HAIT Journal of Science and Engineering B, Volume 2, Issues 1-2, pp. 54-94

Factor analysis in environmental studies

Alexander S. Kaplunovsky

Research Center for Quantum Communication Engineering
Holon Academic Institute of Technology, 52 Golomb Str., Holon 58102, Israel
e-mail: alex_ka@netvision.net.il
Received 10 August 2004, revised 14 October 2004, accepted 5 November 2004

Abstract

In our days, factor analysis has become an important statistical instrument of investigation in modern science, being an adequate tool to investigate the principles of interaction of components and their integration into a system. The paper, not pretending to a complete and detailed review, is intended mainly for a wide community of ecologists, which are interesting in principles of using factor analysis in environmental studies. Properties of factor analysis as a robust method of investigation in environmental studies, are considered and discussed on the examples of using methods of factor analysis in air, water and land ecological systems.

1 Introduction

Factor analysis has become in our days a principal statistical method of investigation in life sciences. One can ask what are the reasons of its wide dissemination in almost all scientific fields. Why so much efforts are directed towards new modifications and development of factor analysis methods? What are the benefits that factor analysis can achieve in environmental studies?

While for most exact sciences, the using of differential equations, algebra, set theory, mathematical logic, and operational research is typical and usually sufficient, in such sciences as biology, psychology, sociology - methods intended to the analysis of multiple processes distribution and based on the probabilistic, rather then functional, homomorphism of the model and the object, are used. Therefore, the primary role passes on to the methods of mathematical statistics, theory of information, theory of random processes, etc. However, most of these methods has one crucial defect: they are not integral, and can not give an answer to the question about the specifics and the reasons of organization of analyzed elements. Factor analysis, being an integral statistical method, meets both above-mentioned requests with an opportunity to define and evaluate the structural-functional organization of the system.

Environmental sciences deal with systems characterized by inherent variability (natural, anthropogenic, spatial/temporal) and multivariate origin [1], by relations of developing and control and acting under laws of probability. Therefore, the revealing relations, their limits and hierarchy should be a principal goal in analysis of complicated environmental systems. This can explain a wide use of factor analysis in environmental studies.

Factor analysis is a standard technique in multivariate analysis. So far as the mathematics of factor analysis is not in the scope of this work, I would like just to mention the best, to my knowledge, classical books of Harman [2] and Lawley and Maxwell [3].

A hundred years ago, Spearman [4] published in the American Journal of Psychology the paper about the factor of general intelligence, based on fulfillment of all tests connected with intellectual functions. The main goal of the factor analysis was the control of conformity of *a priory* given factor structure to the experimental data, and the analysis of quantitative differences between tests. Only the multiple-factor analysis, proposed in 1935 by Thurstone [5], has allowed to pick out factors not defined *a priory*. However, all methods of multiple-factor analysis, the centroid method [5], and the method of maximum likelihood [6], very cumbersome in calculations and, mainly, leading to different factor structures, caused the wave of disappointments in the factor analysis, especially in attempts of pithy interpretation of the factors, and therefore stipulated the domination of positivistic, operationalistic ideas in it. Failures in the interpretation led even such "functionalist" as Thurstone to the view on the factor analysis only as a scientific method for the confirmation or rejection of hypotheses concerning the nature of processes [7]. Rotation of the factor matrix allowed overcoming the uncertainty of interpretation, but criteria of the rotation itself were based on the very vague signs of "simple structure", or on the agreement with data obtained by other methods, other investigators and on the agreement to common principles of the concrete science.

The exit of this deadlock was shown by Hotelling proposed the method of principal components analysis (PCA) permitting calculation of the unique matrix of the orthogonal factors [8]. Although this method required many mathematical computations and could be used in practice only with a progress of computers, it immediately got an appreciation of many investigators. Thurstone first pointed that even the most powerful method of factor analysis - the centroid method - is not more then "calculating compromise" of the principal component method, which was later proved by Anderson and Rubin [9]. After the appearance of this method, factor analysis has got his second birth and has had a right to be considered as a method of the structure search in all fields of science.

Let us consider factor logic as a principle of analysis. Suppose that elements of a system can be observed or measured on any finite and unique set for the whole system, for example, on the time axis or/and on the set of some homogeneous objects. This set of components in the factor analysis got the name of the matrix of individuals. After the components have been chosen and the matrix has been set, the matrix of correlations between parameters can be calculated. Factor analysis transforms this matrix to the matrix of factors, where each of them represents a causal connection of elements. It is important to note that by using the technique of principal components, all factors become orthogonal and caused by different properties of the system.

2 Data base

Let us try to evaluate the total number of publications in periodical journals on the use of factor analysis in environmental studies.

Table 1 shows the relative distribution of papers concerning factor analysis in various fields of science and industry. The data have been chosen from the author's collection hosted on http://www.magniel.com/fa/data. The collection consists of 3460 papers hosted in the Internet by May 2004, which include such expressions as *factor analysis*, *principal component analysis,* etc. in the title of paper. Each paper was marked according to its belonging to the certain type of these fields, sometimes more then one, therefore, the sum of numbers along a column is greater or equal to the figure in the last row.

One can be easy convinced that the probability to find the name of the method in the title of publication is about 5%. For example, the expression *factor analysis* could be found in the text of 33834 papers and in the titles of 1622 papers according to http://highwire.stanford.edu or, respectively, 17376 and 702 papers according to http://www.scirus.com. The table shows that 138 papers of the data base have been related to environmental studies. Therefore they represent about 2500 publications in the Internet. Taking into account that for papers on environmental studies published since 1980's not more then half is hosted in the Internet, we come to the figure of at least 5000 publications on using factor analysis in environmental studies (not including biological and medical aspects) which were published in periodical journals.

Naturally, links to the most of reviewed works, could be found on http://www.magniel.com/fa/data/ecology.html. This can help a reader to find details of using factor analysis in a reviewed work.

	1904 -1980	1981 -1985	1986 -1990	1991 -1995	1996 -2000	2001 -2004	Total
Biology	18	17	20	23	47	41	166
Chemistry	12	14	36	53	88	77	244
Chromatography	4	7	16	22	24	15	88
Economics	14	12	9	4	20	26	85
Environmental Studies	2	4	11	15	61	45	138
Food	1	4	5	2	17	21	50
Geriatry	8	5	10	9	25	31	88
Image Processing	2	7	22	27	38	51	151
Industry	4	0	2	6	38	28	78
Magnetic Resonance	1	1	3	6	25	13	49
Medicine	30	32	64	67	109	116	418
Methodology	10	25	31	49	125	151	391
Operational Research	1	1	1	9	42	41	95
Physiology	20	26	38	39	51	29	203
Psychiatry	15	14	39	61	137	99	365
Psychology	93	86	159	219	379	344	1287
Spectroscopy	11	27	40	50	108	90	326
(a) Total FA-papers	196	242	408	545	1065	1002	3460
(b) All papers($*10^3$)	5186	1518	1890	2117	2430	1999	14707
(c) FA/All ($*10^{-6}$)	38	159	216	257	438	501	235

Table 1. Distribution of papers on factor analysis in the Internet
(from the author's collection http://www.magniel.com/fa/data)
The bottom rows show
(a) - total numbers of papers on factor analysis per the given time interval;
(b) - total number of publications from http://highwire.stanford.edu/;
(c) - ratio of papers on factor analysis to all papers in the Internet (c=a/b$\times 10^{-6}$).

Fig. 1 shows common tendencies of ongoing growth not only an absolute, but also a relative number of publications on factor analysis. We can see that percent of publications on using factor analysis has increased more then in three times for the last 20 years. In the field of environmental studies this parameter shows the growth of 5-6 times.

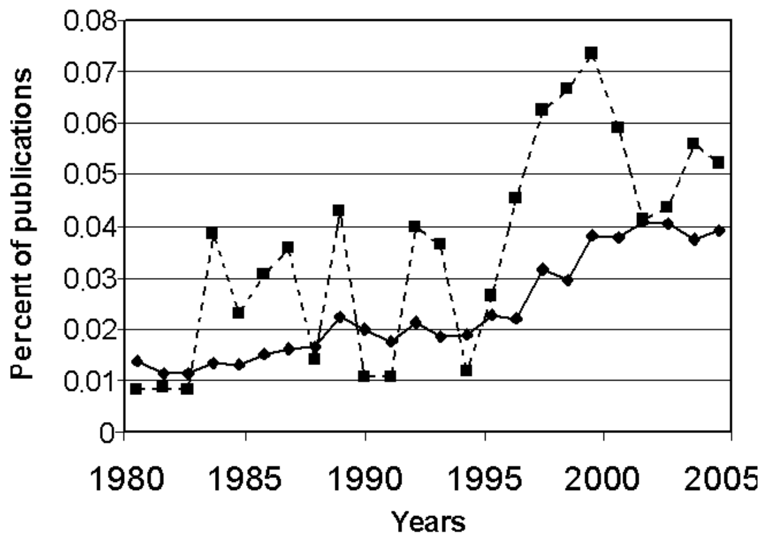

Figure 1: Percentage of publications on using factor analysis within a period of 1981-2004. Solid line - all papers, dashed line - papers on environmental sciences (from http://www.scirus.com and author's collection http://www.magniel.com/fa/data).

3 Factor analysis in environmental studies: common tendencies

Although the distribution of papers in the Internet does not fully correspond to their real publications, one can see a lot of interesting tendencies on the background of ongoing increase not only in the absolute, but also in the relative part of publications concerning various aspects of applications and development of factor analysis. We can see, for example, a wide dissemination of using factor analysis in environmental studies only since 1980's, when other sciences already have accumulated a rich experience in application of various methods of factor analysis and their development.

As psychology is a native home for factor analysis, it is not surprising, that the most experience has been revealed in psychological studies [2, 3, 5–7, 10]. On the early stages of development, its methodology also was developed according to specific requirements of psychology. Meanwhile, other sciences using similar methods of investigations began to adopt these methods. Therefore, the factor analysis has spread through psychiatry to medicine, through biology to chemistry, through sociology to political sciences and economics. Everywhere, new aspects of using factor analysis have been investigated and new methods have been developed. Thus, for instance, one should note schools of Malinowski in chemistry [11, 12], Rummel [13] in political sciences, Jöreskog [14] in geography, which have enriched a treasury of factor analysis with such methods as theory of errors, structural equation models, etc. Factor analysis in neurophysiology was applied by the author in 1971-1976 in studies of human brain organization and mechanisms of memory [15–19]. Such methods of factor analysis as dynamic, informational, hierarchical, etc., were developed in these works. Some of these methods are used in our days, although most of them were later independently rediscovered by other researchers (e.g., dynamic by Barber [20, 21], informational by Browne [22], hierarchical by Becker [23]).

While the first attempts using factor analysis in environmental studies began in 1960's (Garrett *et al.* [24]), serious attempts of using factor analysis in environmental sciences were undertaken only in 1980's, being based on the factor methodology developed, first of all, in chemistry [11, 12] and geology [14].

One of the first who accepted the concept of factor analysis in environmental sciences was Hopke [25–31] who used, adopted and developed factor technique for air quality analysis. For example, the method of three-mode factor analysis has primarily been employed in the social sciences providing the opportunity to examine data that are collected in form of a three-way matrix. With this method, one can simultaneously examine system variations in the three dimensions to determine the causal factors that control the system. On the basis of this method, *Fantasia* – a complex for target transformation factor analysis was developed in [25] to apportion sources in environmental samples and this approach has been applied to the receptor modeling problem that attempts to relate ambient air quality to sources of pollution [26]. Positive matrix factorization - a least squares approach to factor analysis was originally developed especially for environmental data and applied to several problems in resolving sources of environmental pollutants

[29–31].

Works of Aruga [32, 33] were directed to the specifics of environmental data to be studied by the logic and technique of factor analysis. The author analyzed which kind of data (normally distributed, standardized, transformed) are the best for a realistic factor analysis in environmental studies and showed effectiveness of factor analysis on the examples of study the causes of Po River pollution in the Piedmont region.

Numerous works of Einax *et al.* [1, 34–40] consider factor analysis as one of the methods of chemometrics. These works present the state of the art in environmental analysis and studies. Case studies show the enormous possibilities, and the limits, of chemometric methods, demonstrating the possibilities of factor analysis to detect spatial and temporal structures in data sets.

School of Simeonov and Tsakovsky [40–44] is known by applying methods of factor analysis to different aspects of environmental studies: atmosphere, hydrosphere, and pedosphere. The results of these studies are important not only in a local aspect as they allow quick response in finding solutions and decision making but also in a broader sense as a useful environmetrical methodology.

Jackson and Chen [45] applied robust principal component analysis for outlier detection with ecological data and identifying atypical observations. Because environmental studies frequently involve large numbers of variables and observations, and these are often subject to various errors, they tend to bias the interpretation and conclusion of an ecological study to identify atypical observations, that was very difficult using standard statistical approaches. Only the application of robust statistical methods could help in identifying atypical observations.

The effect of simulated outliers was studied by Chan and Shi [46] in application of PCA to climate studies. They showed how the method of projection-pursuit principal component analysis can be applied to analyze regional monthly sea surface temperature and rainfall. Comparisons were made with results derived from the traditional empirical orthogonal function method. The principal component analysis is shown to be much more robust than the empirical orthogonal function method and should be considered as an alternative in many of the climate studies.

Okuhara *et al.* [47] proposed using factor analysis to environmental data with probabilistic neural networks. They analyze observation data which consist of environmental factors as the explanatory variables and a population number of a creature (firefly) as the explained variable. The proposed system incorporates probabilistic neural networks which can acquire an unknown nonlinear mapping from the explanatory variables to the explained variable. The proposed system can estimate the effect of the explanatory variables on the explained variable, that is, it can solve the inverse problem. To realize the desired environment for the selected creature, authors showed that the proposed system can suggest an adequate strategy for the controllable explanatory variables.

The problem of effective dimensionality of environmental indicators is discussed in the work of Yu *et al.* [48]. In this paper, PCA is performed on 14 selected environmental indicators with "bootstrapped" confidence intervals. The term "bootstrap" refers to the process of randomly re-sampling the original sample set to generate new data sets and using these new data sets to make estimates of the statistic of interest. The objective is to derive some quasi-confidence intervals for the statistics when the underlying statistical distributions of the statistics are unknown. The analysis indicates that the first four principal components, which together account for more than 60% of the total variance in the original 14 variables, appear to be statistically significant based on the bootstrapped eigenvalue method, although the bootstrapped eigenvector method seems to be more conservative by identifying only the first two components as the significant ones. The first four principal components have large coefficients (eigenvectors) in absolute values with air, biodiversity, land, and water indicators, respectively. All these facts suggest that there is large redundancy in the existing environmental indicators. Consequently, to avoid overwhelming and confusing indicator-users including decision makers and the general public, developing four sub-indices representing air, water, land, and biodiversity should be the primary focus, which would probably capture the most important aspects of the environment.

While receptor modeling techniques have been repeatedly shown to be useful in quantifying the sources of urban aerosols, the estimation of contributions from distant pollution transport was proved to be more difficult. In the paper of Thurston and Lioy [49] was shown that Chemical Mass Balance (CMB) and multivariate receptor oriented models (e.g. *Principal Component Analysis*) each have their own strengths and shortcomings when addressing aerosols which have undergone significant transport. In particular, multivariate methods are preferable when doubt exists as to the identity and nature of sources influencing a monitoring site, while CMB models are most appropriate when all important sources and their downwind characteristics are known. As a result, it is concluded that these two approaches might best be used sequentially, with multivariate methods preceding CMB in transported aerosol assessments.

One should mention also the monograph of Preisendorfer and Mobley, [50], considering the use of principal component analysis in meteorology and oceanography.

Works on using factor analysis in specific fields of environmental sciences are reviewed in Sect. 4 (Air), Sect. 5 (Water), and Sect. 6 (Land). The fourth traditional part of environment sciences - Biodiversity - is out of the scope of this review.

4 Air

In this section, works concerned with chemical and physical characteristics of the air, air and ozone pollution, precipitation, and other aspects of climatology and meteorology, are reviewed.

4.1 Air pollution

Particulate matter (PM) components are the matter of the most factor analyses of air quality. Source apportionment of PM components in five major Chilean urban areas was investigated by Kavouras *et al.* [51]. Samples of mass and elemental concentrations of particles with diameter less than 10 μm (PM10) and 2.5 μm (PM2.5) were collected. For each of the five cities, factor analysis was applied to identify and quantify the sources of PM10 and PM2.5. The agreement between calculated and measured mass and elemental concentrations was excellent in most of the cities.

Advanced factor analysis of spatial distributions of PM was applied in the work of Paatero *et al.* [28]. This work analyzes PM2.5 24-h average concentrations measured every third day at over 300 locations in the eastern United States during 2000. The non-negative factor analytic model, positive matrix factorization (PMF), has been enhanced by modeling the dependence of PM2.5 concentrations on temperature, humidity, pressure, ozone concentrations, and wind velocity vectors. The model comprises 12 general factors, augmented by 5 urban-only factors intended to represent excess concentration present in urban locations only. The computed factor components or concentration fields are displayed as concentration maps, one for each factor, showing how much each factor contributes to the average concentration at each location. Three-mode factor analysis was used also by Zeng and Hopke [26] to the receptor modeling problem to relate ambient air quality to sources of pollution.

Identification of source nature and seasonal variations of Arctic aerosol by positive matrix factorization was studied in the works of Xie *et al.* [29–

31]. Week-long samples of airborne particulate matter were obtained at Alert (Northwest Territories, Canada). The concentrations of 24 particulate constituents have strong, persistent seasonal variations that depend on the transport from their sources. In order to explore the nature of the cyclical variation of different processes that give rise to the measured concentrations, the observations were arranged into both a two-way matrix and a three-way data array. For the latter, the three modes consist of chemical constituents, weeks within a year, and years. The two-way bilinear model and a three-way trilinear model were used to fit the data and a new data analysis technique, positive matrix factorization (PMF), has been used to obtain the solutions. PMF utilizes the error estimates of the observations to provide an optimal pointwise scaling data array for weighting, which enables it to handle missing data, a common occurrence in environmental measurements. It can also apply nonnegative constraints to the factors. Five factors have been obtained that reproduce the data quite well for both two-way and three-way analyses, where each factor represents a probable source with a compositional profile and distinctive seasonal variations: an acid photochemical factor, a soil factor, an anthropogenic factor, a sea salt factor, and a biogenic factor. The results also help to confirm the hypotheses regarding the origins of the Arctic aerosol.

A new approach to apportioning mass among various PCA source components: the calculation of Absolute Principal Component Scores, and the subsequent regression of daily mass and elemental concentrations on these scores was suggested and developed by Thurston and Spengler [52]. This method was applied to a quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston and allowed the estimation of mass and source particle characteristics for an unconventional source category: transported (coal combustion related) aerosols. This particle class was estimated to represent a major portion of the aerosol mass, averaging roughly 40 per cent of the fine mass and 25 per cent of the inhalable particle mass at the Watertown, MA site. About 45 per cent of the fine particle sulfur was ascribed to this one component, with only 20 per cent assigned to pollution from local sources.

Factor analysis of temporally and spatially resolved ambient samples was done by Jeon *et al.* [53]. The four main source patterns of organic PM components observed in solvent extraction (SX)-gas chromatography/mass spectrometry profiles of both temporally and spatially resolved receptor samples obtained in the El Paso/Juarez border airshed during the study period were interpreted to represent vehicular emissions plus resuspended urban dust; biomass combustion; native vegetation detritus and resuspended agricultural dust; and waste burning.

Statheropoulos *et al.* [54] used factor analysis for examining air pollution in the city of Athens. Five years data on CO, NO, NO_2, O_3, smoke and SO_2 concentrations were analyzed using PCA. Separate analyses were undertaken for summer and winter periods. It was found that the main principal components extracted from the air pollution data were related to gasoline combustion, oil combustion, and ozone interactions.

In the work of Chen *et al.* [55] origins of fine aerosol mass in the Baltimore-Washington corridor were investigated by factor analysis of chemically speciated fine particulate matter (PM2.5) and trace gases. A six-factor model, including regional sulfate, local sulfate, wood smoke, copper/iron processing industry, mobile, and secondary nitrate, was constructed and compared with reported source emission profiles.

Testing and optimizing two factor-analysis techniques on aerosol was presented by Huang *et al.* [56]. Elemental data for aerosol at Narragansett (RI, USA), were used to compare the source-identification power of positive matrix factorization (PMF), a new variant of factor analysis, with that of conventional factor analysis (CFA) and to investigate how much each technique can be "tuned" for best results. PMF was harder to use than CFA but resolved crustal and marine components up to an order of magnitude better. But the most important consideration was found to be the choice of elements, which outweighed all differences between techniques.

Assessment of air pollution sources in an industrial atmosphere was investigated by Pio *et al.* [57] using principal component and multilinear regression analysis. Aerosol samples collected in the industrial area of Estarreja (Portugal), were used to assess the source classes responsible for the particulate levels observed in the local atmosphere. PCA was applied separately to the concentrations of aerosol constituents and meteorological variables to obtain the number of factors and to verify the influence of weather conditions on ambient air quality. The technique led to the conclusion that soil and transport emissions represent important aerosol sources even in this industrial environment.

Resolution of air pollution from regional aerosol components was presented in the work of Fujimura *et al.* [58]. Aerosol was sampled in Western Japan during a kosa (yellow sand) dust event, with dust and pollution transport from the Asian mainland across the Sea of Japan. Factor analysis was applied for elemental concentrations and discovered two significant factors, representing soil and sea salt rich aerosols.

To order to identify sources of ambient air pollutants, Guo *et al.* [59] applied PCA to the data on non-methane hydrocarbons measured at toxic air pollutants monitoring stations in Hong Kong. This multivariate method enabled the identification of major air pollution sources along with the quantitative apportionment of each source to pollutant species. The PCA identified four major pollution sources. The extracted pollution sources included vehicular internal engine combustion with unburned fuel emissions, use of solvent particularly paints, liquefied petroleum gas or natural gas leakage, and industrial, commercial and domestic sources such as solvents, decoration, fuel combustion, chemical factories and power plants.

One should also mention a work of Brumelis *et al.* [60] using an artificial model of monitoring data to aid interpretation of PCA, where an artificial data matrix of element concentrations at sampling locations was created including six simulated gradients of correlated variables (Ca+Mg, Ni+V, Pb+Cu+Zn, Cd, Fe and K). This model represented a simplified model of a National Latvian survey.

4.2 Ozone pollution

Selection of the scenarios of ozone pollution at southern Taiwan using PCA was presented by Yu and Chang [61] The monitoring data analysis performed in this investigation focuses mainly on selecting statistically representative scenarios of ozone pollution. Evaluating the backward trajectories and spatial ozone profiles revealed that weak westerly sea breeze is the dominant factor affecting the production of the high ozone event for most stations.

The sources of photochemical precursors for ozone were evaluated by Buhr *et al.* [62] using PCA of concurrent measurements of $[NO_X]$, $[NO_Y]$ (total reactive oxidized nitrogen species), [CO], $[SO_2]$, $[C_3H_8]$, $[C_6H_6]$, and $[O_3]$ collected at a rural Alabama field site. The results of the analysis indicated that the major sources of NOY in the region are: coal-fired power plants and biomass burning and/or paper mills.

A climatology of total ozone was studied by Eder *et al.* [63]. The spatial and temporal variability of total column ozone obtained from the total ozone mapping spectrometer was examined. The rotated principal component analysis facilitated identification of the probable mechanisms responsible for the variability in homogeneous subregions. The mechanisms were either dynamic in nature (i.e., advection associated with baroclinic waves, the quasi-biennial oscillation, or El Niño-Southern Oscillation) or photochemical in nature (i.e., production of odd oxygen (O or O_3) associated with the annual progression of the Sun).

4.3 Atmospheric measurements

Li *et al.* [64] estimated primary and secondary production of HCHO in eastern North America based on gas phase measurements and principal component analysis. Based on atmospheric measurements of multiple species at Egbert, a rural site in Ontario (Canada), the emission ratios for area sources were estimated using a modified PCA technique. The technique yields three principal components that represent a photochemically aged air mass, a diurnal cycle, and fresh area emissions.

Apportionment of atmospheric aerosols was studied by Borbély *et al.* [65] using target transformation factor analysis. Wind-sector related regional signatures revealed a major contribution of Middle-East Europe to the atmospheric aerosol loading in Europe. The characteristics of the local aerosols were given in terms of source profiles and source scores.

PCA of the elements determined in the mosses was used by Berg *et al.* [66] to identify atmospheric trace element deposition. Dominant factors represented long-range atmospheric transported elements, windblown mineral particles, local emission sources, transport from the marine environment, and contribution from higher plants.

Local atmospheric dynamics studied by using of PCA of organic aerosols was presented in the work of Veltkam *et al.* [67]. The organic constituents of atmospheric aerosols collected at Niwot Ridge (Colorado), along with various physical and meteorological data, were measured during a collaborative field study. Volatile organic compounds were thermally desorbed from aerosol particles, separated by gas chromatography, and identified by mass spectrometry. For each of 48 samples, organic compounds in aerosol particles, organic and inorganic compounds in the vapor phase, wind direction, and time of day were measured. Relationships among the variables were analyzed by principal component analysis in order to examine the covariations within the data set. The 31 variables were grouped into seven factors, and individual compounds, as well as the factors, served as molecular markers for biologic and anthropogenic emission sources. In addition, factor scores were used to illustrate how several organic compounds vary with respect to local atmospheric dynamics.

Motelay *et al.* [68] showed the influence of meteorological parameters (temperature, precipitation amount) on polycyclic aromatic hydrocarbons (PAHs) concentration. PCA was used for PAHs deposition at a suburban site of Evreux located 100 km west of Paris (France).

PCA was used by Nagendra *et al.* [69] to analyze one-year traffic, emission and meteorological data for an urban intersection in Delhi. Data included meteorological, traffic and emission variables. In urban intersections the complexities of site, traffic and meteorological characteristic may result in a high cross correlation among the variables. Authors showed that in such situations PCA can provide an independent linear combination of the variables.

In the above mentioned work of Statheropoulos *et al.* [54], studying air pollution in the city of Athens, principal component analysis was also applied to meteorological data concerning relative humidity, temperature, sunshine duration, wind velocity and wind direction. The most prominent principal components from the meteorological data were related to dry conditions (summer period) and high-speed south-western winds (for both periods). The main relationship was found between total pollution and high humidity in combination with the low-velocity wind.

4.4 Precipitation

In 1980 Boutron and Martin [70] applied the principal component analysis to sources of trace metals in Antarctic snows. More than 3000 concentrations data obtained by the analysis of 12 trace metals in 250 snow samples collected in various locations in Antarctica were processed through PCA. The interpretation of the groups of so obtained covariant metals allowed estimation of relative contributions of the various aerosol sources to the trace metals content of Antarctic aerosols.

PCA was applied by Zhang *et al.* [71] to data obtained from the chemical analysis of rainwater for interpretation of rainwater composition. On the basis of the correlation between variables, the classification of samples into groups was investigated, and sources of correlation were identified.

Works of Adzuhata *et al.* [72, 73] were directed to chemical characterization and evaluation of ionic pollutants of acid fog and rain in Northern Japan using oblique rotational factor analysis. Fog/cloud and rain water were collected at the mountain side of Hachimantai range in northern Japan, and rain water was also collected at Akita City in order to investigate the air pollutant scavenging mechanism. The concentrations of various ions in these samples were analyzed, and the fog drop size and the wind direction were measured at each fog event. Three factors were extracted as the air pollutants, well-known as the cloud condensation nuclei, where the contribution of first factor was closely connected to the long-range transportation of anthropogenic or natural aerosol in air masses of continental origin.

Application of PCA in studies of precipitation in Tricity (Poland) was presented by Astel *et al.* [74]. Authors showed the results of monitoring and environmental pollution assessment for Gdańsk-Sopot-Gdynia Tricity (Poland). The obtained results allowed to determine the relations between variance of average concentration levels of analytes in rainfall samples, depending on the season and the amount of rainfall. Chemometrical analysis confirmed that the composition of inorganic pollution present in the precipitation samples taken within the Tricity was affected by the location of the area and proximity of the Baltic Sea. A characteristic feature of the region is the presence of industrial plants producing granulated phosphatic fertilizers, which contributes to the average level of PO_4^{3-} and F^- ion concentration in rainwater.

Ion concentrations in cloud water was also studied by Deininger and Saxena [75]. The PCA was used to study the relationship between the ionic constituents of the cloud water and the type of air mass in which the cloud formed. Using PCA, authors could identify the most significant acids and salts dissolved in the cloud water.

Long-term trend analysis of daily precipitation in Switzerland are presented in the work of Widmann and Schär [76]. Daily precipitation patterns over Switzerland were investigated by rotated and unrotated principal component analysis for the periods 1901-1990. Empirical orthogonal functions were utilized to homogenize the precipitation series and to optimally transform series of the long-term record into a few variables. Several statistically significant linear trends were detected. This includes, in particular, a wintertime increase in precipitation by up to 30% per 100 years in the western and northern parts of Switzerland.

Tsakovski, Simeonov *et al.* [41] applied PCA to seasonal and multivariate modelling study of wet precipitation data from the Austrian Monitoring Network. The aim of this work was to analyze the data structure of a large data set from rainwater samples collected during a long-term interval. Sampling sites from the network were chosen as data sources (chemical concentrations of major ions only) covering various location characteristics (height above the sea level, rural and urban sampling positions, Alpine rim and Alpine valley disposition, etc.). Several latent factors, named "anthropogenic", "crustal" and "mixed salt", were revealed by the multivariate modelling procedure (PCA) possessing a similar structure for most of the sites. The aim of the multivariate statistical study of simultaneously monitored cloud water, aerosol and rainwater data from different elevation levels in an alpine valley [44] was to extract information about latent factors determining the data structure in all of the cases in order to compare and interpret similarities and dissimilarities with respect to the elevation or the type of the atmospheric event. Four latent factors seem to explain over 85%

of the total variance for almost all sites and events but the factors have different identification for the different events or sites (e.g., “anthropogenic”, “crustal”, “neutralization”, “salt”). Thus, a comparison between sites and between events becomes possible. It was found that cloud water and aerosol events are much more similar with respect to data structure (relevant to emission sources or processes of formation) than the same events and rainwater.

The main purpose of the work of Benzi *et al.* [77] was the characterization of temperature and precipitation fields for better understanding of the Sardinia’s climate. PCA was applied to maximum and minimum temperature and to cumulative daily precipitation data collected in Sardinia. It was shown that the most significant temporal and spatial portion of temperature fields are described by the first principal components and that precipitation fields were well represented by the first three principal components.

Factors linking regional monthly sea surface temperature and rainfall were investigated by Chan and Shi [46] by the projection-pursuit PCA.

Forecasting all-India summer monsoon rainfall can be presented by pre-monsoon principal components of circulation fields covering the South Asian subcontinent. Cannon and McKendry [78] analyzed predictors for all-India summer monsoon rainfall using a bootstrap-based resampling procedure. Monsoon precursor signals represented by principal components were investigated and comparison made with a recent observational and general circulation modelling study. Pre-monsoon principal components formed a compact, interpretable, and significant set of predictors for all-India summer monsoon rainfall.

Principal components of monsoon rainfall in normal, flood and drought years over India were analyzed by Singh [79] using empirical orthogonal function analysis. It was found that during normal, flood and drought years, the first four (most dominated) principal component explains 73%, 77% and 100% of the variance, with all India seasonal mean rainfall.

The summer monsoon circulation shows various spatial and temporal oscillations and often a combination of systems produces an integrated effect. To investigate it, De and Mazumdar [80] used the principal component analysis of rainfall and southwest monsoon over India. In this study phases of the southwest monsoon were identified in an objective manner with the help of T-mode PCA of weekly rainfall anomalies. Mean composite charts were prepared utilizing all available upper air data for each category of the monsoon epochs identified by the principal component analysis. These sets of charts have been constructed for both the strong and weak phases associated with the first four significant principal components. The study suggests an objective method of interpretation of principal components by utilizing synoptic data.

Paper of Jury [81] should be also mentioned in this Section. This work presents the study of intra-seasonal convective variability over southern Africa by PCA of satellite outgoing-longwave radiation departures. Results of analysis were used to characterize the space and time scales of terrestrial cloudiness.

5 Water

This Section deals with applications of factor analysis to all aspects of water on the Earth, including lakes, rivers and oceans, ground- and sewage water, and also problems of pollution and purification.

5.1 Rivers and streams

Evaluation of water ambient quality was studied by Topalián *et al.* [82–84] using PCA. The Reconquista River (Buenos Aires Province, Argentina), one of the most polluted watercourses of Latin America, receives agrochemicals as well as domestic and industrial (mostly untreated) effluents. Physical and chemical water variables were determined [83], unispecies algal bioassays were carried out in laboratory; and density and structure of phyto- and zooplankton were analyzed as well [84]. Multivariate analyses of data showed a clear difference between stations for ammonium, ortophosphate, pH, hardness, chloride, phenols, and other values. In the multivariate analyses between seasons, the concentration of phenols appeared as an important feature. It was concluded that the deterioration of this water body was progressive downriver. The relatively best water quality was recorded whenwhere dissolved oxygen concentration, algal diversity and planktonic crustacean density were higher. The worst water quality corresponded to the lack of cladocerans and lowest crustacean density, and higher: organic and industrial pollution, major nutrients, hardness, conductivity, algal biomass in bioassays, phytoplankton density, dominance of a single algal species, and rotifer proportion in zooplankton.

The work of Haag and Westrich [85] demonstrates the usefulness of PCA in condensing and interpreting multivariate time-series of water quality data. In a case study the water quality system of the lock-regulated part of the River Neckar (Germany) was analyzed, with special emphasis on the oxygen budget. The analysis yielded four stable principal components, explaining 72% of the total variance of 11 parameters. The four components could be interpreted confidently in terms of underlying processes: biological activity, dilution by high discharge, seasonal effects, and the influence of wastewater.

River pollution data were interpreted by means of chemometric methods, including factor analysis, in papers of Einax *et al.* [36, 39, 40]. Sediments and suspended particulate matter taken from sampling sites along the River Elbe from the source to the mouth were subsequently processed by means of multivariate statistics in order to characterize the charge of the River Elbe with inorganic pollutants to elucidate pollution trends [39]. Environmetric study of Simeonov, Einax *et al.* [40] deals with modeling and interpretation of river water monitoring data from the basin of Saale River. Important information was revealed about the ecological status of the region of interest: identification of two separate patterns of pollution (upper and lower stream of the rivers); identification of six latent factors responsible for the data structure with different content for the two identified pollution patterns; and determination of the contribution of each latent factor (source of emission) to the formation of the total concentration of the chemical burden of the river water.

Using PCA to monitor spatial and temporal changes in water quality is presented in the work of Bengraïne and Marhaba [86]. Chemical, biological, and physical data monitored along the Passaic River (New Jersey), are analyzed. PCA was used to extract the factors associated with the hydrochemistry variability and to obtain the spatial and temporal changes in the water quality. Solute content, temperature, nutrients and organics were the main patterns extracted. This study showed the importance of environmental monitoring associated with simple but powerful statistics to understand better a complex water system.

Process identification by PCA of river water-quality data was studied in the work of Petersen *et al.* [87]. Time series of nutrient concentrations and related water quality parameters taken at several locations along the River Elbe were subjected to multivariate statistical analysis. The main question underlying in the study is concerned with whether known interactions between water quality variables can be recovered as statistically significant covariance patterns. Raw data and deviations from an estimated seasonal cycle were analyzed; two leading patterns of covariance were obtained, one discharge-dependent and the other related to biological activities

The high salinization in some reservoirs of the Contas River Basin (Bahia, Brazil) has been erroneously attributed only to concentration by evaporation. However, evaluation of the salt accumulation process during inundation applying PCA, presented in the work of dos Santos *et al.* [88], showed that in period of the intense rainfalls, the saline concentration in the flowing rivers of the reservoirs increases, and this fact can be attributed to the discharge of saline waters from small reservoirs of every drained area, provoked by inundation, is also an important factor in the salinization process. Thus the study of the geochemical variables: Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-} and CO_3^{2-}, showed one group formed by Na^+ and Cl^- and attributed to the discharge of saline water provoked by inundation from a small reservoir, and a second group constituted by Ca^{2+}, Mg^{2+}, K^+ and SO_4^{2-}, caused by an increase provoked by the evaporation in the salinization process.

Montes-Botella and Tenorio [89] studied water characterization and seasonal heavy metal distribution in the Odiel River (Huelva, Spain). PCA showed that variables related with the products of the pyrite oxidation and the salts that are solubilized by the high acidity generated in the oxidation of sulfides, grouped in the first component, accounted for 40% of total variance and were the main influential factor in physicochemical water sample properties. The second influential factor was minority metals (nickel, cobalt, cadmium).

The combined effects of multiple indices were analyzed by Yu *et al.* [90]. In this article, based on the environmental monitoring data, the water quality of the Songhua River (North East China) was analyzed using factor analysis, which comprehensively considered six indices of water quality of each monitoring section. The results showed that the main pollutants had changed to nitrogenous pollutants originated from nonpoint sources, and water quality was variable in different hydrological periods. The results also showed that the method was comprehensive and efficient in analyzing the dynamics of water quality.

Wayland *et al.* [91] analyzed relationships between baseflow stream geochemistry and land use. The purpose of this study was to examine the usefulness of the synoptic sampling approach for identifying the relationship between complex land use configurations and stream water quality. This study compares biogeochemical data from three synoptic sampling events representing the temporal variability of baseflow chemistry and land use using R-mode factor analysis. Separate R-mode factor analyses of the data from individual sampling events yielded only two consistent factors. Agricultural activity was associated with elevated levels of Ca^{2+}, Mg^{2+}, alkalinity, and frequently K^+, SO_4^{2-}, and NO_3^-. Urban areas were associated with higher concentrations of Na^+, K^+, and Cl^-. Other retained factors were not consistent among sampling events, and some factors were difficult to interpret in the context of biogeochemical sources and processes. When all data were combined, further associations were revealed such as an inverse relationship between the proportion of wetlands and stream nitrate concentrations.

A study of heavy metal pollution in the Tinto-Odiel estuary in southwestern Spain is presented in the work of Grande *et al.* [92]. The application of the factor analysis techniques on the nutrients and heavy metal concentrations in 46 water samples taken from 32 different sampling stations located along the estuary, allowed three groups of elements and compounds with a distinct origin to be determined. So, Cu and Zn have a clear fluvial provenance, whereas PO_4 and As are clearly industrial wastes and Cl, K, Ca, Li, Rb and Sr come from the sea. The existence of two agents controlling the behavior of the analyzed elements was deduced from factor analyses, which are: the tidal exchange with the open sea and the fluvial supply.

Evans *et al.* [93] applied factor analysis to investigate processes controlling the chemical composition of four streams in the Adirondack Mountains (New York). Four streams were monitored intensively over a 2 year period. Factor analysis was used to identify interrelationships between dissolved species during this period, and to determine physical processes controlling their behavior. Analysis of the full data set identified species which varied predominantly on an episodic time scale, and species which were subject to seasonal cycles.

5.2 Lakes and reservoirs

Factor analytical study on water quality in Lake Saimaa is presented in works of Mujunen *et al.* [94] and Reinikainen *et al.* [95]. Effects of various chlorinated and non-chlorinated organic compounds and some heavy metals discharged from pulp and paper mills into water, sediment and aquatic animals were studied in a recipient area of southern Lake Saimaa (Finland). The main aim of the project was to find an empirical link between chemical emission parameters and ecotoxicological effects expressed in the ecosystem. By using the multilinear model, three interpretable factors representing natural and anthropogenic processes could be extracted. The natural long-term variation, seasonal fluctuation and dilution of discharges in the recipient area could be extracted into their own factors, which could be easily visualized. The variation could be also presented with estimated variation in the water quality parameters caused by each of these natural or anthropogenic processes.

Loska and Wiechula [96] applied PCA for the estimation of source of heavy metal contamination in surface sediments from the Rybnik Reservoir (Poland). The bottom sediments are very heavily loaded with zinc, manganese, copper, nickel, phosphorus and lead (percentage enrichment factor), and cadmium, phosphorus and zinc (index of geoaccumulation). The increase of cadmium, lead, nickel, and zinc concentrations was connected with the inflow of the contaminated water of the river Ruda and long-range transport. The contamination of the reservoir with copper and manganese resulted mainly from atmospheric precipitation. The variability of the bottom sediment loading with metals during the investigations was affected in the first place by changes in the concentration of iron, but also those elements whose concentrations in the bottom sediment were elevated compared to the concentrations in shale - cadmium, nickel and lead.

Mackin *et al.* [97] studied elemental associations in the surface microlayer of Lake Michigan and its fluvial inputs, which were subjected to separate R mode factor analyses to define the geochemical phases and mechanisms which influence the composition of the surface microlayer. The associations revealed by these analyses indicated that the composition of the fluvial microlayers was controlled by localized factors related to the geology of individual drainage basins, while open lake microlayers were influenced by broad scale physicochemical interactions.

The factor analysis was used by Reisenhofer *et al.* [98] to verify the associations among variables and to separate factors responsible for the observed increase of the eutrophication of a shallow, temperate lake (San Daniele, North Eastern Italy). Ammonia, nitrite and nitrate nitrogen, dissolved oxygen, pH, temperature, total hardness, transparency and Zooplankton abundance were determined in lake water samples in order to monitor the eutrophication process. Nitrate fertilizers from surrounding farmland appear to be the main source of pollution. A secondary factor is constituted by the whole biomass, which is both origin and effect of the increasing productivity of the lake: the anaerobic decomposition of the organic debris in the hypolimnion is relevant.

Rachdawong *et al.* [99] used a factor analysis model with nonnegative constraints to apportion historical records of polycyclic aromatic hydrocarbon (PAH) sources in seven sediment cores from the central Lake Michigan area. The same approach was applied later by Bzdusek *et al.* [100] to apportion the sources of polycyclic aromatic hydrocarbons found in sediments of Lake Calumet and surrounding wetlands in southeast Chicago. Source profiles and contributions, with uncertainties, are determined with no prior knowledge of sources. The model includes scaling and backscaling of data with average PAH concentrations without sample normalization. Factor analysis resulted for a two-source solution indicating coke oven (45%) and traffic (55%) as the primary PAH sources to Lake Calumet sediments and providing new insights since wood burning and secondary coke oven profiles were not recognized in the Chemical Mass Balance model.

5.3 Aquifers and groundwater

A methodology of analyzing deep aquifers with scarce data was developed by Melloul [101] in 1995. This methodology was applied to the Nubian Sandstone aquifer beneath the Sinai Peninsula (in Egypt) and the Negev Desert (in Israel) to improve understanding of the hydrology of a deep aquifer with scarce data. PCA was used to combine various multidisciplinary data in order to identify chemical and physical groups, which were used to define groundwater flow paths. The findings of this study were in accord with the generally accepted hydrogeological conceptual model of an aquifer. However, through this study, new insights were obtained by the use of PCA method concerning: the description of a complex flow system by grouping various qualitative and quantitative parameters; the delineation of optimal operational zones for aquifer exploitation; and the definition and characterization of six main groundwater flow paths from their outcrops in the southern part of the Sinai to its discharge zones in the Arava River Valley and Dead Sea area in the Negev desert. These flow paths are defined by their water categories, which are unified expressions of such properties as salinity, origin, and age of groundwater.

Evolution of groundwater composition in an alluvial aquifer was studied by Helena *et al.* [102]. A set of quantitative analytical data from the alluvial aquifer of the Pisuerga river, located at the north-east of Valladolid (Spain), was processed by multivariate statistical techniques in order to investigate the evolution of the groundwater composition between two surveys. The original matrix consisted of 16 physicochemical variables; the exploration of the correlation matrix allowed to uncover strong associations between some variables as well as a lack of association between the others. PCA showed the existence of up to five significant principal components which account for 71% of the variance. Two of them can be initially assigned to "mineralization" whereas the other components are built from variables indicative of pollution. Varimax rotation allowed to "break up" the "mineralization" principal components into two varimax rotated principal components, assigned respectively to "natural" mineralization and to "saline" man-made contamination (sodium and chloride).

The high salinization in some sectors of the Castellon Plain aquifer (Spain) has been erroneously attributed to seawater intrusion, because of the high and increasing contents of chloride ions. However, application of principal components analysis by Morell *et al.* [103] showed that the chemical characteristics of groundwater are the result of three different components: intruding seawater, freshwater from rainfall infiltration and saline water with a characteristic sulphate-calcium-magnesium facies, derived from bordering aquifers.

The paper of Join *et al.* [104] describes the use of multivariate statistical analysis to trace groundwater circulation in volcanic terrains using PCA based on both structural and hydrochemical parameters of 243 springs of Reunion (Western Indian Ocean). This analysis was consistent with a geological and hydrogeological conceptual model developed from a combined hydrochemical and geological reconnaissance of 27 springs.

Characterization of groundwater contamination using factor analysis was presented by Subbarao *et al.* [105]. The effluent contamination of groundwater at two industrial sites at Visakhapatnam (India) was studied using factor analysis. Thirty groundwater samples near a zinc smelter plant and 19 from the polymers plant were analyzed for specific conductance, chloride, bicarbonate, sulfate, calcium, magnesium, sodium, and potassium. The data were subjected to R-mode factor analysis and the factor scores transferred to areal maps. While magnesium and sulfate are the dominant contaminants at the zinc site, sodium, chloride, and bicarbonate from the effluent are affecting groundwater in the polymers area. Contour maps for each factor suggest the areal extension of the contaminants.

Liu *et al.* [106] applied factor analysis to the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Factor analysis was applied to 28 groundwater samples collected from wells in the coastal Blackfoot disease area of Yun-Lin (Taiwan). Correlations among 13 hydrochemical parameters were statistically examined. A two-factor model was suggested and explained over 77.8% of the total groundwater quality variation. Factor 1 (seawater salinization) included concentrations of EC, TDS, Cl^-, SO_4^{2-}, Na^+, K^+ and Mg^{2+}, and factor 2 (arsenic pollutant) included concentrations of Alk, TOC and arsenic.

Factor analysis was applied by Jayakumar and Siraz [107] to hydrogeochemistry of coastal aquifers. R-mode factor analysis performed on major ion data from a hydrogeochemical survey over the coastal Quaternary deltaic aquifer of the Cauvery Basin (Tamil Nadu, India). Seven major ions (Ca, Mg, Na, K, HCO_3, Cl, and NO_3,) were analyzed from each of the 126 water samples collected in two seasons (pre- and post-monsoon, 63 samples for each). A set of factors was found both in pre-monsoon and post-monsoon data which explained the source of the dissolved ions and the chemical processes which accompany the intrusion of seawater.

5.4 Seas and oceans

Source input elucidation in polluted coastal systems was studied by Grimalt *et al.* [108–110]. The sedimentary hydrocarbon composition of a coastal system that receives the discharges of six rivers were studied by PCA and factor analysis. This study confirmed that factor analysis is consistent with the theoretical background of organic geochemistry in which the molecular composition of the environmental systems is interpreted to originate from a small number of sources. The usefulness of PCA and factor analysis for source input elucidation in environmental studies using molecular markers for sample description was evaluated in [110]. A case study involving the determination of aliphatic and chlorinated hydrocarbons, fatty acids, alcohols, chlorophylls, and some detergent indicators in water particulates from a deltaic system, was selected as a representative testing data set. PCA afforded useful results to differentiate between major groups of samples but not between geochemical sources. In contrast, factor analysis provided a direct correspondence between factor loadings and marker groups defining geochemically consistent organic matter contributions. For autochthonous compounds, factor analysis allowed an even more precise characterization of input sources than that obtained by the common "qualitative" molecular marker approach.

Barbieri *et al.* [111] modeled biogeochemical interactions in the surface waters of the Gulf of Trieste by three-way PCA. Data of temperature, salinity, dissolved oxygen, nutrients and chlorophyll measured on samples of surface seawater and collected monthly during 2 years in different sites of the Gulf of Trieste were modeled by means of three-way PCA. Physicochemical parameters were described by three different components that explained the effect of the river input on the seawater pattern, the effect of temperature, and metabolic-catabolic activity of the phytoplankton, respectively. One spatial component accounted for the gradient of influence of the estuarine waters in the Gulf, and three temporal components characterized three main seasonal conditions.

An environmental study of surface seawaters in the Gulf of Valencia was presented by Morales *et al.* [112] A study was made on the quality of coastal waters in the Gulf of Valencia (Spain) in terms of contamination markers including microbiological agents, toxic heavy metals and nutrients that adversely affect the environment. Relationships were also established between these factors and other physical and chemical parameters. PCA allowed the characterization of the coastal water quality of the study zone, establishing the sources and types of contamination, and identifying the littoral areas associated to the different types of contamination.

Multivariate statistical analysis of sediment data collected from the western coastline of the USA and analyzed for 15 analytes by Simeonov *et al.* [43] indicated that the data structure could be explained by four latent factors. These factors are conditionally named "anthropogenic", "organic", "natural", and "hot spots". They explain over 85% of the total variance of the data system, which is an acceptable value for the PCA model. A study of metal pollution in [42] was based on multivariate statistical modeling of "hot spot" sediments from the Black Sea. Application of PCA to separate zones of the marine environment with different levels of pollution by interpretation of the sediment analysis. The extraction of four latent factors offered a specific interpretation of the possible pollution sources and separates natural from anthropogenic factors, the latter originating from contamination by chemical, oil refinery and steel-work enterprises. Finally, the PLSs modeling gave a better opportunity in predicting contaminant concentration on tracer element as compared to the one-dimensional approach of the baseline models.

Maes and Behringer [113] used PCA to estimate the upper oceanic variability of salinity of western tropical Pacific Ocean. The method was based on combined vertical modes of temperature and salinity, and reconstructs salinity profiles via a weighted least-squares procedure. The modes were defined as the empirical orthogonal functions along the water column.

Bottom water formation in the Weddell Sea was analyzed by Lindegren and Josefson [114]. The general mixing situation in the southern Weddell Sea was studied by PCA and applied to the formation of Antarctic Bottom Water in the southern part of the Weddell Sea. Three source waters were indicated and the fractions of the two source waters, ice shelf water and warm deep water were calculated.

PCA of satellite passive microwave data over sea ice was presented by Rothrock *et al.* [115]. The 10 channels of scanning multichannel microwave radiometer data for the Arctic were examined by principal component analyses. Only the first two principal components contained variance due to the mixture of surface types. Three component mixtures (water, first-year ice, and multiyear ice) could be resolved in two dimensions.

Wensnahan *et al.* [116] used PCA in a case study of special sensor microwave imager from the Bering Sea. Winds from the north formed thin ice areas which were interpreted as large amounts of open water and multiyear ice. With PCA, these same areas are interpreted as 20-30 open water near the lee shores but otherwise as consisting almost entirely of thin ice. Authors came to conclusion that thin ice can be detected using satellite data.

5.5 "Mussel watch"

PCA monitored by observations of mussels ("mussel watch") was applied first in 1984-1989 by Favretto *et al.* [117–119] in the Gulf of Trieste. From six to ten trace elements were determined in the dissolved ash of the edible part of wild mussels from a polluted site by electrothermal atomic absorption spectrometry. The correlation matrix around the mean was used as a starting matrix for PCA. All variables were reduced to two principal components, accounting for 77% of the total variance. The orthogonally rotated factor matrix indicates that Co and Ni are bonded to the first principal component and Cd and Pb to the second principal component. Al, Cr, Mn, Fe, Zn, Cd, and Pb are all positively associated with the first principal component and form a cluster of variables, indicating a common terrigenous origin.

Data analysis of heavy metal pollution in the sea was studied by Piepponen and Lindström [120] by using PCA. Environmental heavy metal pollution in the coastal waters of the Bothnian Sea near the city of Pori was monitored by observations of mussels. The concentrations of heavy metals in the soft tissues of the mussels and in different fractions of the shells were determined by wet digestion and atomic absorption spectrometry. The study showed that the elemental variables Fe, Ti and V were most strongly related to the titanium oxide industry and Al, Co, Hg and Mn to the river Kokemäenjoki.

Trace metals on the Algarve coast were analyzed by Machado *et al.* [121] Cadmium, copper, iron, manganese, nickel, and zinc concentrations were determined by atomic absorption spectrophotometry in samples of soft tissues of the mussels. Metal concentrations increased near urban centres and sources of industrial effluents. Consequently, the areas influenced by the Arade River, the Guadiana estuary and the Formosa Ria lagoon presented the highest metal concentrations. Metal concentrations in mussels from the west coast of the Algarve were higher than in those from the east coast. Cd and Cu concentrations in mussels from the different sampling points have increased over the last 10 years, while Fe, Mn, and Ni concentrations along the Algarve Coast have fallen. These results were discussed in relation to the variation of human impact, some environmental factors, and other natural phenomena.

Stella *et al.* [122] used polycyclic aromatic hydrocarbons analyses for source identification. Polycyclic aromatic hydrocarbons (PAHs) biomonitored in the aquatic environment by means of caged mussels were compared by site and by season. Moreover, their fingerprints were compared to marine sediments and atmospheric airborne PAHs. The characterization of the sampling stations by means of PCA allowed distinguishing the prevalence of pyrogenic or petrogenic types of pollution and between two kinds of combustibles.

5.6 Water supply, waste and sewage water

A three-way principal factor analysis for assessing the time variability of freshwaters related to a municipal water supply was applied by Barbieri *et al.* [123]. Chemical analyses, physical data and biological monitors constitute the 15 parameters, on freshwaters sampled at seven sites in a karstic area of northeastern Italy. The data set was used for a three-way principal factor analysis aimed at exploring the pattern of information about the environmental quality of the monitored freshwaters, since four wells are feeding the municipal water supply of the Province of Trieste.

Vidal *et al.* [124] used factor analysis for the study of water resources contamination due to the use of livestock slurries as fertilizer. Authors investigated the effects of slurry application on water quality in wells, pasture-drainage conduits and rivers. The first axis extracted by PCA of the samples-by-variables matrix represented the degree of dilution of the water strongly related with saline content; the second axis represented redox conditions, affecting organometallic component. In general, the positions of the samples in this factor space reflected the major contamination processes affecting water resources of that type (wells, conduits or rivers).

A methodology of supervisory control of wastewater treatment plants was presented in works of Rosen and Yuan [125] and Lennox and Rosen [126]. In these papers a methodology for integrated multivariate monitoring and control of biological wastewater treatment plants during extreme events was presented. To monitor the process, on-line dynamic PCA was performed on the process data to extract the principal components that represent the underlying mechanisms of the process. Fuzzy c-means (FCM) clustering is used to classify the operational state. Performing clustering on scores from PCA solved computational problems as well as increases robustness due to noise attenuation. The class-membership information from FCM was used to derive adequate control set points for the local control loops. The methodology was illustrated by a simulation study of a biological wastewater treatment plant, on which disturbances of various types are imposed. The results showed that the methodology can be used to determine and coordinate control actions in order to shift the control objective and improve the effluent quality.

The ecological hygiene assessment of the water environment was studied by Krasovskii and Vorob'eva [127] using factor analysis. Regularities of the formation and spread of pollution of water objects in the area of sewage discharges of the cellulose-paper works were revealed. The bottom nature of water pollution and their change in their contribution in relation to the distance were determined. The bottom pollution of water was comparable with that due to sewage and varied from 46 to 69%. The transformation of organic matter makes a significant contribution to the total water pollution.

Critto *et al.* [128] applied PCA to characterization of contaminated soil and groundwater surrounding. The characterization of a hydrologically complex contaminated site bordering the lagoon of Venice (Italy) was undertaken by investigating soils and groundwaters affected by chemical contaminants originated by the wastes dumped into an illegal landfill.

6 Land

Geochemical and geophysical characteristics of the Earth surface are the object of research in this Section.

6.1 Soil

One of the major interests in soil analysis is the integrated evaluation of soil properties, which might be indicators of soil quality. Factor analysis is a powerful tools for this integrated assessment and can help soil researchers to extract much more information from their data. A multivariate study was carried out by Sena *et al.* [129] in three farms from Guaíra, State of São Paulo, Brazil. Conventionally managed plots that intensively utilized pesticides and chemical fertilizers were compared with both non-disturbed forest areas and alternatively managed plots. The latter were under ecological farming employing effective microorganisms integrated with crop residues. Eight soil parameters were determined for each plot. The multivariate approach of principal component analysis allowed to distinguish the areas as a function of the soil management and determine which are the most important parameters to characterize them. The forest areas presented higher microbial biomass with lower cellulolytic population than at cultivated sites. The alternative plots were characterized by higher microbial biomass and polysaccharide content with lower phosphate solubilizers and cellulolytic microorganisms colony counts than at the conventional areas.

Barona and Romero [130] studied the distribution of metals in soils and relationships among fractions. PCA was used as a method for data treatment to establish general relationships among metal amounts accumulated in different fractions and general soil properties, which were expected to govern the metal distribution pattern. On the basis of the results obtained from the position maps of variables and samples, the carbonate content was the soil property with the greatest number of statistically significant correlations with metal contents in fractions, so it can be considered as a relevant parameter in the distribution of some metals such as Pb, Ni, Zn, and Cu in the soils.

Factor analysis of soil heavy metal pollution was also used by Lin *et al.* [131] in study landscape indices of 55 sampling sites in Changhua county in Taiwan to characterize the factor patterns of eight soil heavy metals and the interrelation patterns of these soil heavy metals, landscape and human activities. Factor analyses revealed that soil heavy metals and data concerning landscape data could be grouped into a six-factor model that accounts for 82% of all the variation of data. Moreover, the first factor included the concentration of Cd, Cr, Cu, Ni, and Zn, and urbanization and industrialization landscape indices.

R-mode factor analysis was applied by Kumru and Bakaç [132] to the distribution of elements in soils from the Aydin basin, Turkey to describe the relationship among 15 remotely sensed, geochemical and industrial variables.

Maiz *et al.* [133] studied the evaluation of heavy metal availability in polluted soils by two sequential extraction procedures using factor analysis. Superficial soil and grass samples from 13 locations affected by several anthropogenic sources (mining, metal factory, traffic emissions) were collected in Gipuzkoa, northern Spain. Factor analysis was used to check the associations between the total metal contents in soil and grass, as well as between the levels of the different sequential fractions and the total content in grass.

The complexity and the large variance of environmental data sets limit the use of common statistical methods for the assessment of the state of pollution. Therefore, the application of geostatistical and multivariate statistical methods for the assessment of polluted soils was applied by Einax *et al.* [37, 38]. Geostatistical and multivariate methods of data analysis were used to describe patterns of soil pollution with inorganic contaminants in Celje County, Slovenia. Groups of contaminants and polluted sites were identified using cluster analysis and confirmed with multidimensional variance and discriminant analysis. Factor analysis yields an identification of not directly observable relationships between the contaminants.

Relationships between soil bulk electrical conductivity and the principal component analysis of topography and soil fertility values are presented in the paper of Officer *et al.* [134]. Authors showed that PCA can be applied to create meaningful field scale summaries of groups of attributes and to decrease the estimation error of maps of the summarized attributes.

Factor analysis of the properties of volcanic soil constituents was applied by Meijer and P. Buurman [135]. The variation of fourteen soil chemical and physical properties of twenty soil samples from Andosols was decomposed into the contributions of seven soil constituents or end-members. The samples were from the slopes of the andesitic Turrialba volcano in Costa Rica. Factor analysis of the data explained 98% of the variance by six orthogonal factors.

Analysis of the relationship between soil and vegetation in forest biogeocenoses was studied by Koptsik *et al.* [136]. Coordinated soil-geobotanical studies revealed a close correlation between the species diversity of phytocenoses and soil properties in the Russkii Sever National Nature Park (Vologda oblast).

Factor analysis of nutrient distribution patterns under shrub live-oak in two contrasting soils was applied by Brejda [137]. The objectives of this research were to identify underlying patterns in soil properties using factor analysis, and analyze factor scores to determine how the factor patterns varied between soils, canopy covers, and depth. Factor analysis provided a statistical tool for grouping the 11 correlated soil variables into three uncorrelated factors. Analysis of factor scores allowed independent assessment of soils, shrub cover, depth, and their interactions on soil properties.

6.2 Geomorphology, geophysics and geochemistry

An advantage in regional geochemistry would be that instead of presenting maps for 40-50 (or more) elements only maps of 4-6 factors may have to be presented, containing a high percentage of the information of the single element maps. Factor analysis was first applied to this problem in 1969 by Garrett and Nichol [24] and Garret [138] to the interpretation of regional geochemical stream sediment data. Chork [139] studied exploration cheochemical data from sheeted-vein tin mineralization near Emmaville (N.W.S, Australia), and Chork and Salminen [140] interpreted geochemical data from Outokumpu (Finland).

Filzmoser [141] and Reimann [142] developed the methodical aspects of using factor analysis in the geostatistical treatment of environmental data by the usage of robust multivariate statistical methods in geostatistics. They emphasized using robust principal component and factor analysis for the preliminary investigation of the data to reduce the dimension. Geostatistical methods were applied afterwards to the estimated factor scores. The final results showed the influence of certain combinations of variables in the considered region. Moreover, the estimated factor scores with the robust procedure indicate outlying observations in a much better way.

Problems of evaluating and weighting of geophysical data were presented in the work of Chung and Nigam [143] using PCA. The purpose of this paper was to encourage the use of field weighting over field interpolation in achieving grid-area parity in principal component analysis of geophysical data. While field weighting was shown to achieve exact parity in unrotated analysis, authors showed that it is highly effective in rotated analysis, too.

Pereira *et al.* [144] proposed a methodology for study on geochemical anomaly identification. Based on a case study in which a single geochemical anomaly was located in the vicinity of an abandoned mine in Central Portugal, a recursive methodology for anomaly/background separation was developed. This methodology relies on the supplementary projection of each of the samples taken from a subset of "anomaly candidates" onto the axes provided by PCA of the background subset. The concept of "anomaly intensity", defined by the average of the distances from the original to the supplementary projections, is the basis for final anomaly identification.

PCA was applied by Cuadrado and Perillo [145] to geomorphologic evolution in the study of a sector of the main channel of the entrance to Bahía Blanca system harbour (Argentina). El Toro channel, characterized by recurrent accumulation processes, had to be dredged periodically to maintain the nominal depth of 10 m. Detailed surveys of the reach are made regularly to check the navigation conditions. A set of survey charts made within two dredging operations and covering about 1 year was analyzed by means of PCA. The first principal component obtained from PCA describes the mean depth of the area, while the second principal component explains the morphological variations over time. From it, the accumulation periodicity can be detected, and, therefore, the time of necessary dredging indicated.

Principal components of the topographical environment were studied by Wotling *et al.* [146] by analyzing regionalization of extreme precipitation distribution and dealing with the data of extreme rainfall intensities in the volcanic island of Tahiti (French Polynesia). The paper showed how the method automatically takes into account the topographical relief features.

Maldonado *et al.* [147] showed that the land use dynamics can be characterized by PCA. The multitemporal analysis of changes in the Caatinga land cover (Pernambuco State, North-East Brazil) provided sufficient information about the dynamics of this typical land use. Within this frame, PCA was applied in combination with field survey data, which permitted estimation of point samples of new recovery/degradation. Five classes of changed and unchanged multitemporal effects were discriminated.

7 Conclusion

As we could see, factor analysis is successfully used in various fields of environmental sciences. Some of discussed methods have been especially developed for environmental studies.

The common approach of using factor analysis in environmental studies leads to the synthesis of the following successive logical structure:

1. Dividing the system into sets of “elementary” components.
2. Analysis of the relations of these components in space or in time.
3. Revealing system-forming relations.
4. Description of the structure of the system (model) and its properties (forecast).

Let us illustrate this scheme on the example of study of water quality by the work of Bengraïne and Marhaba [86]. In this work various chemical, biological, and physical data of the studied river were choosen as elementary components. Matrix of realization of elementary components was observed in space (along the river). Cross-correlations between these components were chosen as objects of analysis. The purpose of analysis was to define system-forming relations - factors associated with the hydrochemistry variability, which are in these research - solute content, temperature, nutrients and organics were the main patterns extracted.

Such or similar scheme was applied in the most of reviewed methods. It demonstrates applicability and usefulness of factor analysis in environmental studies. Taking into account a wide dissemination of using factor analysis in these studies, one can confirm that in our days this approach should be considered as one of the main techniques intended to the investigation of the structural-functional organization of the system.

One should not, however, forget that factor analysis does not always give a possibility of the pithy interpretation of factors. The interpretation must be based on the data of a nature and properties of elements of the system, obtained by other methods. Factor analysis in this sense is only a link among the other stages of investigation; the connection with these links must be always maintained, and only the whole chain can lead to the solution of a problem [10]. Only the breadth of erudition of researchers and knowledge of principles of the functional integration of investigated systems are able to create a necessary basis for the objective interpretation of revealing factors.

I gratefully acknowledge very constructive remarks, fruitful comments and practical advices of Prof. E. Levner and anonymous reviewers.

References

[1] J.W. Einax, H.W. Zwanziger, and S. Geiß, *Chemometrics in Environmental Analysis* (Wiley-VCH, Weinheim, 1997).

[2] H.H. Harman, *Modern Factor Analysis* (U. of Chicago Press, Chicago, 1960).

[3] D.N. Lawley and A.E. Maxwell, *Factor Analysis as a Statistical Method* (2-nd ed., Butterworths, London, 1971).

[4] C. Spearman, Am. J. Psychol. **25**, 201 (1904).

[5] L.L. Thurstone, *Vectors of the Mind* (U. of Chicago Press, Chicago, 1935).

[6] D.N. Lawley, Proc. Roy. Soc. Edinburgh A **60**, 82 (1940).

[7] L.L. Thurstone, *Multiple-Factor Analysis* (U. of Chicago Press, Chicago, 1947).

[8] H. Hotelling, J. Educ. Psychol. **24**, 417 (1955).

[9] T.W. Anderson and H. Rubin, In: Proc. Third Berkeley Symp. Math. Stat. Probab., Ed. J. Neyman, vol.**5**, p.111 (U. of California Press, Berkeley, CA, 1956).

[10] M. Reuchlin *Les Méthodes Quantitatives en Psychologie* (PUF, Paris, 1962).

[11] E.R. Malinowski and D.G. Howery, *Factor Analysis in Chemistry* (Wiley-Interscience, N.Y., 1980).

[12] E.R. Malinowski, *Factor Analysis in Chemistry*, 2nd ed. (Wiley, Chichester, 1991).

[13] R.J. Rummel, *Applied Factor Analysis* (Northwestern U. Press, Evanston, 1970).

[14] K.G. Jöreskog, J.E. Klovan, and R.A. Reyment, *Geological Factor Analysis (Methods in Geomathematics)* (Elsevier, 1976).

[15] P.V. Bundzen, N.N. Vasilevsky, A.S. Kaplunovsky, and V.V. Shabaev, Sechenov Fiziol. Zh. SSSR **57**, 969 (1971).

[16] A.S. Kaplunovsky and M.M. Bogoslovsky, Sechenov Fiziol. Zh. SSSR **59**, 1291 (1973).

[17] P.V. Bundzen, Yu.L. Gogolitsin and A.S. Kaplunovsky In: *Biological Diagnosis of Brain Disorders*, Ed.: S.Bogoch, p.275 (Spectrum Publ., N.Y., 1973).

[18] A.S. Kaplunovsky, In: *Memory in Normal and Pathological Reactions of Organism*, p.161 (Leningrad, 1974).

[19] P.V. Bundzen, A.S. Kaplunovsky, V.M. Klimenko, and E.A. Korneva, In: *Methodological Problems of Theoretical Medicine* Ed.: N.P. Bekhtereva, p.25 (Medicina, Leningrad, 1975).

[20] D.C. Barber, Phys. Med. Biol. **21**, 792 (1976).

[21] D.C. Barber, Phys. Med. Biol. **25**, 283 (1980).

[22] M.W. Browne, Linear Algebra Appl. **70**, 51 (1985).

[23] G. Becker, Psychol. Meth. **1**, 341 (1996).

[24] R.G. Garrett and I. Nichol, In: *Int. Geochem. Explor. Symp.*, Ed.: F.C. Canney, p.245 (Quarterly Colorado School of Mines, Golden, CO, 1969).

[25] P.K. Hopke, D.J. Alpert and B.A. Roscoe, Comput. Chem. **7**, 149 (1983).

[26] Y. Zeng and P.K. Hopke, Chemometr. Intell. Lab. Syst. **7**, 237 (1990).

[27] Y.L. Xie, P.K. Hopke, and P. Paatero, J. Chemom. **12**, 357 (1998).

[28] P. Paatero, P.K. Hopke, J. Hoppenstock, and S.I. Eberly, Environ. Sci. Technol. **37**, 2460 (2003).

[29] Y.L. Xie, P.K. Hopke, P. Paatero, L.A. Barrie, and S.M. Li, J. Atmos. Sci. **56**, 249 (1999).

[30] Y.L. Xie, P.K. Hopke, P. Paatero, L.A. Barrie, and S.M. Li, Atmos. Environ. **33**, 2229 (1999).

[31] Y.L. Xie, P.K. Hopke, P. Paatero, L.A. Barrie, and S.M. Li, Atmos. Environ. **33**, 2549 (1999).

[32] R. Aruga, Anal. Chim. Acta **354**, 255 (1997).

[33] R. Aruga, Ann. Chim. - Rome **93**, 1013 (2003).

[34] In: *The Handbook of Environmental Chemistry series*, Ed.: O. Hutzinger. *Statistical Methods* 2G, Ed. J. Einax, (Springer, Berlin, 1995); *ibid*, *Applications* 2H, Ed. J. Einax, (Springer, Berlin, 1995).

[35] M. Kreig, J. Einax, Fresenius J. Anal. Chem. **348**, 490 (1994).

[36] J.W. Einax, D. Truckenbrodt, and O. Kampe, Microchem. J. **58**, 315 (1998).

[37] J.W. Einax and U. Soldt, Chemom. Intell. Lab. Syst. **46**, 79 (1999).

[38] M. Zupan, J. Einax, J. Kraft, F. Lobnik, and V. Hudnik, Environ. Sci. Pollution Res. **7**, 89 (2000).

[39] A. Aulinger, A. Prange, R. Niedergesass, S. Schmolke, and J.W. Einax J. Environ. Monit. **4**, 942 (2002).

[40] V. Simeonov, J.W. Einax, I. Stanimirova, and J. Kraft, Anal. Bioanal. Chem. **374**, 898 (2002).

[41] S. Tsakovski, H. Puxbaum, V. Simeonov, M. Kalina, H. Loffler, G. Heimburger, P. Biebl, A. Weber, and A. Damm, J. Environ. Monit. **2**, 424 (2000).

[42] V. Simeonov, D.L. Massart, G. Andreev, and S. Tsakovski, Chemosphere **41**, 1411 (2000).

[43] V. Simeonov, I. Stainimirova, and S. Tsakovski, Fresenius J. Anal. Chem. **370**, 719 (2001).

[44] V. Simeonov, M. Kalina, S. Tsakovski and H. Puxbaum, Talanta **61**, 519 (2003).

[45] D.A. Jackson and Y. Chen, Environmetrics **15**, 129 (2004).

[46] J.C.L. Chan and J-E. Shi, Int. J. Climatol. **17**, 103 (1997).

[47] K. Okuhara, Y. Matsubara, W. Shiraki, and M. Gotou, Electron. Comm. Jpn. Pt 3, **85**, 18 (2002).

[48] C-C. Yu, J.T. Quinn, C.M. Dufournaud, J.J. Harrington, P.P. Rogers, and B.N. Lohani, J. Environ. Manag. **53**, 101 (1998).

[49] G.D. Thurston and P.J. Lioy, Atmos. Environ. **21**, 687 (1987).

[50] R.W. Preisendorfer and C.D. Mobley, *Principal Component Analysis in Meteorology and Oceanography* (Elsevier, 1988).

[51] I.G. Kavouras, P. Koutrakis, F. Cereceda-Balic, and P. Oyola, J. Air Waste Manag. Assoc. **51**, 451 (2001).

[52] G.D. Thurston and J.D. Spengler, Atmos. Environ. **19**, 9 (1985).

[53] S.J. Jeon, H.L. Meuzelaar, S.A. Sheya, J.S. Lighty, W.M. Jarman, C. Kasteler, A.F. Sarofim, and B.R. Simoneit, J. Air Waste Manag. Assoc. **51**, 766 (2001).

[54] M. Statheropoulos, N. Vassiliadis, and A. Pappa, Atmos. Environ. **32**, 1087 (1998).

[55] L.A. Chen, B.G. Doddridge, R.R. Dickerson, J.C. Chow, and R.C. Henry, Atmos. Environ. **36**, 4541 (2002).

[56] S. Huang, K.A. Rahn, and R. Arimoto, Atmos. Environ. **33**, 2169 (1999).

[57] C.A. Pio, T.V. Nunes, C.S. Borrego, and J.G. Martins, Sci. Total Environ. **80**, 279 (1989).

[58] M. Fujimura, H. Suzuki, Y. Sekine, and J.W. Winchester, Nucl. Instr. Meth. Phys. Res. B **75**, 292 (1993).

[59] H. Guo, T. Wang, and P.K.K. Louie, Environ. Pollut. **129**, 489 (2004).

[60] G. Brumelis, L. Lapina, O. Nikodemus, and G. Tabors, Environ. Modell. Softw. **15**, 755 (2000).

[61] T-Y. Yu and L-F.W. Chang, Atmos. Environ. **34**, 4499 (2000).

[62] M. Buhr, D. Parrish, J. Elliot, J. Holloway, J. Carpenter, P. Goldan, W. Kuster, M. Trainer, S. Montzka, S. McKeen, and F. Fehsenfeld, J. Geophys. Res. D **100**, 22853 (1995).

[63] B.K. Eder, S.K. LeDuc, and J.E. Sickles II, J. Geophys. Res. D **104**, 3691 (1999).

[64] S-M. Li, K.G. Anlauf, H.A. Wiebe, and J.W. Bottenheim, Geophys. Res. Lett. **21**, 669 (1994).

[65] I. Borbély-Kiss, E. Koltay, and G.Y. Szabó, Nucl. Instr. Meth. Phys. Res. B **75**, 287 (1993).

[66] T. Berg, O. Røyset, E. Steinnes, and M. Vadset, Environ. Pollut. **88**, 67 (1995).

[67] P.R. Veltkamp, K.J. Hansen, R.M. Barkley, and R.E. Sievers, J. Geophys. Res. D **101**, 19495 (1996).

[68] A. Motelay-Massei, D. Ollivon, B. Garban, and M. Chevreuil, Atmos. Environ. **37**, 3135 (2003).

[69] S.M.S. Nagendra, and M. Khare, Transport. Res. D **8**, 285 (2003).

[70] C. Boutron and S. Martin, J. Geophys. Res. C **85**, 5631 (1980).

[71] P. Zhang, N. Dudley, A.M. Ure, and D. Littlejohn, Anal. Chim. Acta **258**, 1 (1992).

[72] T. Adzuhata, J. Inotsume, T. Okamura, R. Kikuchi, T. Ozeki, M. Kajikawa, and N. Ogawa, Anal. Sci. **17**, 71 (2001).

[73] T. Adzuhata, T. Okamura, J. Inotsume, R. Kikuchi, T. Ozeki, M. Kajikawa, and N. Ogawa, Water Air Soil Poll. **130**, 337 (2001)

[74] A. Astel, J. Mazerski, Z. Polkowska, and J. Namiesnik, Adv. Environ. Res. **8**, 337 (2004).

[75] C.K. Deininger and V.K. Saxena, Atmos. Environ. **31**, 295 (1997).

[76] M. Widmann and C. Schär, Int. J. Climatol. **17**, 1333 (1997).

[77] R. Benzi, R. Deidda, and M. Marrocu, Int. J. Climatol. **17**, 1231 (1997).

[78] A.J. Cannon and I.G. McKendry, Int. J. Climatol. **19**, 1561 (1999).

[79] C.V. Singh, Int. J. Climatol. **19**, 639 (1999).

[80] U.S. De and A.B. Mazumdar, Theor. Appl. Climatol. **64**, 213 (1999).

[81] M.R. Jury, Theor. Appl. Climatol. **62**, 133 (1999).

[82] M.L. Topalián, P.M. Castañé, M.G. Rovedatti, and A. Salibián, Bull. Environ. Contam. Tox. **63**, 484 (1999).

[83] M.L. Topalián, M.G. Rovedatti, P.M. Castañé, and A. Salibián, Water Air Soil Poll. **114**, 287 (1999).

[84] H.F. Olguín, A. Puig, C.R. Loez, A.Salibián, M.L. Topalián, P.M. Castañé, and M.G. Rovedatti, Water Air Soil Poll. **155**, 355 (2004)

[85] I. Haag and B. Westrich, Hydrological Processes **16**, 3113 (2002).

[86] K. Bengraïne and T.F. Marhaba, J. Hazard. Mater. **100**, 179 (2003).

[87] W. Petersen, L. Bertino, U. Callies, and E. Zorita, Ecol. Model. **138**, 193 (2001).

[88] J.S. dos Santos, E. de Oliveira, R.E. Bruns, and R.F. Gennari, Water Res. **38**, 1579 (2004).

[89] C. Montes-Botella and M.D. Tenorio, Arch. Environ. Con. Tox. **45**, 436 (2003).

[90] S. Yu, J. Shang, J. Zhao, and H. Guo, Water Air Soil Poll. **144**, 159 (2003).

[91] K.G. Wayland, D.T. Long, D.W. Hyndman, B.C. Pijanowski, S.M. Woodhams and S.K. Haack, J. Environ. Quality **32**, 180 (2003).

[92] J.A. Grande, J. Borrego, and J.A. Morales, Environ. Geol. **39**, 1095 (2000).

[93] C.D. Evans, T.D. Davies, P.J. Wigington Jr., M. Tranter, and W.A. Kretser, J. Hydrol. **185**, 297 (1996).

[94] S-P. Mujunen, P. Minkkinen, B. Holmbom, and A. Oikari, J. Chemom. **10**, 411 (1996).

[95] S.P. Reinikainen, P. Laine, P. Minkkinen, and P. Paatero, Fresenius J. Anal. Chem. **369**, 727 (2001).

[96] K. Loska and D. Wiechula, Chemosphere **51**, 723 (2003).

[97] J.E. Mackin, R.M. Owen, and P.A. Meyers, J. Geophys. Res. C **85**, 1563 (1980).

[98] E. Reisenhofer, A. Picciotto, and D. Li, Anal. Chim. Acta **306**, 99 (1995).

[99] P. Rachdawong, E.R. Christensen, and J.F. Karls, Water Res. **32**, 2422 (1998).

[100] P.A. Bzdusek, E.R. Christensen, A. Li, and Q. Zou, Environ. Sci. Technol. **38**, 97 (2004).

[101] A.J. Melloul, Hydrogeol. J. **3**, 19 (1995).

[102] B. Helena, R. Pardo, M. Vega, E. Barrado, J.M Fernandez, and L. Fernandez, Water Res. **34**, 807 (2000).

[103] I. Morell, E. Giménez, and M.V. Esteller, Sci. Total Environ. **177**, 161 (1996).

[104] J-L. Join, J. Coudray, and K. Longworth, J. Hydrol. **190**, 1 (1997).

[105] C. Subbarao, N.V. Subbarao, and S.N. Chandu, Environ. Geol. **28**, 175 (1996).

[106] C-W. Liu, K-H. Lin, and Y-M. Kuo, Sci. Total Environ. **313**, 77 (2003).

[107] R. Jayakumar and L. Siraz, Environ. Geol. **31**, 174 (1997).

[108] J.O. Grimalt, J. Olivé, and J.I. Gómez-Belinchón, Int. J. Environ. Anal. Chem. **38**, 305 (1990).

[109] J.O. Grimalt, L. Canton, and J. Olivé, Chemometr. Intell. Lab. Syst. **18**, 93 (1993).

[110] J.O. Grimalt and J. Olivé, Anal. Chim. Acta **278**, 159 (1993).

[111] P. Barbieri, C.A. Andersson, D.L. Massart, S. Predonzani, G. Adami, and E. Reisenhofer, Anal. Chim. Acta **398**, 227 (1999).

[112] M.M. Morales, P. Martí, A. Llopis, L. Campos, and S. Sagrado, Anal. Chim. Acta **394**, 109 (1999).

[113] C. Maes and D. Behringer, Comptes Rendus Acad. Sci. II A **329**, 597 (1999).

[114] R. Lindegren and M. Josefson, Chemometr. Intell. Lab. Syst. **44**, 403 (1998).

[115] D.A. Rothrock, D.R. Thomas, and A.S. Thorndike, J. Geophys. Res. C **93**, 2321 (1988).

[116] M. Wensnahan, G.A. Maykut, T.C. Grenfell, and D.P. Winebrenner, J. Geophys. Res. C **98**, 12453 (1993).

[117] L. Favretto, and L.G. Favretto, Z. Lebensm. Unters. For. **179**, 377 (1984).

[118] L. Favretto, L.G. Favretto, and L. Felician Z. Lebensm. Unters. For. **184**, 101 (1987).

[119] L. Favretto, L.G. Favretto, G.P. Marletta, and M. Saitta, Anal. Chim. Acta **220**, 144 (1989).

[120] S. Piepponen and R. Lindström, Chemometr. Intell. Lab. Syst. **7**, 163 (1989).

[121] L.M. Machado, M.J. Bebianno, T. Boski, and D. M. Moura, Bol. Inst. Esp. Oceanogr. **15**, 465 (1999).

[122] A. Stella, M.T. Piccardo, R. Coradeghini, A. Redaelli, S. Lanteri, C. Armanino, and F. Valerio, Anal. Chim. Acta **461**, 201 (2002).

[123] P. Barbieri, G. Adami, S. Piselli, F. Gemiti, and E. Reisenhofer, Chemometr. Intell. Lab. Syst. **62**, 89 (2002).

[124] M. Vidal, A. López, M.C. Santoalla, and V. Valles, Agr. Water Manag. **45**, 1 (2000).

[125] C. Rosen and Z. Yuan, Water Sci. Technol. **43**, 147 (2001).

[126] J. Lennox and C. Rosen, Water Sci. Technol. **45**, 227 (2002).

[127] G.N. Krasovskii and L.V. Vorob'eva, Gig. Sanit. **4**, 19 (1998).

[128] A. Critto, C. Carlon, and A. Marcomini, Environ. Pollut. **122**, 235 (2003).

[129] M.M. Sena, R.T.S. Frighetto, P.J. Valarini, H. Tokeshi, and R.J. Poppi, Soil Till. Res. **67**, 171 (2002).

[130] A. Barona and F. Romero, Soil Technol. **8**, 303 (1996).

[131] Y-P. Lin, T-P. Teng, and T-K. Chang, Landscape Urban Plan. **62**, 19 (2002).

[132] M.N. Kumru and M. Bakaç, J. Geochem. Explor. **77**, 81 (2003).

[133] I. Maiz, I. Arambarri, R. Garcia, and E. Millán, Environ. Pollut. **110**, 3 (2000).

[134] S.J. Officer, A. Kravchenko, G.A. Bollero, K.A. Sudduth, N.R. Kitchen, W.J. Wiebold, H.L. Palm, and D.G. Bullock, Plant Soil **258**, 269 (2004).

[135] E.L. Meijer and P. Buurman, Geoderma **80**, 129 (1997).

[136] S.V. Koptsik, G.N. Koptsik, S.Y. Livantsova, N.A. Berezina, and M.G. Vakhrameeva, Russian J. Ecol. **34**, 34 (2003).

[137] J.J. Brejda, Soil Sci. Soc. Am. J. **62**, 805 (1998).

[138] R.G. Garrett, J. Geochem. Explor. **32**, 319 (1989).

[139] C.Y. Chork, J. Geochem. Explor. **37**, 205 (1990).

[140] C.Y. Chork and R. Salminen, J. Geochem. Explor. **48**, 1 (1993).

[141] P. Filzmoser, Environmetrics **10**, 363 (1999).

[142] C. Reimann, P. Filzmoser, and R.G. Garrett, Appl. Geochem. **17**, 185 (2002).

[143] C. Chung and S. Nigam, J. Geophys. Res. D **104**, 16925 (1999).

[144] H.G. Pereira, S. Renca, and J. Saraiva, Appl. Geochem. **18**, 37 (2003).

[145] D.G. Cuadrado and G.M.E. Perillo, Estuar. Coastal Shelf Sci. **44**, 411 (1997).

[146] G. Wotling, Ch. Bouvier, J. Danloux, and J-M. Fritsch, J. Hydrol. **233**, 86 (2000).

[147] F.D. Maldonado, J.R. Dos Santos, and V.C. De Carvalho, Int. J. Remote Sens. **23**, 5005 (2002).