PERGAMON

Atmospheric Environment 34 (2000) 4073–4084

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

An intervention analysis of air quality data at Santiago, Chile

Héctor Jorquera[a,*], Wilfredo Palma[b], José Tapia[b]

[a]Departamento de Ingeniería Química y Bioprocesos, P. Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago 6904411, Chile
[b]Departamento de Estadística, Facultad de Matemáticas, P. Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago 6904411, Chile

Received 18 October 1999; accepted 28 February 2000

Abstract

Air quality data at Santiago, Chile (PM_{10}, $PM_{2.5}$ and ozone) from 1989 to 1998 are analyzed with the goal of estimating trends in and impacts of public policies on air quality levels. Those policies, in effect since the late 1980s, have been essentially aimed at PM_{10} pollution abatement. The analyses show that fall and winter air quality has been improving consistently, specially the $PM_{2.5}$ levels. The estimated trends for the monthly averages of PM_{10} concentrations range from -1.5 to -3.3% per annum, whereas the trends for monthly averages of $PM_{2.5}$ concentrations range from -5 to -7% per annum. The monthly averages of ground ozone daily maxima do not have a significant trend for two of the downtown monitor sites; at the other three monitoring sites (including the one with the highest impacts) there is a clear downward trend between -5 and -3% per annum. The seasonal averages of a declimatized ozone production rate show a downward trend from 1988 through 1995, and no additional improvements have occurred thereafter. These mixed results for ground ozone levels are ascribed to a shift in the magnitude and spatial distribution of emissions in the city, and so there is a need for additional ozone abatement policies and further research on air pollution abatement options.

Keywords: Air pollution trend; Ground ozone levels; Ttime-series models; Parameter estimation

1. Introduction

One of the goals in air quality management is to quantitatively assess the impacts of different initiatives upon air quality levels in a given area. To achieve this objective, it is necessary to discriminate among long-term (climate- and policy-related), seasonal and short-term (weather-related) components in the environmental data. In order to detect trends in data, researchers have used tools like seasonal decomposition, regression analysis and intervention analysis (Roch and Pellerin, 1982; Young et al., 1991; Xu et al., 1996, Milanchus et al., 1998).

Xu et al. (1996) analyzed the data taken at several sites in southern Ontario, Canada, from 1980 to 1990 to estimate trends in ground ozone levels. A linear model for the log of the ozone daily maximum as the response variable was fitted. The explanatory variables were a trend term, a seasonal term and three meteorological variables: daily maximum temperature, hours of sunshine and relative humidity at the time of highest ozone. The effect of adding those three input variables was an improvement in the trend estimate. The estimated ozone trends range from –0.7 to 2.9% annually across southern Ontario and they are all statistically significant.

Bloomfield et al. (1996) have analyzed ground ozone levels in the Chicago metropolitan area from 1981 through 1991. The data consisted of daily maximum ozone levels measured at 45 stations for the ozone season (1 April–31 October); the response variable chosen was the median estimate across the network (corrected for missing values). Bloomfield et al. developed a nonlinear model using several input variables: temperature, relative humidity, wind speed, and opaque cloud cover. A seasonal term consisting of annual and semi-annual frequencies was added to improve the model fit. The final model accounts for 80% of the variance in the ozone concentration data, and the 95% confidence interval for the ozone

* Corresponding author. Tel.: + 562-686-442; fax: + 562-686-5803.

E-mail address: jorquera@ing.puc.cl (H. Jorquera).

1352-2310/00/$ - see front matter
PII: S1352-2310(00)00161-8

trend was estimated as $(-10.3, 4.9)\%$ per decade. By contrast, a simpler model that only included a seasonal and trend components explains just 30% of the variance in ozone data and its trend estimate is positive and thus biased according to the full analysis.

Milanchus et al. (1998) have proposed a different approach. They have analyzed the data from several metropolitan areas in the USA for the period 1984–1995 according to the following methodology. The ozone (log scale) and meteorological data are filtered to extract the long-term trend, a seasonal (annual) component and a short-term component. In this way, the relationship between ozone and meteorological variables is studied at all relevant time scales. The results for the different cities showed that solar radiation and specific humidity yielded the highest correlation among their baselines (the sum of seasonal and trend components), whereas temperature and dew point depression had the highest correlation with ozone on the shortterm, weather-related time scale. The outcome of this analysis is meteorologically adjusted ozone time series at which trend changes can be selected for further analyses, like comparing them with emission abatement policies and related initiatives (Smith and Adamski, 1998).

Young et al. (1991) advocate a different approach, one in which seasonal and trend components are decomposed into quasi-orthogonal components (the so-called "component time-series model") by applying forward and backward Kalman filters to the joint trend and seasonal models. The estimation of model parameters and series decomposition is achieved by the recursive, smoothing algorithm based on the state-space model for the whole system. The methodology is clearly more complex than that of the other approaches currently in use in the literature, but it has the potential to explore and extract meaningful information from dynamic environmental data (see e.g. Young et al., 1991; Ng and Yan, 1998; Schlink et al., 1997).

1.1. Declimatizing ground ozone levels

From the above discussion, it is clear that some meteorological input variables must be considered when ground ozone level variations are studied to detect trends in these data. Here we use an approach originally developed to forecasting ground ozone levels (Acuña et al., 1996; Jorquera et al., 1998a), but this method is also suited for intervention analysis, as we shall see below.

The starting point is the atmospheric diffusion equation (Seinfeld and Pandis, 1998)

$$\frac{\partial C}{\partial t} = -V \cdot \nabla C + \nabla \cdot (K \nabla C) + Q + R + S, \tag{1}$$

where C is the concentration, V the wind vector, K the eddy diffusivity, Q the emission rate, R the net generation term and S stands for physical removal processes such as wet deposition; in the above equation a time average is implicit in Eq. (1), so that the above equation is applicable to hourly averages, for instance.

Let us consider a 3D domain that includes the planetary boundary layer; assuming that the dominant horizontal transport is the advection term, the dominant vertical transport is the turbulent dispersion and since ozone is not directly emitted, Eq. (1) becomes

$$\frac{\partial C}{\partial t} = -V_H \cdot \nabla_H C + \frac{\partial}{\partial z}\left\{K_Z \frac{\partial C}{\partial z}\right\} + R + S, \tag{2}$$

where the subscript H means that only the x and y coordinates are included in the advection term. Now, let us consider the continuity equation for the vertical average of ozone concentration (henceforth angular brackets mean vertical averages)

$$\langle C \rangle = \frac{1}{L}\int_0^L C(x, y, z, t)\, \mathrm{d}z, \tag{3}$$

where L is the height of the 3D domain. The integration of Eq. (2), using definition (3), assuming a total reflection at the upper boundary of the domain, and introducing the deposition flux at the ground leads to

$$\frac{\partial \langle C \rangle}{\partial t} = -\langle V_H \rangle \cdot \nabla_H \langle C \rangle - \frac{v_D C_0}{L} + \langle R \rangle + \langle S \rangle, \tag{4}$$

where v_D is the ozone deposition velocity and C_0 the ground ozone concentration. In the derivation of Eq. (4) the vertical average of the horizontal advection term has been approximated by the decoupled, first term on the right-hand side of Eq. (4). The next step would be to provide parameterizations for the terms that appear in Eq. (4). Unfortunately, no upper air measurements have been done at Santiago, so we cannot provide representative values of vertical profiles of wind speed, temperature, relative humidity, ozone and so on. Thus we have decided to modify Eq. (4) to include only ground ozone level information available to us. In doing so, we make the assumption that the ground ozone level (C_0) and the vertical average of ozone are linearly correlated, that is

$$\langle C \rangle = \alpha C_0 + \beta. \tag{5}$$

Now this equation is only valid in an average sense for a given ozone season, because clearly there is a day-to-day meteorological variability that would have an effect upon vertical ozone profiles and thus on the coefficients α and β. Introducing Eq. (5) into Eq. (4) leads to the following equation valid for an air parcel that moves with the average wind speed $\langle V_H \rangle$, that is, in a Lagrangean reference frame

$$\frac{\mathrm{d}C_0}{\mathrm{d}t} = -\frac{v_D C_0}{\alpha L} + \frac{\langle R \rangle + \langle S \rangle}{\alpha} \equiv R_{\mathrm{NET}}\{(C_k), x, y, z, t\}. \tag{6}$$

In the above expression, (C_k) stands for a detailed photochemical mechanism, including all the relevant species participating in ozone generation plus all the removal mechanisms: dry and wet deposition, etc. Integration of Eq. (6) from a time before the early morning rush hour (when both ozone and air temperature are minimum) up to the time when the ozone concentration is highest leads to

$$O_3^{t2} - O_3^{t1} = \int_{t1}^{t2} R_{\mathrm{NET}}\{(C_k), x, y, z, t\}\mathrm{d}t. \tag{7}$$

Air temperature increases monotonously during that time interval, so using the change of variable $q = \mathrm{d}T/\mathrm{d}t$, with q always positive, the above expression can be written as

$$O_3^{\max} - O_3^{\min} = \int_{T1}^{T2} R_{\mathrm{NET}}\{(C_k), x, y, z, t\}\frac{\mathrm{d}T}{q} \equiv R^*(T_2 - T_1). \tag{8}$$

It is possible that in some days the temperature profile is not so monotonous, but from the standpoint of an average representation for an ozone season at Santiago, the use of the rate of heating q is appropriate. The reduced ozone generation rate R^* is a lumped parameter that reflects the specific mechanisms that dominate photochemical production at a given urban site: NO_x or VOC limitation, patterns of precursor emissions, long-range transport, boundary conditions, etc. Clearly, ozone formation depends upon the daily meteorological conditions in a complex way, as has been revealed in the published studies – for instance, see Hubbard and Cobourn (1998). We had to simplify the scope of the analysis (in the absence of detailed information about upper air measurements) and consider R^* only as a *seasonal average* over the rest of the meteorological conditions not being measured, so the above analysis can only be regarded as a first approximation to declimatize ground ozone concentrations.

Hence, a linear model of daily maximum ozone levels against the diurnal rise of air temperature $(T_2 - T_1)$ up to the time when ozone is highest should produce seasonal estimates of ozone production rate that are more insensitive to weather fluctuations and so a trend in seasonal ozone production can be estimated. In practice, the value of T_2 is quite close to the highest air temperature recorded the same day, so henceforth we shall adopt this approximation.

Other studies agree with the basic model structure given by Eq. (8). Ryan (1995) has found that temperature explains most of the variance in ground ozone levels in the Baltimore metropolitan area, and that lower morning temperatures were also negatively correlated with high ozone impacts; the same findings were reported by Hubbard and Cobourn (1998) for the Louisville, Kentucky metropolitan area. Comrie (1997) performed statistical analyses of data from several cities in the USA and he found that daily maximum temperature and daily total sunshine were significant predictands for ozone maxima in all cases; daily total sunshine is a surrogate for UV flux, and it should be strongly correlated with $(T_2 - T_1)$ in Eq. (8).

In order to apply Eq. (8) at a given set of data, we have to take into account the weekly cycle of ozone precursor emissions, so weekends must be differentiated from workdays. We also have information on the occurrence of rainy days, which are rather infrequent during the ozone season at Santiago, and we have to consider them as well. As we will show in Section 3, intervention analysis provides a framework to include those effects within the model equation.

For the case of PM_{10} and $PM_{2.5}$, it is more difficult to establish a model, because of a lack of measurements of critical variables such as mixing height and lapse rate, that trigger winter air pollution episodes at Santiago (Rutllant and Garreaud, 1995). In addition, no reliable emission inventory database is available yet, so we had to rely only upon ambient air concentration data.

In summary, the specific objectives of the paper are to use the intervention analysis methodology to detect trends in

(i) The univariate, daily ambient air concentrations of PM_{10} and $PM_{2.5}$.
(ii) The monthly averages of daily ozone maxima, and
(iii) The seasonal average ozone production rates, as derived from the simplified model (8).

2. Database: Santiago, Chile

A mixture of topography, climate, and economic growth (see Table 1) have turned Santiago, Chile (33°27′S, 70°42′W) into one of the most heavily polluted cities in South America. Since 1989, local authorities have enacted several policies oriented at curbing winter PM_{10} and $PM_{2.5}$ episodes at the city. Most of these policies have been oriented to the regulation of stationary and mobile sources. A brief description of the policies is given in Table 2. Santiago (population, 5.8 million inhabitants) is located on a gentle slope in a geographically confined basin and at its latitude the radiative transfer controls vertical mixing of the air with frequent low thermal inversions caused by subsidence. The climate is semiarid with an average annual rainfall of around 320 mm (Rutllant and Garreaud, 1995), although there is considerable inter-annual variability associated with El Niño and La Niña phenomena. Precipitation occurs mostly during fall and winter seasons (April till September), whereas the ozone season (October till March) is rather dry.

Table 1
Economic and demographic indices

Item	1980	1990	1997	% change 1980–1990	% change 1990–1997
Population	4,313,510	5,132,106	5,831,300	19	14
Work force	1,445,095	1,845,675	2,351,253	28	27
No households	825,710	1,048,615	1,320,607	27	26
Private cars	312,610	373,176	735,572	19	97
Buses	9500	13,000	9000	37	− 31
Taxis	16,500	21,425	53,851	30	151

Table 2
Policies aimed at improving air quality indices

Policy	Description	Beginning of implementation
(A) Permanent measures		
Traffic bans	Cars are banned on workdays according to their license plates' last digit 2 different digits each day).	1987
Mandatory vehicle inspections	Cars, trucks and buses are tested for CO, PM and HC emissions	1989
Introduction of catalytic cars	New cars must be equipped with catalytic converters, and they are not subject to traffic bans whatsoever.	September 1993
Emission standards for stationary sources	Industrial stationary sources: 112 ($mg\,m^{-3}$)	January 1993
	Nonindustrial stationary sources: 56 ($mg\,m^{-3}$)	January 1997
	Large industrial, stationary sources: 56 ($mg\,m^{-3}$)	January 1995
	Residential, heating boilers: 56 ($mg\,m^{-3}$)	January 1997
Domestic wood stoves	Open burning wood stoves, without PM abatement equipment, are not allowed during the whole year.	June 1993
Emission trade for stationary sources	Any source that started operation in 1992 must compensate 100% of its PM_{10} emissions by 31 December 1996.	
(B) Short-term, PM_{10} curbing measures		
Phase I contingency measures	Whenever PM_{10} levels reach 240 ($\mu g\,m^{-3}$), the stationary sources that accumulate 30% of the total emissions are shut down for 24 h. Traffic bans are extended to 40% of the old car fleet (without catalytic converters).	1995
Phase II contingency measures	Whenever PM_{10} levels reach 330 ($\mu g\,m^{-3}$), the stationary sources that accumulate 50% of the total emissions are shut down for 24 h. Traffic bans are extended to 80% of the old car fleet (without catalytic converters).	1995

The air quality network in use in Santiago (known as MACAM, its Spanish acronym) consists of four monitor sites surrounding the Downtown area, and a fifth located in the Eastern part of town (see Fig. 1). At each monitor station hourly averages of gaseous pollutants and 24 h sampling of PM_{10} and $PM_{2.5}$ have been recorded on a regular basis since 1988. In addition, a meteorological station (E, on the West Side of downtown) records surface meteorological data (air temperature, net and global radiation, precipitation and 3D components of wind).

3. Methodology of analysis

We choose intervention analysis for the data analyses because (i) the dynamic nature of data is explicitly

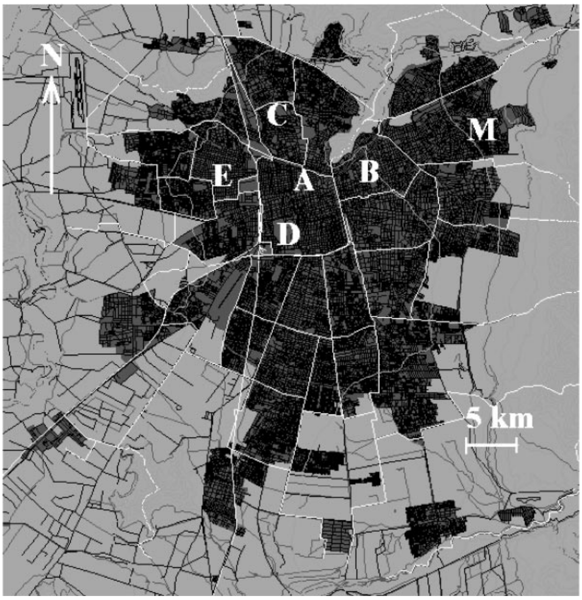

Fig. 1. Map of Santiago, Chile, showing the MACAM air quality monitoring network.

recognized; (ii) the potential impacts are included within the model; and (iii) the presence of outliers is explicitly acknowledged in the model identification process. A description of the technique is given below.

First of all, let us assume that $\{Y_t\}$ is our set of air pollution data, and we want to distinguish the true noise in the data ($\{Z_t\}$) from the different outliers that might also be present within $\{Y_t\}$. The final objective is to identify a model both for the purely random component $\{Z_t\}$ and for the outliers. Since $\{Z_t\}$ is a purely random component, it can be represented by some ARMA model structure, perhaps after differencing the original series to achieve stationary properties (Chatfield, 1996). Following Box and Tiao (1975), the four types of outliers are described below.

Additive outliers (*AO*). An additive outlier is an event that affects a series for one time period only. One example of AO is a recording error, so sometimes the AO are called *gross errors*. If one of these outliers occurs at time $t = T$, the series can be represented by the model

$$Y_t = Z_t + \omega_A P_t(T), \tag{9}$$

where $P_t(T)$ is an indicator function (that is, assumes the value 1 when $t = T$ and is 0 otherwise). The value ω_A stands for the amount of deviation from the true value of Z_t.

Innovational outlier (*IO*). Unlike an additive outlier, an innovational outlier is an event whose effect is propagated according to the noise process model. In this manner an IO affects all values observed after its occurrence. In practice, an IO often represents the onset of an external cause (Tsay, 1988). In time-series notation, the model reads

$$Y_t = \frac{\theta(B)}{\phi(B)}(a_t + \omega_I P_t(T)), \tag{10}$$

where $\theta(B)$ and $\phi(B)$ are the polynomial representation of an ARMA process to model $\{Z_t\}$ (Box and Jenkins, 1970) B is the backward shift operator defined by $BY_t = Y_{t-1}$ and a_t is a white noise sequence. Whereas an AO at $t = T$ only affects Y_T, an IO affects all the values of Y_t for $t \geqslant T$, according to the ψ-weights of the above ARMA process, defined by the identity $\psi(B)\phi(B) = \theta(B)$.

Level shift (*LS*). A level shift (LS) is an event that affects a series at a given time, and whose effect becomes permanent. A level shift could reflect the change of a process mechanism, a change in a recording device or in a measurement instrument or a change in the definition of the variable itself. The model may be represented by

$$Y_t = Z_t + \omega_L S_t(T), \tag{11}$$

where S_t (T) is a step function (i.e. assumes the value 0 before $t = T$ and has the value 1 thereafter). Whenever a level shift is present it affects $\{Y_t\}$ permanently for $t > T$.

Temporary change (*TC*). It is useful to consider an event that has some initial impact on a series and the impact eventually disappears. A temporary (or transient) change (TC) is an event having an initial impact and whose effect decays exponentially according to some dampening factor, say δ. We can represent the observed series as

$$Y_t = Z_t + \frac{1}{1 - \delta B}\omega_C P_t(T), \quad 0 < \delta < 1. \tag{12}$$

Eqs. (9) and (11) are the limiting cases of Eq. (12). In Eq. (9) the dampening factor δ is 0, while in Eq. (11) this factor is 1.

Now we can set the goal of the intervention analysis: to identify the ARMA parameters ($\theta(B)$, $\phi(B)$) that model the disturbance (noise) in the original data and, at the same time, to detect and estimate all relevant outliers occurring in the series $\{Y_t\}$ and their magnitudes $\{\omega_J\}$. This joint estimation method improves the time-series modeling for outliers will be automatically removed along the identification process. Guerard (1989), Harvey (1996) and Sarkar and Kartikeyan (1993) have described applications of this methodology in detail. In the case of our model for the declimatized ozone production rate, effects such as weekend emission reductions or rainy days can be incorporated as additional interventions whose time of occurrence is known, so only their magnitudes $\{\omega_J\}$ need to be estimated.

An additional advantage of this approach is its ability of handling missing values in the data. As pointed out by

Bruce and Martin (1989), the additive outlier (AO) framework can be used to estimate the missing observations in the ARMA context. Since outliers have significant effects in the model parameters, an iterative scheme has to be set. The steps of the algorithm are: (1) outlier detection; (2) outlier adjustment; and (3) parameter estimation, and they are iterated until no further outliers are detected; for additional technical details, see Liu and Chen (1991).

All computations were done using the software package SCA (Liu et al., 1992), which features automatic modeling of seasonal data, joint estimation of parameters and outliers in a model (with automatic missing observation adjustment) and the linear transfer function method. The software uses as threshold criteria the t-value (ratio of parameter estimate and its standard error) greater than 3.0 to classify an observation as being part of an outlier. The specific software version is SCA for Windows, Release 5.2 Professional, running on a Pentium II processor.

4. Results

4.1. Results for $PM_{2.5}$

Fig. 2 shows the monthly averages of $PM_{2.5}$ from 1989 to 1998; all the data show a strong seasonality as expected, and a clear downward trend can be seen at all monitoring stations. In addition, a log transformation of data is a good variance stabilization transformation. The PC-EXPERT module of the SCA software was used to analyze the seasonal data and to obtain sensible parameterizations of the noise process; in all cases it turned out that a MA(1)MA(12) process was adequate to model the noise process. The PC-EXPERT system uses a filtering method to identify seasonal models, when traditional methods (ACF and PACF) do not provide a clear-cut model; for technical details, the reader is referred to the paper by Liu (1986). The results of the seasonal modeling in presence of outliers are shown in Table 3. All data sets were fit to the following model:

$$(1 - B^{12})\log(\mathrm{PM}_{2.5})_t = C + (1 - \theta_1 B) \times (1 - \theta_{12} B^{12}) a_t + N(t), \quad (13)$$

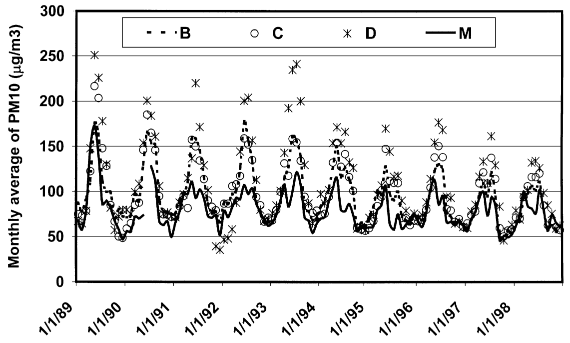

Fig. 2. Monthly averages of $PM_{2.5}$ ($\mu g\ m^{-3}$) at Santiago, Chile, from 1989 till 1998.

where B^{12} is the seasonal extension of the backward shift operator ($B^{12} Y_t = Y_{t-12}$), C is the annual trend, a_t is a white noise process and $N(t)$ is a disturbance term that "stores" all possible outliers according to the four types described by Eqs. (9)–(12). After the model identification step (where a joint outlier and parameter estimation is carried out), ACF and PACF plots for the residuals of the fitting showed no structure at all, thus validating the model given in Eq. (13). According to Table 3, all parameters are significant. Transforming the results for the C parameter back to the original scale (using the log distribution formula) shows that $PM_{2.5}$ levels have evolved annually from -5% at station C to -7% at station B. However, annual averages of $PM_{2.5}$ are still high, reaching about 40 $\mu g\ m^{-3}$ at all stations, above the standards or guidelines proposed in the USA and Canada (between 15 and 20 $\mu g\ m^{-3}$). In order to get better air quality levels at Santiago, the current progress in emission reductions ought to be continued, or additional abatement initiatives should be implemented.

To understand the output of the joint parameter and outlier estimation process, an outlier plot of the data of $PM_{2.5}$ at station D is given in Fig. 3. The continuous line stands for the observed data; the circles represent the adjusted series, that is, a series that follows the model of Eq. (9), but does not have any outlier, and a diamond

Table 3
Modeling of $\log(PM_{2.5})$ data

Station	C	σ_C	t-value	θ_1	$\sigma_{\theta 1}$	t-value	θ_{12}	$\sigma_{\theta S}$	t-value	Outliers detected
B	−0.0743	0.0065	−11.39	−0.2619	0.0892	−2.94	0.7301	0.0683	10.69	None
C	−0.0531	0.0060	−8.92	−0.2963	0.0890	−3.33	0.7780	0.0651	11.96	None
D	−0.0604	0.0067	−9.04	−0.3881	0.0923	−4.21	0.9386	0.0587	15.99	TC at 35, magnitude: −0.779
M	−0.0737	0.0088	−8.40	−0.3535	0.0903	−3.91	0.8332	0.0629	13.25	TC at 110, magnitude: 0.635

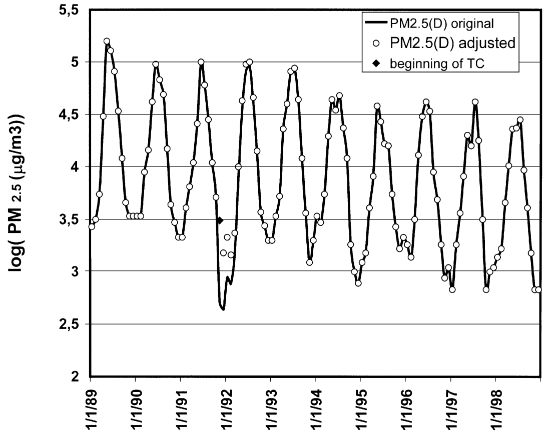

Fig. 3. An outlier plot of the dynamic modeling of $PM_{2.5}$ data at station D.

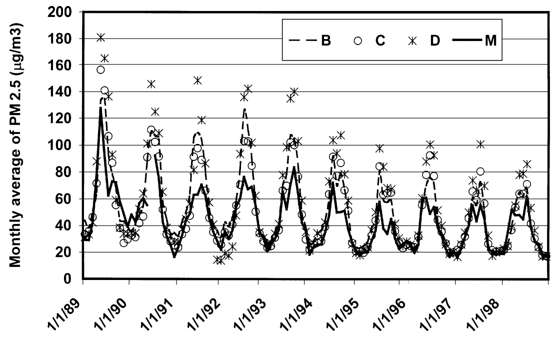

Fig. 4. Monthly averages of PM_{10} ($\mu g\ m^{-3}$) at Santiago, Chile, from 1989 till 1998.

signals the beginning of a temporary change (TC) starting on November 1991 and that lasted about five months. This TC was characterized by unusually low levels at station D, likely a local effect because it was not present at the other stations (see Table 3 and Figs. 2 and 3). The detection of this TC is relevant, because in a simple linear regression of monthly data against time, those low levels would bias the trend estimate for $PM_{2.5}$, a well-known result in ordinary least-squares regression. By removing outliers, the trend estimates are based upon the adjusted series and are thus more reliable.

4.2. Results for PM_{10}

Fig. 4 displays time-series plots of PM_{10} measurements; although a trend is not so clearly appreciated from the figure, all the data sets indeed fit the model described by Eq. (13), and Table 4 gives the resulting parameter estimates. Data for all stations show a decreasing trend in PM_{10} impacts, ranging from -1.5% per annum at station M to -3.3% per annum at station B. Station D possess two significant outliers, one occurring at the same time, the one displayed on its $PM_{2.5}$ data, with unusually low levels measured for the 1991–1992 summer and the other (also of TC type) characterized by high PM_{10} levels in winter 1993 (see Fig. 5). The trend estimate for station M lies outside the range considered being significant at the 95% level, but since there is strong correlation among PM_{10} measurements at the different stations in the network, the -1.5% annual decrease is judged to be relevant anyway. The annual averages of PM_{10} ranged between 76 ($\mu g\ m^{-3}$) at station M to 93 ($\mu g\ m^{-3}$) at station D in 1998. Again, this result shows that additional PM_{10} abatement policies need to be implemented.

It is also clear that coarse particles do not show a downward trend, but rather an increase in the last 10 years. This is due to their emission sources being related to transportation (resuspended street dust, tire and brake wear, etc.) which has steadily increased in the last years; we also have to bear in mind that the city is surrounded by a semi-arid valley with significant sources of wind erosion.

4.3. Results for ozone: monthly average evolution

Fig. 6 shows a plot of the monthly averages of the daily maximum ozone impacts recorded at the different monitors of the MACAM network. It can be seen that the highest impacts are recorded at the eastern site, that is,

Table 4
Modeling of log(PM_{10}) data

Station	C	σ_C	t-value	θ_1	$\sigma_{\theta 1}$	t-value	θ_{12}	$\sigma_{\theta S}$	t-value	Outliers detected
B	-0.0342	0.0059	-5.77	-0.2945	0.0875	-3.36	0.7467	0.0653	11.43	None
C	-0.0222	0.0068	-3.26	-0.3537	0.0846	-4.18	0.7420	0.0606	12.24	None
D	-0.0216	0.0069	-3.14	-0.3747	0.0911	-4.11	0.8341	0.0617	13.51	TC at 35, magnitude: -0.727; TC at 54, magnitude: 0.408
M	-0.0151	0.0089	-1.70	-0.4713	0.0805	-5.85	0.6954	0.0640	10.86	None

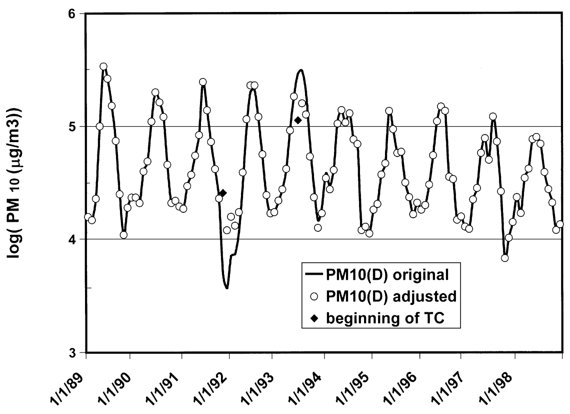

Fig. 5. An outlier plot of the dynamic modeling of PM_{10} data at station D.

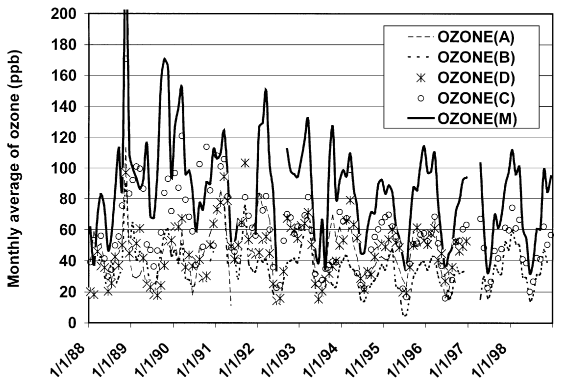

Fig. 6. Monthly averages of daily maximum ground ozone (ppb) at Santiago, Chile, from 1989 till 1998.

downwind of most morning precursor emissions. The site showing the lowest impacts is station B, located between two major traffic lines, so ozone scavenging by fresh NO emissions dominates these lowest measurements. The other stations (A, C and D) display ozone impacts that lie between those two limits. Nevertheless, at stations A, C and D episodic days happen almost every ozone season, although with maxima lower than those recorded at station M.

Table 5 summarizes the results of the ARMA modeling of the monthly ozone data. Trend estimates obtained by a classical linear regression of the log-transformed series were compared with the estimates obtained by means of intervention analysis applied to Eq. (13), and using SCA's PC-EXPERT module to model the noise process. Although the estimated trends are similar, the residuals of the classical linear regression show a clear autocorrelation pattern; thus the variance estimates are underestimated and the t-values overestimated (compare standard errors and t-values in Table 5). This effect is caused because the autocorrelated data possess less information, and so the actual variance estimate ought to be corrected to include autocorrelation (Wilks, 1995). Therefore, the results obtained using ARMA modeling and intervention analysis are more reliable. In summary, we can conclude that stations A and D do not show any trend at all, whereas stations B, C and M display a clear downward trend for the period 1989–1998.

The estimated outliers for the monthly data show no simultaneous behavior across the monitoring sites, so they are ascribed to spatial shifts in patterns of precursor emissions. In terms of the original variable, the results show that there has been a decrease in ozone levels since 1988, but the rate of change has decreased in the last years, and this decrease has not been uniform, with stations A and D showing no trend at all.

Table 5
Modeling of the log(ozone) data

	Classical linear regression			Intervention analysis			
Station	C	σ_C	t-value	C	σ_C	t-value	Outliers detected
A	− 0.0157	0,0157	− 0.10	0.0100	0.0868	0.01	Two AO (at $t = 18$ and 41), one TC at $t = 93$
B	− 0.0362	0.0130	− 2.78	− 0.0516	0.0080	− 6.41	Two TC (at $t = 42$ and 90), one AO at $t = 102$, one LS at $t = 113$, magnitude: 0.308
C	− 0.0498	0.0118	− 4.21	− 0.0554	0.0205	− 2.70	One AO at $t = 32$
D	0.0173	0.0158	1.09	− 0.0052	0.0307	− 0.17	One AO at $t = 45$, one IO at $t = 57$
M	− 0.0338	0.0105	− 3.22	− 0.0268	0.0121	− 2.21	One AO at $t = 68$

4.4. Results for the declimatized ozone production rate

The daily maximum ozone records at station M were used as input for Eq. (8). The air temperatures (maxima and minima) were extracted from the hourly records at the meteorological station E. From both data sets six subsets corresponding to the ozone seasons from 1992 till 1996 were selected for time-series analysis. Some allowance of missing observations was accepted because the method is capable of handling missing observations. Table 6 shows the number of days, missing data, weekend and rainy days in each data subset. The following known interventions were added into Eq. (8):

The weekend emission reduction was modeled as being a binary signal defined by

$$I_1(t) = \begin{cases} 0 & \text{if } t \text{ is a working day,} \\ 1 & \text{if } t \text{ is a weekend.} \end{cases}$$

The rainy day signal was modeled also as a binary signal defined by

$$I_2(t) = \begin{cases} 0 & \text{if no rainfall was recorded on day } t, \\ 1 & \text{if some rainfall was recorded on day } t. \end{cases}$$

In time-series notation, the final model to be identified is given by

$$\mathrm{O}_3^{\max}(t) = A + R\,\Delta T(t) + \sum_{i=0}^{N} \omega_{1i} B^i I_1(t) + \frac{\omega_{\mathrm{P}}}{(1-\delta_{\mathrm{P}} B)} I_2(t) + N(t), \tag{14}$$

where A is a constant term that stands for ozone minima values; R is the seasonal average estimate of the reduced ozone production rate; ΔT is the diurnal rise of air temperature ($T_2 - T_1$); the finite series of powers in B is a polynomial approximation to the transfer function associated with the input intervention signal $I_1(t)$; the transfer function associated to the input signal $I_2(t)$ was modeled as an exponentially damped response; finally, $N(t)$ is an underlying noise that is modeled as an ARMA process and that also includes the outliers identified in each data set. The model identification process is known as the linear transfer function (LTF) method, and specific details are given in Liu and Hanssens (1982), Liu and Hudak (1985) and Pankratz (1991). The first check on the relationship between ozone and ΔT was done using the cross correlation formalism (Liu et al., 1992); this method showed that the relationship was indeed contemporaneous: ozone is related to ΔT measured on the same day and no significant coefficients were obtained for other lags of ΔT.

Table 7 displays the results obtained for the intervention analysis based on Eq. (14). The constant term A is either small (of the order of 15 ppb) or it does not appear in the equation, according to the model prediction for its value (it comes from the minimum ozone ground level recorded on the same day). The structures obtained for the disturbance term $N(t)$ indicate that no differencing of input data was needed (values of the AR parameter are significantly different from 1.0) and so the relationship given by Eq. (14) was a stationary one. Furthermore, the ACF and PACF plots of the model residuals show no structure, hence the model structure has been properly identified, and thus inferences from the fitted results are valid (Milionis and Davies, 1994). The most important parameter is the declimatized, seasonal average ozone production rate. This parameter shows a steady decrease from summer 1992 to fall 1995, but then it shows no clear trend. It can be concluded that photochemical pollution has not improved since 1996, which is consistent with the plot of Fig. 6, the discussion is given in Section 4.3 and the results of a simpler statistic analysis presented elsewhere (Jorquera et al., 1998b).

One issue that remains is how this ozone trend is related to pattern changes in precursor emissions. For instance, the NO concentrations have been showing a steady decrease since 1992, because of the policies detailed in Table 2. However, the composition of volatile organic compounds has not been systematically measured at Santiago, and this composition must have undergone significant changes since the early 1990s because of the introduction of unleaded gasoline in 1993. In the absence of further information, it is not possible to quantify how much of the ozone trend is related to each precursor emission pattern.

Table 6
Data sets for fitting Eq. (14)

Set	Period	No. days	O_3 missing	ΔT missing	Weekend days	Rainy days
1	1/1/92–24/4/92	99	5	2	28	5
2	24/9/92–30/4/93	219	4		68	13
3	18/11/93–30/5/94	194	12	6	57	13
4	1/9/94–30/4/95	242	11	1	72	16
5	1/9/95–30/4/96	243			77	17
6	1/9/96–31/12/96	122			39	9

Table 7
Results of the identification of model Eq. (14)

(A) Constant term, declimatized ozone production rate and noise parameter estimates

Data set	A (ppb)	σ_A (ppb)	t-value	R^* (ppb °C^{-1})	σ_R	t-value	$N(t)$ (ppb)	ϕ_1	σ	t-value
1				11.546	0.624	18.50	AR(1)	0.5564	0.0825	6.74
2				7.764	0.167	46.55	AR(1)	0.3529	0.0651	5.42
3	16.81	7.33	2.29	5.644	0.454	12.44	AR(1)	0.4046	0.0671	6.03
4	12.96	4.15	3.13	5.008	0.242	20.69	White			
5				6.249	0.157	39.92	AR(1)	0.6127	0.0492	12.45
6				5.148	0.139	37.14	AR(1)	0.4456	0.0827	5.39

(B) Transfer functions weights for the weekend emission intervention (ppb)

Data set	ω_{10}	σ_{10}	t-value	ω_{11}	σ_{11}	t-value	ω_{12}	σ_{12}	t-value
1				− 21.40	7.1	− 2.99			
2	− 11.85	3.29	− 3.60	− 10.61	3.3	− 3.22			
3				− 10.30	3.2	− 3.22			
4				− 7.65	2.2	− 3.54	− 5.55	2.2	− 2.58
5				− 13.10	2.4	− 5.54			
6									

(C) Parameters for the rainy day transfer function and significant ouliers detected

Data set	ω_P (ppb)	σ	t-value	δ_P	σ	t-value	Main outliers detected (ppb)
1	− 33.6	14.9	− 2.24	0.6640	0.2616	2.54	One IO
2							Five AO, two IO and one TC
3	− 13.2	5.9	− 2.24	0.6847	0.1872	3.66	Four AO
4	− 10.9	3.2	− 3.38	0.7514	0.0984	7.64	LS = −14.4 at t = 186, three AO, four TC
5							Four AO, one TC, one IO
6							Two AO

Another information about the dynamics of photochemical smog can be obtained from the weights $\{\omega_{1i}\}$ associated with the weekend intervention. It can be seen in Table 7 that weekend emission reductions always cause a decrease in ozone levels, except in data subset 6, where no significant effect was detected. Most ozone reductions show up with a delay of one day, with some exceptions in the data subsets 2 and 4, that display a contemporaneous response and a two-day delay, respectively. These results are consistent with the statistical studies on weekend – workday differences in ground ozone levels reported since the 1970s (Cleveland et al., 1974; Lebron, 1975; Levitt and Chock, 1976, Cleveland and McRae, 1978; Chaum et al., 1978; Karl, 1978; Pryor and Steyn, 1995). The values of the reduction in ozone levels also show a trend, varying from -22 ppb in 1992 to -13 ppb in 1995 and no decrease in 1996, which may be attributed to a shift in magnitude and spatial distribution of precursor emissions in the city, linked to the demographic indices given in Table 2.

The response associated with the rainy day interventions gives additional insight into the mechanisms of ozone wet deposition and urban airshed washout caused by intrusion of cold weather fronts and the related passage of stronger winds. For the three data sets that showed significant effects caused by rainy days, the polynomial approximations to the linear transfer function exhibited coefficients with a slow decay, suggesting a rational approximation to the transfer function model, similar to a TC outlier model. This structure was the one chosen in Eq. (14), and the modeling results are discussed below.

The ω_P parameter is always negative, indicating a decrease in ambient ozone levels due to wet scavenging. The magnitude has been decreasing from 1992 to 1995, that is, a behavior similar to the case of the weekend emission intervention. The δ_P parameter is always positive, indicating a monotonic decrease to zero of the response function after the rain moves away from the city. From the values of Table 7 it turns out that it takes about three days for the impact to fade away, and this is an estimate of the time of reload of precursors into the urban airshed.

Finally, examination of the outliers detected in the model identification process shows that most of them are short-term events: AO that affect the daily maximum ozone at just one day, TC that decay quickly (in a couple of days) and IO that are propagated with the ARMA structure of $N(t)$ (decay in one or two days); all of these can be ascribed to synoptic scale motions, lasting two to three days in average. Regarding long-term trends, the method identified only one-level shift in ground ozone levels: one starting in 5 March 1994, with a permanent decrease of 14.4 ppb. This estimate is highly significant, with a t-value of -7.50 (recall that the software uses as threshold criteria the t-value of 3.0 to classify an observation as being part of an outlier). It seems that this permanent change in ground ozone levels may be ascribed to a shift in the magnitude and spatial distribution of emissions in the city. In fact, since the early 1990s many industries have been moved from the south and west part of the city to the northern side. Transportation sources, though, have been steadily increasing their emissions in the last years, so the net trend is harder to assess, because reliable emission inventory databases are yet to be developed.

5. Conclusions

The results of the analyses for PM_{10} and $PM_{2.5}$ are that policies presented in Table 2 did improve fall and winter air quality levels at Santiago for the period 1989–1998. This improvement was steady, and was essentially achieved by tighter emission standards for stationary and mobile sources, periodic enforcement of the standards and partial shut down of those sources during episodic days. In order to achieve air quality goals similar to those set at other countries worldwide, the current trends must be maintained or even increased. For instance, according to data in Table 2, there are too many taxis on the streets, so a regulation of this part of the fleet seems sensible.

In the case of the ground ozone levels, intervention analysis was applied to the univariate time series of monthly averages and to a linear stochastic model derived from basic principles (see Eq. (14)). In both cases there was an initial decrease in ground levels, but the trend seems to have stopped during the 1996–1998 period; further studies (like VOC ambient monitoring campaigns and updated emission inventories) are required to project future trends of photochemical oxidants.

In summary, unless additional initiatives are enacted, there is a long way to go until air quality levels at Santiago reach levels deemed as satisfactory by international standards. There have been additional measures that started in the last two years, namely introduction of compressed natural gas, cleaning up of dust street, paving roads, and tree seeding along the valley. Since these measures are more recent, they are not expected to show results until a few years ahead, with a new assessment of air pollution abatement strategies scheduled for 2005.

Acknowledgements

Funding for this work came from Grant FONDECYT 198-0972. We are indebted to Mr. Ignacio Olaeta (http://www.sesma.cl) for providing the air quality data. The referees' comments helped us improving the scope and presentation of this paper.

References

Acuña, G., Jorquera, H., Pérez, R., 1996. Neural network model for maximum ozone concentration prediction. In: von der Matsburg, C., von Seelen, W., Vor brüggen, J.C., Sendhoff, B. (Eds.), Lecture Notes in Computer Science, Vol. 1112. Springer, Berlin, pp. 263–268.

Bloomfield, P., Royle, J.A., Steinberg, L.J., Yang, Q., 1996. Accounting for meteorological effects in measuring urban ozone levels and trends. Atmospheric Environment 30 (17), 3067–3077.

Box, G.E.P., Jenkins, G.M., 1970. Time-Series Analysis. Forecasting and Control. Holden-Day, San Francisco.

Box, G.E.P., Tiao, G.C., 1975. Intervention analysis with applications to economic and environmental problems. Journal of American Statistical Association 70, 70–79.

Bruce, A.G., Martin, R.D., 1989. Leave-k-out diagnostics for time series (with discussion). Journal of the Royal Statistical Society Ser. B 51, 363–424.

Chatfield, C., 1996. The Analysis of Time Series — An Introduction, 5th Edition. Chapman & Hall, London.

Chaum, D., Elkus, B., Wilson, K.R., Rice, J.R., 1978. A statistically tested short term oxidant control strategy. Atmospheric Environment 12, 1777–1783.

Cleveland, W.S., Graedel, T.E., Kleiner, B., Warner, J.L., 1974. Sunday and workday variations in photochemical air pollutants in New Jersey. Science 186, 1037–1038.

Cleveland, W.S., McRae, J.E., 1978. Weekday-weekend ozone concentrations in the Northeast United States. Environmental Science and Technology 12, 558–563.

Comrie, A.C., 1997. Comparing neural networks and regression models for ozone forecasting. Journal of Air and Waste Management Association 47, 653–663.

Guerard, J.B., 1989. Automatic time series modelling, intervention analysis, and effective forecasting. Journal Statistical Computation and Simulation 34, 49–53.

Harvey, A., 1996. Intervention analysis with control groups. International Statistical Review 64, 313–328.

Hubbard, M., Cobourn, W.G., 1998. Development of a regression model to forecast ground-level ozone concentration in Louisville, KY. Atmospheric Environment 32 (14/15), 2637–2647.

Jorquera, H., Pérez, R., Cipriano, A., Espejo, A., Letelier, M.V., Acuña, G., 1998a. Forecasting ozone daily maximum levels at Santiago, Chile. Atmospheric Environment 32 (20), 3415–3424.

Jorquera, H., Olguín, C., Ossa, F., Pérez, R., Solar, I., Encalada, O., 1998b. An assessment of photochemical pollution at Santiago, Chile. In: Brebbia, C.A., Ratto, C.F., Power, H. (Eds.), Air Pollution, Vol. VI. Computational Mechanics Publications, Boston, pp. 743–753.

Karl, T.R., 1978. Day of the week variations of photochemical pollutants in the St. Louis area. Atmospheric Environment 12, 1657–1667.

Lebron, F., 1975. A comparison of weekend-weekday ozone and hydrocarbon concentrations in the Baltimore-Washington metropolitan area. Atmospheric Environment 9, 861–863.

Levitt, S.B., Chock, D.P., 1976. Weekday-weekend pollutant studies of the Los Angeles basin. Journal of Air Pollution Control Association 26, 1091–1092.

Liu, L.M., Hanssens, D.M., 1982. Identification of multiple-input transfer function models. Communications in Statistics A 11, 297–314.

Liu, L.M., Hudak, G.B., 1985. Unified econometric model building using simultaneous transfer function equations. In: Time Series Analysis: Theory and Practice, Vol. 7. Elsevier, Amsterdam, pp. 277–288.

Liu, L.M., 1986. Identification of time series models in the presence of calendar variation. International Journal of Forecasting 2, 357–372.

Liu, L.M., Chen, C., 1991. Recent developments of time series analysis in intervention in environmental impact studies. Journal of Environmental Science and Health 26 (10), 1217–1252.

Liu, L.M., Hudak, G.B., Box, G.E.P., Muller, M.E., Tiao, G.C., 1992. Forecasting and Time Series Analysis using the SCA Statistical System. SCA Corp., IL.

Milanchus, M.L., Rao, S.T., Zurbenko, I.G., 1998. Evaluating the effectiveness of ozone management efforts in presence of meteorological variability. Journal of Air and Waste Management Association 48 (3), 201–215.

Milionis, A.E., Davies, T.D., 1994. Regression and stochastic models for air pollution. I. Review, comments and suggestions. Atmospheric Environment 28 (17), 2801–2810.

Ng, C.N., Yan, T.L., 1998. Recursive modelling and adaptive forecasting of air quality data. In: Brebbia, C.A., Ratto, C.F., Power, H. (Eds.), Air Pollution VI. Computational Mechanics Publications, Boston, pp. 869–877.

Pankratz, A., 1991. Forecasting with Dynamic Regression Models. Wiley, New York.

Pryor, S.C., Steyn, D.G., 1995. Hebdomadal and diurnal cycles in ozone time series from the Lower Fraser Valley, B.C. Atmospheric Environment 29 (9), 1007–1019.

Roch, R., Pellerin, J., 1982. On the long term air quality trends and intervention analysis. Atmospheric Environment 16 (2), 161–169.

Rutllant, J., Garreaud, R., 1995. Meteorological air pollution potential for Santiago, Chile: towards an objective episode forecasting. Environmental Monitoring and Assessment 34, 223–244.

Ryan, W.F., 1995. Forecasting severe ozone episodes in the Baltimore metropolitan area. Atmospheric Environment 29 (17), 2387–2398.

Sarkar, A., Kartikeyan, B., 1993. Intervention analysis with nonlinear dependent noise variation. Statistics and Probability Letters 18, 91–103.

Seinfeld, J.H., Pandis, S.N., 1998. Atmospheric Chemistry and Physics. Wiley, New York (Chapter 17).

Schlink, U., Herbarth, O., Tetzlaff, G., 1997. A component time-series model for SO_2 data: forecasting, interpretation and modification. Atmospheric Environment 31 (9), 1285–1295.

Smith, B.E., Adamski, W.J., 1998. Eight-hour ozone trends at sites in Lake Michigan ozone nonattainment areas. Journal of Air and Waste Management Association 48 (12), 1204–1206.

Tsay, R.S., 1988. Outliers, level shifts and variance changes in time series. Journal of Forecasting 7, 1–20.

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press, San Diego.

Xu, D., Yap, D., Taylor, P.A., 1996. Meteorologically adjusted ground level ozone trends in Ontario. Atmospheric Environment 30 (7), 1117–1124.

Young, P.C., Ng, C.N., Lane, K., Parker, D., 1991. Recursive forecasting, smoothing and seasonal adjustment of non-stationary environmental data. Journal of Forecasting 10 (1), 57–89.