1986).
6. L. M. Pavlichenko and S. M. Shapiro, *Vodn. Resur.* **4,** 77 (1982).
7. N. Ferrer, Ph.D. Thesis, Barcelona University (1987).
8. P. K. Hopke, *Receptor Modeling in Environment Chemistry,* John Wiley and Sons, New York (1985).
9. D. L. Massart, B. G. M. Vandeginste, S. N. Deming, V. Michotte and L. Kaufman, *Chemometrics: A Textbook,* Elsevier (1988).
10. E. R. Malinowski and D. G. Howery, *Factor Analysis in Chemistry,* John Wiley, New York (1980).
11. APHA-AWWA-WPCF *Standard Methods for the Examination of Water and Wastewater,* 16th ed. (1985).
12. N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbbrenner and D. H. Bendt, *Statistical Package for the Social Sciences,* McGraw-Hill, New York, 2nd ed. (1975).