Aerosol and Air Quality Research, 14: 1455–1465, 2014

ISSN: 1680-8584 print / 2071-1409 online
doi: 10.4209/aaqr.2013.05.0160

Cluster Analysis for Daily Patterns of SO_2 and NO_2 Measured by the DOAS System in Xiamen

Peng Shi[1], Pin-Hua Xie[2], Min Qin[2], Fu-Qi Si[2], Ke Dou[2], Ke Du[1,3*]

[1] *Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China*
[2] *Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei 230031, China*
[3] *Department of Mechanical and Manufacturing Engineering, University of Calgary, Calgary, AB, Canada*

ABSTRACT

Daily patterns of air pollutants are important to improve measurement retrievals and to model the regimes of local air quality. In this study, the daily patterns of SO_2 and NO_2 as well as their association with visibility and meteorological conditions in a suburban area of Xiamen are investigated. To achieve this goal, continuous field measurements were collected with a Differential Optical Absorption Spectroscopy (DOAS) system in 2011. The K-means clustering is used to classify the daily variation cycles of these measurements associated with different visibility and meteorological conditions such as temperature, relative humidity, wind speed and direction. The Davies-Bouldin index strategy is used to determine the optimal number of clusters. The regime of each cluster associated with visibility and meteorological conditions was then explored and compared. The comparative analyses show that both the maximum hourly average concentrations and the maximum daily average concentrations of SO_2 and NO_2 occurred in spring. Only 0.04 percent and 3.19 percent of the days with SO_2 and NO_2, respectively, did not comply with the latest national ambient air quality standards of China (GB 3095-2012). Moreover, the clustering results highlighted three representative patterns of daily SO_2 concentrations and four representative patterns of daily NO_2 concentrations. Both similarities and differences were presented among these clusters. The consistent changes in aerosol concentration with the changes in the measurements of NO_2 and SO_2 in each cluster provided supplemental evidence for the presence of the daily patterns of SO_2 and NO_2.

Keywords: NO_2; SO_2; DOAS; Cluster analysis.

INTRODUCTION

The rapid urbanization and industrial development in the past two decades greatly deteriorated air quality in China (Shao *et al.*, 2006; Chan and Yao, 2008; Kan *et al.*, 2012). Air pollution now has become one of the top environmental concerns in China (Tian *et al.*, 2007; Chan and Yao, 2008; Yang *et al.*, 2008; Han *et al.*, 2011; Kan *et al.*, 2012). The primary air pollutants, namely NO_2 and SO_2, in Chinese cities, are mainly emitted from industrial and domestic energy production, biomass burning and transportation (Shao *et al.*, 2006; Tian *et al.*, 2007; Yang *et al.*, 2008; Zhao *et al.*, 2008; Wang *et al.*, 2010). As a consequence, they induce a number of environmental issues. For instance, NO_2 is attributable to the acidification of terrestrial ecosystems, eutrophication of lakes and the marine environment, formation of tropospheric O_3 and degradation of human health and agricultural productivity (Grennfelt *et al.*, 1994; Placet *et al.*, 2000; Yang *et al.*, 2008). Similarly, SO_2 greatly affects public health and the habitat suitability of plants and animals. In addition, SO_2 is a precursor of acid rain and atmospheric particulates (Fisher *et al.*, 2011). The major anthropogenic source of SO_2 is the burning of sulfur-containing fossil fuels for domestic heating, power generation and industrial activities (Shao *et al.*, 2006; Tian *et al.*, 2007; Zhao *et al.*, 2008; Lu *et al.*, 2010).

To implement effective air pollution control strategies in polluted urban areas, it's essential to properly identify the local air quality regimes based on the levels and behaviors of air pollutants. Carefully mining a large scale of pollution data is beneficial for improving measurement retrievals and modeling on both regional and global scales (Kuebler *et al.*, 2002; Gariazzo *et al.*, 2007). While the existing studies (e.g., Zhuang, 2007; Wu *et al.*, 2012; Du *et al.*, 2013; Li *et al.*, 2013) on identifying the pollutants variation patterns are frequently constrained by the use of seasonal or monthly average values. This is because an averaged analysis cannot appropriately reflect the actual variation cycles when multiple

* Corresponding author.
Tel.: 1-403-220-7883
E-mail address: kddu@ucalgary.ca

natural variation types present in the studied period. Other than the sources and sinks of air pollutants, meteorological factors also affect the transport, transformation, reaction and removal of the emitted pollutants, and thereby affect the variation of their concentrations. Similar meteorological conditions can occur on the same day of the month regardless of which month or season (Baxla *et al.*, 2009; Rana *et al.*, 2009; Wu *et al.*, 2010). Therefore, grouping pollutants according to similar mixing ratios and daily evolutions is more effective in identifying the pollutants' daily variation properties and influences based on meteorological factors (Flemming *et al.*, 2005; Beaver and Palazoglu, 2006; Adame *et al.*, 2012; Austin *et al.*, 2012). For instance, Adame *et al.* (2012) applied a K-means clustering method as the grouping algorithm to investigate the daily patterns of air pollutants and obtained several distinct daily cycles of surface ozone, SO_2 and NO_2, in a heavily industrialized area of Puertollano, Spain. The K-means clustering method was also used in characterizing the classes of ozone episodes in the San Francisco Bay (Beaver and Palazoglu, 2006) and identifying distinct multi-pollutant profiles in the air (Austin *et al.*, 2012). Despite the impressive features, the application of the K-mean clustering method, however, still needs more independent validation under different atmospheric conditions and at different locations because temperature, relative humidity (RH), wind speed, wind direction and other parameters have significant impact on the daily variation of air pollutants.

Additionally, to the best of our knowledge, the existing studies, which examined the local air pollutants in Xiamen, are all based on point measurements collected at the ground level (e.g., Zhuang, 2007; Deng *et al.*, 2012; Du *et al.*, 2013; Li *et al.*, 2013; Wang *et al.*, 2013). Due to the very complex topography and urban street configuration in Xiamen City, the distribution of pollutant emissions is not homogeneous. These point measurements only represent conditions at specific surveying sites.

The Differential Optical Absorption Spectroscopy (DOAS) was introduced by Ulrich Platt in late 1970s. It's a non-contact air pollution monitoring technique using long path absorption method and can measure multiple trace gases simultaneously in real time (Platt and Stutz, 2008). As DOAS allows large-area monitoring for an open path over hundreds of meters, it can measure path-averaged concentration of gaseous pollutants. That's very useful to overcome the spatial sampling limitations of those point measurements (Platt and Stutz, 2008).

In this work, we present for the first time a DOAS measurement of SO_2 and NO_2 in 2011 in Xiamen. Firstly, the pollution levels of SO_2 and NO_2 in 2011 are evaluated at the observation area. Secondly, air quality regimes are characterized by using the K-means clustering of the daily patterns of air pollutants associated with meteorological conditions. To appropriately select the optimal K value, the Davies–Bouldin index method is applied. The aim of this work is to investigate the daily patterns of SO_2 and NO_2 as well as their association with visibility and meteorological conditions in a suburban area in Xiamen.

METHODOLOGY

Site Description

Xiamen, a rapidly urbanizing coastal city in southeastern China, has a population of approximately 3.61 million and an urban area of 300 km^2. The climate is affected by typical subtropical oceanic monsoons. During 2011, the air temperature ranged from 4.0°C in January, to 37.7°C in June, with an annual average of 20.8°C. 86.6% of the total precipitation occurred in the summer from May to August and in November in 2011. The prevailing wind direction is northeast with a strength level of 3 to 5 (Xiamen Meteorological Bureau, 2011).

The measurements were continuously collected on the 9^{th} floor (approximately 27 m above ground) of a building (24.61°N,118.06°E) located on the campus of the Institute of Urban Environment at the Chinese Academy of Sciences in Xiamen from January 20 to December 16, 2011. The building is located in an urban educational and residential area, approximately 14 km northwest of the city center. All dates and time reported are local time (LT) with 8 hours ahead of the Coordinated Universal Time (UTC).

DOAS System and Data Collection

In this study, the DOAS system used was developed by AIOFM (Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences). The set-up of the DOAS system has been described in detail in our previous studies (Qin *et al.*, 2006; Qin *et al.*, 2009), so here we confine ourselves to a brief description. The DOAS system consists of a light source, a suite of emitter and receiver, a spectrometer combined with detector array and an electronic control system. The length of the light path is 340 m. The instrument was periodically calibrated using the Hg emission spectrum. The spectral resolution was 0.4 nm (FWHM) at around 334 nm.

The concentrations of trace species are quantified by recording and fitting their differential absorption features with reference absorption cross-sections. The differential absorption features are separated from the broad band features mainly caused by atmospheric scattering and broad band absorption over an open light path (Qin *et al.*, 2006; Qin *et al.*, 2009). To prepare the corresponding reference absorption cross-sections, the literature-based absorption cross sections with high revolution were used in DOAS data retrieval (Voigt *et al.*, 2001; Vandaele *et al.*, 1994; Voigt *et al.*, 2002). The wavelength regions for retrieving NO_2 and SO_2 are 338–367 nm and 294–320 nm, respectively. The mean detection thresholds were calculated according to Stutz and Platt, (1996). For SO_2 and NO_2, the mean detection limits are 2.60 $\mu g/m^3$ and 3.74 $\mu g/m^3$, respectively, with an uncertainty of less than 10% (Stutz and Platt, 1996; Qin *et al.*, 2006; Qin *et al.*, 2009).

The path-averaged concentrations of SO_2 and NO_2 were continuously measured by the DOAS system. The temporal resolution of the measurements varied between 2 minutes and 5 minutes, depending on ambient visibility. There were occasions when the DOAS system stopped collecting data because of very low visibility, malfunction of the hardware or power shortage. Thus, only days with more than 18 hours of valid measurements were extracted for cluster analysis.

To investigate the mechanism corresponding to the variation patterns of chemical species, meteorological parameters and horizontal visibility were simultaneously monitored using an automatic weather station (VAISALA MAWS301) operated by the Xiamen Meteorological Bureau. The weather station, located 270 m away from the DOAS system, was installed on the ground to continuously measure wind speed and direction, temperature, relative humidity and horizontal visibility since March 15, 2011. The heights of the sensors to measure the proposed parameters were 10 m, 2 m, 2 m and 4 m respectively, with data being collected every 5 minutes. Several previous studies have suggested that horizontal visibility is well correlated to the concentration of aerosols (Wu *et al.*, 2005; Deng *et al.*, 2008). Low values of horizontal visibility imply relatively heavy aerosol loading and vice versa, mainly because of the extinctions of aerosol. So we can use horizontal visibility as a proxy for the aerosol content. However horizontal visibility may also be dependent on the RH, since the hygroscopicity of aerosols and water vapor can affect visibility values under high RH conditions (Rosenfeld *et al.*, 2007; Wang *et al.*, 2011). Therefore, the visibility used in this study was revised by Eq. (1) in case RH > 40% (Rosenfeld *et al.*, 2007)

$$VIS = \frac{VIS_{original}}{0.26 + 0.4285 \times \log_{10}(100 - RH)} \quad (1)$$

where *VIS* is the adjusted visibility value, $VIS_{original}$ is the value directly measured by the visibility sensor, and RH is the average daily relative humidity in percentage.

K-means Clustering Algorithm

The K-means clustering algorithm was introduced by Hartigan (Hartigan, 1975). It employs an iterative search procedure to allocate N observations into K clusters, in which each observation belongs to the cluster with the nearest center. In the observation period, the days with similar diurnal patterns of pollutants were grouped together by applying the K-means clustering algorithm, where K is an integral number of clusters subjectively predetermined by prior knowledge of the observed data set. However, no prior knowledge or references are available to define the optimal number of clusters (K) for the specific case in this study. Therefore, a two-step K value determination process was proposed: first, all possible values of K in the range of 2 to 8 clusters are chosen to perform grouping (Ramze *et al.*, 1998; Austin *et al.*, 2012); second, the preferred K value is determined using the Davies–Bouldin index (DBI), and an internal metric for evaluating clustering algorithms, as described in Eq. (2)

$$DBI = \frac{1}{K}\sum_{i=1}^{K} \max_{j:i\neq j}\left(\frac{s_i + s_j}{M(c_i, c_j)}\right) \quad (2)$$

where K is the number of clusters, s_i is the average distance of observations in cluster i to the center c_i and s_j is the average distance of observations in cluster j to the center c_j. Small values of DBI represent clusters that are compact and their centers are far from each other. Hence, the clustering values of K that minimize the DBI are the preferred options. In order to reduce the statistical error, the K-means algorithm was run 100 times with multiple initial seeds selected randomly. The DBI calculation and k-means algorithm used were conducted using SPSS v.18.00.

RESULTS AND DISCUSSION

Pollution Levels of NO_2 and SO_2

In order to assess the pollution levels of NO_2 and SO_2 for the study period at the observation site, the hourly mass concentration values have been analyzed and compared with the latest version of National Ambient Air Quality Standards of China (GB 3095-2012) (MEP, 2012), which was promulgated in 2012 and to be implemented from 2016 by the Ministry of Environmental Protection, China. As shown in Table 1, the standard deviations of the concentration values for SO_2 and NO_2 were both relatively high, compared with their annual average values. The maximum hourly values of NO_2 and SO_2 concentration were much higher than the annual average values, but the frequencies of occurrence were very low, as shown in Fig. 1. The occurrence frequency of concentration values for SO_2 between 10 μg/m^3 and 50 μg/m^3 was 91.9%, and that for NO_2 between 40 μg/m^3 and 80 μg/m^3 was 82.1%. Both frequency distribution profiles for the two species were single-peaked. These characteristics indicate that the sources for NO_2 and SO_2 at the observation site were relatively few.

According to GB 3095-2012 criteria (MEP, 2012), the annual average value of sulfur dioxide (SO_2), as shown in Table 1, was slightly higher than the threshold of grade I (20 μg/m^3) but lower than that of grade II (60 μg/m^3), meanwhile the annual average concentration of nitrogen dioxide (NO_2) slightly exceeded the thresholds of grade I and grade II (both threshold values for grade I and grade II are 40 μg/m^3). In addition, the exceedance rates of NO_2 and SO_2 in other averaging time scales (1-hour and 24-hour) are summarized in Table 2. Similarly, the exceedance rates of hourly and daily (24 hours) average SO_2 concentration values were higher than the threshold of grade I, but were lower than grade II. The concentration of nitrogen dioxide slightly exceeded both the thresholds of grade I and grade II. The pollution level of NO_2 is slightly higher than SO_2. Compared to the GB 3095-2012 criteria (MEP, 2012), the ratios of exceedance days to the total days for SO_2 and NO_2 were only 0.04% and 3.19%, respectively, during the observation period. In general, the air pollution of SO_2 and NO_2 at the observation site in Xiamen is not serious.

Table 1. The annual average, standard deviation and extreme values of NO_2 and SO_2 concentration.

Pollutant	SO_2 (μg/m^3)	NO_2 (μg/m^3)
Annual average value	23.18	46.48
Standard Deviation	18.10	23.69
Maximum value	155.46	231.50
Minimum value	< 2.60	< 3.74

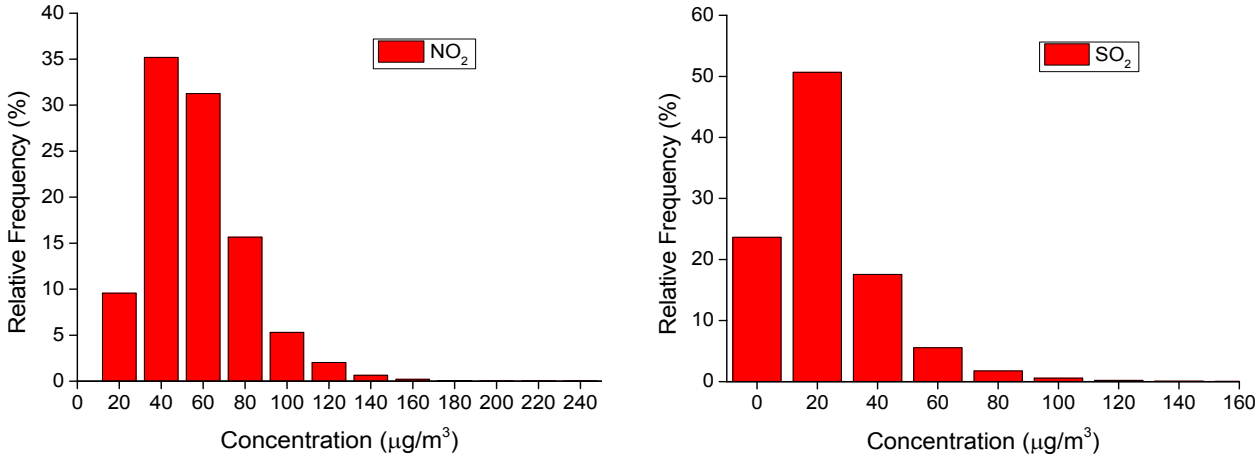


Fig. 1. Frequency distributions of the hourly concentrations of NO_2 and SO_2.

Table 2. The exceedance rates of NO_2 and SO_2 at the observation site during the study period.

Pollutant	Averaging time	Exceedance rate		Reference (Unit: μg/m^3)	
		Grade I	Grade II	Grade I	Grade II
SO_2	1-hrs	0.04%	0	150	500
	24-hrs	1.42%	0	50	150
NO_2	1-hrs	0.04%	0.04%	200	200
	24-hrs	3.19%	3.19%	80	80

Daily-Monthly Variations of NO_2 and SO_2

In Fig. 2, an overview of daily-monthly variations for NO_2 and SO_2 from January 20 to December 16, 2011 at the observation site is shown. The mass concentration values of NO_2 were between 23 μg/m^3 and 90 μg/m^3. The minimum value was 23.6 μg/m^3, occurring between 12:00 and 16:00. The mass concentration values of NO_2 which were higher than 87 μg/m^3 were observed between 18:00 and 23:00, while the mass concentration values which were higher than 74 μg/m^3 were observed between 8:00 and 9:00. High NO_2 concentration values were measured in the cold weather from February to April. The concentrations of SO_2 were lower than 48μg/m^3 for the majority of the observed 24-hour period. The maximum value was around 75 μg/m^3 from 7:00 to 10:00 in April. In general, the hourly levels of SO_2 were lower than NO_2.

As observed in this figure, the highest NO_2 and SO_2 hourly concentrations were observed in the spring, especially in April. There were two peak values for the monthly diurnal variation of NO_2, one of them was observed around 09:00 and the concentration value decreased during the day. This could be due to the photochemical reactions involving NO_2 during the day. The other was observed at night. Only one peak of SO_2 occurred simultaneously with the NO_2 peak in the morning.

Diurnal Patterns of NO_2 and SO_2

K-means cluster algorithm was used to obtain the ensemble of clusters for the diurnal variations of NO_2 and SO_2 measured during the studied period. The average diurnal evolution of each species was evaluated in combination with the meteorological parameters measured simultaneously. Considering the automatic weather station was employed on March 15, 2011, the period for cluster analysis was set from March 15 to December 16, 2011. After removing the days of which the data were missed due to instrument failure or measured at the time with too poor quality of original spectrum in the reporting period, the total number of days used for the grouping analysis was 235.

Before applying the algorithm it is necessary to assign the number of clusters (k), which generally depends on pre-existing characteristics of the data set. The Davies–Bouldin index (DBI) was analyzed to identify the most possibly correct K value, as shown in Fig. 3.

Besides minimizing the DBI value, the results (i.e., the number of days) obtained in the classification should also be considered when determining the most optimal K value. For high K values (e.g., $K > 8$), the number of days in each cluster was very small, which would cause the cluster to lose representativeness, and therefore this classification could not provide new diurnal patterns compared to classifications with smaller k values (Ramze *et al.*, 1998; Austin *et al.*, 2012). However, if the K value is too small (e.g., ≤ 2), there would be a loss in information. For NO_2, it suggests that the most compact clusters were obtained for the solution of $K = 3$. Hence, the 235 days were grouped into three clusters according to NO_2 hourly concentrations. Similarly in Fig. 3(b), the DBI value for solution of $K = 4$ is lowest compared to other solutions. The 235 days were grouped into four clusters according to SO_2 hourly concentrations.

Daily Pattern for NO_2

The daily variation patterns of NO_2, SO_2, and

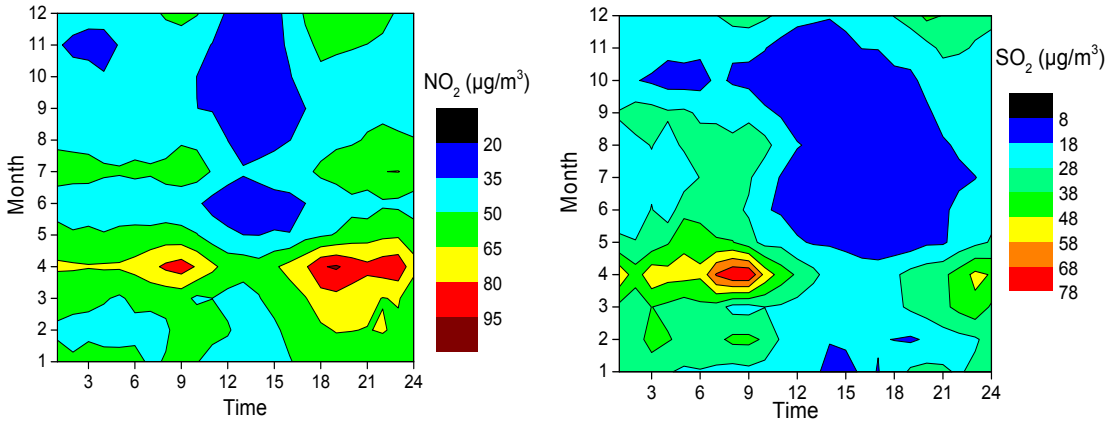


Fig. 2. Daily-monthly variations of hourly-mean concentrations of NO_2 and SO_2.

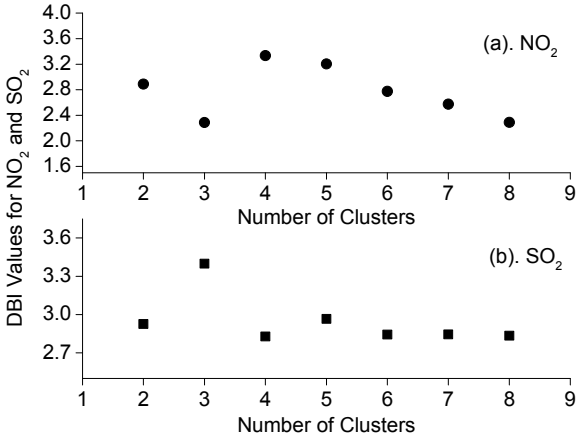
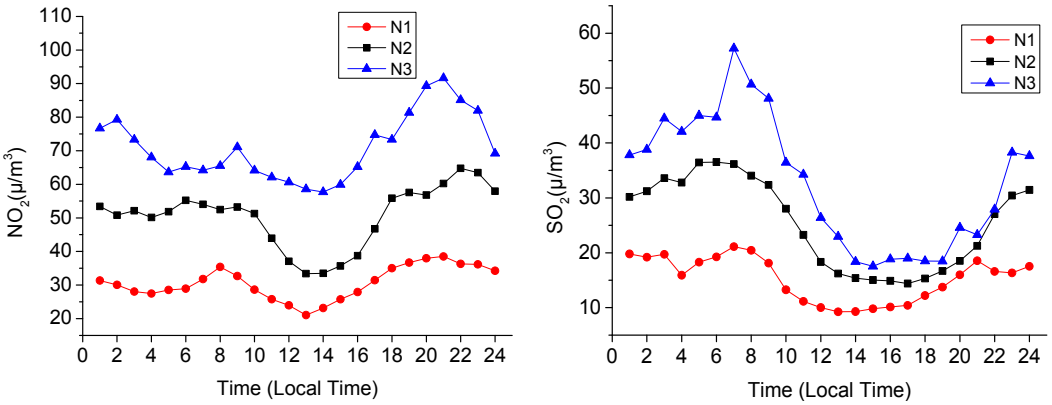


Fig. 3. Davies and Bouldin indices of NO_2 and SO_2 for difference cluster numbers.

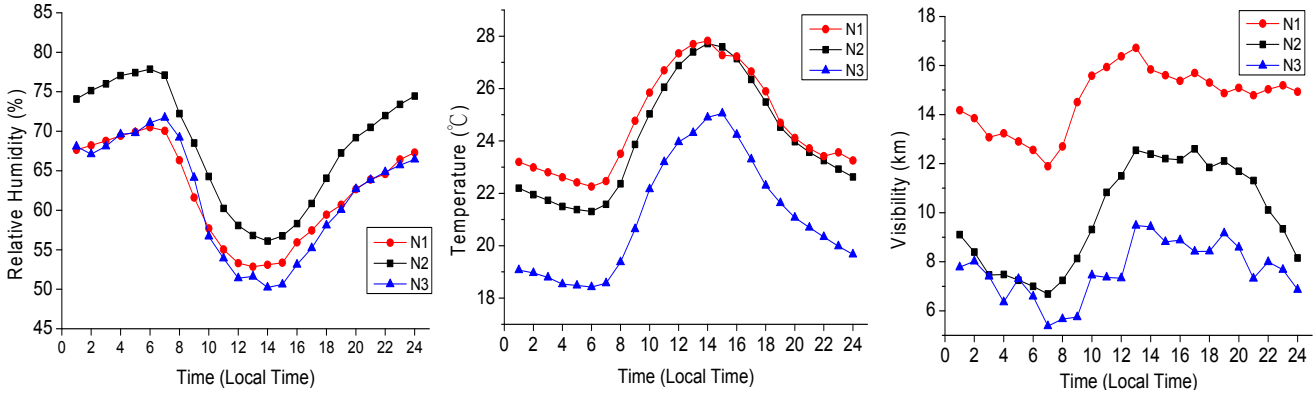
meteorological factors are shown for the three clusters (N1, N2, and N3), as shown in Fig. 4. The clusters were sorted from the lowest hourly average concentration of NO_2 to the highest value.

Fig. 4 shows that the daily variations of NO_2, N1, N2 and N3, have different concentration levels but show a similar daily behavior with high mass concentration values during the night and decreasing during the day, then reaching a minimum around 12:00–15:00 (Local Time). The amplitude (difference between the maximum and the minimum concentrations of NO_2) for each cluster is relatively different. The night concentration values varied from 27.46 μg/m^3 in cluster N1 to 91.66 μg/m^3 in cluster N3, while the day values varied between 21.07 μg/m^3 in cluster N1 to 74.74 μg/m^3 in cluster N3.

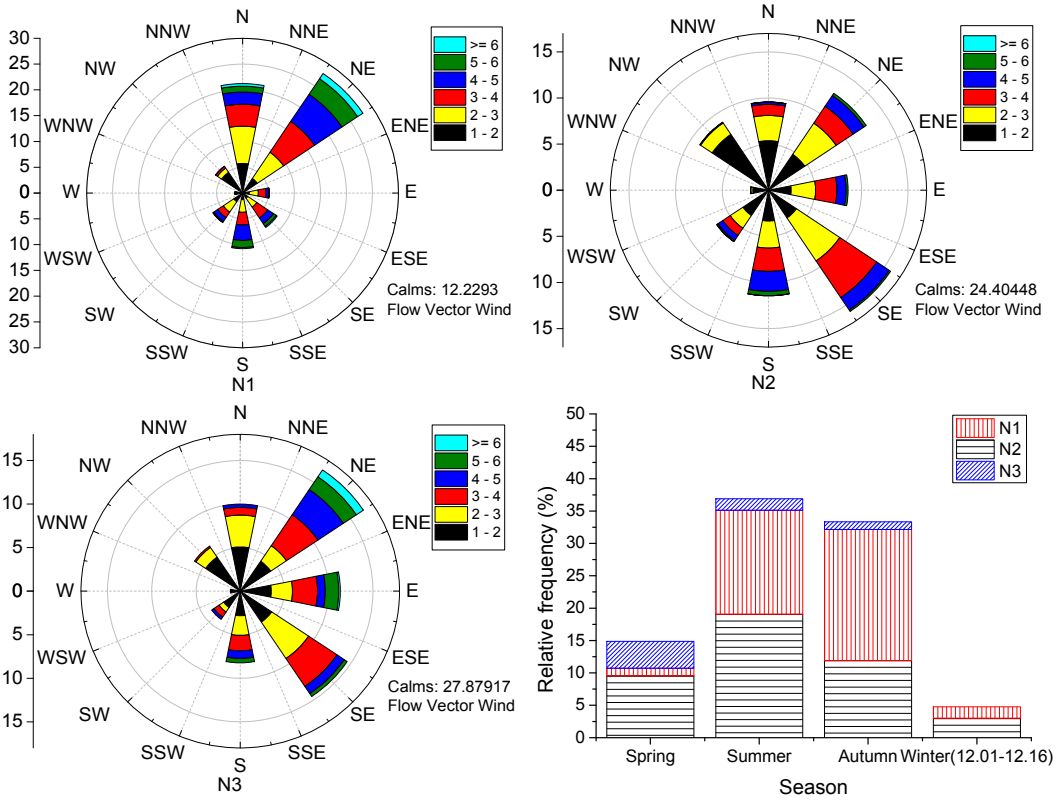
Cluster N1 represents a daily cycle with the lowest concentration values for NO_2 and SO_2, and includes 39.3% of the total days, which makes it the second largest group among the three clusters. For NO_2, cluster N1 has the lowest daily amplitude of 17 μg/m^3. The night values oscillated between 27.46 μg/m^3 and 38.53 μg/m^3 with amplitude of 11.1 μg/m^3. During the daytime, the concentration gradually declined to a minimum concentration of 21.07 μg/m^3 around 13:00, when the ozone (O_3)'s photochemical activity was high, and then rose gradually in the afternoon. The amplitude of NO_2 concentration in the daytime was 14.3 μg/m^3. The daily cycle of NO_2 has two peaks, which is similar to the cases in N2 and N3. Similar to NO_2, N1 also has the lowest SO_2 concentrations among the three clusters. SO_2 concentration in N1 oscillated between 9.24 μg/m^3 and 21.11 μg/m^3. The average concentration value at night is higher than that during the daytime for SO_2. The peak value of 21.11 μg/m^3 was observed around 8:00. N1 has the highest visibility level among the three clusters, with an average visibility of 13.66 km. The temperatures during the period of days for N1 were generally higher than those in N2 and N3, ranging between 22.26°C and 27.82°C. The RH values of N1 were at medium-low level, ranging from 52.9% to 70.5%. During the night, the RH rose to 70.51%, and then declined during the day to a minimum of 52.87%. Most of the time the wind speed was below 3 m/s on the days of N1, with average wind speeds of 3.6 m/s and 2.2 m/s during the day and night, respectively. In addition, most of the days in cluster N1 were in the summer and autumn, with a massive burst of precipitation and active convection. These conditions do not chemically favor the accumulation of air pollutants, including particles, though the main wind direction in cluster N1 was N-NE, blowing from an industrial area. The concentration level of NO_2 is also affected by photochemical activity involving O_3 during the day.



(a) NO_2 daily patterns (left) and associated daily variations of SO_2 (right).



(b) Associated daily variations for relative humidity, temperature and visibility for each cluster.



(c) Associated variations for wind rose and seasonal relative frequency (%) for each cluster.

Fig. 4. NO_2 daily patterns for the three clusters (N1, N2 and N3) and daily evolutions associated for SO_2 (a), relative humidity, temperature, visibility, seasonal relative frequency (%) and daily variations for and wind rose for each cluster (b) and (c).

The daily pattern for cluster N2 occurred most frequently during the observation period. The days in this cluster were distributed over all the months of the study period. The daily amplitude and average concentration values were 31.38 μg/m^3 and 50.43 μg/m^3 for NO_2, and 22.14 μg/m^3 and 25.40 μg/m^3 for SO_2, respectively. The minimum concentration values were 33.41 μg/m^3 for NO_2 around 13:00 and 14.38 μg/m^3 for SO_2 around 17:00. The concentration amplitudes of NO_2 were 14.7 μg/m^3 and 24.2 μg/m^3 during the night (19:00–23:59 and 00:00–08:00) and day (08:00–19:00), respectively. For SO_2, the concentration amplitudes were 14.7 μg/m^3 and 22.4 μg/m^3 during the night and day, respectively. The daily pattern profiles are similar to those in N1, but with higher concentration levels overall. The cycles and values of the temperature and its amplitude in N2 are similar to the cases of N1. The values of relative humidity in N2 are higher than those in N1, but both have a similar daily pattern. There was not a clear prevailing wind direction in N2, but the wind blowing from S to SE was slightly more frequent than other directions. The wind speed was below 3 m/s 79.2% of the time in N2. The daily average wind speed was 2.01 m/s. The average wind speeds during the day and night were 2.6 m/s and 1.5 m/s, respectively. Similar to N1, the smaller amplitudes and higher average concentrations of NO_2 and SO_2 in N2 are related to unfavorable meteorological conditions for air dispersion, e.g., lower temperature, low wind speed, weak convection and solar radiation.

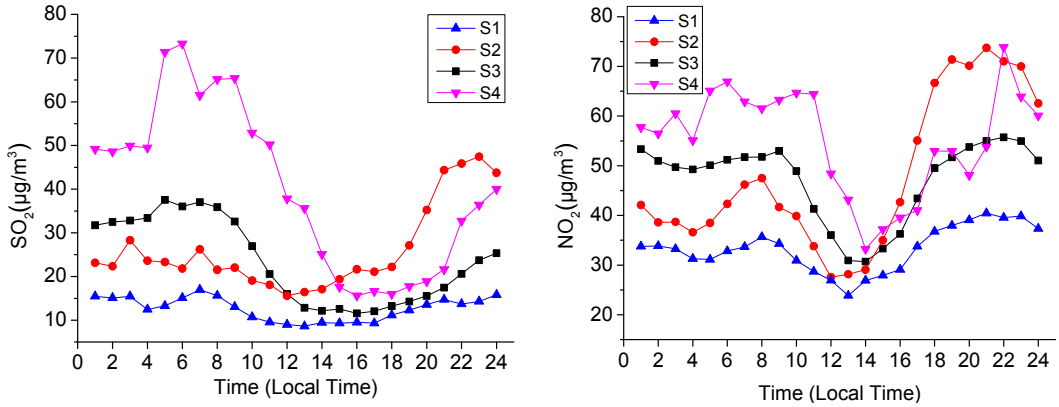
The daily pattern for cluster N3 occurred the least frequently during the period of observation, but N3 had the highest NO_2 concentration values among the three clusters. Its daily cycle is similar to the cases in N1 and N2, but with the highest daily amplitude. During the night, the concentration values of NO_2 oscillated between 63.6 μg/m^3 and 91.7 μg/m^3, while the maximum value of 91.7 μg/m^3 occurred around 21:00 (Local time). The nighttime average value of NO_2 was 77.1 μg/m^3. During daytime, the variation amplitudes of NO_2 concentration were 17.0 μg/m^3 with the maximum value of 71.1 μg/m^3 occurring around 9:00 (Local time), and the day average values were 64.8 μg/m^3. In cluster N3, SO_2 showed the highest levels of concentration among the three clusters. Its daily amplitude was 39.7 μg/m^3 with the minimum of 17.5 μg/m^3 observed from 15:00 to 18:00 (Local time), and a peak value of 57.2 μg/m^3 observed around 8:00. The visibility in N3 is the worst among the three clusters, with the average value of 7.6 km. The daily pattern, values and amplitude of RH in N3 were all similar to those of N2. The maximum temperature amplitude in N3 was 6.6°C. The average temperature in N3 was 21.2°C, which is lower than that of N1 and N2. The prevailing wind directions were SE and E-SE with the average value of 2.1 m/s.

Daily Pattern for SO_2

A cluster analysis was also conducted based on the temporal series of SO_2 concentration data. Four types of daily pattern of SO_2, NO_2, and meteorological conditions have been obtained, as illustrated in Fig. 5. The average and oscillating amplitude values in each cluster for SO_2, NO_2 and visibility are summarized in Table 3.

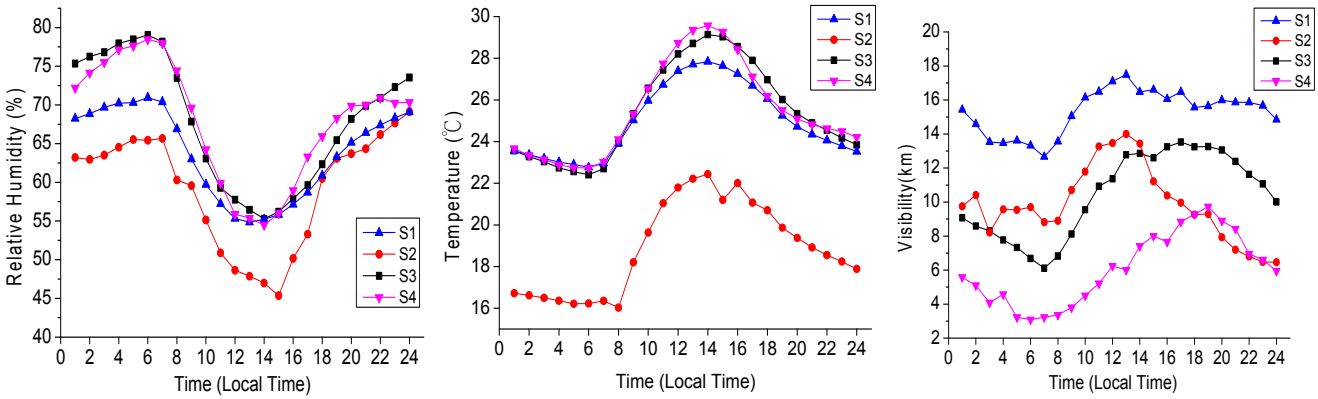
S1 is the largest group among the four clusters (S1, S2, S3 and S4), and includes 40.3% of days in the entire period. Most of the days in S1 were in warmer weather (between July and October). Parameters such as the average concentration, the oscillating amplitude for SO_2, NO_2 and the horizontal extinction of aerosols in this cluster are all the lowest among the four clusters. The average relative humidity and temperature in S1 were 63.9% and 25.0°C, respectively. The oscillating amplitudes were 16.1% and 5.1°C, respectively. The main wind direction was N-NE, and the mean wind speed was 2.8 m/s. Meteorological conditions in S1 were very similar to those in N1. Similar to N1, S1 also had the smallest amplitude and lowest average concentration among S1, S2, S3 and S4. The daily pattern in S1 or N1 can be taken as the background levels for SO_2 or NO_2. Compared with other clusters, these background levels are less likely to be affected by industrial emissions.

The daily variation trends during the days in S3 and S4 for NO_2 and SO_2 were similar to those in S1, but had different concentration levels. S3 is the second largest group after S1, and it consists of days in every month in the

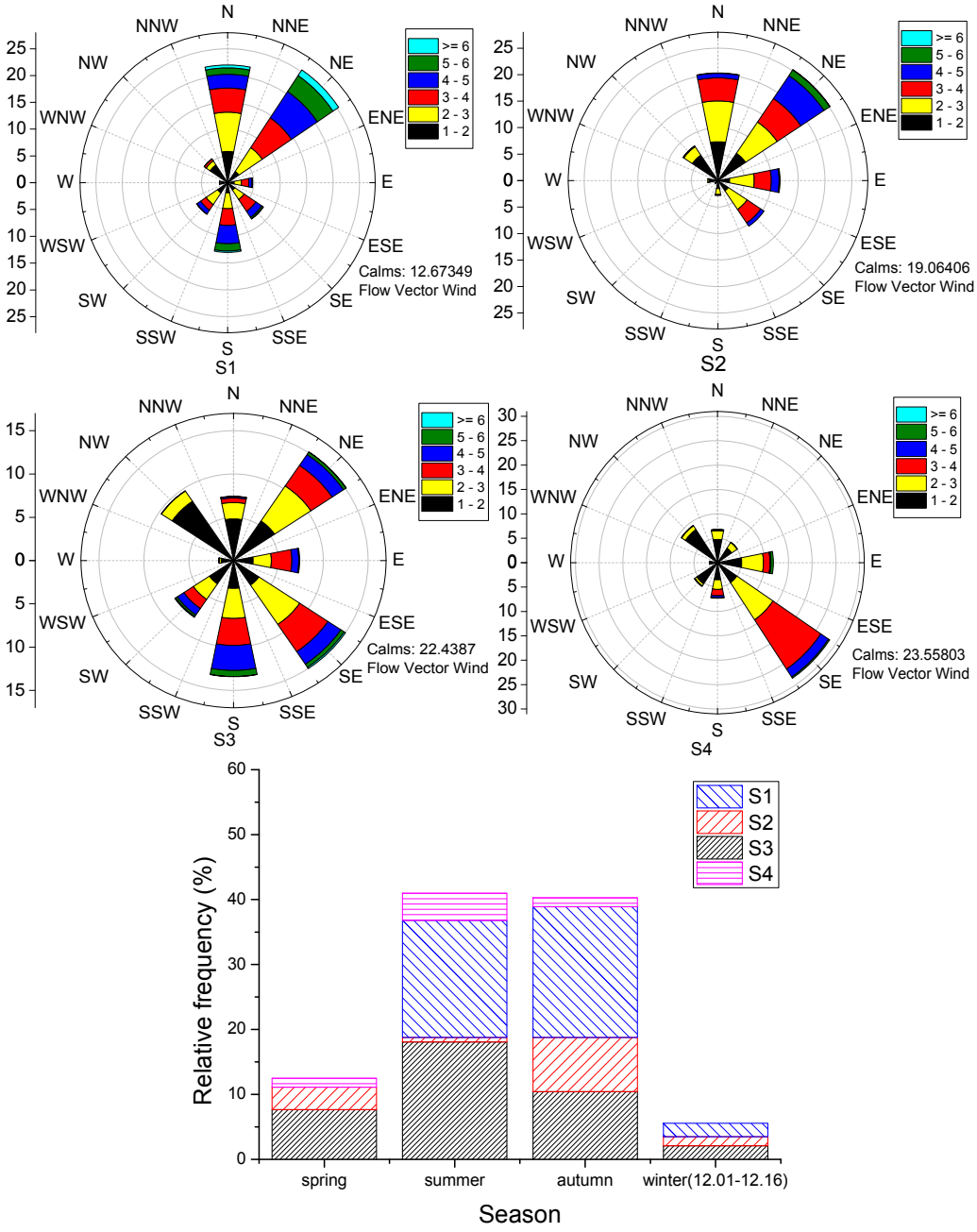


(a) SO_2 daily patterns (left) and associated daily variations of NO_2 (right).

Fig. 5. SO_2 daily patterns for the four clusters (S1,S2, S3 and S4) and daily evolutions associated for NO_2 (a), relative humidity, temperature, visibility, seasonal relative frequency (%) and daily variations for and wind rose for each cluster (b) and (c).



(b) Associated daily variations of relative humidity, temperature and visibility for each cluster.



(c) Associated variations for wind rose and seasonal relative frequency (%) for each cluster.

Fig. 5. *(continued).*

Table 3. The average and amplitude values for NO_2, SO_2 and visibility in clusters of S1–S4.

Species	Value	S1	S2	S3	S4
SO_2 (μg/m^3)	Average	12.7	26.1	23.5	40.4
	Amplitude	8.3	31.8	26	57.7
NO_2 (μg/m^3)	Average	33.3	47.9	47.2	55.3
	Amplitude	16.6	46.2	25	40.6
Visibility (km)	Average	14.1	9.4	9.5	5.5
	Amplitude	5.6	7.9	7.7	6.2

study period. The average concentrations of NO_2 or SO_2 during the night were higher than those during the day for S3, and the average concentration difference between day and night in S3 was higher than that in S1. The S4 is the smallest group among the four clusters. Its daily pattern was similar to the cases of N1 and N2, but had the highest SO_2 and NO_2 average and oscillating amplitude concentration values among the four clusters. Like S3, the average concentrations of NO_2 or SO_2 during nighttime were higher than those during the daytime in S4.

The daily patterns of relative humidity and temperature in S3 and S4 are very similar, but the wind regimes were apparently different. During the day (08:00–19:00), the wind in S3 and in S4 mainly blew S-SE with an average speed of 2.76 m/s and E-SE with an average speed of 2.5 m/s, respectively. However, in the nighttime (00:00–07:59 and 19:00–23:59), the winds associated with S3 and in S4 mainly blew NW-N-NE with the average speed of 1.43 m/s and NNW-N-NE with the average speed of 1.2 m/s, respectively. An obvious shifting of wind directions from NNW-N-NE to S-SE was also observed in S4.

To the north of observation site is the Tongan industrial district. Considering that SO_2 is a representative primary chemical species, the high concentrations in S3 and S4 could result from very unfavorable dispersion conditions and emissions from industrial area during in the night. As shown in Fig. 5(a), the variation curve of daily pattern in S4 is less smooth than that in S3, and during in the night the average concentration difference between S3 and S4 is much larger than that in the daytime, which implies that S4 was more affected by abrupt pollutants emissions during the nighttime. However, the daily pattern in S4 was much less frequent than that in S3. During the day the wind was primarily from the ocean with relative higher average speed, which greatly improved dispersion conditions, and the concentrations of SO_2 and NO_2 decreased in S3 and S4. Meanwhile, the variation of NO_2 was also affected by local photochemical reactions and vehicle emissions.

S2 had higher concentrations of SO_2 and NO_2 than S1. The variation trends in S2 were similar to those in S1 before 16:00, but the concentrations show an unusually fast increase from 16:00 to 24:00. In this time period, the main wind direction in S1 and S2 are both NE, where the Tongan industrial district is located. However the average concentration in S2 was much higher than that in S1 during these 8 hours. The 24-hour average of wind speeds in S1 and S2 were 2.7 m/s and 2.2 m/s, respectively, while the average speeds from 18:00 to 24:00 in S1 and S2 were 2.6 m/s and 1.8 m/s, respectively. In addition, considering the weather during the days in S2 are relatively drier and colder than S1 (Fig. 5(b)), the favorable dispersion conditions in S1 could be the reason for lower concentrations of NO_2 and SO_2 compared to S2.

Discussion of Visibility

The horizontal visibility was measured with an independent instrument and adjusted to eliminate the influence of high relative humidity using Eq. (1). Therefore it can be used as an indicator of the concentration of aerosols. It is demonstrated that the relative levels and variation trend of aerosol concentration are similar to those of SO_2 or NO_2 in each cluster in Figs. 4 and 5. A typical example was observed in S2, where the rapid decrease in visibility was associated with the rapid increase of SO_2 and NO_2 concentrations during the time period from 16:00 to 24:00. The major sources of NO_2 and SO_2 are considered as vehicle and industrial emissions, which are also important sources of aerosols in China. The variation of visibility in a short time-scale reflects that of air species (e.g., NO_2 and SO_2) to some degree (Wang *et al.* 2012). Thus the synchronous variations of visibility here provide supplementary evidence for the existence of the SO_2 and NO_2 daily patterns observed using K-means cluster analysis. Similarly, Wang *et al.* (2012) studied the weekly cycle of horizontal visibility and PM_{10} and found that negative correlation existed between them in southeastern China. In addition, the weekly cycles of PM_{10} were in phase with that of SO_2 and NO_2. These findings are in consistence with the anti-phase variation of visibility with NO_2 and SO_2 observed in this study.

CONCLUSIONS

Continuous measurements of SO_2 and NO_2 by the DOAS system at a suburban site in the north of Xiamen (24.61°N, 118.06°E) were conducted from January 2011 to December 2011. The results showed that both SO_2 and NO_2 concentrations measured by DOAS system had the maximum of 1-hour and 24-hour mean values in the spring (from March to May), with maximal pronounced values in April. Compared to the latest version of National Ambient Air Quality Standards (GB 3095-2012) in China (MEP, 2012), the ratios of exceedance days to the total days for SO_2 and NO_2 were only 0.04% and 3.19%, respectively, during the observation period, indicating very slight pollution of SO_2 and NO_2 at the observation site.

In this study, a K-means cluster analysis has been employed to classify observation days into groups based on the daily concentration profiles of NO_2 and SO_2. To get optimum

cluster numbers based on SO_2 and NO_2 measurements, DBI results obtained from different cluster numbers were compared using the Davies–Bouldin index analysis. A total of three clusters (N1–N3) were obtained based on NO_2 data while four clusters (S1–S4) based on SO_2 data. For each cluster of days the pollution levels and daily patterns for NO_2 and SO_2 were evaluated to explore the reasons of differences among clusters, taking into consideration the meteorological parameters associated with that cluster. In addition, the consistent changes of visibility with the changes of the SO_2 and NO_2 measurements in each cluster provided supplemental evidence for the presence of the daily patterns of SO_2 and NO_2. The results obtained from the proposed cluster analysis approach, will provide meaningful information for air quality regime and pollution control in the Xiamen region.

ACKNOWLEDGEMENTS

The authors thank the support from Fujian Distinguished Young Scholar Career Award (2011J06018), Strategic Priority Research Program of the Chinese Academy of Sciences (XDB05040400), and Knowledge Innovation Program of the Chinese Academy of Sciences (KZCX2-EW-408).

REFERENCES

Adame, J.A., Notario, A., Villanueva, F. and Albaladejo, J. (2012). Application of Cluster Analysis to surface Ozone, NO_2 and SO_2 Daily Patterns in an Industrial Area in Central-Southern Spain Measured with a DOAS System. *Sci. Total Environ.* 429: 281–291.

Austin, E., Coull, B., Thomas, D. and Koutrakis, P. (2012). A Framework for Identifying Distinct Multi Pollutant Profiles in Air Pollution Data. *Environ. Int.* 45: 112–121.

Baxla, S.P., Roy, A.A., Gupta, T., Tripathi, S.N. and Bandyopadhyaya, R. (2009). Analysis of Diurnal and Seasonal Variation of Submicron Outdoor Aerosol Mass and Size Distribution in a Northern Indian City and Its Correlation to Black Carbon. *Aerosol Air Qual. Res.* 9: 458–469.

Beaver, S. and Palazoglu, A. (2006). Cluster Analysis of Hourly Wind Measurements to Reveal Synoptic Regimes Affecting Air Quality. *J. Appl. Meteorol. Climatol.* 45: 1710–1726.

Chan, C.K. and Yao, X. (2008). Air Pollution in Mega Cities in China. *Atmos. Environ.* 42: 1–42.

Deng, J.J., Du, K., Wang, K., Yuan, C.S. and Zhao, J. (2012). Long-Term Atmospheric Visibility Trend in Southeast China, 1973-2010. *Atmos. Environ.* 59: 11–21.

Deng, X., Tie, X., Wu, D., Zhou, X., Bi, X., Tan, H., Li, F. and Jiang, C. (2008). Long-term Trend of Visibility and Its Characterizations in the Pearl River Delta (PRD) Region, China. *Atmos. Environ.* 42: 1424–1435.

Du, K., Mu, C., Deng, J.J. and Yuan, F. (2013). Study on Atmospheric Visibility Variations and the Impacts of Meteorological Parameters Using High Temporal Resolution Data: an Application of Environmental Internet of Things in China. *Int. J. Sustainable Dev. World Ecol.* 20: 238–247, doi: 10.1080/13504509.2013.783886.

Fisher, J.A., Jacob, D.J., Wang, Q., Bahreinic, R., Carougeb, C.C. and Cubisonc, M.J. (2011). Sources, Distribution and Acidity of Sulfate-Ammonium Aerosol in the Arctic in Winter-Spring. *Atmos. Environ.* 45: 7301–7318.

Flemming, J., Stern, R. and Yamartino, R.J. (2005). A New Air Quality Regime Classification Scheme for O_3, NO_2, SO_2 and PM_{10} Observations Sites. *Atmos. Environ.* 39: 6121–6129.

Gariazzo, C., Silibello, C., Finardi, S., Radice, P., Piersanti, A., Calori, G., Cecinato, A., Perrino, C., Nussio, F., Cagnoli, M., Pelliccioni, A., Gobbi, G.P. and Di Filippo, P. (2007). A Gas/Aerosol Air Pollutants Study over the Urban Area of Rome Using a Comprehensive Chemical Transport Model. *Atmos. Environ.* 41: 7286–7303.

Grennfelt, P., Hov, O. and Derwent, D. (1994). Second Generation Abatement Strategies for NO_x, NH_3, SO_2 and VOC. *Ambio* 23: 425–433.

Han, S.Q., Bian, H., Feng, Y.C., Liu, A.X., Li, X.J., Zeng, F. and Zhang, X.L. (2011). Analysis of the Relationship between O_3, NO and NO_2 in Tianjin, China. *Aerosol Air Qual. Res.* 11: 128–139.

Hartigan, J.A. (1975). *Clustering Algorithms*, Wiley & Sons, New York.

Kan, H.D., Chen, R.J. and Tong, S.L. (2012). Ambient Air Pollution, Climate Change, and Population Health in China. *Environ. Int.* 42: 10–19.

Kuebler, J., Russell, A.G., Hakami, A., Clappier, A. and van den Bergh, H. (2002). Episode Selection for Ozone Modelling and Control Strategies Analysis on the Swiss Plateau. *Atmos. Environ.* 36: 2817–2830.

Li, T.C., Wu, C.Y., Chen, W.H., Yuan, C.S., Wu, S.P., Wang, X.H. and Du, K. (2013). Diurnal Variation and Chemical Characteristics of Atmospheric Aerosol Particles and Their Source Fingerprints at Xiamen Bay. *Aerosol Air Qual. Res.* 13: 596–607.

Lu, Z., Streets, D.G., Zhang, Q., Wang, S., Carmichael, G.R., Cheng, Y.F., Wei,C., Chin, M., Diehl, T. and Tan, Q. (2010). Sulfur Dioxide Emissions in China and Sulfur Trends in East Asia Since 2000. *Atmos. Chem. Phys.* 10: 6311–6331.

MEP (The Ministry of Environmental Protection) (2012). National Ambient Air Quality Standards (GB 3095–2012). Accessed on: http://kjs.mep.gov.cn/hjbhbz/bzwb/dqhjbh/dqhjzlbz/201203/t20120302_224165.htm.

Placet, M., Mann, C., Gilbert, R. and Niefer, M. (2000). Emissions of Ozone Precursors from Stationary Sources: A Critical Review. *Atmos. Environ.* 34: 2183–2204.

Platt, U. and Stutz, J. (2008). *Differential Optical Absorption Spectroscopy: Principles and Applications*. Springer Verlag, Heidelberg.

Qin, M., Xie, P.H., Liu, W.Q., Li, A., Dou, K., Fang, W., Liu, H.G. and Zhang, W.J. (2006). Observation of Atmospheric Nitrous Acid with DOAS in Beijing, China. *J. Environ. Sci.-China* 18: 69–75.

Qin, M., Xie, P.H., Su, H., Gu, J.W., Peng, F.M., Li, S.W., Zeng, L.M., Liu, J.G., Liu, W.Q. and Zhang, Y.H. (2009). An Observational Study of the HONO-NO_2 Coupling at an Urban Site in Guangzhou City, South China. *Atmos.*

Environ. 43: 5731–5742.

Ramze, R.M., Lelieveldt, B.B.F. and Reiber, J.H.C. (1998). A New Cluster Validity Index for the Fuzzy c-mean. *Pattern Recognit. Lett.* 19: 237–246.

Rana, S., Kant, Y. and Dadhwal, V.K. (2009). Diurnal and Seasonal Variation of Spectral Properties of Aerosols over Dehradun, India. *Aerosol Air Qual. Res.* 9: 32–49.

Rosenfeld, D., Dai, J., Yu, X., Yao, Z., Xu, X., Yang, X. and Du, C. (2007). Inverse Relations between Amounts of Air Pollution and Orographic Precipitation. *Science* 315: 1396–1398.

Shao, M., Tang, X.Y., Zhang, Y.H. and Li, W.J. (2006). City Clusters in China: Air and Surface Water Pollution. *Front. Ecol. Environ.* 4: 353–361.

Stutz, J. and Platt, U. (1996). Numerical Analysis and Estimation of the Statistical Error of Differential Optical Absorption Spectroscopy Measurements with Least-Squares Methods. *Appl. Opt.* 35: 6041–6053.

Tian, H.Z., Hao, J.M., Hu, M.Y. and Nie, Y.F. (2007). Recent Trends of Energy Consumption and Air Pollution in China. *J. Energy Eng.* 133: 4–12.

Vandaele, A.C., Simon, P.C., Guilmot, J.M., Carleer, M. and Colin, R. (1994). SO_2 Absorption Cross-Section Measurement in the UV Using a Fourier-Transform Spectrometer. *J. Geophys. Res.* 99: 25599–25605, doi: 10.1029/94JD02187.

Voigt, S., Orphal, J. and Burrows, J.P. (2002). The Temperature and Pressure Dependence of the Absorption Cross-Sections of NO_2 in the 250–800nm Region Measured by Fourier-Transform Spectroscopy. *J. Photochem. Photobiol., A* 149: 1–7.

Voigt, S., Orphal, J., Bogumil, K. and Burrows, J.P. (2001). The Temperature Dependence (203–293K) of the Absorption Cross Sections of O_3 in the 230-850 nm Region Measured by Fourier-Transform Spectroscopy. *J. Photochem. Photobiol., A* 143: 1–9.

Wang, H.K., Fu, L.X., Zhou, Y., Du, X. and Ge, W.H. (2010). Trends in Vehicular Emissions in China's Mega Cities from 1995 to 2005. *Environ. Pollut.* 158: 394–400.

Wang, W.S., Gong, D.Y., Zhou, Z.Y. and Guo, Y.X. (2012). Robustness of the Aerosol Weekly Cycle over Southeastern China. *Atmos. Environ.* 61: 409–418.

Wang, Y., Liu, S., Shi, P., Li, Y., Mu, C. and Du, K. (2013). Temporal Variation of Mass Absorption Efficiency of Black Carbon at Urban and Suburban Locations, *Aerosol Air Qual. Res.* 13: 275–286.

Wang, Y., Wan, Q., Meng, W., Liao, F., Tan, H. and Zhang, R. (2011). Long-term Impacts of Aerosols on Precipitation and Lightning over the Pearl River Delta Megacity Area in China. *Atmos. Chem. Phys.* 11:12421–12436.

Wu, D., Tie, X., Li, C., Ying, Z., Kai-Hon Lau, A., Huang, J., Deng, X. and Bi, X. (2005). An Extremely Low Visibility Event over the Guangzhou Region: A Case Study. *Atmos. Environ.* 39: 6568–6577.

Wu, S.P., Wang, X.H., Yan, J.M., Zhang, M.M. and Hong, H.S. (2010). Diurnal Variations of Particle-bound PAHs at a Traffic Site in Xiamen, China. *Aerosol Air Qual. Res.*10: 497–506.

Wu, S.P., Yang, B.Y., Wang, X.H., Hong, H.S. and Yuan, C.S. (2012). Diurnal Variation of Nitrated Polycyclic Aromatic Hydrocarbons in PM_{10} at a Roadside Site in Xiamen, China. *J. Environ. Sci.-China* 24: 1767–1776.

Xiamen Meteorological Bureau (2011). Xiamen Climate Bulletin, Website: http://www.xmeteo.xm.fj.cn/mh/. Accessed on 2013-1-15.

Yang, D.X., Wang, Z.Q. and Zhang, R.J. (2008). Estimating Air Quality Impacts of Elevated Point Source Emissions in Chongqing, China. *Aerosol Air Qual. Res.* 8: 279–294.

Yang, S.J., He, H.P., Lu, S.L., Chen, D. and Zhu, J.X. (2008). Quantification of Crop Residue Burning in the Field and Its Influence on Ambient Air Quality in Suqian, China. *Atmos. Environ.* 42: 1961–1969.

Zhao,Y., Wang, S.X., Duan, L., Lei, Y., Cao, P.F. and Hao, J.M. (2008). Primary Air Pollutant Emissions of Coal-fired Power Plants in China: Current Status and Future Prediction. *Atmos. Environ.* 42: 8442–8452.

Zhuang, M.Z. (2007). Chemical Characteristics of Atmospheric Aerosol in Xiamen. *J. Grad. School Chin. Acad. Sci.* 24: 657–660.

Received for review, May 15, 2013
Accepted, October 30, 2013