Chemometric Studies in the Lagoon of Venice, Italy. 1. The Environmental Quality of Water and Sediment Matrices

G. Scarponi,*,†,‡ C. Turetta,§ G. Capodaglio,†,§ G. Toscano,† C. Barbante,†,§ I. Moret,†,§ and
P. Cescon†,§

Dipartimento di Scienze Ambientali, Università Ca' Foscari di Venezia, Dorsoduro 2137,
I-30123 Venezia, Italy, Facoltà di Scienze Matematiche, Fisiche e Naturali, Università di Ancona,
Via Brecce Bianche, I-60131, Ancona, Italy, and Centro di Studio sulla Chimica e le Tecnologie per
l'Ambiente, CNR, Dipartimento di Scienze Ambientali, Università Ca' Foscari di Venezia, Dorsoduro 2137,
I-30123 Venezia, Italy

Water, particulate, and sediments were collected at a few sites in the Venice Lagoon (Italy) in order to evaluate the quality of these components through a number of mainly chemical measurements. Trace elements (mainly heavy metals), nutrients, PCBs, a toxicity test, and other general constituents were the main features considered. Principal component analysis was applied as a multivariate statistical technique for dimensionality reduction, for the formulation of a multivariate quality criterion, and for identification of the most significant variables. The first few (2−3) principal components have proven useful in summarizing the information available in all the observed variables, supplying a link between the different classes of chemical substances detected in the water matrix and the different polluting inputs (industrial, agricultural, urban) to which the Venice Lagoon is subject (water matrix) and between the chemical composition (speciation of trace elements) of the sediment matrix and its toxicity.

INTRODUCTION

Several studies carried out in the field of environmental analytical chemistry concern the chemical characterization of the various components of ecosystems with the aim of evaluating their quality. In particular, investigations tend to identify particular crisis states caused either by a lack or shortage of essential substances for the normal development of biological activity or by an abnormally high content of such substances or else by the presence of toxic and harmful substances in quantities higher than those naturally present (or not present at all) in uncontaminated ecosystems.[1−4]

Modern chemical analytical methodologies, which take advantage of sophisticated instrumental tools, often capable of simultaneous, multiple determinations and with automatic capabilities, enable a high number of chemical species to be detected, also at the ultratrace level, thus giving a wide database which contains the required information in single variables or spread in several variables.[5−8]

Though the univariate statistical analysis of collected data makes it possible to analyze the actual contents of constituents individually and to compare them with optimal, predetermined intervals, it does not permit full evaluation of the overall quality of samples, which indeed depends on all the observed variables, which must therefore be considered simultaneously. The methods of multivariate statistical analysis allow the original variables to be combined according to well-defined criteria, to obtain one or more transformed variables which collect and synthesize most of the information present in all the observed variables in few dimensions. Finally this allows samples to be compared and their quality to be evaluated through an integrated tool which uses the information contained in the collected data and spread in the numerous variables observed.[9−18]

In recent years chemometric methods have become an integral part of environmental studies. In particular principal component analysis (PCA) has frequently been used to characterize environmental matrices, to define their quality (e.g., soil,[19,20] rainwater,[21,22] surface water,[23] groundwater,[24,25] rivers,[26,27] and lakes[28−30]), and to identify various sources of pollution, sometimes attempting the apportionment of their contributions (e.g., in oil spills,[31] sediments,[32,33] rainwater,[34,35] and atmospheric particulate[36−40]). PCA and/or discriminant analysis have also been applied to study the relationship of toxic response to the chemical composition of various environmental matrices.[41−45]

In this work a (mainly) chemical data set referring to water, particulate and sediments of the Venice Lagoon ecosystem was considered with the aim of evaluating the quality of these environmental matrices. Trace elements (mainly heavy metals), nutrients, polychlorinated biphenyls (PCBs), a toxicity test, and other general constituents and measurements were the main features considered. PCA was applied as a multivariate statistical technique for dimension reduction, for the formulation of a multivariate quality criterion, and for the identification of the most significant variables and their association with pollution sources.

The Venice Lagoon ecosystem (Figure 1), located in Northern Italy, comprises three parts which are closely

---

* Corresponding author: Prof. Giuseppe Scarponi, Department of Environmental Sciences, Ca' Foscari University of Venice, Calle Larga Santa Marta 2137, I-30123 Venice, Italy. Tel: ++39 41 2578547. Fax: ++39 41 2578549. E-mail: scarponi@unive.it.
† Dipartimento di Scienze Ambientali, Università Ca' Foscari di Venezia.
‡ Università di Ancona.
§ Centro di Studio sulla Chimica e le Tecnologie per l'Ambiente, CNR, Dipartimento di Scienze Ambientali, Università Ca' Foscari di Venezia.

Figure 1. Satellite image of the Lagoon of Venice (© ESA 1991, © CNES 1991; concessed by Nuova Telespazio; processed by Geogramma, Olbia, Italy).
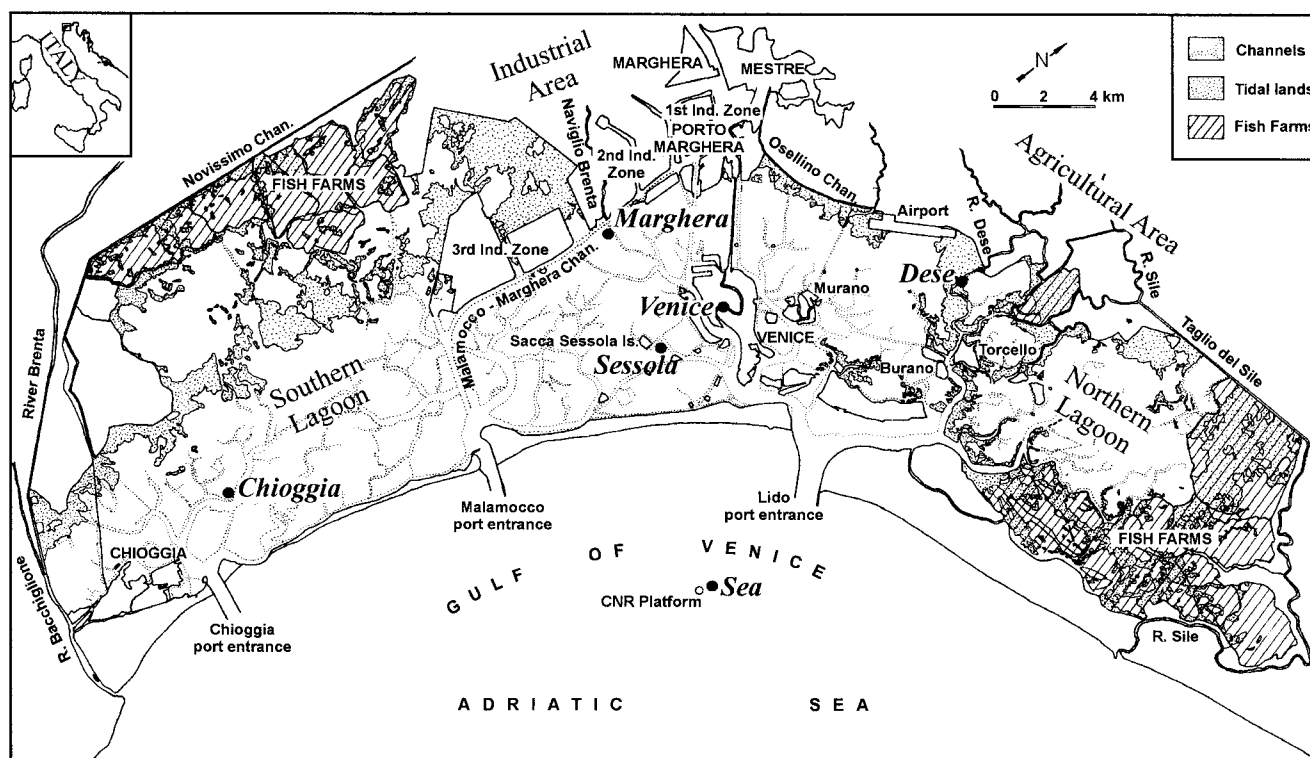
Figure 2. Sampling locations.

connected to each other through their reciprocal water exchanges: the drainage basin, i.e., that part of the mainland which conveys freshwater to the lagoon, the Lagoon itself with a mean water depth of 1.2 m, and the Northern Adriatic Sea whose tides influence the Lagoon with ebb and flow of seawater through the three inlets. Agriculture and zootechny (with 234 000 head of cattle and pigs) occupy about 62% of the surface of the drainage basin, and their influence is mainly reflected in the northern part of the lagoon where the River Dese flows into the Lagoon (the only one to do so). About one and a half million inhabitants live in the drainage basin, while 200 000 reside in lagoon centers, mainly located in the central area, with 100 000 in Venice, though this figure is swelled daily by commuters and tourists. Immediately inland in the central area, directly facing the lagoon, is a vast industrial area (subdivided into three adjacent zones) which reached peak development in 1950–1970 and includes mainly steelworks, chemical factories, oil refineries, and a power station. The southern basin of the lagoon has a substantially lower human impact; it is considered the less polluted area of the Lagoon and many parts are used for fishing and fish farming activities.

Under the Venetian Republic, a few rivers flowing into the Lagoon were diverted to prevent it becoming silted up. At present the ecosystem is principally threatened by pollution and the phenomenon of exceptionally high tide levels; the government is committed to a complex program of research and action to restore and defend the Lagoon and to safeguard the city of Venice and its cultural heritage. The present work is a part of this program.

EXPERIMENTAL SECTION

Samples. In a pilot investigation on the (principally) chemical characterization of water, particulate, and sediments of the Venice Lagoon, samples of the three matrices were collected at the six sites indicated in Figure 2.[46,47] The sites were selected in order to represent quite different areas subjected to different inputs. Site 1 (here referred to as Marghera), located in front of Marghera harbor and the industrial area, is principally influenced by industrial input; site 2 (referred to as Dese), located near the mouth of the River Dese, is under the direct influence of the agricultural area; site 3 (Venice), located in the Grand Canal of Venice (near Ca' Foscari University), is in the urban zone; site 4 (Sessola), located near the Island of Sacca Sessola, represents the central area of the Lagoon; site 5 (Chioggia), located centrally in the southern basin of the Lagoon (near the Petta di Bò *casone*), is the area of the Lagoon considered as the cleanest; site 6 (Sea), in the open sea outside the Lagoon and near the platform of the Italian National Research Council (CNR), is considered as an external reference site.

Samples were collected in a quadrature period (July 1992 from the sixth to the 10th, during the summer algal bloom) in order to have minimum variations caused by tidal effects. Due to the logistical impossibility of effecting it simultaneously, sampling was carried out on successive days but always in flow tide conditions which were as similar as possible (minimum). No particular atmospheric events were observed during, or immediately before, the sampling period.

Chemical Analysis. The complete list of the observed chemical variables and other measurements, together with the related symbols and units, is given in Tables 1 and 2 (see below). For the water matrix (filtered through 0.45 μm pore size filter) the measurements include a set of trace elements, among which there are a number of heavy metals, PCBs as organic contaminants, nutrients and related substances, and a set of variables of general characterization of the matrix. It is to be noted that speciation measurements are included for Cu, Cd, and Pb, for which the data set shows

CHEMOMETRIC STUDIES IN THE LAGOON OF VENICE, ITALY. 1

J. Chem. Inf. Comput. Sci., Vol. 38, No. 4, 1998 **555**

the total metal content, the ASV-labile fraction, and the concentration of ligands which show affinity for each metal.[48-50] One class of ligands is detected for Cd and Pb, while two classes of ligands are detected for copper.

Trace elements, mainly heavy metals, are determined in the particulate and sediment matrices. In the case of the sediments, speciation data are available for the total metal amount (complete solubilization by digestion in an acid mixture of HF, HNO₃, HCl in a Teflon bomb at 180 °C) and for three fractions obtained by the following sequential extraction procedure: CH_3COONH_4 1 M, pH 5, 15 h (ionic exchangeable form and carbonate fraction, A), $NH_2OH \cdot HCl$ 1 M and CH_3COOH 25% v/v, 15 h (metals in the reductive phase bound to Fe−Mn oxides, fraction B); HNO_3 8 M, 3 h (metals bound to sulfides and to the organic phase, fraction C).[51] From the experimental data a residual fraction (D) was calculated by subtracting the sum of the A, B, and C fractions from the total content. A measurement of special interest available for sediments regards the Microtox toxicity test based on the reduction of light emission from photolumi-nescent bacteria (*Vibrio fisheri*) when the sediment is added to a 2% solution of NaCl containing the bacteria (temperature 15 °C, contact time 20 min).[52-54] The effect concentrations (in percentage of the sediment w/v) for 50, 20, 10% luminescence reduction (EC50, EC20, EC10) are reported.

The data set was obtained from an investigation carried out in the framework of the Program "Sistema Lagunare Veneziano" (lines 1.2 and 1.8), and details of the analytical methodologies can be found in the activity technical reports.[46,47]

Statistical Analysis. The multivariate statistical method used was principal component analysis (PCA).[55-58] This procedure was used mainly to achieve a reduction of dimensionality, i.e., to fit a k-dimensional subspace to the original p-variate ($p \gg k$) observations. The statistic used to summarize the most important results was the proportion of the total variation explained by the first k principal components. The interpretation of PCA results was based on the diagrams of coefficients of variables and the scatter plot of samples with reference to the first few (usually two) principal components.

Summary statistics of results, correlation analysis, and the principal component analysis were carried out using the SIMCA-S (version 5.1a)[59] and the UNISTAT (version 4.51)[60] statistical packages. In PCA standardized data were used because of the lack of homogeneity both in the measurement units and in the variances of the observed variables. Significance of principal components was established according to the SIMCA-S cross-validation criteria.[59,61] Briefly these are based on the computation of the quantity $Q^2 = 1 - PRESS/SS$, where PRESS, the prediction error sum of squares for k dimensions, is the sum of squared differences between observed and predicted values by the cross valida-tion procedure (prediction of data previously deleted in model development), and SS is the residual sum of squares for prediction with $k-1$ dimensions. Subsequently Q^2 is com-pared with a significance limit which depends on the degrees of freedom. When Q^2 is larger than the appropriate significance limit for the whole data set (Rule 1) or for $\sqrt{p}$ variables (Rule 2), the tested dimension is considered significant.

RESULTS AND DISCUSSION

Summary Statistics. Tables 1 and 2 report the summary statistics of the complete data set for the water matrix (soluble and particulate fractions) and sediments, respectively. The mean, standard deviation (SD), median, and minimum−maximum of each variable is given. Comparison of the relative standard deviations indicates that in the case of the water matrix the highest scatter about the mean is observed for all nutrients, Cd and Se, for the soluble fraction, and for Cd and Cu in the case of particulate fraction. In the case of the sediments major variations are observed for Cd, Zn, Pb, Cu, Hg, and the EC variables.

Water Matrix (Soluble and Particulate Fractions). The correlation matrix (Table 3) highlights two main sets of intercorrelated variables. Chlorophyll, nitrates, phosphates, heavy metals both in soluble and particulate fractions (excluding Mn), and PCBs are grouped in the first, while ammonia nitrogen, nitrites, silicates, the hydrogen ion, together with elements Ba (soluble and particulate), Mn, Se, and particulate As are included in the second. As regards the speciation variables of Cu, Cd, and Pb it can be observed that labile and total concentrations are generally correlated with each other, while the ligand contents do not correlate either with the related metals (total and labile concentration)

Table 1. Water Matrix (Soluble and Particulate Fractions)[a]

variable, symbol (unit)	mean	SD	median	min.−max.
Soluble Fraction				
transmittance, T (%)	40.7	22.5	34.0	15.2−73.9
oxygen, O2 (mL/L)	4.5	0.6	4.3	3.9−5.5
redox potential, Eh (mV)	65	53	68	9−132
chlorophyll, Chl (μg/L)	1.5	0.5	1.4	1.0−2.4
salinity, S (‰)	32.8	1.3	32.9	31.0−34.1
conductivity, k (mS/cm)	48.3	1.2	48.1	47.0−49.9
hydrogen ion, H (nM)	11.2	7.5	9.0	4.7−25.1
ammonia nitrogen, N-NH3 (μM)	18.6	26.0	9.8	0.7−70.2
nitrites, N-NO2 (μM)	0.79	1.07	0.23	0.08−2.72
nitrates, N-NO3 (μM)	19.6	29.6	1.3	0.7−67.8
phosphates, P-PO4 (μM)	0.79	1.03	0.27	0.05−2.61
silicates, Si-SiO4 (μM)	18.1	19.6	11.3	0.7−55.8
barium, Ba (μg/L)	14.9	6.9	14.0	8.3−27.6
chromium, Cr (μg/L)	3.1	1.8	3.4	0.5−5.1
manganese, Mn (μg/L)	5.6	4.4	4.3	1.4−14
iron, Fe (μg/L)	6.5	2.2	6.0	4.0−10.5
nickel, Ni (μg/L)	5.0	1.7	4.8	2.3−7.1
selenium, Se (μg/L)	0.94	1.07	0.58	0.25−3.06
copper total, Cu-tot (μg/L)	0.52	0.18	0.50	0.33−0.87
ligand 1 for Cu, Lig1-Cu (nM)	32	49	16	0.2−132
ligand 2 for Cu, Lig2-Cu (nM)	112	30	114	71−158
cadmium total, Cd-tot (ng/L)	32	46	15	7−125
cadmium labile, Cd-lab (ng/L)	13	26	2	1−65
ligand 1 for Cd, Lig1-Cd (nM)	0.16	0.14	0.10	0.05−0.40
ligand 2 for Cd, Lig2-Cd (nM)	1.5	0.7	1.3	1.1−2.8
lead total, Pb-tot (ng/L)	120	90	99	32−285
lead labile, Pb-lab (ng/L)	29	15	25	12−55
ligand for Pb, Lig-Pb (nM)	1.2	0.5	1.3	0.33−1.8
polychlorinated biphenyls, PCB (ng/L)	4.8	2.0	5.2	2.0−7.5
Particulate Fraction				
cadmium, Cd-p (ng/L)	6.5	7.8	2.7	0.2−16.8
copper, Cu-p (μg/L)	0.238	0.282	0.113	0.004−0.713
nickel, Ni-p (μg/L)	0.132	0.083	0.139	0.020−0.248
chromium, Cr-p (μg/L)	0.114	0.088	0.095	0.009−0.257
iron, Fe-p (μg/L)	77.7	53.4	82.6	1.4−139.9
aluminum, Al-p (μg/L)	32.9	22.3	32.4	0.8−61.7
arsenic, As-p (μg/L)	0.08	0.07	0.07	0.02−0.19
barium, Ba-p (μg/L)	0.66	0.37	0.57	0.20−1.18
tin, Sn-p (μg/L)	0.22	0.08	0.21	0.13−0.33

[a] Summary statistics for the observed variables.

Table 2. Sediments[a]

variable, symbol (unit)	mean	SD	median	min.–max.	variable, symbol (unit)	mean	SD	median	min.–max.	variable, symbol (unit)	mean	SD	median	min.–max.
chromium A, Cr-A (μg/g)	0.30	0.11	0.25	0.20–0.45	aluminum C, Al-C (mg/g)	3.95	1.34	3.75	2.60–6.30	copper A, Cu-A (μg/g)	5.3	5.4	3.1	1.7–16.1
chromium B, Cr-B (μg/g)	0.46	0.11	0.46	0.30–0.63	aluminum D, Al-D (mg/g)	11.1	4.4	11.1	4.1–17.5	copper B, Cu-B (μg/g)	3.9	1.4	3.4	2.8–6.5
chromium C, Cr-C (μg/g)	8.4	4.4	6.1	5.4–16.3	zinc A, Zn-A (μg/g)	77.5	87.6	44.7	32.5–256.0	copper C, Cu-C (μg/g)	25.6	35.4	11.5	8.2–97.6
chromium D, Cr-D (μg/g)	17.4	9.4	14.7	7.0–31.2	zinc B, Zn-B (μg/g)	35.1	16.6	30.5	21.6–67.2	copper D, Cu-D (μg/g)	6.5	7.3	4.0	0.7–20.8
manganese A, Mn-A (μg/g)	71.9	38.6	79.4	8.5–116.0	zinc C, Zn-C (μg/g)	39.1	28.7	28.2	10.6–83.5	iron A, Fe-A (μg/g)	92.5	116.8	34.5	10.5–310.0
manganese B, Mn-B (μg/g)	25.5	11.8	24.9	8.4–42.8	zinc D, Zn-D (μg/g)	38.2	44.4	23.0	4.5–117	iron B, Fe-B (μg/g)	439	233	493	170–776
manganese C, Mn-C (μg/g)	74	24	67	50–110	lead A, Pb-A (μg/g)	4.8	5.2	3.0	1.5–15.2	iron C, Fe-C (mg/g)	8.42	2.81	7.80	5.10–12.40
manganese D, Mn-D (μg/g)	92	36	94	40–131	lead B, Pb-B (μg/g)	6.0	3.5	5.1	3.2–12.8	iron D, Fe-D (mg/g)	2.3	0.9	2.3	0.9–3.7
cadmium A, Cd-A (μg/g)	0.36	0.44	0.19	0.07–1.22	lead C, Pb-C (μg/g)	11.7	13.1	7.1	3.9–38.2	mercury, Hg (μg/g)	0.56	0.56	0.42	0.03–1.65
cadmium B, Cd-B (μg/g)	0.08	0.09	0.04	0.01–0.26	lead D, Pb-D (μg/g)	22.3	29.3	10.2	3.2–80.8	selenium, Se (μg/g)	0.18	0.11	0.16	0.04–0.34
cadmium C, Cd-C (μg/g)	0.09	0.13	0.03	0.01–0.34	nickel A, Ni-A (μg/g)	1.1	0.7	1.0	0.2–2.2	50% effect concn, EC50 (g/L)	6.6	10.0	2.5	1.6–27.0
cadmium D, Cd-D (μg/g)	0.02	0.02	0.01	0–0.06	nickel B, Ni-B (μg/g)	0.84	0.60	0.76	0.24–1.80	20% effect concn, EC20 (g/L)	2.8	4.3	1.1	0.8–11.6
aluminum A, Al-A (μg/g)	15.1	9.3	15.5	1.1–28.6	nickel C, Ni-C (μg/g)	9.2	3.9	8.2	4.9–15.2	10% effect concn, EC10 (g/L)	1.7	2.6	0.7	0.4–7.0
aluminum B, Al-B (μg/g)	137	45	152	52–180	nickel D, Ni-D (μg/g)	2.5	1.8	2.1	1.2–6.0					

[a] Summary statistics for the observed variables. For explanation of fractions A, B, C, and D in speciation measurements see the text.

Table 3. Water: Correlation Matrix (Underlined Values ≥ |0.70|)

The table is a lower-triangular correlation matrix among the following variables (listed along both the row and the diagonal column headers):

Chl, N-NO3, P-PO4, Cr, Cr-p, Fe, Fe-p, Ni, Ni-p, Cu-tot, Cu-p, Cd-tot, Cd-lab, Cd-p, Pb-tot, Pb-lab, Pb-p, Al-p, PCB, N-NH3, N-NO2, Si-SiO4, H, Ba, Ba-p, Mn, As-p, Se, Sn-p, Lig1-Cd, Lug2-Cd, Lig-Pb, Lig1-Cu, Lug2-Cu, T, O2, Eh, s, k
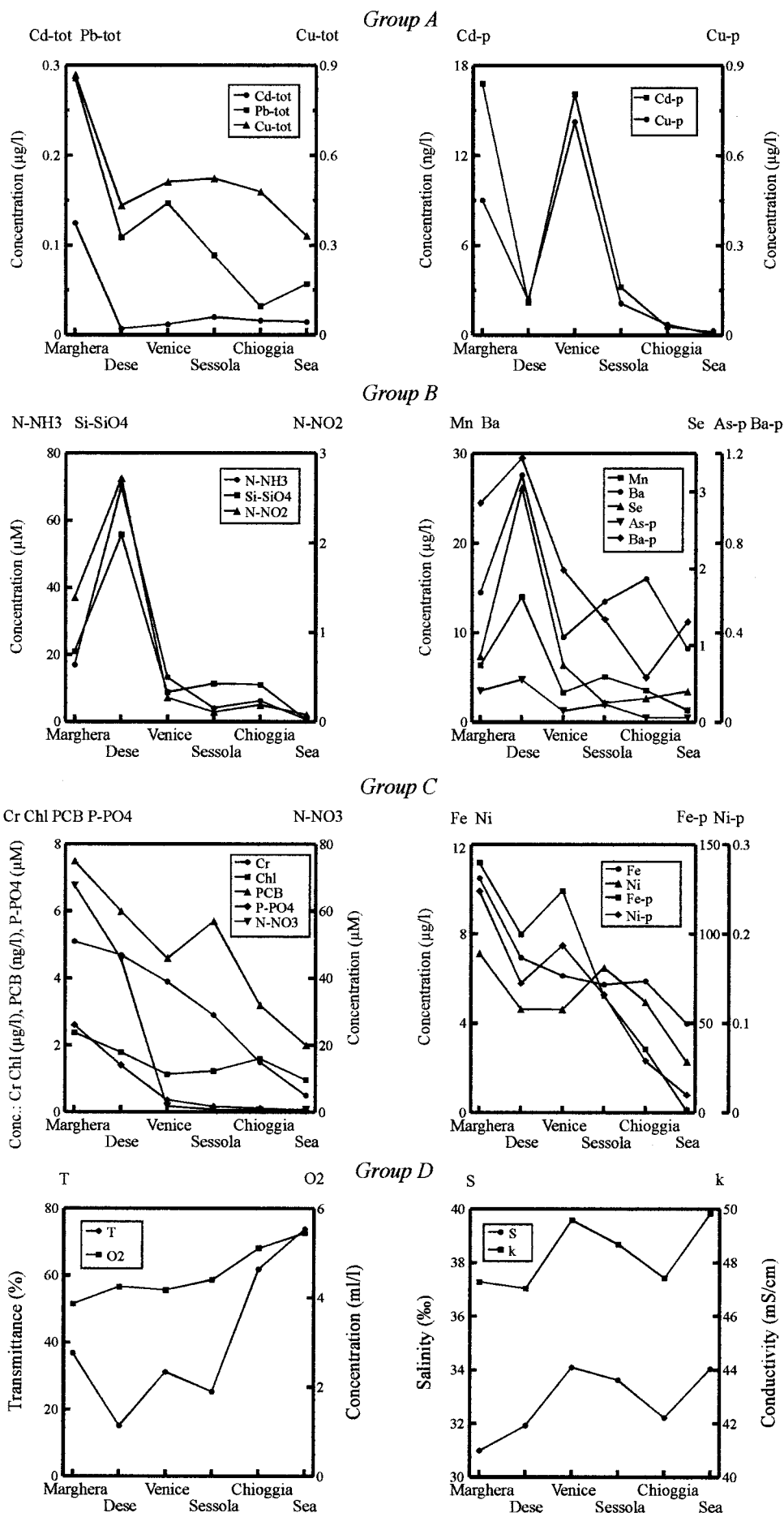
Figure 3. Water matrix: distribution of selected variables in the soluble and particulate phases. Evidence of groupings.

or with each other. This last finding could be considered as a proof of the different nature of ligands which complex different metals (see also the negative correlation between the Cd and Pb ligands, and the correlation of the copper ligand 1 with variables of the second group reported above). Finally it is to be noted that a few high correlations emerge between the variables giving general characterization of the water matrix, namely T, O2 and Eh, on one side, and k and S on the other. Moreover it is significant that general negative correlations are observed between these variables (for which high values mean better quality) and all the other chemical variables (indicators of pollution).

The univariate distribution of variables along the stations, ordered as in Figure 3 (approximately from inside to outside the Lagoon), shows three principal groups corresponding to three distinct profiles. Such groupings comprise substances which, depending on their concentration, constitute pollutants of the environment under investigation. The first group (A in Figure 3), which shows a maximum for Marghera station, and Venice station for particulate variables, comprises substances (the typical anthropogenic heavy metals Cu, Pb, Cd) which can be defined as "industrial" (and "urban") pollutants. The second set (B in Figure 3), characterized by a peak in the Dese station, groups substances which can be considered pollutants of "agricultural" origin (N-NH3, N-NO2, Si-SiO4, Se, Mn, soluble and particulate Ba and particulate As). The uses of As and Se (in pesticides and feed additives),[62–64] Ba (in pesticides),[65] and NH_3 or NH_4^+ salts (in fertilizers)[66] as well as the NH_3 emission from animal excreta[66] are well-known in agriculture. The third group (C in Figure 3), which presents the highest values in both the Marghera and Dese stations, includes more general pollutants, such as N-NO3, P-PO4, PCBs and Cr, Ni, Fe, Al in soluble and particulate phases.

The differentiation between the three groups of variables and their environmental significance is evident especially if one considers the distinct situations which characterize the Marghera and Dese stations; the first receives a considerable input of polluting substances from the industrial activities operating in the immediate hinterland, while the second is subject to agricultural input, arriving principally via the River Dese. Of course the general urban impact can be foreseen for the station located in the Grand Canal of Venice.

Principal component analysis (three components significant according to rule 1, first two, and rule 2, third, of the SIMCA-S cross-validation criteria[59,61]) confirms the existence of the groups already recognized (see biplots in Figure 4). The plot of the first two components (cumulative proportion of total variation about 80%, Figure 4a) shows groups of variables which reflect the distinct profiles observed in the univariate analysis, in particular the group A from Pb-lab to Pb-tot, the group B from Lig1-Cd to Ba-p, and the group C from N-NO3 to Cd-p. Moreover the same plot makes it possible to define a fourth, more dispersed group, which includes a set of variables indicative of the general environmental conditions (salinity, conductivity, oxygen, redox potential, and transmittance). The variables in this group (positively related to water quality) show trends in contrast to the general behavior observed above, with minimum values for the inner stations of the Lagoon (Marghera and Dese) and an increasing trend going toward the sea (Figure 3, group D).
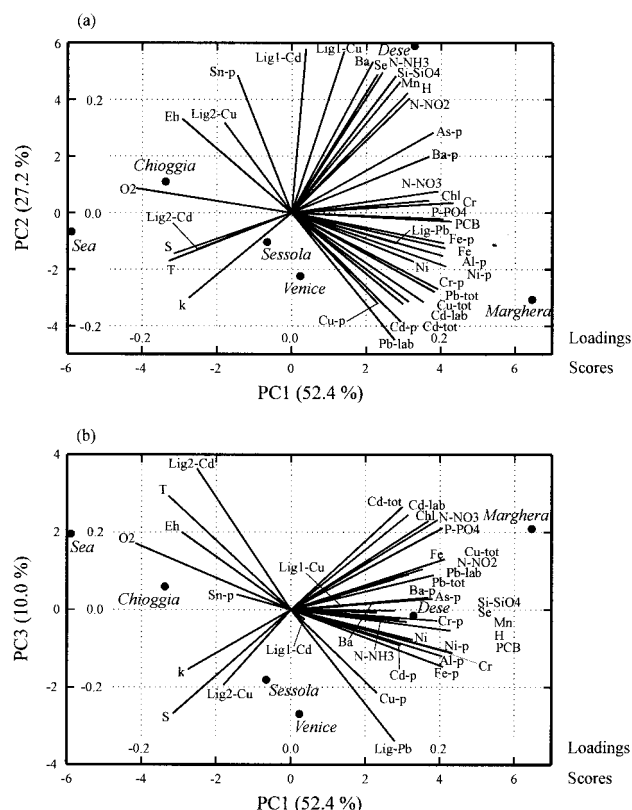

Figure 4. Water matrix: principal component biplots: (a) PC1 vs PC2 and (b) PC1 vs PC3. In parentheses explained variance.

The scatter plot of samples on the first two principal components (Figure 4a) highlights the strong multivariate distinction of the Marghera and Dese stations with respect both to each other and to the other sites. In this respect the first component seems to be indicative of the general pollution level; the variables which have the highest positive weight are the pollutant substances and variables with negative weight are those which evince good characteristics such as oxygen, transmittance and redox potential. From the PC biplot, which shows associations between samples and variables, it can be deduced that the overall (multivariate) contamination level is maximum for Marghera and Dese, intermediate for Venice, Sessola and, to a lesser degree, Chioggia, and minimum for the Sea station (CNR Platform). The second component, on the other hand, indicates the kind of pollution involved: "industrial" for Marghera (and "urban" for Venice), shown by positive values of the component which, in turn, correspond to high content of heavy metals Cu, Pb, and Cd, and "agricultural" for Dese, with negative values of the component, i.e., corresponding to, e.g., high values of N-NH3, N-NO2, Si-SiO4, As, Se, Ba. The remaining stations are not significantly differentiated by the second component.

On the whole, the same groups of variables independently recognized above by univariate analysis can now be observed distributed in the bidimensional PC biplot. This plot synthesizes most of the multivariate information present in the data in two dimensions, whose interpretation in terms of general pollution (first PC) and specific, contrasting contributions, industrial-urban vs agricultural (second PC), is really straightforward.

Perhaps it is useful to observe that the third principal component (10.0% explained variance, see Figure 4b), not
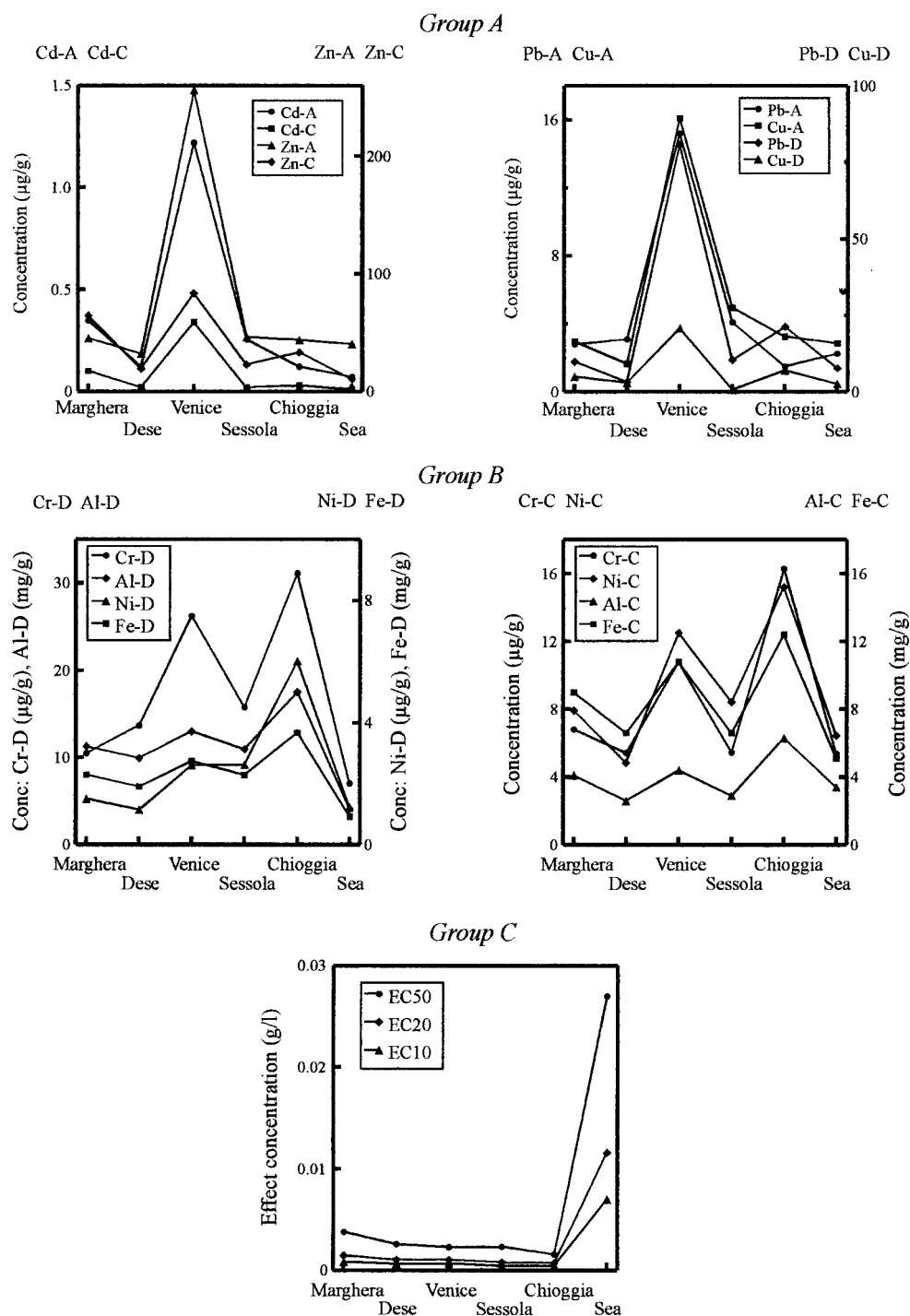
Figure 5. Sediments: distribution of selected variables. Evidence of groupings.

considered in the above discussion, is dominated by the concentration of the ligand 2 for cadmium (positive weight) and the lead ligand concentration (negative weight) both obtained from speciation measurements. In this third PC dimension a separation between the Sea and Marghera stations on the one side (high Cd ligand 1 content) and Venice and Sessola on the other (high Pb ligand content) can be observed. This behavior is, however, difficult to interpret due to the presently unknown nature of the ligands.

Sediments. The data referring to this matrix include, on one hand, a set of concentration measurements of elements, principally heavy metals, most of them (all except Hg and Se) fractioned on the basis of different treatments (specia-

tion), and on the other a toxicity test based on photoluminescence produced by microorganisms (Microtox test). Effect concentrations for 50, 20, and 10% luminescence reduction (EC50, EC20, EC10) are reported. Regarding the toxicity test it is to be noted that there is an inverse relation between the toxicity and the EC value, i.e., an increase in sediment toxicity will involve a decrease in the effect concentration (concentration of sediment in the tested solution).

A few interesting observations can be made with reference to the correlation matrix (not given here to save space). With regard to speciation, a different behavior is observed between the group of Pb, Cu, Cd, and Zn (referred to below as the
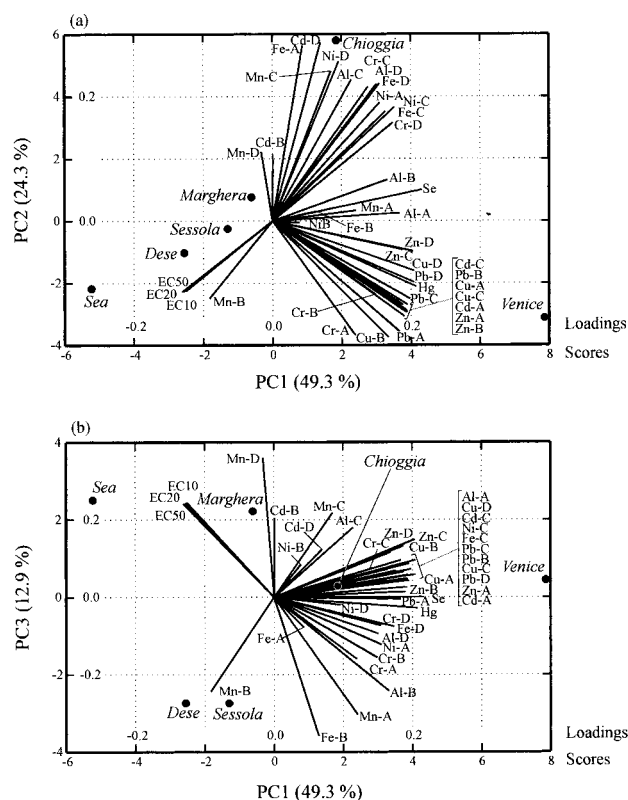
Figure 6. Sediments: principal component biplots: (a) PC1 vs PC2 and (b) PC1 vs PC3. In parentheses explained variance.

group of Pb), on one side, and the group of Al, Fe, Cr, Ni, and Mn (the group of Al), on the other side. For the first group (Pb) the four speciation fractions are all highly correlated (both within each metal and between metals), with the exception of Cd-B and Cd-D ($0.74 \leq corr \leq 1.00$). For the second group (Al) a good correlation is again present between the C and D fractions, again both within metal ($0.61 \leq corr \leq 0.93$) and between metals (with a few exceptions), but low correlations are generally observed for A and B fractions. The variables measuring toxicity (in inverse proportion) at different sensitivities (EC variables) are all highly correlated with each other (corr > 0.999). Moreover they are negatively correlated with all the other variables except Mn-D and Mn-B.

From univariate statistical analysis, three distinct groups have been recognized. The first, where maximum values are observed in the Venice station, brings together the heavy metals Pb, Cu, Cd, Zn, and Hg (group A in Figure 5). The second group (B in Figure 5), which unexpectedly shows peak values in the Chioggia station (and the second highest values in Venice) includes Al, Cr, Ni, and Fe. The third group consists essentially of the EC variables (group C in Figure 5), indicators in inverse proportion of sediment toxicity, with very high maxima at the Sea station and, among the stations inside the Lagoon, a (surprisingly) progressive increase going toward inner sites.

Principal component analysis (three components cross-validated according to rule 2 of SIMCA-S package,[59,61] see biplots in Figure 6) very clearly highlights the three groups recognized above (Pb group, Al group, EC variable group) and shows the contrast between the first two groups (trace elements) and the third (EC variables), particularly with reference to the first principal component, which, in this case

too could represent a reference scale for overall quality evaluation. Considering the speciation measurements, for the metals of the Pb group all the different fractions of the same metal appear well grouped together, while a net separation (particularly evident in the second and third components) is observed for the Al-group metals between the A and B fractions, on one side, and the C and D fractions (better grouped) on the other.

The scatter plot of samples on the biplot of the first two principal components (Figure 6a) shows that increasing values of the first component correspond to a gradual shift from less polluted to more polluted sites, particularly in the order Sea, Dese, Sessola, Marghera, and principally Chioggia and Venice. The second component, which is more influenced by the variables Al, Fe, Cr, Ni, and Mn (particularly the mineralogical fractions C and D), separates the Chioggia station from the others. In this respect the Chioggia station, which had appeared as the least polluted of those inside the Lagoon when we considered the water matrix (the entire basin of Chioggia is also in fact considered as the cleanest in the Lagoon), surprisingly shows a generally higher metal content when the sediment is considered and, consequently, a low quality. At present no further explanation can be given for this contrasting result, which, however, requires more extensive investigation to be considered as conclusive.

Considering the third component (Figure 6b) it is principally the speciation of Mn that appears to separate the stations, with extreme, opposite situations observed for stations Sea and Marghera (high C and D fractions) on one side, and Dese and Sessola (high A and B fractions together with Fe-B and Al-B) on the other.

It is to be stressed that PCA components, although they account for a high proportion of the total variation, do not provide a direct, quantitative relationship between sediment chemical composition and its toxicity, the EC variables being treated as simple variables included in the data set. This aspect will be treated in a subsequent work dealing with a partial least squares (PLS)[67,68] analysis of the same data set.

CONCLUSIONS

The application of the multivariate statistical technique of principal component analysis to the environmental data collected in the Lagoon of Venice has made it possible to confirm the general characteristics of the different sites as shown in the univariate study of each variable. In particular several variable associations have enabled characterization of the different polluting inputs to which the Lagoon is subject, i.e., industrial, agricultural, and urban inputs. In this respect the first few principal components have proven to be useful to summarize the information present in all the variables, discriminating between the different pollutant categories and, very interestingly, giving an indication of the overall quality of the matrices. Principal components do in fact appear well-suited for use in the development of an overall quality index for environmental matrices.

The reduced number of stations considered in this pilot investigation suggests that a cautious approach should be taken to the generalization of results obtained here. A more extensive sampling program should confirm the findings of the present study and give a more in-depth interpretation of the results with respect to the general ecological/environmental knowledge of the Lagoon.

A possible future topic of investigation, not dealt with here due to the reduced number of samples analyzed, concerns the identification of possible redundant variables, which could be eliminated to reduce the chemical analytical work load without any loss of essential information.

REFERENCES AND NOTES

(1) *Water Quality Measurement. The Modern Analytical Techniques*; Mark, H. B., Jr., Mattson, J. S., Eds.; Marcel Dekker: New York, 1981.

(2) Waite, T. D. *Principles of Water Quality*; Academic Press: Orlando, FL, 1984.

(3) Ward, C. H.; Giger, W.; McCarty, P. L. *Ground Water Quality*; Wiley: New York, 1985.

(4) Merian, E.; Haerdi, W. *Metal Compounds in Environment and Life. Interrelation Between Chemistry and Biology*; Science and Technology Letters: Northwood, 1992.

(5) *Analytical Techniques in Environmental Chemistry*; Albaiges, J., Ed.; Pergamon Press: Oxford, 1980; Vol. 1; 1982; Vol. 2.

(6) *Water Analysis*; Minear, R. A., Keith, L. H., Eds.; Academic Press: Orlando, FL, 1982; Vol. I, Inorganic species, Part 1; 1984; Vol. II, Inorganic species, Part 2; 1984; Vol. III, Organic species.

(7) *Water Analysis*; Fresenius, W., Quentin, K. E., Schneider, W., Eds.; Springer-Verlag: Berlin, 1988.

(8) *Analytical Aspects of Environmental Chemistry*; Natush, D. F. S., Hopke, P. K., Eds.; Wiley: New York, 1983.

(9) Hopke, P. K. In *Analytical Aspects of Environmental Chemistry*; Natush, D. F. S., Hopke, P. K., Eds.; Wiley: New York, 1983; pp 219–262.

(10) *Environmental Applications of Chemometrics*; Breen, J. J., Robison, P. E., Eds.; American Chemical Society: Washington, D.C., 1985.

(11) Baasel, W. D. *Economic Methods for Multipollutant Analysis and Evaluation*; Dekker: New York, 1985.

(12) Knudson, E. J.; Dnewer, D. L.; Christian, G. D.; Larson, T. V. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; American Chemical Society: Washington, D.C., 1977; pp 80–116.

(13) Smith, R. W.; Bernstein, B. B.; Cimberg, R. L. In *Marine Organisms as Indicators*; Soule, D. F., Kleppel, G. S., Eds.; Springer-Verlag: New York, 1988; pp 247–329.

(14) Papers presented at the Symposium on Chemometrics with Environmental Applications. Stapanian, M. A., Guest Ed.; *J. Chemometrics* **1991**, *5*, 121–332.

(15) Funtowicz, S.; Munda, G.; Paruccini, M. The Aggregation of Environmental Data Using Multicriteria Methods. *Environmetrics* **1990**, *1*, 353–368.

(16) Chemometrics in Environmental Chemistry. Statistical Methods. Einax, J., Ed.; In the series *Handbook of Environmental Chemistry*; Hutzinger, O., Ed.; Springer: Berlin, 1995; Vol. 2, Part G.

(17) Chemometrics in Environmental Chemistry. Applications. Einax, J., Ed.; In the series *Handbook of Environmental Chemistry*; Hutzinger, O., Ed.; Springer: Berlin, 1995; Vol. 2, Part H.

(18) *Chemometrics in Environmental Analysis*; Einax, J. W., Zwanziger, H. W., Geiss, S., Eds.; VCH: Weinheim, 1997.

(19) Brown, D. A.; Mathur, S. P.; Brown, A.; Kusher, D. J. Relationships Between Some Properties of Organic Soils from the Southern Canadian Shield. *Can. J. Soil Sci.* **1990**, *70*, 363–377.

(20) Samukawa, K. Chemical Characterization on Soil Organic Matter by Curie-Point Pyrolysis-Gas Chromatography. *Chem. Express* **1992**, *7*, 445–448.

(21) Vong, R. J.; Frank, I. E.; Charlson, R. J.; Kowalski, B. R. In *Environmental Applications of Chemometrics*; Breen, J. J., Robinson, P. E., Eds.; American Chemical Society: Washington, D.C., 1985; pp 34–52.

(22) Zhang, P.; Dudley, N.; Ure, A. M.; Littlejohn, D. Application of Principal Component Analysis to the Interpretation of Rainwater Compositional Data. *Anal. Chim. Acta* **1992**, *258*, 1–10.

(23) Librando, V. Chemometric Evaluation of Surface Water Quality at Regional Level. *Fresenius' J. Anal. Chem.* **1991**, *339*, 613–619.

(24) Mas, F.; Estela, J. M.; Cerda, V.; Tomas, X.; Comas, M. A.; Gallego, L. Application of Chemometric Techniques to the Analysis of Undergroundwater from the Majorca Island. *J. Environ. Sci. Health, Part A* **1993**, *A28*, 339–351.

(25) Andrade, J. M.; Prada, D.; Alonso, E.; Lopez, P.; Muniategui, S.; de la Fuente, P.; Quijano, M. A. Selection of Analytical Variables to Optimize Laboratory Efforts in Future Groundwater Studies. *Anal. Chim. Acta* **1994**, *292*, 253–261.

(26) Brown, S. D.; Skogerboe, R. K.; Kowalski, B. R. Pattern Recognition Assessment of Water Quality Data: Coal Strip Mine Drainage. *Chemosphere* **1980**, *9*, 265–276.

(27) Poletti, A.; Murgia, S. M.; Santucci, A. Chemometric Techniques Applied to the Analysis of Quality of River Waters. *Ing. Ambientale* **1989**, *18*, 510–515; *Chem. Abstr.* **1990**, *113*, 138151h.

(28) Henrion, G.; Nass, D.; Michael, G.; Henrion, R. Multivariate 3-Way Data Analysis of Amino Acid Patterns of Lakes. *Fresenius' J. Anal. Chem.* **1995**, *352*, 431–436.

(29) Reisenhofer, E.; Picciotto, A.; Li, D. A Factor Analysis Approach to the Study of the Eutrophication of a Shallow, Temperate Lake (San Daniele, North Eastern Italy). *Anal. Chim. Acta* **1995**, *306*, 99–106.

(30) Camarero, L.; Catalan, J.; Boggero, A.; Marchetto, A.; Mosello, R.; Psenner, R. Acidification in High Mountain Lakes in Central, Southwest, and Southeast Europe (Alps, Pyrenees, Pirin). *Limnologica* **1995**, *25*, 141–156.

(31) Duewer, D. L.; Kowalski, B. R.; Schatzki, T. F. Source Identification of Oil Spills by Pattern Recognition Analysis of Natural Elemental Composition. *Anal. Chem.* **1975**, *47*, 1573–1583.

(32) Wenning, R.; Paustenbach, D.; Johnson, G.; Ehrlich, R.; Harris, M.; Bedbury, H. Chemometric Analysis of Potential Sources of Polychlorinated Dibenzo-p-dioxins and Dibenzofurans in Superficial Sediments from Newark Bay, New Jersey. *Chemosphere* **1993**, *27*, 55–64.

(33) Cash, G. G.; Breen, J. J. Principal Component Analysis and Spatial Correlation: Environmental Analytical Software Tools. *Chemosphere* **1992**, *24*, 1607–1623.

(34) Abbas, M. Z. M.; Bruns, R. E.; Scarminio, I. S.; Ferreira, J. R. A Multivariate Statistical Analysis of the Composition of Rainwater Near Cubatao, SP, Brazil. *Environ. Pollut.* **1992**, *79*, 225–233.

(35) Vong, R. J. Atmospheric Chemometrics for Identification of Trace Element Sources in Precipitation. *Anal. Chim. Acta* **1993**, *277*, 389–404.

(36) Artaxo, P.; Rabello, M. L. C.; Watt, F.; Grime, G.; Swietlicki, E. Nuclear Microprobe Analysis and Source Apportionment of Individual Atmospheric Aerosol Particles. *Nucl. Instrum. Methods Phys. Res., Sect. B* **1993**, *B75*, 521–525.

(37) Michaud, D.; Baril, M.; Perrault, G. Characterization of Airborne Dust from Cast Iron Foundries by Physico-Chemical Methods and Multivariate Statistical Analyses. *Air Waste* **1993**, *43*, 729–735.

(38) Gao, N.; Cheng, M.-D.; Hopke, P. K. Receptor Modeling of Airborne Ionic Species Collected in SCAQS. *Atmos. Environ.* **1994**, *28*, 1447–1470.

(39) Mizohata, A.; Ito, N.; Masuda, Y. Quantitative Determination of the Airborne Particulate Matter Sources by Motor Vehicles Using TTFA. *Taiki Kankyo Gakkaishi* **1995**, *30*, 243–255.

(40) Mendez, J.; Quejido, A. J.; Perez-Pastor, R.; Perez-Garcia, M. Chemometric Study of Organic Pollution in the Aerosol of Madrid. *Anal. Chim. Acta* **1993**, *283*, 528–537.

(41) Giesy, J. P.; Hoke, R. A. Bioassessment of the Toxicity of Freshwater Sediment. *Verh.-Int. Ver. Theor. Angew. Limnol.* **1991**, *24*, 2313–2321.

(42) Ahlf, W.; Wild-Metzko, S. Bioassay Responses to Sediment Elutriates and Multivariate Data Analysis for Hazard Assessment of Sediment-bound Chemicals. *Hydrobiologia* **1992**, *235–236*, 415–418.

(43) Niemi, G. J.; Bradbury, S. P.; McKim, J. M. The Use of Fish Physiology Literature for Predicting Fish Acute Toxicity Syndromes. *ASTM Spec. Technol. Publ.* **1991**, *1124*, 245–260.

(44) Klappa, S. A.; Long, G. R. Computer Assisted Determination of the Biological Activity of Polychlorinated Biphenyls Using Gas Chromatographic Retention Indexes as Molecular Descriptors. *Anal. Chim. Acta* **1992**, *259*, 89–93.

(45) Ma, C. Y.; Bayne, C. K. Differentiation of Aroclors Using Linear Discrimination for Environmental Samples Analyzed by Electron Capture Negative Ion Chemical Ionization Mass Spectrometry. *Anal. Chem.* **1993**, *65*, 772–777.

(46) Frache, R. Study of the quality of some components of the ecosystem using procedures of statistical correlation of data or other methodologies allowing definition of rigorous quality indices. Workshop of the Project Sistema Lagunare Veneziano, Venice, 1992.

(47) Cescon, P. Characterization and speciation of heavy metals and other persistent chemical pollutants also in order to evaluate their ecotoxicity more completely. Workshop of the Project Sistema Lagunare Veneziano, Venice, 1992.

(48) Capodaglio, G.; Toscano, G.; Scarponi G.; Cescon, P. Lead Speciation in the Surface Waters of the Ross Sea (Antarctica). *Ann. Chim. (Rome)* **1989**, *79*, 543−559.

(49) Capodaglio, G.; Scarponi, G.; Toscano, G.; Cescon, P. Cadmium Complexation in Surface Seawater of Terra Nova Bay (Antarctica). *Ann. Chim. (Rome)* **1991**, *81*, 279−296.

(50) Capodaglio, G.; Toscano, G.; Scarponi, G.; Cescon, P. Copper Complexation in the Surface Seawater of Terra Nova Bay (Antarctica). *Int. J. Environ. Anal. Chem.* **1994**, *55*, 129−148.

(51) Cosma, B.; Frache, R.; Mazzucotelli, A.; Soggia, F. Trace Metals in Sediments from Terra Nova Bay-Ross Sea, Antarctica. *Ann. Chim. (Rome)* **1991**, *81*, 371−382.

(52) Bulich, A. A. In *Toxicity Testing Using Microorganisms*; Bitton, G., Dutka, B. J., Eds.; CRC Press: Boca Raton, FL, 1986; Vol. I, pp 57−74.

(53) True, C. J.; Heyard, A. A. Relationships between Microtox Test Results, Extraction Methods, and Physical and Chemical Compositions of Marine Sediment Samples. *Toxicity Assessment* **1990**, *5*, 29−45.

(54) Volpi Ghirardini, A.; Ghetti, P. F.; Pantani, C.; Arizzi Novelli, A. L'impiego di *Photobacterium Phosphoreum* nella Determinazione del Rischio Tossico di Sedimenti. Proceed. VI Meeting Italian Society of Ecology (S.I.T.E.), Venice, 1994; Edizioni Zara, Parma, Italy, 1995; pp 385−388.

(55) Rao, C. R. The Use and Interpretation of Principal Component Analysis in Applied Research. *Sankhya A* **1964**, *26*, 329−358.

(56) Mardia, K. V.; Kent, J. T.; Bibby, J. M. *Multivariate Analysis*; Academic Press: New York, 1979.

(57) Jolliffe, I. T. *Principal Component Analysis*; Springer-Verlag: New York, 1986.

(58) Krzanowski, W. J. *Principles of Multivariate Analysis. A User's Perspective*; Oxford University Press: Oxford, 1988.

(59) SIMCA-S, Umetri, Umeå, Sweden, 1994.

(60) UNISTAT Statistical Package, Unistat Limited, London, 1997.

(61) Wold, S. Cross-Validatory Estimation of the Number of Components in Factor and Principal Component Models. *Technometrics* **1978**, *20*, 397−405.

(62) Baird, C. *Environmental Chemistry*; W. H. Freeman and Co.: New York, 1995; Chapters 6 and 9.

(63) Merian, E., Ed. *Metals and their Compounds in the Environment*; VCH: Weinheim, 1991; Chapters II.3 and II.25.

(64) Newland, L. W. In *The Handbook of Environmental Chemistry*; Hutzinger, O., Ed.; Springer-Verlag: Berlin, 1982; Vol. 3, Part B, pp 27−67.

(65) *The Merck Index. Tenth Edition*; Windholz, M., Ed.; Merck & Co.: Rahway, NJ, 1983; pp 139−142.

(66) Jaffe, D. A. In *Global Biogeochemical Cycles*; Butcher, S. S., Charlson, R. J., Orians, G. H., Wolfe, G. V., Eds.; Academic Press: London, 1992; Chapter 12.

(67) Geladi, P.; Kowalski, B. R. Partial Least-squares Regression: a Tutorial. *Anal. Chim. Acta* **1986**, *185*, 1−17.

(68) Wold, S. In *Chemometric Methods in Molecular Design*; van de Waterbeemd, H., Ed.; VCH: Weinheim, 1996; pp 195−218.