

Statistical approach for the characterization of NO₂ concentrations in Beirut

Charbel Afif · Alain L. Dutot · Corinne Jambert ·
Maher Abboud · Jocelyne Adjizian-Gérard ·
Wehbeh Farah · Pascal E. Perros · Toufic Rizk

Received: 26 June 2008 / Accepted: 18 February 2009 / Published online: 18 March 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Measurement of ambient NO₂ concentrations using diffusion tube samplers is widespread in many countries. Results from a program of measuring ground-level concentration of NO₂ at 20 urban sites within the city of Beirut are presented. In addition, two curbside sites were implemented at different heights. Two-week sampling period measurements were performed over 41 periods for all sites. A study of the sites' behaviors was conducted using principal component analysis and agglomerative hierarchical clustering. Firstly, results indicate that urban sites are equivalent, and one global class is identified. Any further monitoring of the nitrogen dioxide concentration can be conducted by decreasing the number of urban sites. Secondly, nitrogen dioxide concentration levels present a seasonal variation, as expected. Maximum average concentration of 178 μg m⁻³ was observed during

winter (December 2005) and a minimum concentration of 17 μg m⁻³ was observed during summer (May 2006). The annual average concentration measured in 2005 is 67 μg m⁻³, while the recommended value from the World Health Organization is 40 μg m⁻³.

Keywords Passive sampler · Triethanolamine · Beirut · Principal component analysis · Agglomerative hierarchical clustering · Nitrogen dioxide

Introduction

Nitrogen dioxide (NO₂) is a pollutant of the urban atmosphere. It is a primary pollutant (emitted from less than 1% to more than 30% of nitrogen oxides emitted from combustion processes) and a secondary one formed from other pollutants (i.e., ozone and nitrogen monoxide) (Finlayson-Pitts and Pitts 2000). Outdoor concentrations of NO₂ can vary widely, and rapidly, ranging from a few micrograms per cubic meter to peaks of several hundreds of micrograms per cubic meter during particular episodes of high pollution. Personal exposure to NO₂ has effects on lung function, especially on children, the elderly, and asthmatics. Consequently, there is recommendation by the World Health Organization (WHO) aiming to provide a basis for protecting public health from adverse effects of air pollution. Therefore, air quality and, especially, NO₂ is monitored in almost all countries worldwide.

C. Afif (✉) · A. L. Dutot · P. E. Perros · C. Jambert
Laboratoire Interuniversitaire des Systèmes
Atmosphériques (LISA), UMR 7583,
Université Paris VII et XII, Créteil, France
e-mail: charbel.affif@lisa.univ-paris12.fr

M. Abboud · W. Farah · T. Rizk
Faculté des sciences, Université Saint-Joseph,
Beirut, Lebanon

J. Adjizian-Gérard
Faculté des lettres, Université Saint-Joseph,
Beirut, Lebanon

Present Address:
C. Jambert
Laboratoire d'Aérodologie (LA) UMR 5560,
Université Paul Sabatier OMP, Toulouse, France

Diffusion samplers have originally been developed for monitoring pollution at work place atmospheres (Palmes et al. 1976). Being a low-cost, convenient way of mapping spatial distributions and investigating long-term trends of NO₂, these passive diffusion tubes (PDT) are widely used for ambient air monitoring as well (Campbell et al. 1994; Ferm and Svanberg 1998; Gair and Penkett 1995). Since the uptake rate is quite low, this method requires long exposure periods (1 to 4 weeks). Thus, it is not possible to measure short-term concentrations, such as hourly averages of NO₂, or peaks of pollutants, since the concentration is averaged over the entire time of exposure. In order to characterize daily cycles and short-term peaks of pollutants, one should use a method with high frequency measurements like online analyzers. On the other hand, this method is extremely useful to assess long-term concentration trends (e.g., yearly).

The PDT are devices capable of taking samples of gas or vapor pollutants from the atmosphere, without involving active air movement through them. The fixation rate is controlled by physical processes, which can be diffusion through a static air layer or permeation through a membrane (Brown et al. 1984). The average atmospheric concentration is calculated by Fick's first law, using the exposure period (Palmes and Lindenboom 1979). Several passive samplers have been developed since Palmes and Gunnison published the principles of passive sampling in 1973 (Palmes and Gunnison 1973).

Concentrations of atmospheric pollutants, especially NO₂, are extremely variable in space and time, depending on meteorological and topographical conditions and on the spatial distribution of the emission sources. The consequences of a high level of pollution will depend on the population density at the monitoring point and the nature of the ecosystems present. Consequently, air quality sampling sites are chosen relying on criteria given by specialized organizations (e.g., Environmental Protection Agency (EPA) for the USA, *Agence De l'Environnement et de la Maîtrise de l'Energie* (ADEME) for France, etc.)

In the present work, a study on air quality by diffusion samplers over Beirut is presented. It concerns the spatial and temporal variations of nitrogen dioxide concentrations over this city, and is focused on the identification of the different areas with similar variation towards nitrogen dioxide. A passive sampling network of NO₂ covering the whole city was set up in cooperation with the municipality of Beirut. This is the first spatial study in the country since a similar monitoring network for NO₂ has never been established in Lebanon.

Experimental section

Principle of passive diffusion sampling

All passive samplers operate on the principle of diffusion of gases from the atmosphere along a sampler of defined dimensions onto an adsorbing or absorbing (a solid base impregnated with a chemical reagent) medium. The gas molecules are transferred from the ambient air into the sampling device by molecular diffusion, which is a function of temperature and pressure (Delgado-Saborit and Esteve-Cano 2006). Ordinary diffusion is defined as a transfer of matter from one region to another due to a concentration gradient. For constant temperature and pressure conditions, and a laminar fluidic flux, the one-way flux (a single axis) of a molecule from milieu 1 through milieu 2 is covered by Fick's first diffusion law (Eq. 1) (Delgado-Saborit and Esteve-Cano 2006):

$$F_{1,2} = -D_{1,2} \frac{dC_{1,2}}{dx}, \quad (1)$$

where $F_{1,2}$ is the one-way flux of a molecule of interest from milieu 1 through milieu 2 (molec. cm⁻² s⁻¹), $D_{1,2}$ is the molecular diffusion coefficient of the molecule into milieu 2 (cm⁻² s⁻¹), $\frac{dC}{dx}$ is the linear concentration gradient along the diffusion path, and $C_{1,2}$ is the gaseous concentration of the compound (molec. cm⁻³).

Passive samplers (which correspond to milieu 2) have a null concentration regarding the monitored pollutant. Thus, the flux is directly proportional to gaseous concentration in milieu 1 (air), with a constant sampling rate, until milieu 2 is saturated.

Types of passive diffusion samplers

Several samplers have been developed since the original "Palmes-tube" (Palmes et al. 1976), in particular, tube-type and badge-type samplers. The length L and the cross-sectional area A of the sampler control the theoretical uptake rate. The tube-type samplers are usually hollow and cylindrical. Examples of sampling media with passive sampler configurations used for field applications for NO₂ are summarized in Table 1.

The most commonly used sink for NO₂ is triethanolamine (TEA), where the gas is converted into nitrite ions. The reaction-product of TEA has been subjected to several investigations (Aoyama and Yashiro 1983; Gold 1977; Haue-Pedersen et al. 1994; Levaggi et al. 1972; Palmes et al. 1976). Results lead to the identification of TEA N-oxide as a reaction product

Table 1 Some passive samplers used for NO₂ measurement

Sampler type	Dimensions $L(\text{cm}) \times A(\text{cm}^2)$	Inlet	Collection medium	NO ₂ uptake ($\text{m}^3 \text{h}^{-1}$)	Comment	Reference
Palmer-type diffusion tube	7.1×0.95	Open	TEA	$7.17 \times 10^{-5 \text{ a}}$	Protective cover recommended for outdoor sampling	Palmer et al. (1976)
Ferm badge	1.0×3.14	Steel grid + Teflon membrane	KI/NaAsO ₂ NaI/NaOH	$1.74 \times 10^{-3 \text{ a}}$		Ferm (1991)
CSPSS badge	8.0×17.35	PTFE membrane	CHEMIXTM	$1.2 \times 10^{-2 \text{ a}}$	Protective cover for outdoor sampling	Tang et al. (2001)
Krochmal badge	1.0×4.91	Polypropylene membrane	TEA	$2.72 \times 10^{-3 \text{ a}}$		Krochmal and Kalina (1997)
Radiello tube	6.0×0.72	Cylindrical synthesized micro-porous polyethylene	TEA	4.68×10^{-4}	Protective cover for outdoor sampling	Radiello (2006)

^a Rates calculated by Tang et al. (2001)

with the 1:1 conversion of NO₂ to nitrite ions (Glasius et al. 1999; Palmes and Johnson 1987).

Effects of different parameters on nitrogen dioxide sampling

Effects of meteorological factors Different studies were carried out to evaluate the influence of meteorological factors [temperature, relative humidity (RH), and wind velocity] on the sampling rate of NO₂. The greatest effect is observed for the wind velocity with a significant increase of sampling rate (Bush et al. 2001; Campbell et al. 1994; Plaisance et al. 2004). This effect, expressed in terms of reduction in diffusion length, is 47% for air velocities close to 2 m s^{-1} (Plaisance et al. 2004) and can be limited by a physical protection of the tube. Temperature and RH induce weak deviations on the sampling rate in comparison with the one generated by wind velocity. However, Moschandreas et al. (1990) reported that low temperatures in the range of -22°C to 10°C resulted in the underestimation of nitrogen dioxide concentration because of possible anomalous behavior of the TEA below its freezing point of 21°C . In addition, correction functions should be applied under particular conditions ($T > 30^\circ\text{C}$ and $\text{RH} > 80\%$) (Plaisance et al. 2004). Radiello proposes correction of the sampling rate even for ambient temperature (Radiello 2006).

Effects of fluctuations of NO₂ concentration in ambient air Plaisance (2004) conducted a study to estimate the errors of the Palmer tube induced by concentration variations in ambient air. Results indicate that there is a slight influence or no effect of concentration variations on the sample measurement by Palmer tube. Although fast fluctuations of high amplitude can induce a temporal overestimation of the tube, its contribution to the estimation of mean concentration remains weak over the 14-day sampling time. The errors induced by unsteady ambient concentration appear to be of minor importance (1.6% for an overall 14-day sampling time) in comparison with those caused by the meteorological factors.

Effects of within-tube chemical reaction and NO₂ photolysis The effect of trace species such as HONO, peroxyacetyl nitrate (PAN), O₃, and other nitrogen compounds has been investigated by Atkins et al. (1986). They concluded through a campaign conducted in Great Britain that no interference is caused by the presence of PAN or nitrous acid (HONO), although they can both provide nitrite, in a complex with TEA.

A similar conclusion was reached by Gair et al. (1991). However, Hisham and Grosjean (1990) noted that PAN is a “severe interferent” to quantification of NO₂ by a TEA diffusion tube. The latter study was conducted in California, where ambient PAN levels are considerably higher than those in the UK (Heal and Cape 1997). Hedley et al. (1994) also reported a significant overestimation (up to twofold) of NO_z compounds by passive diffusion samplers in their comparative trials. They attributed it to chemical interferences from PAN, other nitrogen containing species, and oxidation of NO by O₃ at the adsorbent.

In an urban environment close to emission sources, NO is the predominant specie in NO_x. Due to chemical reaction with O₃, NO is measured as NO₂ and can give rise to errors of up to 30% for curbside locations (Heal et al. 1999) and greater when the NO₂/NO_x ratio reaches 0.5 at the monitoring site located relatively far from the source (Heal et al. 2000). In rural locations, NO is present in small concentrations relative to NO₂ and considerably lower than O₃. The diffusion tubes will measure total NO_x effectively with small errors upon NO₂ concentration (Heal et al. 2000).

On the other hand, NO₂ is photolyzed during the day with a photolysis frequency (JNO_2) that varies greatly throughout a PDT exposure time. In fact, modeling shows that generation of net excess NO₂ is not particularly sensitive to the condition of zero NO₂ photolysis within the tube (Heal and Cape 1997). The average residence time of NO₂ for diffusion along the tube is 2.8 min (Cao and Hewitt 1991). Chemical reaction between ambient concentrations of NO and O₃ to generate NO₂ occurs on a similar timescale, whereas photolytic lifetime of NO₂ exceeds 17 min for $JNO_2 = 1.10^{-3} \text{ s}^{-1}$ (Heal and Cape 1997). This mismatch in timescales leads to net generation of NO₂, and consequently, NO₂ might be overestimated.

Effects of exposure time When using TEA as adsorbent for NO₂ passive sampling, the nitrogen dioxide is converted to nitrite after reaction with TEA. There is, however, a lack of specificity of TEA towards NO₂ since it is also sensitive to sulfur dioxide (SO₂). In an environment with high concentrations of SO₂, TEA is acidified, and consequently, the sampling efficiency is reduced. There are various reports that NO₂ diffusion tubes coated with TEA and exposed for 1 month give lower NO₂ concentrations than averaged NO₂ concentrations from successive samples exposed for 1- or 2-week periods (Bush et al. 2001; Ferm and Svanberg 1998; Heal and Cape 1997). A 2-week exposure is, thus, a compromise between sufficient sampling and low SO₂ interaction.

Measuring campaign

Area of study Beirut, a 32-km² city, is located at 33°52'10" N 35°30'40" E, on the eastern coast of the Mediterranean Sea. The north and west sides of the city are opened to the sea while the east side is surrounded by Mount Lebanon. The city population density is about 20,167 inhabitants per square kilometer. The main source of NO₂ is high traffic density, the industrial activity being little developed in the surroundings of the city. The number of registered vehicles in Lebanon is greater than one million, with an average age for the vehicle fleet of more than 10 years (El-Zein et al. 2007). The number of cars circulating in Beirut is not known, although a study was launched by the government but no results are published.

Description of passive sampler employed Measurements of NO₂ were made with Radiello passive samplers (provided by Fondazione Salvatore Maugeri). The Radiello sampler is made of a cylindrical collection cartridge. It is housed coaxially inside a cylindrical opaque diffusive body, which minimizes the sensitivity of the system to wind speed, turbulence, and sunlight. These components are screwed onto a supporting plate.

The Radiello device is a passive sampler based on a radial diffusion. Consequently, there is a greater diffusion area than for the axial ones, which therefore facilitates the reaction between NO₂ and the collection cartridge.

The Radiello passive sampler for measuring nitrogen dioxide is based on the principle that NO₂ diffuses across the diffusive body toward the adsorbed TEA on the inner cartridge. The NO₂ is chemisorbed in the TEA as nitrite, and is then extracted. Special care was taken at all times when handling the passive samplers. Except during sampling, all samplers were kept in an airtight bag. Prepared samplers were transported to and from the field on ice, in airtight containers. After exposure, the sampler's cartridges were transferred to plastic tubes in the refrigerator until preparation for analysis. The uncertainty reported by Radiello regarding these tubes is 11.9% at 2σ (Radiello 2006).

Technical analysis The derivatization proceeds through reaction with sulfanilamide (SA)/N-(1-naphthyl)ethylenediamine dihydrochloride (NED) in the laboratory. The nitrite ion reacts with SA and forms a diazonium compound. The coupling of the diazonium compound with NED forms a purple azodye, which is measured with a Vis-spectrophotometer (Delgado-Saborit and Esteve-Cano 2006; Radiello 2006).

For extraction, 5 ml of ultrapure water is poured into the plastic tube containing the cartridge. The solution is stirred vigorously for 5 min to allow the nitrite to dissolve in the water. After stirring, the cartridges are removed from the solution. Sampler extracts are analyzed immediately. For derivatization, 0.5 ml of the nitrite solution is mixed with 5 ml of SA $58.10^{-3} \text{ mol L}^{-1}$ and stirred for 5 min. One milliliter of NED $39.10^{-4} \text{ mol L}^{-1}$ is then added and the whole is stirred again. The NO_2 captured is then determined colorimetrically as nitrite by means of a Nicolet 300 spectrophotometer at 537 nm. The measurement precision within the laboratory analysis and the temperature variation is less than 5% (Radiello 2006).

Periods and sampling sites Since concentrations of atmospheric pollutants are extremely variable in space and time, measuring the quality of ambient air should be performed under normalized conditions. Numerous air-quality monitoring programmes in rural areas occurred in the 1970s, in particular when scientists discovered the problem of long-distance pollution and its harmful effects on the different ecosystems [e.g., the European Monitoring and Evaluation Program, the international programme run by the World Meteorological Organization (WMO): WMO—Global Atmosphere Watch)].

Therefore, the implementation of our sampling sites relied on the criteria of the ADEME (2002). Two types of sampling stations were selected: traffic sites and urban sites. Two traffic sampling sites (S1 and S2) were selected to verify that urban levels of NO_2 are always lower than the traffic ones and no occasional source existed near the sampler. The general objective of traffic sites is to supply data on the concentrations measured in areas that are representative of the maximum level of exposure to which the population located near a road network is likely to be exposed. The site should be under the direct influence of the linear source without any obstacles. It is recommended to avoid obstacles such as hedgerows or walls that might disturb measurements. This type of site should be located either near a roadway with an Annual Daily Traffic Average (ADTA) (in both directions) greater than 10,000 vehicles per day or a canyon-type roadway with a risk of pollution accumulation. The sampler must be located at most 5 m from the road side. The recommended height is between 1.5 and 3 m above the ground (Agence De l'Environnement et de la Maîtrise de l'Energie 2002; Vardoulakis et al. 2003). In this study, S1 is implemented at 3 m, but S2 at 5.5 m above the ground for logistic reasons.

For the urban sites, 20 locations (S3–S22) were selected to monitor the population's average exposure to the phenomena of atmospheric pollution in urban centers. Minimal distance from the site to the nearest road depends on the ADTA and varies from 10 m (for ADTA less than 1,000 vehicles/day) to over 200 m for ADTA greater than 70,000 vehicles/day. Samplers were located at about 3 m above the ground for these sites.

The measurement campaign was carried out from December 2004 till June 2006 over the 22 sites. Figure 1 shows the location of the experimental sites within Beirut. Two-week sampling periods were adopted for both types of site. The results presented in this paper were thus obtained from 41 periods of measurements resulting in 902 determinations of NO_2 concentrations.

Passive samplers were, at first, designed to assess the quality of indoor air. In assessing the quality of air outdoors, it is advisable to protect them with a rain/wind shelter. Even in the absence of rain, the shelter is required to minimize dust contamination and to limit the effects of advection on the diffusive samplers (Roadman et al. 2003). For this study, shelters supplied by the manufacturer were fixed by means of a cable tie. In addition to that, SO_2 was measured by means of badges with 4-week sampling periods. Sulfur dioxide concentrations varied within the range of 4 to $20 \mu\text{g m}^{-3}$ all along the campaign. Since the sampling period for the NO_2 is 2 weeks, the acidification caused by the sulfur dioxide might be limited. Nevertheless, no study was found to propose a correction factor for the reduction of sampling efficiency caused by SO_2 . Temperature corrective factor was applied to the sampling rate based on in situ temperature measurements. Since the diffusive body is opaque, the NO_2 photolysis



Fig. 1 Map of the different sites in Beirut. S1 and S2 are curbside sites. S3 to S22 are urban sites

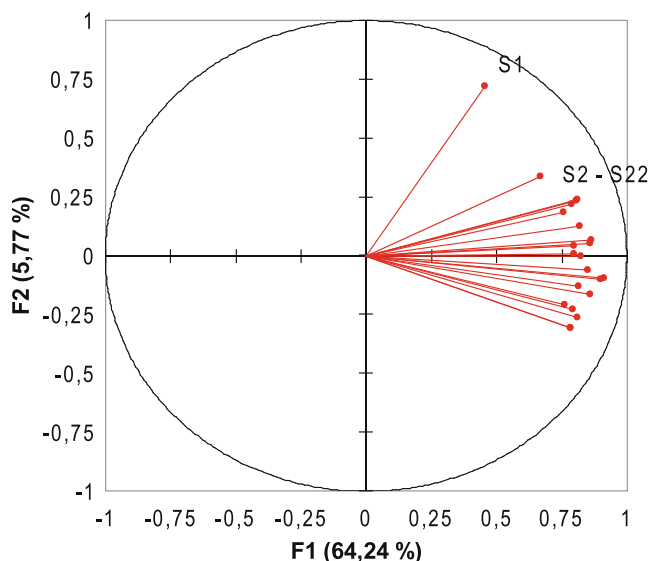


Fig. 2 Correlation circle of the PCA analysis of the data classified by site (Si). F1 and F2 represent 70% of the total variability of the results

is eliminated. Unfortunately, some influencing factors on nitrogen dioxide sampling (like titration reaction of NO by O₃ or RH) stand and can be neither corrected nor controlled.

Method of data analysis

A statistical approach is used to control and validate the implementation of measurement sites. The statistical approach consisted of data analysis by principal component analysis (PCA) and agglomerative hierarchical clustering (AHC). Statistical methods such as PCA and AHC have been used in many studies aimed at the management of water monitoring networks (Kannel et al. 2007; Mendiguchia et al. 2004; Shrestha and Kazama 2007; Singh et al. 2004, 2005) and air quality management (Gramsch et al. 2006; Pires et al. 2008a,b). The study of Gramsch et al. (2006) determined the seasonal trends and spatial distribution of PM₁₀ and O₃ for Santiago de Chile, concluding that the city had four large sectors with dissimilar air pollution variations, whereas the study conducted by Pires et al. (2008a) classified SO₂ sites in six groups and PM₁₀ sites in two groups of dissimilar variation in Oporto Metropolitan Area.

Principal component analysis

Exploratory PCA is a well known data display method in multivariate analysis (Lavine 2000); therefore, only

a summary of the method is described below. The PCA is a widely used technique in meteorology and climatology. It reduces the size of large datasets while minimizing any loss of information, with the aim of better understanding and interpreting the structure of the data.

A typical dataset constituting the original matrix can be viewed as n observations measured for N variables. To simplify the interpretation process, one should select m variables ($m < N$), which express all information contained in the matrix. This can be done by creating new variables that are different from the original ones but constructed from combinations of them. The simplicity of the PCA technique lies in this restriction to linear functions of the original variables.

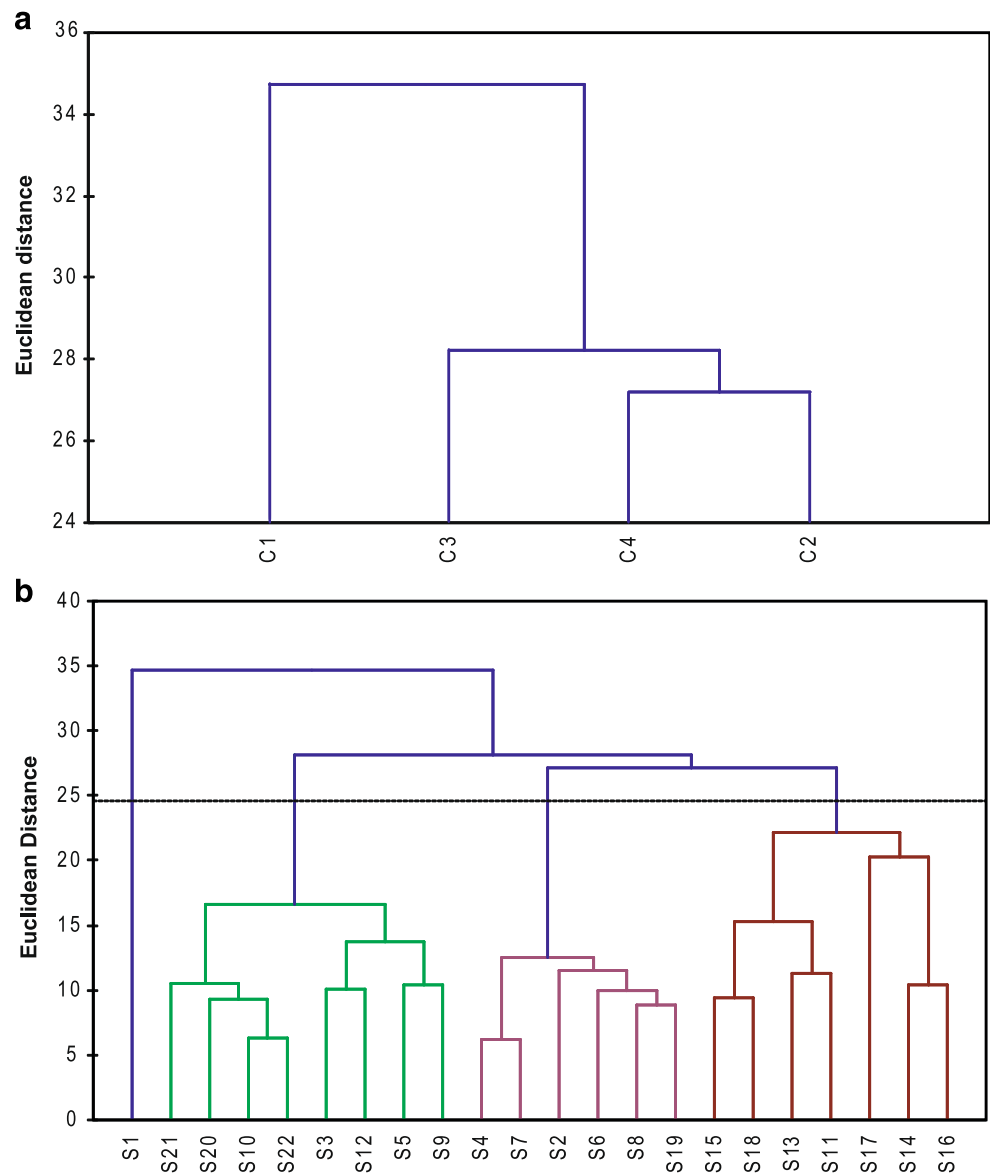
Linear functions of the N variables can thus be obtained. The number of these linear functions is equal to N . The variance of each linear function can be calculated. The first principal component (PC) is the linear function that explains the maximum possible variance. The second PC is a linear function with a maximum possible variance but which is not correlated with the first PC, and so on. The objective of PCA is to find a small number, m , of linear functions of a set of N variables that successively accounts for the maximum amount of variation in the original variables. The eigenvector is the set of coefficients contained in each PC and to which corresponds a score. The eigenvalue corresponds to the variance of each PC and is therefore a measure of its importance in explained variation.

One can compute eigenvalues and eigenvectors of either a covariance, a correlation, or a cross-product matrix of the observations recorded for N variables. If the N initial variables correspond to different sampling sites, this may correspond to a spatial analysis because the eigenvectors have a component for each site and represent an orthogonal spatial pattern (Molteni et al. 1983). PCAs can also be used in other contexts, like a temporal one (Richman 1986).

Commonly, scores are displayed in a two-dimensional PC space facilitating an analysis of the data structure, and this can be assisted by PC coefficients plots. These single plots provide guidance for the recognition of important variables on a PC. Biplots consisting of an overlay of the criteria variables, represented as vectors, over the PC scores plot provide an effective method for studying the object–object, variable–variable, and object–variable relationships.

Generally, if the first factor takes into account a large part of the information, it describes something obvious; the other factors can, however, bring to light interesting phenomena hidden in the data matrix. If we are more interested in the joint variability of the PCs

Fig. 3 Dendrograms presenting the results of the AHC analysis for the data classified by site (Si). **a** The four classes: C1 to C4. **b** The different sites within each class. The *dotted line* represents the truncation level



than in their individual variability, we can rotate them within their subspace and replace them with m -derived variables (rotated components). Wilks (2006) presents a summary of the multitude of techniques that can be used for rotations in a PCA applied to atmospheric sciences. In the analysis presented below, the rotation method was not applied.

Cluster analysis

AHC analysis is a well-established methodology to discriminate different climate regimes (Fovell and Fovell 1993). The aim of hierarchical clustering is to identify

subsets of objects, i.e., clusters, having similar characteristics from within the whole sample (Wilks 2006).

The hierarchical approach starts with all objects being independent elements. The two most similar elements are then combined into one new element. The similarity between two elements can be defined as the distance between these two elements. The “Euclidean distance” method is probably the most commonly chosen type of distance. It corresponds to the geometric distance in a multidimensional space. The distances between the new formed element and the remaining elements are then calculated. The two closest elements are again searched for and grouped; the new distances are calculated. This combination step reduces the num-

ber of elements by one, and it is repeated until only one element remains.

The similarity between clusters consisting of more than one object is often referred to as the cluster method. In accordance with Kalkstein et al. (1987), Ward's cluster method (Ward 1963) was applied because this method is more accurate than others (Lu et al. 2006; Mangiameli et al. 1996) and produces stable and compact clusters (Flemming et al. 2005). Ward's method defines the distance between two clusters as the sum of squared distances.

The cluster merging procedure can be shown graphically as a dendrogram, displaying the hierarchical nesting of the clusters and their distances for merging. The PCA and AHC in this study were carried out by using XLSTAT 2007 software from AddinSoft.

Results and discussion

PCA and AHC described earlier were used (1) to identify sites of similar variation towards NO_2 and (2) to analyze the temporal variability of NO_2 concentration over the city. The biweekly measurements for 18 months at 22 different sampling sites lead to $n = 41$ observations (averaged concentrations) for $N = 22$ variables (sampling sites). The inertia matrix is based on a cross-product matrix of the PCA coefficients. AHC describes the nearness between objects (i.e., site concentrations). A dissimilarity matrix was built using the autoscaled data. The elements of this matrix are the Euclidean distances of one object from the rest. To obtain clusters, the Ward method was used. In each step, this agglomerative method considers the heterogeneity or deviance (the sum of the square of the distance of an object from the centroid of the cluster) of every possible cluster that can be created by linking two existing clusters.

Spatial variability of the data

PCA was used to analyze correlations between the variables of the sites using observations. Figure 2 represents the correlation circle, showing a projection of the initial variables in the factor space. The PCs 1 and 2 are plotted as axes F1 and F2. The calculated eigenvalues reflect the quality of the projection from the N -dimensional initial table to a lower number of dimensions. They show that if we represent the data with only the axis F1, we are still able to see 64% of the total variability of the data. The variables S1 to S22 are all far from the center of the circle. Consequently, if the variables are close to each other, they

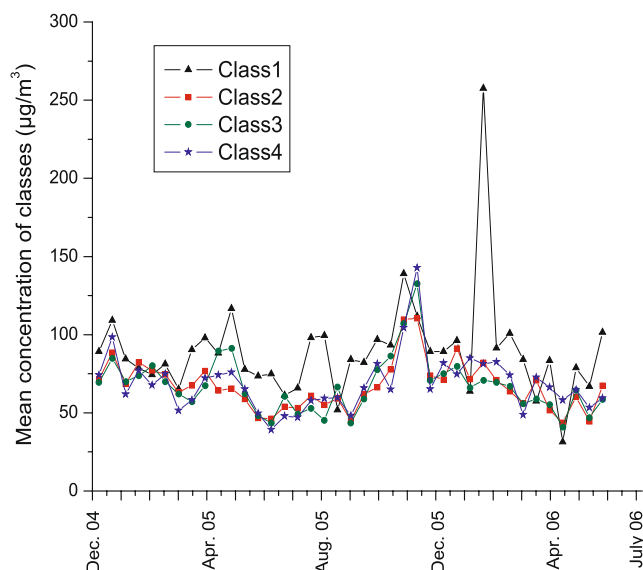


Fig. 4 Temporal variation of centroids for the four classes

are significantly positively correlated (r close to 1); if they are orthogonal (considering each point with the center of the circle), they are not correlated (r close to 0); if they are on the opposite side of the center, then they are significantly negatively correlated (r close to -1). In this study, S2 to S22 are very close to each other, but S1 is relatively far. For the other dimensions of the space (axes F3 to F22), the sum of the squared cosines for each site is much lower than the value of F1 alone (indicating their slight influence). The values determined for F1 are relatively close for the S2 to S22 data and vary within the range of 0.44 to 0.74, while it is about 0.21 for S1. In addition, F2 values are lower than 0.093 for S2 to S22 and about 0.52 for S1. Since S1 is a curbside site and most of the others (S3–S22) are urban sites, we can link the horizontal axis (F1) to urban variability of the pollutant concentration and the vertical one (F2) to curbside sites with much more local influence. Hence, this analysis underlines the discrepancy between curbside sites and urban sites.

In order to confirm this hypothesis and to attempt classifying the different sites, the AHC method was performed. Figure 3a, b presents the corresponding dendrograms. It represents how the algorithm works to group the variables, and then the subgroups of variables. At Euclidean distance of 35, two groups are distinguished: the first one contains class 1 (C1), constituted only of S1. The second group splits into three classes (C2, C3, and C4) for Euclidean distances of 28 and 27. The dotted line represents the truncation, leading to four interhomogeneous classes (C1, C2, C3, and C4). C2 is more homogeneous than C3, which is

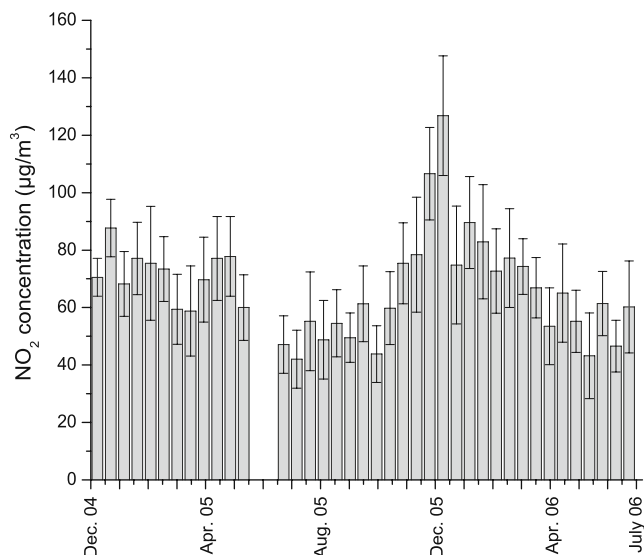


Fig. 5 Temporal variation of mean NO₂ concentration for the 41 periods of observation in urban sites from 2 December 2004 to 29 June 2006 in Beirut. Vertical bars: standard deviation for each period

more homogeneous than C4. This higher homogeneity is indicated by the lower value of within-class variance (7,392 for C2, 7,858 for C3, and 8,255 for C4). An interesting case is that of S2, the second curbside site, which is classified within class C4. This classification is due to the different height of the sampler on this site (5.5 m from the ground compared to 3 m for S1). Since S2 is higher in height, NO₂ is more dispersed at 5.5 m than it is at 3 m of height. Consequently, the measured concentrations at 5.5 m are less and tend to be compared in our case to concentrations of urban sites. This result is in accordance with the recommendation of Agence De l'Environnement et de la Maîtrise de l'Energie (2002) and Vardoulakis et al. (2003). Therefore, traffic sites should be implemented between 1.5 and 3 m of height, otherwise they will not give the desired quality of measurements.

Since no difference could be highlighted within the second group (C2, C3, and C4), we studied the temporal variation of the mean concentration of the four classes trying to find the difference between them. Figure 4

presents the temporal variation of the mean concentration of the four classes. Class 1 presents not only a threefold peak in early 2006, but also relatively higher concentrations of NO₂ all over the period. This can be explained by the nearness of the sampling to the source since S1 is a curbside site. Classes 2, 3, and 4 show weak differences in their temporal variation and a lot of intersections through time.

These results lead to the conclusion that classes C2, C3, and C4 are quite equivalent and can be considered as one global class. Consequently, we suggest minimizing the number of urban samplers for future monitoring of urban NO₂ in Beirut.

Temporal characterization of the NO₂ pollution

The temporal variation (mean and standard deviation) for all of the 20 urban sites is presented in Fig. 5. NO₂ concentrations during the whole period of observation ranged from a minimum value of 17 µg m⁻³ in May 2006 at the site S10 to a maximum value of 178 µg m⁻³ in December 2005 at the site S6, with an urban background average of 66 µg m⁻³. The annual average concentration for the 20 urban sites for 2005 is 67 µg m⁻³, while the level recommended by the WHO is 40 µg m⁻³. The standard deviation of these measurements ranges from 6.6 to 20.8 µg m⁻³. Local authorities should make an effort to improve the air quality of the city.

The period of measurement between 2 December 2004 and 29 June 2006 was divided into groups corresponding to the four seasons. Table 2 shows the statistical analyses of the data by season for these sites over the whole period of measurements. In general, the pollutant concentration is higher during winter than the one during summer. This can be linked to meteorological conditions initiating the formation of a lower inversion layer during winter than in summer in addition to the usage of central-heating burners in winter (Saliba et al. 2006), and to the lower photochemical activity, leading to lower photolysis of NO₂. In summer, the sink of NO₂ due to its reaction with OH to produce HNO₃ is more efficient than in winter, where OH concentrations are lower.

Table 2 Statistical analyses of the data classified by season

Season	Mean	S.D.	Min	Max
Winter 2005	71.7	9.3	34.6	110.3
Spring 2005	71.1	9	35.5	102.7
Summer 2005	51.1	7.9	17.8	96.1
Autumn 2005	82.4	8.3	24.5	177.9
Winter 2006	75.4	10.4	43.7	130.3
Spring 2006	54.9	7.9	16.5	90.9

Conclusion

A nitrogen dioxide network is established for the first time in Lebanon, since no air quality network was present in the country or in Beirut. The main NO_x emission source is traffic. Concentrations of nitrogen dioxide are greater than WHO guidelines, indicating that an effort has to be made to ameliorate air quality in Beirut. Aimed at the identification of city areas with similar pollution variations, two statistical methods, PCA and AHC, were applied. Main results show that the 20 urban monitoring sites of the established network can be classified into one quite homogeneous group, presenting lower concentrations in summer than the ones observed in winter. To complete this, a detailed temporal variation would be of great interest. This could be achieved with the use of online analyzers. This possibility will characterize short-term peaks of nitrogen dioxide, which cannot be done with passive samplers. Furthermore, curbside site S1 showed a different behavior from urban sites. However, it is recommended to implement curbside sites at appropriate height (between 1.5 and 3 m) since dispersion plays an important role in controlling concentration (i.e., curbside site S2). The results and data presented in this study could play an important role in the development of a strategy to control the pollution in the city and, consequently, the region.

Acknowledgements Funding for this study was obtained from the University Research Board at Saint Joseph University, CEDRE (*Coopération pour l'Évaluation et le Développement de la Recherche*) program and the Municipality of Beirut. We would like to thank AIRPARIF (France), Ile-de-France regional council, and the department of geography at Saint Joseph University.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Agence De l'Environnement et de la Maîtrise de l'Énergie (2002). Classification and criteria for setting up air-quality monitoring stations. Agence De l'Environnement et de la Maîtrise de l'Énergie, Paris, pp 64
- Aoyama T, Yashiro T (1983) Analytical study of low-concentration gases. Investigation of the reaction by trapping nitrogen dioxide in air using the triethanolamine method. *J Chromatogr* 265:69–78
- Atkins CHF, Sandalls J, Law DV, Hough AM, Stevenson K (1986) The measurement of nitrogen dioxide in the outdoor environment using passive diffusion tube samplers. AERE-R12133, Harwell Laboratory
- Brown RH, Harvey RP, Prunell CJ, Saunders KJ (1984) A diffusive sampler evaluation protocol. *Am Ind Hyg Assoc J* 45:67–75
- Bush T, Smith S, Stevenson K, Moorcroft S (2001) Validation of nitrogen dioxide diffusion tube methodology in the UK. *Atmos Environ* 35:289–296
- Campbell GW, Stedman JR, Stevenson K (1994) A survey of nitrogen dioxide concentrations in the United Kingdom using diffusion tubes, July–December 1991. *Atmos Environ* 28:477–486
- Cao X-L, Hewitt CN (1991) Application of passive samplers to the monitoring of low concentration organic vapours in indoor and ambient air: a review. *Environ Technol* 12:1056–1062
- Delgado-Saborit JM, Esteve-Cano V (2006) Field study of diffusion collection rate coefficients of a NO₂ passive sampler in a mediterranean coastal area. *Environ Monit Assess* 120:327–345
- El-Zein A, Nuwayhid I, El-Fadel M, Mroueh S (2007) Did a ban on diesel-fuel reduce emergency respiratory admissions for children? *Sci Total Environ* 384:134–140
- Ferm M (1991) A sensitive diffusional sampler. IVL Report L91-172. Swedish Environmental Research Institute, Göteborg
- Ferm M, Svanberg PA (1998) Cost-efficient techniques for urban and background measurements of SO₂ and NO₂. *Atmos Environ* 32:1377–1381
- Ferm M, Rodhe H (1997) Measurements of air concentrations of SO₂, NO₂ and NH₃ at rural and remote sites in Asia. *J Atmos Chem* 27:17–29
- Flemming J, Stern R, Yamartino RJ (2005) A new air quality regime classification scheme for O₃, NO₂, SO₂ and PM₁₀ observations sites. *Atmos Environ* 39:6121–6129
- Fovell RG, Fovell MC (1993) Climate zones of the conterminous United States defined using cluster analysis. *J Clim* 6: 2103–2135
- Gair AJ, Penkett SA, Oyola P (1991) Development of a simple passive technique for the determination of nitrogen dioxide in remote continental locations. *Atmos Environ* 25A:1927–1939
- Gair AJ, Penkett SA (1995) The effects of wind speed and turbulence on the performance of diffusion tube samplers. *Atmos Environ* 29:2529–2533
- Glasius M, Carlsen FM, Hansen TS, Lohse C (1999) Measurements of nitrogen dioxide on Funen using diffusion tubes. *Atmos Environ* 33:1177–1185
- Gold A (1977) Stoichiometry of nitrogen dioxide determination in triethanolamine trapping solution. *Anal Chem* 49:1448–1450
- Gramsch E, Cereceda-Balic F, Oyola P, Baer D (2006) Examination of pollution trends in Santiago de Chile with cluster analysis of PM₁₀ and ozone data. *Atmos Environ* 40:5464–5475
- Haue-Pedersen U, Hanse CL, Knudsen KU, Skov H, Lohse C (1994) Measurements of NO₂ by passive sampling. In: Angeletti G, Restelli G (eds) Physico-chemical behavior of atmospheric pollutants, proceedings of the 6th European symposium, Varese 1993. Report EUR 15609/1. European Commission, Luxembourg
- Heal MR, Cape JN (1997) A numerical evaluation of chemical interferences in the measurement of nitrogen dioxide by passive diffusion samplers. *Atmos Environ* 31:1911–1923
- Heal MR, O'Donoghue MA, Cape JN (1999) Overestimation of urban nitrogen dioxide by passive diffusion tubes: a comparative exposure and model study. *Atmos Environ* 33:513–524

- Heal MR, Kirby C, Cape JN (2000) Systematic biases in measurement of urban nitrogen dioxide using passive diffusion samplers. *Environ Monit Assess* 62:39–54
- Hedley KJ, Shepson PB, Barrie LA, Bottenheim JW, Mactavish DC, Anlauf KG, Mackay GI (1994) An evaluation of integrating techniques for measuring atmospheric nitrogen dioxide. *Int J Environ Anal Chem* 54:167–181
- Hisham MWM, Grosjean D (1990) Sampling of atmospheric nitrogen dioxide using triethanolamine-interference from peroxyacetyl nitrate (PAN). *Atmos Environ* 24:2523–2525
- Kalkstein LS, Tan G, Skindlov JA (1987) An evaluation of three clustering procedures for use in synoptic climatological classification. *J Clim Appl Meteorol* 26:717–730
- Kannel PR, Lee S, Kanel SR, Khan SP (2007) Chemometric application in classification and assessment of monitoring locations of an urban river system. *Anal Chim Acta* 582:390–399
- Krochmal D, Kalina A (1997) A method of nitrogen dioxide and sulfur dioxide determination in ambient air by use of passive samplers and ion chromatography. *Atmos Environ* 31:3473–3479
- Lavine BK (2000) Clustering and classification of analytical data. *Encyclopedia of analytical chemistry: instrumentation and applications*. Wiley, New York, pp 9689–9710
- Levaggi DA, Siu W, Feldstein M, Kothny EL (1972) Quantitative separation of nitric oxide from nitrogen dioxide at atmospheric concentration ranges. *Environ Sci Technol* 6:250–252
- Lu HC, Changb CL, Hsieha JC (2006) Classification of PM10 distributions in Taiwan. *Atmos Environ* 42:1452–1463
- Mangiameli P, Chen SK, West D (1996) A comparison of SOM of neural network and hierarchical methods. *Eur J Oper Res* 93:402–417
- Mendiguchia C, Moreno C, Galindo-Riano MD, Garcia-Vargas M (2004) Using chemometric tools to assess anthropogenic effects in river water: a case study: Guadalquivir River (Spain). *Anal Chim Acta* 515:143–149
- Molteni E, Bonelli E, Bacci E (1983) Precipitation over Northern Italy: a description by means of principal component analysis. *J Clim Appl Meteor* 22:1738–1752
- Moschandreas DJ, Relwani SM, Taylor KC, Mulik JD (1990) A laboratory evaluation of a nitrogen dioxide personal sampling device. *Atmos Environ* 24A:2807–2811
- Palmes ED, Gunnison AF (1973) Personal monitoring device for gaseous contaminants. *Am Ind Hyg Assoc J* 34:78–81
- Palmes ED, Gunnison AF, DiMattio J, Tomaczyk C (1976) Personal sampler for nitrogen dioxide. *Am Ind Hyg Assoc J* 37:570–577
- Palmes ED, Lindenboom RH (1979) Ohm's law, Fick's law, and diffusion samplers for gases. *Anal Chem* 51:2400–2401
- Palmes ED, Johnson ER (1987) Explanation of pressure effects on a nitrogen dioxide (NO₂) sampler. *Am Ind Hyg Assoc J* 48:73–76
- Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MCM, Martins FG (2008a). Management of air quality monitoring using principal component and cluster analysis-Part I: SO₂ and PM10. *Atmos Environ* 42:1249–1260
- Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MCM, Martins FG (2008b) Management of air quality monitoring using principal component and cluster analysis-Part II: CO, NO₂ and O₃. *Atmos Environ* 42:1261–1274
- Finlayson-Pitts BJ, Pitts JN Jr (2000) *Chemistry of the upper and lower atmosphere*. Academic, London
- Plaisance H (2004) Response of a Palmes tube at various fluctuations of concentration in ambient air. *Atmos Environ* 38:6115–6120
- Plaisance H, Piechocki-Minguy A, Garcia-Fouque S, Galloo JC (2004) Influence of meteorological factors on the NO₂ measurements by passive diffusion tube. *Atmos Environ* 38: 573–580
- Radiello M (2006) Fondazione Salvatore Maugeri - IRCCS. http://www.radiello.it/english/download_en.htm
- Richman MB (1986) Rotation of principal components (review article). *J Climatol* 6:293–335
- Roadman MJ, Scudlark JR, Meisinger JJ, Ullman WJ (2003) Validation of ogawa passive samplers for the determination of gaseous ammonia concentrations in agricultural settings. *Atmos Environ* 37:2317–2325
- Saliba NA, Moussa S, Salame H, El-Fadel M (2006) Variation of selected air quality indicators over the city of Beirut, Lebanon: assessment of emission sources. *Atmos Environ* 40:3263–3268
- Shrestha S, Kazama F (2007) Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environ Model Softw* 22: 464–475
- Singh KP, Malik A, Mohan D, Sinha S (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Res* 38:3980–3992
- Singh KP, Malik A, Sinha S (2005) Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. *Anal Chim Acta* 538:355–374
- Tang H, Lau T, Brassard B, Cool W (1999) A new all-season passive sampling system for monitoring NO₂ in air. *Field Anal Chem Technol* 3:338–345
- Tang YS, Cape JN, Sutton MA (2001) Development and types of passive samplers for monitoring atmospheric NO₂ and NH₃ concentrations. In: *Proceedings of the international symposium on passive sampling of gaseous air pollutants in ecological effects research*. *Sci World* 1:513–529
- Vardoulakis S, Fischer BEA, Pericleous K, Gonzalez-Flesca N (2003) Modelling air quality in street canyons: a review. *Atmos Environ* 37:155–182
- Ward JH (1963) Hierarchical grouping to optimize and objective function. *J Am Stat Assoc* 58:236–244
- Wilks DS (2006) *Statistical methods in the atmospheric sciences*, 2nd edn. Academic, London