PERGAMON

Central and South America

Atmospheric Environment 33 (1999) 2575–2585

ATMOSPHERIC ENVIRONMENT

Statistical distributions for air pollution applied to the study of the particulate problem in Santiago

Benoit Morel[a,*], Sonia Yeh[a], Luis Cifuentes[b]

[a]*Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburg, P 15213-3890, USA*
[b]*Dpto, Ing. Industrial, Universidad Catolica de Chile, Chile*

Received 29 June 1998; accepted 9 October 1998

Abstract

An expression relating the distribution of concentrations of air pollutants to the emission levels is derived and applied to the study of the air pollution problem of Santiago de Chile. Features specific to the area, like the geography and meteorology are included in a stochastic framework. The resulting distribution is simple. It has two parameters which can be easily estimated from the data. It fits well the data from Santiago and is a powerful tool to analyze pollution data and design a mitigation strategy. The methodology developed here can be applied to any city.

Keywords: Stochasticity; Particulate; Probability distribution; Response function; Emission level

1. Introduction

Air pollution is a universal phenomenon. But every area is a special case. The relation between anthropogenic pollutant emissions and air pollutant concentration is stochastic and area dependent. As a result the translation of air quality standards (AQS) into emission requirements is not the same everywhere.

Although "it is largely agreed that there is no a priori reason to expect that atmospheric distribution should adhere to a specific probability distribution" (Seinfeld and Pandis, 1998), this paper seeks to suggest that in some cases, there may be a natural probability distribution which relates emission levels to pollution levels. More precisely, treating air pollution as a stochastic process, we derive a probability distribution for air pollutant concentration. Then we observe that this distribution reproduce naturally data from Santiago on particulates. Since this distribution relates air pollutant concentrations to emission levels, it provides a response function, i.e. a basis to translate AQS into emission targets.

This paper starts with a little excursion in the theory of stochastic processes (Karlin and Taylor, 1981) to derive our expression relating pollutant emission levels to the statistical distribution of air pollutant concentration. The rest of the paper is devoted to the analysis and use of that expression.

We first try to identify where this distributions belongs in the world of univariate continuous distributions (Kotz and Johnson, 1970) and relate it some of the distributions, like the fashionable stable distributions

*Corresponding author. Tel.: +1 412 268 2670; fax: +1 412 268 3757; e-mail: bm1v@andrew.cmu.edu, syeh@andrew.cmu.edu.

1352-2310/99/$ - see front matter
PII: S1352-2310(98)00380-X

(Samorodnitsky and Taqqu, 1994), or to the distributions which are being used in the context of air pollution, namely Log-Normal, Weibull, Gamma, and Beta distributions (Georgeopoulos and Seinfeld, 1982).

We note that our distribution for some values of the parameters, would be indistinguishable from one or the other of these distributions. The major merit of our distribution is the fact it relates emission level to pollution level. That is, if one can assume that our distribution is the correct one, it provides a way to know by how much emission levels have to be abated to meet AQS. This level of validation of our distribution is not accomplished in this paper, but we are working on it. That requires data on emission levels we do not have. The goal of this paper is to suggest the potentials of our distribution as a tool for policy analysis.

We use our distribution to analyze air pollution data collectively daily over several years in Santiago for the following pollutants: particulates (PM_{10} and $PM_{2.5}$), nitrogen dioxide (NO_2) and sulfur dioxide (SO_2).

Our analysis of the data from Santiago with this distribution yields interesting insights on the pollution problem of Santiago, its evolution and how challenging it will be to meet US standards for some pollutants like PM_{10} and $PM_{2.5}$.

2. Method

In this section we derive the expression for the statistical distribution of air pollutant concentrations. It corresponds to the stationary distributions of a stochastic process relating emission level to air pollutant concentrations.

2.1. Air pollution as a stochastic process

The fundamental assumption underlying the derivation of the distribution is that the concentrations of air pollutants are the combined results of pollutant emissions and stochastic effects like meteorological conditions. In mathematical terms, Eq. (1) summarizes our assumptions: it implies that the atmospheric concentration x_i and the level of anthropogenic emission E_i of precursors to pollutant "i" are related dynamically by the stochastic differential equation (Dixit and Pindyck, 1993, chap. 2)

$$\mathrm{d}x_i = \{\delta_i(\mu_i - x_i) + \beta_i E_i\}\,\mathrm{d}t + \eta_i x_i\,\mathrm{d}z_i, \tag{1}$$

where μ_i is what the equilibrium concentration would be in absence of anthropogenic contribution, i.e. if $E_i = 0$.

δ_i measures the rate at which the pollutant concentrations return to equilibrium. It is related to the inverse of the residence time.

β_i is the "loading factor". It measures the amount by which emissions increase the atmospheric concentration. It depends on the weather (e.g. the altitude of the inversion temperature), the geography and location of the sources with respect to the points of measurement.

$\eta_i x_i$ is the infinitesimal variance, assumed to be proportional to the concentration level. In Ito's notation, $\mathrm{d}z_i = \varepsilon_i\sqrt{\mathrm{d}t}$, where ε_i is a random variable normally distributed with mean 0 and variance 1. This implies $\langle \mathrm{d}z_i\rangle = 0$ and $\langle \mathrm{d}z_i^2\rangle = \mathrm{d}t$.

The "drift" $\alpha_i = \delta_i(\mu_i - x_i) + \beta_i E_i = \delta_i(x_i^{\mathrm{eq}} - x_i)$ implies that, in the absence of stochasticity, the concentration of pollutant "i" would be

$$x_i^{\mathrm{eq}} = \mu_i + \frac{\beta_i E_i}{\delta_i}. \tag{2}$$

Since its dynamics is given by a stochastic differential equation, the pollutant concentration x_i has a distribution of values $\varphi(x_i, t, E_i)$. The distribution $\varphi(x_i, t, E_i)$ is solution of the Kolmogorov forward equation (Karlin and Taylor, 1981, Eq. (5.25), p. 219):

$$\frac{\eta_i^2}{2}\frac{\partial^2[x_i^2\varphi(x_i,t,E_i)]}{\partial x_i^2} - \frac{\partial[[\delta_i(\mu_i - x_i) + \beta_i E_i]\varphi(x_i,t,E_i)]}{\partial x_i} = \frac{\partial\varphi(x_i,t,E_i)}{\partial t}. \tag{3}$$

Although emissions change during the day, the same pattern of emission is repeated most days. In a similar way, stochastic effects like weather are unpredictable but statistically they do not change very much from one year to the next. Yearly distributions of pollutant concentrations therefore correspond to the stationary solution $\varphi_\infty(x_i, E_i)$ of the Kolmogorov forward equation (3), namely they are the solution of

$$\frac{\eta_i^2}{2}\frac{\partial^2[x_i^2\varphi_\infty(x_i,E_i)]}{\partial x_i^2} - \frac{\partial[[\delta_i(\mu_i - x_i) + \beta_i E_i]\varphi_\infty(x_i,E_i)]}{\partial x_i} = 0. \tag{4}$$

Eq. (4) characterizes completely our distribution. Eq. (4) can easily be solved analytically by noticing that it is an equality between derivatives. Integrating Eq. (4) once with respect to x_i, introducing $\rho_i = 2\delta_i/\eta_i^2$, plus some trivial algebraic manipulations leads to

$$\frac{\partial\log[x_i^2\varphi_\infty(x_i,E_i)]}{\partial x_i} = \frac{\rho_i(x_i^{\mathrm{eq}} - x_i)}{x_i^2}. \tag{5}$$

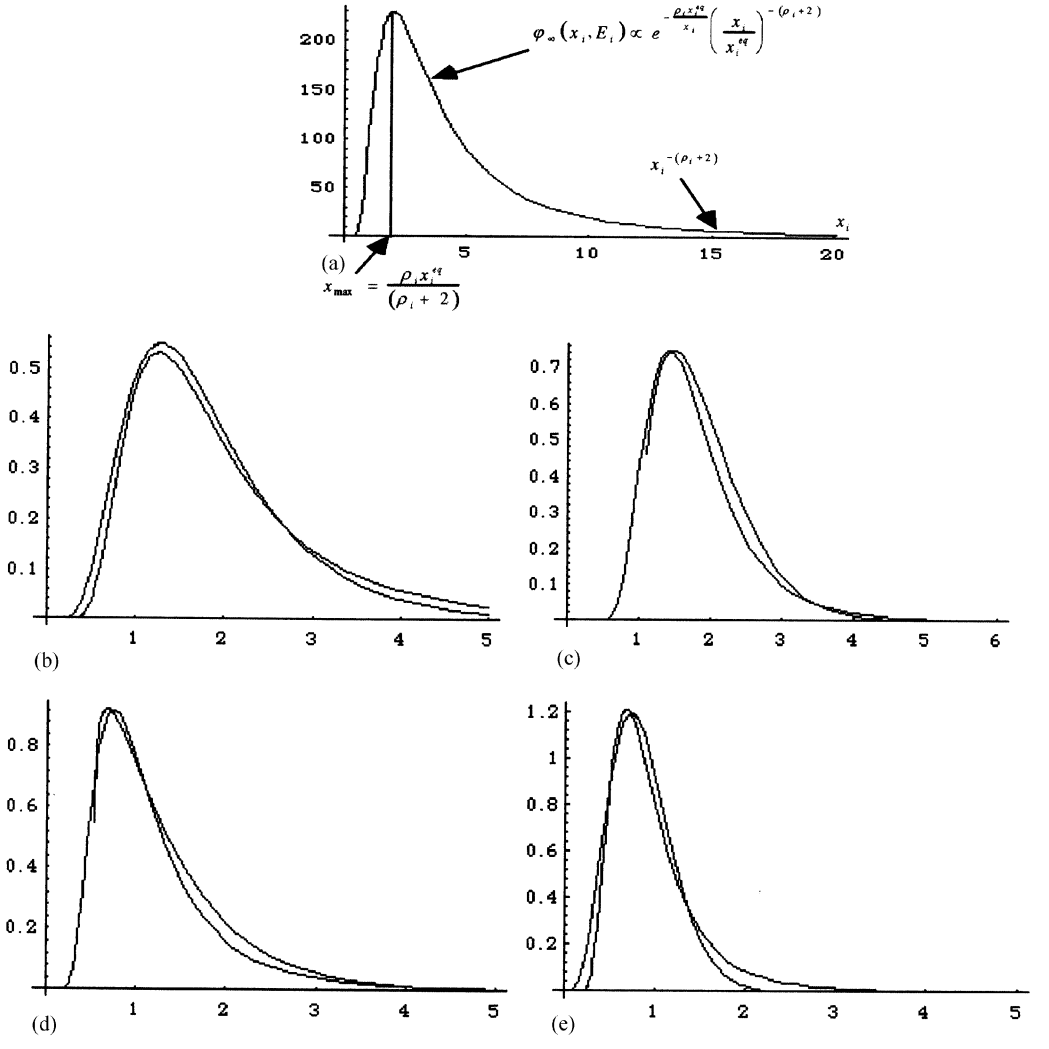

Fig. 1. (a) General shape of $\varphi_\infty(x_i, E_i)$. The position of the peak at $x_{\max}$ and the asymptotic power law behavior for the long tail are used to estimate the data. (b) Comparison of lognormal ($\mu = 0.5$, $\sigma = 0.5$) and $\varphi_\infty(x_i; x_i^{eq} = 2.2, \rho_i = 2.7)$. (c) Weibull: $W(x_i; c = 1.5, \alpha = 1, \xi_0 = 1) = c/\alpha\{x - \xi_0/\alpha\}^{c-1} e^{-\{x-\xi_0/\alpha\}^c}$ with $\varphi_\infty(x_i; x_i^{eq} = 1.2, \rho_i = 7)$. (d) Gamma: $G(x_i; \alpha = 1.2, \beta = 0.7, \gamma = 0.53) = (x-\gamma)^{\alpha-1}e^{-(x-\gamma/\beta)}/\beta^\alpha\Gamma(\alpha)$ with $\varphi_\infty(x_i; x_i^{eq} = 1.27, \rho_i = 2.8)$. (e) Beta distributions: Beta $(p = 5, q = 30) = (1/B(p,q)))x^{p-1}(1-x)^{q-1}$ and $\varphi_\infty(x_i; x_i^{eq} = 1, \rho_i = 4)$.

Eq. (5) can be integrated to yield the expression we will be using abundantly throughout this paper:

$$\varphi_\infty(x_i, E_i) = \frac{x_i^{eq}(\rho_i)^{\rho_i+1}}{\Gamma(\rho_i+1)} e^{-\rho_i x_i^{eq}/x_i} \left(\frac{x_i}{x_i^{eq}}\right)^{-(\rho_i+2)}. \quad (6)$$

2.2. Properties of the $\varphi_\infty(x_i, E_i)$

The general shape of $\varphi_\infty(x_i, E_i)$ is shown in Fig. 1. It has two parameters: $\rho_i = 2\delta_i/\eta_i^2$ and $x_i^{eq} = \mu_i + (\beta_i E_i/\delta_i)$ x_i^{eq} is in fact the average value of the concentration and ρ_i controls the size of the fluctuations. The smaller ρ_i, the larger the statistical fluctuations.

2.2.1. What kind of distribution is $\varphi_\infty(x_i^{eq}, \rho_i)$?

(i) $\varphi_\infty(x_i^{eq}, \rho_i)$ is a special case of a type V Pearson curve (Johnson and Kotz, 1970, Eq. (28), p.12). Pearson curves are by definition, the probability distributions $p(y)$ solutions of

$$\frac{1}{p(y)}\frac{dp(y)}{dy} = \frac{\alpha y + \beta}{y^2 + \gamma y + \delta} \quad (7)$$

They are classified according to the number of real roots of the denominators. Type V corresponds to the case where the denominator is a perfect square. Our distribution corresponds to the case $\gamma = \delta = 0$ (cf. Eq. (5)).

(ii) $\varphi_\infty(x_i^{eq}, \rho_i)$ is related to, but different from α-stable non-Gaussian distributions.

Using Eq. (6), $Pr[x_i \geq u_i]$ can be computed explicitly in terms of an incomplete gamma function

$$Pr[x_i \geq u_i] = \frac{\gamma(1+\rho_i, \rho_i x_i^{eq}/u_i)}{\Gamma(1+\rho_i)}. \tag{8}$$

Asymptotically (i.e. when $u_i \to \infty$), Eq. (8) yields the power law behavior that non-Gaussian or Paretian-stable distributions have (Samorodnitsky and Taqqu, 1994):

$$Pr[x_i \geq u_i] \propto u_i^{-(\rho_i+1)}. \tag{9}$$

To be α-stable non-Gaussian, a distribution should, for some value of the parameters μ, $0 < \alpha \leq 2$, $|\beta| \leq 1$ and σ, be the inverse Fourier transform of the expression (which in fact represents their characteristic function, cf. Samorodnitsky and Taqqu (1994 Eq. (1.2.8), p. 16):

$$\exp\left\{ i\mu\vartheta + \begin{cases} -\sigma^\alpha|\vartheta|^\alpha\left(1 - i\beta\dfrac{|\vartheta|}{\vartheta}\tan\left(\dfrac{\alpha\pi}{2}\right)\right) & \text{if } \alpha \neq 1 \\ -\sigma\left(1 + i\beta\dfrac{|\vartheta|}{\vartheta}\dfrac{2}{\pi}\log|\vartheta|\right) & \text{if } \alpha = 1 \end{cases} \right\}. \tag{10}$$

The parameter α appearing in those distributions corresponds to our parameter ρ_i. For the specific value $\rho_i = \frac{1}{2}$, $\varphi_\infty(x_i)$ is indeed the inverse Fourier transform of the expression Eq. (9), with: $\alpha = \frac{1}{2}$, $\beta = 1$, $\mu = 0$ and $\sigma = x_i^{eq}$. It corresponds to the well-known Lévy distribution.

But the values taken by the parameter ρ_i (cf. Table 1) imply that the air pollutant concentration distributions fall outside of the category of α-stable non-Gaussian distributions.

2.2.2. Relation with Lognormal, Weibull, Gamma and Beta distributions

In Fig. 1b–e examples are provided of lognormal distributions, Weibull, Gamma and Beta distributions which would give a fit of data comparable in quality to the distribution $\varphi_\infty(x_i; x_i^{eq}, \rho_i)$, for corresponding values of the parameters. In other words there are situations where several distributions may fit equally well the data.

We do not have a general criterion to decide which is the best. Collateral information is required. We can simply point to the fact that our distribution express a certain physical situations reflected in the assumptions underlying the stochastic equation from which it derives. When these assumptions reflect the physics of the situation, the distribution $\varphi_\infty(x_i; x_i^{eq}, \rho_i)$ offer more promising prospects, for policy analysis, that distributions which do not provide the same kind of link between emission level and air pollutant concentration.

2.2.3. Moments

ρ_i and x_i^{eq} can be directly estimated from the first two moments of the data distributions. The general form of the moments is

$$\langle x_i \rangle_n = \int_0^\infty (x_i)^n \varphi_\infty(x_i)\,dx_i = \frac{x_i^{eq}(\rho_i x_i^{eq})^{n-1}\Gamma(\rho_i + 1 - n)}{\Gamma(\rho_i)}. \tag{11}$$

Because of the fact that the large x_i^{eq} dependence of $\varphi_\infty(x_i, E_i)$ is like a power law, only the small order moments are well defined. When $n \geq \rho_i + 1$, the moments are infinite as can be seen from Eq. (11). x_i^{eq} corresponds to the first moment, i.e. the average value of the concentration:

$$\langle x_i \rangle_1 = x_i^{eq}. \tag{12}$$

The second parameter ρ_i can be estimated from the second moment with the knowledge of x_i^{eq}:

$$\langle x_i \rangle_2 = (x_i^{eq})^2 \frac{\rho_i}{(\rho_i - 1)}. \tag{13}$$

The estimate of the individual parameters like E_i, μ_i, β_i, δ_i and η_i is a different matter, which we discuss later.

3. Results

3.1. The data

We use for this paper, measurements of PM_{10}, $PM_{2.5}$, NO_2, and SO_2, taken over several years in five different sites in Santiago. We used the daily averages and daily maxima from the five stations. The typical form of the data aggregated over the five stations, are shown for the case of particulates ($PM_{2.5}$ and PM_{10}) in Figs. 2a and b and compared with the US AQS. The histograms of the data aggregated over time are shown in Figs. 3a, a′, b and b′.

The particulates data reveal a few interesting features:

(i) They have a strong seasonal variability (Figs. 2a and b). In the winter, the pollutant concentrations tend to be much higher. The major cause for this difference is not attributable to changes in emission patterns, but to the weather and to the fact that the particulates are trapped in a smaller volume because the inversion temperature is at a much lower level in winter.

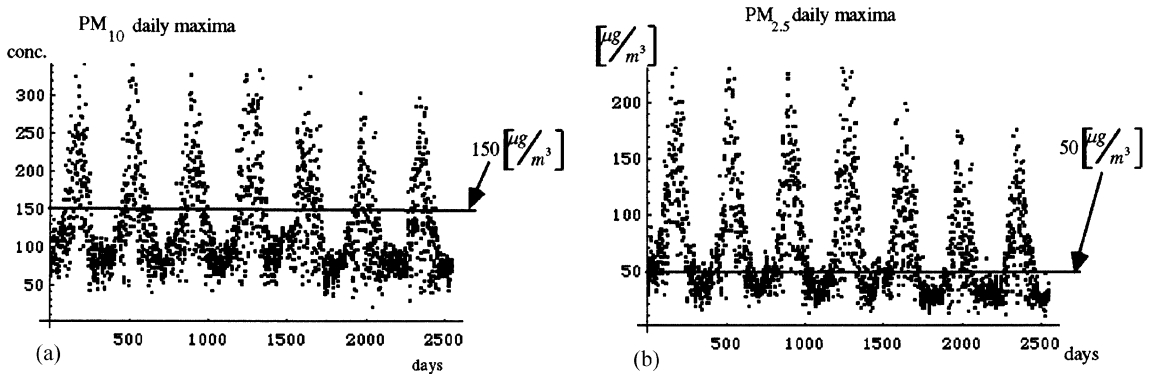

Fig. 2. (a) Daily maxima for PM_{10}: 150 $\mu g\,m^{-3}$ with probability $> 99.7\%$ (i.e. except for at most 24 h in a yr), corresponds to the US standards prior the most recent revision. (b) Daily maxima for $PM_{2.5}$: 50 $\mu g\,m^{-3}$ with a probability 99.7% was the 1996 US standards.

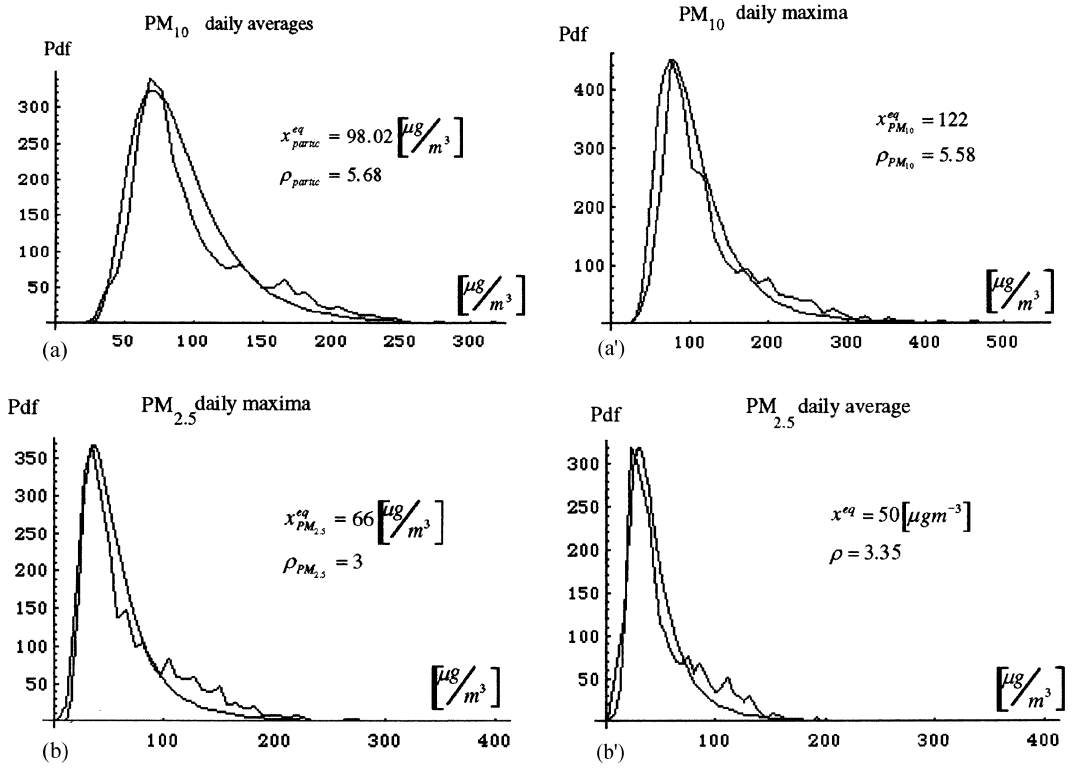

Fig. 3. (a) Fit of the daily average PM_{10} data using Eq. (8). (a′) Fit of the daily maxima PM_{10} data using Eq. (8). (b) Fit of the daily maxima $PM_{2.5}$ data using Eq. (8). (b′) Fit of the daily averages $PM_{2.5}$ data using Eq. (8).

In terms of the physical parameters introduced in this paper, this means that the parameter δ_i which is inversely proportional to the residence time, is smaller in winter and the "loading factor" β_i is larger. As a result, $\rho_i = 2\delta_i/\eta_i^2$ is smaller (the fluctuations of the distributions are larger in winter) and $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$, the average value of the concentration, is significantly larger. In order to appreciate the size of the effect of the seasonal variability we have divided the data in two seasons (summer and winter) and compute the values of the two parameters $\rho_i = 2\delta_i/\eta_i^2$ and $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ for each season for each year separately. The winter and summer

values for the variables of largest policy interest (i.e. $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$) are shown in Figs. 6a, a′; b, b′; c and d. The annual value is the average between the extreme values.

The values of x_i^{eq} shown in Table 1 correspond to the averages over the years of the annual averages.

(ii) The data reveals also some evolution with time. Fig. 6a suggests a reduction over the years of the per annum average of the daily maxima of PM_{10} concentration. The same trend is more pronounced on the $PM_{2.5}$ data (Fig. 6b). The present $PM_{2.5}$ daily averages of concentrations $\approx 35\ \mu g\,m^{-3}$ are significantly smaller than what Table 1 would imply (namely $50\ \mu g\,m^{-3}$), but still significantly above the US 1996 standards of $15\ \mu g\,m^{-3}$.

Table 1
Values for the parameters used in the fits

Pollu tant t	$\rho_i = \frac{2\delta_i}{\eta_i^2}$	$x_i^{eq} = \mu_i + \frac{\beta_i E_i}{\delta_i}$	USAQS
$SO_2^{average}$	3	18 (ppb)	30 (ppb)
SO_2^{maxima}	1.8	47 (ppb)	140 (ppb)[a]
$NO_2^{average}$	4.2	31 (ppb)	50 (ppb)
NO^{maxima}	3.5	87 (ppb)	no limit
PM_{10}^{maxima}	5.58	122 ($\mu g\,m^{-3}$)	150 ($\mu g\,m^{-3}$)[b]
$PM_{10}^{average}$	5.68	98 ($\mu g\,m^{-3}$)	50 ($\mu g\,m^{-3}$)
$PM_{2.5}^{maxima}$	3.07	66 ($\mu g\,m^{-3}$)	50 ($\mu g\,m^{-3}$)[c]
$PM_{2.5}^{average}$	3.35	50 ($\mu g\,m^{-3}$)	15 ($\mu g\,m^{-3}$)

[a]Limit should not be passed more than 8 h per yr.
[b]Limit should not be violated more than 24 h per yr.
[c]Corresponds to the 98th percentile over 3 yr.

3.2. Fitting the data with $\varphi_\infty(x_i; x_i^{eq}, \rho_i) = x_i^{eq}(\rho_i)^{\rho_i}/\Gamma(\rho_i)\, e^{-\rho_i x_i^{eq}/x_i}(x_i/x_i^{eq})^{-(\rho_i+2)}$

To "fit" the data, we used the moments in the simplest possible form. i.e. we computed the first and second moment from the data and used Eqs. (12) and (13) to translate them into values for the parameters x_i^{eq} and ρ_i. The figure shows the plot of the distribution $\varphi_\infty(x_i; x_i^{eq}, \rho_i)$ on top of the data. The values for the parameters are given in Table 1. The Mathematica code we used (Mathematica™ version 3.0.1, Wolfram, 1996) is in Appendix A. Figs. 3a, a′, b, b′; 4a, b; 5a, b show that our simple two parameters expression reproduces naturally reasonably well the histograms of the data from Santiago with the value of the parameters shown in Table 1.

More sophisticated methods of parametric estimation are possible. No attempt was made to use any sophisticated parametric estimation method like maximum likelihood or variations of it. A more refined parametric estimation is best set-up with a precise purpose like finding the values of the parameters which capture as well as possible the tail of the distribution, or around a season. The seasonal variability of the data implies that the average x_i^{eq} changes value during the year (cf. Fig. 6a–d for example). Furthermore, in the case of $PM_{2.5}$ for example, the annual average has changed with time (cf. Figs. 6b and b′). A subtle quantitative estimate would be different depending on what aspect of the data is under scrutiny.

3.3. Validation of our distribution

The fact that the two parameters general expression:

$$\varphi_\infty(x_i; x_i^{eq}, \rho_i) = \frac{x_i^{eq}(\rho_i)^{\rho_i}}{\Gamma(\rho_i)} e^{-\rho_i x_i^{eq}/x_i} \left(\frac{x_i}{x_i^{eq}}\right)^{-(\rho_i+2)}$$

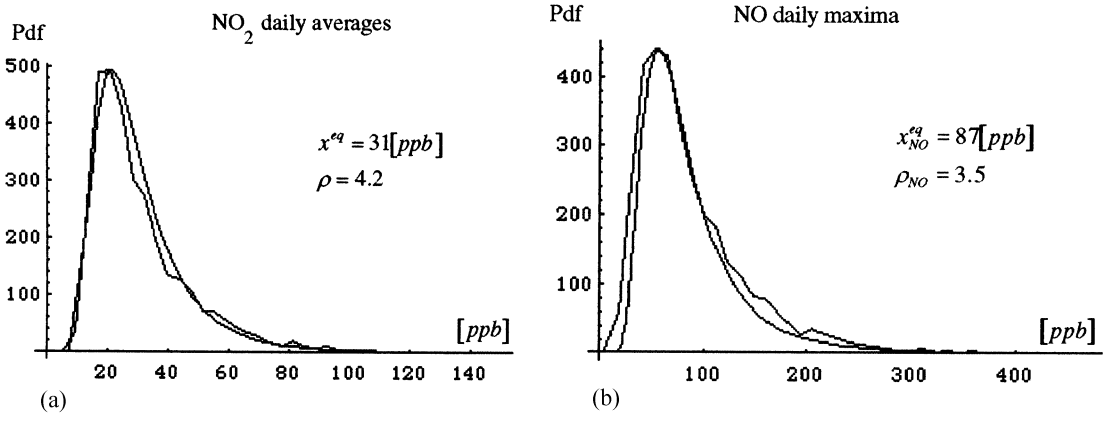

Fig. 4. (a) Fit on the histogram of the NO_2 daily average data. The average in Santiago is 31 ppb. The US legal limit for the average is 50 ppb. (b) Fit on the histogram of the NO daily maxima data. The average in Santiago is $x_{NO}^{eq} = 87$ ppb and $\rho_{NO} = 3.5$. There is no US legal limit.

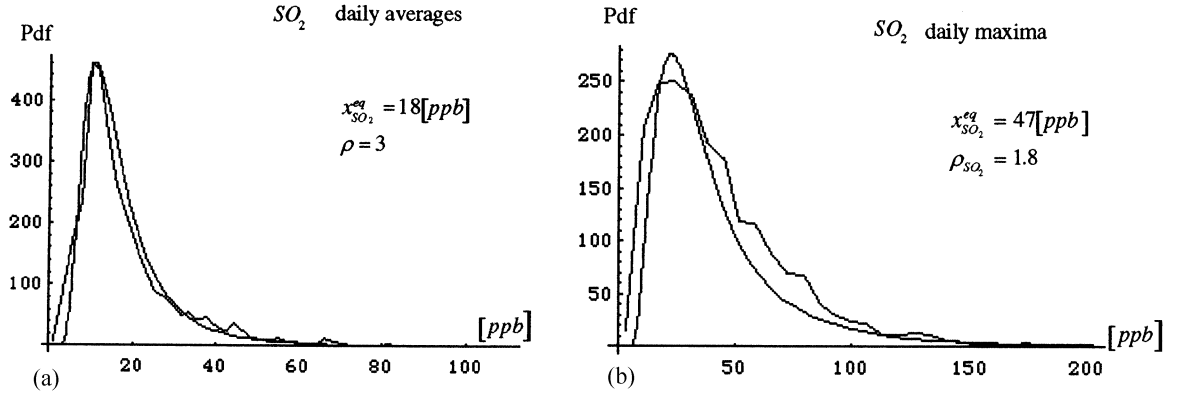

Fig. 5. (a) The fit on the histogram of the daily average SO_2 concentration data. The 1974 US standards are that it should not go over 30 ppb in average. The average in Santiago is 18 ppb. (b) The fit on the histogram of the daily peak SO_2 concentration data.. The 1974 US standards is 140 ppb for at most 24 h yr^{-1}.

can be made to reproduce qualitatively well the empirical air pollution data of Santiago is not in and of itself a proof that it is the "right" expression. To be able to study directly the relation between emission levels and concentration distribution would provide a more powerful ground to "validate" the expression

$$\varphi_\infty(x_i;\, x_i^{eq}, \rho_i) = \frac{x_i^{eq}(\rho_i)^{\rho_i}}{\Gamma(\rho_i)} e^{-\rho_i x_i^{eq}/x_i} \left(\frac{x_i}{x_i^{eq}}\right)^{-(\rho_i+2)}.$$

We do not have the kind of data needed for that. But we plan to study other cases like Los Angeles where such an analysis may be possible.

3.4. Application to policy

When estimates of the parameters ρ_i and x_i^{eq} have been found, the power of our approach for policy becomes more obvious. The value of $\rho_i = 2\delta_i/\eta_i^2$ which controls the size of the fluctuations, cannot be influenced by policy decisions. But the value of x_i^{eq} depends explicitly on the emission level ($x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$) and can therefore be influenced by policy.

AQS either prescribed values for x_i^{eq} or for the probability of going over a limit $Pr[x_i \geq u_i]$, which using Eq. (8) has the form

$$Pr[x_i \geq u_i] = \frac{\gamma(1+\rho_i, \rho_i x_i^{eq}/u_i)}{\Gamma(1+\rho_i)}. \quad (8)$$

Knowing ρ_i and $Pr[x_i \geq u_i]$, it is possible to estimate numerically from Eq. (8) what the value of x_i^{eq} should be to comply with AQS and compare with what it is.

The dependence of the probability of exceeding US 1996 standards for particulate, on the value of the parameters $x^{eq}_{PM_{10}}$, $x^{eq}_{PM_{2.5}}$ and $x^{eq}_{SO_2}$ are displayed in Fig. 7a–c. From these diagrams it is possible to estimate the values that these parameters must have in order to comply with US 1996 standards.

Once the desired value of x_i^{eq} is known, the next step is to translate that into values for the anthropogenic emissions E_i.

3.5. Numerical analysis

For SO_2, the daily average concentration of SO_2 in Santiago is $x^{eq}_{SO_2} = 18$ ppb, well within the US standard of $x^{aver}_{SO_2} \leq 30$ ppb (Fig. 5a). But one message of Fig. 7c, is that Santiago violates US standards for peak concentrations of SO_2. The US standards to compare with is $x_{SO_2} \leq 140$ ppb, except for periods shorter than 8 h in a year, less than 0.1%. Fig. 7c shows that with the average maximum concentration we estimated of $x^{eq}_{SO_2} \approx 47$ ppb, the probability that the concentration exceeds 140 ppb during a whole day is about 3%, i.e. exceeds the US standards.

This is partially due to the geography of Santiago, which is responsible for trapping the pollutants. In another city where the parameter ρ is three times larger, the same average peak concentration $x^{eq}_{SO_2} \approx 47$ ppb would correspond to a significantly smaller value of the probability of exceeding the US limit. This illustrates the area dependence of the problem and also how the area dependence is captured in our approach.

With the geographical conditions of Santiago, in order to comply with US standards, $x^{eq}_{SO_2}$ (cf. Fig. 7c) should about or be less than 25 ppb, almost half smaller than our estimated value today $x^{eq}_{SO_2} \approx 47$ ppb.

- The particulate problem of Santiago is much more severe than SO_2. The overall mean of the daily average concentrations of particulate of 98 $\mu g\, m^{-3}$ for PM_{10} and 50 $\mu g\, m^{-3}$ for $PM_{2.5}$ is to be compared with the

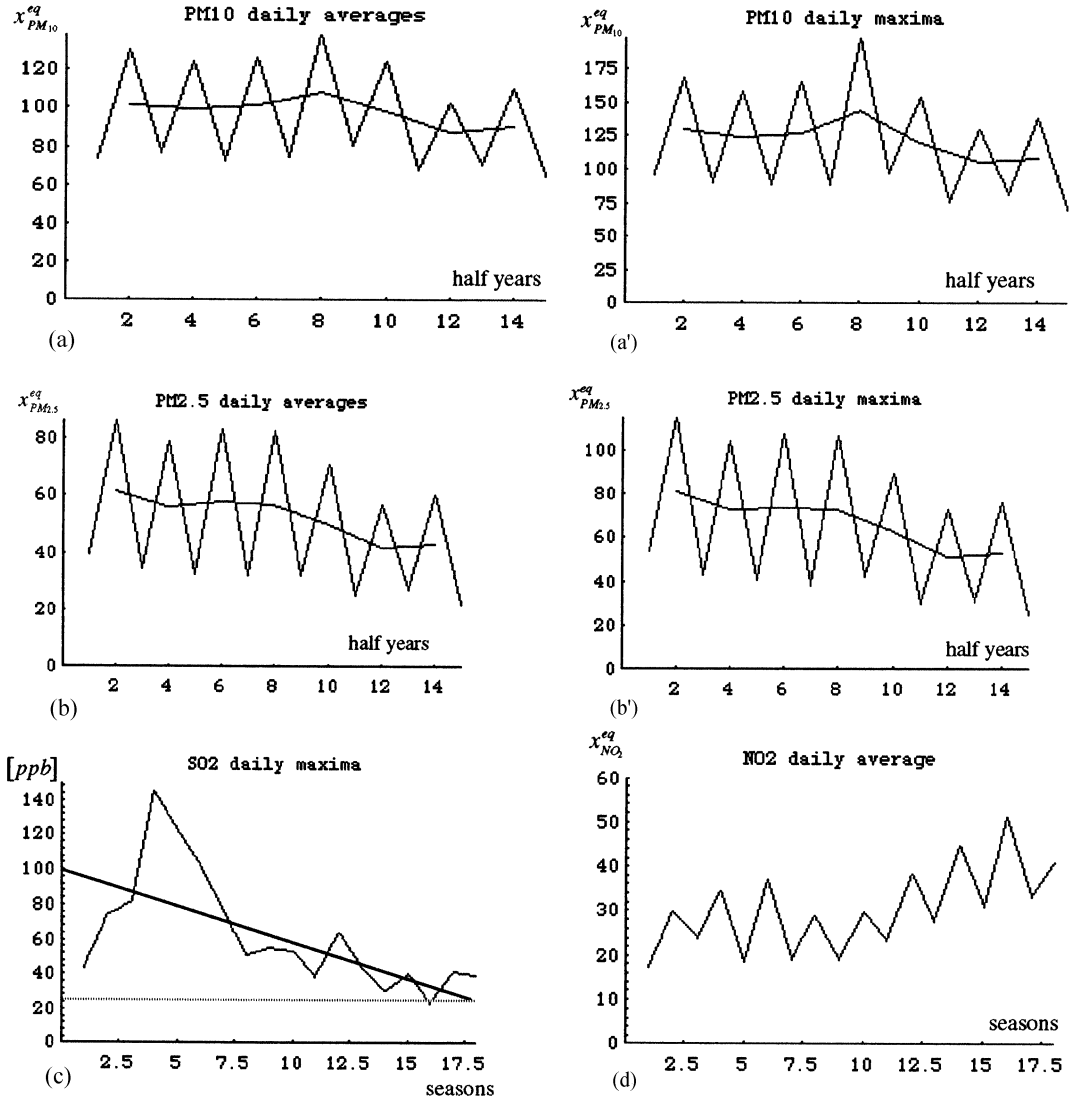

Fig. 6. (a) Seasonal variability of the value of the daily average $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ in the case of PM_{10}. Superimposed is a line joining the average value for each year. The annual average decreased from 101 to 90 $\mu g\,m^{-3}$. The effective value for all the years combined is $\approx 98\ \mu g\,m^{-3}$ (cf. Table 1). (a′) Seasonal variability of the value of the daily maxima $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ in the case of PM_{10}. Superimposed is a line joining the average value for each year. The annual average decreased from 130 $\mu g\,m^{-3}$ to 107 $\mu g\,m^{-3}$. The effective value for all the years combined is: $x^{eq}_{PM_{10}} = 122\ \mu g\,m^{-3}$ (cf. Table 1). (b) Seasonal variability of the value of the daily average $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ in the case of $PM_{2.5}$. Superimposed is a line joining the average value for each year. The annual average decreased from 61 to 43 $\mu g\,m^{-3}$. The effective value for all the years combined is: 50 $\mu g\,m^{-3}$ (cf. Table 1). (b′) Seasonal variability of the value of the daily maxima $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ in the case of $PM_{2.5}$. Superimposed is a line joining the average value for each year. The annual average decreased from 81 to 52 $\mu g\,m^{-3}$. The effective value for the whole set of years is 66 $\mu g\,m^{-3}$ (Table 1). (c) Seasonal variability and evolution with time of the value of $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ in the case of SO_2. The US standard is not to exceed 140 ppb more than 24 h yr^{-1} (0.3% probability). For Santiago, this means $x^{eq}_{SO_2} < 20$ ppb (cf. Fig. 7c). The line is a guide to the eye. (d) Seasonal variability of the value of $x_i^{eq} = \mu_i + \beta_i E_i/\delta_i$ in the case of NO_2. The US standard is 50 ppb. Shows an upward trend.

corresponding US standards 50 and 15 $\mu g\,m^{-3}$. Today the PM_{10} daily maxima exceed 150 $\mu g\,m^{-3}$ almost 10% of the time (Fig. 7a). A simplified reading of the US AQS is that, this should happen at most 24 h per annum, or with a probability of 0.3%. (The old standard of "one day a year exceedance" has been replaced by a more complicated rule: "98th percentile over three years", Federal Register, 1998).

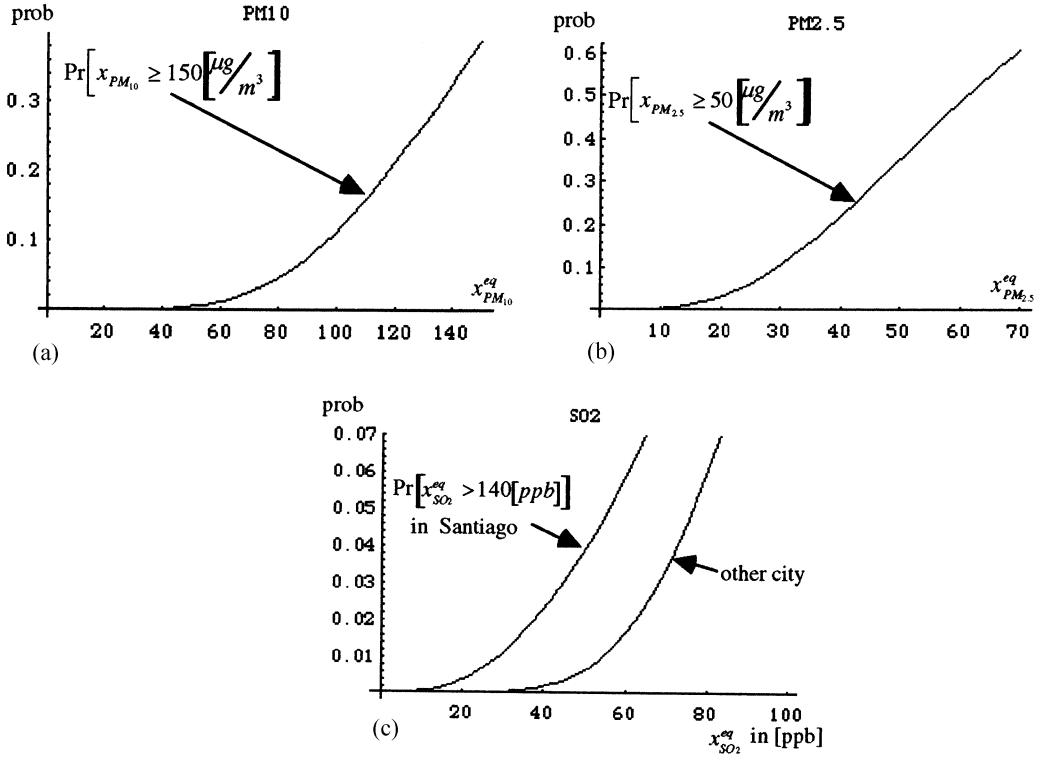

Fig. 7. (a) The case of particulates PM_{10}: The 1996 US standards for PM_{10} maxima were 150 $\mu g\,m^{-3}$ for all but at most one day in a year. In order to meet US 1974 standards for daily maxima, the value of $x^{eq}_{PM_{10}}$ should be about 50 $\mu g\,m^{-3}$. Today's value of $x^{eq}_{PM_{10}}$ is $\approx 100\ \mu g\,m^{-3}$ (cf. Fig. 6a). (b) The case of $PM_{2.5}$: The US standards allow exceeding 50 $\mu g\,m^{-3}$ at most 24 h yr^{-1}. To comply with US standards, $x^{eq}_{pm.2.5}$ should be around 15 $\mu g\,m^{-3}$. According to our estimate, it is today around 40 $\mu g\,m^{-3}$ (Fig. 6b). (c) The case of SO_2: US standards are: $x_{SO_2} \leq 140$ ppb, except for periods shorter than 24 h. In Santiago, this requires $x^{eq}_{SO_2} \approx 15$ ppb. According to our estimates (Fig. 6c), today in Santiago, $x^{eq}_{SO_2} \approx 25$ ppb. The curve shows a hypothetical city with a different geography (different value for the parameter ρ). In that case the same air quality standards translate in a less stringent requirement on the SO_2 emissions.

Daily maxima of $PM_{2.5}$ exceed the US AQS limit of 50 $\mu g\,m^{-3}$ also about 10% of the time (Fig. 7b). 50 $\mu g\,m^{-3}$ corresponds to the $PM_{2.5}$ concentration over the whole period in Santiago (Table 1).

Using Eq. (8), it is possible to estimate the amount of abatement needed in Santiago to meet US AQS. The result is shown in Figs. 7a and b. The message from these figures is that to meet US standards, for PM_{10}, the value of $x^{eq}_{PM_{10}}$for the daily maxima should be reduced to the range $\approx$ 30–50 $\mu g\,m^{-3}$, down from today's 100 $\mu g\,m^{-3}$ (Fig. 6a). And for $PM_{2.5}$, the value of $x^{eq}_{PM_{2.5}}$should be reduced to the range 15 $\mu g\,m^{-3}$, down from today's 40 $\mu g\,m^{-3}$ (Fig. 6b), i.e. by a factor of larger than two.

4. Discussion

What the use of our expression

$$\varphi_\infty(x_i;\, x_i^{eq},\, \rho_i) = \frac{x_i^{eq}(\rho_i)^{\rho_i}}{\Gamma(\rho_i)}\, e^{-\rho_i x_i^{eq}/x_i}\left(\frac{x_i}{x_i^{eq}}\right)^{-(\rho_i+2)}$$

adds to the study of the well-documented particulate problem of Santiago is a response function. i.e. for example in the case of $PM_{2.5}$, it can be used to translate AQS requirements into values for $x^{eq}_{PM_{2.5}}$ which with the expression $x^{eq,ave}_{PM_{2.5}} = \mu_{PM_{2.5}} + \beta^{ave}_{PM_{2.5}} E_{PM_{2.5}}/\delta^{ave}_{PM_{2.5}}$, can be expressed as requirements for the emission level $E_{PM_{2.5}}$. The last step is not completely straightforward as in that expression $\beta^{ave}_{PM_{2.5}}$ and $\delta^{ave}_{PM_{2.5}}$ are not known.

In fact it is not necessary to know the values of $\beta^{ave}_{PM_{2.5}}$ and $\delta^{ave}_{PM_{2.5}}$ to estimate what relative change of emission $E_{PM_{2.5}}$ is needed to meet AQS requirements, as we show now. In the process we note that the physical interpretation of $E_{PM_{2.5}}$ is more complex than only the direct emission. $E_{PM_{2.5}}$ measures the rate of production of particulates. It is sensitive to the chemical composition of the atmosphere (i.e. to the concentration in the other pollutants).

Through their seasonal variability, the data for $PM_{2.5}$ show evidence of a reduction of atmospheric load above Santiago over the years (Fig. 6b and b′) a bit more

pronounced than the data for PM_{10} (Fig. 6a and a′). Over the seven years of measurement, the annual mean of the daily average $x^{eq}_{PM_{2.5}}$ decreased from 61 to 43 $\mu g\,m^{-3}$ (Fig. 6b), a significant improvement, but still a long way from the US standards of 15 $\mu g\,m^{-3}$.

To meet US AQS, $x^{eq}_{PM_{2.5}}$ should reach the value smaller than 15 $\mu g\,m^{-3}$ (a reduction by more than a factor 3) and $x^{eq}_{PM_{10}}$ has to be reduced to 50 $\mu g\,m^{-3}$, i.e. by a factor of 2. This suggests that Santiago has still a long way to go.

The average value $x^{eq,max}_{PM_{2.5}}$ of the daily maxima during that period went from 81 to 53 $\mu g\,m^{-3}$. To meet the US standards, $x^{eq,max}_{PM_{2.5}}$ should be less than 15 $\mu g\,m^{-3}$ (Fig. 7b).

In order to translate the change of atmospheric load into change of value of $E_{PM_{2.5}}$, we use the fact that by subtracting $x^{eq,ave}_{PM_{2.5}} = \mu_{PM_{2.5}} + \beta^{ave}_{PM_{2.5}} E_{PM_{2.5}}/\delta^{ave}_{PM_{2.5}}$ from $x^{eq,max}_{PM_{2.5}} = \mu_{PM_{2.5}} + \beta^{max}_{PM_{2.5}} E_{PM_{2.5}}/\delta^{max}_{PM_{2.5}}$, one gets

$$\frac{(x^{eq,max}_{PM_{2.5}} - x^{eq,ave}_{PM_{2.5}})}{E_{PM_{2.5}}} = \frac{\beta^{max}_{PM_{2.5}}}{\delta^{max}_{PM_{2.5}}} - \frac{\beta^{ave}_{PM_{2.5}}}{\delta^{ave}_{PM_{2.5}}}. \quad (14)$$

Writing Eq. (14) for two different times and taking the ratio, it is possible to translate the reduction in the $PM_{2.5}$ load into a relative reduction of emission level:

$$\frac{E^1_{PM_{2.5}}}{E^2_{PM_{2.5}}} = \frac{(x^{eq,max}_{PM_{2.5}} - x^{eq,ave}_{PM_{2.5}})|_1}{(x^{eq,max}_{PM_{2.5}} - x^{eq,ave}_{PM_{2.5}})|_2}. \quad (15)$$

If in Eq. (15), "1" refers to the final time, and "2" refers to the initial time, then $E^1_{PM_{2.5}} = E^2_{PM_{2.5}} - \Delta E_{PM_{2.5}}$ and $E^1_{PM_{2.5}}/E^2_{PM_{2.5}} = 1 - \Delta E_{PM_{2.5}}/E^2_{PM_{2.5}}$.

The reduction of the $PM_{2.5}$ load by a third that took place in the seven years of measurements could be explained by an emission reduction of

$$\frac{E^1_{PM_{2.5}}}{E^2_{PM_{2.5}}} = 1 - \frac{\Delta E_{PM_{2.5}}}{E^2_{PM_{2.5}}} \approx 0.5. \quad (16)$$

The same computation with PM_{10} leads to the conclusion that the relatively modest decrease of atmospheric load of about 10% ($x^{eq,ave}_{PM_{10}}$ went from 101 to 90 $\mu g\,m^{-3}$ (Fig. 6a) and $x^{eq,max}_{PM_{10}}$ went from 130 to 107 $\mu g\,m^{-3}$ (Fig. 6a′)) corresponds to a sizable emission reduction of about 40%

$$\frac{E^1_{PM_{10}}}{E^2_{PM_{10}}} = 1 - \frac{\Delta E_{PM_{10}}}{E^2_{PM_{10}}} \approx 0.58. \quad (17)$$

In both cases, E_{PM} does not represent strictly the emission but the source of $PM_{2.5}$ and PM_{10}. The direct emission abatement which took place during that period were presumably much smaller. Other factors affected the value of E_{PM}, like the changes in the chemical composition of the atmosphere. In the case of $PM_{2.5}$, for example one expects a strong contribution from the secondary production of particulates associated with the presence of compounds like sulfates, nitrates and ammonium (Ansari and Pandis, 1998). It is therefore noteworthy that the reduction of the $PM_{2.5}$ atmospheric load in Santiago coincided with a parallel and significant reduction on the SO_2 atmospheric load (Fig. 6c). A reduction of the concentration of SO_2 leads to reduction of sulfates, which in general leads to a reduction of particulates.

Noteworthy also is the slow but steady increase of the concentration of NO_2 during the same period (Fig. 6d). Particulates interact also with nitrates although with less affinity than sulfates.

Acknowledgements

Cifuentes acknowledges the support of Fondecyt Project 197011. We acknowledge the support from the grant NSF, SBR-9521914. We want to express our appreciation to Profs M. Small and S. Pandis for their interest and numerous helpful comments.

Appendix A. Mathematica code

```
      data = Table["..."];(* data as a series of numbers*)
str1 = StringToStream[data];
plad = ReadList[str1,Number];

ni = nbin;(* number of bins for histogram of data*)
xmin = 1;
xmax = Max[plad];

For[i = 1, i < = ni, i + + ,
     a = (i − 1)*Max[plad]/ni;
     b = i*Max[plad]/ni;
     x[i] = 0.5*(a + b);
     num[i] = Length[Select[plad, a < # < = b & ]]];
hist = Table[num[i], {i, 1, ni}];
histod = Table[{x[i], num[i]}, {i, 1, ni}];

     (*histogram of data*)
histogram = ListPlot[histod, PlotJoined- > True,
       PlotRange- > All]

total = Apply[Plus, hist]; (*total number of data*)

     (*computation of moments*)
sum1 = 0; (*first moment*)
sum2 = 0; (*second moment*)
For[i = 1, i < = ni, i + + ,
     sum1 = sum1 + x[i]*num[i]/total;
     sum2 = sum2 + num[i]*x[i]^2/total];

xeq = sum1;
lamb = sum2/sum1^2;
```

```
rho = sum2/(sum2-sum1^2);
xtop = rho*xeq/(rho + 2);
h[x_]: = Exp[-rho*xeq/x]*x^(-rho-2);
plotexp = Plot[Max[hist]*h[x]/h[xtop],
        {x, xmin, xmax}, PlotRange- > All];
Show[histogram, plotexp]
```

References

Ansari, A.S., Pandis, S.N., 1998. On the response of inorganic PM to precursor concentrations. Environmental Science and Technology, submitted.

Crovella, M., 1997. IEEE/ACM Transactions on Networking 5, 835–846.

Dixit, A., Pindyck, R., 1993. Investment Under Uncertainty. Princeton University Press, Princeton, NJ.

Georgeopoulos, P., Seinfeld, J., 1982. Environmental Science and Technology 16, 401A–415A.

Johnson, N.L., Kotz, S., 1970. Continuous Univariate Distributions. Wiley, New York.

Karlin, S., Taylor, H., 1981. A Second Course in Stochastic Processes. Academic Press, New York.

Mandelbrot, B., 1982. The Fractal Geometry of Nature. Freeman, New York.

Samorodnitsky, G., Taqqu, M.S., 1994. Stable Non-Gaussian Random Processes. Chapman & Hall, London.

Seinfeld, J., Pandis, S.N., 1998. Atmospheric Chemistry and Physics. Wiley, New York.

Teverosvky, V., Taqqu, M., 1997. Journal of Times Series Analysis 18, 279–304.