Chemosphere 63 (2006) 1261–1272

CHEMOSPHERE

www.elsevier.com/locate/chemosphere

# Seasonal variation of air pollution index: Hong Kong case study

Xie-Kang Wang [a], Wei-Zhen Lu [b,*]

[a] *State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, PR China*
[b] *Department of Building and Construction, City University of Hong Kong, Kowloon Tong, Hong Kong*

Received 17 May 2005; received in revised form 5 October 2005; accepted 17 October 2005
Available online 1 December 2005

**Abstract**

Air pollution is an important and popular topic in Hong Kong as concerns have been raised about the health impacts caused by vehicle exhausts in recent years. In Hong Kong, sulphur dioxide $SO_2$, nitrogen dioxide ($NO_2$), nitric oxide (NO), carbon monoxide (CO), and respirable suspended particulates (RSP) are major air pollutants caused by the dominant usage of diesel fuel by goods vehicles and buses. These major pollutants and the related secondary pollutant, e.g., ozone ($O_3$), become and impose harmful impact on human health in Hong Kong area after the northern shifting of major industries to Mainland China. The air pollution index (API), a referential parameter describing air pollution levels, provides information to enhance the public awareness of air pollutions in time series since 1995. In this study, the varying trends of API and the levels of related air pollutants are analyzed based on the database monitored at a selected roadside air quality monitoring station, i.e., Causeway Bay, during 1999–2003. Firstly, the original measured pollutant data and the resultant APIs are analyzed statistically in different time series including daily, monthly, seasonal patterns. It is found that the daily mean APIs in seasonal period can be regarded as stationary time series. Secondly, the auto-regressive moving average (ARMA) method, implemented by Box–Jenkins model, is used to forecast the API time series in different seasonal specifications. The performance evaluations of the adopted models are also carried out and discussed according to Bayesian information criteria (BIC) and root mean square error (RMSE). The results indicate that the ARMA model can provide reliable, satisfactory predictions for the problem interested and is expecting to be an alternative tool for practical assessment and justification.


*Keywords:* Air pollutant index; Auto-regressive moving average; Bayesian information criteria; Classification; Root mean square error; Time series

## 1. Introduction

As the most predominant source of air pollution in urban area, air pollutants from vehicle emissions received more attention than ever before with the continuous increase of vehicle demand world widely in recent decades (Chovin, 1967; Jacobs, 1974; Kent and Mudford, 1979; Black et al., 1985; Williams, 1987; USEPA, 1991a,b; Kenneth, 1994; Larsolov, 1994; Jorgensen, 1996; Bradley et al., 1999; Singer and Harley, 2000; Ye et al., 2000; Charron and Harrison, 2003; Schifter et al., 2003). Hong Kong faces similar problem and

* Corresponding author. Tel.: +852 27844316; fax: +852 27887612.
*E-mail address:* bcwzlu@cityu.edu.hk (W.-Z. Lu).

0045-6535/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chemosphere.2005.10.031

**Nomenclature**

| | |
|---|---|
| HKSAR | Hong Kong Special Administrative Region |
| CityU | City University of Hong Kong |
| HKEPD | Hong Kong Environmental Protection Department |
| HKAQO | Hong Kong Air Quality Objective |
| API | air pollution index |
| ARMA | auto-regressive moving average |
| AR | auto-regressive |
| MA | moving average |
| ACF | autocorrelation function |
| PACF | partial autocorrelation function |
| AIC | Akaike information criteria |
| BIC | Bayesian information criteria |
| MSE | mean square error |
| RMSE | root mean square error |
| CO | carbon monoxide |
| $O_3$ | ozone |
| NO | nitric oxide |
| $NO_2$ | nitrogen dioxides |
| $NO_x$ | nitrogen oxides |
| $N_2O$ | nitrous oxide |
| RSP | respirable suspended particulate |
| $SO_2$ | sulphur dioxide |

has, unexpectedly, the highest population density in the world (roughly 6000 persons/km$^2$). With continuous economy development and population increase, a series of severe problems relating to the environmental protection and sustainable development has addressed much concern than ever before, in particulate, the air pollution resulted from vehicle emission, which has direct impact on human health and city image. According to the air quality records in Hong Kong (HKEPD, 1999–2003), the main pollutant sources came from the vehicle emissions during past decade since the northern shifting of major industry to Mainland China in 1980s. The reports from Transport Department (1994–2000) indicate that, the total vehicle number and vehicle mileage were recorded as 385342 and 21.88 million kilometer in 1991, increased by 462410 and 25.6 million km in 1994, and reached to 516358 and 28 million km in 2000; the percentages of respirable suspended particulate (RSP) and $NO_x$ emissions from vehicle to the total corresponding emissions were increased from 40% and 22% in 1993 to 58% and 37% in 2000 respectively. In addition, the percentages of carbon monoxide (CO) and volatile organic compounds (VOC) emitted from vehicles have been retained about 89.3% and 92.5% of corresponding items during 1997–2000. These pollutants have, in varying degrees, harmful effect and/or potential danger to human health by direct inhalation or other ways of infection (Calvert, 1984; WHO, 1987; Hewitt and Sturges, 1993; Dockery and Pope, 1994; Pope et al., 1995; Peters et al., 1996; Lu et al., 2002a,b; Lu et al., 2003a,b; Wang et al., 2003a,b; Lu and Wang, 2004).

In general, health impact of air pollution results from a combination of the concentration of air pollutants and the period of time one is exposed to the pollutants. For example, the eye and throat irritation are the most frequent effects of $O_3$ exposure, occurring on 16 and 17 days per capita each year, respectively (Hall, 1996); a number of clinical studies have focused on multiday exposure (100–800 μg/m$^3$) to ozone, which shows that, during repeated daily exposures to ozone, lung function decrement increases after the first exposures, followed by decrease on subsequent exposures (Hackney et al., 1977; Horvath et al., 1981; Bedi et al., 1989; Christian et al., 1998). On the other hand, Ostro (1994) estimated from a review of dose–response relationships for PM10 and indicated that a 10 μg/m$^3$ change in PM10 concentration was associated with a 1% change immortality. The UK Department of Health Committee on the Medical Effects of Air Pollution (Department of Health, 1998) concluded that there are +0.75% per 10 μg/m$^3$ PM10 (24-h mean) for deaths (all causes) and +0.80% per 10 μg/m$^3$ PM10 (24-h mean) for acute respiratory hospital admittances, and a dose–response relationship of 2.5% per 50 μg/m$^3$ for $NO_2$. Rooney et al. (1998) obtained about 190 excess deaths associated with ozone exposure and 175 associated with PM10 during a 5-day photochemical episode in midsummer 1995. Stedman (2004) believed that there were between 423 and 769 excess deaths in England and Wales during the first two weeks of August 2003 associated with the elevated ambient ozone and PM10 concentrations. To provide timely information of air pollution to public, the air pollution index (API) is stipulated/reported by Hong Kong Environmental Protection Department (HKEPD) since 1995. In Hong Kong case, the API is converted from the data of five types of pollutants by certain weighting systems and ranges from 0 to 500. Similar systems can be found in other places like USA, Singapore, Malaysia, Philippines, and Taiwan region. The purpose of API index is to help citizen understand how local air quality is and change in time series. Therefore, the general API is more relevant to us as it represents the air pollution level, which we shall be exposed to for most of the time. The roadside API mainly aims at the air pollution degree by the close proximity to vehicle emission sources, which way be worse to those spending several hours each day close to busy roads. The general API level at or below 50 means that all pollutant levels are in the satisfactory range over 24-h period, however, air pollution consistently at 'High' levels (API of 51–100) in a year implies that the annual Hong Kong Air Quality Objective (HKAQO) for protecting long-term health effects could be violated. An API level exceeding 100 means that levels of one or more pollutant(s) is/are within the unhealthy range. In this study, the variation of API time series and the corresponding air pollutant concentrations are analyzed and discussed at an urban roadside station in Hong Kong. The study would focus on the prediction of daily mean API time series using the auto-regressive moving average (ARMA) method (Box and Jenkins, 1976; Box et al., 1994; Gareth and Louise, 1993; Spyros and Michele, 1997) and the performance assessment of the adopted models. The API forecast may serve as an alert to the public before the onset of serious air pollution episodes. It helps the public, especially susceptible groups such as those with heart or respiratory illnesses, to consider taking precautionary actions if necessary.

## 2. Materials

### 2.1. Sample location and available data

The data measured at Causeway Bay roadside air quality monitoring station (Fig. 1), established since January 1998, are selected as samples to analyze the variation of major pollutants from vehicle emission and API index. The monitoring station is set up at the height of 3 m above ground inside a busy commercial area surrounded by many high-rise buildings. The available database includes respirable suspended particulate (RSP), sulphur dioxide ($SO_2$), nitrogen dioxide ($NO_2$), and carbon monoxide (CO) for the period of 1999–2002, and daily API indeces from July 1999 to October 2003, provided by Hong Kong Environmental Protection Department (HKEPD, 1999–2003). According to the original database, $SO_2$ was observed by UV fluorescence (TECO Model 43A, Monitor Laboratories 8850), $NO_2$ by Chemiluminescence analyzer (API 200A, Monitor Laboratories 8840), CO by Non-dispersive infra-red absorption with gas filter correlation (TECO Model 48, 48C), and RSP by Gravimetric or Oscillating microbalance (Graseby Andersen PM10 R&P TEOM Series 1400a-PM10) respectively. The resultant air pollutants contain the hourly mean concentrations. The hourly APIs are calculated by comparing these concentrations with the corresponding air quality objectives (AQOs) established under the Air Pollution Control Ordinance shown in Table 1. Concerning the air quality on roadside, only four air pollutants mentioned are considered and the APIs for each of above four pollutants are calculated by certain weighting methods. The highest API value is reported as the API of the relevant hour, then, the maximum, the minimum, and the mean API values of that day can also be obtained through analyzing all APIs of 24-h period.

### 2.2. Variations of main air pollutant levels

Four highest hourly levels and two highest daily levels of main pollutants monitored at Causeway Bay

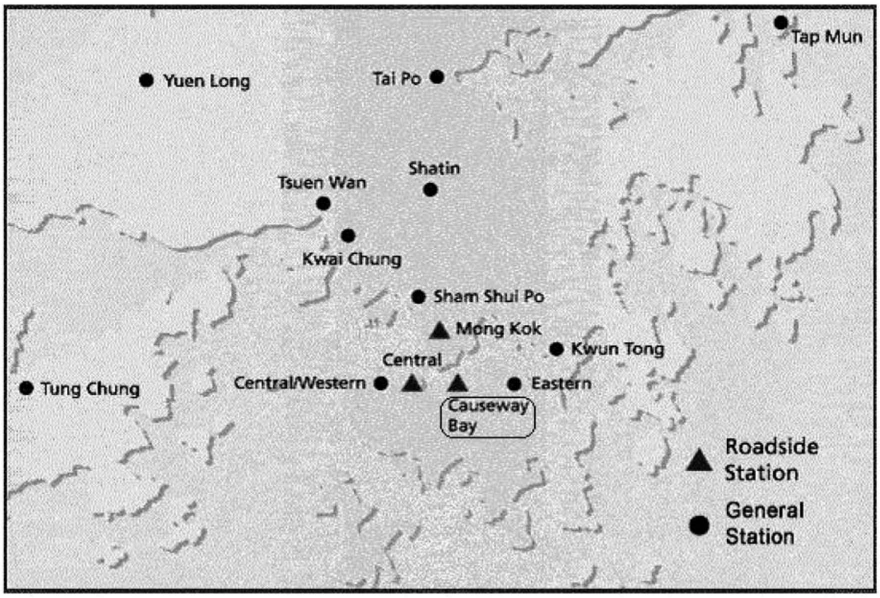


Fig. 1. Location of Causeway Bay Air monitoring station and others in Hong Kong.

Table 1
API sub-index levels and their corresponding air pollutant concentrations

| API sub-index | Relationship with the AQO | Concentration (μg/m$^3$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | RSP 24-h | $SO_2$ 24-h | $SO_2$ 1-h | $NO_2$ 24-h | $NO_2$ 1-h | CO 8-h | CO 1-h | $O_3$ 1-h |
| 0 | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 | 50% of annual AQO or 25% of short-term AQO | 27.5 | 40 | 200 | 40 | 75 | 2500 | 7500 | 60 |
| 50 | Annual AQO or 50% of short-term AQO | 55 | 80 | 400 | 80 | 150 | 5000 | 15000 | 120 |
| 100 | Short-term AQO | 180 | 350 | 800 | 150 | 300 | 10000 | 30000 | 240 |
| 200 | – | 350 | 800 | 1600 | 280 | 1130 | 17000 | 60000 | 400 |
| 300 | – | 420 | 1600 | 2400 | 565 | 2260 | 34000 | 90000 | 800 |
| 400 | – | 500 | 2100 | 3200 | 750 | 3000 | 46000 | 120000 | 1000 |
| 500 | – | 600 | 2620 | 4000 | 940 | 3750 | 57000 | 150000 | 1200 |

Table 2
Highest pollutant levels measured during 1999–2002

| Pollutant | Years | Four highest hourly levels | | | | Two highest daily levels | |
|---|---|---|---|---|---|---|---|
| | | 1st high | 2nd high | 3rd high | 4th high | 1st high | 2nd high |
| $SO_2$ | 1999 | 202 | 154 | 148 | 147 | 90 | 69 |
| | 2000 | 186 | 173 | 148 | 135 | 68 | 62 |
| | 2001 | 151 | 150 | 150 | 147 | 76 | 72 |
| | 2002 | 238 | 224 | 223 | 209 | 71 | 68 |
| $NO_2$ | 1999 | 335 | 332 | 331 | 321 | 209 | 207 |
| | 2000 | 374 | 295 | 290 | 280 | 213 | 185 |
| | 2001 | 300 | 293 | 293 | 293 | 197 | 195 |
| | 2002 | 283 | 278 | 278 | 271 | 208 | 198 |
| CO | 1999 | 5290 | 5180 | 5060 | 4950 | 4789 | 4731 |
| | 2000 | 4140 | 4140 | 4030 | 3910 | 3525 | 3453 |
| | 2001 | 4950 | 4600 | 4490 | 4370 | 3623 | 3594 |
| | 2002 | 4950 | 4950 | 4830 | 4830 | 3680 | 3665 |
| RSP | 1999 | 302 | 297 | 288 | 285 | 226 | 209 |
| | 2000 | 329 | 312 | 282 | 279 | 191 | 190 |
| | 2001 | 275 | 273 | 273 | 267 | 182 | 178 |
| | 2002 | 247 | 240 | 235 | 234 | 172 | 154 |

Note: All concentration units are in μg/m$^3$. 1 h-AQOs for $SO_2$, $NO_2$ and CO are 800, 300, 30000, 8 h-AQO is 10000 for CO; 24 h-AQOs are 350, 150, 180 for $SO_2$, $NO_2$ and RSP respectively.

during 1999–2002 (Table 2). Table shows that $SO_2$ and CO levels are below the relevant 1-h, 8-h or 24-h HKAQOs, however, the violations of 1-h and 24-h HKAQOs are recorded for $NO_2$ during 1999–2002 and RSP during 2000–2001. The percentile properties of main air pollutants are calculated and listed in Table 3. Over viewing the different percentage levels (from 10% to 95%), the percentiles of hourly $NO_2$, CO, $SO_2$ and RSP change in different varieties during 1999–2002.

Fig. 2 shows statistically averaging 24-h variations of major pollutants, i.e., $SO_2$, $NO_2$, CO and RSP, at Causeway Bay monitoring station during 1999–2002, which are used as examples to specify the typical hourly pollutant levels during 24-h period. The figure indicates that $SO_2$, $NO_2$, CO and RSP levels generally present three changing phases, i.e., the early morning phase (00:00–5:00 am) for low pollution levels, the daytime phase (6:00 am–18:00 pm) with increasing pollution levels in general, and the evening phase (18:00 pm–00:00) during which all pollutant levels present descending trends. Besides, the variations of $NO_2$, RSP and CO concentrations shown in Fig. 2 almost follow the same diurnal pattern.

Fig. 3 presents the averaging monthly variations of $SO_2$, $NO_2$, CO and RSP during the period of 1999–2002. It can be seen that, the varying patterns of $NO_2$ are almost the same during the said period, i.e., low $NO_2$ levels during summer (June, July and August)

Table 3
Hourly statistics of main air pollutants during 1999–2002

| Pollutant | Year | Hours | Data capture rate % | Percentiles | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | 10 | 25 | 50 | 75 | 90 | 95 |
| $SO_2$ | 1999 | 8533 | 97.4 | 9 | 14 | 21 | 31 | 44 | 57 |
| | 2000 | 8546 | 97.6 | 15 | 19 | 24 | 31 | 43 | 57 |
| | 2001 | 8586 | 98.0 | 8 | 10 | 14 | 20 | 34 | 46 |
| | 2002 | 7333 | 83.7 | 6 | 7 | 10 | 17 | 34 | 54 |
| $NO_2$ | 1999 | 8482 | 96.8 | 53 | 71 | 102 | 133 | 158 | 174 |
| | 2000 | 8511 | 97.2 | 56 | 72 | 96 | 119 | 140 | 154 |
| | 2001 | 8595 | 98.1 | 62 | 81 | 104 | 127 | 147 | 162 |
| | 2002 | 7335 | 83.7 | 52 | 68 | 89 | 116 | 143 | 161 |
| CO | 1999 | 8487 | 96.9 | 800 | 1030 | 1380 | 1730 | 2180 | 2530 |
| | 2000 | 8549 | 97.6 | 920 | 1150 | 1490 | 1840 | 2180 | 2410 |
| | 2001 | 8439 | 96.3 | 920 | 1150 | 1380 | 1730 | 2070 | 2300 |
| | 2002 | 7329 | 83.7 | 800 | 1030 | 1270 | 1610 | 1960 | 2180 |
| RSP | 1999 | 8483 | 96.8 | 53 | 74 | 103 | 133 | 159 | 178 |
| | 2000 | 8436 | 96.3 | 48 | 69 | 98 | 129 | 156 | 172 |
| | 2001 | 8364 | 95.5 | 46 | 67 | 94 | 122 | 151 | 170 |
| | 2002 | 8477 | 96.8 | 36 | 56 | 79 | 100 | 123 | 138 |

Note: All concentration units are in μg/m$^3$.

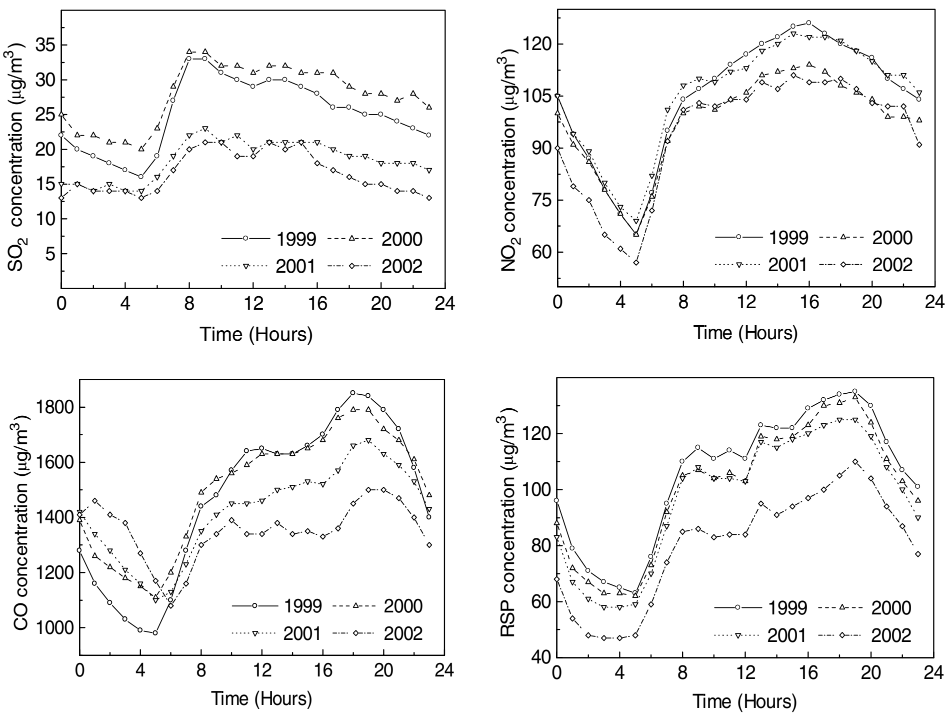


Fig. 2. Average diurnal variations of major air pollutant levels in Causeway Bay during 1999–2002.

and high in other months; the monthly RSP levels in 2002 are generally lower than those in other 3 years and, further, the RSP concentrations in spring (March, April, May) and autumn (September, October, November) are higher than that in summer (June, July, August) and winter (December, January, February); the

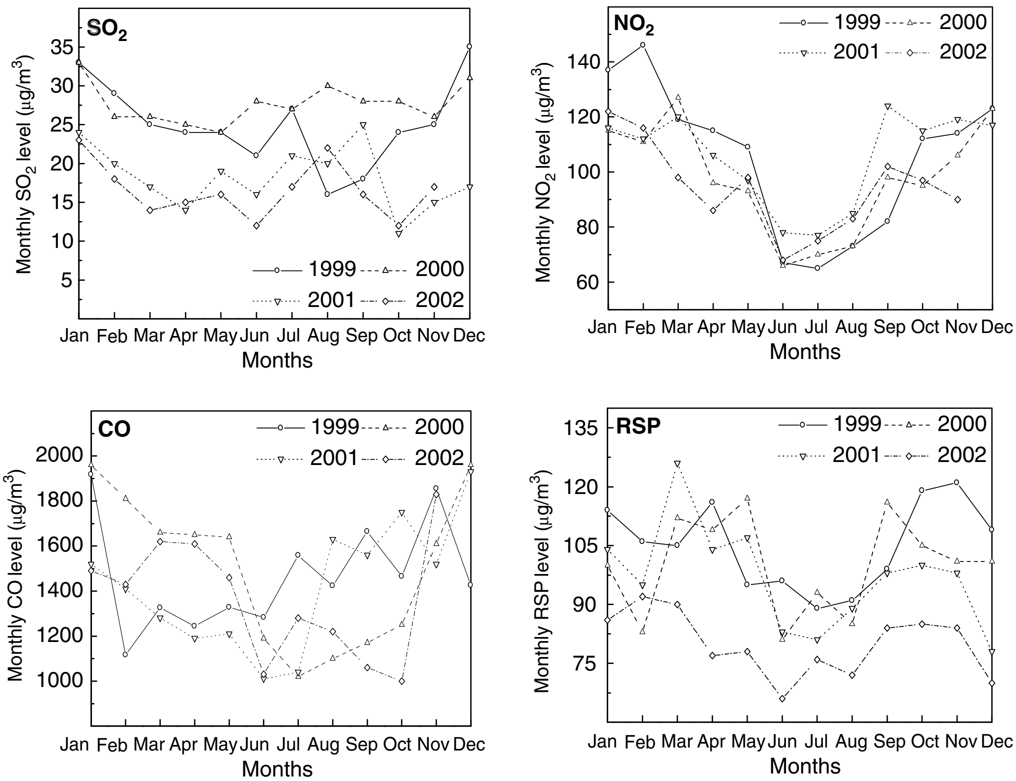


Fig. 3. Monthly variations of main air pollutant levels during 1999–2002.

concentration variations of $SO_2$ and CO appear randomly for the period of 1999–2002 but with descending trends generally.

### 2.3. Variations of air pollution index

Based on the available database of major air pollutants during July 1999–October 2003, the relevant parameters of air pollution index (API) in Causeway Bay area can be obtained through statistical analysis. The results are shown in Figs. 4–10. Fig. 4 describes the correlation analyses among daily mean API $\sim$ maximum API, daily mean API $\sim$ minimum API. The profiles in the figure implies that good correlations exist between daily mean API and the maximum, the minimum APIs with correlation coefficients of $R^2 = 0.9143$ and $R^2 = 0.9303$ respectively.

The alterations of daily mean and maximum APIs based on monthly averaging periods and corresponding standard deviations are depicted in Figs. 5 and 6. It is observed that the monthly variations of APIs demonstrate "V" shape curves, which, again, indicate the low

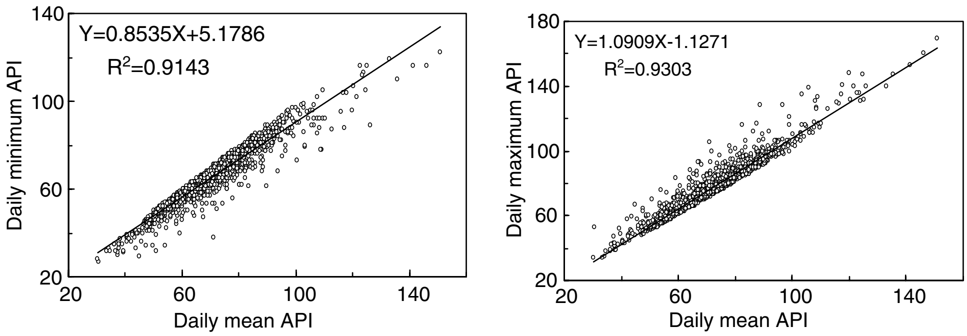


Fig. 4. Correlations between daily maximum, minimum and mean APIs 1999–2003.

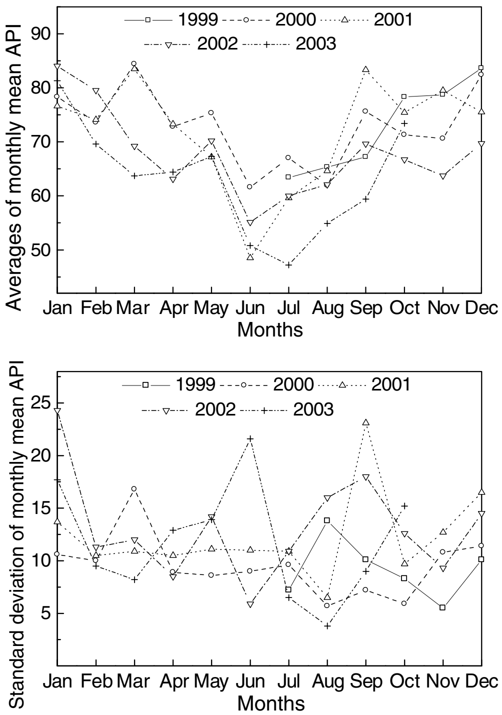


Fig. 5. Averages and standard deviations of monthly mean API during 1999–2003.

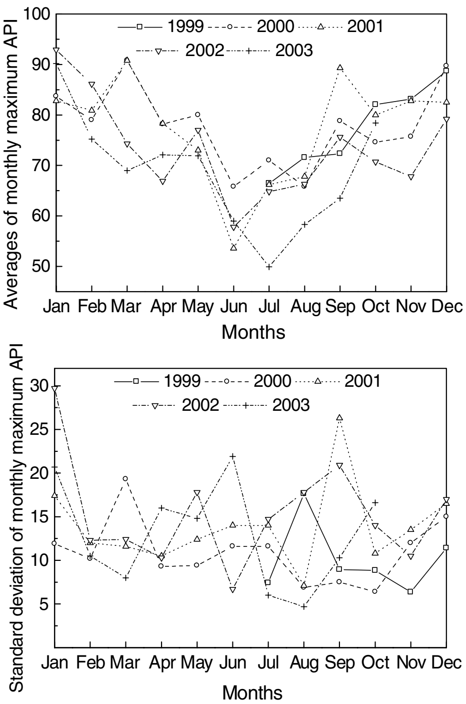


Fig. 6. Averages and standard deviations of monthly maximum API during 1999–2003.

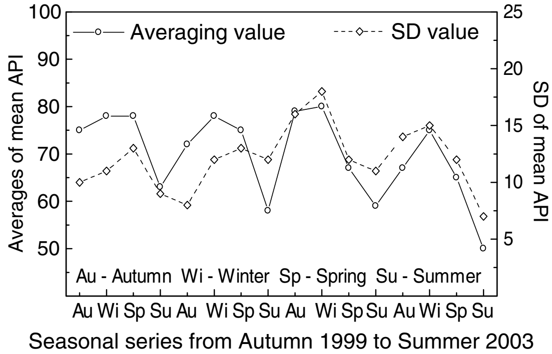


Fig. 7. Averages and standard deviations of mean API during 1999–2003.

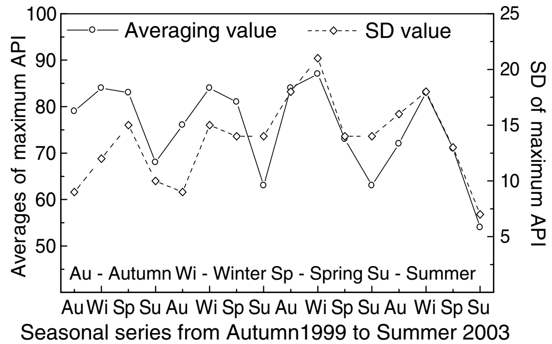


Fig. 8. Averages and standard deviations of maximum API during 1999–2003.

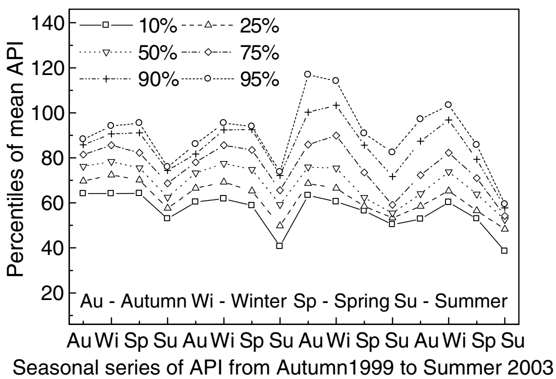


Fig. 9. Percentile variations of seasonal mean APIs during 1999–2003.

pollution levels in summer and high levels on both sides. The average standard deviations for both mean API and maximum API are about 10.

The seasonal variations of APIs and the relevant statistical properties from Autumn 1999 to Summer 2003

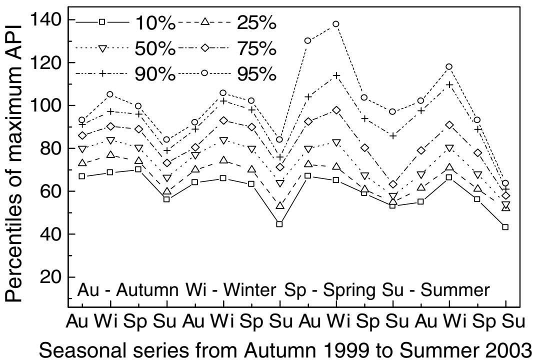


Fig. 10. Percentile variations of seasonal maximum APIs during 1999–2003.

are shown in Figs. 7–10. It is noticed that both mean and maximum APIs in seasonal series present periodically changes with lower APIs in summer and higher ones in other seasons, especially winter. Such phenomena are mainly affected by Hong Kong weather cycle, i.e., heavy rainfall and southeast wind dominant during the summer, dry and cool air and northwest wind leading in the winter. The figures also indicate that the air pollution situation in year 2001 show more severe than other years during the studied period.

## 3. Models and results

### 3.1. Classification and samples of API data

Based on the profiles in Figs. 5–8, the API indeces in summers during the studied period are obviously different from those in other seasons. Hence, it would be better to classify the API prediction in time series into two categories, i.e., the summer periods during 2000–2003 and other seasons from 1999 to 2003. In the predictions, the summer API data (data length = 184) in 2000 and 2001 were specified as training set for model training. The trained model is then used to forecast the summer API data for the period of June–August in 2002 (data length = 92). On the other hand, the model for predicting API time series in other seasons contains the training data from autumn 1999 to spring 2002 (data length = 820) and is used to predict the API levels from September 2001 to May 2002 (data length = 273).

### 3.2. Selected model and prediction results

In the simulation model, the data of the daily mean APIs in all seasons are regarded as a stationary time series because of the time-independence of the statistical properties such as the mean, the standard deviation, the percentile, etc. It means that auto-regressive and moving average (ARMA) method can be used to simulate such parameters. The general mathematical expression is given below:

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \cdots + \phi_p(y_{t-p} - \mu) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}, \qquad (1)$$

where $y$ is the time series variable, $\phi_i$ and $\theta_j$ are the $i$th and the $j$th order of auto-regressive (AR) and moving average (MA) parameters respectively, $\mu$ is the mean value of the time series studied, $e_t$ is the term of white noise.

Generally speaking, the order of the ARMA model can be found by examining the decay trends of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the stationary series. Figs. 11 and 12 present the profiles of ACF and PACF for both summer period and other season periods during the period of interest. However, the ACF and the PACF values do not always provide a clear indication of the suitability of the selected ARMA model, e.g., the decay properties shown in Figs. 11 and 12 do not possess obvious tail-off patterns within less lag number.

Considering the calculation method of API (i.e., a general index based on five air pollutants, namely, sulphur dioxide ($SO_2$), nitrogen dioxide ($NO_2$), nitric oxide (NO), carbon monoxide (CO), and respirable suspended particulates (RSP)), the feasibility and reliability of common neural network (NN) models, there is a limitation to predict the API trend by using anyone of these NN

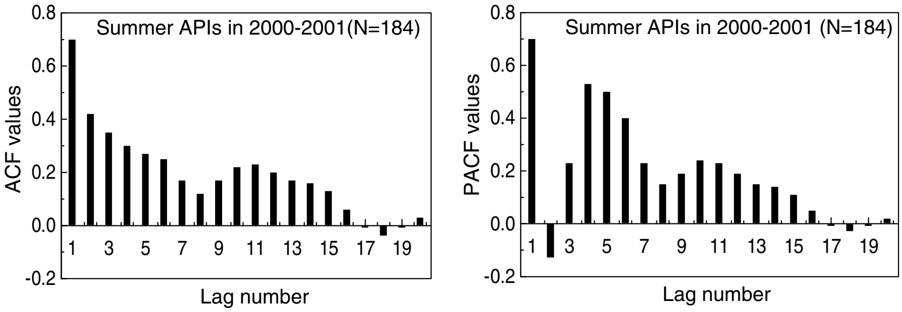


Fig. 11. ACF and PACF values of daily mean APIs in summer during 2000–2001.

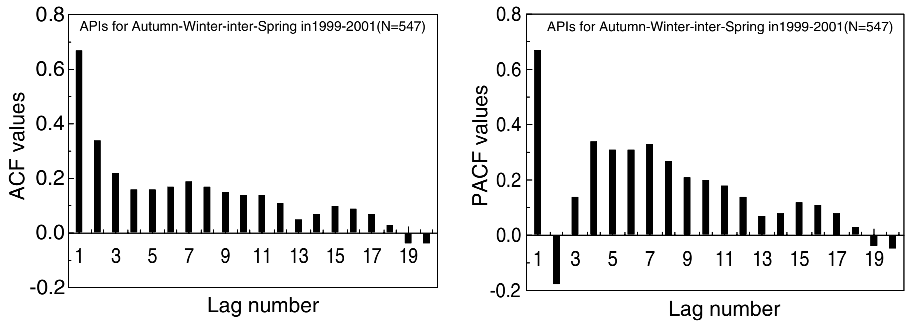


Fig. 12. ACF and PACF values of daily mean APIs in other seasons during 1999–2001.

Table 4
Testing results of selected ARMA models for both time series

| Structure | Coefficient values and RMSE | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Summer API model | | | | Other seasons API model | | | |
| | $\phi$ | $\theta$ | BIC | RMSE | $\phi$ | $\theta$ | BIC | RMSE |
| AR(1) | 0.697 | | 1721 | 0.901 | 0.666 | | 5808 | 0.663 |
| AR(2) | 0.7833, −0.1285 | | 1722 | 0.877 | 0.7889, −0.1839 | | 5795 | 0.670 |
| AR(3) | | | | | 0.815, −0.2966, 0.1434 | | 5791 | 0.655 |
| AR(4) | | | | | 0.8187, −0.3026, 0.1605, −0.0230 | | 5796 | 0.658 |
| ARMA(1,1) | 0.50 | −0.383 | 1704 | 0.872 | 0.458 | −0.382 | 5780 | 0.662 |
| ARMA(1,2) | 0.845 | 0.014, 0.361 | 1705 | 0.893 | 0.730 | −0.0848, 0.2607 | 5775 | 0.675 |
| ARMA(2,1) | 0.1434, 0.2935 | −0.725 | 1712 | 0.962 | 0.1993, 0.2027 | −0.629 | 5783 | 0.678 |
| ARMA(2,2) | 0.6911, 0.0877 | −0.1404, 0.3136 | 1710 | 0.884 | 1.1525, −0.2243 | 0.3383, 0.3861 | 5773 | 0.653 |

models although such models are used for air pollution forecast in parallel (Lu et al., 2002b, 2003a,b, 2004; Lu and Wang, 2005; Wang et al., 2003a,b). In simulations reported here, totally eight ARMA models with different orders of model parameters were used to analyze the cases with two time series, i.e., summer period and other seasons mentioned above. The model parameters including the ARMA coefficients denoted with $p$ and $q$ order are estimated according to the Box–Jenkins method (Box and Jenkins, 1976; Box et al., 1994; Gareth and Louise, 1993; Spyros and Michele, 1997). The relevant parameters used in prediction are listed in Table 4. To further assist the identification of suitable ARMA model, two general information criteria are available for justification, i.e., Akaike information criteria (AIC) (Akaike, 1974) and Bayesian information criteria (BIC) (Sawa, 1978). Considering that the BIC more emphasizes on the parsimony of the model than the AIC does (Christian and Chrisian, 2002), the BIC criteria is used in the study listed as below:

$$\mathrm{BIC} = N\log(\mathrm{MSE}) + (p+q+1)\log N, \tag{2}$$

where MSE is the mean square error, $N$ is number of training sample (184 and 547 samples are adopted for both time series in the simulations respectively).

For evaluating the forecasting capability of ARMA models, the root mean square error (RMSE) can be defined as follows:

$$\mathrm{RMSE} = \frac{1}{M}\sqrt{\sum_{t=1}^{M}(y_t-\hat{y}_t)^2}. \tag{3}$$

Here, $M$ is the number of the forecasting samples and adopts 92 for summer period and 273 for other seasons respectively in the simulations. The $y_t$ and $\hat{y}_t$ represent the actual value and the forecast value for time $t$ respectively. The resultant RMSE values of the relevant ARMA models are shown in Table 4.

According to BIC criteria, the smallest the BIC value, the best the performance of the ARMA model does. Hence, the ARMA(1,1) and ARMA(2,2) shown in Table 4 are selected to forecast the daily mean APIs in summer and other seasons respectively. The forecasting evaluations in Table 4 indicate that the RSME corresponding to the selected model also possesses the smallest value among all experimental models. The comparisons between observations and predicted daily mean APIs produced by the corresponding ARMA(1,1) and ARMA(2,2) models are shown in Fig. 13. The predictions comply well with the relevant observations. The results

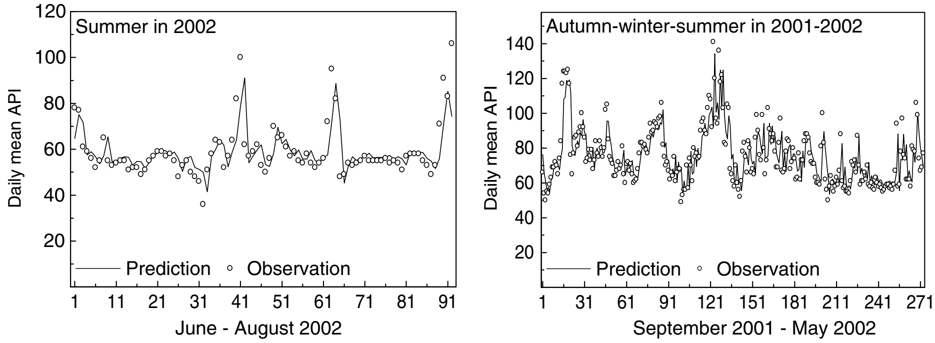


Fig. 13. Comparisons between predictions and observations of daily mean API for two time series.

prove that both ARMA(1,1) and ARMA(2,2) models can provide reliable, satisfactory predictions for both time series. The ARMA method may be an alternative tool for analyzing similar problems in different time series.

## 4. Conclusion

A detailed study on variations of major air pollutants and daily air pollution index (API) in Causeway Bay area during the period of 1999–2003 is reported in this paper. Based on the statistical analyses, the diurnal variations of $SO_2$, $NO_2$, CO and RSP levels three basic phases, i.e., the early morning phase (00:00–5:00 am) with low pollution levels, the daytime phase (6:00 am–18:00 pm) with increasing pollution levels, and the evening phase (18:00 pm–0:00) with descending pollution trends. The monthly varying processes of main pollutants present different patterns during the studied period but generally with lower levels in summer and higher levels in other seasons. Concerning the variations of daily API time series, the daily APIs can be regarded as ststionary time series because the statistical parameters such as the mean, the standard deviations and the percentile are independent of the time. Therefore, the auto-regressive and moving average (ARMA) method can be used as a cost-effective toll to forecast the API trends in different time series. In this study, the most suitable ARMA models for summer period and other seasons are ARMA(1,1) and ARMA(2,2) respectively. Both models can produce reliable and satisfactory results comparing with the corresponding observations.

## Acknowledgements

The work was partially supported by research grants from the Research Grants Council of the Hong Kong Special Administrative Region (HKSAR), China [Projects No. CityU 1185/05E and CityU 1253/04E]. The provision of original data from HKEPD is also appreciated.

## References

Akaike, H., 1974. A new look at statistical model identification. IEEE Transactions on Automated Circuits 19, 716–723.

Bedi, J.F., Horvath, S.M., Drechsler-Parks, D.M., 1989. Adaptation of older individuals repeatedly exposed to 0.45 ppm ozone for two hours. Journal of the Air Pollution Control Association 39, 194–199.

Black, F., Braddock, J., Bradow, R., Ingalls, M., 1985. Highway motor vehicles as sources of atmospheric particles: Projected Trends 1977 to 2000. Environment International 11 (2–4), 205–233.

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control, Second ed. Holden-Day, San Francisco CA.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 1994. Time Series Analysis: Forecasting and Control, third ed. Prentice-Hall, Inc., A Paramount Communications Company, Upper Saddle River, New Jersey 07458, USA.

Bradley, K.S., Stedman, D.H., Bishop, G.A., 1999. A global inventory of carbon monoxide emissions from motor vehicles. Chemosphere: Global Change Science 1, 65–72.

Calvert, J.G., 1984. $SO_2$, NO and $NO_2$ Oxidation Mechanism. In: John, I.T. (Ed.), Atmospheric Considerations, Acid Precipitation Series, Vol. 3. Butterworth publishers, Boston, London Sydney Wellington Durban Toronto.

Charron, A., Harrison, R.M., 2003. Primary particle formation from vehicle emissions during exhaust dilution in the roadside atmosphere. Atmospheric Environment 37 (29), 4109–4119.

Chovin, P., 1967. Carbon monoxide: analysis of exhaust gas investigation in Paris. Environmental Research 1, 198–216.

Christian, D.L., Chen, L.L., Scannell, C.H., Ferrando, R.E., Welch, B.S., Balmes, J.R., 1998. Ozone-induced inflammation is attenuated with multiday exposure. American Journal of Respiratory Critical Care Medicine 158, 532–537.

Christian, H., Chrisian, H., 2002. Index forecasting and model selection. International Journal of Intelligent Systems in Accounting, Finance and Management 11, 119–135.

Department of Health, 1998. Quantification of the Health Effects of Air Pollution in the United Kingdom Committee on the Medical Effects of Air Pollution. The Stationary Office, London.

Dockery, D.W., Pope, C.A., 1994. Acute respiratory effects of particulate air pollution. Annual Revision Public Health 15, 107–132.

Gareth, J., Louise, S., 1993. Time Series: forecasting, simulation, applications. Ellis Horwood Limited, Market Cross House, Cooper Street, Chichester, West Sussex, PO19 1EB, England.

Hackney, J.D., Linn, W.S., Mohler, J.G., Collier, C.R., 1977. Adaptation to short-term respiratory effects of ozone in men exposed repeatedly. Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology 43 (1), 82–85.

Hall, J.V., 1996. Assessing health effects of air pollution. Atmospheric Environment 30 (5), 143–746.

Hewitt, C.N., Sturges, W.T., 1993. Global Atmospheric Chemical Change, Environmental Management Series. In: Cairns, J., Harrison, R.M. (Eds.). Elsevier Applied Science, London and New York.

Hong Kong Environmental Protection Department, 1999–2003. Air Quality in Hong Kong, The Government of the Hong Kong Special Administrative Region.

Horvath, S.M., Gliner, J.A., Folinsbee, L.J., 1981. Adaptation to ozone: duration of effect. American Review of Respiratory Disease 123, 496–499.

Jacobs, P.A., 1974. A random measure model for the emission of pollutants by vehicles on a highway. Stochastic Processes and their Applications 2 (2), 163–176.

Jorgensen, K., 1996. Emissions from light and medium goods vehicle in Denmark. The Science of the Total Environment 189/190, 131–138.

Kenneth, T.K., 1994. On-road vehicle emissions: US studies. The Science of the Total Environment 146/147, 209–215.

Kent, J.H., Mudford, N.R., 1979. Motor vehicle emissions and fuel consumption modeling. Transportation Research Part A 13 (6), 395–406.

Larsolov, O., 1994. Motor vehicle air pollution control in Sweden. The Science of The Total Environment 146–147, 27–34.

Lu, W.Z., Wang, X.K., Wang, W.J., Leung, A.Y.T., Yuen, K.K., 2002a. A preliminary study of ozone trend and its impact on environment in Hong Kong. Environment International 28 (6), 503–512.

Lu, W.Z., Wang, W.J., Fan, H.Y., Leung, A.Y.T., Xu, Z.B., Lo, S.M., Wong, J.C.K., 2002b. Prediction of pollutant levels in Causeway Bay area in Hong Kong using an improved neural network model. ASCE Journal of Environmental Engineering 128 (12), 1146–1157.

Lu, W.Z., Fan, H.Y., Lo, S.M., 2003a. Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. NeuroComputing 51 (1), 387–400.

Lu, W.Z., Wang, W.J., Wang, X.K., et al., 2003b. Using Improved Neural Network Model to Analyze RSP, $NO_x$ and $NO_2$ Levels in Urban Air in Mong Kok, Hong Kong. Environmental Monitoring and Assessment 87, 235–254.

Lu, W.Z., Wang, X.K., 2004. Interaction patterns of major air pollutants in Hong Kong territory. The Science of the Total Environment 324, 247–259.

Lu, W.Z., Wang, W.J., Wang, X.K., Yan, S.H., Lam, J.C., 2004. Potential assessment of a neural network model with PCA/RBF approach for forecasting pollutant trends in Mong Kok urban air, Hong Kong. Environmental Research 96, 79–87.

Lu, W.Z., Wang, W.J., 2005. Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends. Chemosphere 59, 693–701.

Ostro, B., 1994. Estimating health effects of air pollution: a method with an application to Jakarta, Working Paper 1301, Policy Research Department, World Bank, Washington, DC.

Peters, J., Hedley, A.J., Wong, C.M., 1996. Effects of an ambient air pollution intervention and environmental tobacco smoke on children's respiratory health in Hong Kong. International Journal of Epidemiology 25, 821–828.

Pope, C.A., Thun, M.J., Namboodira, M., Dockery, D.W., 1995. Particulate air pollution as a predictor of mortality in a prospective study of US adults. American Journal of Respiratory Critical Care Medicine 151, 669–674.

Rooney, C., McMichael, J., Kovats, R.S., Coleman, M.P., 1998. Excess mortality in England and Wales, and in Greater London, during the 1995 heatwave. Journal of Epidemiology and Community Health 52, 482–486.

Sawa, T., 1978. Information criteria for discriminating among alternative regression models. Econometrica 46, 1273–1291.

Schifter, I., Díaz, L., Vera, M., Guzmán, E., López-Salinas, E., 2003. Impact of sulfur-in-gasoline on motor vehicle emissions in the metropolitan area of Mexico City. Fuel 82 (13), 1605–1612.

Singer, B.C., Harley, R.A., 2000. A fuel-based inventory of motor vehicle exhaust emissions in the Los Angeles area during summer 1997. Atmospheric Environment 34, 1783–1795.

Spyros, M., Michele, H., 1997. ARMA models and the Box–Jenkins methodology. Journal of Forecasting 16, 147–163.

Stedman, J.R., 2004. The predicted number of air pollution related deaths in the UK during the August 2003 heatwave. Atmospheric Environment 38, 1087–1090.

Transport Department, 1994–2000. Annual transport digest 1994, The Government of the Hong Kong Special Administrative Region.

USEPA, 1991a. United States Environmental Protection Agency, Draft 1991 Transportation Air Quality Planning Guidelines.

USEPA, 1991b. United States Environmental Protection Agency, National Air Policy and Emissions Trends Report.

Wang, W.J., Lu, W.Z., Wang, X.K., Leung, A.Y.T., 2003a. Prediction of maximum daily ozone level using combined neural network and statistical characteristics. Environment International 29, 555–562.

Wang, X.K., Lu, W.Z., Wang, W.J., Leung, A.Y.T., 2003b. A study of ozone variation trend within area of affecting human health in Hong Kong. Chemosphere 52 (9), 1405–1410.

WHO, 1987. Nitrogen Dioxide? Air quality guidelines for Europe, WHO Regional publications, European series No. 23, World Health Organization Regional Office for Europe, Copenhagen, 297–314.

Williams, M.L., 1987. The impact of motor vehicles on air pollutant emissions and air quality in the UK? An overview. The Science of The Total Environment 59, 47–61.

Ye, S.H., Zhou, W., Song, J., Peng, B.C., Yuan, D., Lu, Y.M., Qi, P.P., 2000. Toxicity and health effects of vehicle emissions in Shanghai. Atmospheric Environment 34 (3), 419–429.