AIR POLLUTION AND HEALTH EFFECTS IN SÃO PAULO, BRAZIL:

A TIME SERIES ANALYSIS

Nelson Da Cruz Gouveia

Thesis submitted to the Faculty of Medicine of the University of London for the degree of Doctor of Philosophy

London School of Hygiene and Tropical Medicine

1997

ABSTRACT

A time series study was conducted to investigate the association between variations in daily levels of air pollution and health effects in the city of São Paulo, Brazil. This study was prompted by positive associations reported in other time series studies, principally in North America and Europe, and preliminary results from some limited analyses reported for São Paulo. Its aims were to examine specific causes of mortality and hospital admissions, to identify more vulnerable subgroups defined in terms of age, to assess the role of socio-economic conditions in modifying the association and to detail the impact of other potential risk factors, especially meteorological.

Daily measurements of air pollutants (PM_{10}, SO_2, NO_2, O_3 and CO) for 12 monitoring stations across the city and several meteorological variables, along with daily counts of mortality for all ages during 3 years and hospitalisations for children during 23 months were available. The time series models used Poisson regression analysis and were adjusted for effects of trend, cyclical patterns (including season), weekday, holidays, meteorological factors, and autocorrelation.

Increases in PM_{10} and SO_2 were associated with a 3-4% increase in daily deaths for all causes in the elderly (results are presented for an increase from the 10^{th} to the 90^{th} centile of pollution measurements). Cardiovascular deaths were additionally associated with CO (4% increase). Respiratory deaths in the elderly showed higher increases (6%) associated with PM_{10}. No significant effects for children's mortality were observed. Nevertheless, respiratory or pneumonia hospital admissions for children showed significant increases associated with O_3 (5-8%), NO_2 (9%), and to a lesser extent with PM_{10} (9%). There was a significant trend of increasing risk of death according to age with effects only evident for older subjects. However, this age effect was more evident for all cause mortality. There was a weak suggestion of larger effects on mortality for areas economically more affluent. Some indication was found of a harvesting effect occurring in the mortality and hospital admission series in São Paulo.

Results are broadly consistent with those previously reported but somewhat smaller in magnitude. In contrast with an earlier preliminary analysis in São Paulo, there were no effects on mortality for children. However, new analyses for hospital admissions indicated that children are at an increased risk of non-fatal illness in relation to air pollution.

TABLE OF CONTENTS

CHAPTER 3. METHODS

CHAPTER 4. DESCRIPTIVE RESULTS

CHAPTER 5. TIME SERIES: MORTALITY

CHAPTER 6. TIME SERIES: HOSPITAL ADMISSIONS

CHAPTER 7. SENSITIVITY ANALYSIS

CHAPTER 8. DISCUSSION

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 8

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

ACKNOWLEDGEMENTS

This work would have been impossible without the help, support and collaboration of several people to whom I want to express here my gratitude.

Above all, I want to thank my supervisor, Dr. Tony Fletcher, head of the Environmental Epidemiology Unit of the London School of Hygiene and Tropical Medicine, who promptly accepted me as a student after my upgrading and has since been working closely with me, giving essential support and incentive throughout all the different phases of this project. His warm character and relaxed approach made it a pleasure to work with him, and the study certainly gained very much in quality and interest with his high-quality scientific level contributions.

I would like also to acknowledge the help received from Dr. Ben Armstrong who gave me invaluable statistical advice and made the very useful suggestions concerning the presentation of academic work. In addition, I want to thank the whole staff of the Environmental Epidemiology Unit for providing the most friendly environment in which to work and for innumerable tips from statistical and computing to clerical support. In particular I want to thank Bharat Thakrar, Peter Egger, Carolyn Stephens and Paul Wilkinson for their assistance.

The complex and intricate mysteries of time series analysis would not have been deciphered without the inestimable help from Antonio Ponce de Leon and Giota Touloumi, who had the patience and disposition to introduce me to this new area. I also owe my gratitude to Stephen Bremner not only for statistical advice and endless discussions on time series analysis, but also for providing unfailing support which made my journey throughout this PhD much more supportable.

In Brazil, I am especially grateful for the collaboration of Carlos E. Negrão and Masayuki Kuromoto from the Companhia de Tecnologia de Saneamento Ambiental (CETESB) in obtaining the air pollution data; Fernão Dias de Lima from the School of Public Health in São Paulo for help in decoding the mainframe tapes containing these data; Dr. Paulo Saldiva and his group at the Faculdade de Medicina in making available part of the air pollution data also used in this work; Marisa C. Salle from the Centro de Informações em Saúde (CIS) for providing the hospital data; and the staff from the Instituto Astronômico e Geofísico (IAG) of the University of São Paulo for providing the meteorological data. My thanks also to Marcos Drumond of PRO-AIM in São Paulo who provided the mortality data and other important information used in this work.

Many other people, directly or indirectly, helped me enormously during different stages of my studies in London. Carlos A. Monteiro, my first mentor, will always have my greatest appreciation.

I am also grateful to Regina Rodrigues for help on innumerable occasions and Mara D. Gomes for her sound advice.

I could state here an endless list of names of colleagues and friends who, in one way or another, had shared with me moments of joy or anguish during the past few years. To mention only a few of them: Mike, Julio, Denilson, Eunice, Tete, Nani&Ligia, Jane, Isabellita, Romulo, Solange, Ana Ludermir, Regininha, Vicky, Malu, Mario, Karina, Clymene, Ciça&Cris, Ronaldo, Roberto, Carminha, and many other colleagues at the London School of Hygiene and Tropical Medicine.

I would like also to acknowledge the financial support provided throughout all my stay in London by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of the Ministry of Science and Technology of Brazil, without which this project would not have been possible.

Last, but by no means least, I want to express my deepest gratitude to my mother, my father (in memorium) and my brothers for their continuous support, their encouragement and their most precious love. To them I dedicate this work.

CHAPTER 1. INTRODUCTION

One of the milestones in our history was the discovery of fire. It was also at that moment that man started to produce air pollution. Ever since, even during pre-industrial times, activities such as cooking and heating have given rise to air pollution. However, until the 19th century, air pollution was not a widespread problem because the atmosphere was able to dilute it easily and pollution did not build up over densely populated areas. It was the advent of the industrial revolution in the western world that rapidly increased the amount and variety of chemical contaminants in the atmosphere.

Industrialisation jointly with the process of urbanisation brought so-called progress at the cost of high consumption of energy to supply the industrial machinery, public transport systems, and domestic electrical goods. It has raised the standard of living of the population, leading to improvements in life expectancy, with a concomitant introduction of new threats to health. Human activities have resulted in the production and release of a variety of air contaminants and have thus become major agents of environmental change on global,

regional and local scales. This contamination of the environment due to human activities has nowadays reached global proportions [UNEP,93].

Nevertheless, it was only earlier in this century, when a series of severe short-term episodes of extremely high levels of air pollution occurred that concern was raised about possible adverse effects on human health caused by such air pollution. But while this concern was previously restricted to developed nations, in more recent decades, industrialisation and urbanisation has spread to many other countries around the world. This rapid and, in many cases disorganised process of urbanisation for many developing countries has resulted in some of the problems experienced by developed countries earlier this century, being repeated.

Pollution episodes of proportions similar to the ones in the first half of this century do not happen frequently anymore. However, urban population is on an increase and it is estimated that, currently, nearly 50% of the human population is living in urban areas [WRI,1996]. In addition, there has been a sharp increase in the number of vehicles and in automobile traffic, creating new pollution problems. At the same time, concern has been growing again in more recent years in relation to adverse health effects due to exposure to air pollution, at levels of exposure much lower than previously.

Mechanisms to explain effects on the mortality and morbidity of the population remain uncertain, as does the magnitude of the relationships. This thesis aims to contribute to the body of knowledge on the quantitative relationship between exposure to air pollution and adverse human health effects, in a developing country context where good quality data is lacking. The main part of the research is an epidemiological investigation carried out in the city of São Paulo, in Brazil.

Chapter 2 introduces a general picture of the current situation of air pollution worldwide and in the city of São Paulo. In addition, an overview of the epidemiological literature with respect to the relationship between air pollution and health effects, the possible biological mechanisms for such effects, the rationale, the hypotheses and the objectives of the study are presented. Chapter 3 describes in detail the design and the methods utilised for the epidemiological study. An overview of the characteristics of the city of São Paulo, the sources of data and the data collection procedures are also included. In addition, a

comprehensive description of the analytical methods employed and of the modelling strategy used in this study is presented.

The main descriptive results for the exposures, the two different outcomes and the most important confounders are outlined in Chapter 4. Chapters 5 and 6 present the analytical results for the two different groups of outcomes examined in this study, mortality and hospital admissions. In addition, Chapter 5 provides more detail and exemplifies each step in the modelling procedures. Some of the key analytical assumptions of this study are examined in Chapter 7 in order to check how sensitive the findings were to alternative modelling strategies.

The results found in this study are then considered throughout Chapter 8 in relation to the whole body of evidence derived from different studies. Issues related to design, conduct, analysis and interpretation of these results are also comprehensively examined. Finally, Chapter 9 aims to draw up some general and specific conclusions about this study and the relationship between air pollution and adverse health effects.

CHAPTER 2. BACKGROUND AND RATIONALE

This chapter aims to provide an overview of some of the main issues related to air pollution and its associated human health effects. To understand the widespread and growing concern in relation to air pollution, trends in emissions and ambient levels for many cities worldwide are discussed in the context of the process of industrialisation and urbanisation. This is mostly based on data from the United Nations Environment Program (UNEP) and the air quality monitoring project managed by UNEP-World Health Organisation (WHO) which includes sites in many different countries worldwide. Then, the past and current situations of air pollution in São Paulo, Brazil, are summarised and discussed.

The epidemiological evidence of harmful effects on health due to exposure to air pollution is then reviewed as well as the main problems and limitations of those studies. Following that, some pathophysiological considerations about the effects of each pollutant are briefly

overviewed. A conceptual framework, as well as the hypothesis and objectives for the present study are then introduced.

2.1. AIR POLLUTION: TRENDS IN EMISSIONS AND LEVELS WORLDWIDE.

Air pollution in urban areas arises from a multitude of sources. The combustion of fossil fuels for domestic heating, for power generation, in motor vehicles, in industrial processes and in the disposal of solid wastes by incineration are generally the principal sources of air pollution emissions to the atmosphere [UNEP,1992].

For industrialised countries where the emissions from power plants have been subject to increasingly strict control, the main source of air pollutants comes from motor vehicles. Cities of the developing world, in contrast, exhibit a greater variety in the nature of their air pollution sources. For example, cities in Latin America, tend to have higher vehicle densities than those in other developing regions. This source is less important in cities with lower levels of motorization (e.g. cities in Africa) and in cities located in temperate regions that are still dependent on coal or biomass fuel for space heating and other domestic purposes (e.g. cities in China and parts of Central and Eastern Europe) [UNEP,1992].

Traditionally the most important pollutants present in the atmosphere and which are also more subject to monitoring programmes include Sulphur Dioxide (SO_2), the Oxides of Nitrogen (NO and NO_2, collectively termed NO_x), Carbon Monoxide (CO), Ozone (O_3), and Suspended Particulate Matter (SPM). In this review a more detailed description of emissions is provided for Sulphur Dioxide and Oxides of Nitrogen because more complete data was available for these pollutants.

Acid gases, such as SO_2 and NO_x are significant atmospheric pollutants at both local and regional scale. Apart from the direct health effects associated with them, these pollutants are also precursors of aerosol particle formation [UNEP,1993]. In addition, these pollutants have in common the fact that their emissions are extremely linked to the economical and industrial processes of the countries. The overview on trends and current ambient levels includes most of the traditional air pollutants.

2.1.1. Sulphur Dioxide

Emissions of SO_2 originate from natural or anthropogenic sources. According to UNEP [1989,1993] natural emissions are approximately equally divided between developed and developing countries. Meanwhile, it is estimated that around 70% of man-made emissions are produced in developed nations. In fact, Europe and North America are by far the largest contributors to the release of SO_2 into the atmosphere, although China and former USSR follow very closely [UNEP,1989].

On the other hand, rates of change for such emissions between 1970 to 1986 showed a remarkably contrasting picture. Increases during this period were greater for Asia where emissions were almost doubled. Latin American and African countries followed with considerable increases while in Europe emissions of SO_2 in 1986 were practically the same as in 1970. North America (USA and Canada) showed SO_2 emissions falling, particularly for low-level emission sources [UNEP,1993].

The general picture emerges of emissions of SO_2 increasing in developing countries while decline or stabilisation is evident in more developed nations. The rise in emissions might be due to the rapid economic and industrial development and consequent urbanisation underway in many developing nations. It might also reflect less strict or absent pollution control in those countries. On the other hand, declines in emissions in developed countries were a result of the implementation of varied emissions control policies. Typical strategies included fuel switching (from high sulphur coal and oil to low sulphur fuels such as gas) and the increasing use of electricity and natural gas for domestic heating. Many countries have also had marked success in reducing industrial emissions by installing air pollution control equipment at enterprises. It should be emphasised that decreases in emissions also happened in some developing countries.

As emissions have fallen significantly in most industrialised countries since the late 70's, declining trends in urban levels of SO_2 have been observed. However, levels well above the guidelines proposed by WHO (40-60$\mu g/m^3$ for annual mean or 100-150$\mu g/m^3$ for daily mean) are still evident in some cities in developing countries. According to the United Nations Environment Programme [UNEP,1991], nearly one third of the monitored cities in the world had annual averages of SO_2 above WHO guidelines in the period between 1980-1984. Even in some places where mean concentrations of SO_2 were typically within WHO

guidelines, there may be days or even months when levels substantially exceed those guidelines.

2.1.2. Suspended Particulate Matter

Different patterns are observed for Suspended Particulate Matter (SPM) which is one of the most widely measured pollutants. In fact, most countries measure just Total Suspended Particulate Matter (TSPM) and only few can provide measurements for its respirable fraction, PM_{10} (which is constituted by particles smaller than 10 μm of diameter). No global trends in SPM emissions have yet been established and national emission inventories are limited. For some industrialised countries where data are available, emissions of SPM from 1970 to 1984 showed a decline [UNEP,1987]. In addition, the decrease in domestic coal use in many countries over the past 10-20 years has led to a progressive reduction in the emissions of certain particles. Indeed, diesel vehicles have now overtaken domestic coal use as the major source of particles in many cities [UNEP,1993]. Nevertheless, some countries still rely on coal or oil for domestic heating and cooking and therefore, their emissions continue to be high.

In terms of urban levels, suspended particulate matter is one of the most prevalent pollutants worldwide. According to the inventory of WHO, out of 41 monitored cities around the world, roughly 60% had annual average concentrations above the WHO guidelines (60-90 $\mu g/m^3$ for total suspended particles - TSP) [UNEP,1991]. It can also be observed that concentrations in developing countries are noticeably higher than those of developed countries [UNEP,1993]. This is probably due to the process of industrialisation and urbanisation in the absence of emission controls, which has meant that emissions from industry and motorised vehicles are increasingly causing air quality problems.

In spite of limited data, it appears that for many countries trends in urban levels of SPM have not presented marked increases or decreases in the past few years, although in some countries where levels were extremely high a declining trend can be observed.

2.1.3. Oxides of Nitrogen

Combustion of fossil fuel produces two oxides of nitrogen: nitric oxide (NO) and nitrogen dioxide (NO_2) which are collectively termed oxides of nitrogen (NO_x). NO is generally produced first but, once in the atmosphere, photochemical oxidation may occur to produce NO_2.

Emissions of NO_x have a pattern very similar to that presented for SO_2 (Table 2.1). Europe and North America together are responsible for nearly 70% of the total release of this pollutant into the atmosphere. However, while emissions have increased approximately one third between 1970 to 1986 in nearly all continents, North America exhibited relatively stable trends in emissions [UNEP, 1993].

This upward trend in emissions for NO_x principally reflects the increase in the number of vehicles in most countries. Vehicle exhaust fumes are a major source of this pollutant and there has been less widespread application of emission control procedures compared to SO_2.

Table 2.1 - Emissions of Oxides of Nitrogen from anthropogenic sources (in 10^6 t per year), by continent around 1979.

	Emissions	(%)
Africa	0.5	2.3
North America	7.0	32.1
South America	0.8	3.7
Asia	5.0	22.9
Europe	8.0	36.7
Oceania	0.5	2.3

source: adapted from UNEP [1989]

In terms of urban levels there are much less data available for recent trends in NO_x because these gases are not part of the WHO monitoring scheme (GEMS). Nevertheless, it appears that a number of cities in both North America and Europe have experienced either stationary or slightly upward trends in urban levels of NO_2 during the past 10-15 years [UNEP,1991]. Although none of the cities for which data are available showed excessive NO_2

concentrations it could well be that in places with high traffic volumes, levels of NO_2 are high at least in some parts of the city [UNEP,1992].

2.1.4. Other Pollutants

Ozone is generated by two processes: the downwards mixing of stratospheric O_3 and, more significantly, as a result of chemical reactions involving the absorption of solar radiation by NO_2 in the presence of volatile organic compounds (VOCs) and carbon monoxide (CO). The principal sources of VOCs, CO, and NO_2 in urban areas are motor vehicles and high concentrations of O_3 are set to occur in places which combine high precursor emissions and insolation (sunshine).

Ozone levels are subject to monitoring systems in only few places worldwide. Nevertheless, there is evidence that surface levels have been increasing in the past 2-3 decades as a result of the rise in its precursors from man-made combustion sources [UNEP,1993]. Typically high concentrations of this gas are found in places with heavy traffic and with a high degree of sunshine, such as cities in hot sunny climates, e.g. some tropical developing countries.

The major source of carbon monoxide is vehicle emissions. It is calculated that 90-95% of CO emissions are generated by vehicles. Concentrations of this pollutant depends particularly on the site where the monitoring station is located. Levels are expected to be higher near main traffic corridors.

Trends worldwide shows that urban levels of CO are elevated where there is high traffic density. Even in some cities monitored by WHO where levels are not that high, air quality standards for CO (9 ppm for a 8-hours moving average) had been exceeded quite frequently [UNEP/WHO,1988].

2.1.5. Conclusions

The comparisons of air pollution emissions and levels between different countries or cities must be regarded with caution. Firstly, because air-quality data and information on emissions sources are collected in a variety of ways in different countries. Secondly, the number and criteria for positioning of monitoring stations or the time period over which

they collect data are not the same for all countries. Despite these problems the previous comparisons are helpful in providing a general picture of the air pollution problem worldwide.

It could be seen that developed countries in general present the highest emissions of pollutants although these emissions have decreased or at least stabilised in the past few years. Decreases have been achieved in most cases despite an increase in fuel consumption and economic output through the adoption of a mix of pollution control strategies, energy conservation measures and through fuel switching. Moreover, the transition to an industrialised and urbanised society happened gradually for these countries which has allowed them to progressively enforce environmental policies and control of emissions from stationary sources.

Recently, these countries have experienced a shift in the type of their air pollution sources. Economic development and increasing personal wealth of their populations has resulted in an increase in motor vehicle traffic. This in turn has meant that levels of air pollutants associated with vehicle emissions, especially NO_x, CO and hydrocarbons, have increased in many places, although the increases are offset by measures for abatement of car-related pollution.

On the other hand, less developed countries have presented the highest growth in emissions of pollutants over the past few decades because of rapid population growth, industrialisation and increasing demand for energy [UNEP,1992]. For most of these countries emissions are rising because they are going through a very rapid and disorganised transition from a rural to an industrial way of production while still relying on combustion of high-sulphur coal for cooking and domestic heating. In addition, in order to conform to global economic forces, many of these countries have to either implement flexible or weaker environmental regulations or to completely neglect such legislation.

Therefore, while in more developed countries there was a shift in sources of air pollution from combustion of fossil fuels to motor vehicles, less developed countries have been facing a double burden in terms of pollution with increasing emissions from both sources. Many of the problems faced by developed countries in the past in relation to industrial and domestic pollution are being repeated. Additionally, the contribution to air pollution due to motor vehicle emission is increasing much faster in these countries. In fact, it is in developing

countries and in eastern Europe that the greatest increases in the number of motor vehicles are expected. For countries such as India and China the number of passenger cars in use increased by three to four times between 1981 and 1991 while most developed countries showed relatively small increases over the same time period [UNEP,1992;UNEP,1993; WRI,1996].

As a result, the uneven pattern of emissions and environmental control policies among developed and developing countries has been reflected in uneven levels for the traditional pollutants as exhibited in Figures 2.1 and 2.2. Urban air pollution in cities of developing countries have been and still are in general much higher than in their counterparts in the developed world.

The figures show the annual average levels of SO_2 and SPM for three recent time periods for an aggregate of countries based in their income. It can be observed that high income countries not only present the lowest levels for both pollutants but they also show a declining trend in levels of SO_2 and (to a lesser extent) SPM levels. In contrast, mainly low-income countries have been presenting constant or even slightly increasing levels of pollutants over the same period, widening a gap that already existed between developed and developing countries.

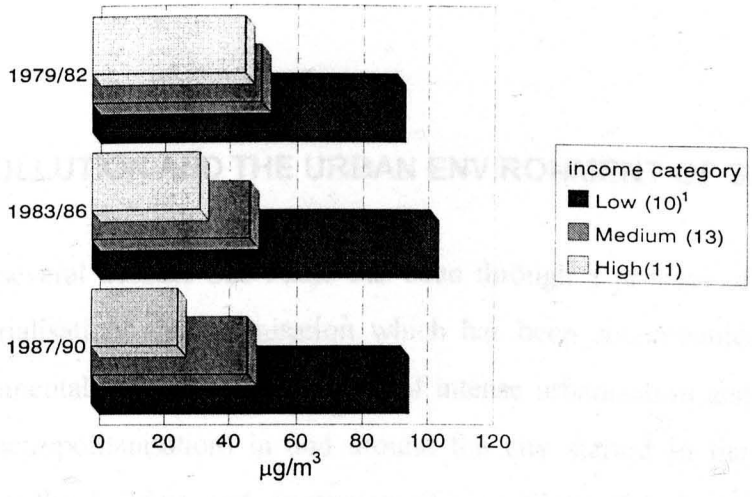


Fig 2.1 - Comparison of levels of Sulphur Dioxide in three different time periods among countries grouped by their income. [1] Figures in brackets refer to number of countries used to calculate the mean average. Data for each country was used when available for four years or more and for all three time periods. Source: modified from World Bank [1992].

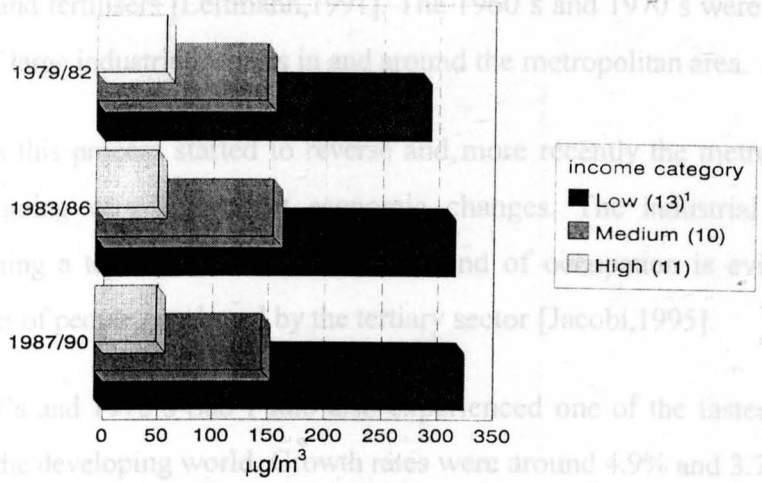


Fig 2.2 - Comparison of levels of Suspended Particulate Matter (SPM) in three different time periods among countries grouped by their income. Only data derived from cities with gravimetric measurements of SPM. See notes on Fig 2.1. Source: modified from World Bank [1992].

In conclusion, it seems that although the burden of exposure to air pollution is an almost inevitable feature of urban living throughout the world, the affliction of pollution is relatively more important for developing countries where levels are significantly higher and industrial emissions coexist with traffic related emissions. Environmental inequalities might also be observed within many of the these countries with the worst of the environmental pollution experienced by the politically weaker and socially more deprived groups. Nevertheless, only few studies on the effects of such deterioration of air quality have been conducted in developing countries and measures for control of emissions have also not yet been satisfactorily implemented.

2.2. AIR POLLUTION AND THE URBAN ENVIRONMENT OF SÃO PAULO

Over the past several decades São Paulo has been through a process of rapid population growth, industrialisation, and urbanisation which has been accompanied by increasingly serious environmental problems. The process of intense urbanisation and formation of the conurbation (metropolitanisation) in and around the city started in the 1940's and was characterised by the building and expansion of a number of industries which emitted significant pollutants, e.g., automobile manufactures, electrical appliances, steel,

petrochemicals and fertilisers [Leitmann,1991]. The 1960's and 1970's were marked by the establishment of large industrial centres in and around the metropolitan area.

After the 1970's this process started to reverse and more recently the metropolitan region has been undergoing very significant economic changes. The industrial metropolis is gradually becoming a tertiary one. This shift in trend of occupation is evidenced by the increased number of people employed by the tertiary sector [Jacobi,1995].

During the 1960's and 1970's São Paulo also experienced one of the fastest growth rates among cities of the developing world. Growth rates were around 4.9% and 3.7% per year for those decades. Fertility rates were also extremely high during that period. However, a reduction in the population growth more recently brought this rate to nearly 1.0% per year for the period between 1980 to 1991 [Jacobi,1995; FSEADE,1993a].

With increasing industrialisation the problem of air pollution in São Paulo soon became a concern for citizens and government and was introduced on to the political agenda. Emission control started in 1976 by the Company of Environmental Technology and Sanitation (CETESB) of the State Secretary for Environment and it has been relatively successful in reducing overall levels of industrial air pollution, especially during the period of industrial growth. By enforcing emission standards, penalising the industries when their emissions were above permitted levels, and mandating use of best available practice for the largest sources of industrial pollution, CETESB has achieved important reductions in industrial emissions of SO_2, CO, and NO_x.

However, while emission abatements were relatively successful for stationary sources, and the urbanisation-industrialisation process slowed down, the motor vehicle fleet in São Paulo expanded from 1.6 to 3.8 million vehicles. Vehicle numbers increased by 140% between 1980 and the early 1990's when population growth had stabilised (Fig 2.3). Recently, the vehicle fleet in the city of São Paulo was estimated to be 4.5 million cars and 12,000 buses and represents nearly 20% of the country's total [CET,1992].

This huge increase in the vehicle fleet has made motor vehicle emissions the most important source of air pollution in the city São Paulo. Currently, vehicles and related activities (fuel evaporation, crankcases, tyres) account for 94% of CO emissions, 89% of hydrocarbons, 92% of NO_x, 64% of SO_2, and 40% of particulate emissions (Table 2.2).

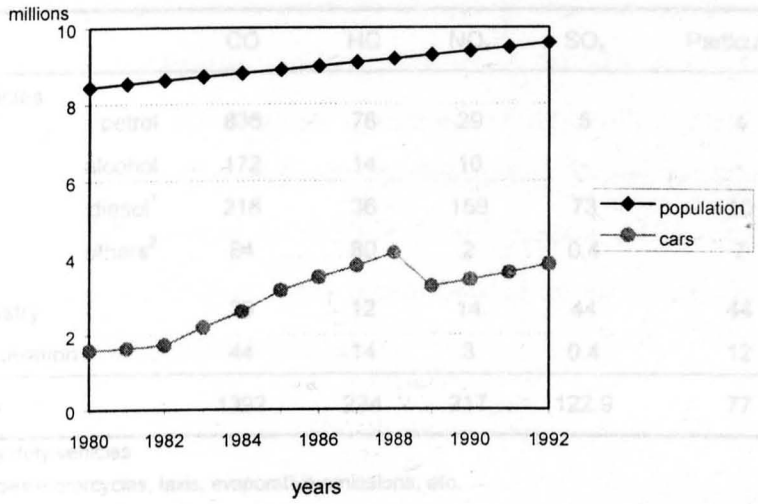


Fig 2.3 - Trends in population and number of motor vehicles (in millions) in São Paulo, Brazil, 1980-1992. Due to changes in registration of all vehicles from 1988 to 1990 there was an artificial decrease in the fleet during this period. Source: CET [1992].

Control of emissions from these mobile sources had also been enforced by CETESB. Since 1976 there is a control programme for emissions of black smoke from heavy duty diesel vehicles and in 1986 a more complete set of control measures on emissions from petrol cars was also implemented. An important measure adopted was the addition of alcohol to petrol, replacing lead as an additive. This led to a reduction in the emissions of CO, while hydrocarbons and NO_x remained at the same levels. In addition, this modification in the petrol mixture has resulted in a considerable reduction in lead levels in São Paulo [CETESB,1993].

During the 1980's there was a substantial change in the vehicle fleet with the introduction of the alcohol engine cars replacing the petrol engine ones. In 1980 the first alcohol engine cars were introduced and by 1989 they accounted for 49% of the fleet. The petrol mixed with 22% of alcohol and the alcohol itself are two fuels with lower pollutant potential. Their pioneering introduction in Brazil in the early 1980's allowed a considerable success in the control of vehicle emissions, comparable to the USA and Europe but in half the time [CETESB,1993].

Table 2.2 - Estimate of emissions from different sources of air pollution in São Paulo, Brazil, circa 1985-90 (in 1,000 ton/year)

		CO	HC	NO_x	SO_x	Particulate
vehicles						
	petrol	835	78	29	5	4
	alcohol	172	14	10	-	-
	diesel[1]	218	36	159	73	10
	others[2]	84	80	2	0.4	7
industry		39	12	14	44	44
incineration		44	14	3	0.4	12
total		1392	234	217	122.9	77

[1] heavy duty vehicles

[2] includes motorcycles, taxis, evaporative emissions, etc..

However, despite the fact that half of the cars run on less-polluting alcohol, currently, there are more than 4.5 million motorised vehicles in São Paulo, a fleet which is growing at the rate of 5% annually [Leitmann,1991]. Efforts for emission reduction have been largely counterbalanced by the increase in the number of vehicles. Not only because there are more cars releasing pollutants into the atmosphere but also because of the increase in frequency of traffic jams and reduction on the average speed of the cars which increase the emissions per km of each individual vehicle.

For example, in 1992, the traffic congestion indicator at peak times averaged 36 km* in 1995 it grew to 94 km and in 1996 it reached 190 km, becoming the second most serious problem identified by the population, after violence. The rate of occupancy is 1.5 persons per car, and the number of cars estimated to circulate per day is 3,200,000 [CETESB,1995; Sobral,1995].

As a consequence of the growth in the number of motor vehicles, in combination with the topography and climatic factors which do not help the dispersion of pollutants, air pollution levels in São Paulo still exceed the international and national air quality standards for most of the air pollutants, for at least a few days a year.

* The Traffic Engineering Company (CET) of the city of São Paulo systematically measures the total length of traffic jams at a number of critical points around the city.

Trends in annual averages of PM_{10} and SO_2 in recent years are presented in Figure 2.4. It can be noted that levels of SO_2 have been declining with no violations of the WHO standards (40-60μg/m^3 for annual average). The trend for SO_2 is comparable with the situation in many other countries where emissions of this pollutant have been at least partially controlled. Meanwhile, levels of fine Particulate Matter (PM_{10}) were stable or gradually increasing after a decline up to 1989. Further, annual means of PM_{10} during this period were 20 to 70% higher than levels recommended by the Environmental Protection Agency (EPA-USA) for annual mean (50μg/m^3).

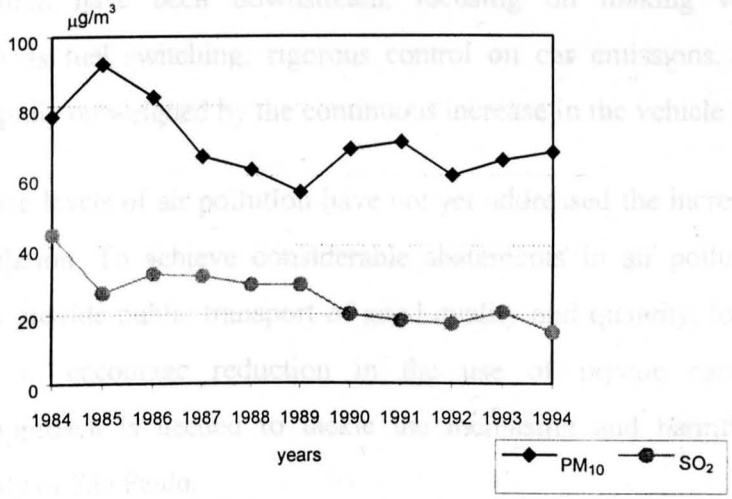


Fig 2.4 - Annual mean levels (μg/m^3) of Sulphur Dioxide and Inhalable Particles for São Paulo, Brazil, 1984-1994.

For O_3 and NO_2 there were no clear general trends as levels for these pollutants have remained fairly constant in the past few years. Yet, violations of guideline standards proposed either by WHO or EPA have been frequent for both pollutants on a daily basis. On the other hand, levels of Carbon Monoxide (CO) have been steadily increasing over the past few years.

Therefore, despite several measures applied to improve the city's air quality, the air of São Paulo still has excessive levels of Particulate Matter, Carbon Monoxide, Nitrogen Dioxide and Ozone which exceed the Brazilian National Air Quality Standards (NAQS) and other international guidelines. These excesses occur every year in at least one or more monitoring sites by a factor of nearly 2 for PM_{10} and Ozone and 1.5 for Oxides of Nitrogen.

As a result, health warnings due to air pollution from carbon monoxide, ozone, and particulate material are frequently issued for the whole metropolitan area of São Paulo [Leitmann,1991]. The harmful and inconvenient effects of air pollution are nowadays a matter of public concern. A household survey showed that air pollution was perceived as having the greatest impact among environmental problems. It is seen as a relevant problem for the majority of the population in the city of São Paulo [Jacobi,1995].

Initially associated with industrial production, which has now reduced its impact significantly, air pollution nowadays is mostly produced by motor vehicles and related activities. Most of the environmental policies and programmes elaborated so far to reduce car-related pollution have been downstream, focusing on making vehicles cleaner. Approaches such as fuel switching, rigorous control on car emissions, etc., have been palliatives and rapidly outweighed by the continuous increase in the vehicle fleet.

Measures to reduce levels of air pollution have not yet addressed the increasing number of vehicles in circulation. To achieve considerable abatements in air pollution levels it is important also to provide public transport of good quality and quantity, to alleviate traffic congestion, and to encourage reduction in the use of private cars. Therefore, a comprehensive approach is needed to tackle the increasing and harmful levels of air pollution in the city of São Paulo.

2.3. A REVIEW OF THE EPIDEMIOLOGICAL EVIDENCE.

The adverse effects on human health associated with air pollution were vividly illustrated following some severe short-term episodes of extremely high levels of air pollution that occurred in Meuse Valley, Belgium in 1930 [Firket,1931], Donora, Pennsylvania in 1948 [Ciocco,1961] and in London, England in 1952 [Ministry of Health,1954] which resulted in serious effects on mortality and morbidity.

In one of these episodes, in London during the winter of 1952/1953, concentrations of SO_2 and smoke (a measure of particulate matter) rose much above 500$\mu g/m^3$. Excess mortality was mainly due to Respiratory and Cardiovascular diseases and despite being much more severe in adults, also affected children [Ministry of Health,1954]. In the same period a huge

increase in the number of applications for the London Emergency Bed Service showed that the effects of air pollution were also observed on morbidity. In addition, studies of that episode showed that the most affected were the age-groups >45 and <5 and especially for respiratory diseases (4 times greater than the usual number of applications) and for cardiovascular diseases (3 times greater) [Abercrombie,1953].

This and subsequent evidence of adverse health effects caused by air pollution led to the adoption of policies such as the Clean Air Acts published in the 1950s in the UK and later in the USA and other western countries. These policies, in parallel with changes from domestic coal heating to gas and electricity, successfully reduced levels of air pollution in several urban areas. Air pollution was then not widely viewed as an important cause of adverse health effects. Studies conducted during the 1960s and 1970s did not provide strong evidence for an effect of air pollution on mortality, and research emphasis shifted to the effect of air pollution on morbidity [HEI,1995].

Initially studies were also more general focusing on excess cases of all cause mortality or morbidity. More recently however, the focus has turned to more specific outcomes like cardiovascular diseases or diseases of the respiratory system for which a biological air pollution link is most plausible. In addition, investigators have been interested not only in high peak episodes of smoke and SO_2 but also on the possible effects on health of long-term exposure to lower levels of an extensive range of air pollutants. On the same lines, attention has been turned to acute effects of air pollution measured on a daily basis as well as effects across different seasons.

Due to the ubiquitous characteristic of air pollution, the exposure of individuals has been generally assessed in an ecological approach, i.e. levels of exposure in the area of residence of the individual are used as the exposure value. This limitation had consequently influenced the choice of design for epidemiological studies. Most frequently, ecological or correlational designs were used although cross-sectional and cohort studies were also employed. A particular type of ecological study, the time-series design, has more recently been widely employed in this field.

2.3.1. Studies of Mortality

After the studies of high level episodes, several other studies were conducted on more common levels of air pollution in order to assess possible adverse health effects. In the following sections, first studies which investigated exposure contrasts between geographical areas are reviewed. Secondly, studies investigating changes in exposure over time at various different resolutions are examined. These studies have also been separated into studies carried out in developed countries and in developing countries.

- **Geographical studies**

Studies of geographical comparisons of mortality rates and levels of air pollutants are typically cross-sectional, comparing rates within or among a number of communities. At the end of the 1970s, an extensive review of studies, mostly correlational or cross-sectional, concluded that acute increases in mortality were only found when daily concentrations of particulate matter (measured as British Smoke) were above 500μg/m^3 and simultaneously having SO_2 around 700μg/m^3 [Holland et al.1979], in other words that the then-current levels of air pollution had little or no adverse impact on health.

Nevertheless, a few years later, some ecological cross-sectional investigations have found associations between mortality rates and sulphate or fine particulate air pollution levels in U.S. metropolitan areas [Ware et al.1981]. Using multiple regression analysis and adjusting for several covariates, positive and statistically significant associations were found. However, the methodological uncertainties in these studies made it difficult to quantify precisely the exposure-response relationships [Ware et al.1981].

In the Czech Republic, infant mortality rates among 46 administrative districts were compared according to their levels of air pollution [Bobak and Leon,1992]. A positive significant association was found between respiratory infant deaths and annual levels of air pollution. The highest to lowest quintile risk ratios for postneonatal respiratory mortality were 2.41 for PM_{10} and 3.91 for SO_2. Levels of PM_{10} averaged 68.5μg/m^3 and SO_2 31.9μg/m^3 among the areas studied. Nevertheless, in addition to its ecological design and inherent limitations, the exposure to air pollution was poorly assessed in this study. Annual averages of air pollution for large areas of the country, sometimes including urban and rural

settlements, were used as a measure of exposure assessment. Nevertheless, it was one of the few studies that focused on infant mortality and air pollution.

These studies have been criticised because in most cases they did not control directly for important confounding variables such as cigarette smoking, socioeconomic status, meteorological variables, and other covariates thought to be important in the relationship between air pollution and health effects. In an attempt to fill such gaps, cohort studies were employed. These studies have the advantage of overcoming some of the limitations of ecological designs.

Two major prospective cohort studies estimated the effect of air pollution in a multivariate analysis while controlling for individual risk factors (cigarette smoking, age, sex, education and occupational exposure). In the Harvard 6 cities study [Dockery et al.1993], long-term exposure to particulate and sulphates (including recurring episodes of relative high pollution) was positively associated with death from lung cancer and cardiopulmonary disease but not with other causes considered together. The adjusted mortality rate-ratio for the most polluted of the cities as compared with the least polluted was 1.26 (1.08-1.47).

Pope [Pope et al.1995] in a larger cohort which represented a wider geographical area of the USA found also that sulphate and fine particulate air pollution were associated with mortality. A difference of approximately 15 to 17% between mortality risks in the most polluted cities and those in the least polluted cities was found. In addition, the authors found that the effect of air pollution was nearly the same in both ever smokers and among never smokers.

However, in both studies no control for temperature or other meteorological variables was performed. Likewise, acute short term effects were impossible to assess since mortality could be the result of a cumulative chronic exposure. They also assumed that recent measurements represented long-term averages. Finally, although individual data was obtained for the health outcome and several covariates, the exposure was still assessed in an ecological approach using measurements from stationary monitory stations.

Studies assessing the health effects of air pollution in developing countries, even using simple ecological design are rare. In a study in Santiago, Chile, geographical comparisons of standardised mortality ratios (SMR) controlling for socioeconomic levels were examined in

relation to air pollution levels [Salinas and Vega,1995]. A clear pattern in the geographical distribution of risk of death, both for all causes and respiratory mortality was found for fine particulate matter and CO. In order words, higher risks of death were observed in more polluted areas of the city.

Of the few studies carried out in Brazil, Penna and Dulchiade [1991] looked at pneumonia mortality in children living in Rio de Janeiro. They used linear multiple regression analysis for the levels of particulate pollution in different areas of the city and their respective infant mortality rate for pneumonia. They found a significant association between infant mortality and annual average levels of total suspended particulate matter (TSPM). However, the ecological design of this study is an important limitation. In addition, annual averages of air pollution were used making it difficult to assess a temporal relationship and to explore short term effects and possible lag periods.

- **Time studies**

Although studies employing different designs can be found in the published literature on air pollution and health effects, it is noticeable the preference of authors for studies looking at the temporal variations in levels of pollution and associated health effects. One of the first investigations of this type was an analysis of the London winter of 1958-59 [Martin,1960]. In this study a significant positive association between the concentrations of Total Suspended Particulate Matter (TSPM) and daily counts of deaths for all causes was observed. A weaker but still significant association was also found for SO_2. However, this study, as well as the analysis of previous episodes, was mostly a correlational study and limited by the state of art of the statistical analysis.

Mazumdar et al [1982], studying 14 London winters of 1958-59 to 1971-72 corroborated previous findings of a positive and significant association between mortality and pollution. They used 3 different types of analysis: year-by-year multiple regression, stratification using nested quartiles of one pollutant within quartiles of the other, and multiple regression of a subset of highly polluted days. Attention was paid to the collinearity between pollutants and to the possible non-linear shape of the association.

In Athens, Greece, a design contrasting high and low exposure days was chosen where 'index days' with high pollution levels and 'control days' with low pollution levels were

matched by temperature, season, and day-of-week [Katsouyanni et al.1990]. Using analysis of variance for randomised blocks the results showed a short term association between air pollution and overall mortality, particularly for respiratory diseases. However, index days were chosen based in the level of one pollutant only (SO_2) making it difficult to discern which pollutant was causing the observed effects. In addition, an assessment of a dose-response relationship was impossible to obtain as well as the examination of lag patterns.

More recently there has been an increasing number of investigations applying time-series methodology. The decrease in levels of air pollution in more industrialised nations and the development of new analytical approaches for time series data has led investigators to examine the relationship between air pollution and health effects in situations where levels of pollution were much lower than traditionally considered at risk. The new methodological tools associated with increasing computational capacity and the low levels of exposure have facilitated the development of time series analysis in air pollution epidemiology.

One of the earlier studies employing such methodology was a re-analysis of the same 14 London winters studied by Mazumdar [Schwartz and Marcus,1990]. Graphical, multivariate time-series, and Poisson analysis was applied to assess the effect of pollution on all cause mortality. The authors concluded that particulates were strongly associated with mortality rates in London. However, when this study included both pollutants in the same model and adjusted for the confounding effects of temperature and humidity the association with particles remained significant but disappeared for SO_2.

This time series analysis was one of the earliest to discuss the problems of collinearity with weather and other long-term variables that could artificially enhance the correlation of air pollution with mortality if omitted. Since then, a large number of time series studies have been conducted in several locations, focusing on different outcomes and agegroups.

Several time-series studies conducted in developed countries were very unspecific and examined the association of air pollution and mortality from all causes excluding accidents and violent deaths in the total population. In general, they found significant associations between mortality and increased levels of air pollution [Schwartz,1993; Schwartz and Dockery,1992a; Schwartz,1991a; Dockery et al.1992; Touloumi et al.1994; Spix and Wichmann,1996; Sartor et al.1995]. The magnitude of the effects found in most of these studies are quite similar. Schwartz conducted two meta-analysis including most of the daily

time-series produced to date. He reported Relative Risks for a 100μg/m^3 increase in Total Suspended Particle (TSP) concentration of 1.06 and 1.09 [Schwartz,1994d; Schwartz,1997] for all non-accidental deaths. He also demonstrated a dose-response relation between the concentrations of TSP and the risk of death, without any evidence of a threshold value.

Other studies explored the association of air pollutants with specific causes of death. They observed that this association seemed to be slightly more pronounced for respiratory mortality [Fairley,1990; Pope et al.1992; Schwartz,1994a; Anderson et al.1996] or even more specifically for pneumonia or COPD mortality [Schwartz,1994e; Schwartz and Dockery,1992b]. In these studies Relative Risks for all non-accidental causes were around 1.06 to 1.16 while for respiratory mortality it varied from 1.18 to 1.43 (for 100μg/m^3 increase in pollution).

Dockery and Pope [1994] also conducted an extensive review of studies on the effects of the respirable fraction of the total particulate matter (PM_{10}) focusing more specifically on respiratory mortality. They estimated that respiratory deaths, which were 2% to 8% of the total deaths in those studies, had increases between 1.5% and 3.7% (weighted mean 3.5%) for each 10μg/m^3 increase in PM_{10}.

Although there is considerably more evidence for effects of particulate matter pollution even when adjusting for the collinearity between this pollutant and SO_2, some studies found also positive and significant associations between SO_2 and daily mortality [Ballester et al.1996; Schwartz,1993; Spix and Wichmann,1996]. Moreover, some studies have also found 'independent' effects of SO_2 when adjusting for particulate [Touloumi et al.1996; Touloumi et al.1994]. Other pollutants like O_3 [Sartor et al.1995; Anderson et al.1996; Burnett et al.1997; Sartor et al.1997] were also found in some studies to have statistically significant associations with daily mortality.

In spite of early evidence indicating that the effect of air pollution was possibly higher in young children and the elderly [Abercrombie,1953], only few studies have focused specifically on children. Therefore, it is difficult to have a more precise measurement of the magnitude of the effect in this agegroup. On the other hand, some studies have shown that the association of air pollution and adverse health effects was stronger for the elderly [Verhoeff et al.1996; Schwartz and Dockery,1992b; Schwartz,1994e; Ballester et al.1996].

For example, Schwartz [1994e] found a Relative Risk of mortality in the elderly of 1.09 in comparison with 1.06 for all ages. In Holland, the risk for all ages was 1.19 while for over 65 years old it was 1.26 [Verhoeff et al.1996].

More recently, a collaborative European project (Air Pollution and Health, an European Approach - APHEA) applied the time-series methodology in a very detailed protocol to investigate the effects of air pollution in 15 different cities across Europe. In summary it was found significant associations between air pollutants (mainly particulate matter and SO_2 and less commonly for O_3 and CO) with mortality for all causes, cardiovascular and respiratory diseases [Sunyer et al.1996; Vigotti et al.1996; Bacharova et al.1996; Wojtyniak and Piekarski,1996]. The effects found for particulate pollution in these studies (PM_{10} or British Smoke-BS) were relatively lower than the ones found in the USA. A meta-analysis of these results showed increases in daily mortality of 2-3% for a 50$\mu g/m^3$ increase in particulate pollution. In addition, the authors found a East-West difference in effects of SO_2 and BS. In Western cities the Relative Risk for an increase in SO_2 or BS was estimated as 1.029 while for Central-Eastern cities the Relative Risk were 1.008 and 1.006 respectively [Katsouyanni et al.1997b].

There are a few time-series studies conducted in developing countries. Ostro [Ostro et al.1996], in a time series analysis controlling for weather and temporal patterns found also an association between PM_{10} and mortality for all causes or for respiratory diseases in Santiago, Chile. A 1% increase in mortality was associated with 10$\mu g/m^3$ change in PM_{10}. Another time series study, this conducted in Mexico City, found also associations of daily mortality and exposure to O_3, SO_2, and TSP (Relative Risks for a 100ppb or 100$\mu g/m^3$ change in pollution was equal to 1.024, 1.024, and 1.050, respectively) [Loomis et al.1996]. Effects in this study were higher for the elderly (RR=1.059 for TSP) and more consistent for TSP in a multi-pollutant model. In China, an effect of SO_2 and TSP on all cause mortality and on cause-specific mortality was found [Xu et al.1994]. The association was more consistent for SO_2, for respiratory mortality, and during the summer.

Saldiva et al [1994] performed a preliminary time series study of respiratory mortality in children under 5 years old and air pollution in the city of São Paulo, Brazil, from May 1990 to April 1991. They found a significant association between daily respiratory mortality and levels of NO_x but not with PM_{10}. Although multiple regression analysis was used, no

adjustment was performed for autocorrelation usually present in this type of data. In addition, adjustment for confounding, especially for temperature and seasonal patterns, was rather crude employing only simple indicator variables. In a second study by the same group, they focus on the effects of air pollution on elderly mortality [Saldiva et al.1995]. It was found that for an increase of $100\mu g/m^3$ in levels of PM_{10}, a 13% increase in all cause mortality was observed. However, as in their previous study, only one year of data was available for analysis and insufficient control for temporal and meteorological effects was undertaken . Moreover, only all cause mortality excluding external causes was explored in this second study.

Neither of the two studies by Saldiva et al looked at indicators of morbidity such as hospital admissions nor attempted to define vulnerable population sub-groups by age or socio-economic status. Therefore there is a clear need to undertake a new study that employs adequate adjustment for temporal and meteorological confounding, uses more than a single year of data to adjust for long-term trends in the data, explores cause-specific as well as all cause mortality, examines morbidity as well as mortality, and attempts to define by age and socio-economic status the population sub-groups which may be most sensitive to short-term health effects of air pollution.

In summary, many studies from developed and (to a lesser extent) developing countries have shown statistically significant associations between daily levels of air pollution and daily mortality for all causes or for specific causes. The magnitude of the effects (expressed as Relative Risks per unit change in ambient concentration) assessed by most time series studies is quite similar. However, when plotting the estimated Relative Risks found in some of the reviewed studies against the annual mean level of particulate pollution in each location there is a suggestion that smaller effects (per $\mu g/m^3$) are observed in places where levels of pollution are higher (Fig 2.5). Since higher mean values of pollution are usually found in cities of developing countries, it could well be that this pattern is attributable to differences between developed and developing countries in weather or temperature influences or, more generally speaking, in population susceptibility.

Although this hypothesis seems rather interesting, the exploratory nature of such a trend analysis should be stressed. The studies summarised in the plot comprised only the ones for which a Relative Risk for all ages and all causes except violent deaths could be calculated

for a $10\mu g/m^3$ change in particulate pollution. Moreover, not all of these studies employed an identical analytical approach and additionally, transformations of measurements of BS and TSP had to be made in order to make them comparable. The linear regression (unweighted) for these 19 points (drawn in the graph) was not statistically significant (p=0.24).

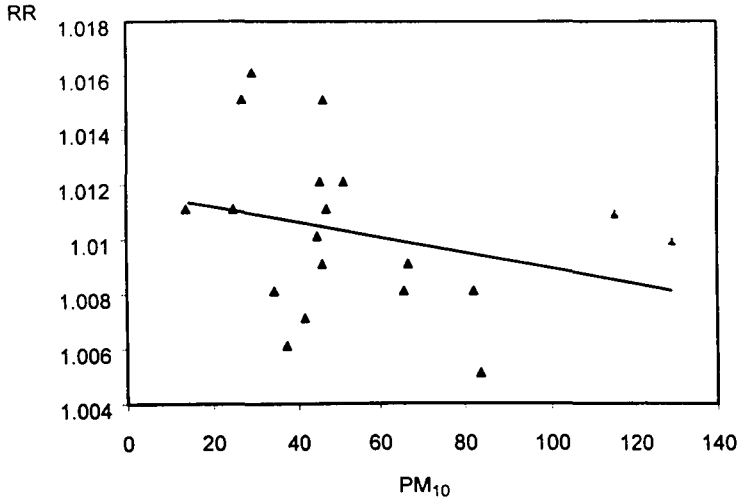


Fig 2.5 - Plot of the estimated Relative Risks of all cause mortality in all ages for a $10\mu g/m^3$ change in levels of PM_{10} from 19 time series analysis against the mean level of PM_{10} in each location during the period of the study. Some studies provided data for Total Suspended Particulate (TSP) or British Smoke (BS). In these cases a transformation to PM_{10} was performed assuming $PM_{10} \cong TSP*0.55$ and $PM_{10} \cong BS$. The two studies in developing countries cities (Santiago and Mexico City) are showed in grey.

Consequently, more evidence on the association between air pollution and mortality is needed, especially from studies conducted in places with higher levels of air pollution and different climate conditions in order to confirm such pattern.

2.3.2. Studies of Morbidity

Morbidity is considered a useful indicator in health studies because it is usually more sensitive than mortality in terms of expressing variations in the health of the population and also because it provides important information for health policy and planning. However, morbidity can be difficult to measure in epidemiological investigations. In the review presented here, attention was generally focused on respiratory morbidity. This was measured by different indicators: reported respiratory symptoms and illnesses, variations in pulmonary

function, and hospital or emergency room admission for respiratory diseases. Studies of daily admissions to hospitals have the advantage of being based on existing data, usually easily accessible. In addition, they make possible the examination of acute relationships with daily variations in air pollution.

Once again studies that looked at morbidity effects of air pollution are divided into studies which investigated exposure contrasts between geographical areas and studies examining changes in exposure over time.

- **Geographical studies**

In one of the first studies looking at morbidity effects of air pollution, Douglas and Waller [Douglas and Waller,1966] followed a cohort of British children living in cities with different levels of air pollution assessed by an index of domestic coal consumption. They found a significant association between incidence of Acute Lower Respiratory Infections (ALRI) and not for Acute Upper Respiratory Infections (AURI) and living in cities classified as highly polluted. Despite its pioneering contribution to the assessment of health effects of air pollution, this study relied on a very crude measure of exposure.

A later study in the UK [Melia et al.1981a; Melia et al.1981b] using cross-sectional and longitudinal design did not find consistent evidence that annual levels of smoke or SO_2 were causing respiratory illness in children. In the cross-sectional study an association was found for smoke after adjusting for a series of confounders. However, those findings were not confirmed by the longitudinal study conducted in the same communities.

Additional evidence of a statistically significant association between air pollution and upper respiratory infections in children have been demonstrated in Finland [Jaakkola et al.1991] using a cross-sectional study. Adjusted odds ratios (OR) for one or more upper respiratory infection of residents in more polluted areas against those living in more clean areas ranged from 1.6 to 2.0 for levels of particulate ranging from 17 to 33$\mu g/m^3$ and SO_2 from 9 to 29$\mu g/m^3$.

The cohort study involving 6 cities in USA also explored the association of air pollution and respiratory effects in school children [Ware et al.1986]. They found that respiratory symptoms, rates of bronchitis and a composite measure of lower respiratory illness were

statistically associated with annual average levels of particulate and SO_2. Illness and symptom rates were higher by approximately a factor of two in the community with the highest air pollution concentration compared with the community with the lowest concentration. A second assessment of the same cohort study [Dockery et al.1989] confirmed the findings of the first study. In both cases, pulmonary function was not related to levels of air pollution.

Few studies have addressed the same question in developing countries using geographical comparisons. In Brazil, Sobral [Sobral,1989] in a cross-sectional study of school children in São Paulo found higher rates of respiratory symptoms in children living in more polluted areas. However, no measure of effect was provided by this study as well as no adjustment for any potential confounders was made.

- **Time studies**

Studies investigating morbidity effects and changes in air pollution levels over time usually refer to panel studies, studies of pulmonary function or analysis of hospital admissions. In Switzerland [Braun-Fahrlander et al.1992], a diary study of children aged 0 to 5 years old showed that the incidence and duration of respiratory symptom episodes were likely to be associated with particulates and the duration of the episodes might also be associated with NO_2.

Decreases in pulmonary function in children associated with levels of air pollution, especially for PM_{10}, SO_2, and O_3 were also reported in short-time follow-up or panel studies in Europe and USA [Roemer et al.1993; Hoek et al.1993; Pope and Dockery,1992]. These studies found statistically significant decreases in values of peak expiratory flow (PEF) or other measures of pulmonary function for moderate levels of PM_{10}, SO_2, and O_3. Peters and colleagues [Peters et al.1996] studying a panel of asthmatic children and adults reported a weak same-day effect and stronger cumulative effect of air pollution on peak expiratory flow (PEF). This study reported that the results were more consistent for SO_2 and Sulphate concentration and that the effects were smaller and less consistent for adults.

As for mortality, several studies applied the time series methodology to assess possible morbidity effects of air pollution. Using school absenteeism as an indicator of health effect, Ransom and Pope [Ransom and Pope,1992] found an statistically significant association

with daily or weekly values of PM_{10} for children in Utah Valley, USA. In this study, absenteeism increased 2% for each $100\mu g/m^3$ increase in PM_{10}. Other study conducted in the same area showed that PM_{10} was also associated with increases in respiratory symptoms and decreases in pulmonary function in school children and asthmatic patients of all ages [Pope et al.1991].

Other time series studies conducted in the Utah Valley showed that levels of PM_{10} were correlated with hospital admissions for respiratory diseases [Pope,1989; Pope,1991]. In addition, these studies found that PM_{10} was more strongly correlated with children's admission than with adults and more correlated with bronchitis and asthma than with admissions for pneumonia and pleurisy [Pope, 1989]. However, no comparable estimates of the effect of air pollution on admission of children was provided making the comparison with adults difficult.

In Canada, Burnett and colleagues [Burnett et al.1994; Burnett et al.1995; Burnett et al.1997] found positive and significant associations between respiratory hospital admissions for all ages and particulate sulphates and O_3. Results were more consistent for levels of pollutants recorded on the same day and up to 3 days prior to admission. The largest impact of air pollution was on admissions for children [Burnett et al.1994], and besides respiratory diseases, for adults there was also an affect on cardiac admissions [Burnett et al.1995].

Several other authors reported associations of ambient levels of air pollution and hospital admissions for adults or the elderly in time-series studies [Ponce de Leon et al.1996; Schwartz,1994b; Schwartz,1994c; Schwartz and Morris,1995; Sunyer et al.1993; Schwartz,1996]. These investigations, however, found more consistent associations with particles and sulphates and less commonly with other pollutants like O_3 and CO. For example, Schwartz found that PM_{10} was a significant risk factor for admissions for pneumonia in the elderly in 2 different cities in USA [Schwartz,1994b; Schwartz,1994c]. Associations were stronger for PM_{10} and with COPD (excluding asthma), but they were also present for O_3. The effects reported in these studies range from Relative Risks of 1.12 to 1.19 for pneumonia admissions and 1.20 to 1.27 for COPD. These RR were calculated for a $100\mu g/m^3$ change in PM_{10}.

After controlling for seasonal and other long-term temporal trends and temperature, PM_{10} was associated with daily admissions for cardiovascular diseases in the elderly in Detroit [Schwartz and Morris,1995]. However, no significant associations were found for SO_2, CO or O_3. In Barcelona, Spain [Sunyer et al.1993], a 6 to 9% increase in emergency room admissions for COPD was observed for a 25μg/m^3 increase in SO_2 or Black Smoke.

Unlike other cities, little or no effect of particulate was observed in a study of respiratory admissions for all ages in London [Ponce de Leon et al.1996]. In this study O_3 levels showed a small but significant association with admissions for respiratory diseases in all ages except for 0-14 years old.

Schwartz [Schwartz,1997] summarised the results of recent studies that have examined the association between air pollution and hospital admissions. The averaged values he calculated represented a meta-analysis weighted by the inverse of the variances of each study. He found that for a 100μg/m^3 increase in levels of PM_{10}, the Relative Risk for respiratory, COPD, and pneumonia admissions for all ages were 1.13, 1.19, and 1.13 respectively. For the same increase in levels of O_3, the risks were 1.06, 1.10, and 1.07. Although some studies have observed stronger effects of air pollution on admission of children, no similar estimates were available for comparison.

Once more, studies of morbidity effects of exposure to air pollution are relatively rare in developing countries. In Mexico, a short follow-up study of pre-school children showed that exposure to high O_3 levels for two consecutive days was associated with a 20% increased risk of respiratory illness [Romieu et al.1992]. In Beijing, China, a time series study suggested that TSP and SO_2 were associated with hospital outpatient visits [Xu et al.1995].

A preliminary report of the effects of heavy industrial pollution on respiratory health of children from Cubatão, Brazil, indicates that levels of PM_{10} were associated with significant reductions in pulmonary function [Spektor et al.1991]. Although effects on pulmonary function in these children were approximately similar to what has been found in Europe or USA, levels of air pollution in Cubatão were much higher than levels experienced nowadays in developed countries. In addition, in this study they found no evidence for a threshold in that the decrease in pulmonary function per unit pollution was similar for less exposed and more heavily exposed children.

No studies to date assessed the effects of air pollution on hospital admissions for respiratory diseases in São Paulo. The only study examining hospitalisations in this setting focused on cardiovascular diseases in the elderly [Rumel et al.1993] and employed very simple analytical techniques. They found that emergency room admissions for myocardial infarction (MI) were significantly associated with temperature and CO.

2.3.3. Main Limitations of Epidemiological Studies on Air Pollution

The primary limitation in environmental epidemiology is the inability or lack of resources to accurately measure environmental exposures in large number of individuals. This explains the widespread use of an ecological approach to assess the exposure of individuals in studies of the health effects of air pollution, no matter which design is used. An average of values obtained from a network of stations located in different sites throughout a specified area is used to express the level of air pollution in that area. Then, all individuals that live in the area are considered to have that level of exposure. It is clear, however, that there are limitations in this approach.

Air pollution is a widespread condition which does not respect any geographical boundaries. Thus, this ecological assessment of exposure does not take into account the differences in levels of air pollution that can exist between places in the same area. It also does not consider the movement of people from one place to another which can result in each having substantially different exposure to air pollution. However, alternative and more accurate methods of assessing exposure are not easily available. For example, personal monitoring is not feasible for large epidemiological studies.

Other important problem in air pollution epidemiology is the fact that pollutants are not completely isolated in the air. On the contrary, they constitute a highly complex mixture where the effects of one can be potentiated by the presence of another. Also the high degree of collinearity between these pollutants implies great difficulties to properly distinguish the ones more hazardous to human health.

But apart from these general limitations pertaining to most epidemiological studies of air pollution, there are other methodological problems specific to each of the most common designs. For example, pure geographical comparisons of health endpoints and levels of air

pollution, even if some adjustment is done, is quite prone to the effect of confounders, the so-called 'ecological fallacy'. Individual data on important confounders such as cigarette smoking, socioeconomic status, age, and etc. can be collected in cross-sectional studies. However, the major methodological limitation of this design is that the temporal dimension of exposure and effect is impossible to be assessed.

In addition, most of the cross-sectional, ecological and cohort studies reviewed here that explored the relationship between exposure to air pollution and health effects could not evaluate the short-term acute effects of air pollution.

To overcome some of these problems and in face of an urgent need to assess the relatively small health effects of air pollution in situations where levels are generally below traditional guidelines, the use of time series methodology has been increasing in epidemiology in the past few years. This approach has advantages such as permitting the search for delayed effects of air pollution and easy adjustment for meteorology and seasonal patterns, variables known to be important confounders in the relationship of air pollution and health effects. However, this methodology requires careful attention to model specification.

Some of the time series studies reviewed here have considered health endpoints that are clearly not Gaussian in their distribution, and concern can been raised regarding the appropriate transformations of the data and specific analytic methods employed. In addition, data that are ordered in time are likely to be correlated, a phenomenon known as autocorrelation. After appropriate modelling there should be no residual autocorrelation in the data. However, only few of the studies presented had tested for this residual autocorrelation or used approaches to adjust for such autocorrelation in the analysis.

The studies considered have also no uniform method for removing seasonal and other long-term patterns from the data before evaluating short-term effects of air pollution. Some have employed complex trigonometric functions while others categorised the data and included indicator variables. Moving averages have also been employed in such studies although there is no consensus about the appropriate time window. Virtually all published studies have adjusted using one or other method, for these patterns which are known to be very important in time series data. In addition, the use of indicator variables is risky since it can lead to either under-specification or over-specification of the model depending on the set of indicators included [Schwartz et al.1996].

Finally, careful attention should be devoted to model the weather effects, especially temperature which are one of the most important covariates to enter the model. A study in Holland [Mackenbach et al.1993] showed that after adjustment for a range of weather variables, the SO_2 effect changed sign into a significant negative effect on mortality. The authors concluded that confounding by weather variables would explain most of the associations between mortality and air pollution. However, as noted by Brunekreef [Brunekreef et al.1995], so many weather variables were entered into the model in that study that the results of the adjusted analysis became rather difficult to interpret.

Temperature is known to be associated with mortality and morbidity [Saez et al.1995; Kalkstein,1993; Kalkstein,1995] and it is also strongly correlated with air pollution. Therefore, it constitutes one of the most important confounders in this relationship. In conclusion, properly adjusting for the effects of weather is vital for any time series study of air pollution and health effects.

2.4. AIR POLLUTION AND THE PATHOGENESIS OF ILLNESS

The health effects associated with exposure to air pollution observed in the preceding review need to be interpreted in the context of toxicological evidence provided by human or laboratory animal studies. Air pollutants present in ambient air can affect human health through different biological mechanisms. In an attempt to assess the biological plausibility that these pollutants can cause adverse health effects, the following sections will discuss what is known about these mechanisms for each of the most common air pollutants.

2.4.1. Particulate Matter

The extent of adverse health effects attributed to particulate matter are believed to be determined by the physical and chemical nature of the particle itself, the physics of deposition and distribution in the respiratory tract, and the biological events occurring mainly in the lungs in response to the particle [Bascom,1996a].

The size distribution influences the deposition in the respiratory system since large particles (> 10μm diameter) are retained in the upper respiratory airways while fine particles (<

10μm) can reach deepest parts of the respiratory system, including the alveolus if the particles are in the ultra-fine range [CETESB,1993]. Once inside the respiratory system, these particles can produce adverse health effects through different mechanisms:

- impairment of the clearance of inhalable particles retained in the lungs,
- carcinogenic effect since the particles may contain or have adsorbed in the surface, carcinogenic substances such as polycyclic aromatic hydrocarbons (PAH),
- ability of the small particles to potentiate the physiologic effects of irritant gases also present in the inhalable air [CETESB,1993],
- possible alveolar inflammation, with release of mediators capable, in susceptible individuals, of causing exacerbation of lung disease and of increasing blood coagulability or plasma viscosity [Seaton et al.1995; Peters et al.1997a].

To explain the toxic effects of particles, more recently, emphasis has been put on the role of particle acidity and the induction of inflammation at sites of contact [Bascom,1996a]. In addition, many of the health effects of particles are thought to reflect the combined action of the diverse components of the pollutant mixture. This combined action is seen as an important mechanism in particulate air pollution since it means that the effects of a mixture of PM and SO_2 for example, are bigger than the isolated effect of each pollutant alone [CETESB,1993; Bascom,1996a].

2.4.2. Sulphur Dioxide

The solubility of SO_2 in the aqueous surface of the respiratory airways controls the amount of this pollutant that is able to reach the deepest parts of the respiratory tract. Because SO_2 is highly soluble in water, it is mostly absorbed in the nose and upper respiratory airways and very little reaches the lungs directly. However, increasing ventilation results in absorption in deeper portions of the lung [CETESB,1993; WHO,1979; Bascom,1996a].

SO_2 is a respiratory irritant that does not cause substantial acute or chronic toxicity in animals exposed to ambient concentrations. When absorbed, however, it can cause a variety of acute responses in cellular elements and sensory nerves ending on or very near the surface of the large airways. This leads to an increase in the airways resistance, increase in the

production of mucus, and irritant responses directly on smooth muscles or via sensory afferent nerve fibres that lead to reflex bronchoconstriction [Bascom,1996a; CETESB,1993].

In addition small aerosol particles containing SO_2 can reach deeper parts of the lungs and may lead to serious damage in the pulmonary tissues [CETESB,1993]. There are also some evidence that it can produce inflammation, increased allergic sensitisation, and produce pulmonary function changes in asthmatics subjects [Bascom,1996a].

It has also been shown that exposure to SO_2 may also increase bronchial reactivity to other agents. The possible mechanism of biological action on the respiratory system may be the impairment of the mucociliary clearance on which cells lining in the respiratory tract sweep out invading bacteria and inert particles from the lung [WHO,1979; Read,1991].

2.4.3. Ozone

Exposure to O_3 produces morphological and inflammatory changes in the lung parenchyma of several laboratory animal species. It can also cause functional, biochemical and structural changes to the small airways analogous to the changes caused by ageing [CETESB,1993; Read,1991]. Increased bronchial reactivity observed after exposure to O_3 may be related to this inflammatory process or to the airway injury [Bascom,1996a].

Pulmonary damages such as emphysema, atelectasis, focal necrosis, bronchopneumonia and fibrosis, accompanied by cellular alterations, have been reported after prolonged and repeated exposure to low levels in laboratory animals. The degree of morphological injury seems to be proportional to the product of the concentration and duration of exposure [WHO,1978].

It appears that exposure to O_3 may lead to impairment of the defence mechanism against disease, making subjects more susceptible to respiratory infections. Decreased resistance to streptococcal pneumonia was evidenced in mice after exposure to various mixtures of NO_2 and Ozone [Erlich et al.1979]. The mechanism of this impairment seems to include both the immune system and the mucociliary clearance [Read,1991]. Studies in human volunteers with controlled exposure to Ozone showed increased airways resistance and decreased ventilatory performance [WHO,1978].

2.4.4. Oxides of Nitrogen

The term Oxides of Nitrogen (NO_x) is understood to include Nitric Oxide (NO) and Nitrogen Dioxide (NO_2). Other oxides of nitrogen which exists in the atmosphere are not known to have any biological significance [WHO,1977]. The predominant oxide of nitrogen is nitric oxide which is readily converted to NO_2 by chemical reactions in the atmosphere. Among the oxides of nitrogen, nitrogen dioxide is the most important in terms of health effects [WHO,1977].

The toxicity of NO_2 is generally attributed to its oxidative capabilities, although as an oxidant it is less reactive than O_3 [Bascom,1996b]. Due to its low solubility in aqueous surface NO_2 is able to reach the deepest parts of the respiratory system. There it can produce nitrosamines which can be highly carcinogenic [CETESB,1993].

Similarly to O_3, exposure to high levels of NO_2 for weeks can injure the smallest air passages of the lung, causing emphysema-like changes in the lungs of animals. In addition, exposure to levels slightly above current ambient levels stimulates the production of inflammatory substances and increases host susceptibility to both bacterial and viral respiratory infection [Read,1991; Erlich,1979; WHO,1977; Bascom,1996b]. This latter effect is clearly dose-related.

The exposure to NO_2 also interferes with the lung's ability to remove inhaled deposited particles efficiently by altering the phagocytic, enzymatic and functional process of the alveolar macrophages and the ciliated epithelial cell [WHO,1977]. Laboratory studies on controlled human exposure provide evidence of an increase in airways resistance and decrease in lung function in asthmatic individuals, who are particularly sensitive to NO_2 [WHO,1977; Bascom,1996b]. Exposures to moderate levels, similar to that occasionally found at the roadside shows that bronchial epithelial cells may be impaired in their function, in theory making the tissue more susceptible to infections [Read,1991].

2.4.5. Carbon Monoxide

The most widely known physiological effect of CO is interference with Oxygen (O_2) transfer. CO has high affinity for hemoglobin (Hb) the O_2-carrier molecule that transports oxygen from the lungs around the body. The combination of CO gas with Hb, leads to the

formation of carboxihemoglobin, COHb, which can no longer transport O_2. Hence the effects of CO exposure are often manifested first in the most oxygen-sensitive organs since the oxygen supply is reduced.

Although CO has a more than 200-times greater affinity for combination with blood hemoglobin than oxygen, the reaction is reversible, and exposure to clean air removes most of the gas from the body [Colls,1997]. High exposures are relatively rare but can cause acute poisoning; coma and collapse begins to occur at COHb levels >40%. However, most exposures to CO in urban settings are several orders of magnitude lower than those associated with intoxication and poisoning. Yet some exposures during urban activity may adversely affect the heart, brain, and the central nervous system. [Colls,1997; Bascom,1996b]

CO competes with O_2 for binding of hemoglobin, but, in addition, it binds other proteins such as myoglobin (Mb) in cardiac and skeletal muscle. A decrease in the concentration of O_2 in these tissues (e.g. during exercise) will enhance COMb formation and can cause CO to shift rapidly out of the blood into the muscle compartment. Sequestration of Mb as COMb might limit the O_2 uptake by these tissues and impair O_2 delivery to intracellular contractile processes [Bascom,1996b].

Individuals with ischemic heart disease (IHD) are particularly sensitive to exposure to CO. Earlier onset of angina pectoris during exercise and signs indicative of myocardial ischemia have been documented in patients with IHD at COHb levels in the range of 2 to 6% [Bascom,1996b]. Other documented outcomes include neurobehavioral effects such as visual and auditory perception, motor and sensorimotor skills, and vigilance [Bascom,1996b]. More recently, evidence was reported for an increase in plasma viscosity similar to that observed for exposure to particulate matter [Peters et al.1997a].

2.5. GENERAL CONCLUSIONS

Urban levels of air pollution are increasing worldwide. While in the developed world there is an emerging problem related to photochemical oxidants and acidic aerosols, in less developed countries problems still exist with classical pollutants such as SO_2 and Suspended

Particulate Matter. As a result of the major short-term episodes of the 1950's and 1960's public health measures have been implemented leading to substantial reductions in air pollution levels in developed countries. However, such measures have not been as strict in less developed countries where environmental pollution has escalated more recently due to increasing urbanisation, industrialisation and car ownership. The gap between developed and developing countries concerning air pollution levels has therefore widened.

The health effects caused by acute short-term and very high episodes have already been clearly demonstrated since their magnitude was relatively high. However, in the last decade, there has been an extensive discussion about the short-term and also long-term health effects due to considerably lower levels, that is lower than the air quality guidelines set up by national and international organisations.

Epidemiological studies have been implicating air pollution with mortality and morbidity effects. There is evidence of short-term pollution-associated mortality for all non-accidental causes, for cardiovascular and for respiratory causes. For morbidity, studies have showed associations with respiratory symptoms, declines in pulmonary function both in children and in adults, and hospitalisation for cardiovascular and respiratory diseases as a consequence of exposure to low-level air pollution.

A plausible biological link between several pollutants and the development of health hazards has already been established. This link is more evident with disorders of the respiratory system since it is the first target of pollutants once they are inhaled. There are also some indications that air pollutants can interfere with blood coagulability which would explain the associations found with cardiovascular diseases. However, the exact mechanisms of such health effects are still not very clear.

The evidence brought by these studies supports the hypothesis that exposure to commonly observed ambient levels of air pollution have an effect on mortality and morbidity. However, the precise magnitude of the effect of air pollution on respiratory diseases remains unclear. In addition, some of the early evidence comes from ecological studies with important methodological weaknesses.

The past few years witnessed the development of new statistical and computational tools that allowed the examination of the relatively small effects of air pollution, and in situations

where levels were frequently below traditional guidelines. The application of new methods, such as time series, in air pollution studies has been contributing to the understanding of the health effects due to exposure to air pollution.

The advantage of time series studies is that individuals are examined over a period of time over which many factors that may confound the relationship between air pollution and health effects are likely to be constant. Therefore, factors such as cigarette smoking, occupation, and indoor residential exposure, cannot confound the short-term temporal relationships because they have no short-term variation in time.

These studies have contributed to the assessment of the magnitude of the associations with air pollution. However, most of them were conducted in developed countries, where levels of pollution have been generally lower than most underdeveloped nations. It is important to have more consistent evidence of the effects of urban air pollutants in the population of cities from the developing world.

In addition, most of these developed countries are in the Northern Hemisphere (USA and Europe) where weather conditions and indoor sources of pollutants, among many other characteristics, are very different from developing countries. The question of whether air pollution levels measured outdoor could be affecting people indoor might be less relevant in cities with warmer climates since homes and buildings are not tightly sealed from the outside.

Moreover, the main source of air pollution is vehicle emissions and much of the expected growth in vehicle number is likely to occur in developing countries [UNEP,1992] which might consequently increase even further their urban levels of air pollution. Therefore, more studies should be conducted in such cities in order to assess the impact of the increasing levels of air pollution. Time series studies usually rely on routinely collected data which is most times available but still under-used in developing countries. Utilisation of such data would improve its quality and help environmental policy formulation for these countries.

Another important reason for studying the health effects caused by exposure to air pollution in developing countries refers to the fact that experimental studies in animals have provided evidence that the effects of such exposure can be aggravated by the poor nutritional status of the exposed subject [Saldiva et al.1992]. That means the effect of urban air pollution might

be worse for developing countries since they have a higher prevalence of malnourished children whose chance of dying might be potentiated by elevated levels of air pollution. Socioeconomic or other differences within individuals might characterise groups with higher risks of effects from air pollution exposure.

Finally, it should be pointed out the relatively scarcity of well designed time series studies that investigated effects on mortality and hospital admissions for children. Most of the published literature had focused on the elderly or, as in many cases, presented results only for total population without investigating this important vulnerable group. Some work has been previously done in Sao Paulo, however, as discussed above, they were limited in the analytical approach and in depth. Only mortality for all causes was studied for the elderly and no previous examination of the effect of air pollution on hospital admissions has been conducted in São Paulo.

Therefore, this study proposes to investigate the effect of exposure to outdoor levels of air pollution in the occurrence of mortality and hospital admissions for children and the elderly in São Paulo, Brazil. São Paulo is one of the largest cities in the developing world and air pollution levels are usually above National and International guidelines. More specific indicators of the health outcomes, a rigorous control of the confounding effects of temporal and weather variables, and more precise indices of exposure to air pollution should aid a more complete assessment of the impact of air pollution on health in this setting.

2.6. HYPOTHESES

- Outdoor levels of Particulate Matter (PM_{10}), Nitrogen Dioxide, Sulphur Dioxide, Ozone, and Carbon Monoxide each are associated with the risk of being admitted to hospital.

- Outdoor levels of Particulate Matter (PM_{10}), Nitrogen Dioxide, Sulphur Dioxide, Ozone, and Carbon Monoxide each are associated with the risk of death.

- The impact of these pollutants on mortality and hospitalisations is stronger for more vulnerable subgroups of the population defined in terms of age and socioeconomic status.

2.7. OBJECTIVES

1. To investigate the association between exposure to air pollutants and the occurrence of deaths and hospital admissions in individuals of different agegroups living in the urban area of the city of São Paulo, Brazil, overall, taking into account categories of cause of death and lag periods between exposure and effect.

2. To identify which of the measured pollutants are more important in the relationship with mortality and hospital admissions.

3. To assess the detailed impact of other potential risk factors, especially meteorological, and adjust for them in the analysis of the effect of air pollution.

4. To identify specific groups that might be more susceptible to such exposures.

5. To compare the magnitude of the effects of air pollution on mortality and hospital admissions in Sao Paulo with what has been found in other cities and with the preliminary study of mortality in children and the limited study in the elderly conducted in Sao Paulo.

CHAPTER 3. METHODS

3.1. STUDY SITE CHARACTERISTICS AND POPULATION

This study was carried out in the city of São Paulo, Brazil, capital of the state with the same name. São Paulo is the most urbanised, industrialised and affluent city in Brazil. It is also the largest among other 38 adjoining municipalities which constitutes the Metropolitan Region of São Paulo (MRSP) or also known as Greater São Paulo. The population of São Paulo, according to the 1991 national census was about 9.5 million inhabitants.

The city occupies a land mass of 1,512 km^2 at an average altitude of 750 metres. It is divided into 58 administrative districts and subdistricts which are the traditional geographical bases for census data and vital registration (Fig 3.1).

Within the city of São Paulo, 43% of its land is residential, 37% is vacant, 9% is commercial, 8% is industrial and the remainder is used for other purposes (recreational, agricultural, etc.) [Leitmann,1991].

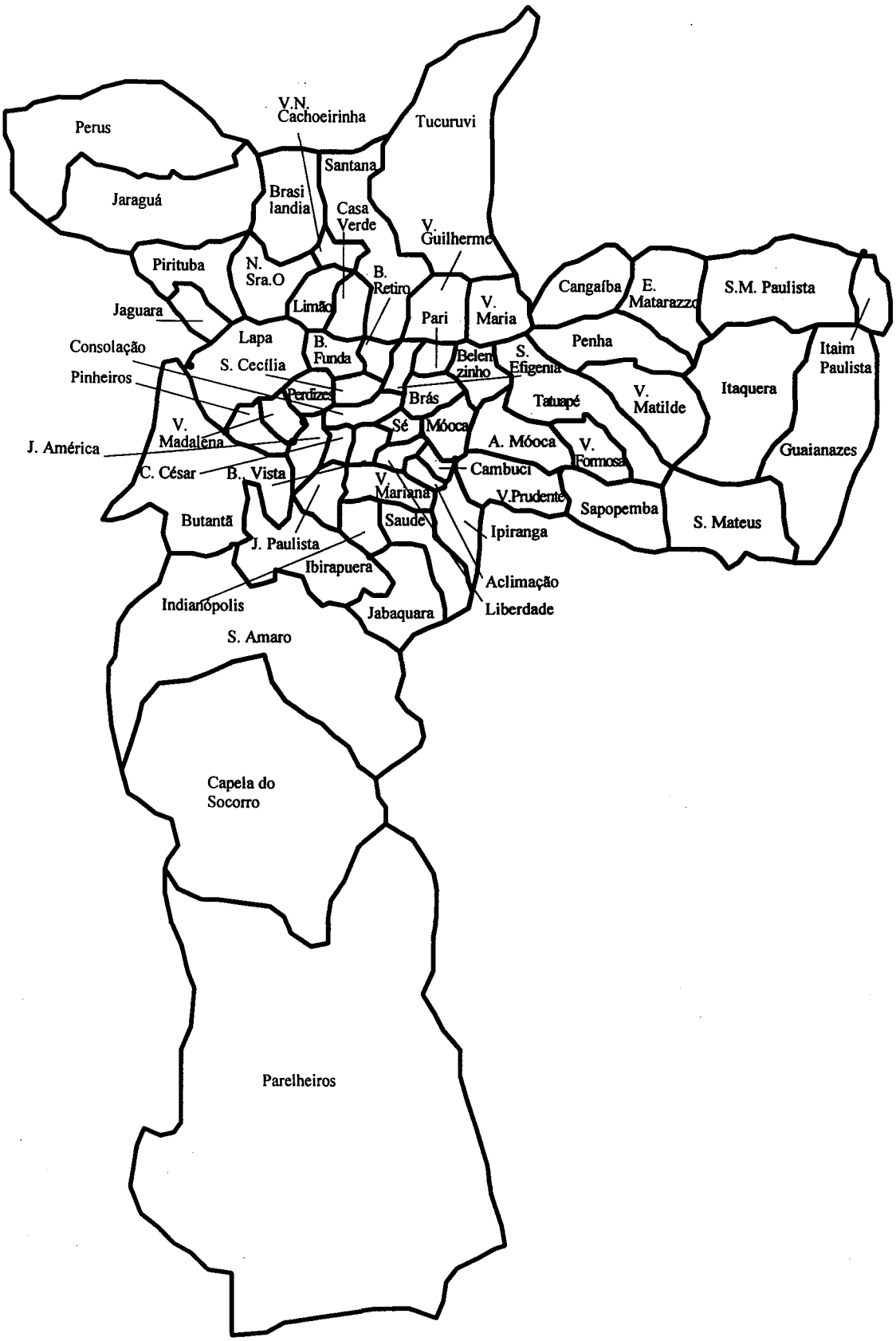


Fig 3.1 - Map of the city of São Paulo showing its 58 districts.

Although there are no available data to quantify the trends in the amount of green areas, there is a perception that they have been diminishing. Nowadays the city has only 2.8% of public green areas in the urban zone compared to the minimum of 12% required by law [PMSP,1992]. Electricity is predominantly supplied by hydropower stations.

São Paulo maintains a comprehensive air quality monitoring network that provides data on all major air pollutants. This network incorporates adequate quality control procedures to ensure that the data are demonstrably valid.

Air pollution monitoring in São Paulo started in 1973 with the systematic measurement of the daily average levels of sulphur dioxide (SO_2) and smoke. Nowadays, the Environmental Department and its Company for Environmental Technology and Sanitation (CETESB) are responsible for the monitoring of air quality through two sampling networks. The manual network has been in operation measuring SO_2 and smoke since 1973, Carbon monoxide (CO) since 1976 and Total Suspended Particulate Matter since 1983. The automatic network in operation since 1981 measures five regulated pollutants (Particulate matter with less than 10μm of diameter - PM_{10}, Oxides of Nitrogen - NO and NO_2, collectively called NO_x, Ozone - O_3, SO_2, and CO) as well as hydrocarbons and meteorological parameters [CETESB,1993]. Within the city of São Paulo there are 13 stations of the automatic network distributed around the most urbanised area (Table 3.1).

The information collected by the equipment in each fixed station is immediately transmitted by private telephone lines to a central station where the data are processed by computer. This system provides daily and hourly information on levels of several pollutants. The location of each station can be seen on the Figure 3.2.

Additional meteorological information concerning temperature, humidity, precipitation, atmospheric pressure and wind is also routinely collected on a daily and hourly basis by the Astronomic and Geophysics Institute of the University of São Paulo.

Data on mortality and morbidity in São Paulo are in most cases available and sometimes readily accessible. Morbidity data from hospitals are particularly adequate. São Paulo has about 150 hospitals with a variety of capacities and specialities. About 66% of the beds in

those hospitals are part of the National Health System and, therefore, are funded by the Government [FSEADE,1993a]. To obtain reimbursement for each admission, the hospitals must submit standardised forms to the State Health Authority. There, these forms are processed in order to provide data for administrative purposes. Among other information those reports contain the date of admission and the International Classification of Diseases, Ninth Revision (ICD-9) code for the discharge diagnosis.

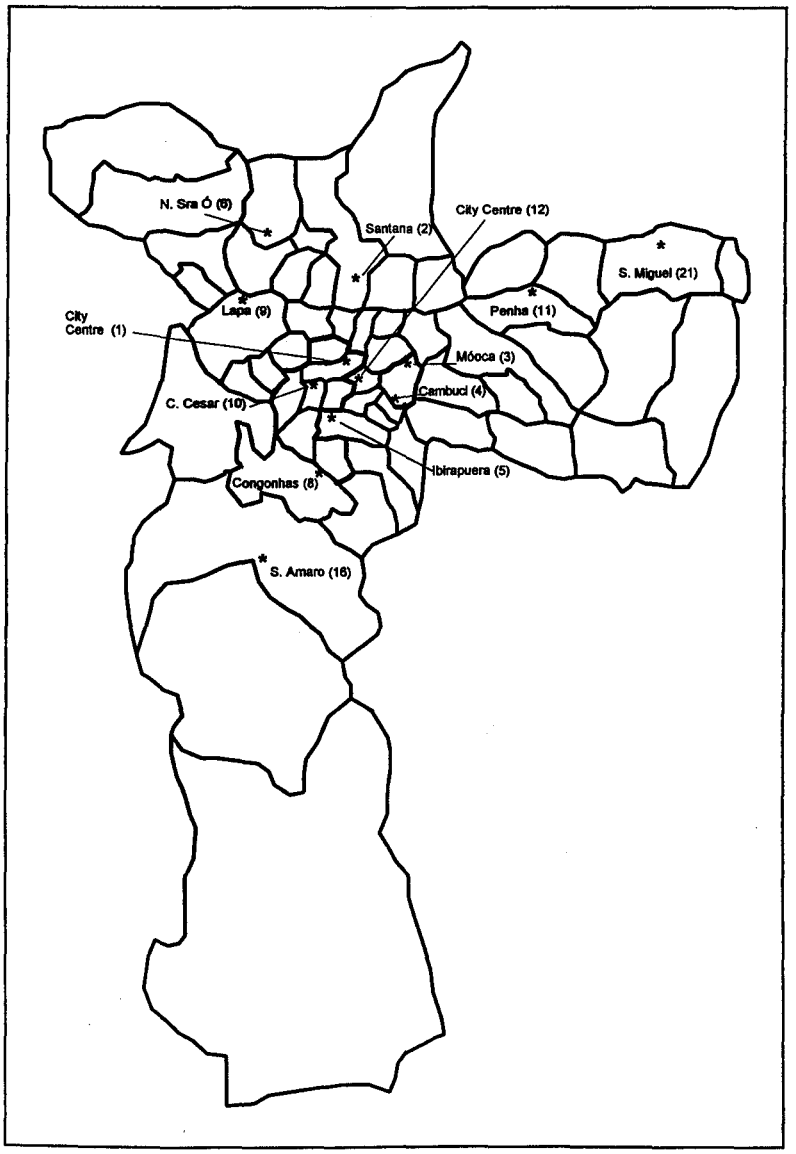


Fig. 3.2 - Map of São Paulo showing the location, name and numeric code of the 13 automatic monitoring stations of air quality.

Table 3.1 - Characterisation of the automatic network of air quality monitoring stations.

station	Name/Area	parameters*								
		PM_{10}	SO_2	NO_x	CO	O_3	HR	Temp	WS	WD
1	City Centre	X	X	X	X	X	X	X	X	X
12	City Centre		X		X					
2	Santana	X	X						X	X
3	Moóca	X	X	X	X	X			X	X
4	Cambuci	X	X							
5	Ibirapuera	X	X						X	X
6	N. Sra. Ó	X	X							
8	Congonhas	X	X	X	X	X				
9	Lapa	X	X			X			X	X
10	C. Cesar	X	X	X	X					
11	Penha	X	X							
16	S. Amaro	X	X						X	X
21	S. Miguel	X	X						X	X

* HR = Relative Humidity; Temp =Temperature; WS = Wind Speed; WD = Wind Direction.

Routine vital statistics in the city of São Paulo are also considered to be complete and of high quality. The great majority of deaths are assigned by medically qualified personnel and under-registration is less than 5%. According to the criteria of the United Nations Organisation, this register can be accepted as complete [FSEADE,1990]. In addition, São Paulo has its own mortality information system called PRO-AIM (Programa de Aprimoramento das Informações de Mortalidade) of the City's Department of Health. PRO-AIM processes and analyses all deaths that occur within the city bounds. The deaths of city residents that occur outside the city can be traced through other systems and, according to current estimates represent only about 6% of all deaths and in most cases they are deaths due to external causes [FSEADE,1993b; PROAIM,1992]. All information contained in each death certificate is computerised and available for public use about 2 weeks after the death.

3.2. STUDY DESIGN

This study investigated the association between outdoor levels of air pollution with variations in mortality and hospitalisations using data available from secondary sources. The principal aim of the study was to explore the temporal aspect of this association. Therefore, time series analyses of daily counts of deaths and hospital admissions were carried out. Time series like this can be seen as a longitudinal study using aggregated data. The same population was used as control and exposed and the unit of analysis was days. Daily data were gathered for several air pollutants, for the health outcomes, and for potential confounders. These data were in most cases collected routinely by public administrative organisations.

In a daily time series analysis of this type, factors such as socio-economic status or occupational exposure are not likely to confound the results, since they are not expected to vary concurrently with air pollution. However, other factors which vary on a day-by-day basis and are correlated in time with air pollution can obscure the estimation of its effects. Two groups of variables fall into this category: meteorological variables such as temperature and humidity and chronological or time-related variables such as day-of-week and other cyclical patterns. These potential confounders were, on the whole, easily measured and were taken into account in the analysis.

The study encompassed all but one of the 58 districts of the city of São Paulo. The district of Parelheiros was excluded because it has a very low population density and it is the one located furthest from a monitoring station of air quality. The outcomes of the study were both mortality (daily counts of deaths) and morbidity (expressed as daily counts of hospital admissions). Hospital admissions were defined as stays of at least 24 hours in a hospital bed. The specific causes of mortality and hospital admissions and the agegroups investigated were:

Mortality

I. elderly (individuals with 65 or more years of age)

- due to all causes excluding external causes (International Classification of Diseases, Ninth Revision - ICD-9 - chapter XVII - Injury and Poisoning)

- due to cardiovascular diseases (ICD-9 chapter VII)
- due to respiratory diseases (ICD-9 chapter VIII)

II. children (up to 5 years of age)

- due to respiratory diseases (ICD-9 chapter VIII)
- due to pneumonia infections (ICD-9 codes 480-487)

Hospital Admissions

I. children (up to 5 years of age)

- due to respiratory diseases (ICD-9 chapter VIII)
- due to pneumonia infections (ICD-9 code 480-487)
- due to asthma and bronchitis (ICD-9 codes 466, 490, 491, 493)

II. infants (up to 1 year of age)

- due to pneumonia infections (ICD-9 code 480-487)

Non-specific diarrhoea diseases for children up to 5 years old (ICD-9 code 009) were also selected for the analysis of hospital admission as a control outcome. Total number of deaths have been considered a suitable health outcome for studies of air pollution. Even though it may be argued that they include a substantial proportion of 'irrelevant' causes, they are completely recorded and avoid the problem of diagnosis or cause of death misclassification [Katsouyanni et al,1995]. In addition, specific causes of death and admission - mainly diseases of the respiratory system and cardiovascular diseases - have been chosen as the most 'relevant' outcomes of air pollution exposure. This was based on biological plausibility as well as on results from previous studies.

3.2.1. Period of the Study

Data on hospital admissions were available for a period of 23 months (November 1992 to September 1994). Data on mortality was compiled retrospectively from computer files for a period of 3 years (1991 to 1993). This procedure was designed to allow for seasonal variations in the incidence of diseases and also to have a sufficient number of events for analysis. Data on air pollution was available for the period 1991-1994 from CETESB. Meteorological data were also collected for the same period. Demographic and socio-economic characteristics of the population under investigation were available from the last census carried out in 1991. These information was assumed to be constant over the entire study period.

3.3. DATA COLLECTION, HANDLING, ENTRY AND CLEANING

3.3.1. Mortality

Data on all deaths that occurred in the city of São Paulo for years 1991 to 1993 were provided by PRO-AIM from routine vital statistics. Copies of all death certificates completed in the city are sent to PRO-AIM on a daily basis where they are processed and checked for discrepancies. Certificates with inconsistent diagnoses are selected and the appropriate hospital/doctor contacted in order to clarify or correct these diagnoses. In addition, coding of place of residence of every deceased is carried out and all this information is entered into computer files.

These files in the form of individual records contained all information available on the death certificate plus some extra provided by PRO-AIM itself. These included the date of death, age, sex, the residential address of the deceased, a code for the district or subdistrict of residence, where the death occurred (hospital, at home, in the streets), and the underlying cause of death coded according to the International Classification of Diseases, Ninth Revision (ICD-9).

Files for the specific years of this study were then assembled and deaths of non-residents, of residents in the subdistrict of Parelheiros as well as of unknown place of residence, were excluded. Moreover, eligible deaths were only deaths from all causes except external

causes (chapter XVII - Injury and Poisoning - ICD-9). This final dataset was then used to construct files with the total number of deaths per day. These files comprised daily counts of all deaths or specific causes of death by agegroups and were the files used in the analysis.

3.3.2. Hospital Admissions

Information on hospital admissions are collected routinely by the State Health Authority. Hospitals which are part of the National Health System produce a form called AIH (Autorização de Internação Hospitalar) for reimbursement for every admission. All forms for a month's period are entered into computer files either at the hospital or at the Health Authority. In either case files for all hospitals are assembled at the Health Authority.

This database contains information on every admission for all hospitals of the NHS in the whole state of São Paulo. Included are the code of the hospital, date of birth, age and sex, date of admission and discharge, how many days spent in intensive care, if died or not during admission, the main diagnosis for the admission (coded according to ICD-9), hospital status (private, public, etc.), type of hospital (paediatric, orthopaedic, cardiac, etc..), and reason for admission (treatment, diagnosis, other purposes, etc.).

A separate file including admissions only for hospitals in the city of São Paulo was produced covering the period between November 1992 and September 1994. In addition, only children under 12 years old admitted to hospitals during the period mentioned above were included. Again, this final dataset was used to construct files where the daily counts of admission per day for different causes and agegroups were used for the analysis.

3.3.3. Meteorological Variables and Socioeconomic Data

Data on meteorological variables were provided by the Astronomic and Geophysics Institute of the University of São Paulo (IAG-USP). Data were available for the whole period of the study and included daily measurements of temperature (mean, maximum and minimum daily level in ^{o}C), humidity (mean, maximum and minimum daily levels expressed as percentages), atmospheric pressure (daily mean in mmHg), precipitation (daily total in mm), wind speed (daily mean in km/h), and wind direction (daily resultant).

Data from the 1991 census and from Municipal and Regional Departments were collected on a number of socioeconomic indicators for each of the 58 administrative districts of the city of São Paulo: mean household income, mean level of education of the head of the household, proportion of households connected to a sewage system, proportion of households with piped water and mean number of individuals per household. In a previous study carried out in São Paulo, Stephens et al. [1994] had generated a composite index based on these five socioeconomic indicators and this was applied in the present study*. The administrative districts of the city of São Paulo were classified into quartiles of the composite socioeconomic index. This classification was then used in the present study to characterise the socioeconomic conditions of each subject based on their district of residence.

3.3.4. Air Pollution

Data on air pollution was available from CETESB, the environmental agency in São Paulo. They provided files containing information for all measured pollutants and for all monitoring stations that belong to the automatic network for the period of the study. These data comprised daily values of Sulphur Dioxide (SO_2), Respirable Particles (PM_{10}), Oxides of Nitrogen (NO_x), Carbon Monoxide (CO) and Ozone (O_3). Data on PM_{10} and SO_2 were available for all 12 and 13 monitoring stations respectively. Other pollutants were only collected in some of the stations (see Table 3.1 for a detailed description of pollutants and monitoring stations).

The information provided for years 1991 to 1993 came directly from the Air Quality Bulletin of CETESB. These data had been pre-cleaned and checked for inconsistencies. A 24-hour period starting at 4 p.m. the previous day up to 4 p.m. on the current day was used to calculate the daily indicators for each pollutant. For PM_{10} and SO_2 the indicators were daily means (24-hour averages) expressed in $\mu g/m^3$. For CO the highest 8-hours moving average expressed in ppm (parts per million) was used. The maximum hourly mean observed during the 24-hour period and expressed in $\mu g/m^3$ was used for O_3 and NO_2.

* For each variable a numerical value was assigned from zero (for the value indicating the worst conditions) to one (for the best conditions). The composite index was the mean of these 5 values.

Although some stations have large amounts of missing data, it was possible to calculate city-wide daily averages for all pollutants for most of the days for the period 1991-1993. For SO_2 and PM_{10} at least 4 monitoring stations were used to calculate the daily average for the city and there were no missing days. For the other pollutants, at least one station was used as the mean value for the city and missing days accounted for 0.6%, 0.8% and 9.1% of the data for CO, O_3 and NO_2 respectively.

For 1994 the data came directly from CETESB and it consisted of hourly measurements for all stations. The same sequence of procedures used by CETESB to obtain the averaged daily values of each station used in the Air Quality Bulletin were then applied. Details of the methods utilised are described in the next section (Quality Control for Air Pollution Measurements). Subsequently, these data with daily values of pollutants for each station were checked for inconsistencies and for the presence of outliers. After this procedure, all the stations were again averaged to provide daily mean values for the whole city.

During 1994 there was a larger proportion of missing data for all pollutants in this dataset provided by CETESB, especially for NO_2. However, only the months of January to September of this year were used for the analysis of Hospital Admissions. Therefore, assembling the periods from November, 1992 to December, 1993, plus the 9 months of 1994 there were 1.1% missing days for PM_{10} and SO_2, 1.6% for CO, 7.7% for O_3 and 46.9% for NO_2.

3.4. QUALITY CONTROL FOR AIR POLLUTION MEASUREMENTS

Due to its vital importance in the air quality monitoring in the city of São Paulo, the measurements provided by CETESB were subjected to a series of quality control procedures to ensure that the results obtained were of the best quality possible. These included inspections of the equipments and following a strict set of standard procedures to calculate the average values for each hour, day and station. In addition, to assess the validity of the data, a comparison was performed with measurements conducted independently.

CETESB have their own methods to ensure the quality of air pollution measurements of their stations. Every 3 months each piece of equipment undergoes preventive maintenance. In the case of PM_{10}, the filters and other components of the equipment are cleaned, tested and some checking procedures are conducted. After that an audit is carried out in order to check if all procedures executed at the service were correct. This includes constructing a calibration curve for each piece of equipment. Lastly, a final quality control check of the data is performed to ensure that the information provided by the equipment is of a high quality. For other pollutants similar procedures are carried out over the same timespan. These activities are conducted by different staff each time in order to avoid systematic errors or bias in the execution of the procedures.

All the 13 monitoring stations in the city of São Paulo have approximately the same characteristics. The measurement points are all at the same height of 2.5m. There are separate inlets for PM_{10} and for gases. In total, 8 of these stations are at street level and practically at the same distance from the street. Four other stations are placed in a semi-background situation, close to streets but in very residential areas. One station is located in a background situation inside a park. The concentrations of SO_2 are measured by colorimetry, CO is measured by non-dispersive infra-red, PM_{10} by a beta gauge aerosol monitoring instrument, and O_3 and NO_x are measured by chemoluminescence.

The data collected at each minute for SO_2, NO_2, CO, and O_3 at these monitoring stations are then summarised as hourly mean values. For PM_{10}, measurements are collected each hour. These measurements have to comply with a defined protocol with criteria for completeness of measurements. For example, for the calculation of 24-hour average levels of PM_{10} and SO_2, it is required that at least 67% of the hourly measurements (16 out of 24) be available. For O_3 and NO_2 the same rule applies with the extra requirement that the 16 valid values of O_3 need to include the daytime period. In the case of CO to calculate each 8-hour moving average it is necessary to have at least 6 hourly values. In addition, in order to select the highest 8-hour moving average, at least 16 valid values are required.

When these conditions are not met the corresponding measurement is set to missing. Additionally, if measurements computed at each minute (or hour for PM_{10}) are below the detection limit for that pollutant, they are set to zero. These values are not discarded and are used for the calculations of daily means. If for PM_{10} the daily average is still below

the detection limit, then that value is set to missing. For SO_2 and CO if daily averages were under the detection limit they are accepted. For NO_2 and O_3 the daily maximum 1-hour value is used, thus daily means are not calculated. This is a standard procedure applied by CETESB and it could have given rise to errors in the calculation of daily values for some pollutants. However, this error is unlike to be differential with respect to rates of disease. Finally, measurements from all the station fulfilling the completeness criteria were averaged.

In order to validate the measurements conducted by CETESB, data from a monitoring station of the automatic network was compared to independent measurements performed by the Air Pollution Study Group of the Department of Physics of the University of São Paulo [Artaxo et al,1994]. They collected data for PM_{10} from June to August 1994, with a 12-hour sampling time, from 8:00 a.m. to 8:00 p.m. daily. The sampler was located about 200 metres away from the CETESB's monitoring stations at a height of about 20 metres.

The sampler used a gravimetric method to measure PM_{10} (GEPA) in contrast to the beta gauge used by CETESB. Moreover, the inlets of the 2 pieces of equipments were different. Nevertheless, results for this sampling period revealed an excellent agreement between the two measurements (Fig 3.3).

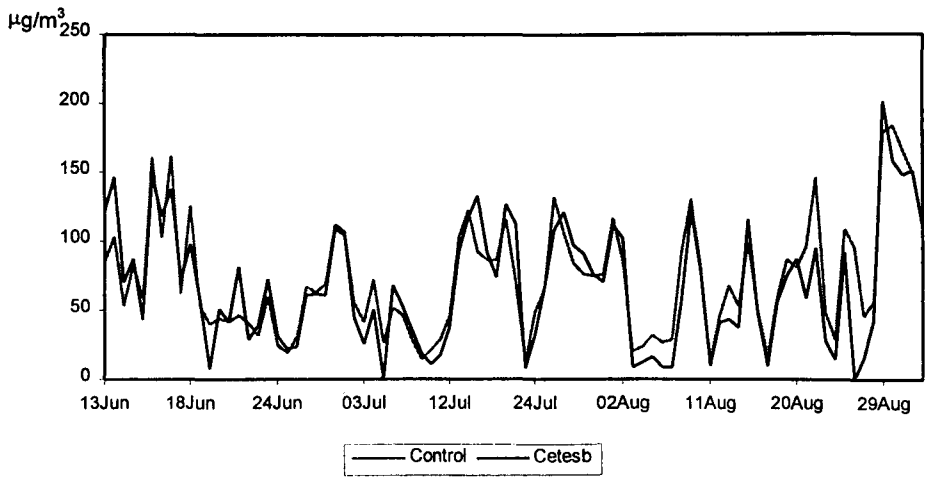


Fig 3.3 - Daily values of PM_{10} measured by CETESB and by the Air Pollution Study Group of the Department of Physics of the University of São Paulo (Control). Daily means ($\mu g/m^3$) were measured for the months of June to August, 1994. Modified from Artaxo et al. [1994].

The regression equation for the measurements for the two instruments calculated by this Study Group [Artaxo et al,1994] was:

$$PM_{10}\ (CETESB) = PM_{10}\ (Control)*0.992 - 2.214.$$

The R^2 value for the regression was 0.86. It can be concluded that this parallel sampling has satisfactorily validated the measurements from CETESB against the gravimetric instrument.

3.5. ANALYTICAL METHODS

The data available for this study can be viewed as a collection of observations made sequentially in time. Data such as this, in which the unit of observation is a period of time as opposed to a person as in most epidemiological studies, constitutes what is called a 'time-series' [Chatfield,1989; Diggle,1990]. One of the approaches largely used in epidemiology to assess the possible relationship between air pollution and health effects is to analyse if short-term fluctuations in one are associated with the other. In other words, to assess if they exhibit a temporal association. In order to ascertain such effects it is more convenient to consider daily observations, for if there are short-term relationships, say an effect of air pollution on mortality lagged by zero, one or two days, these can only be detected through daily analysis. Therefore, time-series analysis of daily air pollution and mortality and hospital admissions were conducted to assess the temporal effects of air pollution on health.

A considerable advantage of the time series approach is that time-constant covariates such as sex, age, occupational exposure, cigarette smoking habits, or any other factor whose distribution does not vary from day to day in the study population, are unable to confound the short-term association between air pollution and health. On the other hand, other influences which are temporally dependent such as meteorological conditions are potential confounders and should be controlled for [Katsouyanni et al,1997a].

Another important issue that has to be addressed when analysing time series of count data is that only a small proportion of the population dies or is admitted to hospital on a given

day. Therefore, these counts are counts of rare events and can be thought of as following a Poisson distribution. A transformation of the data is required to approximate Normality, otherwise use of Poisson models is mandatory. If the number of daily health events is large enough, a Gaussian distribution can be assumed, however, a logarithmic transformation may still help not only to improve the approximation to Normality but also to stabilise the variance [Schwartz et al,1996; World Bank,1995].

In this study, Poisson regression models were used in the analysis. As in classic Poisson regression, the model assumes

$$\text{Log (expected } Y) = \beta_0 + \beta_1 X_1 + \ldots\ldots + \beta_p X_p$$

where Y is the count of deaths or hospital admissions on a given day, $X_1 \ldots X_p$ are the explanatory variables or predictors of Y, and $\beta_1 \ldots \beta_p$ the parameters or regression coefficients for those predictors.

3.5.1. Adjusting for Confounding.

A basic issue in the analysis of epidemiological studies is to control adequately for potential confounders. Time series data such as daily counts of deaths or hospital admissions have some unique features in this regard. For example, they often exhibit substantial long-term trends and/or other systematic variation over the course of a year such as seasonal and cyclical patterns. Since any two variables that show a long term trend or other systematic variation over time must be correlated, when attempting to identify correlations that are more likely to be causal, one needs to exclude this source of variability. In addition, series of health events may also present systematic short-term fluctuations that can bias the results of an analysis. These fluctuations, for example a day-of-week pattern or a public holiday effect, may not be necessarily present in all data, but they occur often enough that they should be checked [Schwartz et al,1996]. Another important source of confounding factors in this type of analysis is meteorological variables. It has been shown, for example, that temperature and humidity have an association with mortality [Kalkstein,1993; Kalkstein,1995; Saez et al,1995; Anonymous,1997]. In addition, periods with high air pollution levels tend to occur under specific meteorological conditions, and hence are also highly correlated with them.

Therefore, when attempting to assess the health effects of exposure to air pollution in time series data it is necessary to remove these patterns from the data beforehand in order to avoid biased estimates.

To account for confounding factors a 'hierarchical' approach was used. Adjustments were first performed allowing for longer term patterns in the data, e.g. trend, then for seasonal and cyclical variations, then for short-term systematic effects (calendar effects) and in a final step adjusting for short-term unsystematic effects (meteorological variables). The confounding variable selection for outcomes that presented a large number of daily events (mortality in the elderly and hospital admissions for children) was conducted using multiple linear regression with the outcome variable being log transformed so as to mimic the scale of a loglinear model. In this situation, the contribution of any additional parameter in the model building was assessed by an F-test at every step, where:

$$F_{p,n} = \frac{\frac{SS_{\text{expanded model}} - SS_{\text{initial model}}}{p}}{MSE_{\text{expanded model}}}$$

with p,n degrees of freedom. In the above formula, SS is the Sum of Squares, p is the number of additional parameters in the expanded model, MSE is the Mean Square Error and n is the degrees of freedom for the MSE.

Some of the outcomes analysed in this study have a low number of daily occurrences (mortality in children). Hence, they present several days with zero occurrences. For these it would be inappropriate to carry out the model building in a multiple linear regression with log transformed outcome. In this situation Poisson regression models were employed. The contribution of any additional parameter included in the model was assessed by examining the deviance/log-likelihood of the extended and the initial models and obtaining χ^2-statistics.

Both statistical tests employed were not used at this stage for hypothesis examination but rather for confounder control. Therefore, a cut-off point of 10% was chosen to decide on

the inclusion or not of any tested variable. This conservative approach was an attempt to ensure adequate adjustment for confounding.

In addition to statistical tests, other analytical and graphical methods were employed at the modelling steps. These included the use of time plots, spectral analysis and periodograms and plots of the autocorrelation function. These methods were used to verify the adequacy of the model and also to update the initial selection of variables used for control of confounding.

- **Adjustment for long-term trends.**

In the scope of this study, long-term trends may be defined as long-term changes in the mean daily mortality or number of hospital admissions over a certain period of time. It is expected that such events be found in populations which are growing in size over a certain period. To adjust for long-term trends, a variable taking the values of 1 to T, where T is the number of the days in the study period, was included in the model. The square of this linear time term was also introduced if it contributed to the model fit. Indicator variables for the years of the study were also examined in order to capture any remaining year-to-year fluctuation not described by the linear and quadratic terms. They were tested and left in the model if they made a significant contribution to the fit.

- **Seasonality and other cyclical patterns.**

To remove the seasonality and other cyclical patterns from the data with a lesser risk of over-specifying the model it was decided to introduce filters into the regression by adding sinusoidal terms (harmonic waves). This can control for fluctuations in an extensive range of periods. While the pattern of mortality or hospital admissions over time is unlikely to be a pure sine wave, the sum of sine waves of increasing frequencies can fit more complex functions [Schwartz et al,1996]. In addition, not all sinusoidal terms begin at zero on January 1st. Hence, if a sine wave of a given frequency is fitted in a regression, a cosine wave of the same frequency must be fitted to adequately account for the phase of the pattern [Schwartz et al,1996].

An annual sinusoidal wave can be described by the function:

$$f(t) = \alpha \sin(2 \pi t/365) + \beta \cos(2 \pi t/365), \quad t = 1 \text{ to } T, \text{ day of study.}$$

The argument $w=2 \pi t/365$ determines the frequency or period of the sinusoidal wave, in this example a cycle of length one year. Further terms with $w=2 \pi t \mathbf{k} /365$, k=2,3 etc. were included to describe other wavelength periods. For this study it was decided to include terms up to the 6^{th} order ($k \leq 6$), as this would pick up periodicities of wavelength down to a two months period. A further cycle of length 2 years was examined for its possible contribution to filtering biannual cycles. This level of filtering was decided based on a previous multicentric study [Katsouyanni et al,1996] which recommended not to go further than two months otherwise one can filter out relatively short term fluctuations including the effects we are seeking to explain in terms of covariates.

The appropriate amount of filtering was confirmed by inspection of periodogram plots of the data. These plots also identified other wavelength periods not a priori covered by the standard functions that needed to be filtered. These additional terms were created but their inclusion into the model was still dependent on statistical tests.

The seasonal and other cyclical variations were thus accounted for by the inclusion of sine and cosine waves spanning periods from two years down to two months. The curves were fitted after the adjustment of the long-term trends. A stepwise approach was used to fit each pair of variables (the sine and cosine wave of the same frequency). When a significant improvement in the model's fit was obtained, the pair of variables was left in the model and a subsequent one, with a higher frequency or shorter wavelength was tested.

- **Short-term (calendar) effects.**

After controlling for long and medium-term fluctuations in the data it is still necessary to adjust for short-term ones. This is in order to separate out those short-term fluctuations due to characteristics of the weekly routine of people and medical care, from those of primary interest. Mortality and especially hospital admissions for example, often exhibit a strong day-of-week pattern. This pattern was modelled on its own by including indicator variables for day-of -week.

examined at the same day and with each series lagged up to five days in order to identify the relative lag at which the two series were most strongly related. Filtering the outcome, and when necessary, including autocorrelated errors beforehand was important because remaining autocorrelation could produce artificial patterns in the cross-correlation [Katsouyanni et al,1997a].

After identifying the best lags, the residuals of the outcomes were then plotted against the four measures of daily temperature at the appropriate lags in order to help determine the exposure-response shape and therefore aid the choice of appropriate transformations which would provide the best model fit. Splines, which are piecewise polynomials, were used as a non parametric smoothed line in these plots to help identify the meteorological relationships. Since this study was not interested in estimating the coefficients for these relationships, precision did not matter and splines were a good way of addressing this issue.

Therefore, based on these plots, different shapes could be observed including linear, two-piece linear, three-piece linear, and non-linear shapes. These functions were then constructed and were tested either in a Poisson or linear regression model of the different outcomes, comprising the adjustments for seasonality, trends and other cyclical patterns. The variable or function that provided the best improvement in fit, i.e., the largest F-test or the maximum change in deviance, whether minimum, maximum or mean temperature, was then selected and included in the core model. This procedure made it possible to restrict the search for the best explanatory variables to control for the effect of temperature on mortality and hospital admissions.

For Relative Humidity, single linear and quadratic terms were constructed for the daily maximum, minimum and mean level. They were evaluated and the one which provided the best fit was chosen regardless of the significance of the coefficient. They were also lagged if the temperature term in that model was lagged. Atmospheric pressure and precipitation were also explored in a stepwise approach using single linear and quadratic terms.

Schwartz and others [Schwartz et al,1996] have observed that these variables should be considered in time series studies even if their inclusion does not have a major impact on the coefficients for pollution. Therefore, for this study they were all included in the model independent of model fit. Further indicator variables for bank holidays, strikes in health services and strikes in the transport system were also tested and included in the model when the improvement in fit was statistically significant.

- **Meteorological variables**

After seasonality and temporal trends, weather conditions are the most important predictors and potential source of confounding in time series analysis of air pollution and health. However, amongst the whole range of meteorological variables, it is widely agreed that temperature and humidity are the ones more likely to be associated with health outcomes [Schwartz et al,1996]. They have been shown to be associated with mortality and hospital admissions and some studies have described that this association is non-linear, or at least piecewise linear usually taking the form of a U-shaped relationship [Kalkstein,1991; Schwartz,1991a ;Schwartz and Dockery,1992a].

In this study the meteorological variables available were temperature (the daily mean, maximum and minimum level), relative humidity (the daily mean, maximum and minimum level), atmospheric pressure, and precipitation. An additional variable was created to account for the daily variations in temperature. This was simply the difference between the maximum and the minimum temperature within a single day.

To adjust for the effects of meteorological variables on mortality and hospital admissions within the scope of this study it was decided to check the contribution to the fit of the model of all the available indicators. However, due to its importance as highlighted in the literature about air pollution and health effects, emphasis was put on modelling temperature and a more comprehensive analysis was performed for that indicator.

To model the temperature effect, initially, cross-correlation functions were produced to characterise at which delay the effect of temperature was greater on the health outcome. The residuals of mortality and hospital admissions after seasonal and calendar corrections were correlated with the residuals of all four different measures of temperature after adjusting also for season and trend using the same set of covariates. Correlations were

3.5.2. Autocorrelation

Another aspect that has to be addressed in the statistical analysis of time series data is the role of serial correlation (also known as autocorrelation) in the outcome variable. Measurements connected in time are likely to be correlated and not independent. This is the case in situations where number of deaths or hospital admissions on two days close together are more alike than in two randomly chosen days. This may be due to imperfectly control for weather or some omitted covariates, producing serial correlation in the residuals of the model [World Bank,1995; Schwartz et al,1996]. It can also be caused by, e.g. a viral epidemic lasting more than two days. Autocorrelation is measured by the correlation coefficient between current observations and observations from p periods before the current one. It can be identified by using an autoregressive integrated moving-average (ARIMA) model that is based on the Box-Jenkins strategy for time series [SAS,1991b]. This procedure can identify the type of the process and estimate its parameters.

The identification of the process is made by means of comparing plots of the data to theoretical patterns for Autoregressive processes. Parameter estimates and diagnoses of the fit are calculated using conditional least-square estimation methods [SAS,1991b].

For most times series data autocorrelation is removed after adequate adjustment for seasonality and cyclical patterns. Remaining autocorrelation can be controlled for by the inclusion of autoregressive parameters in the regression model. Its effect does not bias the regression coefficient, but will bias the estimated standard errors [Schwartz et al,1996; SAS,1991b].

3.5.3. Overdispersion

A distinct issue in the analysis of the effect of air pollution on mortality and hospital admissions is the problem of general overdispersion. This refers to the increase in the variance above the mean of the distribution of the health outcome counts that can occur when the underlying population (here of person-days) is not homogeneous with respect to risk of death/admission. To overcome this problem the estimated variance is inflated by an overdispersed factor. In Poisson regression this factor was obtained by dividing the

Pearson Chi square by its residual degrees of freedom as proposed by McCullagh and Nelder [1989]. For the autoregressive loglinear regression models with Poisson errors a program to deal with overdispersion using generalised estimating equations was used.

3.5.4. Diagnostic Plots

During the model building steps it is important to consider some of the properties of the series as assessed by visual examination of the data. Diagnostic plots provide guidance in the adequacy of the adjustments for long and short-term trends, seasonal and other cyclical patterns, and they proved critical to evaluating the success of the approach adopted. Time plots, spectral analysis or periodogram plots and autocorrelation function plots (ACF) were used throughout this study. They were used in different phases of the model building process to confirm if confounding control was satisfactory and if necessary to revise the initial selection of variables.

- **Time plots**

Visual examination of the original series of data plotted against the time period helped to identify seasonal and other cyclical patterns that needed to be accounted for in the analysis. They were also useful to check if there were any aperiodic patterns or unusual events that could have occurred, e.g. in one year only, which would have been missed if approached by a different method.

Plots of the predicted values over time after regressing a potential confounding factor were helpful in checking if the fitted model provided an adequate description of the data in relation to that specific confounder. They were also helpful to indicate when additional filtering was beginning to predict shorter term patterns that should be left for the explanatory variables in the model [Schwartz et al,1996].

As with most statistical models, the plot of residuals (residual = observation - fitted value) versus time was employed to examine if the patterns seen in the original series have been effectively removed. The examination of the residual time plot, particularly when a smooth curve was also fitted, could often identify if long wavelength patterns remained in the data. Moreover, if structures appeared in both the original series and in the residuals

this would have suggested insufficient fitting. On the other hand, if structures appeared in the residuals but were not present in the original series this would have suggested overfitting [Katsouyanni et al,1997a].

- **Spectral analysis or periodogram plots**

By examining a time plot of the data it is often possible to detect patterns in a time series. However, random fluctuations in the data and combinations of patterns make many cyclical patterns difficult to detect with the naked eye. Spectral analysis is a statistical approach to detect regular cyclical patterns or periodicities, in transformed time series [SAS,1991b]. It consists of decomposing the data into a sum of sine and cosine waves of various wavelengths and amplitudes. Then, the time series is regressed on the sine and cosine components for various frequencies and the sum of squares for each regression model is calculated. The plot of the sum of squares from each regression model against the frequencies, ω_k is called the periodogram [SAS,1991b]. These sums of squares are called periodogram ordinates. They can be interpreted as the amount of variation in Y at each frequency. The periodogram, therefore, provides a summary description of the data showing the amount of variation that occurs in a given frequency range.

The dominant peaks in a periodogram indicate cyclic patterns in the data, while the position of the peaks suggests the frequency at which these occur. Large peaks at frequencies well above 1 year indicates the presence of long-term trends in the data. Peaks around 365 days indicates a one year cyclical pattern. Evidence of a day-of-week effect is exhibited by large peaks at frequencies around 7 days.

During the model building process, the periodogram was useful in checking whether specific periodicities had been adequately accounted for and if remaining pattern appeared to be random noise.

- **Autocorrelation (ACF) and partial autocorrelation function (PACF)**

Usually, three types of processes can be identified underlying a time series data. They can be Autoregressive-AR, Moving Average-MA or mixed ARMA-process. Generally, autoregressive processes (AR) are more commonly found in biomedical phenomena [SAS,1991b; Schwartz et al,1996]. To identify the type of the process, and more

importantly, the order of the process underlying the data in this study, autoregressive integrated moving-average (ARIMA) models that are the basis of the Box-Jenkins strategy for time series modelling [SAS,1991b], were used.

This procedure computed the autocorrelation function (ACF) and the partial autocorrelation function (PACF), producing plots and statistics for groups of correlation coefficients. The ACF plot or correlogram is the plot of the autocorrelation coefficients at each lag unadjusted for previous lags. The PACF describes the serial correlation of a time series at lags 1, 2, etc., with the value at each lag corrected for the previous lag [Katsouyanni et al,1996]. It is usually difficult to assess the order of an AR process from the ACF alone and the use of the PACF plot is mandatory.

The ACF of an AR(x) process is a mixture of damped exponential and sinusoids and dies out (or attenuates) slowly. The PACF of an AR(x) process 'cuts off' at lag x so that the 'correct' order is assessed as that value of x beyond which the sample values of the autocorrelation coefficients does not show an statistically significant difference from zero [SAS,1991b]. In other words, if the first x values of a PACF-diagram of appropriately detrended data are greater than two standard errors for the sample autocorrelation, this points to an AR(x) model, i.e. a model where the value of a given day is (linearly) dependent on the values of the last x days [Chatfield,1989; Spix and Wichmann,1996].

The visual examination of these plots helped to provide evidence for the presence or absence of an autocorrelation coefficient greater than two standard errors at a certain lag. In addition, the ARIMA models provided an estimate of the parameter for the autocorrelation at that lag and checked whether the residuals became random noise when this parameter was incorporated into the model.

After control for season and trend, the magnitude of the serial correlation in the residuals of mortality or hospital admission data was low since strong autocorrelation is also a consequence of inadequate control for temporal confounding. This was evidenced by decreasing levels of autocorrelation, as showed by ARIMA models applied in the residuals at each step in the model building.

The residuals of the final adjusted model for each outcome in this study were inspected for remaining autocorrelation using ARIMA models and PACF plots. When it was present, an estimate of the parameter for the autocorrelation was obtained and then incorporated into the final Poisson models.

3.5.5. Model Building Process

Different outcomes were examined in this study. These outcomes were counts of daily mortality or hospital admissions for distinct causes and agegroups. For each different outcome the model was fitted in a stepwise forward strategy. Thus, terms to adjust for potential confounders were included first in order to obtain a baseline or 'core' model. The contribution of air pollution variables to the prediction of daily mortality or hospital admissions was only examined after establishing these core models.

Variables were added to the core model according to the statistical procedures described in section 3.6.1. Periodograms, plots of residuals and of predicted values were used at each step to check the adequacy of the modelling. When the long term trends, seasonality, other cyclical patterns, and meteorological variables were adequately controlled for, autocorrelation and partial autocorrelation function of the residuals were examined to assess the presence of any remaining autocorrelation in the data. If no autocorrelation was present, then Poisson regression models allowing only for general overdispersion were used to estimate the effect of air pollutants. If there was autocorrelation, the order of the process was established and autoregressive errors were implemented in the regression model. In this case the model was based on statistical methods developed by Zeger and applied to the analysis of air pollution by Schwartz and others [Schwartz,1991a; Schwartz,1993; Schwartz,1994c; Spix et al,1993; Dockery et al,1992]. In summary, this model is an autoregressive log-linear regression with Poisson errors which also deals with overdispersion. In any case, air pollution variables were only included in the models after a baseline or 'core' model was established. Regression coefficients and 95% confidence intervals for each specific exposure were then obtained.

All the statistical procedures were implemented using the statistical program SAS with special emphasis on usage of the Econometric and Time Series (ETS) module.

CHAPTER 4.
DESCRIPTIVE RESULTS

4.1. AIR POLLUTION

Daily city-wide levels of air pollution for the city of São Paulo were obtained by averaging the values from all the monitoring stations that provided valid values for that pollutant and on that day. Daily values for the period 1991-1993 were used for the mortality analysis. The analysis of hospital admission data covered the period of November 1992 to September 1994. For PM_{10} and SO_2 there were valid values for most of the study period (1369 days) with only 8 days missing (0.6%). NO_2 was not available for 25% of the days whereas CO was missing for 1.2% and O_3 in 4.2% of the days (Table 4.1).

Levels of air pollution in São Paulo were relatively high for some pollutants during the period of this study. Inhalable particles (PM_{10}), for example, exhibited on average a value

of 65.1 $\mu g/m^3$ for this 4 year period. Although annual means have lately been between 20% to 40% higher than the values proposed as guideline by the Environmental Protection Agency-USA (Table 4.2), only 3.2% of the days were beyond the daily limit of 150 $\mu g/m^3$.

Table 4.1 - Descriptive statistics for the air pollution data

			Percentiles						
	n° of days	mean (SD)	min	5	25	50	75	95	max
PM_{10}	1361	65.1 (30.9)	18.4	31.6	43.9	56.6	77.8	130.8	231.8
SO_2	1361	18.2 (8.5)	3.2	7.7	12.1	16.6	22.3	34.7	61.1
NO_2	1027	165.3 (89.1)	26	67.5	104.4	144	204.1	326.4	692.9
O_3	1312	65.5 (41.9)	2.7	14.4	34.2	56	87.6	146.3	272
CO	1353	5.8 (2.2)	1.3	2.8	4.3	5.6	7.0	9.5	22.8

For O_3 the recommended maximum level was exceeded on about 18% of the days if we consider the lower limit of 100 $\mu g/m^3$ or only 1% for a limit of 200 $\mu g/m^3$. In the case of CO, although we have seem that in São Paulo the levels for this pollutant have been rising in the last 11 years, only 7% of the days showed a maximum 8-hours above the recommended daily maximum of 9 ppm. On the other hand, levels of NO_2 are relatively much higher than recommended annually by EPA and hourly by WHO. Annual levels were 50 to 75% higher than EPA guidelines and daily maximum levels exceeded the WHO guideline (190$\mu g/m^3$) in more than 20% of the days.

Table 4.2 - Recommended guidelines for air pollution levels

	Annual maximum	Daily maximum	Source
PM_{10}	50	150	EPA-USA
SO_2	40-60	100-150	WHO
	80	365	EPA-USA
CO	--	9 ppm (8 hours)	EPA-USA
NO_2	--	190-320 (1 hour)	WHO
	100	--	EPA-USA
O_3	--	100-200 (1 hour)	WHO
	--	235 (1 hour)	EPA-USA

EPA = Environmental Protection Agency
WHO = World Health Organisation

Looking at the daily variations in pollution levels during this 4 years period we can notice that PM_{10}, SO_2 and NO_2 exhibit a strong seasonal variation (Fig 4.1). For these pollutants levels are generally higher during winter periods (June to September). It can also be seen that levels of PM_{10} were comparatively higher during the winter of 1994 when reports of deteriorating levels of air pollution started increasing in the media. There was no available information on levels of NO_2 for the last 5 months of the period of this study. This was due to equipment being out of order.

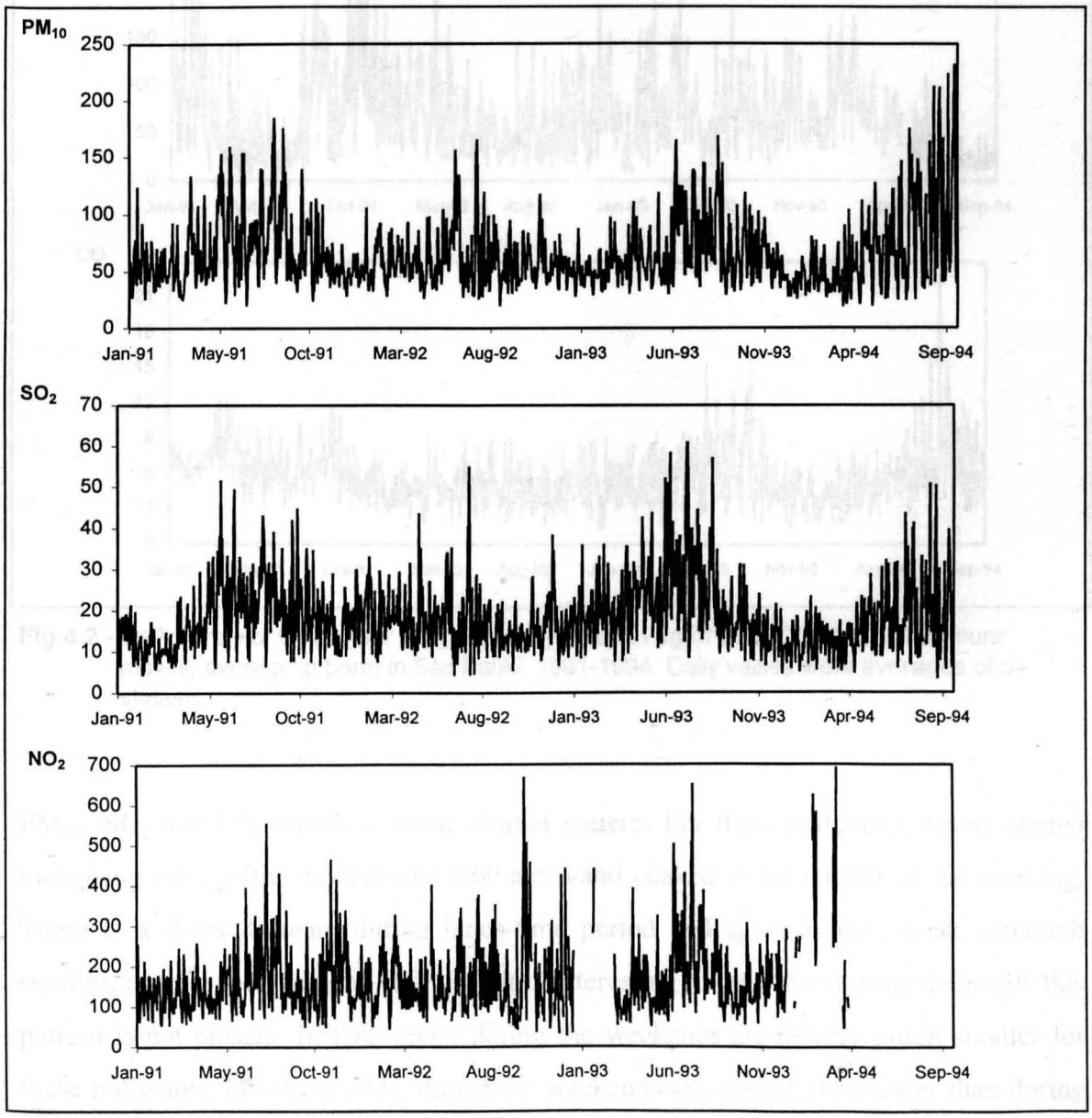


Fig 4.1 - Daily levels of PM_{10} and SO_2 (24-hours average in $\mu g/m^3$) and NO_2 (1-hour maximum daily value in $\mu g/m^3$) in São Paulo, 1991-1994. Daily values were averages of all stations

In Figure 4.2 it can be seen that levels of Carbon Monoxide did not present any strong seasonal pattern across the years. Nevertheless, its levels were on increase throughout the period with also much higher levels during the winter of 1994. On the other hand Ozone showed a seasonal pattern but in an opposite direction to other pollutants. Its levels were higher during summer periods when sunshine is more abundant.

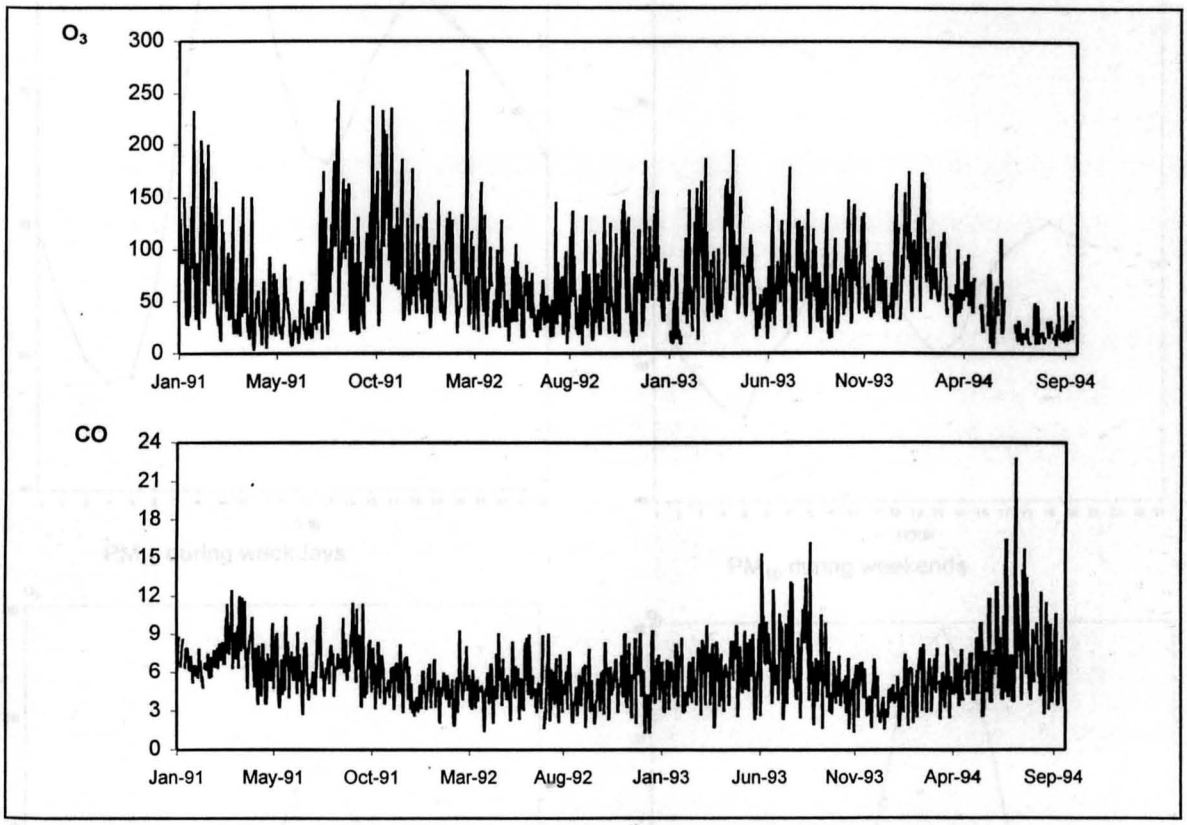


Fig 4.2 - Daily levels of O_3 (1-hour maximum daily value in $\mu g/m^3$) and CO (highest 8-hours moving average in ppm) in São Paulo, 1991-1994. Daily values were averages of all stations

PM_{10}, SO_2 and CO exhibit a strong diurnal pattern. For these pollutants, levels started increasing early in the day (around 6:00 a.m.) and peaked in the middle of the morning. There is a slight decrease during lunch-time period and again another peak, although smaller, later in the afternoon (Fig 4.3). It is interesting to note that during weekends this pattern is not present. In fact, levels during the weekends are usually much smaller for these pollutants. Levels of PM_{10} during the weekends are around 30% lower than during the middle of the week. This is probably because these pollutants are strongly traffic-related. For NO_2 there is also an increase from early hours in the morning but after

peaking the levels stay in a relatively flat plateau up to 8:00 p.m. and then starting decreasing.

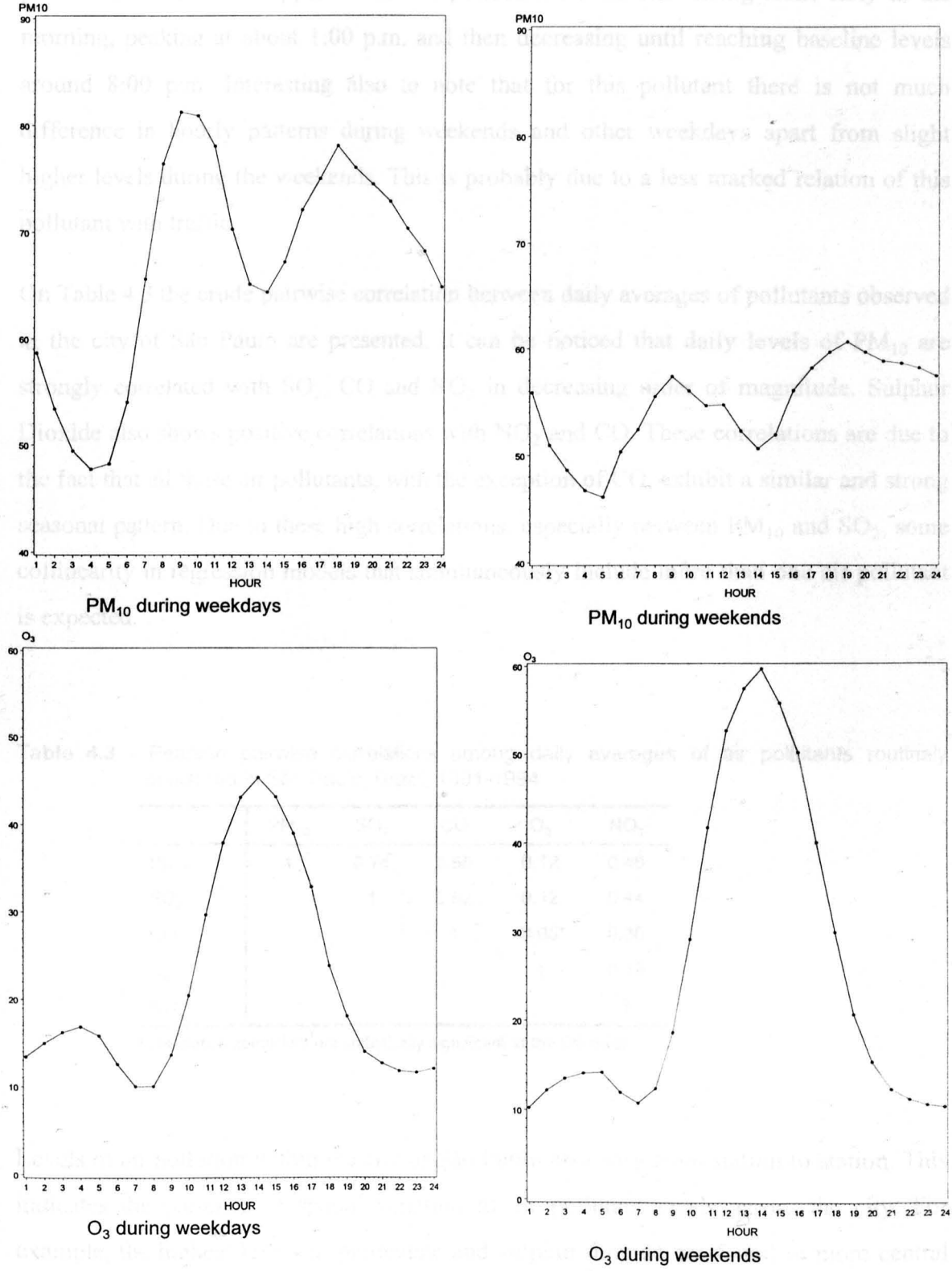


Fig 4.3 - Hourly levels of PM_{10} and O_3 in São Paulo, for weekdays and weekends averaged over all days between 1991-1994.

As O_3 is a photochemical pollutant, its presence in the atmosphere depends on the presence of daylight and it is expected that levels should rise as soon as sunshine is present. That is what happens with this pollutant. Levels start rising from early in the morning, peaking at about 1:00 p.m. and then decreasing until reaching baseline levels around 8:00 p.m. Interesting also to note that for this pollutant there is not much difference in hourly patterns during weekends and other weekdays apart from slight higher levels during the weekends. This is probably due to a less marked relation of this pollutant with traffic.

On Table 4.3 the crude pairwise correlation between daily averages of pollutants observed in the city of São Paulo are presented. It can be noticed that daily levels of PM_{10} are strongly correlated with SO_2, CO and NO_2 in decreasing order of magnitude. Sulphur Dioxide also shows positive correlations with NO_2 and CO. These correlations are due to the fact that all these air pollutants, with the exception of CO, exhibit a similar and strong seasonal pattern. Due to these high correlations, especially between PM_{10} and SO_2, some collinearity in regression models that simultaneously include more than one air pollutant is expected.

Table 4.3 - Pearson pairwise correlations among daily averages of air pollutants routinely measured in São Paulo, Brazil, 1991-1994

	PM_{10}	SO_2	CO	O_3	NO_2
PM_{10}	1	0.74	0.55	0.12	0.46
SO_2		1	0.52	0.12	0.44
CO			1	-0.05*	0.30
O_3				1	0.19
NO_2					1

* correlation coefficient not statistically significant at the 5% level

Levels of air pollution within the city of São Paulo also vary from station to station. This indicates the presence of spatial variation in air pollutant levels across the city. For example, the highest levels of particulate and sulphur dioxide are found in more central areas of the city (stations 10-C. Cesar and 8-Congonhas) where traffic of vehicles are more intense or in more industrial areas (stations 9-Lapa and 3-Móoca). Lower levels for

these pollutants can be observed in monitoring stations located at more peripheral areas (stations 21-S. Miguel, 11-Penha and 6-N. Sra. Ó, and 16-S. Amaro). These stations are situated close to streets but far from traffic, therefore, for this study they are considered semi-background stations.

Differences in measured levels of air pollution among monitoring stations can also be attributed exclusively to their different location. Monitoring stations at the street level are more likely to exhibit higher levels of air pollution, especially if the street where they are placed has heavy traffic of vehicles. Background stations, on the other hand, are expected to present lower averages of air pollution. The only one background station in São Paulo (Ibirapuera) is centrally located but is situated inside a park, more than 100 metres away from the nearest road.

In Table 4.4 the averaged values of air pollution for all the stations and for the station with the highest and lowest levels during 1991-1994 are shown. In addition, mean averages for the roadside, semi-background and background station are presented for PM_{10} and SO_2 to illustrate the differences in levels of pollution according to the location of the station. Ozone, NO_2, and CO are measure by fewer stations and they are all located at the roadside level.

It can be observed that the differences in mean levels between the station providing the highest and lowest values for the same pollutant can be as high as a 3.3 fold difference for SO_2 to only 1.4 fold difference for Ozone. As expected, the roadside stations provided the highest measurements for SO_2 and PM_{10}. Semi-background and the background stations exhibited lower average levels with the background being slightly lower than the mean for the semi-background stations. Nevertheless, the station providing the lowest value for both SO_2 and PM_{10} is not the background station but one of the stations located at semi-background situation.

Table 4.4 - Means of air pollution for the station with the highest and lowest value for measured pollutants in São Paulo, Brazil, during the years 1991-1994.

Air Pollutant	stations	Mean
CO	all stations	5.8
	lowest (3-Móoca)	3.0
	highest (10-C. Cesar)	6.9
NO_2	all stations	165.3
	lowest (3-Móoca)	93.2
	highest (1-City Centre)	172.1
O_3	all stations	65.5
	lowest (1-City Centre)	58.7
	highest (3-Mooca)	81.7
SO_2	all stations	18.2
	roadside	22.4
	semi-background	12.2
	background	10.7
	lowest (21-S. Miguel)	9.1
	highest (4-Cambuci)	30.4
PM_{10}	all stations	65.1
	roadside	71.0
	semi-background	58.3
	background	52.6
	lowest (21-S. Miguel)	48.9
	highest (10-C. Cesar)	90.8

Differences in spatial distribution of each air pollutant can also be appreciated from pairwise correlations among monitoring stations for the same pollutant. For example, the measurements of particulate matter among monitoring stations are highly correlated over time (mean correlation coefficient of 0.62 between the 12 stations that measure this pollutant). In contrast, SO_2 exhibits much greater variation from one station to the other. The mean correlation coefficient for this pollutant is only 0.35 between 13 stations. This is an indication that PM_{10} is more uniformly distributed around the city's urban area whereas SO_2 is more dependent on local sources. Other pollutants showed an intermediate correlation. However, they were measured in fewer stations.

4.2. MORTALITY

Daily records of all deaths were obtained from PROAIM - The São Paulo Municipal Government Death Registry. During the period of 1991 to 1993 there were 182,611 deaths registered in this death registry including all causes and ages. This provides an annual mean crude mortality rate of 6.43 per 1000 individuals. From these deaths, nearly 0.5% (955) had missing information for age and 4.7% (8,518) had unknown place of residence. These observations were therefore excluded. Additionally, deaths due to external causes (injury and poisoning - chapter 17 of the ICD-9) which represented 14.2% (25,984) of the total deaths were later also excluded from the dataset used for analysis. Consequently, a final number of 151,756 deaths were available for the period of the study.

The proportional distribution of these deaths according to agegroup is presented in Figure 4.4. It can be noted that, as expected, the majority of deaths (43.2%) occurred in the elderly group (individuals aged 65 years or older).

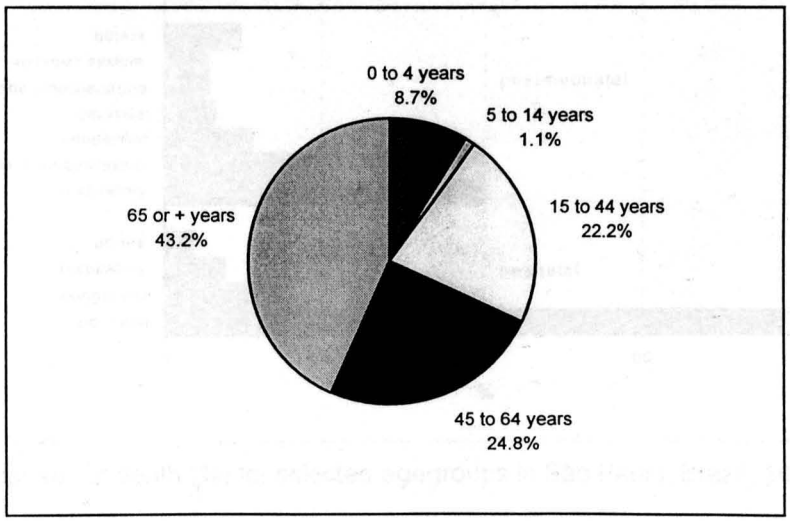


Fig 4.4 - Proportional distribution of deaths for all causes among different agegroups, São Paulo, Brazil, 1991-1993

The distribution of causes of death is also distinct for each different agegroup. Among children who die up to 28 days old (neonatal deaths) nearly 80% were due to conditions which originated during the perinatal period. The second main cause for this agegroup were the congenital anomalies. In post-neonatal deaths (after 28 days and up to one year

old) and in pre-school children (1 to 5 years old) respiratory diseases were the most important cause of mortality followed by infectious and parasitic diseases. The most important causes of deaths for all different agegroups are summarised in Fig 4.5.

For school aged children (5-14 years old) the overall number of deaths were very low. Nevertheless, for this agegroup and for others up to 65 years of age the main cause of mortality was the external causes (injury and poisoning) with usually respiratory diseases as the second most frequent cause.

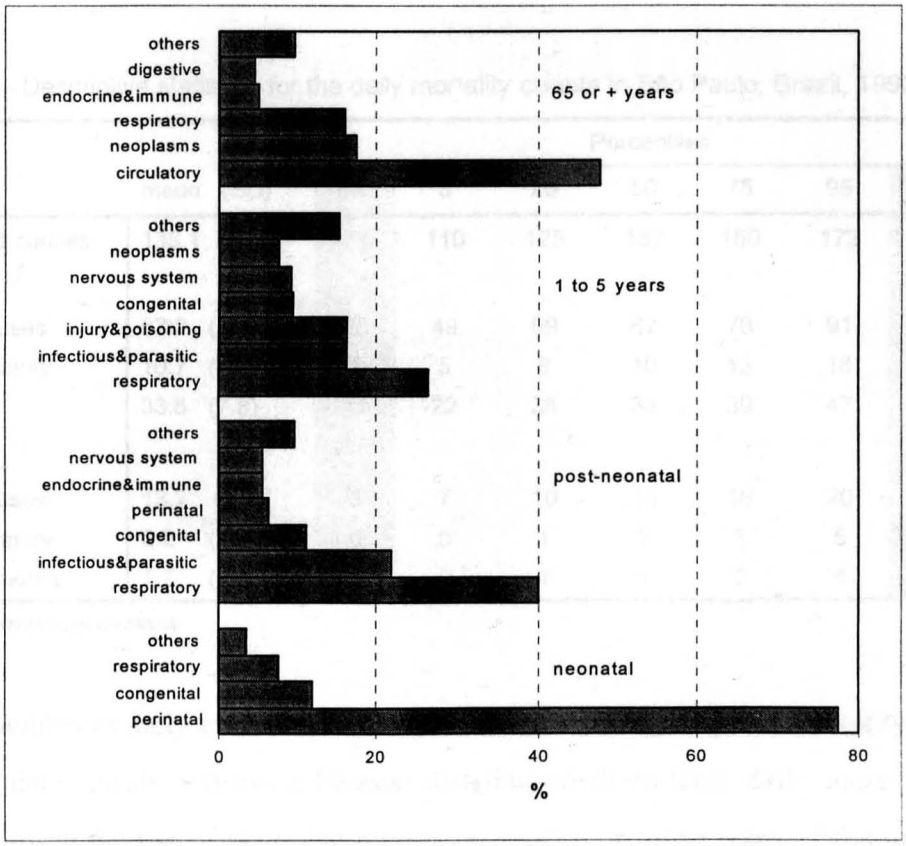


Fig 4.5 - Main causes of death (%) for selected agegroups in São Paulo, Brazil, 1991-1993

For elderly people (aged 65 years old or more), the diseases of the circulatory system represented the most prevalent group with 49% of all mortality. These were mainly cardiovascular diseases (ischaemic heart disease and cerebrovascular disease). The second group of causes of death was neoplasms with 17.8% and the third was diseases of the respiratory system with 16.3%. It should be mentioned that pneumonia infections forms

the bulk of respiratory causes for all agegroups with an average of 80% of all respiratory deaths.

Selected aspects of daily mortality in São Paulo for the period 1991-1993 are shown in Table 4.5. The median daily number of deaths for all causes (excluding external causes) in the total population was 137, of which 67 were among individuals 65 years or older, and 13 were among children under 5 years of age. In elderly people, cardiovascular diseases accounted for nearly half of the daily deaths whereas among children only about 20% of all daily deaths were due to respiratory diseases.

Table 4.5 - Descriptive statistics for the daily mortality counts in São Paulo, Brazil, 1993-1994.

		Percentiles						
	mean (SD)	min	5	25	50	75	95	max
All ages & causes	138.1 (18.9)	71	110	125	137	150	172	209
over 65								
all causes	67.8 (12.7)	28	49	59	67	76	91	113
respiratory	10.7 (4.2)	1	5	8	10	13	18	27
CVD	33.6 (7.8)	11	22	28	33	39	47	64
under 5								
all causes	13.2 (3.9)	3	7	10	13	16	20	26
respiratory	2.2 (1.7)	0	0	1	2	3	5	10
pneumonia	1.7 (1.5)	0	0	1	1	2	4	9

CVD = cardiovascular diseases

A large number of daily events were observed for some subgroups of mortality. Although counting data usually follows a Poisson distribution these large daily mean number of deaths allowed the use of Gaussian regression models in some steps of the analysis. On the other hand, deaths in children for example, with much lower daily means necessitated the use of Poisson models throughout the analysis.

During the period of this study (3 years for mortality) there was a slight increase in the population of São Paulo. This was reflected in the upward trend in the daily total number of deaths observed (Fig. 4.6). In addition, these deaths showed an important seasonal pattern with all cause mortality peaking during the winter months of June to August.

Similar patterns were also observed for all cause mortality in the elderly, cardiovascular and respiratory mortality in the elderly.

Deaths in São Paulo did not exhibit a strong weekly pattern. For example, all cause mortality in the elderly had mean number of deaths ranging from 65.6 on Saturdays to 69.3 on Thursdays while the overall mean was 67.8 deaths per day.

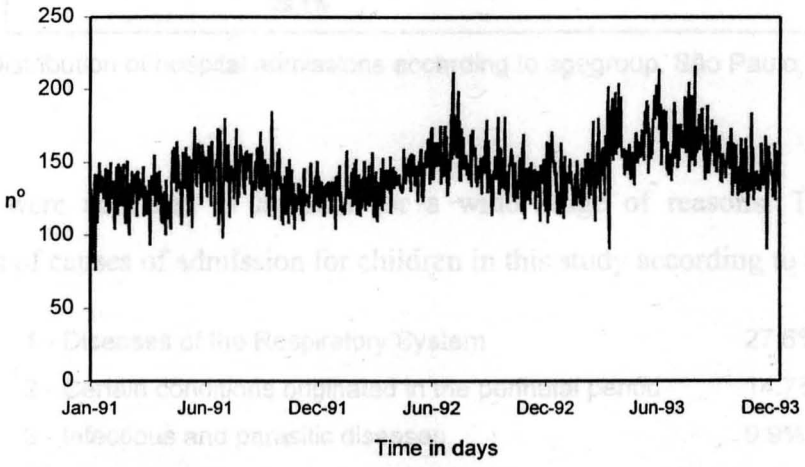


Fig 4.6 - Daily number of deaths for all causes (excluding external causes) in São Paulo, Brazil, 1991-1993

4.3. HOSPITAL ADMISSIONS

Hospital admission data were extracted from files of the Health Authority of the State of São Paulo for the period between November 1992 to September 1994. During these 23 months 188,299 records were available corresponding to hospital admissions of children under 12 years old. Of these, 7,469 (4%) had missing information for age and 1,045 (0.6%) had no information on cause of admission and were therefore, excluded from the dataset. The distribution of these admissions according to age can be inspected in Fig 4.7.

It can be noted that neonatal admissions, i.e. admissions of children up to 28 days of life represented nearly 20% of all admissions. Nevertheless, pre-school children (aged between 1 to 5 years old) are the largest group with nearly 30% of all admissions.

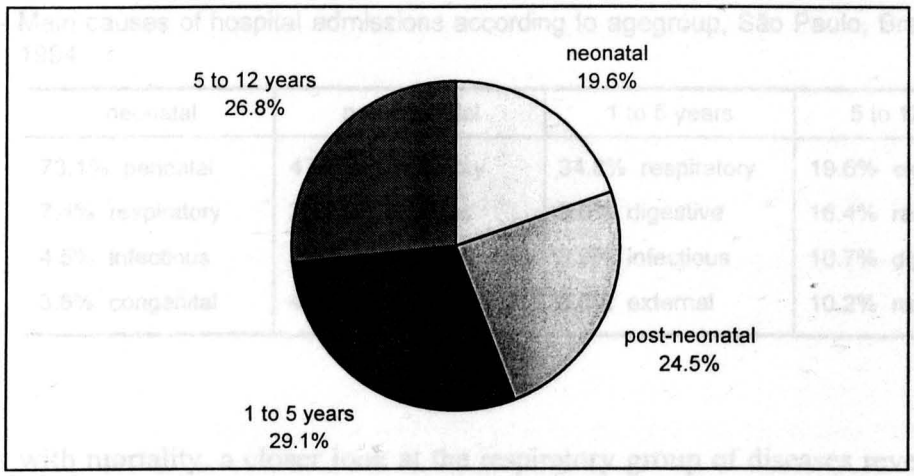


Fig 4.7 - Distribution of hospital admissions according to agegroup, São Paulo, Brazil, 1992-1994

Children were admitted to hospital for a wide range of reasons. The first 10 main categories of causes of admission for children in this study according to the ICD-9 were:

1 - Diseases of the Respiratory System	27.6%
2 - Certain conditions originated in the perinatal period	14.7%
3 - Infectious and parasitic diseases	9.9%
4 - Injury and poisoning	8.4%
5 - Diseases of the digestive system	7.0%
6 - Diseases of the nervous system and sense organs	6.4%
7 - Congenital anomalies	6.2%
8 - Diseases of the genitourinary system	5.1%
9 - Neoplasm	3.6%
10 - Symptoms, signs and ill-defined conditions	2.4%

As it can be seen, respiratory diseases represented the largest group with 27.6% of all admissions followed by perinatal, and infectious and parasitic diseases. It should be noted that the main contributor in this latter group was diarrhoeal diseases. However, as it would be expected, the distribution of causes of admission changed according to different agegroups (Table 4.6). During the neonatal period the majority of admissions were due to perinatal conditions while from age one month up to 5 years old respiratory conditions were the main contributor. Note that in school aged children (5 to 12 years old) external conditions (injury and poisoning) surpassed respiratory diseases as the main reason for hospitalisation.

Table 4.6 - Main causes of hospital admissions according to agegroup, São Paulo, Brazil, 1992-1994.

neonatal	postneonatal	1 to 5 years	5 to 12 years
73.1% perinatal	47.5% respiratory	34.8% respiratory	19.6% external
7.4% respiratory	22.1% infectious	9.0% digestive	16.4% respiratory
4.5% infectious	5.9% congenital	8.9% infectious	10.7% digestive
3.6% congenital	4.9% nervous system	8.6% external	10.2% nervous

As well as with mortality, a closer look at the respiratory group of diseases revealed that pneumonia infections again played a major role. They were responsible for nearly 2/3 of all admissions due to respiratory conditions. Other specific causes of respiratory disease admissions and their respective contribution are:

- Chronic obstructive pulmonary diseases and allied conditions 10.8 %
- Other diseases of upper respiratory tract 8.8 %
- Acute respiratory infections 7.5 %
- Other diseases of respiratory system 7.5 %
- Pneumoconiosis and other lung diseases due to external agents 0.2 %

To evaluate the effects of changes in exposure to air pollution on hospital admissions on a daily basis it is important to have an understand of the daily variations of these admissions. This study focused on the association of air pollutants and hospital admissions in children under 5 years of age. Therefore, the following analysis will focus on that specific agegroup.

In Table 4.7 the number of hospital admissions per day are presented in detail. The median daily number of admissions for all causes (excluding external causes) in children under 5 years old was 137, of which 54 were due to respiratory conditions, and of those, 39 were due to pneumonia infections. The distribution of theses admissions for children under 1 year of age is also described.

Along with mortality, hospital admissions also showed an important seasonal pattern. However, this was more noticeable for admissions due to respiratory diseases which peaked during the months of April to July (Fig 4.8). When looking at admissions for all

causes except external conditions the plot shows a much less strong pattern. In fact, the daily admissions for children under 12 years old for all causes exhibited a pattern much more stable with admissions almost in the same range throughout the year (Fig 4.9). In addition, hospital admissions exhibited a strong weekly pattern with admissions low during weekends and rising on Monday and remaining high throughout the week.

Table 4.7 - Descriptive statistics for the daily Hospital Admissions counts in São Paulo, Brazil, 1992-1994.

		Percentiles						
	mean (SD)	min	5	25	50	75	95	max
all causes	130.9 (40.3)	45	68	92	137	162	190	221
Respiratory								
<5	56.1 (18. 7)	18	29	42	54	68	91	117
<1	30.0 (14.1)	5	13	20	27	37	60	77
Pneumonia								
<5	40.8 (15.0)	6	20	30	39	50	70	94
<1	24.0 (11.4)	3	9	16	21	30	48	69

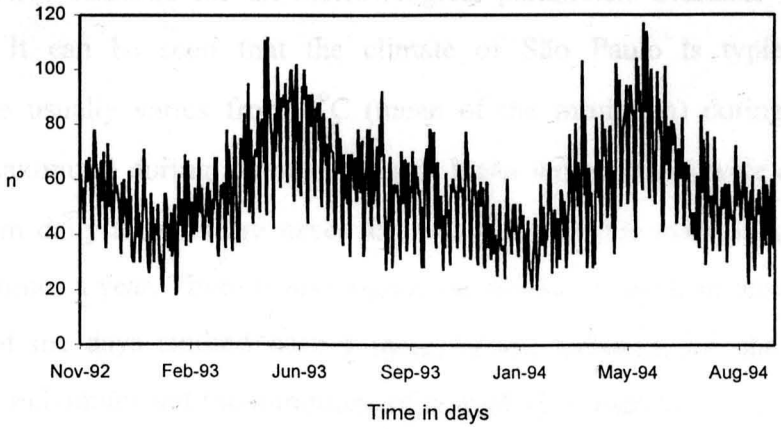


Fig 4.8 - Daily number of hospital admissions for respiratory diseases in children under 5 years of age in São Paulo, Brazil, 1992-1994.

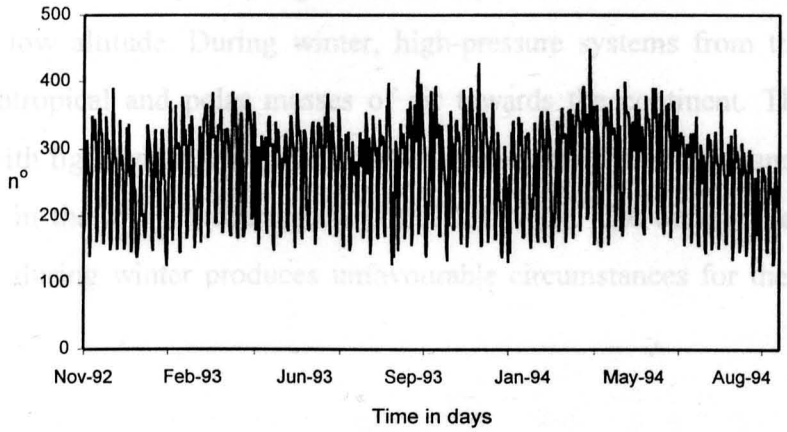


Fig 4.9 - Daily number of hospital admissions for all causes (excluding external causes) in children under 12 years of age in São Paulo, Brazil, 1992-94.

4.4. METEOROLOGICAL VARIABLES

Daily measures of meteorological variables were obtained from the University of São Paulo's Instituto Astronomico e Geofísico. They comprise: temperature (mean, maximum and minimum daily level in °C), humidity (mean, maximum and minimum daily levels expressed in percentage), atmospheric pressure (daily mean in mmHg), precipitation (daily total in mm), wind speed (daily mean in km/h), wind direction (daily resultant).

The descriptive statistics for all meteorological parameters available are presented on Table 4.8. It can be seen that the climate of São Paulo is typically subtropical. Temperature usually varies from 8°C (mean of the minimum) during winter to 30°C (mean of maximum) during summer. Annual mean temperatures were around 19°C and the minimum daily temperature never fell below zero while maximum were over 30°C only a few times a year. There is also a great variability in levels of temperature during a day. Half of the days studied have a range of temperatures, i.e. the daily difference between the maximum and the minimum, of at least 10.3 degrees.

The weather in São Paulo can be summarised as dry in winter and humid during summer. From September to April the area where the city is located is dominated by a southern humid wind that brings frequent frontal systems. In addition, there is higher incidence of solar irradiation and displacement of the centre of the polar and subtropical masses of air

towards the sea. Consequently a great instability over the continent with precipitation and clouds of low altitude. During winter, high-pressure systems from the Atlantic Ocean moves subtropical and polar masses of air towards the continent. This produces high stability with light winds, many hours of calms, a reduction in rain, and frequent thermal inversions in the low-altitude region [CETESB,1993]. The occurrence of such weather conditions during winter produces unfavourable circumstances for the dispersion of air pollutants.

Table 4.8 - Descriptive statistics for meteorological parameters in São Paulo, Brazil, 1991-1994.

		Percentiles						
	mean (SD)	min	5	25	50	75	95	max
Temperature								
mean	19.2 (3.3)	8.6	13.3	16.97	19.4	21.7	23.9	26.3
maximum	25.2 (4.6)	10.6	16.7	22.30	25.7	28.4	31.7	35.2
minimum	15.1 (3.2)	1.9	9.2	12.80	15.4	17.8	19.6	22.4
range	10.1 (3.7)	1.4	3.6	7.3	10.3	12.8	16	22.2
Humidity								
mean	78.7 (9.5)	42.5	60.6	73.20	79.5	85.5	92.9	98
maximum	94.4 (3.9)	67	88	93	95	97	98	100
minimum	52.6 (17.8)	7	25	39	51	65	85	97
Atmospheric Pressure	95.8 (2.8)	88.4	91.4	93.8	95.6	97.6	100.8	106
Precipitation	4.3 (10.4)	0	0	0	0	3.2	23	121

The daily variations in mean temperature and relative humidity can be visualised in Figures 4.10 and 4.11. Note that Relative Humidity is lower during the winter periods which characterises a dry winter.

In the past few years changes were observed in temperature, humidity, visibility and other meteorological parameters measured at the city of São Paulo. For example, in spite of lower levels of solar radiation due to atmospheric pollution the temperatures observed within the city are generally higher than in the neighbouring rural areas [CETESB,1993]. This is probably caused by the huge built-up area in the metropolitan region of São Paulo (MRSP).

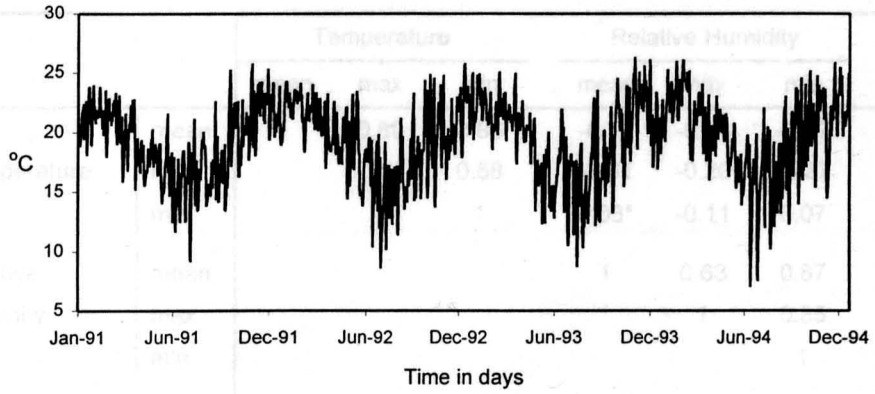


Fig 4.10 - Daily mean values of Temperature (°C) in São Paulo, Brazil, 1991-1994

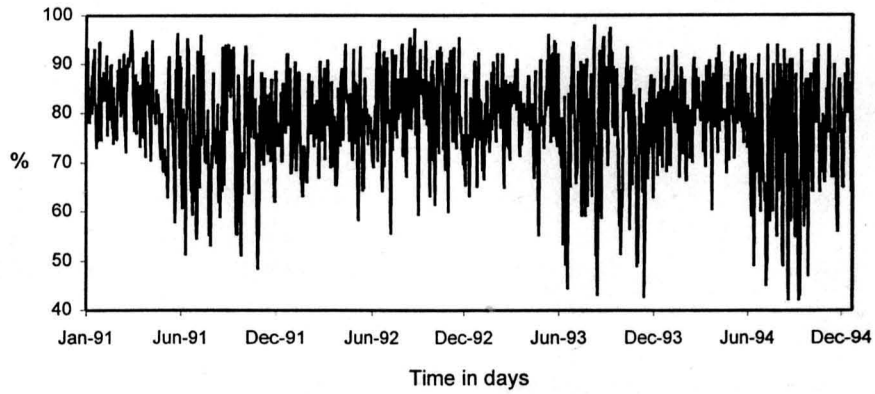


Fig 4.11 - Daily mean values of Relative Humidity (%) in São Paulo, Brazil, 1991-1994

The crude pairwise correlation between meteorological parameters are presented on Table 4.9. As expected there is a high correlation between mean, maximum and minimum daily values for both temperature and humidity. Besides that, the only correlations that call attention are the negative correlation between atmospheric pressure and temperature, and between mean relative humidity and temperature (especially maximum).

Table 4.9 - Pearson correlation matrix for meteorological parameters.

		Temperature			Relative Humidity				
		mean	max	min	mean	max	min	PA	PT
Temperature	mean	1	0.89	0.86	-0.39	-0.26	-0.37	-0.68	0.08
	max		1	0.58	-0.61	-0.26	-0.37	-0.52	-0.4*
	min			1	0.03*	-0.11	0.07	-0.67	0.23
Relative Humidity	mean				1	0.63	0.87	0.10	0.33
	max					1	0.35	0.12	0.21
	min						1	0.04*	0.28
	PA							1	-0.26
	PT								1

PA = atmospheric pressure PT = precipitation
* correlation coefficients not significant (p>0.05)

CHAPTER 5.
TIME SERIES: MORTALITY

The methods utilised for the time-series analysis of the effects of air pollution in the city of São Paulo were fully described in chapter 3. The present chapter will introduce in more detail the results for each procedure employed in order to build the 'core model' as well as the main findings for the mortality data. As referred to in section 3.2 - Study Design, the time-series analyses of mortality and hospital admissions in São Paulo were carried out only for the selection of outcomes chosen beforehand. This chapter will concentrate on the results of the mortality data and the next chapter will present the results for the hospital admissions data.

5.1. MODEL SELECTION PROCESS

As discussed previously, for each mortality outcome a complete set of procedures were performed in order to build the core model. The selection of variables was carried out in

regression models and using statistical tests to assess improvements in fit. In general, results were very similar for all the different causes of mortality. Nevertheless, to illustrate these procedures in detail, descriptions will be provided for 'all cause mortality in individuals aged 65 years or more' used as the outcome of reference. For other outcomes, just the final 'core model' will be described and any specific steps in the model building process will only be outlined if they were substantially different from those obtained for this outcome. Other endpoints also explored in this analysis were respiratory and cardiovascular deaths in the elderly (65 years or older), respiratory deaths in children under 5 years, and pneumonia deaths in children under 5 years of age.

5.1.1. Long, medium and short term variations and other cyclical patterns

The adjustment for temporal patterns, especially long and medium term ones, was aided by the use of spectral analysis of the residuals. This procedure helped identify patterns or periodicities not considered *a priori* and also to assess the adequacy of the periodicities selected.

The plot of the raw mortality data shown in Fig 4.9 (Chapter 4) illustrates the substantial seasonal periodicity present in the data which needs to be filtered. The periodogram of the unfiltered log-transformed data shows large spikes for periods above one year very close to the y-axis and a larger annual spike (Fig 5.1-(a)). These spikes indicate the presence of long term trends in the data as well as the one year cycle. The long term trends were then modelled by including a linear time term, a quadratic time term, and indicator variables for the years of the study. A biennial cycle also contributed to correct for these long term trends. These variables were entered into the models for different outcomes according to the significance of their contribution to the fit. The periodogram of the residuals after fitting these long term trends shows only the large spike corresponding to a one year period (Fig 5.1-(b)). Note that the vertical dotted line in the plots refers to a frequency equivalent to a period of two months, therefore, the periodicities that need to be removed are concentrated in the left-most part of the plot.

The seasonal and cyclical medium term patterns were then modelled by using harmonic waves as described in section 3.6.1.

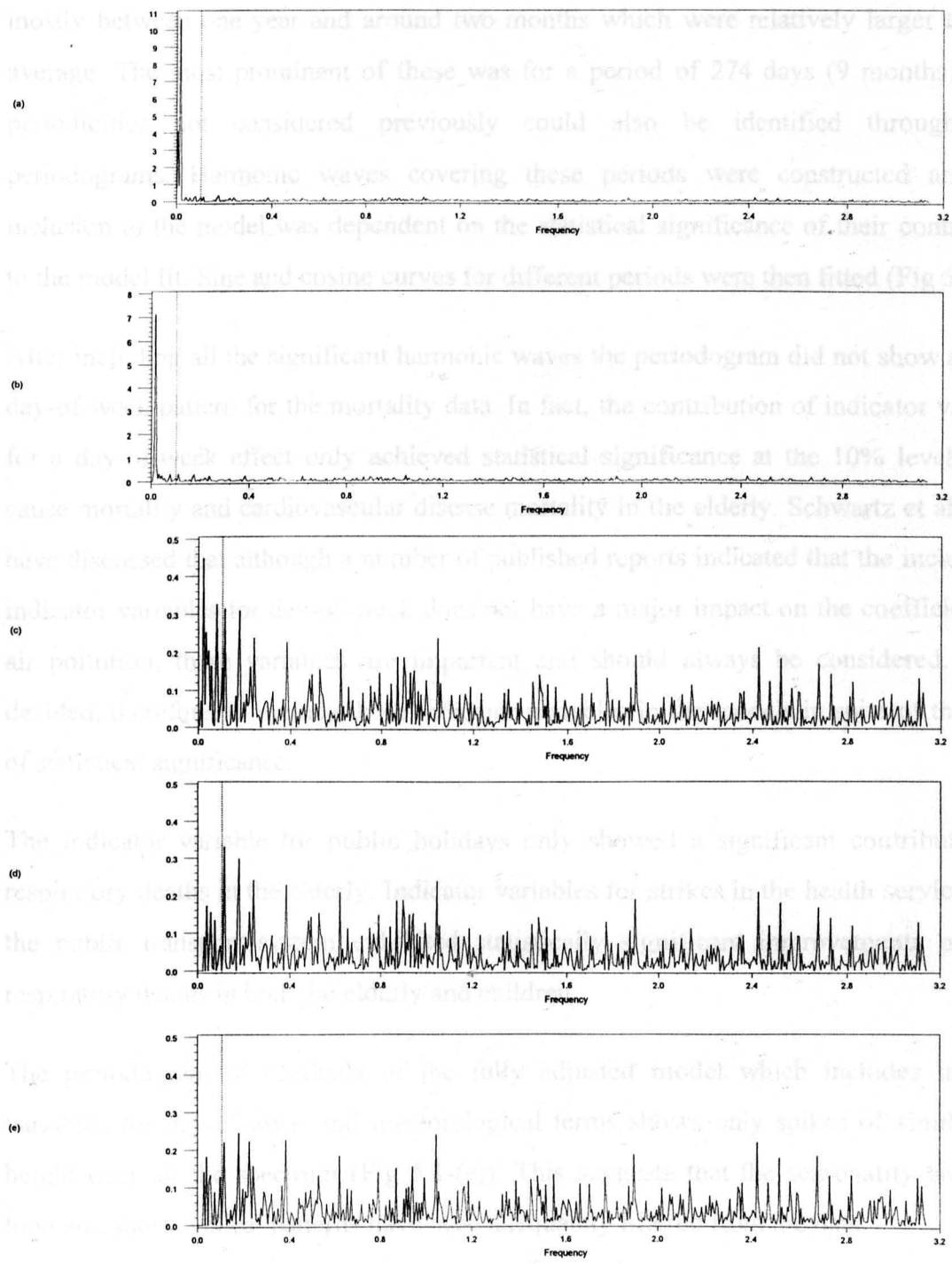


Fig 5.1 - Periodograms of the residuals of all cause mortality in the elderly for every step of the model building. (a) unfiltered log-transformed data. (b) after adjustment for long term trends. (c) after inclusion of harmonic waves for one year period. (d) after inclusion of all harmonic waves down to two months. (e) Periodogram of the fully adjusted model including indicator variables for day of week and terms for meteorological variables. The vertical dotted line refers to a frequency equivalent to a period of two months.

As soon as the one year cycle curve was fitted there was a dramatic change in the scale of the periodogram. This made it possible to identify other important spikes (Fig 5.1-(c))

mostly between one year and around two months which were relatively larger than the average. The most prominent of these was for a period of 274 days (9 months). Other periodicities not considered previously could also be identified through these periodograms. Harmonic waves covering these periods were constructed and their inclusion in the model was dependent on the statistical significance of their contribution to the model fit. Sine and cosine curves for different periods were then fitted (Fig 5.1-(d)).

After including all the significant harmonic waves the periodogram did not show a strong day-of-week pattern for the mortality data. In fact, the contribution of indicator variables for a day-of-week effect only achieved statistical significance at the 10% level for all cause mortality and cardiovascular disease mortality in the elderly. Schwartz et al [1996] have discussed that although a number of published reports indicated that the inclusion of indicator variables for day-of-week does not have a major impact on the coefficients for air pollution, these variables are important and should always be considered. It was decided, therefore, to include these indicator variables in the models in spite of their lack of statistical significance.

The indicator variable for public holidays only showed a significant contribution for respiratory deaths in the elderly. Indicator variables for strikes in the health services or in the public transport system exhibited statistically significant improvements only for respiratory deaths in both the elderly and children.

The periodogram of residuals of the fully adjusted model which includes indicator variables for day-of-week and meteorological terms shows only spikes of similar, low height over all the spectrum (Fig 5.1-(e)). This suggests that the seasonality and other long and short cyclical patterns have been adequately filtered from the data.

This was also confirmed by the inspection of plots of the predicted values over time after including each term in the model. The plots showed that the fitted model for seasonality and other temporal patterns was providing a good description of the data. The plots for all cause mortality in the elderly are displayed in Appendix A.

5.1.2. Meteorological Variables

Cross-correlations of the de-trended and de-seasonalized mortality residuals against residuals of temperature with lags up to five days were examined. The highest correlation for mortality in the elderly was for temperature with lag=0, i.e. with measures of temperature on the same day. For respiratory and pneumonia deaths in children under five years of age the best cross-correlations were observed for temperature lagged two days.

Subsequently, plots of the adjusted residuals against the four measures of daily temperature at the appropriate lags were considered. The crude relationship between temperature and mortality has been demonstrated to be non-linear or at least piecewise linear [Touloumi et al,1994; Schwartz et al,1996]. It was observed in this study that after filtering the seasonality and long term patterns from the data there were still signs of non-linearity. For example, for all cause mortality in the elderly there was some indication that a piecewise function would better describe its relationship with mean temperature although a quadratic relation should also be considered (Fig 5.2). It also seems that the relationship between minimum temperature and all cause mortality in the elderly could be approximated by three-piecewise shape whereas for maximum and the difference in temperature the pattern suggests a linear or quadratic relationship.

Thus, for each different indicator of temperature and according to the shape of the exposure-response relationship described by visual examination of the plots, a series of different functions was created. For example, a two-piece linear function was obtained by constructing two complementary variables, 'hot' and 'cold' to measure daily temperature. They were defined as:

```
cold=1 if temp<=20; else cold=0;
hot=1 if temp>20; else hot=0;
cold1=temp*cold; hot1=temp*hot;
```

The indicators 'hot' and 'cold' are defined to allow for different intercepts of the relationship between temperature and health outcomes. The terms 'cold1' and 'hot1' are defined to allow for different slopes. The choice of the changing-point was again based on the visual examination of the plots. Two out of these four variables (for example cold and

hot1) should be included into the model as well as ('temp') the measure of temperature to which they are referring (mean, maximum, etc.). This procedure assures sufficient flexibility in temperature control without overspecifying the model.

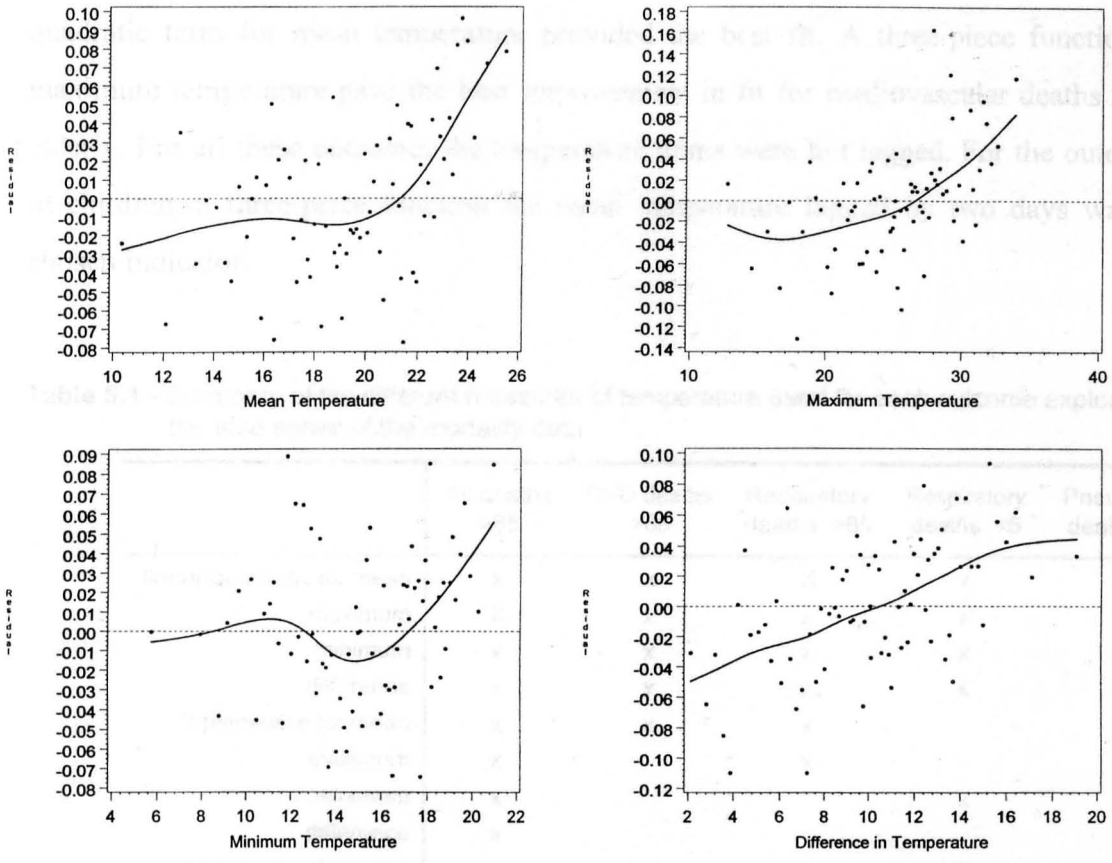


Fig 5.2 - Plots of the residuals of all cause mortality in the elderly, after adjustment for long term trends, seasonal and calendar effects, against four measures of daily temperature. Each point represents the mean of 15 adjacent observations in order of increasing levels of temperature. A nonparametric smoothed line was fitted to the data using a spline routine [SAS,1991a] to indicate the shape of the relationships. Difference in temperature refers to the daily difference between maximum and minimum temperature.

Following the same approach, a three-piece linear function was also produced. Additionally, indicator variables were created using eight equally-spaced intervals of temperature. This approach is also considered to adequately control for the temperature effect mainly when the exact shape of the relationship is non-linear and inadequately summarised by a piecewise linear function [Schwartz,1994c]. These different functions were produced for all four measures of daily temperature. Single linear and quadratic terms were also investigated. A summary of the different variables constructed to take into account the effect of temperature on mortality are listed in Table 5.1.

For all cause mortality in the elderly the best fit according to F-tests was achieved by the inclusion of a linear plus a quadratic term for maximum temperature. For respiratory deaths in the elderly, the residuals after adjusting for seasonal and temporal patterns showed a nearly linear relationship with temperature. Yet, for this outcome a linear plus quadratic term for mean temperature provided the best fit. A three-piece function for maximum temperature gave the best improvement in fit for cardiovascular deaths in the elderly. For all these outcomes the temperature terms were not lagged. For the outcomes in children, a three-piece function for mean temperature lagged by two days was the chosen indicator.

Table 5.1 - Summary of the different measures of temperature used for each outcome explored in the time series of the mortality data.

	All deaths >65	CVD deaths >65	Respiratory deaths >65	Respiratory deaths <5	Pneumonia deaths <5
linear/quadratic for mean	x	x	**X**	x	x
maximum	**X**	x	x	x	x
minimum	x	x	x	x	x
difference	x	x	x	x	x
2 piecewise for mean	x	x	x		x
maximum	x		x		
minimum	x			x	x
difference	x				x
3 piecewise for mean	x	x	x	**X**	**X**
maximum	x	**X**	x	x	x
minimum	x	x	x	x	x
difference	x	x	x	x	x
8 indicators for mean	x	x	x	x	x
maximum	x	x	x	x	x
minimum	x	x	x	x	x
difference	x	x	x	x	x

Note: capital 'X' indicates the temperature term with the best fit, and used in subsequent modelling.

After selecting the best temperature variable to fit the data, the residuals now also adjusted for the appropriate measure of temperature were once more plotted against temperature in order to check if adequate adjustment had occurred. The plot for all cause mortality (Fig 5.3) exhibits an almost straight line with zero gradient although some observations for extreme temperatures give the impression that inadequate adjustment has

been made. This demonstrated that the adjustment was successful in controlling for the effect of temperature on mortality.

The next step was to include a measure of Relative Humidity in the model. Measures of the daily mean, maximum and minimum humidity were tested as a linear function at the same lag used for temperature. Maximum humidity gave the best improvement in the model's fit for the elderly and minimum humidity likewise in children. Other meteorological variables such as atmospheric pressure and precipitation did not contribute significantly to any model's fit and were therefore, not included in the final model for any outcome.

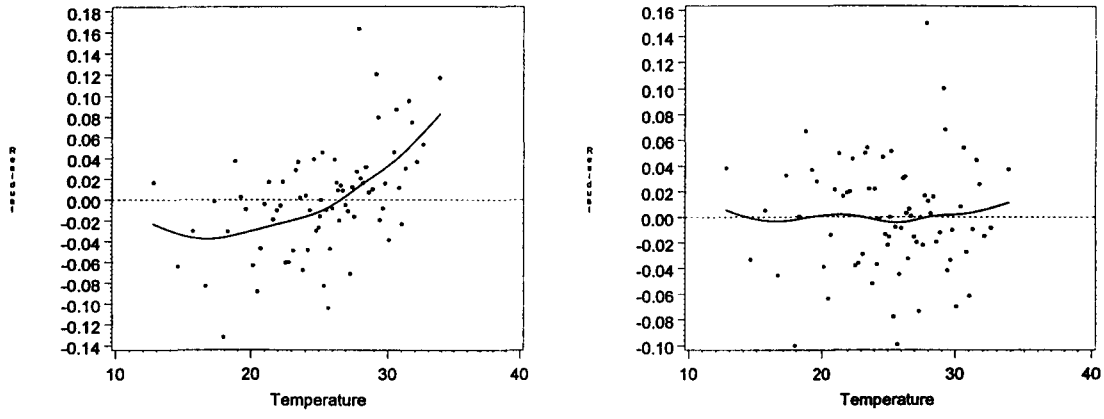


Fig 5.3 - Plots of the residuals of all cause mortality in the elderly against mean temperature. In the plot on the left, the residuals are adjusted for long term trends, seasonal and calendar effects but not temperature. In the plot on the right, the residuals are additionally adjusted for temperature using a linear plus a quadratic term for maximum temperature. Each point represents the mean of 15 adjacent observations in order of increasing levels of temperature. A nonparametric smoother was run-through the data using a spline routine [SAS,1991a] to indicate the shape of these relationships.

5.1.3. Autocorrelation

The final step in developing the core model was to allow for any remaining autocorrelation in the data. Autoregressive models were used to identify the order of the process and to estimate the parameters. The statistics and the Partial Autocorrelation Function (PACF) plots produced by those models were examined first for the raw data, and subsequently for the residuals after each stepwise adjustment. Taking the all cause mortality in the elderly as an example we can observe that the PACF plot of the raw data

showed significant autocorrelation up to lag 8 suggesting an AR(8)-process at least (Fig 5.4 (a)). As different patterns were filtered from the data the PACF plots change (Fig 5.4 (b) to (f)). When long term trends and the 1 year cycle are filtered the PACF is down to an AR(2)-process with small coefficients (Fig 5.4 (c)). After adjustment for season, trend and calendar effects most of the serial correlation in the residuals has disappeared. However, the partial autocorrelation function still points to an AR(1) process but with a small coefficient of approximately 0.15 (Fig 5.4 (e)).

The final model allowing also for meteorological variables still indicates an AR(1)-process with a coefficient of 0.11, statistically significant at the 5% level (Fig 5.4 (f)). Therefore, the final model included an autoregressive parameter of order one (AR-1) to allow for this remaining autocorrelation.

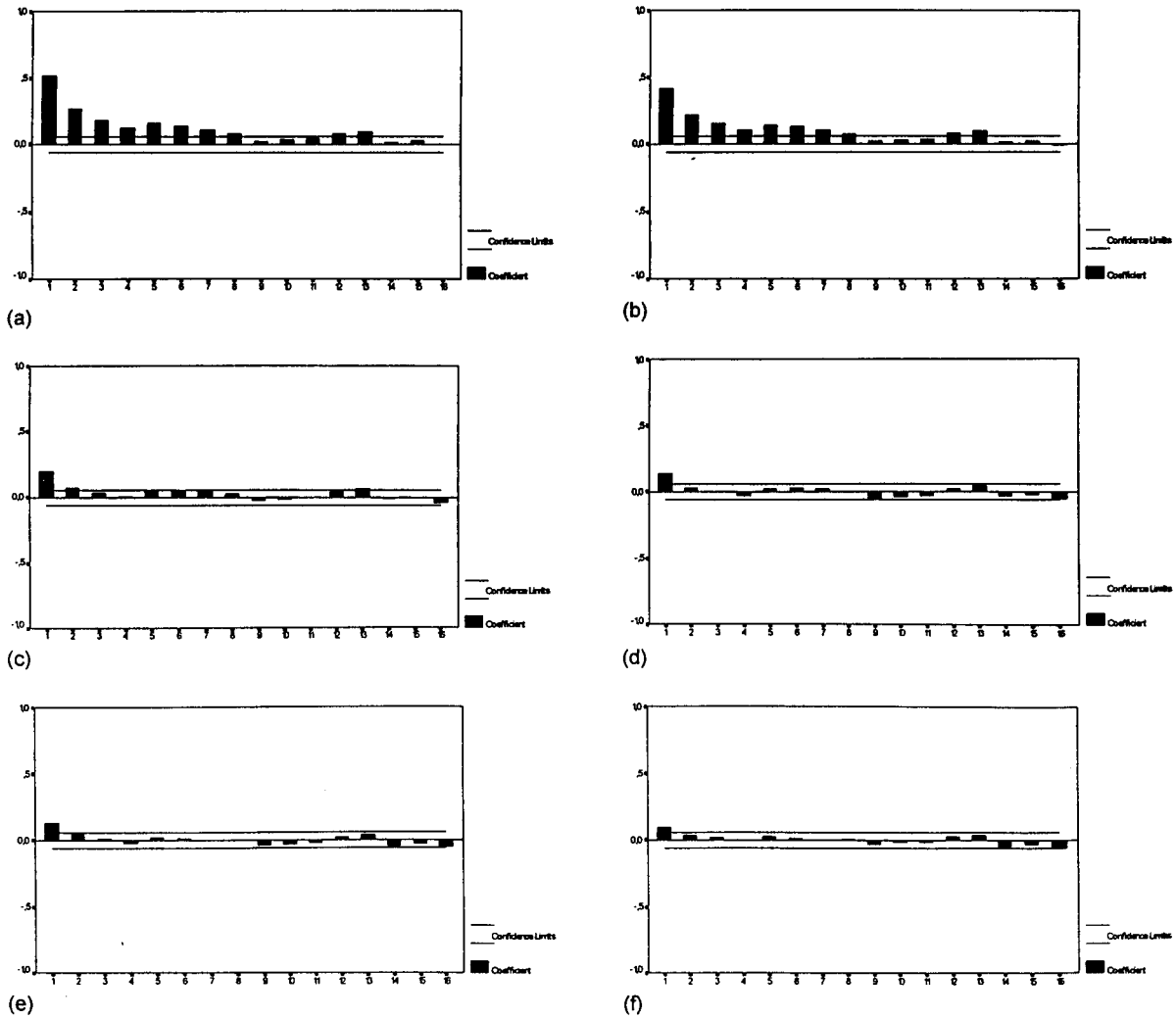


Fig 5.4 - Partial autocorrelation function for all cause mortality in the elderly. (a) unadjusted data. (b) adjusted for long term trends. (c) adjusted for 1 and 2 years cycles. (d) adjusted for all sine and cosines. (e) adjusted for day of week, holidays and strikes. (f) fully adjusted model (including meteorological variables).

For the other outcomes the PACF plots were very similar with decreasing degrees of autocorrelation as the model was fitted. However, for these other outcomes, the adjustment for trend, seasonality, calendar effects and meteorology was enough to completely account for the autocorrelation. This means that all the autocorrelation originally observed for these outcomes before adjustment for confounding was explained by those extraneous factors now included in the model. Therefore, no additional correction for serial correlation was needed for these outcomes.

A summary of the core model for each outcome is presented in Table 5.2. The table includes all variables used to adjust for long, medium and short term temporal patterns present in the data plus the meteorological variables and highlights the variables that were selected for each outcome. The residuals of these models were then plotted against the time period to examine if there were signs of temporal and seasonal patterns still not adjusted for. As observed in the example of Figure 5.5 the core model was successful in controlling for all these patterns.

Table 5.2- Variables selected for the core model of each mortality outcome.

	> 65 years old			< 5 years old	
	all causes	respiratory	cardiovascular	respiratory	pneumonia
indicators for years		X		X	X
trend (linear+quadratic)	X	X	X		
2 years	X	X	X		
1 year	X	X	X	X	X
9 months	X	X	X	X	X
7 months				X	X
6 months	X		X	X	X
5 months	X		X	X	X
4 months		X			
3 1/2 months		X			
3 months					X
79 days	X		X		
74 days		X			
69 days		X		X	X
2 months	X			X	
days of week	X	X*	X	X*	X*
holidays		X			
transport strikes		X			
health services strikes				X	X
temperature	L+Q (maximum)	L+Q (mean)	3-piecewise (maximum)	3-piecewise (mean)	3-piecewise (mean)
humidity	minimum	maximum	maximum	minimum	minimum

* not statistically significant at the 10% level
L+Q = Linear plus quadratic terms

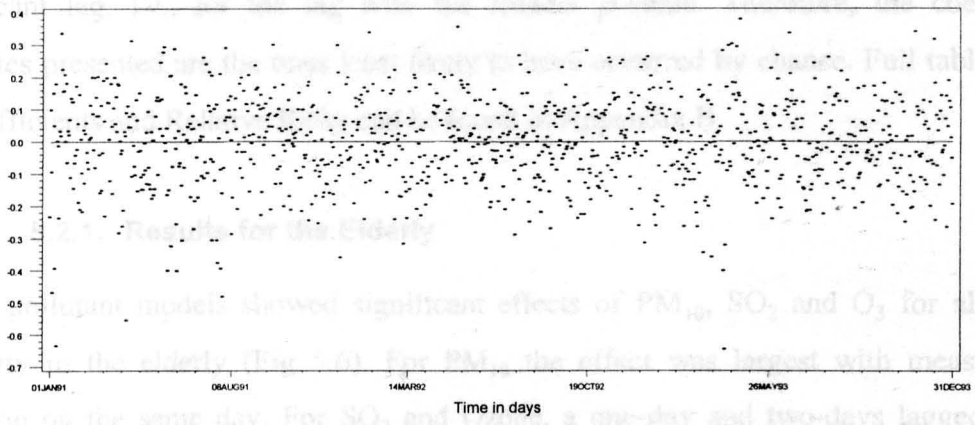


Fig 5.5 - Plot of the residuals of all causes mortality in the elderly from the core model adjusting for long and short temporal trends, plus meteorological variables versus time. A non-parametric smooth curve was run-through the data to indicate any remaining seasonal variability in the data.

5.2. MAIN FINDINGS

Poisson regression models allowing for overdispersion and, in the case of all cause mortality in the elderly, for autocorrelated errors were used to estimate the coefficients of air pollution variables for mortality for different causes and ages. In addition, Relative Risks (RR) of death were calculated by exponentiating the coefficients with natural base *e*. These RR were calculated in relation to a change from the 10^{th} to the 90^{th} centile in levels of each air pollutant. The actual values representing these changes for the three year period of the mortality analysis were:

- PM_{10} $\Rightarrow$ 64.2 $\mu g/m^3$
- SO_2 $\Rightarrow$ 20.2 $\mu g/m^3$
- NO_2 $\Rightarrow$ 180.9 $\mu g/m^3$
- O_3 $\Rightarrow$ 106.0 $\mu g/m^3$
- CO $\Rightarrow$ 5.1 ppm

Some Relative Risks were also calculated for a 100 $\mu g/m^3$ increase in levels of some of these pollutants (or 10 ppm increase in levels of CO) in order to have results comparable to the published literature.

All the analyses were performed for levels of pollutants on the same day and lagged one and two days. The graphs displayed in this section will show the results for all lags studied. However, the tabulated results presented here are only for the most statistically

significant lag, i.e., for the lag with the smaller *p*-value. Therefore, the coefficient estimates presented are the ones least likely to have occurred by chance. Full tables with all coefficients and Relative Risks can be found in Appendix B.

5.2.1. Results for the Elderly

Single pollutant models showed significant effects of PM_{10}, SO_2 and O_3 for all cause mortality in the elderly (Fig 5.6). For PM_{10} the effect was largest with measures of pollution on the same day. For SO_2 and Ozone, a one-day and two-days lagged effect respectively gave the largest coefficient. CO and NO_2 showed positive relationships with mortality for all causes (for some lags) although these were not statistically significant.

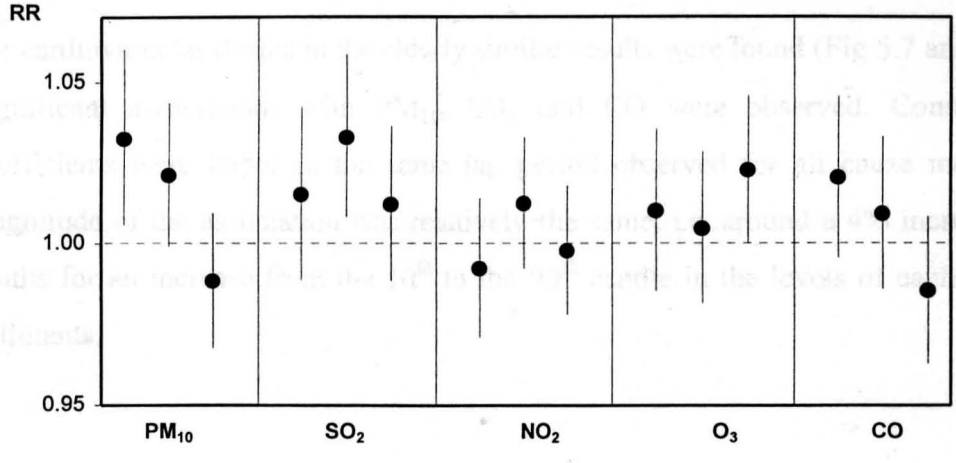


Fig 5.6 - Relative Risks (RR) and 95% CI for all causes mortality excluding external causes in the elderly for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most).

After adjusting for long term trends, seasonality, calendar effects and for meteorology, an increase from the 10th to the 90th centile in the levels of PM_{10} or SO_2 was associated with a 3.3% increase in daily mortality in the elderly (Table 5.3). A slightly lower effect was observed for O_3 (2.3%), followed by CO (2.0%) and for NO_2 (1.3%).

In order to have comparable results with previous published studies, the Relative Risks were also estimated for a 100 $\mu g/m^3$ increase in levels of PM_{10} and SO_2. In this case, the additional mortality due to this increase in PM_{10} was 5% and for SO_2 it was 17%. In contrast with the published literature concerning health effects of outdoor air pollution, SO_2 exhibited a 'stronger' effect in all cause mortality in the elderly in São Paulo.

Table 5.3 - Poisson regression coefficients for unit change in single pollutant models and Relative Risk (RR) of death for all cause mortality (excluding external causes) in people aged 65 years or older for an increase from the 10th to the 90th centile in levels of air pollutants.

	coefficient	Std Err	RR	95% CI
PM_{10}	0.499	0.209	1.033	(1.006 - 1.060)
SO_2 (lag 1)	1.611	0.612	1.033	(1.008 - 1.058)
NO_2 (lag 1)	0.069	0.057	1.013	(0.992 - 1.033)
O_3 (lag 2)	0.213	0.108	1.023	(1.000 - 1.046)
CO	3.967	2.459	1.020	(0.996 - 1.046)

coefficients and Standard Error (SE) X 10^3
Results at lag zero otherwise stated

For cardiovascular deaths in the elderly similar results were found (Fig 5.7 and Table 5.4). Significant associations with PM_{10}, SO_2 and CO were observed. Consistently, the coefficients were larger in the same lag period observed for all cause mortality. The magnitude of the association was relatively the same, i.e. around a 4% increase in daily deaths for an increase from the 10th to the 90th centile in the levels of each of those air pollutants.

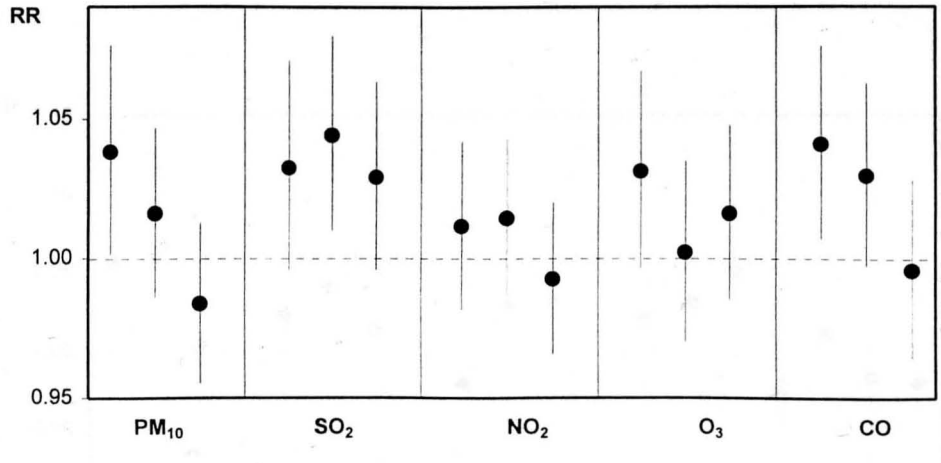


Fig 5.7 - Relative Risks (RR) and 95% CI for cardiovascular mortality in the elderly for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most).

Table 5.4 - Poisson regression coefficients for a unit change in single pollutant models and Relative Risk (RR) of death for cardiovascular mortality in people aged 65 years or older for an increase from the 10th to the 90th centile in levels of air pollutants.

	coefficient	Std Err	RR	95% CI
PM_{10}	0.584	0.287	1.038	(1.001 - 1.076)
SO_2 (lag 1)	2.136	0.846	1.044	(1.010 - 1.080)
NO_2 (lag 1)	0.078	0.079	1.014	(0.986 - 1.043)
O_3	0.288	0.165	1.031	(0.996 - 1.067)
CO	7.839	3.329	1.041	(1.007 - 1.076)

coefficients and Standard Error (SE) X 10^3
Results at lag zero otherwise stated

The risks for a 100 $\mu g/m^3$ increase in PM_{10} and SO_2 was respectively 6% and 23%. Ozone and NO_2 also showed positive associations with Relative Risks around 1.02 or 1.03 respectively. However, these risks were not statistically significant at the 5% level.

Deaths due to respiratory diseases in the elderly also exhibited positive associations with all the major air pollutants (Fig 5.8 and Table 5.5). The increase in daily respiratory mortality was around 6% for PM_{10}, 5% for O_3, 4% for SO_2, and 3% for NO_2. However, these associations were mostly not statistically significant with the exception of PM_{10}. For this pollutant the Relative Risk was 1.06 and it was slightly larger when levels of PM_{10} were lagged by one day.

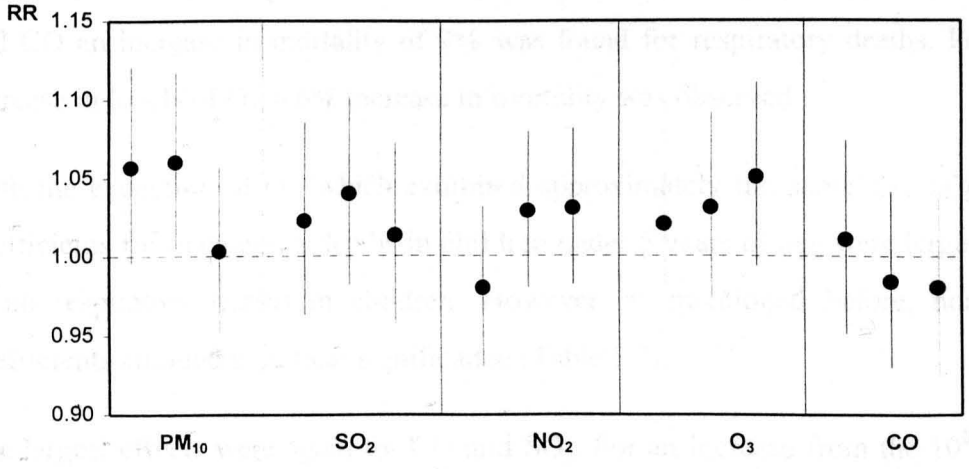


Fig 5.8 - Relative Risks (RR) and 95% CI for respiratory diseases mortality in the elderly for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most).

Although the coefficients were not significant for SO_2, the pattern of the relationship for this pollutant was similar to other causes of death (all causes and cardiovascular diseases), i.e. the strongest effect was present when SO_2 was lagged one day. Relative Risks for CO at all lags were close to 1.0. The *p*-value for 2 days lag was slightly lower (p=0.5), so the estimate at lag 2 is presented in table 5.5.

Table 5.5 - Poisson regression coefficients for single pollutant models and Relative Risk (RR) of death for respiratory mortality in people aged 65 years or older for an increase from the 10th to the 90th centile in levels of air pollutants.

	coefficient	Std Err	RR	95% CI
PM_{10} (lag 1)	0.910	0.423	1.060	(1.005 - 1.118)
SO_2 (lag 1)	1.937	1.456	1.040	(0.981 - 1.101)
NO_2 (lag 2)	0.168	0.136	1.031	(0.982 - 1.082)
O_3 (lag 2)	0.467	0.268	1.051	(0.994 - 1.111)
CO (lag 2)	-3.854	5.634	0.981	(0.927 - 1.037)

coefficients and Standard Error (SE) X 10^3
Results at lag zero otherwise stated

5.2.2. Results for Children

For the effect of air pollutants on mortality due to respiratory diseases, or more specifically due to pneumonia, in children under 5 years of age none of the results were statistically significant at the 5% level. The coefficients showed stronger and positive patterns for SO_2 and CO (Figures 5.9 and 5.10). Ozone had also a positive effect when lagged by one day. For an increase from the 10th to the 90th centile in the levels of SO_2 and CO an increase in mortality of 9% was found for respiratory deaths. For the same increase in levels of O_3, a 6% increase in mortality was observed.

With the exception of O_3, which exhibited approximately the same size of effect, most coefficients for pneumonia deaths in children under 5 years of age were larger, compared to all respiratory deaths in children. However, as mentioned before, none of these coefficients attained statistical significance (Table 5.7).

The largest effects were again for CO and SO_2. For an increase from the 10th to the 90th centile in levels of CO, an increase in mortality for pneumonia in children of 14% was predicted by the model. For SO_2 the increase in daily mortality was 11%.

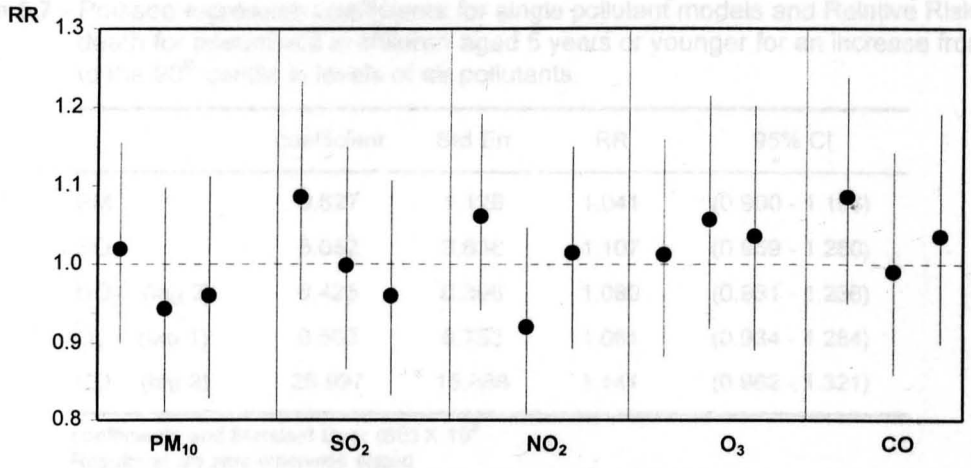


Fig 5.9 - Relative Risks (RR) and 95% CI for respiratory diseases mortality in children under 5 years of age for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most).

Table 5.6 - Poisson regression coefficients for single pollutant models and Relative Risk (RR) of death for respiratory diseases in children aged 5 years or younger for an increase from the 10th to the 90th centile in levels of air pollutants.

	coefficient	Std Err	RR	95% CI
PM_{10} (lag 1)	-0.902	1.210	0.944	(0.810 - 1.098)
SO_2	4.071	3.234	1.086	(0.955 - 1.233)
NO_2 (lag 1)	-0.455	0.367	0.921	(0.807 - 1.047)
O_3 (lag 1)	0.525	0.674	1.057	(0.918 - 1.215)
CO	16.145	13.277	1.086	(0.950 - 1.238)

coefficients and Standard Error (SE) X 10^3
Results at lag zero otherwise stated

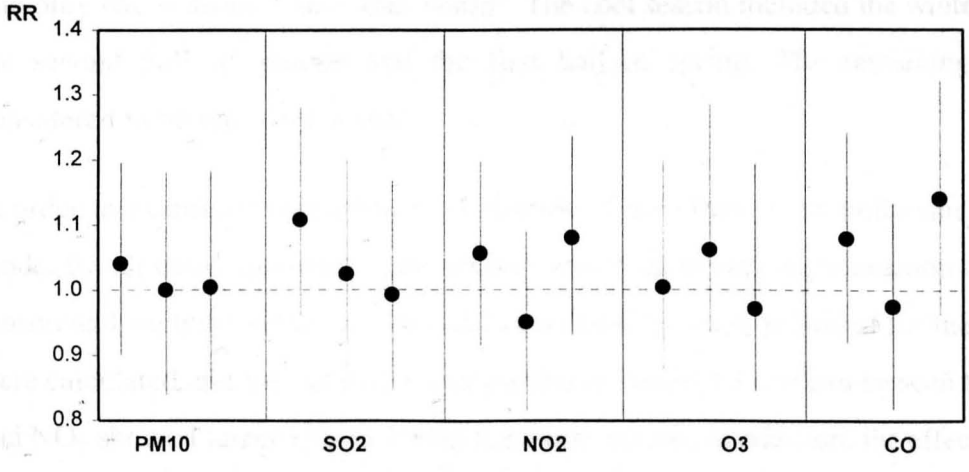


Fig 5.10 - Relative Risks (RR) and 95% CI for pneumonia deaths in children under 5 years of age for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most).

Table 5.7 - Poisson regression coefficients for single pollutant models and Relative Risk (RR) of death for pneumonia in children aged 5 years or younger for an increase from the 10th to the 90th centile in levels of air pollutants.

	coefficient	Std Err	RR	95% CI
PM_{10}	0.627	1.126	1.041	(0.900 - 1.196)
SO_2	5.052	3.636	1.107	(0.959 - 1.280)
NO_2 (lag 2)	0.425	0.396	1.080	(0.931 - 1.236)
O_3 (lag 1)	0.563	0.753	1.061	(0.934 - 1.284)
CO (lag 2)	25.997	15.888	1.141	(0.962 - 1.321)

coefficients and Standard Error (SE) X 10^3
Results at lag zero otherwise stated

In spite of the failure to find a statistically significant effect of air pollutants on mortality in children under 5 years of age for all respiratory diseases or for pneumonia infections, it should be noted the consistency of the results within these two groups of causes of mortality and between them and the mortality in the elderly.

5.2.3. Effects by Season

Although the analysis conducted so far were adjusted for seasonality and weather-related indicators like temperature and humidity, it was decided to investigate whether the effects on mortality associated with exposure to air pollution could still be dependent on season. The four seasons in Brazil, and particularly in São Paulo are not as distinct as they are in more temperate climates. Therefore, the period covered by the mortality data was divided into only two seasons, 'cool' and 'warm'. The cool season included the winter period plus the second half of autumn and the first half of spring. The remaining period was considered to be the warm season.

In order to examine the possible modification of the effect of air pollutants by season, a model for all cause mortality in the elderly was fitted having an interaction term between season and each air pollutant. The effect estimates for each pollutant for the two seasons were calculated and the Relative Risks plotted in Figure 5.11. It can be seen that PM_{10}, O_3 and NO_2 showed larger effects during the warm season. In addition, the effect of O_3 in the warm season was statistically significant at the 5% level.

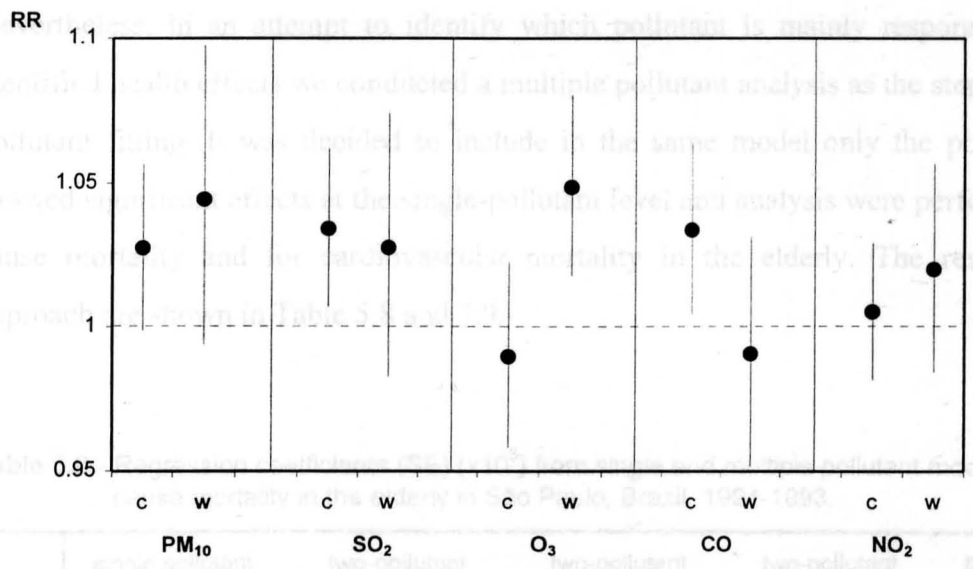


Fig 5.11 - Relative Risks (95% CI) for all cause mortality in the elderly for an increase from the 10th to the 90th centile in levels of pollutants by season (c=cool and w=warm season)

Although the effect of PM_{10} on all cause mortality in the elderly was larger during the warm season, the smaller effect during the cool season was of borderline significance. For SO_2 and CO the effects were higher and statistically significant at the 5% level during the cool season. Nevertheless, it should be mentioned that the effect modification by season was only statistically significant for O_3 (p=0.009) and of borderline significant for CO (p=0.075).

5.2.4. Multiple Pollutant Models

So far we have investigated the effects of air pollutants on health through single pollutant models. It was observed that more than one pollutant showed statistically significant associations with all cause mortality and cardio-vascular mortality in the elderly. However, individuals are not exposed to one pollutant at a time but to a mix of air pollutants present in the air. Trying to isolate the individual effect of each pollutant to distinguish the most damaging agent is a difficult task. Air pollutants in urban areas usually share the same sources of emission and frequently exhibit a high degree of collinearity. Therefore, the inclusion of two or more highly correlated pollutants simultaneously in a model may lead to estimation problems.

Nevertheless, in an attempt to identify which pollutant is mainly responsible for the identified health effects we conducted a multiple pollutant analysis as the step after single pollutant fitting. It was decided to include in the same model only the pollutants that showed significant effects at the single-pollutant level and analysis were performed for all cause mortality and for cardiovascular mortality in the elderly. The results of this approach are shown in Table 5.8 and 5.9.

Table 5.8 - Regression coefficients (SE) ($x10^3$) from single and multiple pollutant models for all cause mortality in the elderly in São Paulo, Brazil, 1991-1993.

	single pollutant model		two-pollutant model		two-pollutant model		two-pollutant model		three-pollutant model	
PM_{10}	0.50	(0.21)	0.30	(0.23)	0.43	(0.22)			0.24	(0.24)
SO_2	1.61	(0.61)	1.26	(0.69)			1.56	(0.63)	1.31	(0.70)
O_3	0.21	(0.11)			0.16	(0.11)	0.16	(0.11)	0.14	(0.11)

When PM_{10} and SO_2 were included in the model for all cause mortality simultaneously, they both have their coefficients decreased with slight increases in the standard errors. However, while SO_2 had a decrease of 20% in its parameter estimate, PM_{10} had a reduction of 40%. When these two pollutants were analysed in the same model with Ozone, SO_2 was nearly unaffected while PM_{10} lost around 15% of its effect estimate. In both cases the parameter estimate for O_3 showed decreases of 24-27%.

In a three-pollutant model we can observe that all pollutants exhibit substantial decreases in their parameter estimates. The reduction is greater for PM_{10} (53%), followed by O_3 (36%) and SO_2 (20%). Note that the standard error increased similarly for PM_{10} and SO_2 while for O_3 it remained unaffected. Similar results were observed in the multiple pollutant models for cardiovascular diseases in the elderly (Table 5.9).

PM_{10} showed greater reductions than SO_2 both when analysed together and when examined against Carbon Monoxide. In the three-pollutant model PM_{10} exhibits the largest decrease in the coefficient estimate, followed by CO. In all cases the standard errors presented increases proportionally smaller than the coefficients.

Table 5.9 - Regression coefficients (SE) ($x10^3$) from single and multiple pollutant models for cardio-vascular mortality in the elderly in São Paulo, Brazil, 1991-1993.

	single pollutant model		two-pollutant model		two-pollutant model		two-pollutant model		three-pollutant model	
PM_{10}	0.58	(0.29)	0.31	(0.32)	0.36	(0.31)		.	0.14	(0.34)
SO_2	2.14	(0.85)	1.72	(0.95)			1.70	(0.89)	1.55	(0.96)
CO	7.84	(3.33)			6.13	(3.64)	5.81	(3.49)	5.33	(3.68)

The above results showed that SO_2 exhibited the smaller reductions either in a two-pollutant or three-pollutant models. This robustness can be interpreted as an indication that the association is stronger for this air pollutant. The large fall in the magnitude of the coefficient estimates for other pollutants suggests that they are less likely to be acting as risk factors for mortality in the elderly.

5.2.5. Harvesting Effects

Another question explored in this analysis was whether at least part of the of daily deaths associated with exposure to air pollution occurred in already compromised persons who would have died anyway in the immediate future without any elevation in air pollution. This assumption is supported by the observation of stronger effects of the high level episodes on higher age groups and in those chronically ill [Abercrombie,1953]. The bringing forward of deaths in these individuals by a few days or weeks has been termed as 'mortality displacement' or sometimes as the 'harvesting effect' [Spix et al,1993; Spix et al,1996].

This would imply that after an episode of air pollution and/or adverse weather circumstances (with resulting excess of deaths), a drop in mortality would occur a few days later since the pool of vulnerable individuals would be depleted. To verify if such a phenomenon was present in the mortality series in São Paulo, even though no extremely high episodes occurred during the studied period, several lags of PM_{10} were included simultaneously in a model for all cause mortality in the elderly. The coefficients obtained were plotted in Figure 5.12.

It can be seen that there is a positive effect of PM_{10} on the same day which persists up to one day. Then, a negative effect for lags 2 and 3 days is observed, which means a deficit of deaths during these days consistent with the harvesting phenomenon. These lagged variables for PM_{10} are obviously highly correlated and therefore, the coefficients are unlikely to be statistically significant when obtained simultaneously from the same model. Nevertheless, the patterns for lags 0 and 1 are quite distinct from lags 2 and 3 days and is unlikely that they have occurred by chance. From the fourth day onwards this previous pattern seems to repeat but point estimates are much closer, there is much variability and chance alone could explain them. Similar results were obtained for cardiovascular and respiratory mortality.

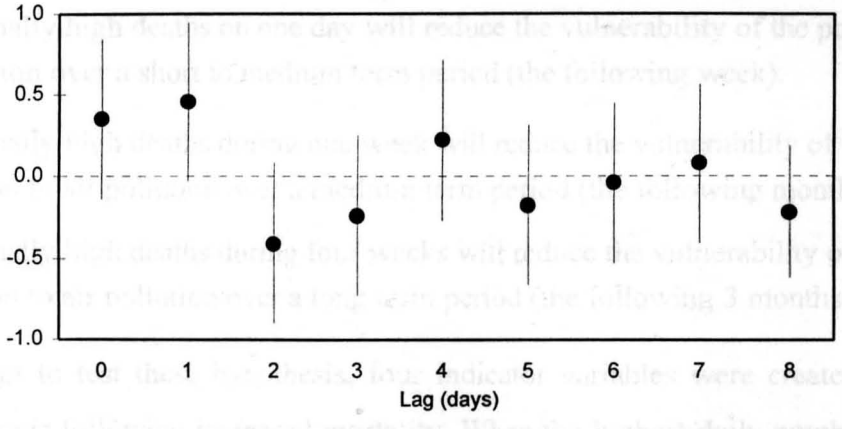


Fig 5.12 - Plot of coefficients of PM_{10} ($x10^3$) for different lags, in a model for all cause mortality in the elderly in São Paulo, Brazil, 1991-1993.

Another mechanism of the harvesting effect could be a change in population vulnerability, or to put it another way, following an episode of increased mortality the size of the susceptible population would become smaller as a result of the depletion of the more vulnerable individuals. The lower coefficient evident for winter deaths (5.2.3) is consistent with such a phenomenon. Consequently, the effect of air pollution on mortality after this period would also be smaller. The exact duration of this subsequent period of lower effect of air pollution is not known but it is very likely to be dependent on the size of the small, highly vulnerable subset of the population we assume the majority of the deaths comes from, and on the intensity of the harvesting effect.

In a study in Germany, Spix and collaborators [Spix et al,1996] investigated the harvesting effect by introducing into the regression model a variable summarising the sum

of deviations of the observed counts from the number of expected deaths. This variable was introduced as an effect modifier of the air pollution variable and they found that the mean of the previous 18 days provided the best fit for such analysis.

In this present investigation, it was decided to explore the duration of the harvesting effect in terms of short, medium and long-term effect. The general hypothesis was that unusually high mortality would reduce the population of highly vulnerable individuals, which would be manifested as a reduced risk of mortality related to levels of air pollution. This was investigated as four secondary hypotheses on different time scales:

- exceptionally high deaths on one day will reduce the vulnerability of the population to air pollution over a short-term period (the following 3 days).
- exceptionally high deaths on one day will reduce the vulnerability of the population to air pollution over a short to medium term period (the following week).
- exceptionally high deaths during one week will reduce the vulnerability of the population to air pollution over a medium term period (the following month).
- exceptionally high deaths during four weeks will reduce the vulnerability of the population to air pollution over a long term period (the following 3 months).

In an attempt to test these hypothesis, four indicator variables were created (F_1-F_4) to indicate periods following increased mortality. When the highest daily number of deaths in the previous 3 days was equal or greater than the 80^{th} percentile, then F_1 was set to 1 (otherwise F_1=0). Similarly, F_2 indicated a high mortality period in the previous 7 days, F_3 indicated a high weekly mortality in the previous month, and F_4 indicated a high monthly mortality in the previous 3 months. All cause mortality in the elderly excluding deaths due to external causes was used for this analysis.

These indicator variables were created after filtering the mortality data for long-term trends and one-year cycle. If unfiltered data were used, these indicators, especially the medium to long term ones, would be reflecting mainly the winter periods, when mortality is usually higher, and the more recent experience due to an expanding population. Filtering the long term trend and one year cycle was sufficient to smooth the data to a convenient degree (Fig 5.13). Additional analysis based on data filtered also for

meteorological variables were explored. However, this extra filtering did not make any substantial change to the earlier analyses.

These indicator variables were then used in a full model for all cause mortality in the elderly with an interaction term between the indicator and PM_{10} to identify if the PM_{10} effect would be reduced during periods following increased mortality. The results of these analyses are summarised in table 5.10.

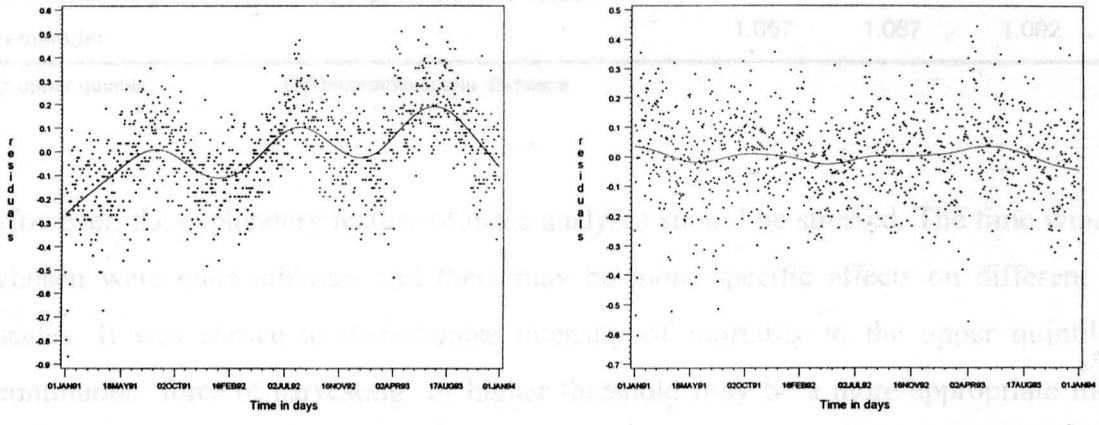


Fig 5.13 - Plot of the residuals of all cause mortality in the elderly before (left) and after (right) filtering for long term trend and one year cycle.

For all cause mortality, it can be observed that the effect of PM_{10} in days following a high mortality period are consistently lower than the effect after a period of normal or low mortality. This effect is observed in a short-term period (3 to 7 days) and medium to long-term period (one to three months). The decrease in coefficients ranges from 25 to 57%. However, none of the interaction terms tested were significant at the 5% level (*p*-values for the interaction terms in each model ranged from 0.4 to 0.75).

For cardiovascular diseases there is no evidence of short-term harvesting effect since RR are approximately similar or even higher following periods of high mortality. It could be however, that the harvesting effect for this deaths happens in a more medium term scale. The coefficient was lower during a month following at least one week with increased mortality. No evidence of harvesting effect could be observed for respiratory diseases mortality in the elderly. There might be an indication on the very short term with coefficients lower when following a period of high mortality in the previous 3 days.

Table 5.10 - Relative Risk of mortality for different causes for a 100μg/m^3 increase in PM_{10} during periods following high mortality or remainder periods.

Mortality indicator	All causes	CVD	Respiratory
at least one day of high* mortality in the past 3 days	1.024	1.069	1.086
remainder	1.053	1.061	1.095
at least one day of high* mortality in the past 7 days	1.025	1.076	1.198
remainder	1.050	1.051	1.046
at least one week of high* mortality in the past month	1.039	1.042	1.106
remainder	1.052	1.073	1.088
at least one month of high* mortality in the past 3 months	1.029	1.060	1.102
remainder	1.057	1.067	1.092

(* upper quintile) CVD=cardiovascular diseases

However, the exploratory feature of these analyses should be stressed. The time windows chosen were quite arbitrary and there may be more specific effects on different time scales. It was chosen to dichotomise intensity of mortality to the upper quintile. A continuous 'force of harvesting' or higher threshold may be a more appropriate model. Finally, the filtering to remove seasonal effects may be filtering out the stronger harvesting component, the winter excess. The interaction terms included in these models were not statistically significant at the 5% level. Therefore, these results should be viewed as only a weak indication of harvesting effect in the mortality for the elderly in the city of São Paulo, but further examination of this hypothesis needs to be considered.

5.2.6. Analysis by Age and Socioeconomic status

The results presented in the previous analysis showed statistically significant associations between exposure to some pollutants and mortality for all non-accidental causes and for respiratory and cardio-vascular causes in the elderly. However, no statistically significant effects at the 5% level could be found for respiratory deaths in children under 5 years of age. In order to explore further how the association between exposure to air pollution and mortality varies according to the age of the exposed individuals, additional analyses were performed in order to assess a possible modification of the pollution-mortality association by age.

To accomplish such an analysis, 11 different agegroups were defined. These included the agegroup of under 5 years of age which was examined previously in this study and additional agegroups defined as having an approximately similar number of individuals (Table 5.11). The outcomes used here were all cause mortality excluding external causes and respiratory mortality. These two outcomes were selected because they were the only ones that could provide sufficient numbers for the analysis of each separate agegroup. Nevertheless, for the agegroup of 5 to 12 years old only a few numbers of deaths were available for this analysis. As examined in chapter 4, neonatal deaths are strongly related to perinatal causes and were, therefore, excluded from this analysis.

The complete model building procedures were then performed for each separate agegroup to find the appropriate control variables to adjust for long, medium and short term temporal and meteorological patterns. It was decided to perform the model building for each agegroup because there might be different patterns of seasonality for people at different ages which could have been missed if the model building were conducted only on the aggregated dataset. A satisfactory degree of filtering for seasonal and temporal patterns was confirmed by spectral analysis of the residuals for each different agegroup and for both outcomes. In addition, the residuals were also examined for autocorrelation. One agegroup for each different outcome exhibited autocorrelation of order 1 after the filtering procedures.

Table 5.11 - Distribution of deaths according to different agegroups and percentage (%) of respiratory deaths in relation to all non-accidental deaths

agegroup	All causes	Respiratory	Respiratory as % of all causes
<5	6031	2351	39.0
5 to 12	737	144	19.5
12 to 40	16520	2004	12.1
40 to 50	15358	1379	9.0
50 to 60	17697	1614	9.1
60 to 65	13317	1329	10.0
65 to 70	15171	1771	11.7
70 to 75	14861	1972	13.3
75 to 80	15443	2381	15.4
80 to 85	13054	2390	18.3
85 or +	14525	2977	20.5

Then, Poisson regression analysis was performed for each separate agegroup using PM_{10} as the exposure variable. The coefficients obtained were transformed into Relative Risks and plotted (Fig 5.14 and 5.15). It can be seen that practically no effect was observed for PM_{10} on all cause mortality up to the age of 65 years old. For the elder groups there were positive associations although for the agegroup 75-80 years old the Relative Risk was close to unity. For respiratory mortality there were positive effects of PM_{10} for the agegroup 5-12 and 40-50 years old and then again for individuals older than 65 years of age in a pattern similar to all cause mortality. The effects for respiratory mortality were somewhat larger than for all cause mortality.

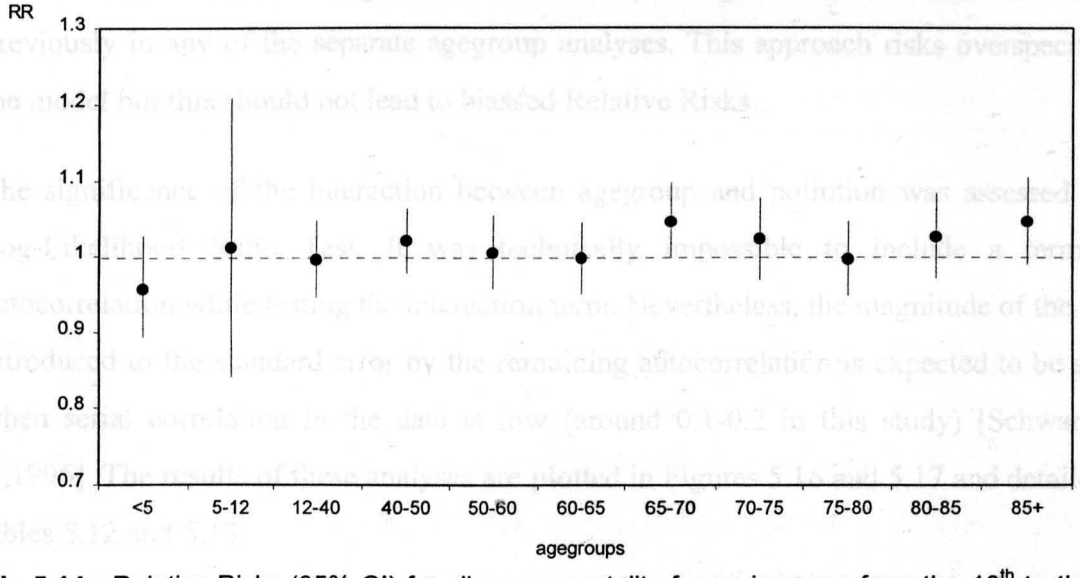


Fig 5.14 - Relative Risks (95% CI) for all cause mortality for an increase from the 10th to the 90th centile in levels of PM_{10} according to agegroup. (Individual fitting)

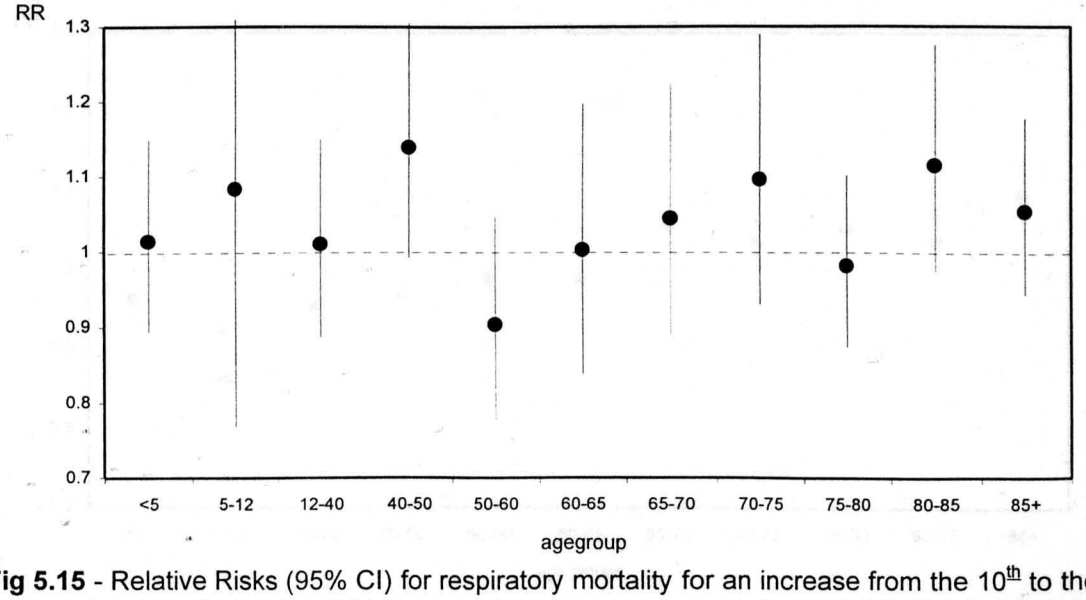


Fig 5.15 - Relative Risks (95% CI) for respiratory mortality for an increase from the 10th to the 90th centile in levels of PM_{10} according to agegroup. (Individual fitting). Upper limits for agegroups 5-12 and 40-50 are 1.50 and 1.30 respectively.

However, these results were obtained from individual analysis of each agegroup. To assess a possible modification of the effect of PM_{10} by age the 11 different agegroups were aggregated into a single dataset with the counts of deaths of each group identified by a categorical variable 'age'. A Poisson regression was then performed on this new dataset with adjustment for overdispersion, having all cause or respiratory mortality as the outcome, using PM_{10} as the exposure, and age group, the potential effect-modifier, treated as a categorical variable.

The model testing the interaction term of PM_{10} with 'age' was adjusted for trend, cyclical patterns and and meteorological confounders by using all control variables selected previously in any of the separate agegroup analyses. This approach risks overspecifying the model but this should not lead to biassed Relative Risks.

The significance of the interaction between agegroup and pollution was assessed by a Log-Likelihood Ratio Test. It was technically impossible to include a term for autocorrelation while testing the interaction term. Nevertheless, the magnitude of the error introduced to the standard error by the remaining autocorrelation is expected to be small when serial correlation in the data is low (around 0.1-0.2 in this study) [Schwartz et al,1996]. The results of these analyses are plotted in Figures 5.16 and 5.17 and detailed in tables 5.12 and 5.13.

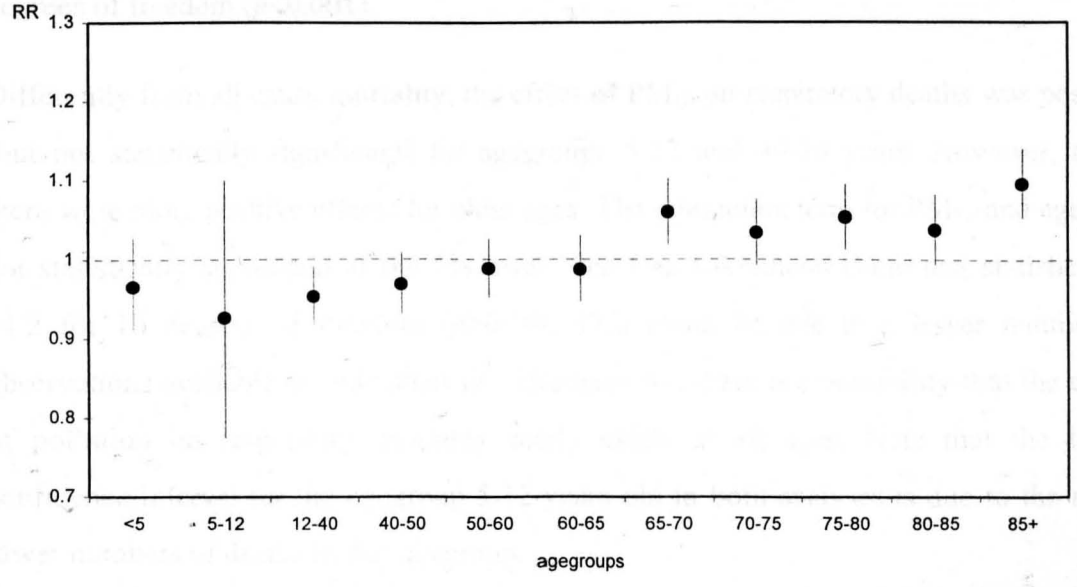


Fig 5.16 - Relative Risks (95% CI) for all cause mortality for an increase from the 10th to the 90th centile in levels of PM_{10} according to agegroup. (Interaction analysis)

Table 5.12 - Poisson regression coefficients and Relative Risk (RR) of death for all causes for an increase from the 10th to the 90th centile in levels of PM_{10} according to agegroup.

	coefficient	Std Err	RR	95% CI
PM_{10}*age <5	-0.549	0.495	0.97	(0.907 - 1.027)
PM_{10}*age 5-12	-1.202	1.395	0.93	(0.774 - 1.100)
PM_{10}*age 12-40	-0.759	0.313	0.95	(0.916 - 0.991)
PM_{10}*age 40-50	-0.503	0.322	0.97	(0.930 - 1.008)
PM_{10}*age 50-60	-0.190	0.301	0.99	(0.951 - 1.026)
PM_{10}*age 60-65	-0.204	0.341	0.99	(0.945 - 1.030)
PM_{10}*age 65-70	0.897	0.316	1.06	(1.018 - 1.102)
PM_{10}*age 70-75	0.511	0.321	1.03	(0.992 - 1.076)
PM_{10}*age 75-80	0.800	0.314	1.05	(1.012 - 1.095)
PM_{10}*age 80-85	0.569	0.340	1.04	(0.994 - 1.082)
PM_{10}*age 85+	1.423	0.319	1.10	(1.052 - 1.140)

coefficients and Standard Error (SE) X 10^3

This analysis confirmed that the effect of PM_{10} on mortality for all non-accidental causes varies according to age. There is a negative effect for the agegroups under 5, 5 to 12, 12 to 40, and 40 to 50 years old although it is only statistically significant for the group 12 to 40 as indicated by the Relative Risk and 95% confidence interval. There is no evidence at all of effect for the agegroups up to 65 years old. Over the age of 65 there is a clear and statistically significant positive effect of PM_{10}. The effect modification of PM_{10} by age was statistically significant with a Log Likelihood Ratio test statistic of 63.5 for 10 degrees of freedom ($p<0.001$).

Differently from all cause mortality, the effect of PM_{10} on respiratory deaths was positive (but not statistically significant) for agegroups 5-12 and 40-50 years. However, again there were more positive effects for older ages. The interaction term for PM_{10} and age was not statistically significant at the 5% level. The Log Likelihood Ratio test statistic was 14.2 for 10 degrees of freedom ($p>0.10$). This could be due to a lesser number of observations available for this analysis. Alternatively, there is a possibility that the effect of pollution on respiratory mortality really exists at all ages. Note that the larger confidence interval for the agegroup 5-12 years old in both analyses is due to the much lower numbers of deaths in this agegroup.

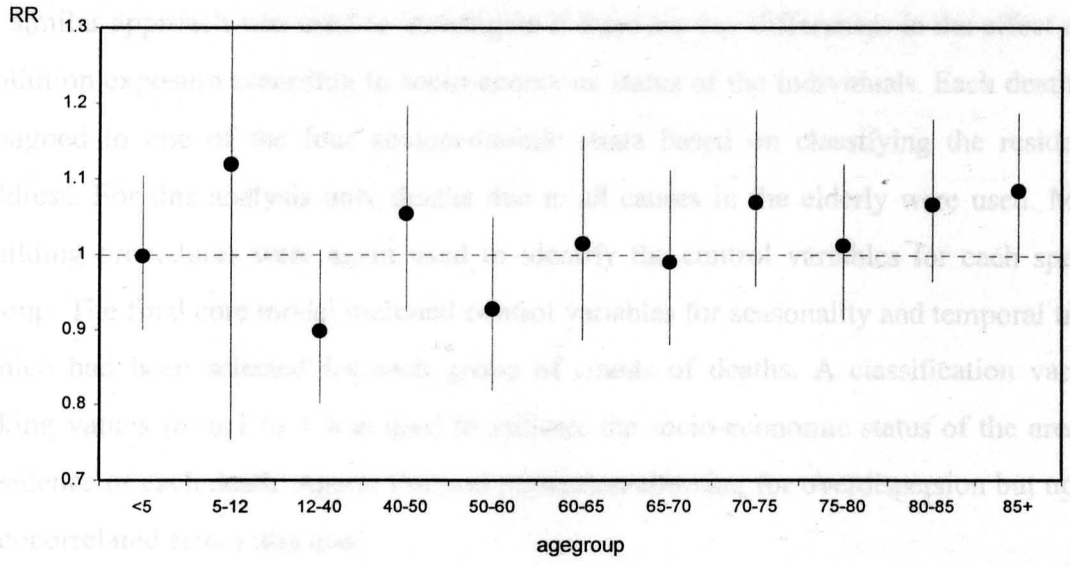


Fig 5.17 - Relative Risks (95% CI) for respiratory mortality for an increase from the 10th to the 90th centile in levels of PM_{10} according to agegroup. (Interaction analysis). Upper limit for agegroup 5-12 is 1.60.

Table 5.13 - Poisson regression coefficients and Relative Risk (RR) of death for respiratory diseases for an increase from the 10th to the 90th centile in levels of PM_{10} according to agegroup.

	coefficient	Std Err	RR	95% CI
PM_{10}*age <5	-0.034	0.819	1.00	(0.899 - 1.105)
PM_{10}*age 5-12	1.753	3.036	1.12	(0.753 - 1.619)
PM_{10}*age 12-40	-1.668	0.906	0.90	(0.801 - 1.006)
PM_{10}*age 40-50	0.814	1.030	1.05	(0.924 - 1.198)
PM_{10}*age 50-60	-1.168	0.991	0.93	(0.818 - 1.050)
PM_{10}*age 60-65	0.224	1.059	1.01	(0.887 - 1.157)
PM_{10}*age 65-70	-0.157	0.933	0.99	(0.879 - 1.112)
PM_{10}*age 70-75	1.040	0.871	1.07	(0.957 - 1.192)
PM_{10}*age 75-80	0.186	0.812	1.01	(0.913 - 1.120)
PM_{10}*age 80-85	1.014	0.799	1.07	(0.964 - 1.179)
PM_{10}*age 85+	1.281	0.723	1.09	(0.991 - 1.188)

coefficients and Standard Error (SE) X 10^3

In conclusion, it seems that there is a modification of the effect of PM_{10} by age with positive and statistically significant effects more frequent for individuals over 65 years of age. However, this effect is more evident when considering mortality for all causes excluding accidents and violent deaths. Identical analysis for respiratory deaths showed some evidence for this pattern but with less consistent results.

A similar approach was used to investigate if there are any differences in the effect of air pollution exposure according to socio-economic status of the individuals. Each death was assigned to one of the four socioeconomic strata based on classifying the residential address. For this analysis only deaths due to all causes in the elderly were used. Model building procedures were again used to identify the control variables for each specific group. The final core model included control variables for seasonality and temporal trends which had been selected for each group of counts of deaths. A classification variable taking values from 1 to 4 was used to indicate the socio-economic status of the areas of residence of each death. Again, Poisson regression allowing for overdispersion but not for autocorrelated errors was used.

It could be seen that there were limited differences in the association of deaths for all causes with air pollution in the elderly according to the socioeconomic conditions of the place of residence (Fig 5.18 and table 5.14). There is a slight increase in risk for individuals who lived in more wealthy areas compared to more deprived ones. These differences were not large enough to provide a statistically significant effect modification by socio-economic status (Log-Likelihood Ratio test statistic = 1.95 for 3 degrees of freedom (p>0.5)).

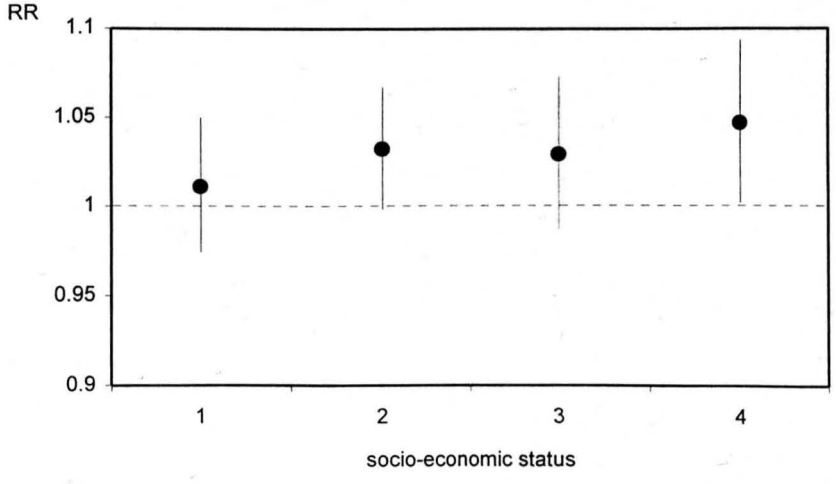


Fig 5.18 - Relative Risks (95% CI) for all cause mortality in the elderly for an increase from the 10^{th} to the 90^{th} centile in levels of PM_{10} according to a socioeconomic classification. Level 1 are the poorer individuals while level 4 are the richer ones. (Interaction analysis)

Table 5.14 - Poisson regression coefficients and Relative Risk (RR) of death for all causes in the elderly for an increase from the 10^{th} to the 90^{th} centile in levels of PM_{10} according to socioeconomic status.

	coefficient	Std Err	RR	95% CI
PM_{10}*SES 1	0.175	0.300	1.01	(0.974 - 1.050)
PM_{10}*SES 2	0.496	0.267	1.03	(0.998 - 1.067)
PM_{10}*SES 3	0.449	0.334	1.03	(0.987 - 1.073)
PM_{10}*SES 4	0.712	0.349	1.05	(1.002 - 1.094)

coefficients and Standard Error (SE) X 10^3

CHAPTER 6.
TIME SERIES: HOSPITAL ADMISSIONS

This chapter presents the results of the time series analysis of daily counts of hospital admission for children in relation to air pollution in the city of São Paulo. The time period covered by these data is different from the one used in the mortality analysis. Consequently, air pollutants and all other covariates comprised a slightly different dataset. These differences have already been highlighted in the methods and descriptive analysis chapters. Nevertheless, the approaches used for this analysis were essentially the same as those used before for mortality. Therefore, the detailed results of each step in the model building procedures are only briefly summarised. As mentioned before, the analysis of hospital admissions was restricted only to children under 5 years of age and for a selection of diagnoses.

6.1. MODEL SELECTION PROCESS

As with some of the mortality analysis, the model selection for hospital admission was performed using simple linear regression with a log-transformed outcome. At each step, the inclusion of a variable in the model was assessed by F-tests. The procedures used were the same as described for the analysis of mortality data, namely spectral analysis of residuals, use of diagnostic plots to confirm the adequacy of the variables selected, cross-correlation functions for temperature and the assessment of autocorrelation. Therefore, this section will only highlight the main differences noted when modelling the hospital admissions outcomes. The variables selected for this part of the study were admissions for all respiratory diseases in children under 5 years old, pneumonia infections in children under 5 and under 1 years old, asthma admissions in children under 5 and a control outcome diarrhoeal diseases in children under 5 years old.

When adjusting for long term trends, seasonal patterns, and short term temporal effects, the first peculiarity noticed was the characteristic pattern of seasonality present in these data. After decomposing the raw series using spectral analysis and plotting the periodogram, large periodicities for periods around and above one year, one week (7 days), 3.5 days and at 2 days were evident (Fig 6.1). These indicated firstly that, although smaller, there was also a long time trend pattern in this series as well as a periodical cycle of one year. Secondly, the weekly periodicity confirmed the suspicion that admissions to hospital would have a stronger day-of-week pattern due to inherent characteristics of the routine of medical care services.

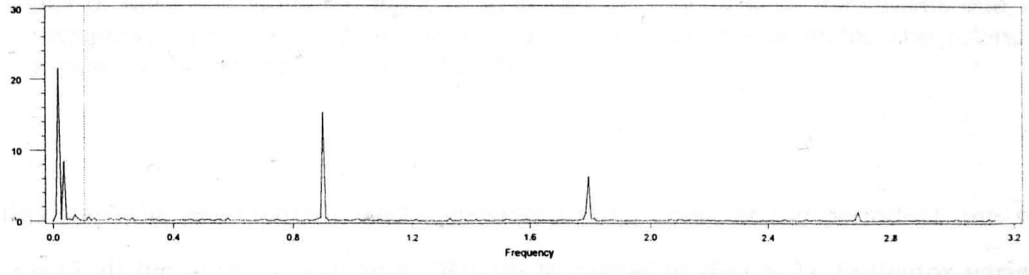


Fig 6.1 - Periodogram of the residuals of unfiltered log-transformed data for hospital admissions from all respiratory causes in children under 5 years old. The vertical dotted line refers to a frequency equivalent to a period of two months. The 5 more prominent spikes identified in this graph refer to (from left to right): long term trends, one year cycle, one week, 3.5 days and 2 days.

When testing and introducing variables to control for such periodicities it was noted that a time trend variable was not necessary in the model (except for pneumonia in <1 year) since its contribution was not statistically significant at the 10% level. Indicator variables for years were allowed into the models for pneumonia (both age-groups) and for diarrhoea. Nevertheless, when introducing the sine and cosine waves for a one year period, the long term periodicity present in the periodogram was then accounted for. Sine and cosine waves down to 2 months were then introduced into each model according to the significance of their contribution to the model fit.

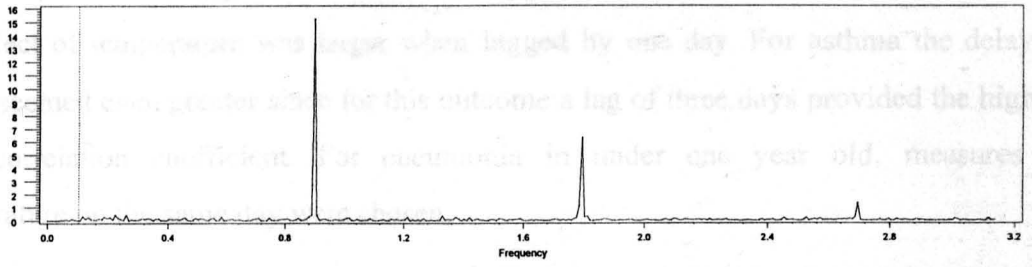


Fig 6.2 - Periodogram of the residuals after adjustment for long term trends and seasonal and temporal patterns down to 2 months for hospital admissions from all respiratory causes in children under 5 years old. The vertical dotted line refers to a frequency equivalent to a period of two months.

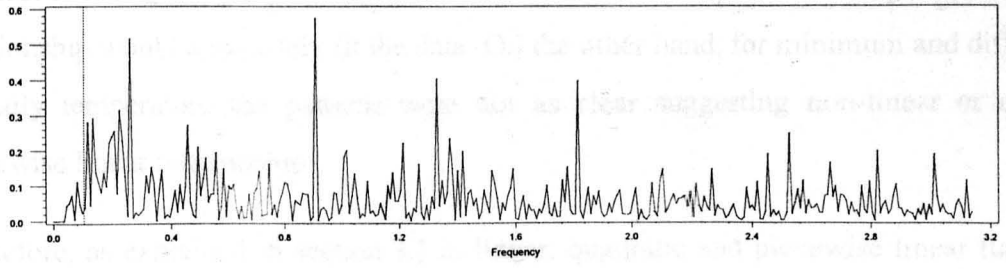


Fig 6.3 - Periodogram of the residuals of the fully adjusted model including indicator variables for day of week and terms for meteorological variables for hospital admissions due to all respiratory causes in children under 5 years old. The vertical dotted line refers to a frequency equivalent to a period of two months.

After this step, the periodograms still exhibited a large spike at the period of one week which made all the other periodicities difficult to visualise (Fig 6.2). Indicator variables for each day of the week were then tested. Their contribution to the model's fit was highly significant and they were entered into the models. Public holidays were also significant for all models tested and were, therefore, included. As expected the inclusion of these

variables were responsible for removing the large weekly periodicities observed earlier as well as the two other smaller ones at 3.5 and 2 days producing a substantial change in the scale of the plot (Fig 6.3). No significant contribution was noted for strikes, either in the public transport provision or in the health services.

Residuals of the models up to this point were adjusted for autocorrelation and then cross-correlated with the residuals of the temperature variables, after applying the same core model. The cross-correlation function plots and tables allowed us to identify at which lag the effect of each indicator of temperature was greatest. Interestingly, lagged effects of temperature seemed to have a greater effect than measures taken on the same day. For admissions from all respiratory causes, from pneumonia in under 5 and from diarrhoea, the effect of temperature was larger when lagged by one day. For asthma the delay in effect seemed even greater since for this outcome a lag of three days provided the highest cross-correlation coefficient. For pneumonia in under one year old, measures of temperature on the same day were chosen.

Subsequently, the adjusted residuals of each outcome were plotted against the four indicators of daily temperature at the appropriate lags. It was possible then to examine individually the shape of the relationship between temperature and hospital admissions. In contrast to mortality, the plots indicated that for mean and maximum temperature a linear relationship would adequately fit the data. On the other hand, for minimum and difference in daily temperature the patterns were not as clear suggesting non-linear or at least piecewise linear relationships.

Therefore, as explained in section 2.1.2, linear, quadratic and piecewise linear functions were created for each measure of temperature based on the plots described above. This was an attempt to adjust for the effects of temperature on hospital admissions with flexibility for the different shapes of this relationship. These different functions were then tested in linear regression models to assess their contribution to the fit. The best contributions (as assessed by F-tests) were obtained by linear terms of maximum temperature. For asthma and pneumonia infections in under one years old, linear mean temperature gave the best fit.

The next step was to select an appropriate measure of relative humidity. Linear and quadratic terms were tested using the same lag structure as for temperature. Linear terms for maximum or minimum relative humidity provided the best fit and were then introduced into the model. Similarly to mortality, there were no significant contributions from atmospheric pressure or precipitation in these models.

The final core model produced for each outcome including controlling variables for time trends, seasonality and meteorological variables is summarised on Table 6.1. The residuals of these models were then plotted against the time period to examine if there were signs of temporal and seasonal patterns still not adjusted for. As observed in the example of Figure 6.4 the core model was successful in controlling for all these patterns.

Table 6.1 - Variables selected for the core model of each hospital admissions outcome.

	all respiratory	pneumonia	pneumonia < 1	asthma	diarrhoea
indicators for years		X	X		X
linear trend			X		
2 years	X	X	X	X	X
1 year	X	X	X	X	X
9 months	X	X	X	X	X
others up to 2 months	X	X	X	X	X
weekdays	X	X	X	X	X
holidays	X	X	X		X
Temperature	maximum	maximum	mean	maximum	mean
Relative Humidity	maximum	maximum	maximum	minimum	maximum

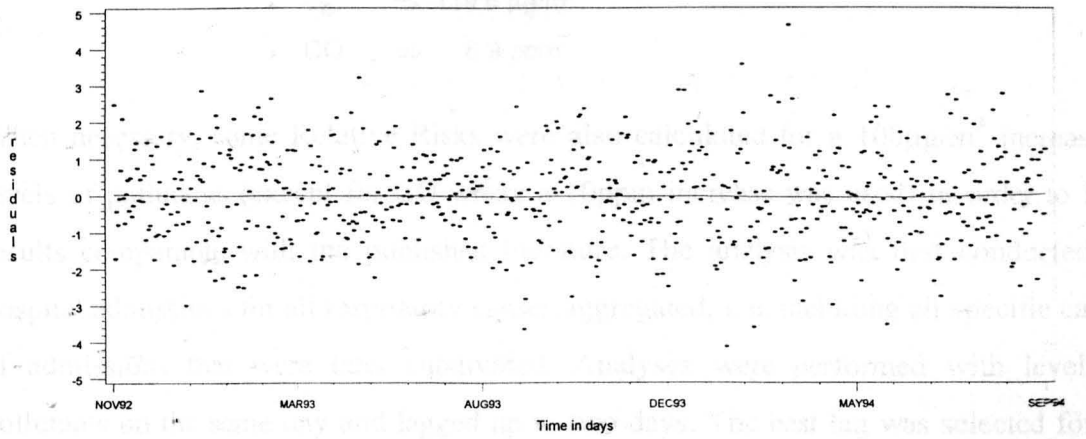


Fig 6.4 - Plot of the residuals of admissions for all respiratory causes from the core model adjusting for long and short temporal trends, plus meteorological variables versus time. A non-parametric smooth curve was run-through the data to indicate any remaining seasonal variability in the data.

After all these adjustments, the residuals of the core model were then checked for the possibility of any remaining autocorrelation. Autoregressive models were again used. The statistics produced and the Partial Autocorrelation Function (PACF) plots helped to identify remaining autocorrelation only for admissions for all respiratory causes and for pneumonia in under 5. For the other outcomes it seemed that the adjustment for trend and season was enough to correct the autocorrelation usually present in this type of data.

The plot of the predicted values after each step in the model building is presented in Appendix C. The plots confirmed the adequacy of the variables chosen as providing a good description of the data.

6.2. MAIN FINDINGS

Finally, the exposure variables were entered into the final core models produced for each outcome. Poisson regression models allowing for overdispersion and when necessary for autocorrelated errors were used. Relative Risks were again calculated for an increment from the 10^{th} to the 90^{th} centile in levels of each air pollutant. These values for the two year period covered by the hospital admission data were:

- $PM_{10} \Rightarrow 98.1\ \mu g/m^3$
- $SO_2 \Rightarrow 27.1\ \mu g/m^3$
- $NO_2 \Rightarrow 319.4\ \mu g/m^3$
- $O_3 \Rightarrow 119.6\ \mu g/m^3$
- CO $\Rightarrow$ 6.9 ppm

When necessary, some Relative Risks were also calculated for a $100\mu g/m^3$ increase in levels of pollutants (except for CO where a 10ppm increase was used) in order to have results comparable with the published literature. The analysis was first conducted for hospital admissions for all respiratory causes aggregated, i. e. including all specific causes of admissions that were later subdivided. Analyses were performed with levels of pollutants on the same day and lagged up to two days. The best lag was selected for the tables presented in this chapter but complete tables can be found in Appendix D.

Daily variations in levels of Ozone and NO_2 had positive effects on hospital admissions for all respiratory diseases in children under 5 years old. The coefficient estimate for O_3 reached the statistically significant 5% level while for NO_2 it was only of borderline significance (Fig 6.5).. The effects of other pollutants were also positive but not statistically significant. The effects were larger at exposures on the same day for PM_{10}, NO_2, O_3 and CO and at lag one day for SO_2.

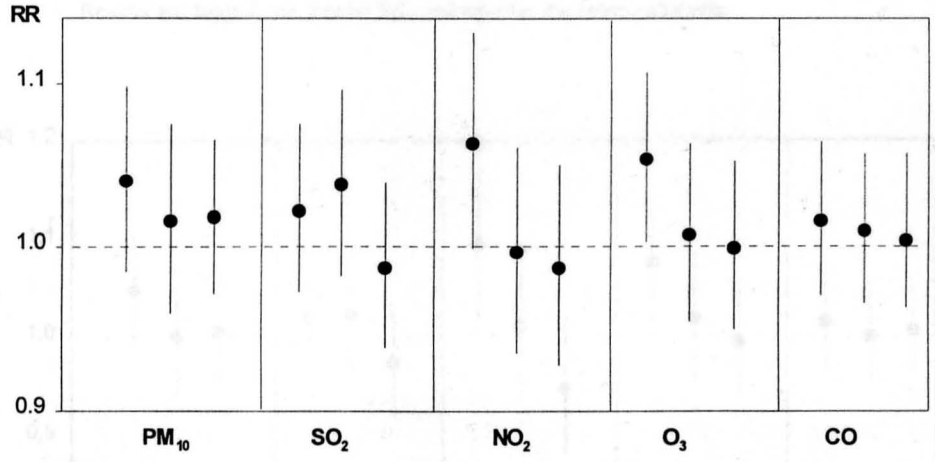


Fig 6.5 - Relative Risks (RR) and 95% CI for admissions from all respiratory causes in under 5 years old for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most, respectively).

An increase from the 10th to the 90th centile in levels of O_3 and NO_2 corresponded to a 5% and 6% increase in admissions for respiratory diseases in children under 5 years old. The Relative Risks for other pollutants were about 1.04 for PM_{10} and SO_2, and 1.02 for CO. The regression coefficients, standard errors, the Relative Risks and 95% Confidence Intervals for the association between air pollutants and hospital admission for all respiratory diseases are summarised in Table 6.2.

The following step was to explore if an effect of air pollution variables could be observed in more specific causes of respiratory admission, namely pneumonia infections and asthma. Figure 6.6 displays the Relative Risks and 95% confidence intervals for admissions from pneumonia infections in children under 5 years of age associated with exposure to air pollution.

Table 6.2 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for all respiratory diseases in children under 5 years old. RR are for an increase from the 10th to the 90th centile in daily levels of each pollutant.

	coefficient	Std Err	RR	95% CI
PM_{10}	0.403	0.286	1.040	0.985 - 1.099
SO_2	1.378	1.026	1.038	0.983 - 1.096
NO_2	0.192	0.099	1.063	0.999 - 1.132
O_3	0.438	0.211	1.054	1.003 - 1.107
CO	2.387	3.397	1.017	0.971 - 1.065

(coefficients and SE) X 10^3
Results are lagged one day for SO_2 and zero for the other pollutants.

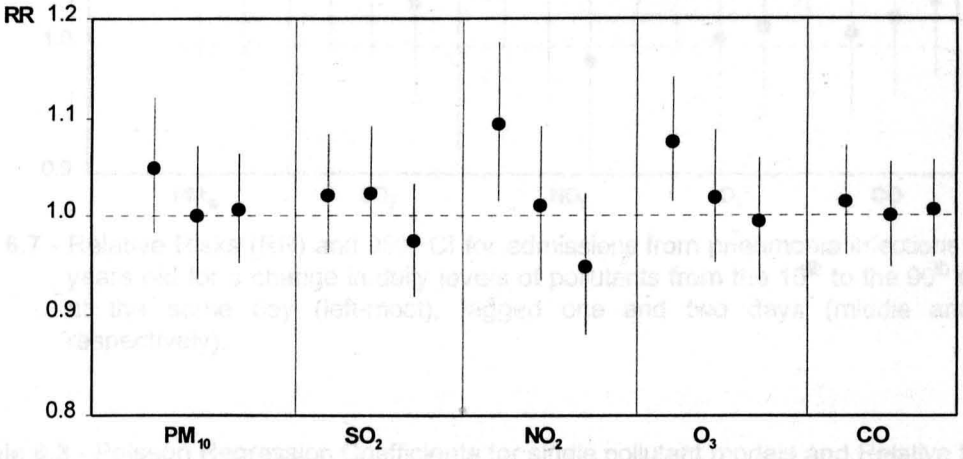


Fig 6.6 - Relative Risks (RR) and 95% CI for admissions pneumonia infections in under 5 years old for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most), lagged one and two days (middle and right-most, respectively).

For pneumonia in under 5 years old, O_3 and NO_2 showed positive and statistically significant associations at the 5% level. It can be noted that the magnitude of the effects and the lag pattern of the Relative Risks were similar to those of all respiratory admissions. Additionally, the estimates of effect for these pollutants were greater than for admissions for all respiratory diseases. An increase in NO_2 from the 10th to the 90th centile was associated with a 9% increase in hospital admissions for pneumonia in children under 5 years old. For O_3 this increase was estimated to be 8% (Table 6.3).

When analysing the admissions for pneumonia only in children under one year old (excluding neonatal cases) it was noted that the risk estimates for PM_{10} and SO_2 were higher demonstrating a possibly stronger effect of these air pollutants on this age-group.

Moreover, the estimate of effect was statistically significant at the 5% level for PM_{10} and of borderline significance for SO_2 (Fig 6.7). For an increase from the 10^{th} to the 90^{th} centile in levels of PM_{10}, a 9% increase in admissions was observed. For SO_2, an increase of 7% in admissions was associated with the same variation in pollution (Table 6.3).

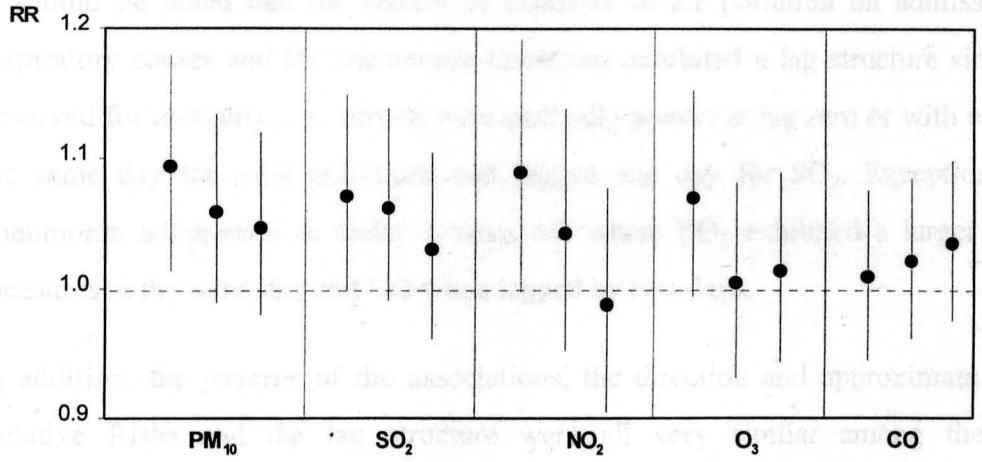


Fig 6.7 - Relative Risks (RR) and 95% CI for admissions from pneumonia infections in under one years old for a change in daily levels of pollutants from the 10^{th} to the 90^{th} centile. Effect at the same day (left-most), lagged one and two days (middle and right-most, respectively).

Table 6.3 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for pneumonia infections in children under 5 and under 1 years old. RR are for an increase from the 10^{th} to the 90^{th} centile in daily levels of each pollutant.

	coefficient	Std Err	RR	95% CI
Pneumonia <5				
PM_{10}	0.498	0.338	1.050	0.984 - 1.121
SO_2	0.873	1.191	1.024	0.961 - 1.091
NO_2	0.279	0.117	1.093	1.016 - 1.177
O_3	0.615	0.253	1.076	1.014 - 1.142
CO	2.087	3.998	1.015	0.961 - 1.071
Pneumonia <1				
PM_{10}	0.913	0.396	1.094	1.013 - 1.180
SO_2 (lag 0)	2.537	1.331	1.071	0.998 - 1.149
NO_2	0.271	0.144	1.091	0.996 - 1.193
O_3	0.566	0.317	1.070	0.993 - 1.152
CO (lag2)	5.028	4.394	1.035	0.975 - 1.099

(coefficients and SE) X 10^3

Note: Results are lagged one day for SO_2 and zero for the other pollutants otherwise stated.

Consistently, the size of the effects of NO_2 and O_3 were approximately the same as for pneumonia admissions in under 5 years old (9% and 7% increase in admissions, respectively). However, both coefficient estimates were now only of borderline significance. CO also exhibited a much greater coefficient for pneumonia in children under one year old, however, it continues to be not statistically significant.

It should be noted that the effects of exposure to air pollution on admissions for all respiratory causes and for pneumonia infections exhibited a lag structure similar to that observed for mortality, i.e., effects were generally greater at lag zero or with exposures on the same day for most pollutants and lagged one day for SO_2. Exceptions were for pneumonia admissions in under 1 years old where SO_2 exhibited a larger effect with measures at the same day and CO when lagged by two days.

In addition, the patterns of the associations, the direction and approximate size of the Relative Risks and the lag structure were all very similar among these hospital admissions. The very small Relative Risks exhibited for CO for all these outcomes should also be pointed out.

The analysis of admissions for asthma and other wheeze conditions revealed a pattern different from that of other respiratory diseases (Fig 6.8). For most pollutants the risk of admissions for asthma is higher for longer lag periods. With the exception of CO which exhibited a higher effect at lag zero, all the other air pollutants had larger coefficients when lagged by two days. However, none of these coefficients was statistically significant at the 5% level.

The results for admissions due to asthma are presented in Table 6.4. It should be noted that the Relative Risks for SO_2 and NO_2 were much higher for this disease than for other respiratory admissions.

The effects of daily variations in levels of air pollutants on hospital admissions for diarrhoeal diseases is summarised on Figure 6.9. It can be seen that most coefficients obtained reflected a negative relationship but were not statistically significant at the 5% level. Exceptions were PM_{10} and SO_2 which exhibited a statistically significant negative association with admissions for diarrhoea when lagged two days.

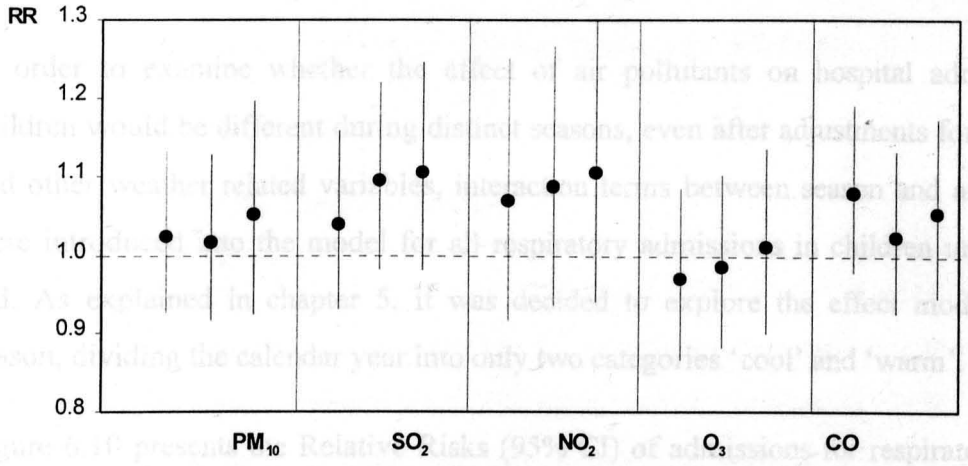


Fig 6.8 - Relative Risks (RR) and 95% CI for admissions from asthma in under 5 years old for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most). lagged one and two days (middle and right-most. respectively).

Table 6.4 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for asthma in children under 5 years old. RR are for an increase from the 10th to the 90th centile in daily levels of each pollutant.

	coefficient	Std Err	RR	95% CI
PM_{10}	0.513	0.677	1.052	0.923 - 1.198
SO_2	3.732	2.254	1.106	0.981 - 1.247
NO_2	0.317	0.259	1.107	0.940 - 1.300
O_3	0.089	0.499	1.011	0.899 - 1.136
CO (lag 0)	11.267	7.214	1.081	0.980 - 1.192

(coefficients and SE) X 10^3

Note: Results are lagged two days for all pollutants except CO.

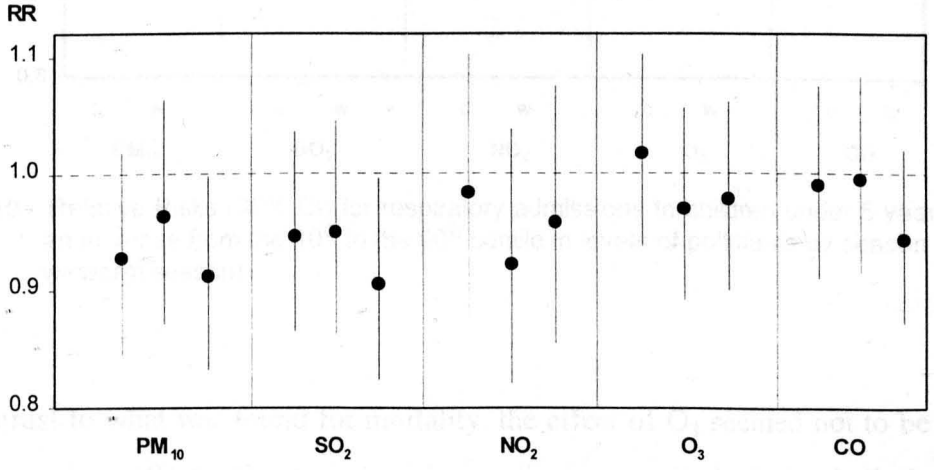


Fig 6.9 -Relative Risks (RR) and 95% CI for admissions from diarrhoea in under 5 years old for a change in daily levels of pollutants from the 10th to the 90th centile. Effect at the same day (left-most). lagged one and two days (middle and right-most. respectively).

6.2.1. Effects by Season

In order to examine whether the effect of air pollutants on hospital admissions for children would be different during distinct seasons, even after adjustments for seasonality and other weather related variables, interaction terms between season and air pollutants were introduced into the model for all respiratory admissions in children under 5 years old. As explained in chapter 5, it was decided to explore the effect modification by season, dividing the calendar year into only two categories 'cool' and 'warm'.

Figure 6.10 presents the Relative Risks (95% CI) of admissions for respiratory diseases associated with all air pollutants studied and according to each season. It can be seen that for PM_{10}, SO_2 and CO the effect estimates were greater in the cool season (winter plus second half of autumn and first half of spring). However, these differences were not statistically significant. For NO_2 the effect was greater during the warm season and it reached statistical significance at the 5% level.

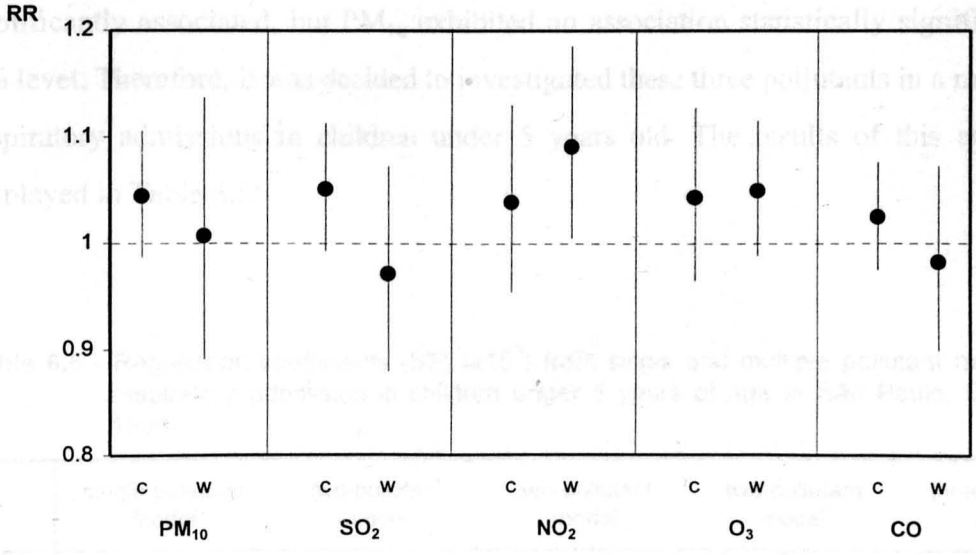


Fig 6.10 - Relative Risks (95% CI) for respiratory admissions in children under 5 years of age for an increase from the 10th to the 90th centile in levels of pollutants by season (c=cool and w=warm season)

In contrast to what was found for mortality, the effect of O_3 seemed not to be influenced by season since effect estimates were very similar in magnitude during both the warm and

the cool season. However, it has to be pointed out that the interactions of each air pollutant and season were not statistically significant at the 5% level.

6.2.2. Multiple Pollutant Models

It was observed that more than one air pollutant exhibited an statistically significant association with hospital admissions for all respiratory diseases and for pneumonia infections in children. In an attempt to distinguish which pollutant is mainly responsible for these associations, a multiple pollutant analysis was performed. Once more, it was decided to include in the same model only the pollutants that exhibited a significant association in the single pollutant models.

In the time series analysis of hospital admissions it was observed that NO_2 and O_3 exhibited close to or statistically significant associations with admissions for all respiratory diseases and for pneumonia infections in children under 5 years of age. For pneumonia admissions in children under one year old the same pollutants were not quite significantly associated, but PM_{10} exhibited an association statistically significant at the 5% level. Therefore, it was decided to investigated these three pollutants in a model for all respiratory admissions in children under 5 years old. The results of this analyses are displayed in Table 6.5.

Table 6.5 - Regression coefficients (SE) ($x10^3$) from single and multiple pollutant models for all respiratory admission in children under 5 years of age in São Paulo, Brazil, 1992-1994.

	single pollutant model		two-pollutant model		two-pollutant model		two-pollutant model		three-pollutant model	
O_3	0.44	(0.21)	0.58	(0.26)	0.37	(0.21)			0.52	(0.27)
NO_2	0.19	(0.10)	0.15	(0.10)			0.13	(0.11)	0.11	(0.12)
PM_{10}	0.40	(0.29)			0.58	(0.33)	0.57	(0.49)	0.44	(0.51)

When both O_3 and PM_{10} were considered in different models with NO_2, their effect estimates increased by a third or more. Moreover, when O_3 and PM_{10} are included in the

same model, O_3 showed a decrease on its coefficient estimate while PM_{10} showed a substantial increase. In a model considering all three pollutants, O_3 and PM_{10} still exhibited increases, though smaller, in their estimated effects.

On the other hand, in all the models studied NO_2 exhibited reduced estimates. It should be noted, however, that decreases in effect estimates for NO_2 were larger when PM_{10} was present in the model. In addition, it has to be pointed out that although PM_{10} had increases on its estimated effect in multipollutant models, the Standard Errors of its estimate also exhibited substantial increases. For O_3, however, the Standard Errors remained almost unchanged and therefore, this pollutant showed statistically significant or borderline significant estimates in multiple pollutant models.

6.2.3. Harvesting Effects

Similarly to the investigation for mortality, it was decided to check if a harvesting effect is present in the series of hospital admissions in children. The principle is more or less the same, i.e. to ascertain if at least part of the hospital admissions in children under 5 years old for respiratory diseases due to exposure to air pollution occurred in children at a higher risk of developing a respiratory disease. These more susceptible children would comprise the fraction that develop a respiratory disease severe enough to require hospitalisation.

Here the concept of harvesting should be understood as depletion of a pool of susceptible individuals leading to subsequent fewer number of cases following a period with a relatively large number of cases. The term 'displacement' used for mortality sometimes nearly interchangeably with harvesting effect, does not entirely apply for hospital admissions.

Again this would mean that after an episode of air pollution and/or adverse weather conditions, a reduction in the daily number of admissions would be observable in the series of hospitalisations due to reduction in the number of more susceptible children outside the hospital. The length of the period with lower than normal admissions is likely

to vary according to the severity of the air pollution/weather episode and to the number of highly susceptible children that comprises the population of under 5 years old.

Including several lags of daily maximum levels of O_3 simultaneously in a model for all respiratory admissions in children under 5 years old it can be observed that the effect of O_3 decreases and becomes slightly negative at lag 2 days (Fig 6.11). However, in contrast to mortality, the pattern of coefficients over time appears much more irregular, and it is thus difficult to draw any specific conclusions.

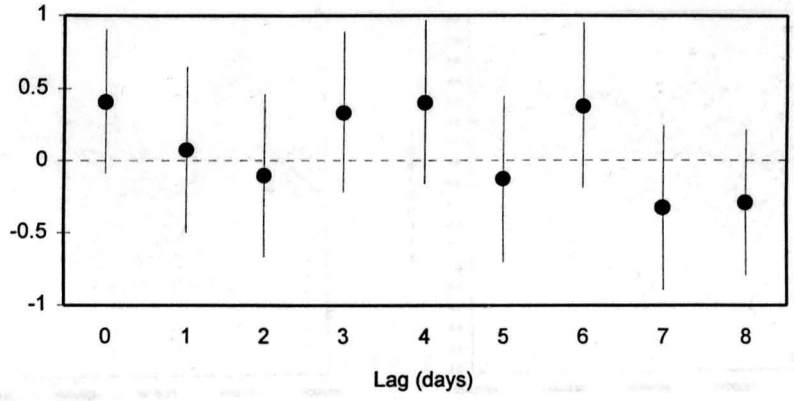


Fig 6.11 - Plot of coefficients of O_3 ($x10^3$) for different lags, in a model for all respiratory admissions in children under 5 years old in São Paulo, Brazil, 1991-1993.

Another way of examining the presence of a harvesting effect in the series of hospital admissions would be that following a period of increased admissions the size of the susceptible population of children would become smaller. Consequently, the effect of air pollution on hospital admissions after this period would also be smaller.

No studies have yet attempted to examine the presence of a harvesting effect for hospital admissions. Therefore, no previous indication of its duration is available. It was decided to explore the duration of this harvesting effect in terms of short, medium and long-term effects. The general hypothesis was that unusually high periods of hospital admissions would reduce the population of highly vulnerable individuals, which would be manifested as a subsequent reduced risk of hospital admissions related to levels of air pollution. This was again investigated as four secondary hypotheses at different time scales similarly to

mortality. In summary, four indicator variables were created to indicate periods of increased hospital admissions.

No substantial long term trend was observed for the series of admissions for respiratory diseases in children. Therefore, filtering only for one year cycle using harmonic waves was sufficient to smooth the data to a satisfactory degree (Fig 6.12). The indicator variables were created based on the filtered data to avoid mis-representation of winter periods, when the daily number of admissions is usually higher than average.

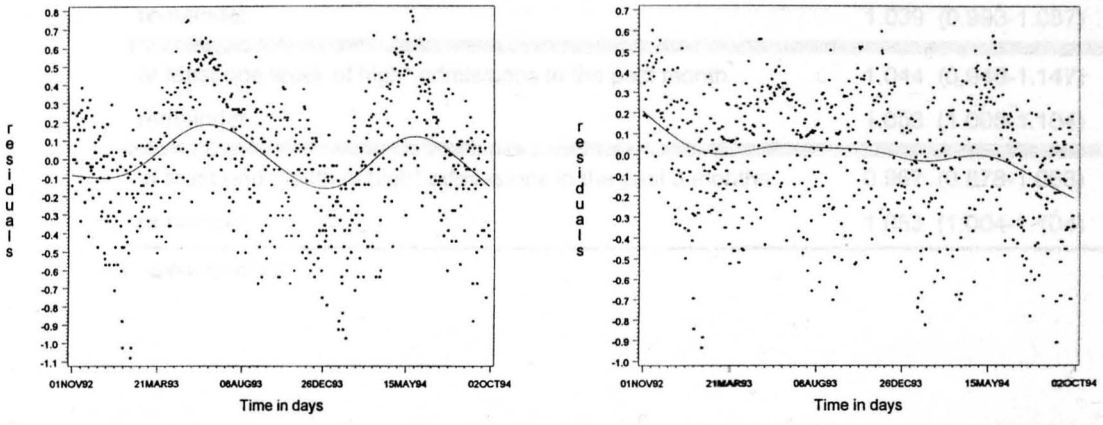


Fig 6.12 - Plot of the residuals of respiratory admissions in children under 5 years old before (left) and after (right) filtering for one year cycle.

The harvesting effect was then assessed by the parameter of an interaction term of these four indicators with daily maximum level of O_3. These interactions terms identified if the effect of O_3 was reduced following periods of high number of admissions. The model used was all respiratory admissions for children under 5 years of age adjusted for seasonal and meteorological patterns. The coefficients obtained were transformed into Relative Risks and are displayed in Table 6.6.

The direction of the point estimates suggests the presence of a harvesting effect on the short term (3 days). For medium term the RR are approximately of the same size but the difference increases again for a long term scale (3 months). However, the *p*-values for the interaction terms of different indicators of periods of high admissions with O_3 ranged from 0.11 to 0.98, i.e. none of them reached the statistical significance level. Therefore,

this analysis provides only some evidence of a possible harvesting effect in the hospital admissions for respiratory diseases in children in São Paulo.

Table 6.6 - Relative Risk (95% CI) of hospital admissions for respiratory diseases in children under 5 years old for a 100μg/m^3 increase in O_3 during periods following high number of admissions or remainder periods.

Period	RR
at least one day of high* admissions in the past 3 days	1.004 (0.908-1.109)
remainder	1.047 (1.000-1.097)
at least one day of high* admissions in the past 7 days	1.038 (0.930-1.156)
remainder	1.039 (0.993-1.087)
at least one week of high* admissions in the past month	1.044 (0.949-1.147)
remainder	1.053 (1.005-1.104)
at least one month of high* admissions in the past 3 months	0.967 (0.878-1.063)
remainder	1.053 (1.004-1.104)

(* upper quintile)

CHAPTER 7. SENSITIVITY ANALYSIS

This chapter addresses the question of how sensitive the observed associations between air pollution and health effects are to several key analytic assumptions and features of the dataset. The objective was to assess if choosing alternative modelling strategies and approaches to considering confounding or modifying variables, such as weather, would have produced different estimates of the effects of air pollution. In addition, we evaluated to what extent some procedures employed had influenced the results obtained.

To accomplish this goal, additional analyses were performed for the effects of PM_{10} on all cause mortality in the elderly. Further analyses were conducted for other air pollutants and outcomes if substantial changes in the effect estimates for PM_{10} were observed. Four main aspects of the analysis were considered in order to evaluate the sensitivity of the results to model specification:

- model *y*-error distributional assumptions / estimation techniques
- model specification for meteorological variables, specifically temperature
- role of other explanatory variables, outliers, or potential residual confounding of temporal and seasonal related factors
- role of spatial variability in levels of air pollution

7.1 RESPONSE DISTRIBUTIONAL ASSUMPTIONS

The final standard model used in the time series analysis of mortality and hospital admissions was a Poisson regression allowing for general overdispersion and when necessary, for autocorrelated errors. For most outcomes examined in this study, the filtering procedures for long term trends, seasonality and meteorology were sufficient to account for the autocorrelation present in the un-filtered series as judge by the PACF. For all cause mortality in the elderly, however, there was still some remaining autocorrelation in the data. The size of the error introduced in the parameter estimates by this autocorrelation can be evaluated by comparing the results of the standard model and of a Poisson regression without adjustment for autocorrelation (Table 7.1). If the autocorrelation was not accounted for, the error introduced would be a decrease of about 6% in the coefficient estimates of air pollution. Practically no change in the standard errors was observed by the exclusion of the autocorrelated terms.

On the other hand, the correction for general overdispersion produces changes in the standard errors (SE) but not in the parameter estimates. As observed in the example in Table 7.1, the inclusion of an overdispersion factor increased the SE of the coefficients for PM_{10} by 9% while keeping the same value for the coefficient. The effect of the adjustment for general overdispersion is that of increasing the standard errors (SE) of the estimates. Therefore, failure to allow for such overdispersion will lead to underestimation of the variance of the coefficient and exaggerated significance levels.

Table 7.1 - Parameter estimates for PM_{10} under alternative model assumptions

	Coefficient*	SE*	*p-value*
Poisson with overdispersion and autocorrelation	0.499	0.209	0.017
Poisson (only overdispersion)	0.469	0.210	0.025
Poisson (no autocorrelation, no overdispersion)	0.469	0.191	0.014
Ordinary Least Squares regression	0.474	0.220	0.032

* $x10^3$

Although the number of daily deaths from all causes in the elderly in this study was large enough to make Gaussian approximation reasonable, mortality data is in the form of

counts, so that the variance of the data is not constant. Therefore, Poisson regression models were used. However, in order to test the sensitivity of the findings to the use of Poisson regression modeling the model was refitted using a different error distributional assumption.

The standard model was re-estimated using Ordinary Least Square (OLS) regression with a log-transformed outcome. The rationale for the log-transformation is for comparability with the log-linear model used in Poisson Regression. Inhalable particles (PM_{10}) remained significantly associated with total daily mortality in the elderly in the OLS model with a slightly smaller coefficient.

In conclusion, in spite of the differences in coefficient estimates from the models described in Table 7.1, they all provided essentially similar Relative Risks for an increase of 100 $\mu g/m^3$ in levels of PM_{10} (RR=1.05). The Poisson model was shown to be little influenced by the degree of autocorrelation present in the data and by the adjustment for general overdispersion. The former produces slightly underestimated parameters and the later provides more conservative tests without changing the parameter estimate itself.

The use of Ordinary Least Squares regression which assumes a Gaussian distribution sufficed in this example. However, the mean number of daily deaths was relatively large (mean of 67 deaths per day) which makes the assumption of a Normal distribution more acceptable. Nevertheless, the use of Poisson models is mandatory for less common outcomes when the daily number of deaths is low.

7.2 MODELING TEMPERATURE

A major concern in the analysis of daily variations in mortality or hospital admissions and pollution is the adequacy of the weather specification. Temperature seems to play a major role among the weather variables since it has been identified as having an important contribution to mortality [Saez et al,1995; Kalkstein,1993; Kalkstein and Smoyer,1993] and the lack of adequate control for its effects has been blamed for leading to spurious associations of air pollution with health effects [Mackenbach et al,1993].

In the present analysis a substantial effort has been made to adjust for different forms and functions of temperature. Three different measures were available, the daily mean, maximum and minimum temperature. A fourth variable measuring the daily variability (daily difference between maximum and minimum) was created. Different lag periods and formats for the temperature-mortality/hospital admission relationship were explored. The final selection of temperature terms was based on improvements in the model's fit, as assessed by F-tests.

The great majority of studies on the association between air pollution and health outcomes have adjusted for temperature effects using only measures of mean temperature (and sometimes dew-point temperature). In this study, however, maximum temperature was shown to provide better improvements in models' fit than mean temperature. This was true for models having outcomes such as all cause and cardiovascular mortality in the elderly, respiratory, pneumonia, and asthma hospital admissions in children under 5. In order to check if the use of mean temperature in these models would have had an effect on the estimated coefficient of air pollution, the terms for maximum temperature were replaced by mean temperature (maintaining the same lag structure and functional form) in a model for all cause mortality in the elderly. The results of such an approach are displayed in Table 7.2.

Table 7.2 - Parameter estimates for air pollutants under an alternative approach to adjust for the temperature effect in a model for all cause mortality in the elderly. Relative Risks (95% CI) were for a 100μg/m^3 change in levels of pollution.

	Temperature variables	Coefficient**	SE**	RR	95% CI
PM_{10}	Standard model*	0.499	0.209	1.05	(1.01 - 1.09)
	model with mean temperature	0.661	0.208	1.07	(1.03 - 1.11)
SO_2	Standard model*	1.611	0.612	1.17	(1.04 - 1.32)
	model with mean temperature	1.853	0.616	1.20	(1.07 - 1.36)
O_3	Standard model*	0.213	0.108	1.02	(1.00 - 1.03)
	model with mean temperature	0.200	0.109	1.02	(0.999 - 1.05)

* linear plus quadratic terms for maximum temperature

** x10^3

The sensitivity of the parameter estimates was assessed only for PM_{10}, SO_2 and O_3 since these were the pollutants that provided statistically significant associations with all cause mortality in the elderly. If mean instead of maximum temperature was used in this model an increased coefficient estimate for the effects of PM_{10} and SO_2 would be obtained. This overestimated coefficient is 32% higher for PM_{10} although the standard error changed minimally. Therefore, this different model provided an increased Relative Risk which continued to be significant. A similar effect was observed for SO_2 although the increase in coefficient estimates was smaller. Replacing maximum by mean temperature would produce a coefficient 15% higher with also slightly higher standard error.

Ozone behaved in an opposite manner. Although very small (only 6%) there was a decrease in the parameter estimate for this pollutant when mean temperature was used in the model instead of maximum temperature. However, as the change in coefficient estimate is small, there was no modification in the estimates of the Relative Risk for a $100\mu g/m^3$ increase in levels of this pollutant.

The sensitivity of the parameter estimates of air pollution to different temperature control measures was also noted for other causes of mortality explored in this study. For cardiovascular deaths in the elderly, models were originally fitted with a 3-piece linear function for maximum temperature. To check if the use of mean temperature in these models would have provided different estimates of the air pollutant effects, maximum temperature was replaced by daily mean temperature as a 3-piece linear function. Once more, the coefficient for PM_{10} was 44% higher and for SO_2 it was 23% higher when daily mean temperature was used.

A question could be raised now as to whether the models that selected mean temperature as the best control variable during the model building steps would have substantial changes on their parameter estimates for air pollution if maximum temperature were used instead. For example, for respiratory deaths in the elderly the best temperature term selected in the model building steps was a linear plus a quadratic term for mean temperature. In this case, mean temperature was replaced by the daily maximum in the form of a linear and quadratic term. It was observed that the reduction in the coefficient for PM10 was very small (5%) with practically no change in the estimate of the Relative Risk. For SO_2 the reduction was slightly more pronounced (14%), however, it should be

noted that the effect of SO_2 in the original model for respiratory mortality was not statistically significant at the 5% level. This alternative approach to modeling temperature also gave a non significant estimate for this air pollutant.

Another sensitivity check explored was to allow more than one temperature term into the model if they still presented statistically significant contributions to the model fit. In other words, after including into the model the temperature term that provided the best improvement in fit, the remaining terms were re-examined and if one or more terms still showed significant contributions to the fit, they were allowed into the final model. In the case of all cause mortality in the elderly, besides the linear and quadratic term for maximum temperature, a two piece linear function for minimum temperature was also allowed into the model. The results of this modeling strategy are summarized in Table 7.3. The use of a more comprehensive temperature model leads to a decrease in the coefficient for PM_{10} of only 5% with no change in the standard error. Therefore, no change in the Relative Risk was observed.

Table 7.3 - Parameter estimates for PM_{10} and SO_2 under alternative approaches to adjust the temperature effect. Relative Risks (95% CI) were for a 100μg/m^3 change in levels of pollution.

	Temperature variables	Coefficient**	SE**	RR	95% CI
PM_{10}	Standard model*	0.499	0.209	1.05	(1.01 - 1.09)
	standard model plus 2-piece linear function for minimum temperature	0.476	0.209	1.05	(1.01 - 1.09)
SO_2	Standard model*	1.611	0.612	1.17	(1.04 - 1.32)
	standard model plus 2-piece linear function for minimum temperature	1.492	0.609	1.16	(1.03 - 1.31)

* linear plus quadratic of maximum temperature
** x10^3

This low sensitivity of the coefficient estimates for PM_{10} was also observed for SO_2 in a model for all cause mortality in the elderly. A decrease of 7% was observed in the parameter estimates when employing a more conservative approach to control for temperature. This produced little variation on the estimate of the Relative Risk for a 100μg/m^3 change in levels of SO_2.

Similarly, for cardiovascular and respiratory mortality in the elderly the use of a more comprehensive approach to modelling temperature produced little change in the coefficients for either PM_{10} or SO_2. No substantial changes in the estimates for air pollutants were observed when allowing more than one temperature term into those models.

The last examination of the sensitivity of these models to the way temperature terms were specified was to allow different lag structures for all four measures of temperature, in different formats. Thus, linear, quadratic, 2-piece linear and 3-piece linear terms were constructed for mean, maximum, minimum and difference in temperature at lags zero, one and two days. All these functions were tested in a stepwise selection in a model for all nonaccidental mortality in the elderly.

The final model achieved using this modeling approach allowed various temperature terms: a linear term for maximum temperature at lag 0, a 2-piece linear term for minimum temperature at lag 0, and 3-piece linear function for mean temperature at lag 1 and at lag 2 and for minimum at lag 1. The coefficient ($x10^3$) for PM_{10} in this model was 0.411 (SE=0.233) which provides a Relative Risk for a $100\mu g/m^3$ change in daily PM_{10} of 1.04 (95% C.I. 0.995 - 1.091). Note that there is greater decrease in the coefficient estimate for PM_{10} (18%) in this model compared to other modeling alternatives described in Table 7.3. Nevertheless, the Relative Risk presented only a small reduction and was then only marginally significant.

In conclusion, little change in the coefficient estimates for air pollutants were observed when using a more conservative approach to model temperature, i.e., allowing as many temperature terms as still contributing significantly ($p<0.10$) to the model's fit. In addition, the strategy of reviewing the effects and choosing between temperature terms including variables for mean, maximum, minimum and difference in temperature was appropriate; in most models maximum temperature was shown to be a more important predictor than mean temperature. Most published studies have only used mean daily temperature in their analysis. It was shown that if in those models mean was used instead, overestimated parameters would be obtained for air pollutants, but mainly for PM_{10} and SO_2. On the other hand, for the outcomes that selected mean temperature as the best

variable to adjust for the temperature effect, little change occurred when maximum temperature was used instead.

7.3 OTHER EXPLANATORY VARIABLES

In order to get a better understanding of the impact of each set of adjusting variables on the air pollution estimates, a different analysis was performed. This consisted of fitting PM_{10} first in a model with all cause mortality and gradually adding other covariates. This exploratory procedure was helpful in examining the changes in PM_{10} estimates in relation to each set of adjusting variables added at different modeling steps and also to get a general idea of the contribution of each of them in achieving a satisfactorily adjusted model.

With no correction for trends, seasonality or meteorology PM_{10} has a quite large coefficient which is slightly increased when variables allowing for long-term trends are included into the model (Table 7.4). A substantial reduction occurs when a sine and cosine wave for one year period are introduced. Including the remaining sine and cosine waves had little further effect on the parameter estimate. This demonstrated that not correcting for seasonal patterns, especially for one year cycles would have led to serious overestimation of the pollution effect.

Table 7.4 - Regression results for all cause mortality in the elderly in relation to each set of adjusting variables

Model (each factor progressively added)	PM_{10}		
	Coefficient*	SE*	% change
no confounders	1.635	0.166	-
trend	1.741	0.164	+6.5
1 year cycle	0.943	0.161	-42.3
full season model	0.999	0.161	-38.9
days of week	0.984	0.164	-39.8
meteorology (full model)	0.499	0.209	-69.5

* $x10^3$

Day of week had little contribution in this analysis of mortality. Including meteorology beyond what was already explained by the seasonal model led to some further correction and another dramatic decrease in the coefficient for PM_{10}. Again, this denotes the importance of meteorological variables and the risk of overestimating the effects of air pollutants when their confounding effect, especially that of temperature, is not properly accounted for.

It is then clear that, besides temperature, temporal patterns have an important role in time series analyses of air pollution effects. Therefore, another sensitivity assessment was for the role of the model specification for temporal confounders. Alternative methods for filtering the long wavelength fluctuations from the data were examined. First, indicator variables for season and year were introduced into the basic model replacing the harmonics. As seen in Table 7.5 the coefficient for PM_{10} is largely overestimated denoting a possibly unsatisfactory adjustment for seasonality and a clear lack of adjustment for the one year cycle. Next, indicator variables for each month of each year (35 variables) were used instead to control for seasonality. The estimate for the effect of PM_{10} is now smaller but still 30% higher than from the standard model. In this case, the one year seasonal cycle and the long-term trend present in the data are, at least, partially accounted for.

Table 7.5 - Regression results for all cause mortality in the elderly for alternative ways of modeling temporal patterns

	Coefficient*	SE*
standard model	0.499	0.209
indicator variables for season and year	1.038	0.208
indicator variables for each month	0.645	0.209

Note: all models were fully adjusted for long term trends, day of week and meteorology
* $x10^3$

Then, an attempt was made to examine whether the decision to control for seasonality filtering cycles up to two months was adequate. This was achieved by first, allowing harmonic waves of up to one month period into the model and, in another step, by excluding the harmonics corresponding to two months (Table 7.6). Sine and cosine waves

down to one month were then introduced into the previous basic model. The coefficient was slightly affected. Then, the harmonic waves for two months were excluded from the basic model. Again the changes were extremely small but in an opposite direction with slight increase in the coefficient for PM_{10}. Nevertheless, in all these situations, the estimate for the Relative Risk for a change of 100μg/m^3 in levels of PM_{10} was the same (RR=1.05).

Table 7.6 - Regression results for all cause mortality in the elderly for alternative ways of modelling temporal patterns

	Coefficient*	SE*
standard model	0.499	0.209
model allowing harmonic waves down to 1 month	0.484	0.209
model excluding harmonic waves for two months	0.520	0.208

Note: all models were fully adjusted for long term trends, day of week and meteorology
* x10^3

The final comparison involved introducing into the basic model for all cause mortality in the elderly all temporal and seasonal variables which were allowed in at least one of the individual models explored in this study. Therefore, apart from linear and quadratic time terms, indicators for day-of-week, and variables allowing for meteorological factors, 11 different pairs of sine and cosine waves, plus indicator variables for each year of the study, were all allowed into the same model. The effect of this possible overadjustment was a 9% increase in the coefficient for PM_{10} (coefficient = 0.543 and SE = 0.212).

Finally, the possible sensitivity of the estimates for air pollution due to influential observations or outliers was tested. The Poisson model was repeated excluding the 18 days (1.6%) that exceed the Brazilian National Ambient Air Quality Standard of 150 μg/m^3. When this analysis was performed there was a decrease of 21% in the magnitude of the effect of PM_{10} and an increase in the Standard Error (SE) of the pollution estimate which were no longer statistically significant (p=0.09). The increase in SE was probably due to having less observations as it would be expected the variability to decrease in such situation.

In summary, the importance of proper adjustment for long-term trends and seasonal patterns was demonstrated for a time series analysis of air pollution and health effects. It was clear that the one year cyclical pattern present in this data is highly important, if not the most important periodicity that has to be accounted for. In addition, it seems that the use of harmonic waves had more satisfactorily corrected for seasonal patterns than would have been achieved using different approaches. In fact, these analyses showed that the model is little sensitive to the exact degree of filtering of the seasonal patterns, i.e., filtering harmonics corresponding to a little more or a little less than two months, or including several different periodicities regardless of the statistical significance of their contribution would produce nearly no change to the coefficient estimates. Nevertheless, it seems that the results are more sensitive to some potentially influential high pollution days, since the exclusion of such days produced a reduction of the estimates of air pollution.

7.4 SPATIAL VARIABILITY IN LEVELS OF AIR POLLUTION

Two different issues were explored in relation to the spatial variability in levels of air pollution. First, it was investigated if excluding from the analysis individuals who live in more peripheral areas, those more distant from a monitoring station of air quality, would have provided different, and possibly higher estimates for the effects of air pollutants. The other issue considered was whether averaged levels of air pollution using all the monitoring stations distributed around the city were providing a reasonable measurement of the exposure of the population.

The automatic network of monitoring stations of air quality in the city of São Paulo comprises 13 stations not evenly distributed around the city. They are mostly concentrated in more central areas (Fig 3.2, Chapter 3). Thus, it could be that an average of these stations would be a poor measure of exposure for individuals living in more peripheral areas, far from such stations. This was the reason why individuals residing in the more distant district of Parelheiros had been excluded from all the analysis of daily mortality.

This sensitivity analysis examined if there would be any extra gains in coefficient estimates when excluding other peripheral areas, which were not as distant as Parelheiros but were also outside the area covered by monitoring stations. The rationale for this is that individuals living in those areas, potentially less polluted, would not be experiencing the same levels of air pollution measured in more central areas and so would not exhibit exposure-related effects. Therefore, excluding those individuals might have resulted in higher estimates for the effects of air pollution.

Figure 7.1. shows a map of the city of São Paulo, the location of each monitoring station of air quality, and the districts that were excluded for this analysis. Note that, apart from Parelheiros, another 13 more peripheral districts were excluded. The deaths of residents of all the remaining districts were aggregated into daily counts spanning the 3 years available for the mortality data. For this analysis, deaths due to all causes except external causes were used. A complete set of model building procedures was performed in order to obtain a model with satisfactory adjustments for seasonality, other temporal patterns, and meteorological variables.

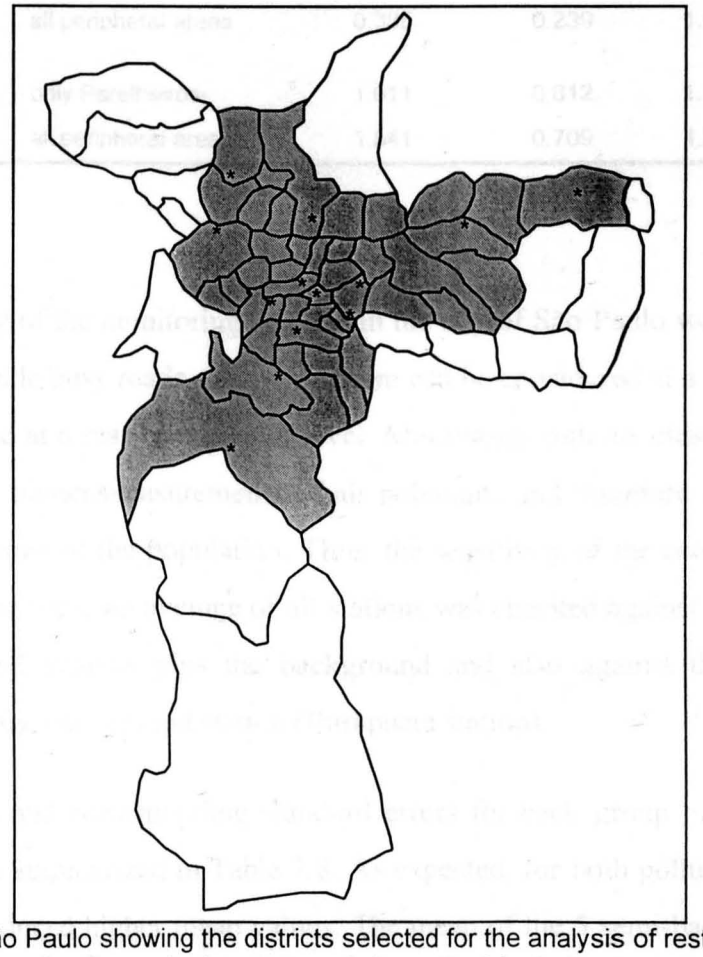


Fig 7.1 - Map of São Paulo showing the districts selected for the analysis of restricted areas and the location of each monitoring station of air quality (black stars).

The final model was then analyzed in a Poisson Regression framework with adjustments for overdispersion. The counts of deaths for all cause mortality in this sub-area of the city of São Paulo did not exhibit any autocorrelation after the filtering procedures.

The results of this analysis are displayed in Table 7.7. As it can be seen, the exclusion of peripheral areas of the city produced a small decrease in the coefficient estimate for PM_{10} and an increase in its standard error. Nevertheless, the reduction in the estimate of the Relative Risk was quite small. On the other hand, for SO_2 there was an increase in its coefficient estimate as well in its SE. Although changes were small in both cases, they happened in opposite directions.

Table 7.7 - Regression results for PM_{10} and SO_2 for all cause mortality in the elderly for the standard model and for a model excluding other peripheral areas of the city of São Paulo. Relative Risks (RR) were calculated for a change of 100μg/m^3 in levels of PM_{10} or SO_2.

	Areas excluded	Coefficient*	SE*	RR	*p-value*
PM_{10}	only Parelheiros	0.499	0.209	1.05	0.017
	all peripheral areas	0.385	0.239	1.04	0.11
SO_2	only Parelheiros	1.611	0.612	1.17	0.009
	all peripheral areas	1.841	0.709	1.20	0.009

* x10^3

The great majority of the monitoring stations in the city of São Paulo were situated at the street level alongside busy roads. Only 4 of them can be considered at a semi-background level and only one at a real background level. Monitoring stations closer to busy streets are likely to have higher measurements of air pollutants and therefore poorly reflect the true average exposure of the population. Thus, the sensitivity of the coefficient estimates of air pollutants for using an average of all stations was checked against an average of the 4 semi-background stations plus the background and also against the measurements provided only by the background station (Ibirapuera station).

The mean values and corresponding standard errors for each group of stations and for PM_{10} and SO_2 are summarized in Table 7.8. As expected, for both pollutants, the average of all stations produced higher mean values. The mean of the 5 semi-background stations

comes intermediate and the mean for the background station comes very close in third place.

Table 7.8 - Mean values and standard errors (SE) for PM_{10} and SO_2 for different

	PM_{10}		SO_2	
all stations	64.3	(27.6)	18.8	(8.5)
5 semi-background	56.4	(28.4)	12.0	(6.1)
1 background	49.7	(27.5)	11.1	(9.0)

Exposure values measured at only one station could not be assumed to represent the exposure of the whole city. Therefore, it was decided to restrict the area to which exposure values would be assigned based on the background station. All districts within an area of 5 Km radius from the background station 'Ibirapuera' were selected (Fig 7.2).

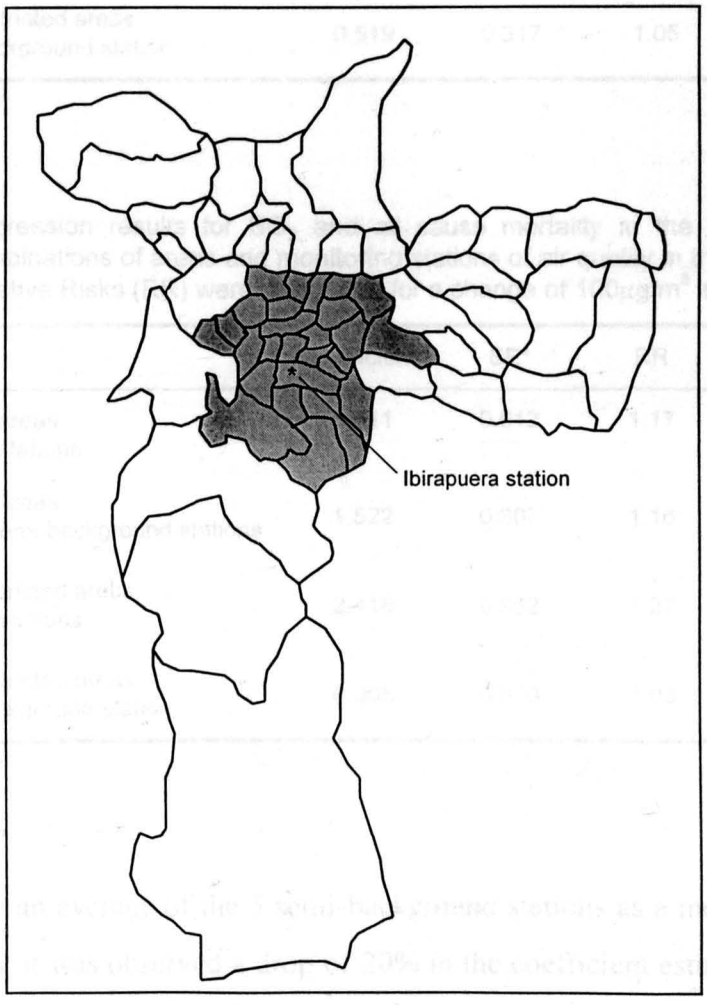


Fig 7.2 - Map of São Paulo showing the districts selected for the analysis restricted to the background station (Ibirapuera station).

These 27 districts represent the most central areas of the city of São Paulo. Once more, daily counts of deaths for all nonaccidental causes were selected for this sub-area and model building procedures were performed in order to find out a core model for this group of deaths. Results of this analysis are displayed in Table 7.9 and 7.10.

Table 7.9 - Regression results for PM_{10} and all cause mortality in the elderly for different combinations of areas and monitoring stations of air quality in the city of São Paulo. Relative Risks (RR) were calculated for a change of 100μg/m^3 in levels of PM_{10}.

	Coefficient*	SE*	RR	*p-value*
all areas all stations	0.499	0.209	1.05	0.017
all areas 5 semi-background stations	0.397	0.199	1.04	0.046
restricted areas all stations	0.581	0.298	1.06	0.05
restricted areas background station	0.519	0.317	1.05	0.10

* x10^3

Table 7.10 - Regression results for SO_2 and all cause mortality in the elderly for different combinations of areas and monitoring stations of air quality in the city of São Paulo. Relative Risks (RR) were calculated for a change of 100μg/m^3 in levels of SO_2.

	Coefficient*	SE*	RR	*p-value*
all areas all stations	1.611	0.612	1.17	0.009
all areas 5 semi-background stations	1.522	0.807	1.16	0.06
restricted areas all stations	2.418	0.952	1.27	0.01
restricted areas background station	0.305	0.870	1.03	0.73

* x10^3

When using only an average of the 5 semi-background stations as a measure of exposure of the population, it was observed a drop of 20% in the coefficient estimate for PM_{10} and a smaller decrease for the standard error of the estimate. Nevertheless, the change in the

Relative Risk was again quite small. A similar pattern was observed for SO_2 when using the average of the 5 semi-background stations. Coefficients changed from clear significance to marginal significance.

Restricting the model to the smaller area (line 3 of Tables 7.9 and 7.10) led to an increase in the SE and in the case of SO_2, to a somewhat larger estimate. Using only one background station to represent the exposure of the individuals, there was little change in coefficient estimate for PM_{10} whereas for SO_2 it was sharply reduced.

Both, PM_{10} and SO_2 are strongly traffic-related pollutants, especially SO_2 for which emissions from motor vehicles (diesel engines) correspond to 64% of total emissions (see Table 2.2). However, the dispersion of these pollutants in the atmosphere is likely to be different. The correlation coefficients between the background station against all the other stations (Table 7.11) show that PM_{10} is highly correlated with most of the stations and particularly higher with the semi-background stations while SO_2 exhibited much lower correlations. This would suggest that these exposures are differently distributed around the city. Levels of PM_{10} seem to be more homogeneous over the city whereas for SO_2, measurements at the street level are not much correlated with those measured at the background level.

Table 7.11 - Correlation coefficients between the background monitoring station of air quality and all the other stations for PM_{10} and SO_2, in São Paulo, Brazil, 1991-1993.

	Street level stations								semi-background			
nº	1	2	3	4	8	9	10	12	6	11	16	21
PM_{10}	0.6	0.8	0.8	0.4	0.7	0.7	0.7	-	0.8	0.8	0.8	0.8
SO_2	0.3	0.4	0.3	0.5	0.5	0.4	0.4	0.3	0.4	0.4	0.3	0.2

As individuals usually spend their time at different environments (which would possibly include the street level for, e.g., commuting to and from work/school), their exposure to SO_2 would be poorly measured by only one station or by stations measuring the same type of environment. Variations in the exposure of these individuals seems not to be

adequately measured by the background station which fail to take into account the variations at the street level.

In summary, it was seen that when excluding areas in the periphery of the city that might not be experiencing the same degree of air pollution measured in more central areas, there was an increase in the standard errors with no change in the estimates for the effects of air pollution. Therefore, it can be assumed that the averaged levels of air pollutants were giving a good measure of the exposure even in those areas far from monitoring, especially for PM_{10}. Similar results were obtained again for PM_{10} when restricting the measure of exposure to only 5 stations, not located at the busy roads level or when using a background station. For SO_2, on the other hand, the use of the background station provided very low coefficients for the effect of this pollutant. This probably indicates a less homogeneous levels of this pollutant across the city.

CHAPTER 8. DISCUSSION

The results found in this study concerning the effects of air pollution on daily mortality and hospital admissions need to be considered in the context of the whole body of evidence provided by different studies carried out all over the world. Before that, however, issues related to the design, conduct, analysis and interpretation of this epidemiological study require careful examination.

The use of time series methodology in epidemiological studies is still a recent approach and it has been developed to a large extent in the last few years. Therefore, issues like overdispersion, serial correlation, time-invariant confounders, etc., are now part of the epidemiological glossary. These issues will be adequately addressed in the section below on Study Design.

In addition, a cautious inspection of an epidemiological study would not be complete if the roles of confounding and bias, both in the selection of study subjects and in the information collected for the study, were not fully addressed. Thus, the sections on

Selection Bias, Information Bias, and Confounding were written to appropriately examine such topics. Finally, this chapter also presents an in-depth discussion of the problem of Exposure Assessment, especially relevant to environmental epidemiology and particularly to air pollution studies.

8.1. STUDY DESIGN AND ANALYTICAL ISSUES

Distinct analytical approaches can be used when examining the relationship between exposure to air pollution and health effects. The purpose of the present study was to investigate if short-term fluctuations in air pollution were associated with effects on health. In other words, to assess if they exhibit an acute (short-term) temporal association. In order to ascertain such effects it is more convenient to consider daily observations, for if there are short-term relationships, say an effect of air pollution on mortality lagged by zero, one or two days, these can only be detected through daily analysis. Therefore, time series analysis was applied to assess the effects of air pollution on health in the city of São Paulo.

Time series studies are a particular type of ecological study in which the unit of aggregation is a time period instead of a geographical area. As ecological studies, they have advantages over other epidemiological designs. For example, they can often be carried out by combining existing data files on large populations. Therefore, they are generally less expensive and take less time than studies involving the individual as the unit of analysis [Morgenstern,1982].

However, these advantages are counter-balanced by some constraints characteristic of ecological studies. The main problem limiting the use of such designs in epidemiology is their great potential for bias, the so called 'ecological fallacy'. Since exposure and disease are measured at the aggregated level, it is impossible to link exposure with disease in particular individuals. For example, one can not be sure if those persons who died or were admitted to hospital due to respiratory infection were the same ones who experienced the air pollution exposure. Another major limitation of ecological studies is the lack of ability to control for the effects of potential confounding factors which can not be measured at

the individual level. In addition, data on these studies represent average exposure levels rather than actual individual values [Hennekens and Buring,1987].

Nevertheless, some of these problems are less important in time series studies of air pollution in comparison to geographical studies. To avoid the ecological fallacy it would be necessary to obtain individual level data on both exposure and health outcome. While measuring health outcomes can be feasibly achieved in epidemiological studies, measuring individual exposure to air pollutants is either too expensive or completely impossible due to logistical reasons. However, the ubiquitous characteristic of the exposure to air pollution makes it easy to assume that a change in mean exposure is well estimated by a change in measured background levels. This approach certainly introduces a degree of error in the exposure estimates and does not also take into account the movement of people across different geographical areas, amount of time spent outdoors, etc. The consequences of this error on exposure assessment will be discussed later.

A distinct issue of daily time series analysis is that counts of deaths or hospital admissions are counts of rare events and can be thought of as following a Poisson distribution. Thus, in the present study all models were implemented in the Poisson Regression framework. However, as it could be observed in the sensitivity analysis, the choice of specific statistical models did not show any major effect on the results obtained. In fact, Thomas [1994] had pointed out that if the model is correctly specified with regard to covariates, neither the choice of statistical model nor the method for fitting exposure-response relationships is likely to have any important influence on the results. However, when the daily average number of health events is relatively small, the use of Poisson models is then mandatory.

But again, the Poisson assumption is only valid if the model for its expectation is correctly specified with no omitted covariates. If the model is misspecified, one might expect that the conditional variance will be larger than its modelled expectation, a phenomenon known as overdispersion [Thomas,1994]. Most of the models in the present study had overdispersion parameters of about **1.2**. This overdispersion was corrected by employing the methods described in chapter 3. While this small degree of overdispersion agrees with the current literature on time series studies of air pollution [HEI,1997; Thomas,1994; Schwartz,1993], it also indicates that the data conformed well with the

Poisson distribution. Or in other words, that the right set of covariates were included in the model. Further reductions in the overdispersion parameter could be attained, as demonstrated by Samet and collaborators [HEI,1997], by the successive addition of more complex variables based on time, weather, and interactions but, the risk of over-specifying the model increases sharply.

In addition to overdispersion, the assumption of independence may also be violated due to serial correlation. Measurements connected in time are likely to be correlated and not independent. However, this autocorrelation is usually due to an extraneous factor such as weather, epidemics or when a few somewhat important temporal-related influential factors are omitted [World Bank,1995; Schwartz et al,1996]. These situations can produce serial correlation in the residuals of the model.

After the inclusion of all covariates, most of the models in the present study exhibited no remaining autocorrelation denoting a successful adjustment for temporal and weather related variables. The few outcomes that exhibited a degree of autocorrelation after adjustments exhibited very small coefficients (in the order of 0.1 or less). This remaining autocorrelation was taken into account in those models although if not, there would be very little impact on the standard errors.

Another problem with ecological analysis is that certain predictor variables tend to be more highly correlated with each other than they are at the individual level. Consequently, the increased correlations between these variables make it particularly difficult to isolate their effects on the disease or health event [Morgenstern,1982]. This also applies to time series studies of air pollution where this multicollinearity, especially between air pollutants, and of these with meteorological variables is one of the biggest problems [Thomas,1994].

In this study it was observed that air pollution and meteorological variables were correlated probably due to long term trends and other seasonal patterns present in the unfiltered series. However, there were still signs of collinearity between these covariates after applying all the filtering procedures. Perhaps, the main consequence of this multicollinearity was to inflate the standard errors of individual parameter estimates, which

in turn made their associated confidence intervals wider, and thus, results less statistically significant [World Bank,1995].

Perhaps the main drawback of multicollinearity is the difficulty to discern which air pollutant are principally responsible, since the collinearity between two or more pollutants make it difficult to isolate their individual effects. Models including highly correlated pollutants could lead to an overestimation of the effects of one pollutant and on average underestimating of the effects of the others [Schwartz and Marcus,1990]. In addition, it has been discussed that attempts to isolate the effect of a single pollutant employing statistical methods, should be regarded with caution. Such a model may not be realistic in face of the possible synergism between pollutants [Loomis et al,1996].

To obtain the main results presented in this study several models were examined. There were five different outcomes for mortality which were tested against five different air pollutants in three possible lag periods. Similar number of scenarios arose for the hospital admissions analyses. Although most of these models were not independent, when multiple comparisons were made, concerns arose about the possibility that chance alone could be responsible for a certain proportion of the associations being statistically significant.

The traditional approach to tackle such a problem has been to make the statistical tests more stringent, either by changing the criterion for a more stringent value (1% instead of 5%), or by actually inflating the calculated *p-values* by some factor that depends on the number of comparisons made [Rothman,1986]. However, it is argued that such approaches decrease type I-error (finding false positive associations) at the expense of increasing type II-error (finding false negative associations) which in turn has nothing to do with multiple comparison, since they apply equally well to a single comparison [Rothman,1986].

Therefore, some researchers maintain that no adjustments are needed for multiple comparisons provided that it is clearly reported how many comparisons have been made and that all non-statistically significant results are reported along with the significant ones [Rothman,1990; Poole,1991; Savitz and Olshan,1995]. In the case of the present study, this argument is strengthened by the fact that most of the statistically significant

associations found were hypothesised beforehand and were also validated by the published literature. However, no prior assumption was made for the lag period of those associations. Multiple lags were tested in order to find out possible delayed effects of air pollution and this approach should be considered in the interpretation of the results.

Nevertheless, it was decided not to pursue any formal adjustment of the *p-values* in the present study. Emphasis has been placed on the coherence of the results in terms of internal patterns by age or season for example, and externally in relationship with other results.

8.2. SELECTION BIAS

In any study using secondary data or routine vital statistics the quality of data being analysed is always a matter of concern. The routinely collected data used here had several advantages such as availability, being inexpensive, and containing information on a large number of subjects and covering long periods of time. However, they also have some drawbacks mostly from the fact that no control over the data collection procedure was possible. The two groups of study subjects observed in this study were: individuals over 65 years of age and children under 5 years old who died and children under 5 years old who were admitted to hospital during the respective study periods in the city of São Paulo.

One of the main concerns in observational studies is the situation where procedures used to select subjects may lead to different magnitude of disease risks or relative risks among those who participate and those who would be theoretically eligible for the study but do not participate. In other words, the situation is that the inclusion of subjects in the study depends in some way on the exposure of interest. This distortion in the design of a study is known in epidemiology as 'selection bias' [Rothman,1986; Hennekens and Buring,1987]. The role of bias in the selection procedures of daily deaths and hospital admissions will be discussed.

8.2.1. Mortality

Routine vital registration data on deaths in São Paulo is considered to be complete and of high quality. Mortality data was provided by PRO-AIM, the City Health Department Programme in charge of processing and analysing all deaths of city residents that occurred in the municipality of São Paulo since May 1990. Deaths of city residents that occurred elsewhere may not be included in the PRO-AIM's statistics. They are, however, included in the death registry of Fundação SEADE. Comparing the two sources of data it was possible to estimate the proportion of missing deaths of city residents in the PRO-AIM dataset. These deaths, which occurred outside the city of São Paulo, represented about 6% of all deaths of city residents [FSEADE,1993b; PROAIM,1992].

A third of these deaths were due to external conditions and perinatal deaths [PROAIM,1992]. External or violent deaths were mostly traffic accidents that occurred outside the city or deaths by drowning that occurred on the beaches nearby. Perinatal deaths occurring outside the city refers to pregnant women looking for medical care outside the city since they live in areas with few or no maternal and child health services. In any case, perinatal and violent deaths were excluded from any analyses performed in this study.

There is, however, an extra proportion of missing deaths in the PRO-AIM's registry. These are estimated to be around 1% [M. Drummond, personal communication]. Therefore, the total missing deaths are estimated to be around 5% (comprising the 1% of missing deaths in the city and 4% of non-violent deaths occurring outside São Paulo). This low proportion of missing data on mortality is not expected to have any major influence in this study. Moreover, it is not very likely that the distribution of this group of missing deaths presents a strong temporal variation, or even if it does, it is not expected that this variation will be correlated with exposure, i.e., the daily levels of air pollution. Thus, bias in the selection of deaths is not likely to have occurred in this study.

8.2.2. Hospital Admissions

Data on health services utilisation has been widely used as a proxy for the underlying morbidity of the population. In this study, to assess the possible effect of exposure to air

pollution on child morbidity, daily counts of hospital admissions for children under 5 years were examined. The data on hospital admissions were extracted from computerised files of the Health Services Information Database of the National Health System (NHS). These files contain information on admissions for all Brazilian hospitals which are part of the NHS (public, philanthropic, university, and private but contracted hospitals).

The main advantages of this database are its extended coverage and the accessibility of the data. In addition, information can be disaggregated temporally to a daily level and spatially to the hospital level. However, this information system was primarily designed for accounting purposes and not for epidemiological studies. Therefore, its relative utility may suffer from problems in the quality of the information and from the fact that this system targets a selected population.

Under-registration probably does not exist or might be negligible since all the hospitals need to get reimbursement for their admissions. The coverage is also relatively good as 66% of all the available hospital beds in general hospitals in the city of São Paulo are part of the system, i.e., they provide information on every admission to the Health Department [M.L.Lebrao, personal communication; FSEADE,1993]. However, the remaining one third of hospital beds refers to private hospitals most times covered by private health insurance. These are the admissions of the wealthier portions of the population and are, therefore, under-represented in the files used for this study.

Another issue that should always be considered when analysing and interpreting data on hospital admissions is the possibility that limitations in the supply of beds could also influence the timing of admission. That means the rates of admission to the hospital can be affected by the availability of beds which may depend on some independent event acting on a long or short-term basis [Bennett,1981]. This constraint that does not exist for mortality data can introduce some degree of selection bias in the study.

However, the under-representation of wealthier subjects can only pose limitations on inferences about socioeconomic differences in risks from air pollution exposure. Although selected, this study population acts as its own control in a time series study, minimising the possibility of bias. The possible constraints in the availability of hospital beds were

also not expected to be correlated with air pollution levels. Thus, neither situation poses a major threat to a time series study.

8.3. INFORMATION BIAS

In addition to the two sources of outcomes used in this study (mortality and hospital admissions), information on exposure and other covariates were also obtained from secondary sources. More specifically, these refer to data on air pollution, socio-demographic and meteorological variables. As mentioned before, routinely collected data might present problems related to the quality of the information and they need, therefore, careful examination.

If information on each death or hospital admission, in the daily levels of air pollutants or meteorological factors, or in socio-demographic characteristics has inaccuracies, this will result in systematic errors in the classification of subjects to either disease, exposure, or covariate status. This information bias is one of the most important limitations of observational studies. Since some degree of inaccuracy in classifying individuals is inevitable, misclassification is always a potential concern [Hennekens and Buring,1987].

The consequences of this bias are different depending on whether the classification error on one axis (say, disease) is independent of classification on the other axis (exposure or other covariates). When misclassification is random or non-differential, for example, when the exposure classification is incorrect for the same proportion of subjects in the groups compared, it always results in an underestimate of the true relative risk since the similarity between the two groups increases. Conversely, if the proportion of subjects misclassified differs between the study groups, the effect of such systematic or differential misclassification is that the observed estimate of effect can be biased in the direction of producing either an overestimate or underestimate of the association [Rothman,1986].

The assessment of the quality of the information provided by each dataset will be outlined in the following sections. This will primarily be based on the completeness and accuracy of the information. A brief review of some of the methodological issues concerning these datasets will also be discussed. The issue of misclassification of exposure, crucial to

studies in environmental epidemiology will have a more in-depth examination in a separate section.

8.3.1. Mortality

The coverage of the mortality register operated by PRO-AIM was shown to be very high. Therefore, a quantitative constraint does not pose a problem for the mortality data used in this study. Nevertheless, some potential problems need to be borne in mind when interpreting results of a mortality analysis.

The principal limitation for the use of any mortality statistics is the quality of the information on the basic cause of death in the death certificate. There have not been any recent studies examining the quality of this information. However, the diagnosis of the great majority of deaths occurring in the city of São Paulo are known to be assigned by medically qualified personnel [FSEADE,1990].

The proportion of ill-defined conditions as the basic cause of death can also be viewed as a measure of the quality of the death certification. Between 1990 and 1995 the proportion of ill-defined conditions among the causes of death in the city of São Paulo has been between 1.1 and 1.3%. In addition, PRO-AIM carries out a checking procedure for all unclear diagnosis contacting the hospital where the death occurred or the doctor who provided the information [M. Drummond, personal communication].

One concern pertaining to the reliability of studies using data from mortality registers is that there might be differences in interpretation among those who assign the diagnosis or those who code the cause of death. Consequently, the same disease can sometimes be classified with different codes. However, most of the analysis performed here were not conducted on individual codes but on groups of ICD codes, which minimises such problems. In addition, for this study there were no changes in ICD and ICD-9 was used throughout the study period. The new ICD-10 started being used in Brazil only from January 1996 [M.Drumond, personal communication].

There are also no studies about the quality of other information recorded on the death certificate. However, the proportion of missing data for variables like sex (0.1%), age (0.5%) and place of residence (4.7%) indicates the reasonable quality of this information.

8.3.2. Hospital Admissions

Although demands of hospital admissions can be a useful end point for quantifying effects on community health, the information on admissions used in this study may have problems with respect to the accuracy and completeness of the information provided by each hospital. Unfortunately, there are no studies examining the quality of these data.

The main concern is the quality of the information related to the diagnosis. Since the information system was designed with the aim of administrative control and payment for health services, it is possible that some hospitals bias the information on diagnosis, favouring diseases that require more expensive treatments. Audits conducted by the Local Health Authority are frequent but the problem can still happen. A related issue refers to diseases that have many different clinical forms, so they can be reported with different diagnoses. Although important, the latter is less of an issue for the present study since the grouping of diseases used here are not much influenced by this type of error.

Nevertheless, both questions can pose crucial problems for epidemiological studies examining patterns of hospital admissions. However, one has to bear in mind the unique characteristics of time series studies. Factors whose distribution does not vary from day to day are unable to act as confounders [Katsouyanni et al,1997a]. It is not expected that the misclassification of diagnosis would have changed over time. Therefore, in spite of their importance, these problems were minimised in the present study.

On the other hand, the proportion of missing data for variables like sex, age, date of birth, date of admission or discharge are all very low for this dataset. This is an indication of the completeness of this dataset. Thus, bearing in mind the flaws in the hospital admissions statistics in Brazil, this dataset can be widely used for epidemiological purposes if appropriate caution is taken in interpreting the data.

8.3.3. Air Pollution

Levels of air pollutants in São Paulo have been systematically measured by CETESB, the local environmental agency, since the late 1960's. The attainments of the monitoring network and the data handling capabilities of CETESB have been considered of good quality by the United Nations Environmental Programme [UNEP,1992]. Nevertheless, there are issues surrounding the measurements and some specific procedures employed in the monitoring of air pollution that deserve further consideration.

CETESB pays careful attention to data collection to ensure that the results obtained are of the best quality. Measurements of each pollutant and for each monitoring station have to comply with a defined protocol with criteria for completeness. Consequently there is great variation among monitoring stations in continuity of data collection with some stations presenting large amounts of missing data. In addition, only measurements from the stations fulfilling the completeness criteria on a given day were used in computing the daily exposure.

Despite this, it was possible to obtain daily city-wide averages for all the days for PM_{10} and SO_2 for the mortality analysis and with a very low number of missing days for the hospital admissions analysis and for the other pollutants in both analyses. However, the calculation of these city-wide daily averages was based on different numbers of stations. For PM_{10} and SO_2 at least 4 monitoring stations could be used to calculate the daily average for the city. For the other pollutants, sometimes only one station was available and was, therefore, used as the mean value for the city. This procedure, although unavoidable, could have introduced errors in the estimation of daily exposure to pollutants.

Additionally, when daily means of PM_{10} in a monitoring station were below the detection limit of the equipment, the value of that day in that station was set to missing. This is a standard procedure applied by CETESB and it could also have given rise to errors in the calculation of citywide values of PM_{10}. However, in both cases, these errors are unlikely to be differential. They are both sources of misclassification of exposure to air pollutants, but of non-differential misclassification. Misclassification due to error of measurement of

the actual exposure is a common problem in studying air pollution epidemiology and probably leads to an underestimation of the true association [Hatch and Thomas,1993].

Although there might be problems in the daily city-wide averages of air pollutants, the actual accuracy of the measurements for each monitoring station is of a high degree. As examined in the validation measurements for PM_{10} conducted by an independent equipment, there was excellent agreement between both measurements.

It has also to be pointed out that during the period covered by the present study there were no changes in the methodology used to measure air pollution as well as no modification in the configuration of the monitoring network.

Another issue that has to be considered is that the distribution of the monitoring network of air quality was not designed with the assessment of the exposure of the population in mind. The great majority of these stations are situated at the street level alongside busy roads and only four are situated at semi-background level and only one at a real background level. In addition, they are mostly concentrated in more central areas of the city (Fig 3.2, Chapter 3). Therefore, one might say that they best describe each monitoring station's immediate surroundings and would be a poor measure of the exposure of the majority of the population.

However, it was seen that when excluding areas in the periphery of the city that might not be experiencing the same degree of air pollution measured in more central areas, there was an increase in the standard errors with little change in the estimates for the effects of air pollution. When using the background and semi-background stations as a measure of exposure of the population, similar results were obtained. Then, it can be inferred that the average values of all the stations provided a reasonable estimate of the exposure of the individuals living in São Paulo. Moreover, the validating measurements produced for PM_{10} were obtained from a portable monitoring equipment located about 100 metres from the street level compared to 10 metres for the networked station. Even so, they provide nearly the same measurements.

Finally, the implications of the procedure adopted by CETESB to calculate daily averages should be discussed. The period of measurement used was from 4pm on the previous day

to just before 4pm on the present day. This produces daily averages that already include a one-third day lag period. Therefore, the fact that levels of PM_{10} have an association with mortality on the same day means that mortality was affected by an average which includes levels from the previous day. Thus, some of the delayed effect of pollution was taken into account in the way the daily averages were calculated.

8.3.4. Meteorology and Socio-Demographic Data

Other data used in this study concern meteorological variables and socio-economic and demographic data. They were only used here for making adjustments in the analysis of air pollution and mortality and hospital admissions, and to perform analysis by different socioeconomic groups.

An important consideration that has to be made concerns the measurement of meteorological parameters. Temperature, humidity, pressure, winds and precipitation were all measured at one location only. As well as air pollution, it is expected that some of these variables, or at least temperature, will vary spatially in an urban area. In fact, Sobral [1996] observed that more central areas of São Paulo exhibited temperatures consistently higher than more peripheral areas. The difference in temperature was on average 2 to 4 ^{o}C but it could be as much as 10 ^{o}C occasionally.

Therefore, it could be that the meteorological measurements were not completely representative of all the areas of the city of São Paulo. Once more, however, this error was constant throughout the study period and could only reduce the estimation of variability in levels of meteorological variables experienced by the population.

Another important point is that the classification of study subjects in different socioeconomic categories was based on variables collected at an ecological level, the area where the individuals reside. This ecological approach suffers the same limitations already described for ecological studies. In addition, the composite index used to classify the administrative districts into different quartiles of socioeconomic conditions was calculated giving equal weight to each of the 5 indicators. This index thus, might not

discriminate appropriately the socioeconomic conditions of each area. Therefore, caution should be taken when interpreting the results of that analysis.

8.4. CONFOUNDING

One of the central issues in any epidemiological study is the concept of confounding. In the present study this refers to a situation where an observed Relative Risk of mortality or hospital admission due to exposure to air pollution is in part due to a third factor that is associated with air pollution and independently affects the risk of mortality or hospitalisation. The presence of confounding in a study can lead to an overestimate or underestimate of the true association between exposure and outcome and can even change the direction of the observed effect [Hennekens and Buring,1987].

Since a time series approach is concerned with searching for changes in disease and exposure status in the same individual over time, it has the considerable advantage of being less subject to confounding because the same population is used for exposed and control. Therefore, time-constant covariates such as sex, age, socio-economic status, occupational exposures, cigarette smoking habits, or any other factor whose distribution does not vary from day to day in the study population, are unable to confound the short-term association between air pollution and health.

On the other hand, other influences which vary on a day-by-day basis and are consequently correlated in time with air pollution can obscure the direct estimation of its effects and should be controlled for [Katsouyanni et al,1997a]. Two main groups of variables fall into this category: meteorological variables such as temperature and humidity and chronological or time-related variables such as day-of-week and other cyclical temporal patterns. These potential confounders were, on the whole, easily measured and could be taken into account in the present analysis.

However, the ecological nature of the measurements used for meteorological variables should be pointed out. Along with air pollution, these confounders were not measured at the individual level and are thus, subject to the constraints of an ecological design. Consequently, it is possible that some underadjustment for their effect could have been

made. In addition, meteorological variables were collected at only one site and those measurements were taken as representative of the exposure of the whole study population. If it is considered that these measurements were a poor proxy of the true experience of the subjects then their confounding effect on the association of air pollution and health effects was underestimated. Moreover, when the exposure and the confounding variable are correlated to any degree then the extent of error in measurement in either the confounding or the exposure can significantly affect the adjusted relative risk estimate [Smith and Phillips,1990].

Nevertheless, residual confounding is an issue here to a lesser extent in mortality studies while it may be more important in morbidity studies. More care should be taken in the context of hospital visits or admissions because several factors related to the health care system with characteristics highly correlated to the local situation may introduce spurious associations [Katsouyanni et al,1997a].

8.5. EXPOSURE ASSESSMENT

Among all methodological problems common to any epidemiological research, the assessment of the exposure of individuals is one of the most difficult and is a key obstacle especially for environmental epidemiology [Rothman,1993]. For exposure to air pollution this is clear since it is widespread and such exposure is frequently estimated from measurements of an air quality monitoring network system. These are usually limited to fixed sites near the main sources of pollutants or in highly polluted areas. An average of values obtained from this network is used to express the levels of air pollution for the whole area. Then, all individuals living in that area are considered to have that level of exposure.

Averaging the values from different monitoring stations in the city of São Paulo implied ignoring a lot of variability in levels of air pollution that exist across different areas within the city. Therefore, the exposure of individuals living in different areas was homogenised and consequently, the estimated risk due to exposure to air pollution was possibly underestimated. However, this approach has the advantage of minimising the inaccuracy

in levels of exposure of those study subjects who frequently move throughout different areas and, hence, are exposed to very different levels of pollutants.

Another important issue when assessing exposure to air pollution on the basis of outdoor levels, is that no account is given to the fact that individuals spend most of their time indoors. Some studies have shown that even for students or workers at least 50% of their time was spent at home [Quackenboss et al,1982; Nitta and Maeda,1982]. On the other hand, many investigations have shown that indoor levels of pollutants are generally lower than outside levels unless there are specific sources indoors like gas cooking stoves, unvented space heating, cigarette smoking, volatile emissions, etc.[Quackenboss et al,1991; Quackenboss et al,1982; Nitta and Maeda,1982]. In these situations levels indoor can be higher than outside levels.

However, in cities like São Paulo with mild climates, the proportion of time spent outdoors is much greater than, for example, in North American cities [Loomis et al,1996]. In addition, it is common for people to keep their windows open for extended periods of time and throughout the whole year making the differences between outdoor and indoor levels presumably smaller. Therefore, the use of outdoor levels in this study seems reasonable. Moreover, the question of whether or not outdoor levels are a good proxy of the exposure of the population may be irrelevant, since it is the outdoor levels rather than indoors that can be regulated.

In conclusion, classifying individuals with respect to their exposure status can be regarded as one of the major problems of this study and it is possible that some degree of misclassification has occurred. In general this misclassification is expected to be random or non-differential and thus, it is likely to make the association weaker and more difficult to be detected. However, it may be possible that the exposure profile of those chronically ill (who are more likely to die or being admitted to hospital) is different than the exposure profile of the general population. Furthermore, the very vulnerable individuals may consciously modify their behaviour due to awareness of the harmful effect of air pollution. In any case, the result of such misclassification is the attenuation of the relationship between air pollution and the health endpoint. Consequently, it is therefore unlikely that the reported associations found in this study were spurious ones.

8.6. RESULTS

8.6.1. Mortality

In this study, evidence was found that current ambient air pollution levels in the city of São Paulo have short term adverse health effects. Analysis of daily mortality showed increases in deaths when air pollution levels were elevated on the same day or, in some cases, in the previous day. Overall, the magnitude of the effects was small but statistically significant.

The association between daily levels of air pollution and mortality was for individuals older than 64 years of age for all nonaccidental causes, for cardiovascular and for respiratory diseases. These associations were statistically significant for PM_{10}, SO_2, O_3 and CO in single pollutant models. For all cause mortality (excluding deaths due to external conditions) 100μg/m^3 increases in levels of PM_{10} and SO_2 were associated with 5% and 17% increases in daily mortality respectively. For O_3 a 100μg/m^3 increase was associated with 2% increase in mortality.

These findings are in accordance with a series of reports published in the past few years based on time series analysis of routinely collected data. These studies have shown statistically significant positive associations between measures of air pollution and daily mortality counts. Associations have been reported for SO_2 [Xu et al,1994; Touloumi et al,1996; Spix and Wichmann,1996], O_3 [Sartor et al,1995; Borja Aburto et al,1997; Anderson et al,1996], CO [Rumel et al,1993], but most frequently for particulate concentration, primarily Total Suspended Particles (TSP) and Particulate Matter with less than 10μm in aerodynamic diameter (PM_{10}) [Wordley et al,1997; Pope et al,1992; Schwartz,1993; Schwartz,1991a; Dockery et al,1992; Schwartz,1994a; Schwartz,1991b; Saldiva et al,1995].

In a meta-analysis of several studies conducted by Schwartz [Schwartz,1994d] he reported an increase of 6% in daily mortality for all nonaccidental causes for a 100μg/m^3 change in levels of PM_{10}. In a second review of studies he reported a slightly higher effect of PM_{10} on mortality (9%) [Schwartz,1997]. Dockery and Pope [1994] also reviewed several mortality and air pollution studies and calculated an averaged effect of PM_{10} to be 1% for a 10μg/m^3 change in levels of pollution. Thurston [1996], reviewing a series of studies of

PM_{10} and mortality confirmed that an acute pollution-mortality association could occur at routine ambient levels. The increase in daily mortality for all causes varied between these studies ranging from 5% to 10%.

The size of the effects found in these studies is slightly higher than what was observed in São Paulo. Moreover, all these studies have performed analysis for mortality in the total population, i.e., all ages combined. Nevertheless, the studies that examined the effects of air pollution for different agegroups report effects of about 7 to 10% increase in mortality in the elderly for a $100\mu g/m^3$ increase in particulates or SO_2 [Schwartz and Dockery,1992b; Ballester et al,1996; Schwartz,1994e; Verhoeff et al.1996]. These results are also slightly higher than the estimates found in the present study. In addition, most of these studies concentrated on only one or two pollutants and very few have a more comprehensive analysis looking at the combined effect of several air pollutants.

One reason for this difference in the magnitude of the effects could be the more rigorous approach used in this study to adjust for the potential confounding effects of temporal and meteorological variables. Also, it is expected that some degree of misclassification of disease or exposure status has occurred leading to attenuation of the effects. Perhaps the attenuation is more pronounced in this study. However, there are several dissimilarities between cities in the North compared to cities in the South. Population in developing country cities, including the elderly or other more vulnerable groups, are experiencing other important competing causes of mortality which are together more important than air pollution. The effect of competing causes of death may act in a comparable way to harvesting by depleting the number of vulnerable individuals and thus reducing the relative risk associated with air pollution.

The size of the effect of air pollution on mortality in São Paulo is consistent with the observation discussed in section 2.3.1 that smaller effects are observed in places where levels of pollution are higher. Besides the considerations described above, it could also be that differences in weather conditions, especially in temperature and humidity, characteristic of cities from North and South, might be influencing the composition of the pollution mix and therefore, its effect on human health.

A previous time series study conducted in São Paulo [Saldiva et al,1995] found a statistically significant association between mortality for all causes in the elderly and levels of PM_{10}, SO_2, NO_x, and CO. In that study when all pollutants were included simultaneously in the model only PM_{10} remained significant. In addition, they found a higher coefficient estimate reporting an increase of 13% in mortality for a $100\mu g/m^3$ increase in levels of PM_{10}.

However, this previous study in São Paulo has some methodological differences compared to the present investigation. The time period covered was only one year (may 1990 to april 1991), the averaging periods of air pollution were 24-hours average from midnight to midnight, and only 10 of the 12 monitoring stations of air quality were used to provide daily citywide measurements. But most importantly, the way temporal and meteorological patterns were adjusted differs substantially from the present investigation.

Indicator variables for months of the year were used in the study of Saldiva et al [1995] to control for seasonal and long-term patterns. Although the period of the study spanned only one year, there might have been long-term trends present in the data that were not adequately adjusted by this approach. To control for the effect of temperature, ranges of mean temperature were used. However, it was shown in the sensitivity analysis of the present data that maximum temperature fitted the data better, and the use of other indicators of temperature could easily overestimate the effect of air pollutants.

Therefore, the careful approach employed by this study to adjust for the confounding effects of temperature could have in part explained the relatively smaller effects observed in comparison to other studies. Other reasons to have variations in effect size of PM_{10} between studies include the differences in composition of the mass of particles, the averaging time period employed to calculate daily averages, and differences in whether other pollutants were considered simultaneously in the mortality-PM_{10} model. In addition, the age bands explored in many of the analysis presented here were not always the same which makes comparison somehow more difficult.

The studies that observed an association between SO_2 and all cause mortality have reported effects around 12% for a $100\mu g/m^3$ change in levels of this pollutant [Xu et al,1994; Ballester et al,1996; Touloumi et al,1996; Spix and Wichmann,1996]. For

Ozone, studies have in general found small effects similar to the ones observed in São Paulo. Burnett [Burnett et al,1997] reported Relative Risks for a 30 ppb increase in O_3 ranging from 1.024 to 1.043 for different cities in Canada. This indicates an increase in daily mortality for all causes of 2.4% to 4.3%. In Mexico City, a city which shares several similarities with São Paulo, it was found an effect of 2.4% increase in mortality for a 100-ppb increment in levels of Ozone [Borja Aburto et al,1997]. However, Schwartz in a meta-analysis of studies on O_3 found little consistency for its effects on mortality [Schwartz,1997]. In his study an overall change in daily deaths for a 100$\mu g/m^3$ in O_3 was estimated to be only 0.8% with 95% CI ranging from 1.7 to -0.01.

Multi-pollutant models

Although there are reports of statistically significant results for different components of air pollution, the question of which pollutant or pollutants are the ones more strongly associated with mortality is not yet resolved. In our study we explored models with more than one pollutant and observed that the effects of PM_{10}, SO_2 and O_3 had substantial reduction when analysed concurrently in the same model. In spite of the high degree of collinearity between PM_{10} and SO_2 (r=0.76), it seems that SO_2 remains more stable than PM_{10} in models that include both pollutants. Ozone had a low correlation with both pollutants (0.26 and 0.16 respectively) but exhibited the highest decrease on its coefficient in a multiple pollutant model.

To interpret such results we need to consider that if one pollutant was acting only as a proxy for the other with no contribution of its own, it would be expected that this parameter will become very small in a two-pollutant model while the other would remain unchanged [Katsouyanni et al,1997a]. Therefore, it seems that only SO_2 and possibly PM_{10}, have independent effects on all nonaccidental mortality in the city of São Paulo.

Nevertheless, other studies could find a consistent effect of O_3 on mortality even in the presence of other pollutants in the model [Anderson et al,1996; Sartor et al,1995]. A study conducted in Athens, Greece which has a warmer climate than most cities in Europe or America also found an association with SO_2 and all cause mortality which was, in that case, not confounded by particulates [Touloumi et al,1996; Touloumi et al,1994]. However, studies in the USA generally agreed in finding an effect of particulates that was

independent of the SO_2 and not the other way round. There is still controversy in this matter with some emphasising the high collinearity between pollutants and the fact that they constitute a complex mixture, the components of which may interact with one another in their effects on human health. These considerations make it difficult to conclude that a single component of air pollution is causally associated with adverse effects on human health [Moolgavkar and Luebeck,1996; Moolgavkar et al,1995].

Particle size

While the debate about which pollutant or mix of air pollutants are the real culprits in the association with health effects is still not resolved, some recent research has called attention to the possible important role of the size of the particles. Seaton [Seaton et al,1995] have pointed out that even a small gravimetric concentration of particles represents a very large number of ultrafine particles, and they have suggested that the number of inhaled particles, rather than the mass, is responsible for the health effects observed. In fact, Peters [Peters et al,1997b] in a panel study of asthmatic subjects could demonstrate that the number of ultrafine particles was a better predictor than the mass of the particles. In addition, they found that the effects of the number of the ultrafine particles on Peak Expiratory Flow were stronger than those of PM_{10}.

In summary, adding to the controversy of which air pollutant is the more relevant one, investigators are also discussing if it is the composition of the mass of particles or the size of the particle itself which are the major determinants of the observed association with health effects. Nevertheless, consistent results from studies conducted in different countries, with very diverse climatic conditions, patterns of daily activities, etc., have agreed that increases in air pollution, as roughly indexed by levels of particulate, SO_2, or even some other substance, are associated with adverse effects on human health.

Cardiovascular diseases mortality

Consistent with results found for all non-accidental deaths, mortality due to cardiovascular diseases in the elderly also exhibited an statistically significant association with levels of PM_{10} and SO_2. In addition, an association with CO was also observed for this group of causes of death. The size of the effect was slightly higher. For a $100\mu g/m^3$ increase in daily levels of PM_{10} and SO_2, 6% and 24% increase in mortality respectively

was expected. For CO, a 10μg/m^3 increase corresponded to 8% increase in daily deaths. Since most of the non-accidental deaths comprise cardiovascular deaths, it was expected that effects similar to the ones reported for all cause mortality would be observed.

Several studies have also found effects of air pollution on mortality for cardiovascular diseases in the elderly. Effects have been reported to be between 5-20% for a 100μg/m^3 increase in particulate levels [Schwartz and Dockery,1992b; Dockery and Pope,1994; Wordley et al,1997; Pope et al,1992; Schwartz,1994e; Sunyer et al,1996; Ballester et al,1996; Borja Aburto et al,1997], and around 10-14% for 100μg/m^3 in SO_2 levels [Sunyer et al,1996; Xu et al,1994]. Some studies have also found an association of O_3 with cardiovascular mortality and reported an effect of 3.6% increase in deaths for 100μg/m^3 increase in O_3 [Borja Aburto et al,1997; Anderson et al,1996]. Although we found a similar effect size for O_3 in this study (2.9%) this result was not statistically significant at the 5% level.

Less evidence, however, is found for the effect of CO on cardiovascular mortality. A study of emergency admissions for ischaemic heart disease (IHD) in the largest hospital in São Paulo found that 2.1% of heart attack admissions were attributed to CO [Rumel et al,1993]. However, little control on seasonality was used in this study. More evidence comes from a recent study conducted in London [Poloniecki et al,1997]. Applying time series techniques the authors found a consistent effect of CO on hospital admissions due to heart attacks.

Respiratory diseases mortality

In the present analysis we found significant associations of mortality due to respiratory diseases only for PM_{10}. The association was stronger with measures of this pollutant lagged by one day. A 100μg/m^3 increase in levels of PM_{10} were associated with a 17% increase in respiratory deaths in the elderly in São Paulo, an effect much larger than for all cause mortality.

In accordance, the literature about health effects of air pollution also relates most of the associations of respiratory mortality with PM_{10} [Pope et al,1992; Borja Aburto et al,1997; Dockery and Pope,1994]. However, reported effect sizes were somewhat larger. Those studies reported effects of 30-40% increase in daily mortality for respiratory diseases

associated with PM_{10}. Analysis of specific causes of respiratory mortality also showed associations with deaths from COPD (increases of 19-25%) and pneumonia (11-18%) for the same change in levels of PM_{10} [Schwartz,1994a; Schwartz and Dockery,1992b; Schwartz,1994e].

Although results of studies examining mortality for respiratory diseases were consistent in finding effects of PM_{10}, there are also reports of associations with other pollutants. A study in London found an statistically significant association of O_3 with this group of causes of death. For an increase from the 10^{th} to the 90^{th} centile in levels of O_3, a 5.4% increase in mortality was observed [Anderson et al,1996]. In China, an effect of SO_2 was also observed for respiratory mortality (29% increase for a doubling in levels of SO_2) [Xu et al,1994]. Although results for other pollutants in this study were consistently positively associated with mortality for respiratory diseases, none of those results reached the 5% level of statistical significance.

Results for children's mortality

The present study was unable to find an statistically significant association between mortality in children less than 5 years of age, either for respiratory diseases or for pneumonia infections, with increases in daily levels of air pollution.

Studies that investigated the association of air pollution with mortality in children are scarce. However, such studies are important since this agegroup is not only believed to be more susceptible to the effects of ambient air pollution, but also children are not usually active smokers nor have occupational exposures (potential confounding variables of the association between exposure to air pollution and health effects among adults). In addition, children remain most of their time near or at home. Therefore, studies on children are less prone to such confounders and can provide more precise estimates of effect.

One of the few studies focusing on mortality in children used time series methodology and was conducted in São Paulo. It found an statistically significant association of daily mortality for respiratory diseases and Oxides of Nitrogen (NO_x) [Saldiva et al,1994]. This study reported that 30% of mortality for respiratory diseases in children was explained by variations in levels of NO_x. But once again, that study employed analytical methods which deserve some caution interpretation. Indicator variables for season and month of

the year were used to adjust for seasonal and temporal patterns in the data. In addition, indicator variables were also used to adjust for the confounding effect of temperature and humidity.

Other studies on air pollution and children's mortality were geographical studies. Pena and Dulchiade [Penna and Duchiade,1991] in a ecological study in Rio found an association between infant mortality due to pneumonia infections and levels of Total Suspended Particles (TSP) and SO_2. In the Czech Republic [Bobak and Leon,1992] it was found that the Relative Risk of postneonatal mortality from respiratory diseases was 2.41 times greater comparing the highest and lowest quintile of TSP levels.

Two main reasons can be argued as possible causes for the inability of the present study to find an association of air pollution and mortality in children. The first is the lack of statistical power to detect an association when it exists, specially if it is a small one. A second explanation is the real absence of an association between air pollution and mortality in children.

Deaths of children under 5 years old are a rare event, specially if excluding neonatal deaths as was the case in this study. The mean number of daily deaths in children under 5 years old for respiratory diseases in São Paulo was only 2.7 deaths a day. Power to detect associations in Poisson regression depends on the count or rate of events. In addition, it is likely that the effect of air pollution on children mortality is small when compared to many other risk factors that are believed to be more important than air pollution, specially for children in developing countries. Therefore, the power to detect such a small effect would need a larger count of daily deaths.

However, it could well be that the effect of air pollution on children's mortality is absent. There are studies showing significant associations of air pollution levels and respiratory symptoms in children [Braun-Fahrlander et al,1992; Sobral,1989; Pope and Dockery,1992], or with decrements in pulmonary function [Pope and Dockery,1992; Roemer et al,1993; Hoek et al,1993; Spektor et al,1991] or even with admissions to hospital for respiratory diseases [Xu et al,1995; Burnett et al,1994; Pope,1991; Pope,1989]. The present study also found an association between daily levels of air pollutants and hospitalisations for respiratory diseases.

Thus, it seems likely that levels of air pollution may have an effect on children's health, especially respiratory health. However, such an effect might not be harmful enough to kill those children. Alternatively, the mechanisms that lead to death after exposure to air pollution might be completely different for children when compared to the elderly, saving them from suffering such detrimental effects.

Results according to age and socioeconomic status

An interesting finding of the present study was the modification of the effect of PM_{10} by age. When examining the effect of air pollution on all cause mortality according to age, it was observed that the effect was only positive and statistically significant for individuals over 65 years of age. This effect modification was statistically significant at the 5% level.

A similar analysis performed for respiratory mortality provided less evidence of such a pattern. Although there were more positive effects for older ages, the effect on agegroup 5-12 was also positive (but not statistically significant). This agegroup is the one with fewer number of deaths and is therefore more subject to errors in effect estimates. Nevertheless, it is plausible that the effect of pollution on mortality for respiratory diseases in children really exists. The evidence of respiratory morbidity effects from the studies cited above supports this hypothesis. However, the interaction analysis for this category of cause of death was not statistically significant.

Some few studies have examined the presence of an age effect by dichotomising the mortality counts (at 65 or 70 years) [Schwartz and Dockery,1992b; Ballester et al,1996; Schwartz,1994e; Verhoeff et al.1996]. They usually found higher effect estimates for the over 65 years old group. This study found some evidence that there is a modification of the effect of PM_{10} by age. This finding has important implications for estimations of the health impact of exposure to air pollution and deserves more investigation.

Using a similar method, there was some indication of a modification of the effect of PM_{10} on all cause mortality in the elderly by socioeconomic status. It was observed that individuals living in wealthier areas had a slightly larger relative risk of death compared to those living in more deprived ones. This could be interpreted in the context of competing causes. Poorer individuals are subject to higher risks of death due to several conditions including infectious diseases and violence. The wealthier could be more

vulnerable to the effects of air pollution since they are more 'protected' from other important causes of death.

However, the differences in risk according to socioeconomic condition were not large enough to provide a statistically significant effect modification. The approach adopted to classify the individuals by differing socioeconomic status was rather limited mainly because the districts are large and socially quite heterogeneous. This precludes a more detail assessment of the role of socioeconomic conditions in the risk of health effects due to exposure to air pollution.

8.6.2. Hospital Admissions

There are some advantages in using hospitalisation statistics to assess the health effects of exposure to air pollution. A visit to hospital is a voluntary response by the affected individual, based on perceived needs [Lipfert,1993]. An advantage of hospital admission over mortality is that the latter deal with an inevitable end point and are thus concerned with determining the degree of 'prematurity' in addition to cause-and-effect relationships [Lipfert,1993]. The prematurity is not an issue for studies of hospital admission. On the other hand, hospital admissions may represent more severe cases of illnesses and are thus less sensitive than visits to emergency room to the effects of air pollution. In addition, they represent a morbidity outcome which is also less frequent.

This study found indications that current ambient air pollution levels in the city of São Paulo have short term adverse effects on children's morbidity. Analysis of daily hospital admissions showed that when air pollution levels were elevated on the same day or, in some cases, in the previous day an increase in hospitalisation of children was observed. These associations were found for admissions for all respiratory diseases or admissions for pneumonia infections in children under 5 years old but not for asthma.

For respiratory admissions increases from the 10^{th} to the 90^{th} centile in levels of O_3 and NO_2 were associated with 5% and 6% rise in admissions. The association with O_3 was statistically significant at the 5% level whereas for NO_2 it was only of borderline significance. Other air pollutants also exhibited positive associations with respiratory

admissions but all estimates failed to reached the significant level. However, consistent with the mortality analysis, the largest effects were at lag zero days for most pollutants except SO_2 for which lag one day provided the largest estimate of effect.

Examining admissions for a subgroup of respiratory diseases, similar patterns of association were found. For admissions due to pneumonia infections in children under 5 years old, O_3 and NO_2 showed statistically significant associations. Effects were also larger than for all respiratory admissions. For an increase from the 10^{th} to the 90^{th} centile in levels of O_3 an 8% increase in admissions was observed. For NO_2 the increase in admissions was 9%. For pneumonia admissions among infants, the effects of O_3 (7% increase in admissions) and NO_2 (9% increase in admissions) were only marginally significant. In addition, the patterns of the associations were slightly different with highest coefficient estimates at lag zero for SO_2 and lag 2 for CO. Moreover, the effect of PM_{10} was statistically significant at the 5% level. A 9% increase in admissions for pneumonia infections in infants was found for an increase in levels of PM_{10} from the 10^{th} to the 90^{th} centile.

Different studies have found evidence of an association between exposure to air pollution and morbidity effects in children. The effects examined in these studies included increases in respiratory symptoms [Braun-Fahrlander et al,1992; Pope and Dockery,1992; Sobral,1989], changes in pulmonary function [Pope and Dockery,1992; Roemer et al,1993; Hoek et al,1993; Spektor et al,1991], absence from school [Romieu et al,1992; Ransom and Pope,1992], and recorded episodes of illnesses [Dockery et al,1989; Jaakkola et al,1991; Sobral,1989]. There are only few studies that assessed health effects through hospital admissions of children. Moreover, time series studies of the acute effects of air pollution on hospital admissions have generally been less consistent than for mortality. Effect sizes and pollutants involved varied considerably in the published literature.

Associations of admissions for respiratory diseases for all ages have been described for SO_2 [Bates and Sizto,1983; Burnett et al,1995; Burnett et al,1994] and CO [Burnett et al,1997]. However, more consistent results have been described for O_3 [Burnett et al,1994; Burnett et al,1997; Bates and Sizto,1983; Ponce de Leon et al,1996; Schwartz,1994b; Schwartz,1994c] and PM_{10} [Burnett et al,1997; Pope,1989; Pope,1991; Schwartz,1996; Schwartz,1994b; Schwartz,1994c] suggesting that these two pollutants

might have positive and independent effects on hospitalisations for respiratory diseases. Burnett [Burnett et al,1994] found significant associations between O_3-sulphate pollution mix and admissions for asthma, COPD, and pneumonia infections in Canada. He reported also that largest impacts on respiratory admissions were found on infants (children less than one year old). In a later report [Burnett et al,1997] he stressed the association between hospital admissions for respiratory diseases and O_3.

Pope [Pope,1989; Pope,1991] found statistically significant associations between respiratory admissions and monthly PM_{10} levels in Utah, USA. These associations were also reported to be stronger for children than for adults. However, in London, it was found that Ozone (lagged one day) was significantly associated with an increase in daily admissions for respiratory diseases among all age groups, except 0-14 years old group [Ponce de Leon et al,1996]. No effects of particulate matter was found in London.

In a metanalysis of several studies conducted mostly in North America, Schwartz [Schwartz,1997] found that there was no significant heterogeneity between the effects observed and that the weighted average relative risk of respiratory admissions for a $100\mu g/m^3$ increase in PM_{10} and O_3 were 1.13 and 1.06 respectively. For pneumonia admissions, these risks were 1.13 and 1.07 for each pollutant. These effects are slightly larger than the ones found in the present study and are for studies that examined admissions for all ages.

In this study, an association of changes in daily levels of NO_2 was consistently found for admissions for respiratory diseases and for pneumonia infections. Some few studies have found similar results. Walters et all [Walters et al,1995] found an association of background urban levels of NO_2 with admission rates for respiratory diseases in children under 5 years old after controlling for socioeconomic deprivation and ethnicity. In Barcelona, NO_2 was associated with adults asthma visits to emergency room [Castellsague et al,1995]. In a meta-analysis of 6 European cities, an association was found for admissions due to chronic obstructive pulmonary disease (COPD) [Anderson et al,1997].

There is also some evidence that NO_2 might be associated with respiratory symptoms and illnesses in children [Samet et al,1993; Dockery et al,1989] or with the duration of the

episode of illness [Braun-Fahrlander et al,1992]. In addition, Lipsett [Lipsett et al,1997] found also an association between Emergency Room visits for asthma in children and levels of NO_2. In São Paulo, an association was observed between NO_x and mortality in infants for respiratory causes [Saldiva et al,1994]. However, in a recent air pollution episode in London [Anderson et al,1995] when levels of NO_2 rose to record levels, no association was found for respiratory admissions for children, although positive and significant results were observed for mortality and for admissions in adults.

It should be pointed out that measurements of NO_2 exhibited the highest proportion of missing data for the analysis of hospital admissions (46.9%). In addition, both pollutants, NO_2 and O_3 were measured in much fewer stations than other pollutants like PM_{10} and SO_2. Therefore, it is expected that measurements of exposure to NO_2 and O_3 in this study might not have been as accurate as for other pollutants. This will principally attenuate the relationship leading to lower significant results.

On the other hand, most of the studies used here for comparison, had analysis for all age groups and sometimes only for individuals older than 65 years old. In addition, the analysis for different subgroups of respiratory diseases have probably been conducted using the same core model, i.e. the same set of control variables used for all respiratory admissions. In this study, each outcome was modelled individually and therefore, control variables for seasonal and meteorological patterns differ from one outcome to another. This methodological approach potentially provides better adjusted models with more power to detect relationships.

In addition, negative and mostly not statistically significant association were observed between air pollutants and admissions to hospital for diarrhoeal diseases used here as a control diagnosis. When PM_{10} and SO_2 were lagged by two days, the association with diarrhoea was significant (and negative) at the 5% level. However, admissions for diarrhoea and temperature have a very high correlation which indicates a strong seasonal pattern for this disease with high number of admission during summer or very hot days (when PM_{10} and SO_2 tend to be lower). Thus, it might be that some residual temperature effect was still present in the data which would explain the significant negative results, since a biological protective mechanism of air pollution on diarrhoeal diseases is very unlikely.

Asthma admissions

This study was unsuccessful in finding statistically significant associations between different air pollutants and asthma admissions in children under 5 years old. Nevertheless, apart from O_3, most of the coefficients were positive for this relationship. In addition, there seems to be a stronger lag pattern for these associations since, apart from CO, effects lagged two days had larger estimates.

Previous studies relating air pollution concentrations to hospitalisations for asthma or wheezy conditions have shown conflicting results. In Paris [Dab et al,1996] an association was found with levels of SO_2 and admissions for respiratory diseases which was higher for asthma. In Holland, a study using identical methodology [Schouten et al,1996] did not find any association with asthma or even with other respiratory diseases. Schwartz [1994c] did not find an association between air pollution and asthma in Detroit, although it was present for all respiratory admissions and for pneumonia. However, all these studies have explored the association between asthma admissions and air pollution for all ages, or only in the elderly.

Contradicting the present study, some recent analysis reported positive associations between O_3 and hospital admissions or Emergency Room (ER) visits for asthma in children [Romieu et al,1995; Buchdahl et al,1996], or in all ages [Burnett et al,1995; Burnett et al,1994]. In addition, Burnett [1994] found that this association had a larger impact for children (15% of admissions associated with pollution) and the least effects on the elderly (4%). In Canada, Stieb et al [1996] found a positive and statistically significant association between O_3 and ER visits for asthma with a lag of 2 days for adults but not for children. However, in a recent study Lipsett [Lipsett et al,1997] found associations of PM_{10} and NO_2 with ER visits for asthma in all ages but not for O_3.

Such association has also been described for PM_{10} [Schwartz et al,1993; Pope,1989]. Pope [1989] found stronger associations of PM_{10} levels with admissions for bronchitis and asthma than with admissions for pneumonia and pleurisy in children, results which are in opposition to the findings of the present study.

The failure to find a statistically significant association between air pollutants and asthma might be attributed to different factors. This study explored such effects on children

admitted to hospital with a diagnosis of asthma or bronchitis (ICD-9 codes 466, 490, 491, 492). While problems of diagnosis misclassification have already been discussed, asthma is a disease that might pose extra difficulties for correct ascertainment of diagnosis. There is still no consensus for clinical or laboratory definition of asthma. Its clinical manifestation varies from one individual to another and also from time to time which makes diagnosis difficult [Seaton et al,1989].

In addition, this study was concerned with acute effects of air pollution on exacerbating asthma attacks, not on the initiation of asthma. However, hospitalisations would essentially pick the more severe cases of asthma attacks. Less severe cases are very likely to be larger in number and treated at the Emergency Room and discharged to home. Such cases, therefore, are not included in the hospital admissions dataset.

Levels of air pollutants, specially PM_{10} and O_3 for which an effect on asthma seems to be more evident, were not particularly low in the city of São Paulo during the study period. Therefore, the absence of association cannot be justified by supposedly low levels of air pollution. In addition, among the different metrics generally used for O_3, the indicator used in this study (maximum hourly mean) has been used in other studies that found significant associations with asthma admissions.

Finally, it could be that air pollution has no real effect on asthma admissions. Recent studies have shown that rapid increases in attendance with asthma in hospitals (asthma 'epidemics') are associated with grass pollen released preceding and during thunderstorms and not to air pollution [Celenza et al,1996; Davidson et al,1996; Bauman,1996]. Such a hypothesis does not preclude the possibility that air pollution has a more subtle contribution to triggering attacks of asthma. However, it could be that such effects were too small for children living in São Paulo to be detected by the present study.

8.6.3. Effects by Season

Although careful attention was paid in this study to adjust all the models for the influences of meteorological variables and season, the effects of air pollution on mortality and hospital admissions could still plausibly be different during each season. This could

be caused by an interaction between air pollutants and temperature which varies across the year since the mix of pollutants is expected to be different according to season. Alternatively, it could be simply the result of higher individual exposure based on modified patterns of outdoor activity which may vary from season to season [Moolgavkar and Luebeck,1996].

In this study, it was found that the effect of PM_{10}, O_3 and NO_2 on all cause mortality was greater during the warm season while SO_2 and CO had stronger effects during the cool season. However, only the interaction between O_3 and season was statistically significant at the 5% level. As O_3 is a photochemical oxidant, its formation in the atmosphere is dependent on sunshine and consequently, its levels are higher during the summer months, when sunshine is more abundant. That could explain why the effect of O_3 on mortality has been previously described as greater during the warm months [Anderson et al,1996; Moolgavkar et al,1995].

Nevertheless, it is interesting to note that PM_{10} and NO_2, had higher effects during periods when their levels were not at their highest since both pollutants peaked during the cool season. Several other studies have also found stronger effects of air pollutants during the warm season [Ostro,1995; Saldiva et al,1995; Anderson et al,1996]. However, in contrast to what was found in this study, some of them found also stronger effects of SO_2 during the warm season [Ballester et al,1996; Katsouyanni et al,1997b; Xu et al,1994; Spix and Wichmann,1996].

In São Paulo levels of SO_2 are higher during the winter months and its effect on mortality is also higher during that period. Evidence of similar findings comes only from a study in Athens [Touloumi et al,1996] which also found a stronger effect of SO_2 during the cool season.

In London [Anderson et al,1996] it was observed that although effects of pollutants, particularly O_3 were stronger during the warm season for all cause and cardiovascular mortality, the opposite happened for respiratory mortality when effects of air pollutants were stronger during the cool season. In São Paulo, however, the seasonal effects of respiratory, cardiovascular or all cause mortality were all similar.

Nevertheless, for hospital admissions of children for respiratory diseases the effect of season on the risk for PM_{10} was in the opposite direction to mortality. It was found that the effect of PM_{10} was usually higher in winter months or during the cool season. Other pollutants like SO_2 or CO exhibited a pattern similar to mortality being also higher during the cool season. These findings disagree with some studies that showed stronger effects of air pollution on respiratory admissions during the warm season [Bates and Sizto,1983; Ponce de Leon et al,1996; Burnett et al,1997].

The reasons why effects of PM_{10} according to season were different for respiratory mortality and hospital admissions in the city of São Paulo were not clear. This could well be explained by the fact that mortality was examined in the elderly while hospital admissions were explored only amongst children. However there are also differences between the seasonal effect of pollutants on hospital admissions in São Paulo and elsewhere.

Most of the studies on the effects of air pollution on health have been carried out in cities of the Northern Hemisphere. With few exceptions, weather conditions in those cities, are particularly different than São Paulo which has semi-tropical climate with hot summers and amenable winters. If outdoor activity patterns differ substantially between winter and summer months for cities in Europe and USA, it is likely that patterns of exposure to outdoor air pollution between seasons will also be greatly different in those cities. This is less probable in São Paulo where activity patterns are not likely to change substantially with weather.

The presence of a stronger effect during warm or cool months indicates that in a given season, the effect of pollutants on mortality and hospital admissions are affected by weather conditions, such as temperature. This could mean that perhaps the way seasonality and meteorological factors were controlled for during the modelling steps was inadequate and this would have left a degree of residual confounding of season in the analysis. Alternatively, it could be that the mix of ambient pollutants varies across the seasons and their effects in São Paulo were really different throughout the year. These differences would indicate an interaction between air pollution and season and not necessarily an effect of more exposure to air pollution.

Alternatively, these seasonal differences could be attributed to the 'harvesting effect' producing a stronger effect of pollution in a certain season and depleting the population of highly susceptible individuals for the following season. Whatever is the answer to these questions, it seems reasonable that season could modify the effect of air pollutants on mortality or hospital admissions. However, theses effects deserve further examination in different settings with different weather conditions.

8.6.4. Harvesting

There is clear evidence from this present study and from a vast published time series literature that exposure to current ambient levels of air pollution produces short-term adverse effects on human health. However and specifically for mortality, many have put forward the hypothesis that the effects seen are mostly due to hastening of the death of persons already debilitated who would have died anyway in a few days or weeks independently of the effect of air pollution [Spix et al,1993]. This so-called 'displacement of deaths' or 'harvesting effect' is biologically plausible but only few studies have attempted to demonstrate this hypothesis [Spix et al,1996; Spix,1997].

This study investigated the possibility of a harvesting effect occurring for the series of mortality and hospital admissions in the city of São Paulo. However, none of the tests performed achieved the statistical significant level of 5%, presumably due to insufficient power to detect this phenomenon. Although not significant, it seems that the effect of PM_{10} on all cause mortality in the elderly was weaker during periods following a course of high mortality. This might mean that after a period of increased number of daily deaths (due to air pollution or any other cause), the number of susceptibles among the population becomes smaller, therefore, the effect of air pollution becomes weaker during these periods and until the pool of susceptible is again refilled.

For respiratory hospital admissions of children under 5 years old similar pattern were observed. The effect of O_3 was lower during periods following a high number of admissions, but only over the very short term 1-3 days. This may be only the effect of children admitted not being eligible for re-admission until they have left the hospital again.

However, the concept of 'displacement', i.e. the hastening of admissions that would have occurred anyway in few days or weeks, seems a little inappropriate for hospital admissions. Here the underlying mechanism is that a period of high daily number of admissions depletes the pool of more susceptible children leaving only few to suffer the effects of air pollution in subsequent days.

The importance of this phenomenon concerns the assessment of the public health impact of exposure to air pollutants. Calculations of the public health or long term impact of air pollution on mortality using estimates provided by these time series risk implying that the health impact is substantial. If it is accepted that some of the deaths caused by air pollution are just displacement of deaths by few days or weeks, the overall years of life lost for a period of, say one year, would not be substantially changed.

On the other hand, the importance of the harvesting effect for hospital admissions is likely to be different since hospitalisation is not a once-only event. After an episode of increasing hospitalisation, children admitted and then discharged returns to being at increased risk of being re-admitted. Therefore, one child can, theoretically contribute with many admissions. In this analysis, each admission was treated as independent from another. Thus, following harvesting the pool of susceptible are refilled by both new individuals 'arriving' and by the return of those recovered from an earlier morbidity event.

Moreover, for morbidity outcomes, the effect of harvesting on the public health impact estimates are likely to be different than for mortality. Since there is no displacement of the admissions but only a drop in the number of admissions after a period of increased adverse weather or air pollution circumstances, each admission and possible re-admission will count in the calculations of the burden of, for example, an air pollution episode.

In conclusion, it is methodologically possible to investigate the magnitude of the harvesting effect in time series studies of health events. However, this investigation is in practice limited by power since this effect is not easily detectable in actual data. This is because those affected by an episode of air pollution or adverse weather conditions will react with different delays. Consequently, some will react on the same day, some the next, and a decreasing fraction with longer delays. The overlay of these reaction phases makes

the harvesting effect difficult to be detected. Nevertheless, the extent of its effects on the impact estimates for mortality and morbidity should be considered and attempts to understand this phenomenon and measure its magnitude should be pursued.

8.7. PUBLIC HEALTH IMPACT

There are still a great number of uncertainties about the magnitude of the public health impact of exposure to air pollution. Appropriate techniques to translate the measures of association found in epidemiological studies of air pollution into measures of public health impact are still under development. Quantification of the potential population impact caused by air pollution is important in order to provide governmental health and environmental agencies with guidance for policy making and air pollution control strategies.

Methods to quantify such effects at the population level still face some problems. Consistent information about the health implications of air pollution exposure is limited, especially for developing countries. Air pollution is characterised by a mixture of pollutants and it is very difficult to separate the effects of each pollutant or to appropriately estimate the joint effects of mixtures of air pollutants.

Another issue when evaluating the public health importance of short term effects of air pollution is the possibility that the time of death (and maybe of hospital admission) for many individuals is advanced by only a few days or weeks. Therefore, those deaths would not have any major influence on the overall mortality of that population. Although the issue of mortality displacement has been addressed in this investigation, ways of incorporating its influence in the health impact assessment are still to be developed.

Despite all these uncertainties, the numbers of premature deaths and hospital admissions attributable to exposure to air pollution in São Paulo were calculated and are shown in Table 8.1. The Relative Risk (RR) for each individual outcome was calculated for the annual mean level of air pollution (PM_{10} or NO_2) compared to zero. This assumes a linear relationship between the death rates and levels of pollution and no threshold in the relationship between air pollution and health effects. However, this involves an

extrapolation below the lowest observed values which is clearly not attainable. These RR were then translated into percentage effects attributable to air pollution. These figures were applied to the observed annual number of events for each health endpoint for the period studied in the city of São Paulo to obtain the annual numbers of health events attributable to air pollution (Table 8.1).

According to these calculations, more than 800 deaths a year in individuals over 65 years of age can be attributed to levels of PM_{10} in the city of São Paulo. For cardiovascular mortality, 479 deaths a year and 238 for respiratory diseases are attributed to PM_{10}. In addition, it is estimated that 655 respiratory hospital admissions in children under 5 years of age are attributed to current levels of NO_2 in São Paulo and 700 due to pneumonia infections.

Table 8.1 - Percentage effects for the annual mean level of air pollutants and attributable annual numbers of health events for different outcomes, for São Paulo, Brazil, 1991-1994.

Effect*	% effect for mean levels of air pollution	annual nº attributable to air pollution
elderly:		
all cause mortality	3.3	814
CVD mortality	3.9	479
respiratory mortality	6.1	238
children < 5 :		
respiratory hospital admissions	3.2	655
pneumonia admissions	4.7	700

* effects for the elderly were calculate for annual mean levels of PM_{10} (65.1μg/m^3) and in children for annual mean levels of NO_2 (165.3 μg/m^3)

It should be stressed that these figures were obtained from different models, with different specifications, therefore, which accounts for the slightly larger number of pneumonia admissions compared to all respiratory admissions in children. However, one can conclude that the overwhelming majority of the variation in admissions for all respiratory diseases in children is due to pneumonia infections. Moreover, these may be crude estimates since they are based on effects of a single pollutant. The effects of other components of the mixture of air pollutants are not considered, however were they to be

included, they should not be simply added, because the effects of pollutants are not well separated and adding may result in overestimation.

Ostro [1996] has proposed a methodology to estimate the health effects of air pollutants and, more specifically, the health benefits of a reduction in annual averages of pollution. He proposed that the product of the dose-response relationship, the size of the exposed population at risk, and a change in air pollution levels, generates the total health impact in that population for that change in air pollution. Although differently formulated, this methodology is equivalent to the calculations performed above.

Applying Ostro's methodology for São Paulo, the calculation of the health benefit in mortality for all causes excluding external causes in the elderly for a hypothetical change in PM_{10} are detailed below. The estimated Relative Risk obtained in this study for a unit increase in PM_{10} for all non-accidental mortality in the elderly in São Paulo is 1.0005, the mortality rate for all non-accidental causes in the elderly is 50.79/1000, and the population in this agegroup is estimated from the last census as 479,470 individuals. Annual ambient levels of PM_{10} during the period of this study averaged approximately 65$\mu g/m^3$. Assuming a hypothetical annual change of PM_{10} down to an annual mean of 30$\mu g/m^3$, the total number of deaths in the elderly saved by such a reduction can be calculated as:

Expected deaths = percent effect of PM_{10} in $\mu g/m^3$ x (1/100)
x baseline mortality rate
x change in PM_{10}
x exposed population
Expected deaths = 0.05 x 0.01 x 0.05147 x 35 x 479470 = 432 deaths a year

Therefore, this reduction in annual levels of PM_{10} is expected to save about half of the deaths that are attributed to this pollutant. Applying the same calculations for cardiovascular mortality the number of deaths saved due to CVD per year would be 258 and for respiratory mortality 130. Applying the same approach to the hospital admission data in children under 5 years old for pneumonia and assuming a reduction of 100$\mu g/m^3$

in levels of NO_2 (which are currently at 165.3μg/m^3), around 450 admissions for pneumonia per year in children less than 5 years old would be prevented.

However, these were estimates based on the short-term effects of exposure to air pollution on health. Since, especially for mortality, there is some evidence that at least part of these deaths were just advanced or displaced by few days or weeks, their impact in terms of loss of life expectancy of the population can be quite negligible. For example, some of the 432 deaths in the elderly, potentially preventable by halving the annual averages of PM_{10}, would have occurred anyway in, say, a few days. Some others would have happened only over a longer period, and most likely falling in the following year.

To provide an impression of the overall impact on life expectancy, the proportionate contribution of the premature deaths of these 432 deaths in the elderly may be estimated. They represent 1.8% deaths in that agegroup, and assuming, for example, a mean loss of life expectancy of 6 months for those deaths, this would represent an average loss of life expectancy of only 1 or 2 days, a negligible amount.

No acute effects were observed in individuals younger than 65 years old and little evidence was found for acute effects in children mortality in this study. However, repeated or lasting exposure to air pollution are likely to produce chronic adverse health effects in these individuals. These chronic effects can possibly shorten their life expectancy by periods much longer than few days or weeks. Nevertheless, their deaths are not necessary related to a short term increase in air pollution levels, and can not be directly estimated in these data.

Evidence from cohort studies conducted in the USA [Dockery et al,1993; Pope et al,1995] supports the hypothesis that air pollution also has chronic health effects. Estimates obtained in those cohort studies were applied by Brunekreef to the life tables of the Dutch male population between 20-75 years old [Brunekreef,1997]. He calculated that long term exposure to moderate level of air pollution (in this case, living in a city with mean average of PM_{10} between 30-50μg/m^3) could reduce the life expectancy by 1.11 years. In the public health point of view, this is a quite substantial effect.

Short and long term exposure to air pollution probably have different mechanisms in causing adverse health effects. They constitute different processes, probably largely independent. Nevertheless, their effects can be seen as complementary.

However, it should again be stressed the limitations of these approaches to assess the public health impact of exposure to air pollution. No account has been given to the effects of multiple pollutants and their possible interaction. In addition, only few more evident effects have been more consistently studied. Other important effects in terms of public health are restricted activity days including work loss and school absenteeism. These have also being shown to be associated with air pollution [Ostro,1994; Romieu et al,1992; Ransom and Pope,1992] and their contribution should also be considered when assessing the public health impact of exposure to air pollution.

CHAPTER 9. CONCLUSIONS AND IMPLICATIONS

Urban air pollution is a worldwide problem. While in the developed world there is an emerging problem related to photochemical oxidants and some acidic aerosols, in less developed countries problems still exists with classical pollutants such as SO_2 and Suspended Particulate Matter. As a result of the major short-term episodes that occurred around the 1950's, environmental health policies and emission control measures have been implemented in many countries which led to substantial reductions in air pollution levels. These measures have been more frequently implemented in developed countries which has contributed to widening the gap between them and the countries of the developing world in relation to air pollution levels.

The health effects caused by such high level episodes were easily visible in the very large numbers of individuals affected. However, in the last decades, there has been a wide discussion about the short-term and also long-term health effects due to considerably lower levels, that is lower than the air quality guidelines set up by national and international organisations.

This study was carried out to assess the short term effects of exposure to air pollution on mortality and hospital admissions in the city of São Paulo, Brazil, using routinely collected data in a time series approach. Careful consideration was given to several design and analytical issues pertaining to time series studies. Special attention was also given to minimise the roles of bias and confounding, crucial issues for any epidemiological investigation. In addressing the objectives of this study, the following conclusions can be drawn :

1. This study confirmed that current ambient levels of air pollution in São Paulo are associated with short term adverse effects on mortality for individuals over 65 years old for all causes, excluding accidents and violent deaths. In addition, it was identified that respiratory and cardiovascular diseases are also associated with air pollution. These associations were observed for PM_{10}, SO_2, O_3, and CO. The magnitude of the excess mortality associated with air pollution was small but statistically significant at the 5% level after adjustments (Objective 1).

2. The effect of air pollution on mortality in the elderly was greater for respiratory diseases and cardiovascular diseases when compared to all non-accidental causes (Objective 1).

3. Current ambient levels of air pollution in São Paulo are also associated with short term increases in hospital admissions for respiratory diseases in children under 5 years old. Data for older agegroups were not available for this study. Levels of O_3 and NO_2 are associated with admissions for all respiratory diseases and for pneumonia infections admissions in children. Both pollutants were also associated with admissions for pneumonia in infants (children under 1 year old), although only at a marginally significant level. Levels of PM_{10} also showed an association with admissions for pneumonia infections in infants. No statistically significant associations were observed for asthma admissions (Objective 1).

4. The associations between exposure to air pollution and the development of adverse health effects (mortality or hospital admissions) were observed to happen with measures of pollution on the same day or lagged by one day depending on air pollutant and specific health effect. However, the daily measures of air pollution used in this study were calculated including one third of a day lag (average from 4pm to 4pm) (Objective 1).

5. For mortality, the multiple pollutant models showed that exposure to SO_2, and to a lesser degree PM_{10}, had a clearer association with mortality than other pollutants measured. However, it is difficult to separate the individual effect of each pollutant since they constitute a highly complex mixture with a high degree of collinearity between them. Thus there is uncertainty in properly distinguishing the more hazardous one (Objective 2).

6. Daily maximum temperature provided the best adjustment for the temperature effects in most models examined in this study. It was shown that the use of mean temperature would overestimate the effect of pollutants such as PM_{10} or SO_2 by 15-30%. It is possible that many studies which failed to examine the contribution of maximum temperature would not have adequately accounted for its confounding effect in the air pollution-health effects relationship (Objective 3).

7. The effect of air pollution on mortality increases with age. No statistically significant effects were observed for children under 5 years old or other agegroups up to the age of 65 years old. Only the elderly (65 years old or more) were found to be at a higher risk of death due to exposure to air pollution in São Paulo. These findings have important implications for pollution health impact estimates (Objective 4).

8. In addition, there was some evidence of a slightly higher risk of death among individuals resident in wealthier areas but this difference was not statistically significant. However, the limitations of the approach adopted to classify the socioeconomic status of individuals preclude a more detailed assessment of the role of

these conditions in the risk of health effects due to exposure to air pollution (Objective 4).

9. The magnitude of the effects on health in the elderly observed in this study was generally smaller than what has been observed in studies of elderly populations in North American and European cities, and less even than results from studies that presented analysis for all ages combined. This is consistent with the observation that smaller effects are observed in places where levels of pollution are higher (Objective 5).

10. The results found in this study contrast with earlier preliminary analyses carried out in São Paulo. An effect was found on mortality in individuals over 65 years of age but the magnitude was somewhat smaller. The earlier suggestion of mortality effects in children were not confirmed. The principal reasons for theses differences are probably the much larger dataset on which the present study is based and the more appropriate adjustment for confounding factors. For children, analysis of hospital admission data (carried out for the first time in the city of São Paulo) do indicate an association with changes in levels of air pollution (Objective 5).

Other conclusions that were also reached by this study includes:

11. The results of this study were only weakly sensitive to alternative modelling approaches and to different ways of adjusting for temporal patterns.

12. There was some evidence of a 'harvesting effect' or 'mortality displacement' among the elderly in São Paulo. This provides limited support for a model where exposure to air pollution causes for some individuals (who are already terminally ill) a quite moderate advancement of the time of death. There was also evidence, although weaker, that the same phenomenon might be occurring for hospital admissions in children. This effect has importance in the assessment of the public health impact of exposure to air pollution.

13. Findings from this study are consistent with the substantial international literature relating air pollution to health effects. In addition, an internal coherence of results was observed with pollutants displaying larger effects at the same lag period for different diseases or agegroups, or the same pollutant exhibiting similar effects for related health outcomes.

14. The associations between air pollution and adverse health effects observed in this study do not necessary imply causation. A critical factor in assessing causality in epidemiology is the finding of similar results in multiple studies, in different locations with varying levels of potential confounding factors. This reduces the risk that confounding could produce similar Relative Risks across the different studies. The associations found in São Paulo contribute to the body of evidence relating air pollution to adverse health effects.

15. The overall magnitude of the impact of exposure to air pollution in terms of loss of life expectancy seems to be small. However, various uncertainties still exist on how to adequately assess the impact of air pollution on health using time series data. Nevertheless, due to the ubiquitous characteristic of the exposure to air pollution, an important implication of these findings is that large numbers of individuals living in urban areas like São Paulo are experiencing levels of air pollution that are capable of producing adverse health effects. In addition, effects on children's morbidity contribute a significant public health impact.

In addition to these conclusions, the present study also raised several other questions which are still unsolved and deserve further examination in future studies. Some of the main implications generated by this study are delineated below:

- The role of the so-called 'harvesting effect' in the estimates derived by time series studies needs further elucidation. In addition, better methods for extrapolating from these time series studies into quantitative health impact assessment on exposed populations need to be developed.

- For mortality, the results showed that only the agegroup of individuals over 65 years of age are at an increased risk of death associated with air pollution. They constitute, therefore, the most susceptible group to such exposure and should be regarded as such for prevention programmes purposes.

- The findings from this epidemiological study indicate that current ambient levels of air pollution experienced by the population of São Paulo are associated with adverse health effects. This health impact of environmental air pollution deserves further recognition by those involved in the urban policy planning process. Most of the air pollution in São Paulo is related to motor vehicle emissions, and the fleet of vehicles has been increasing rapidly in the last few years. It is expected, therefore, that levels of pollution and adverse health impacts in São Paulo will increase if appropriate measures are not developed and enforced.

- This study has demonstrated the usefulness of using existing routinely collected data on outcomes and environmental exposures in São Paulo to carry out an epidemiological study. These data are easily accessible and their future use would be encouraged by improvements in quality. In the developing countries context this is an important source of information that should be further explored and studies similar to this should be replicated. In addition, there is a great scope to use data such as these to monitor future interventions and to evaluate the effectiveness or the impact of other enforced measures. In addition, these data can provide decision-makers with more appropriate tools for policy formulation, helping the process of making a bridge between health, environment and policy.

- Finally, further well designed and conducted research on the health effects of exposure to urban ambient levels of air pollution in developing countries context are needed. Still few such studies have been conducted in these countries. The extent to which factors such as temperature, other meteorological parameters, population susceptibility, etc. are modifying the adverse health effects from exposure to air pollution needs to be more fully understood and quantified. A more complete understanding of the relationship between air pollution and health effects would

greatly benefit from studies conducted in different locations, with varying levels of ambient air pollution, weather conditions, and in populations with different health profiles and socioeconomic characteristics.

BIBLIOGRAPHY

Anonymous (1997). Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. The Eurowinter Group. *Lancet*, 349 (9062):1341-1346.

Abercrombie GF (1953). December fog in London and the Emergency Bed Service. *Lancet*:234-235.

Anderson HR, Limb ES, Bland JM, Ponce de Leon A, Strachan DP, and Bower JS (1995). Health effects of an air pollution episode in London, December 1991. *Thorax*, 50 (11):1188-1193.

Anderson HR, Ponce de Leon A, Bland JM, Bower JS, and Strachan DP (1996). Air pollution and daily mortality in London: 1987-92. *BMJ*, 312 (7032):665-669.

Anderson HR, Spix C, Medina S et al. (1997). Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: Results from the APHEA project. *European Respiratory Journal*, 10:1064-1071.

Artaxo P, Castro Jr. WE, Freitas M, and Longo KM (1994). Receptor modelling of atmospheric aerosols in the urban area of São Paulo. University of São Paulo. São Paulo. Progress Report

Bacharova L, Fandakova K, Bratinka J, Budinska M, Bachar J, and Gud aba M (1996). The association between air pollution and the daily number of deaths: findings from the Slovak Republic contribution to the APHEA project. *J Epidemiol Community Health*, 50 Suppl 1:s19-21.

Ballester F, Corella D, Perez Hoyos S, and Hervas A (1996). Air pollution and mortality in Valencia, Spain: a study using the APHEA methodology. *J Epidemiol Community Health*, 50 (5):527-533.

Bascom R (1996a). Health effects of outdoor air pollution. Committee of the Environmental and Occupational Health Assembly of the American Thoracic Society. *Am J Respir Crit Care Med*, 153 (1):3-50.

Bascom R (1996b). Health effects of outdoor air pollution. Part 2. Committee of the Environmental and Occupational Health Assembly of the American Thoracic Society. *Am J Respir Crit Care Med*, 153:477-498.

Bates DV and Sizto R (1983). Relationship between air pollutant levels and hospital admissions in Southern Ontario. *Can J Public Health*, 74 (2):117-122.

Bauman A (1996). Asthma associated with thunderstorms [editorial]. *BMJ*, 312:590-591.

Bennett AE (1981). Limitations of the use of hospital statistics as an index of morbidity in environmental studies. *J Air Pollut Control Assoc*, 31 (12):1276-1278.

Bobak M and Leon DA (1992). Air pollution and infant mortality in the Czech Republic, 1986-88. *Lancet*, 340 (8826):1010-1014.

Borja Aburto VH, Loomis DP, Bangdiwala SI, Shy CM, and Rascon Pacheco RA (1997). Ozone, suspended particulates, and daily mortality in Mexico City. *Am J Epidemiol*, 145 (3):258-268.

Braun-Fahrlander C, Ackermann Liebrich U, Schwartz J, Gnehm HP, Rutishauser M, and Wanner HU (1992). Air pollution and respiratory symptoms in preschool children. *Am Rev Respir Dis*, 145:42-47.

Brunekreef B (1997). Air pollution and life expectancy: is there a relation? *Occup Environ Med*, 54:781-784.

Brunekreef B, Dockery DW, and Krzyzanowski M (1995). Epidemiologic studies on short-term effects of low levels of major ambient air pollution components. *Environ Health Perspect*, 103 Suppl 2:3-13.

Buchdahl R, Parker A, Stebbings T, and Babiker A (1996). Association between air pollution and acute childhood wheezy episodes: prospective observational study. *BMJ*, 312:661-665.

Burnett RT, Brook JR, Yung WT, Dales RE, and Krewski D (1997). Association between ozone and hospitalization for respiratory diseases in 16 Canadian cities. *Environ Res*, 72 (1):24-31.

Burnett RT, Dales R, Krewski D, Vincent R, Dann T, and Brook JR (1995). Associations between ambient particulate sulfate and admissions to Ontario hospitals for cardiac and respiratory diseases. *Am J Epidemiol*, 142:15-22.

Burnett RT, Dales RE, Raizenne ME et al. (1994). Effects of low ambient levels of ozone and sulfates on the frequency of respiratory admissions to Ontario hospitals. *Environ Res*, 65 (2):172-194.

Castellsague J, Sunyer J, Saez M, and Anto JM (1995). Short-term association between air pollution and emergency room visits for asthma in Barcelona. *Thorax*, 50 (10):1051-1056.

Celenza A, Fothergill J, Kupek E, and Shaw RJ (1996). Thunderstorm associated asthma: a detailed analysis of environmental factors. *BMJ*, 312:604-607.

CET (1992). Acidentes de Trânsito - 1992. Prefeitura do Município de São Paulo. São Paulo.

CETESB (1993). Relatório de Qualidade do Ar no Estado de São Paulo - 1992. Compania de Tecnologia de Saneamento Ambiental. São Paulo.

CETESB (1995). Relatorio de Qualidade do Ar no Estado de São Paulo-1995. Companhia de Tecnologia de Saneamento Ambiental. São Paulo.

Chatfield C (1989). The analysis of Time Series - an introduction. London. Chapman & Hall. 4th

Ciocco A and Thompson DJ (1961). A follow-up on Donora ten years after: methodology and findings. *American Journal of Public Health*, 51:155-164.

Colls J (1997). Air Pollution. An introduction. London. E & FN Spon. First edition.

Dab W, Medina S, Quenel P et al. (1996). Short term respiratory health effects of ambient air pollution: results of the APHEA project in Paris. *J Epidemiol Community Health*, 50 Suppl 1:s42-6.

Davidson AC, Emberlin J, Cook AD, and Venables KM (1996). A major outbreak of asthma associated with a thunderstorm: experience of accident and emergency departments and patients' characteristics. Thames Regions Accident and Emergency Trainees Association. *BMJ*, 312:601-604.

Diggle PJ (1990). Time Series. A Biostatistical Introduction. Oxford, UK. Oxford University Press.

Dockery DW, Pope AC, Xu X et al. (1993). An association between air pollution and mortality in six U.S. cities. *N Engl J Med*, 329:1753-1759.

Dockery DW and Pope CA (1994). Acute respiratory effects of particulate air pollution. *Annu Rev Public Health*, 15:107-132.

Dockery DW, Schwartz J, and Spengler JD (1992). Air pollution and daily mortality: associations with particulates and acid aerosols. *Environ Res*, 59 (2):362-373.

Dockery DW, Speizer FE, Stram DO, Ware JH, Spengler JD, and Ferris BG, Jr. (1989). Effects of inhalable particles on respiratory health of children. *Am Rev Respir Dis*, 139 (3):587-594.

Douglas JBW and Waller RE (1966). Air pollution and respiratory infections in children. *British Journal of Preventive and Social Medicine*, 20:1-8.

Erlich R, Findlay JC, and Gardner DE (1979). Effects of repeated exposures to peak concentration of nitrogen dioxide and ozone on resistance to streptococcal pneumonia. *Journal of Toxicology and Environmental Health*, 5:631-642.

Fairley D (1990). The relationship of daily mortality to suspended particulates in Santa Clara County, 1980-1986. *Environ Health Perspect*, 89:159-168.

Firket J (1931). The cause of the symptons found in Meusa Valley during the fog of December, 1930. *Bull Acad R Med Belg*, 11:683-741.

FSEADE (1990). Município de São Paulo. Fundação Sistema Estadual de Análise de Dados. São Paulo.

FSEADE (1993a). Anuario Estatístico do Estado de São Paulo - 1991. São Paulo. Fundação Sistema Estadual de Análise de Dados.

FSEADE (1993b). Conjuntura Demográfica. Fundação Sistema Estadual de Análise de Dados. São Paulo. 24/25

Hatch M and Thomas D (1993). Measurement issues in environmental epidemiology. *Environ Health Perspect*, 101 Suppl 4:49-57.

HEI (1995). Particulate Air Pollution and Daily Mortality. Replication and Validation of Selected Studies. The Phase I - Report of the Particle Epidemiology Evaluation Project. Health Effects Institute. Cambridge, MA.

HEI (1997). Particulate Air Pollution and Daily Mortality. Analyses of the Effects of Weather and Multiple Air Pollutants. The Phase I.B - Report of the Particle Epidemiology Evaluation Project. Health Effects Institute. Cambridge, MA.

Hennekens CH and Buring JE (1987). *Epidemiology in medicine*. Boston: Little, Brown and Co. First edition.

Hoek G, Fischer P, Brunekreef B, Lebret E, Hofschreuder P, and Mennen MG (1993). Acute effects of ambient ozone on pulmonary function of children in The Netherlands. *Am Rev Respir Dis*, 147 (1):111-117.

Holland WW, Bennett AE, Cameron IR et al. (1979). Health effects of particulate pollution: reappraising the evidence. *Am J Epidemiol*, 110 (5):527-659.

Jaakkola JJ, Paunio M, Virtanen M, and Heinonen OP (1991). Low-level air pollution and upper respiratory infections in children. *Am J Public Health*, 81 (8):1060-1063.

Jacobi P (1995). Environmental Problems Facing Urban Households in the City of São Paulo, Brazil. Stockholm Environment Institute. Stockholm.

Kalkstein LS (1991). A new approach to evaluate the impact of climate on human mortality. *Environ Health Perspect*, 96:145-150.

Kalkstein LS (1993). Health and climate change. Direct impacts in cities. *Lancet*, 342:1397-1399.

Kalkstein LS (1995). Lessons from a very hot summer. *Lancet*, 346:857-859.

Kalkstein LS and Smoyer KE (1993). The impact of climate change on human health: some international implications. *Experientia*, 49:969-979.

Katsouyanni K, Anderson HR, and Hoek G (1997a). Design and analysis of temporal studies in air pollution epidemiology. *Air Pollution Epidemiology Report Series 10* (in press).

Katsouyanni K, Touloumi G, Spix C et al. (1997b). Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from time series data from the APHEA project. *British Medical Journal*, 314:1658-1663.

Katsouyanni K, Karakatsani A, Messari I et al. (1990). Air pollution and cause specific mortality in Athens. *J Epidemiol Community Health*, 44 (4):321-324.

Katsouyanni K, Schwartz J, Spix C et al. (1996). Short term effects of air pollution on health: A European approach using epidemiologic time series data: The APHEA protocol. *J Epidem Community Health*, 50:S12-S18.

Katsouyanni K, Zmirou D, Spix C et al. (1995). Short-term effects of air pollution on health: A European approach using epidemiological time-series data. The APHEA project: Background, objectives, design. *European Respiratory Journal*, 8:1030-1038.

Leitmann J (1991). São Paulo Environmental Profile. Urban Environment Case Study Series. UNDP/World Bank/UNCHS.

Lipfert FW (1993). A critical review of studies of the association between demands for hospital services and air pollution. *Environ Health Perspect*, 101 Suppl 2:229-268.

Lipsett M, Hurley S, and Ostro B (1997). Air pollution and emergency room visits for asthma in Santa Clara County, California. *Environ Health Perspect*, 105 (2):216-222.

Loomis DP, Borja Aburto VH, Bangdiwala SI, and Shy CM (1996). Ozone exposure and daily mortality in Mexico City: a time-series analysis. *Research Report Health Eff Inst*, number 75.

Mackenbach JP, Looman CW, and Kunst AE (1993). Air pollution, lagged effects of temperature, and mortality: The Netherlands 1979-87. *J Epidemiol Community Health*, 47 (2):121-126.

Martin AE and Bradley WH (1960). Mortality, fog and atmospheric pollution: an investigation during the winter of 1958-1959. *Monthly Bull Min Health Public Health Lab Serv*, 19:56-72.

Mazumdar S, Schimmel H, and Higgins IT (1982). Relation of daily mortality to air pollution: an analysis of 14 London winters, 1958/59-1971/72. *Arch Environ Health*, 37 (4):213-220.

McCullagh P and Nelder JA (1989). Generalized Linear Models. London. Chapman & Hall.

Melia RJ, Florey CD, and Swan AV (1981a). Respiratory illness in British schoolchildren and atmospheric smoke and sulphur dioxide 1973-7. I: cross-sectional findings. *J Epidemiol Community Health*, 35 (3):161-167.

Melia RJ, Florey CD, and Chinn S (1981b). Respiratory illness in British schoolchildren and atmospheric smoke and sulphur dioxide 1973-7. II: longitudinal findings. *J Epidemiol Community Health*, 35 (3):168-173.

Ministry of Health (1954). Mortality and morbidity during the London fog of December 1952. Reports on Public Health and Medical Subjects. Ministry of Health. London. Number 95

Moolgavkar SH and Luebeck EG (1996). A critical review of the evidence on particulate air pollution and mortality. *Epidemiology*, 7 (4):420-428.

Moolgavkar SH, Luebeck EG, Hall TA, and Anderson EL (1995). Air pollution and daily mortality in Philadelphia. *Epidemiology*, 6:476-484.

Morgenstern H (1982). Uses of ecologic analysis in epidemiologic research. *Am J Public Health*, 72 (12):1336-1344.

Nitta H and Maeda K (1982). Personal exposure monitoring to nitrogen dioxide. *Environ Int*, 8:243-248.

Ostro B (1994). Estimating the health effects of air pollution. A method with an application to Jakarta. Policy Research Working Paper. World Bank. Washington, DC. Number 1301.

Ostro B (1995). Fine particulate air pollution and mortality in two Southern California counties. *Environ Res*, 70 (2):98-104.

Ostro B (1996). A methodology to estimate air pollution health effects. WHO. Geneva. WHO/EHG/96.5

Ostro B, Sanchez JM, Aranda C, and Eskeland GS (1996). Air pollution and mortality: results from a study of Santiago, Chile. *J Expo Anal Environ Epidemiol*, 6 (1):97-114.

Penna ML and Duchiade MP (1991). Air pollution and infant mortality from pneumonia in the Rio de Janeiro metropolitan area. *Bull Pan Am Health Organ*, 25 (1):47-54.

Peters A, Doring A, Wichmann HE, and Koenig W (1997a). Increased plasma viscosity during an air pollution episode: a link to mortality? *Lancet*, 349 (9065):1582-1587.

Peters A, Wichmann HE, Tuch T, Heinrich J, and Heyder J (1997b). Respiratory effects are associated with the number of ultrafine particles. *Am J Respir Crit Care Med*, 155 (4):1376-1383.

Peters A, Goldstein IF, Beyer U et al. (1996). Acute health effects of exposure to high levels of air pollution in eastern Europe. *Am J Epidemiol*, 144 (6):570-581.

PMSP (1992). Plano Diretor de São Paulo. Prefeitura do Município de São Paulo. São Paulo.

Poloniecki J, Atkinson RW, Ponce de Leon A, and Anderson HR (1997). Daily time series for cardiovascular hospital admissions and previous day's air pollution in London, UK. *Occup Environ Med*, 54:535-540.

Ponce de Leon A, Anderson HR, Bland JM, Strachan DP, and Bower J (1996). Effects of air pollution on daily hospital admissions for respiratory disease in London between 1987-88 and 1991-92. *J Epidemiol Community Health*, 50 Suppl 1:s63-70.

Poole C (1991). Multiple comparisons? No problem! [editorial]. *Epidemiology*, 2 (4):241-243.

Pope CA III, Thun MJ, Namboodiri MM et al. (1995). Particulate air pollution as a predictor of mortality in a prospective study of U.S. Adults. *American Journal of Respiratory and Critical Care Medicine*, 151:669-674.

Pope CA, Schwartz J, and Ransom MR (1992). Daily mortality and PM10 pollution in Utah Valley. *Arch Environ Health*, 47:211-217.

Pope CAII (1991). Respiratory hospital admissions associated with PM10 pollution in Utah, Salt Lake, and Cache Valleys. *Arch Environ Health*, 46:90-97.

Pope CA, 3d (1989). Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *Am J Public Health*, 79 (5):623-628.

Pope CA, 3d and Dockery DW (1992). Acute health effects of PM10 pollution on symptomatic and asymptomatic children. *Am Rev Respir Dis*, 145 (5):1123-1128.

Pope CA, 3d, Dockery DW, Spengler JD, and Raizenne ME (1991). Respiratory health and PM10 pollution. A daily time series analysis. *Am Rev Respir Dis*, 144 (3 Pt 1):668-674.

PROAIM (1992). Programa de Aperfeiçoamento das Informações de Mortalidade. Prefeitura do Município de São Paulo. São Paulo. Number 4.

Quackenboss JJ, Kanarek MS, Spengler JD, and Letz R (1982). Personal monitoring for nitrogen dioxide exposure: methodological considerations for a community study. *Environ Int*, 8:249-258.

Quackenboss JJ, Krzyzanowski M, and Lebowitz MD (1991). Exposure assessment approaches to evaluate respiratory health effects of particulate matter and nitrogen dioxide. *J Expo Anal Environ Epidemiol*, 1:83-107.

Ransom MR and Pope CA (1992). Elementary school absences and PM10 pollution in Utah Valley. *Environ Res*, 58:204-219.

Read C (1991). Air Pollution and Child Health. Greenpeace. London.

Roemer W, Hoek G, and Brunekreef B (1993). Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. *Am Rev Respir Dis*, 147 (1):118-124.

Romieu I, Lugo MC, Velasco SR, Sanchez S, Meneses F, and Hernandez M (1992). Air pollution and school absenteeism among children in Mexico City. *Am J Epidemiol*, 136 (12):1524-1531.

Romieu I, Meneses F, Sienra Monge JJ et al. (1995). Effects of urban air pollutants on emergency visits for childhood asthma in Mexico City. *Am J Epidemiol*, 141 (6):546-553.

Rothman KJ (1986). *Modern Epidemiology*. Boston: Little, Brown and Co.

Rothman KJ (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, 1 (1):43-46.

Rothman KJ (1993). Methodologic frontiers in environmental epidemiology. *Environ Health Perspect*, 101 Suppl 4:19-21.

Rumel D, Riedel LF, Latorre M, and Duncan BB (1993). [Myocardial infarct and cerebral vascular disorders associated with high temperature and carbon monoxide in a metropolitan area of southeastern Brazil]. *Rev Saude Publica*, 27 (1):15-22.

Saez M, Sunyer J, Castellsague J, Murillo C, and Anto JM (1995). Relationship between weather temperature and mortality: a time series analysis approach in Barcelona. *Int J Epidemiol*, 24 (3):576-582.

Saldiva PH, Lichtenfels AJ, Paiva PS et al. (1994). Association between air pollution and mortality due to respiratory diseases in children in São Paulo, Brazil: a preliminary report. *Environ Res*, 65:218-225.

Saldiva PH, Pope CA, 3rd, Schwartz J et al. (1995). Air pollution and mortality in elderly people: a time-series study in São Paulo, Brazil. *Arch Environ Health*, 50:159-163.

Saldiva PHN, King M, and Delmonte VCL (1992). Respiratory alterations due to urban air pollution: an experimental study in rats. *Environmental Research*, 57:19-33.

Salinas M and Vega J (1995). The effect of outdoor air pollution on mortality risk: an ecological study from Santiago, Chile. *World Health Stat Q*, 48:118-125.

Samet JM, Lambert WE, Skipper BJ et al. (1993). Nitrogen dioxide and respiratory illnesses in infants. *Am Rev Respir Dis*, 148 (5):1258-1265.

Sartor F, Demuth C, Snacken R, and Walckiers D (1997). Mortality in the elderly and ambient ozone concentration during the hot summer, 1994, in Belgium. *Environ Res*, 72 (2):109-117.

Sartor F, Snacken R, Demuth C, and Walckiers D (1995). Temperature, Ambiente Ozone Levels, and Mortality during Summer, 1994, in Belgium. *Environ Res*, 70:105-113.

SAS (1991a). SAS/GRAPH Software: Usage. SAS Institute Inc. Cary, NC. First Edition (Version 6)

SAS (1991b). SAS/ETS Software: Applications Guide 1. Time Series Modelling and Forecasting, Financial Reporting, and Loan Analysis. SAS Institute Inc. Cary, NC. First Edition (Version 6)

Savitz DA and Olshan AF (1995). Multiple comparisons and related issues in the interpretation of epidemiologic data. *Am J Epidemiol*, 142 (9):904-908.

Schouten JP, Vonk JM, and de Graaf A (1996). Short term effects of air pollution on emergency hospital admissions for respiratory disease: results of the APHEA project in two major cities in The Netherlands, 1977-89. *J Epidemiol Community Health*, 50 Suppl 1:s22-9.

Schwartz J (1991a). Particulate air pollution and daily mortality in Detroit. *Environ Res*, 56:204-213.

Schwartz J (1991b). Particulate air pollution and daily mortality: A synthesis. *Public Health Reviews*, 19:39-60.

Schwartz J (1993). Air pollution and daily mortality in Birmingham, Alabama. *Am J Epidemiol*, 137:1136-1147.

Schwartz J (1994a). What are people dying of on high air pollution days? *Environmental Research*, 64:26-35.

Schwartz J (1994b). Air pollution and hospital admissions for the elderly in Birmingham, Alabama. *Am J Epidemiol*, 139:589-598.

Schwartz J (1994c). Air pollution and hospital admissions for the elderly in Detroit, Michigan. *American Journal of Respiratory and Critical Care Medicine*, 150:648-655.

Schwartz J (1994d). Air pollution and daily mortality: A review and meta analysis. *Environmental Research*, 64:36-52.

Schwartz J (1994e). Total suspended particulate matter and daily mortality in Cincinnati, Ohio. *Environ Health Perspect*, 102:186-189.

Schwartz J (1996). Air pollution and hospital admissions for respiratory disease. *Epidemiology*, 7:20-28.

Schwartz J (1997). Health effects of air pollution from traffic: ozone and particulate matter. In: Fletcher T, McMichael AJ, eds. *Health at the Crossroads. Transport Policy and Urban Health.* Chichester, England: John Willey & Sons Ltd,

Schwartz J and Dockery DW (1992a). Particulate air pollution and daily mortality in Steubenville, Ohio. *Am J Epidemiol*, 135:12-19.

Schwartz J and Dockery DW (1992b). Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am Rev Respir Dis*, 145:600-604.

Schwartz J and Marcus A (1990). Mortality and air pollution in London: a time series analysis . *Am J Epidemiol*, 131:185-194.

Schwartz J and Morris R (1995). Air pollution and hospital admissions for cardiovascular disease in Detroit, Michigan. *American Journal of Epidemiology*, 142:23-35.

Schwartz J, Slater D, Larson TV, Pierson WE, and Koening JQ (1993). Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis*, 147:826-831.

Schwartz J, Spix C, Touloumi G et al. (1996). Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidem Community Health*, 50:S3-S11.

Seaton A, MacNee W, Donaldson K, and Godden D (1995). Particulate air pollution and acute health effects . *Lancet*, 345 (8943):176-178.

Seaton A, Seaton D, and Leitch AG (1989). *Crofton and Douglas's Respiratory Diseases*. Oxford: Blackwell Scientific Publications. Fourth edition.

Smith GD and Phillips A (1990). Declaring independence: why we should be cautious. *J Epidem Community Health*, 44:257-258.

Sobral HR (1989). Air pollution and respiratory diseases in children in São Paulo, Brazil. *Soc Sci Med*, 29 (8):959-964.

Sobral HR (1995). O meio ambiente e a cidade de São Paulo. São Paulo. Makron Books.

Sobral HR (1996). Ilha de calor na cidade de São Paulo: sua dinamica e efeitos na saúde da população. University of São Paulo.(Thesis).

Spektor DM, Hofmeister VA, Artaxo P et al. (1991). Effects of heavy industrial pollution on respiratory function in the children of Cubatão, Brazil: a preliminary report. *Environ Health Perspect*, 94:51-54.

Spix C (1997). Daily time series of mortality counts: Estimating the harvesting effect. *Statistics in Medicine*.(submitted).

Spix C, Heinrich J, Dockery D et al. (1993). Air pollution and daily mortality in Erfurt, east Germany, 1980-1989. *Environ Health Perspect*, 101:518-526.

Spix C, Heinrich J, and Wichmann HE (1996). Daily Mortality and Air Pollution in Efurt (1980-89) and Thuringen (1985-89), East Germany. In: Englert N, Seifert B, Wichmann HE, eds. *Workshop on Air Pollution Epidemiology: Experiences in East and West Europe*. Luxembourg: Office for Official Publications of the European Communities,

Spix C and Wichmann HE (1996). Daily mortality and air pollutants: findings from Koln, Germany. *J Epidemiol Community Health*, 50 Suppl 1:s52-8.

Stephens C, Timaeus I, Akerman M, et al. (1994). Environment and Health in Developing Countries: an analysis of intra-urban differentials using existing data. London School of Hygiene and Tropical Medicine. London.

Stieb DM, Burnett RT, Beveridge RC, and Brook JR (1996). Association between ozone and asthma emergency department visits in Saint John, New Brunswick, Canada. *Environ Health Perspect*, 104 (12):1354-1360.

Sunyer J, Castellsague J, Saez M, Tobias A, and Anto JM (1996). Air pollution and mortality in Barcelona. *J Epidemiol Community Health*, 50 Suppl 1:s76-80.

Sunyer J, Saez M, Murillo C, Castellsague J, Martinez F, and Anto JM (1993). Air pollution and emergency room admissions for chronic obstructive pulmonary disease: a 5-year study. *Am J Epidemiol*, 137 (7):701-705.

Thomas D (1994). Statistical issues in studies of the association between daily mortality and daily pollution. Technical report. University of Southern California. Number 98

Thurston GD (1996). A critical review of PM10-mortality time-series studies. *J Expo Anal Environ Epidemiol*, 6 (1):3-21.

Touloumi G, Pocock SJ, Katsouyanni K, and Trichopoulos D (1994). Short-term effects of air pollution on daily mortality in Athens: a time-series analysis. *Int J Epidemiol*, 23:957-967.

Touloumi G, Samoli E, and Katsouyanni K (1996). Daily mortality and "winter type" air pollution in Athens, Greece--a time series analysis within the APHEA project. *J Epidemiol Community Health*, 50 Suppl 1:s47-51.

UNEP (1987). The State of the World Environment. World Health Organization. United Nations Environmental Programme. Geneva.

UNEP (1989). Environmental Data Report. Blackwell. Oxford.

UNEP (1991). Environmental Data Report. Blackwell. Oxford.

UNEP (1992). Urban Air Pollution in Megacities of the World. Blackwell. Oxford.

UNEP (1993). Environmental Data Report 1993-94. Blackwell. Oxford.

UNEP/WHO (1988). Assessment of Urban Air Quality. Geneva. World Health Organization.

Verhoeff AP, Hoek G, Schwartz J, and Van Wijnen JH (1996). Air pollution and daily mortality in Amsterdam. *Epidemiology*, 7:225-230.

Vigotti MA, Rossi G, Bisanti L, Zanobetti A, and Schwartz J (1996). Short term effects of urban air pollution on respiratory health in Milan, Italy, 1980-89. *J Epidemiol Community Health*, 50 Suppl 1:s71-5.

Walters S, Phupinyokul M, and Ayres J (1995). Hospital admission rates for asthma and respiratory disease in the West Midlands: their relationship to air pollution levels. *Thorax*, 50 (9):948-954.

Ware JH, Ferris BG, Jr., Dockery DW, Spengler JD, Stram DO, and Speizer FE (1986). Effects of ambient sulfur oxides and suspended particles on respiratory health of preadolescent children. *Am Rev Respir Dis*, 133 (5):834-842.

Ware JH, Thibodeau LA, Speizer FE, Colome S, and Ferris BG, Jr. (1981). Assessment of the health effects of atmospheric sulfur oxides and particulate matter: evidence from observational studies. *Environ Health Perspect*, 41:255-276.

WHO (1977). Oxides of Nitrogen. Environmental Health Criteria. World Health Organization. Geneva. 4

WHO (1978). Photochemical Oxidants. Environmental Health Criteria. World Health Organization. Geneva. 7

WHO (1979). Sulphur Dioxide and Suspended Particulate Matter. Environmental Health Criteria. World Health Organization. Geneva. 8

Wojtyniak B and Piekarski T (1996). Short term effect of air pollution on mortality in Polish urban populations--what is different? *J Epidemiol Community Health*, 50 Suppl 1:S36-41.

Wordley J, Walters S, and Ayres JG (1997). Short term variations in hospital admissions and mortality and particulate air pollution. *Occup Environ Med*, 54 (2):108-116.

World Bank (1992). World Development Report 1992 - Development and the Environment. USA. Oxford University Press.

World Bank (1995). Review of methods proposed for estimating the population health risks of urban air pollution. London School of Hygiene and Tropical Medicine. Consultant's Report Prepared for the World Bank, Environmental Department. World Bank.

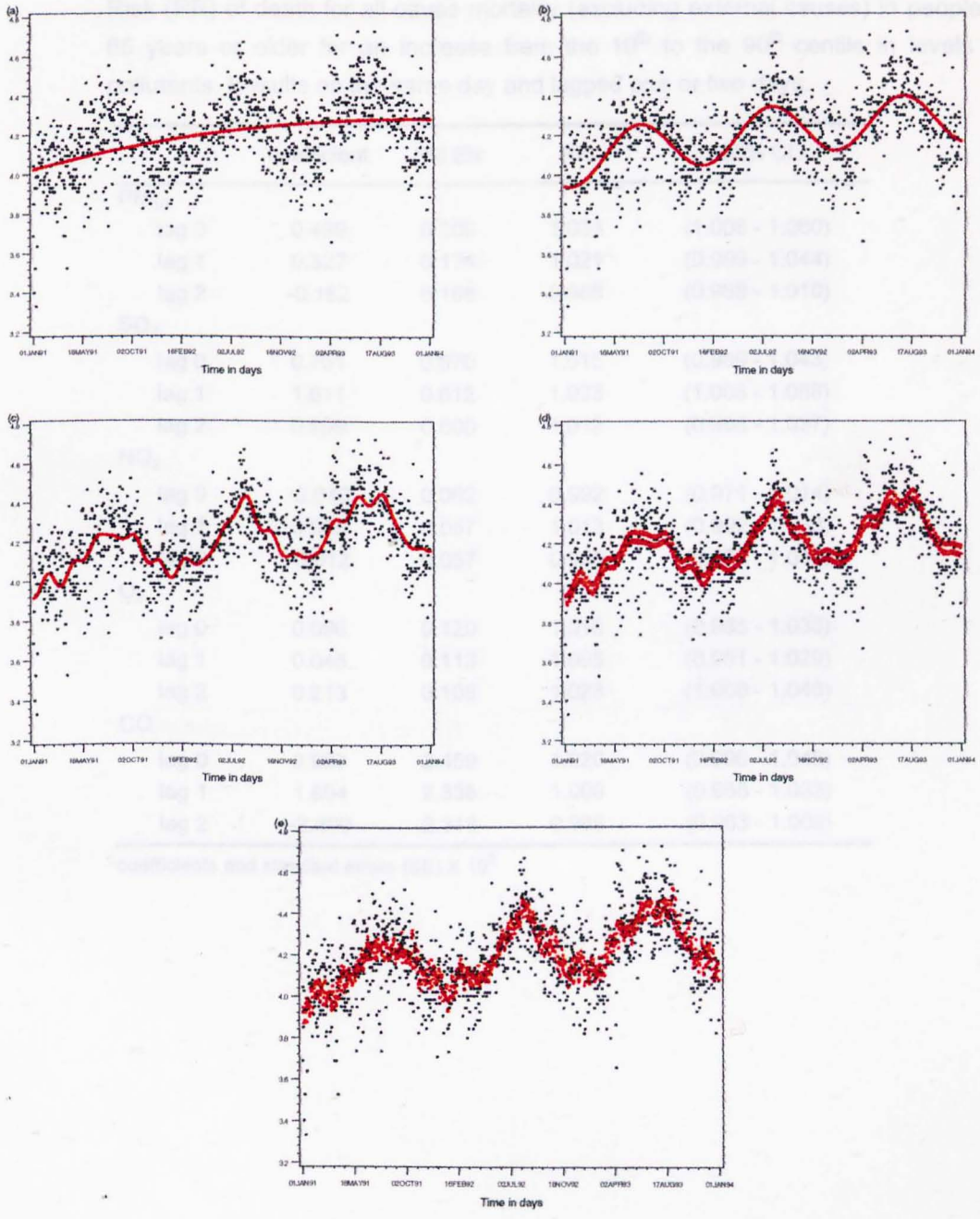
WRI (1996). The Urban Environment. World Resources. A guide to the Global Environment. Oxford University Press. New York.

Xu X, Gao J, Dockery DW, and Chen Y (1994). Air pollution and daily mortality in residential areas of Beijing, China. *Arch Environ Health*, 49 (4):216-222.

Xu X, Li B, and Huang H (1995). Air pollution and unscheduled hospital outpatient and emergency room visits. *Environ Health Perspect*, 103 (3):286-289.

Appendix A

Plot of observed versus predicted values for all cause mortality in the elderly (a) after including linear and quadratic term for trend; (b) after adding sine and cosine waves for one year; (c) after adding all sine and cosine waves; (d) adding indicators for day of the week; (e) adding meteorological terms (full model).



Appendix B

Table B.1 - Poisson regression coefficients for unit change in single pollutant models and Relative Risk (RR) of death for all cause mortality (excluding external causes) in people aged 65 years or older for an increase from the 10^{th} to the 90^{th} centile in levels of air pollutants. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.499	0.209	1.033	(1.006 - 1.060)
lag 1	0.327	0.174	1.021	(0.999 - 1.044)
lag 2	-0.182	0.168	0.988	(0.968 - 1.010)
SO_2				
lag 0	0.751	0.670	1.015	(0.989 - 1.043)
lag 1	1.611	0.612	1.033	(1.008 - 1.058)
lag 2	0.599	0.600	1.012	(0.988 - 1.037)
NO_2				
lag 0	-0.043	0.062	0.992	(0.971 - 1.014)
lag 1	0.069	0.057	1.013	(0.992 - 1.033)
lag 2	-0.012	0.057	0.998	(0.978 - 1.018)
O_3				
lag 0	0.096	0.120	1.010	(0.985 - 1.036)
lag 1	0.045	0.113	1.005	(0.981 - 1.029)
lag 2	0.213	0.108	1.023	(1.000 - 1.046)
CO				
lag 0	3.967	2.459	1.020	(0.996 - 1.046)
lag 1	1.804	2.338	1.009	(0.986 - 1.033)
lag 2	-2.890	2.318	0.985	(0.963 - 1.008)

coefficients and standard errors (SE) X 10^3

Table B.2 - Poisson regression coefficients for a unit change in single pollutant models and Relative Risk (RR) of death for cardiovascular mortality in people aged 65 years or older for an increase from the 10^{th} to the 90^{th} centile in levels of air pollutants. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.584	0.287	1.038	(1.001 - 1.076)
lag 1	0.246	0.239	1.016	(0.986 - 1.047)
lag 2	-0.258	0.233	0.984	(0.955 - 1.013)
SO_2				
lag 0	1.579	0.921	1.032	(0.995 - 1.071)
lag 1	2.136	0.846	1.044	(1.010 - 1.080)
lag 2	1.410	0.831	1.029	(0.996 - 1.063)
NO_2				
lag 0	0.062	0.084	1.011	(0.981 - 1.042)
lag 1	0.078	0.079	1.014	(0.986 - 1.043)
lag 2	-0.042	0.077	0.992	(0.966 - 1.020)
O_3				
lag 0	0.288	0.165	1.031	(0.996 - 1.067)
lag 1	0.018	0.156	1.002	(0.970 - 1.035)
lag 2	0.148	0.149	1.016	(0.985 - 1.048)
CO				
lag 0	7.839	3.329	1.041	(1.007 - 1.076)
lag 1	5.699	3.193	1.029	(0.997 - 1.063)
lag 2	-0.866	3.233	0.996	(0.964 - 1.028)

coefficients and standard errors (SE) X 10^3

Table B.3 - Poisson regression coefficients for single pollutant models and Relative Risk (RR) of death for respiratory mortality in people aged 65 years or older for an increase from the 10^{th} to the 90^{th} centile in levels of air pollutants. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.861	0.471	1.057	(0.996 - 1.121)
lag 1	0.910	0.423	1.060	(1.005 - 1.118)
lag 2	0.051	0.415	1.003	(0.952 - 1.057)
SO_2				
lag 0	1.110	1.517	1.023	(0.963 - 1.086)
lag 1	1.937	1.456	1.040	(0.981 - 1.101)
lag 2	0.661	1.440	1.013	(0.957 - 1.073)
NO_2				
lag 0	-0.113	0.147	0.980	(0.930 - 1.032)
lag 1	0.157	0.138	1.029	(0.980 - 1.080)
lag 2	0.168	0.136	1.031	(0.982 - 1.082)
O_3				
lag 0	0.192	0.277	1.021	(0.963 - 1.081)
lag 1	0.289	0.276	1.031	(0.974 - 1.092)
lag 2	0.467	0.268	1.051	(0.994 - 1.111)
CO				
lag 0	2.131	6.084	1.011	(0.951 - 1.074)
lag 1	-3.135	5.686	0.984	(0.930 - 1.041)
lag 2	-3.854	5.634	0.981	(0.927 - 1.037)

coefficients and standard errors (SE) X 10^3

Table B.4 - Poisson regression coefficients for single pollutant models and Relative Risk (RR) of death for respiratory diseases in children aged 5 years or younger for an increase from the 10^{th} to the 90^{th} centile in levels of air pollutants. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.304	0.998	1.020	(0.899 - 1.156)
lag 1	-0.902	1.210	0.944	(0.810 - 1.098)
lag 2	-0.627	1.172	0.961	(0.828 - 1.112)
SO_2				
lag 0	4.071	3.234	1.086	(0.955 - 1.233)
lag 1	-0.085	3.574	0.998	(0.866 - 1.149)
lag 2	-2.010	3.605	0.960	(0.832 - 1.107)
NO_2				
lag 0	0.328	0.332	1.061	(0.942 - 1.192)
lag 1	-0.455	0.367	0.921	(0.807 - 1.047)
lag 2	0.081	0.356	1.015	(0.893 - 1.150)
O_3				
lag 0	0.116	0.655	1.012	(0.883 - 1.159)
lag 1	0.525	0.674	1.057	(0.918 - 1.215)
lag 2	0.334	0.722	1.036	(0.891 - 1.202)
CO				
lag 0	16.145	13.277	1.086	(0.950 - 1.238)
lag 1	-1.845	14.389	0.991	(0.858 - 1.143)
lag 2	6.849	14.153	1.035	(0.899 - 1.192)

coefficients and standard errors (SE) X 10^3

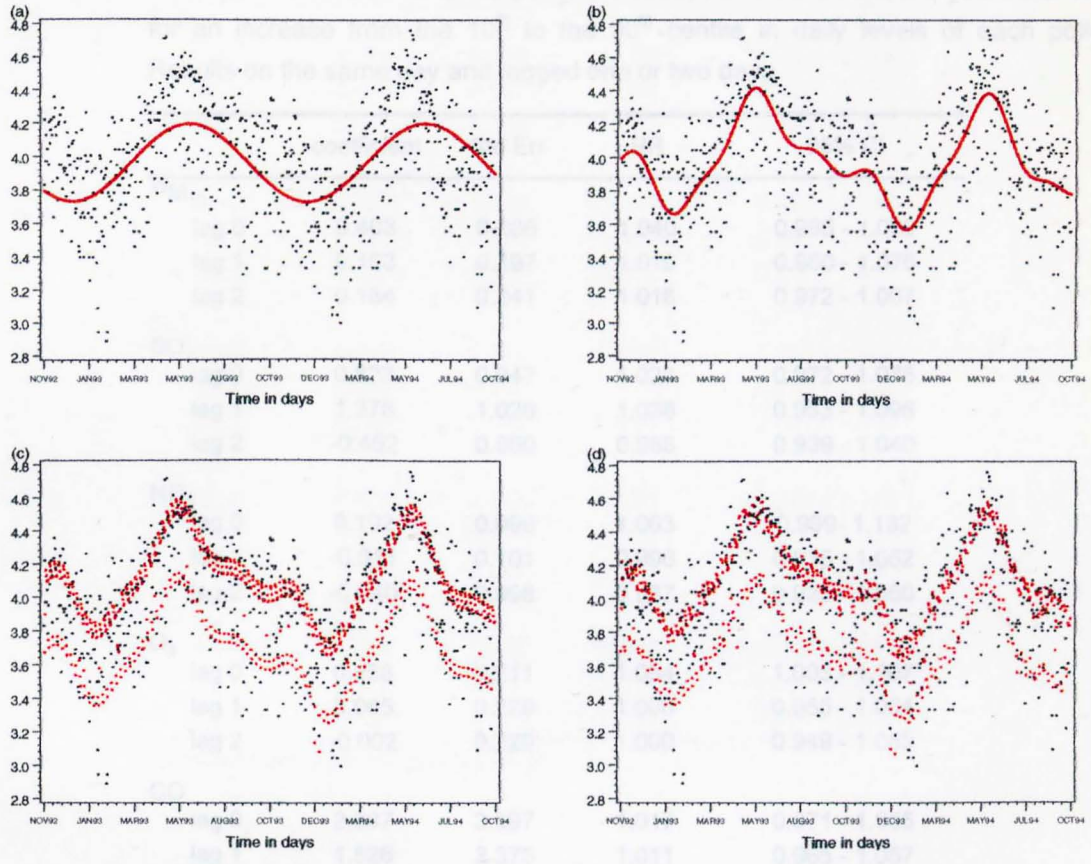
Table B.5 - Poisson regression coefficients for single pollutant models and Relative Risk (RR) of death for pneumonia in children aged 5 years or younger for an increase from the 10^{th} to the 90^{th} centile in levels of air pollutants. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.627	1.126	1.041	(0.900 - 1.196)
lag 1	0.004	1.362	1.000	(0.838 - 1.182)
lag 2	0.078	1.323	1.005	(0.847 - 1.183)
SO_2				
lag 0	5.052	3.636	1.107	(0.959 - 1.280)
lag 1	1.229	4.009	1.025	(0.875 - 1.204)
lag 2	-0.336	4.041	0.993	(0.847 - 1.168)
NO_2				
lag 0	0.302	0.374	1.056	(0.915 - 1.196)
lag 1	-0.278	0.409	0.951	(0.813 - 1.089)
lag 2	0.425	0.396	1.080	(0.931 - 1.236)
O_3				
lag 0	0.043	0.734	1.005	(0.877 - 1.198)
lag 1	0.563	0.753	1.061	(0.934 - 1.284)
lag 2	-0.271	0.817	0.972	(0.845 - 1.193)
CO				
lag 0	14.985	15.123	1.079	(0.918 - 1.242)
lag 1	-4.875	16.333	0.976	(0.818 - 1.133)
lag 2	25.997	15.888	1.141	(0.962 - 1.321)

coefficients and standard errors (SE) X 10^3

Appendix C

Plot of observed versus predicted values for hospital admissions of children under 5 years old for all respiratory causes (a) after adding sine and cosine waves for one year; (b) after adding all sine and cosine waves; (c) adding indicators for day of the week; (d) adding meteorological terms (full model).



Appendix D

Table D.1 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for all respiratory diseases in children under 5 years old. RR are for an increase from the 10^{th} to the 90^{th} centile in daily levels of each pollutant. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.403	0.286	1.040	0.985 - 1.099
lag 1	0.163	0.297	1.016	0.960 - 1.076
lag 2	0.184	0.241	1.018	0.972 - 1.067
SO_2				
lag 0	0.823	0.947	1.023	0.972 - 1.075
lag 1	1.378	1.026	1.038	0.983 - 1.096
lag 2	-0.452	0.960	0.988	0.939 - 1.040
NO_2				
lag 0	0.192	0.099	1.063	0.999- 1.132
lag 1	-0.011	0.101	0.996	0.935 - 1.062
lag 2	-0.040	0.098	0.987	0.928 - 1.050
O_3				
lag 0	0.438	0.211	1.054	1.003 - 1.107
lag 1	0.065	0.229	1.008	0.955 - 1.064
lag 2	-0.002	0.220	1.000	0.949 - 1.053
CO				
lag 0	2.387	3.397	1.017	0.971 - 1.065
lag 1	1.526	3.376	1.011	0.965 - 1.057
lag 2	0.597	3.094	1.004	0.963 - 1.058

coefficients and standard errors (SE) X 10^3

Table D.2 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for pneumonia infections in children under 5 years old. RR are for an increase from the 10^{th} to the 90^{th} centile in daily levels of each pollutant. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.498	0.338	1.050	0.984 - 1.121
lag 1	0.012	0.351	1.001	0.936 - 1.071
lag 2	0.071	0.285	1.007	0.953 - 1.064
SO_2				
lag 0	0.771	1.099	1.021	0.963 - 1.083
lag 1	0.873	1.191	1.024	0.961 - 1.091
lag 2	-0.962	1.119	0.974	0.918 - 1.034
NO_2				
lag 0	0.279	0.117	1.093	1.016 - 1.177
lag 1	0.032	0.124	1.010	0.934 - 1.092
lag 2	-0.164	0.119	0.949	0.881 - 1.023
O_3				
lag 0	0.615	0.253	1.076	1.014 - 1.142
lag 1	0.151	0.278	1.018	0.954 - 1.087
lag 2	-0.042	0.266	0.995	0.935 - 1.059
CO				
lag 0	2.087	3.998	1.015	0.961 - 1.071
lag 1	0.096	3.956	1.001	0.948 - 1.056
lag 2	0.898	3.621	1.006	0.958 - 1.057

coefficients and standard errors (SE) X 10^3

Table D.3 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for pneumonia infections in children under 1 year old. RR are for an increase from the 10^{th} to the 90^{th} centile in daily levels of each pollutant. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.913	0.396	1.094	1.013 - 1.180
lag 1	0.591	0.357	1.060	0.989 - 1.135
lag 2	0.474	0.348	1.048	0.980 - 1.120
SO_2				
lag 0	2.537	1.331	1.071	0.998 - 1.149
lag 1	2.236	1.309	1.062	0.991 - 1.139
lag 2	1.103	1.313	1.030	0.961 - 1.105
NO_2				
lag 0	0.271	0.144	1.091	0.996 - 1.193
lag 1	0.128	0.143	1.042	0.952 - 1.138
lag 2	-0.038	0.139	0.988	0.905 - 1.077
O_3				
lag 0	0.566	0.317	1.070	0.993 - 1.152
lag 1	0.029	0.320	1.004	0.931 - 1.081
lag 2	0.115	0.311	1.014	0.942 - 1.090
CO	0.000	0.000		
lag 0	1.155	4.795	1.008	0.944 - 1.075
lag 1	3.072	4.495	1.021	0.961 - 1.085
lag 2	5.028	4.394	1.035	0.975 - 1.099

coefficients and standard errors (SE) X 10^3

Table D.4 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for asthma in children under 5 years old. RR are for an increase from the 10^{th} to the 90^{th} centile in daily levels of each pollutant. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	0.248	0.527	1.025	(0.926 - 1.133)
lag 1	0.166	0.547	1.016	(0.915 - 1.129)
lag 2	0.513	0.677	1.052	(0.923 - 1.198)
SO_2				
lag 0	1.420	2.061	1.039	(0.931 - 1.159)
lag 1	3.340	2.073	1.095	(0.980 - 1.222)
lag 2	3.732	2.254	1.106	(0.981 - 1.247)
NO_2				
lag 0	0.214	0.251	1.071	(0.914 - 1.251)
lag 1	0.268	0.244	1.089	(0.934 - 1.268)
lag 2	0.317	0.259	1.107	(0.940 - 1.300)
O_3				
lag 0	-0.254	0.484	0.970	(0.866 - 1.086)
lag 1	-0.112	0.484	0.987	(0.881 - 1.105)
lag 2	0.089	0.499	1.011	(0.899 - 1.136)
CO				
lag 0	11.267	7.214	1.081	(0.980 - 1.192)
lag 1	3.536	7.461	1.025	(0.926 - 1.133)
lag 2	7.600	8.050	1.054	(0.944 - 1.175)

coefficients and standard errors (SE) X 10^3

Table D.5 - Poisson Regression Coefficients for single pollutant models and Relative Risks (RR) of hospital admission for diarrhoeal disease in children under 5 years old. RR are for an increase from the 10^{th} to the 90^{th} centile in daily levels of each pollutant. Results on the same day and lagged one or two days.

	coefficient	Std Err	RR	95% CI
PM_{10}				
lag 0	-0.770	0.483	0.927	0.845 - 1.017
lag 1	-0.382	0.519	0.963	0.871 - 1.064
lag 2	-0.945	0.470	0.911	0.833 - 0.997
SO_2				
lag 0	-2.031	1.706	0.946	0.864 - 1.036
lag 1	-1.918	1.814	0.949	0.862 - 1.045
lag 2	-3.640	1.796	0.906	0.823 - 0.996
NO_2				
lag 0	-0.053	0.185	0.983	0.875 - 1.103
lag 1	-0.250	0.189	0.923	0.820 - 1.038
lag 2	-0.132	0.184	0.959	0.854 - 1.075
O_3				
lag 0	0.142	0.347	1.017	0.938 - 1.103
lag 1	-0.249	0.362	0.971	0.892 - 1.056
lag 2	-0.180	0.352	0.979	0.901 - 1.063
CO				
lag 0	-1.472	6.154	0.990	0.910 - 1.076
lag 1	-0.580	6.233	0.996	0.915 - 1.083
lag 2	-8.340	5.776	0.944	0.873 - 1.021

coefficients and standard errors (SE) X 10^3