Mathematical Models for Air Pollution Health Effects

by

Oswaldo Morales Napoles

Thesis submitted in conformity with the requirements for the degree of Master of Science to the Delft University of Technology

Delft, the Netherlands
July, 2004

To Sandra and to my father

Abstract

Mathematical Models for Air Pollution Health Effects

by

Oswaldo Morales Napoles

A wide range of literature on air pollution by particles of aerodynamic diameter less than 10 micrometers and their relation to human health has been developed. Data on population exposure to pollution and its effects on mortality and morbidity have been put forward, and most of these approaches have shown a positive association between increases in concentration levels and increases in mortality.

The Cox proportional Hazards model for survival analysis will be explored and tools for model performance proposed. Regression models that investigate short term health effects of air pollution will be introduced and exemplified with data for a city in central Mexico. Finally expert judgment will be introduced as an element of the set of mathematical devices available for investigating and characterizing the uncertainty regarding air pollution health effects estimates.

Acknowledgments

I want to thank Roger Cook, from whom I learned more than mathematics, for his guidance and support during these two years in Delft. I am grateful to all the members of my graduate committee, John Evans, Leonora Rojas and Dorota Kurowicka, for sharing with me their expertise and knowledge for the completion of this thesis.

I am also grateful to the General Direction of Research on Urban Regional and Global Pollution from the National Institute of Ecology of Mexico for their help with the research from Mexico City and Toluca, and to the Mexican National Institute of Public Health for providing data on mortality.

Finally I also I want to thank my family and friends who have always unconditionally supported me.

Table of Contents.

Chapter 1. Introduction.

1.1 General Remarks on Air Pollution.

The modern world configuration is characterized by the intensive use and production of materials that are not necessarily friendly for the environment. By reaching the air, water, and soil in the earth, these substances may have an effect on humans' health in a greater or lesser degree, depending on their toxicity or concentration for example.

In particular, in the last decades, air pollution and its relationship to human health has become a major concern. Nations around the world have passed laws to prevent certain substances to exceed standard levels that are considered "admissible" and cooperation between countries, states, and different organisms to better understand this problem has been growing.

It is a common practice now days to monitor substances in the air that are considered harmful to humans. The United States of America, for example has six *criteria pollutants* according to the Environmental Protection Agency that are monitored constantly because they are typical across the country and are considered as potentially harmful to humans. These are particulate matter, sulfur dioxide, carbon monoxide, nitrogen dioxide, ozone, and lead. The World Health Organization in setting guidelines for air quality standards for Europe recognizes four *classical air pollutants*: nitrogen dioxide, ozone and other photochemical oxidants, particulate matter and sulfur dioxide. In Mexico the Health Secretary together with the Secretary for Fisheries, Natural Resources and Environment formulated and published in December 1994 the Mexican Official Norms *criteria pollutants* i e. ozone, carbon monoxide, sulfur dioxide, nitrogen dioxide, particles and lead. A brief description of these pollutants is included in Appendix A1.1.

This research presents some of the methods commonly used by epidemiologists to explain the relationship between air pollution and mortality, proposes some tools for analysis and exemplifies when possible with results for studies conducted in Europe, United States and Mexico City. Also, new results are presented with mortality and air pollution data from the Mexican city of Toluca and recent tools for modeling uncertainty are exemplified with an exercise conducted in Mexico City with air pollution experts. Before continuing the reader is referred to appendix A1.2 to familiarize with some common definitions in epidemiology that will be used during the rest of the research.

1.2 Particulate Matter and Mortality

One of the pollutants that has been more studied in the past years is particulate matter (PM). A wide range of literature in air pollution by PM and its relation to human life has been developed. Data in population's exposure to pollution and its effect on mortality and morbidity have been set forward and most of these approaches have shown a positive association between increases in concentration levels and increases in mortality.

Basically, two approaches have been used to study the relationship between $PM_{2.5,10}$ and health effects, both are environmental in the sense that population exposure to air pollution is

measured in central monitors rather than with individual monitors. The first one is referred to as time series analysis, which measure acute effects on health.

The second approach is referred to as cohort studies. These are also environmental in the sense described above, but they investigate the relationship between air pollution and mortality following a population through time and gathering data for the population at hand. In this case, the event of dying is the endpoint of a person's lifetime, and information about the time to death or survival time is available.

These two kinds of models will be discussed in more detail in chapters 2 and 3 for the time being we will be content with showing some results found across the literature. It is important to mention that this section does not pretend to be exhaustive in any way; important differences in methods, results and interpretations can be found in the literature, the purpose of this section is to make the reader familiar with the results found in previous studies and, when possible, make these results comparable with the results from this investigation.

1.2.1 Cohort Studies Results.

Cohort studies use the Cox Proportional Hazards Model or some variation of this to investigate long-term health effects of air pollution. As stated before, it becomes a little difficult to make comparisons across studies basically due to differences in methodology (including data sources and air pollution measures) found in different studies. However, bearing this constraint in mind the reader may look at table A1 in the appendix where some results from well know cohort studies are summarized.

1.2.1.1 Method.

For the Cox Proportional Hazards Model (CPHM) all that matters is the order of event times when the covariates are time invariant, estimates of relative risk or ratio of hazards for individuals with different covariate values (given by coefficients) i.e. the change in risk per unit change in the predictor variable are estimated by maximizing the **ln** of the partial likelihood function:

$$PL(\beta) = \prod_{i=1}^{n} \left[\frac{e^{Z_i \beta}}{\sum_{j \in R(t_i)} e^{Z_j \beta}} \right]^{c_i}$$

In the expression above Z is the vector of covariates and the summation in the denominator runs over the individuals in the risk set, *c* is an indicator variable taking the value of cero when a censored time is observed.

1.2.1.2 The Extended Adventist Health Study of Smog.

The Extended Adventist Health Study of Smog (AHSMOG) study followed about 6000 white, non-Hispanic, non-smoking, long-term California Seventh-Day Adventists residents for 6 to 10 years beginning in 1977 and updated through 1992. The participants were aged 27-

95 and completed lifestyle questionnaires that ascertained different individual variables used in the study. All air quality monitors in the state were used to create individual exposure profiles.

Estimates of monthly ambient concentrations of PM_{10}, ozone, sulfur dioxide, and nitrogen dioxide were formed for study participants for the period 1966-1992 using site monitoring stations maintained by the California Air Resources Board. Concentrations of PM_{10} prior to 1987 were estimated using site and season-specific regression based on TSP.

1.2.1.3 The Harvard Six Cities Study.

In the Harvard Six Cities Study (SCS), beginning in 1974 an 8111 white subjects (aged 25 to 74 at enrolment) cohort was followed for 14 to 16 years. Participants completed a questionnaire assessing variables such as age, sex, weight, height, education level, complete smoking history, occupational exposures, and medical history.

Outdoor concentrations of total suspended particulate matter, sulfur dioxide, ozone and suspended sulfates were measured in each community at centrally located site monitors. Adjusted mortality rate ratios were estimated for a difference in pollution equal to that between the city with highest levels of air pollution and the city with the lowest levels –that is the adjusted rate ratios across the range of exposure for each pollutant among the six cities.

1.2.1.4 The American Cancer Society Study.

On the extended American Cancer Society study (ACSS), ACS volunteers enrolled individual participants in the fall of 1982. Participants were residents across the 50 states of the USA, the District of Columbia and Puerto Rico. All were adults 30 years or older, and completed a questionnaire that elicited sex, age, weight, height, smoking history, alcohol use, occupational exposures, diet, education, marital status, and other characteristics. Mean concentrations of air pollution were compiled from various primary data sources. The quarterly mean values for all stations in each metropolitan area were calculated across the study years using daily averages except for ozone where daily 1-hour maximums were used and calculated for the full year and the third quarter only.

1.2.1.5 Discussion.

For a better understanding of the estimates presented in table A1 the reader is referred to the original paper. All the relative risk estimates (that have a multiplicative effect in baseline mortality for a given increment in air pollution) presented in Table A1 have been found statistically significant in the original report. The reader must also have in mind that even when the estimates are presented for 10 unit increase in concentration levels, these are not always reported in the original studies for an air pollution increase of this magnitude, sometimes they are reported for the difference in concentration between the most and least polluted sub population, or an inter quintile range of concentration estimate.

As the reader may see, central estimates of risk ratio for a 10 microgram per meter cubic increase in the concentration of a given pollutant range from 0.99 to 1.33, and the confidence

bounds range from 0.78 to 1.10 for the lower bound and from 1.05 to 1.62 in the upper bound. These estimates are of course for different causes of mortality and pollution measures but few remarks are clear from a first look at table A1: if the models used in each study are correct, then the health effects are not the same for every pollutant and; second, there is a great deal of uncertainty regarding the true value of the coefficient that is considered as the per cent change in mortality due to 1 unit increase in air pollution if the population is homogeneous.

Next, short term effects of air pollution studies will be summarized in the same way as the results for cohort studies.

1.2.2 Time Series Studies Results.

Generalized Linear Models (GLM), Generalized Additive Models (GAM) or GLM with regression splines, also referred to as "the fully parametric alternative for the GAM with non parametric smoothers" are often used to estimate effects associated to air pollution while accounting for smooth fluctuations in the mortality that confound estimates of air pollution effect; these models will be discussed in some detail in chapter 4. The use of GAMs, however has become common because it allows for non parametric adjustment for non linear confounding effects of seasonality, trends and weather variables and it is a more flexible approach than fully-parametric alternatives like the GLM with cubic splines.

1.2.2.1 Method.

Generally time series studies assume a Poisson distribution for the daily counts of the response variable (non-accidental, cardiovascular and/or respiratory disease related deaths) and regress them against some measure of variables that change in time, for example concentrations of other pollutants such as SO_2, O_3, NO_2, or temperature. An Additive Model is fitted as:

$$g(E(Y|x)) = g(\mu) = \alpha + \sum_{i=1}^{p} f_i(x_j)$$

Where f_i is a nonparametric function estimated from data using smoothing operations, describing the relationship between each of the i predictors and the transformed mean response $g(\mu)=\ln(\mu)$.

Unlike linear regression models, which are fitted by weighted least squares and have an exact solution, the estimation procedure for a GAM (or a GLM with splines) requires iterative approximations to find the optimal estimates.

The coefficient for a particular predictor is the logarithm of the ratio of mortality rates on two days with values of the predictor that differ in one unit. Log linear regression coefficients multiplied by 1,000 can be interpreted as approximately the percentage of change in risk (mortality) associated with a ten-unit change in the predictor variable for a small coefficient.

1.2.2.2 The National Mortality, Morbidity and Air Pollution Study.

Perhaps, the National Mortality, Morbidity and Air Pollution Study (NMMAPS) and the APHEA project are the largest projects that investigate short term effects of air pollution in health. The NMMAPS is the largest multi-site time series study conducted. In 90 USA cities the daily total mortality was regressed on PM_{10} with different lags to estimate the relative increase in mortality rates associated with a 10 $\mu g/m^3$ increase in PM_{10}.

The goal of NMMAPS was to estimate city-specific, regional and national effects of PM_{10} on mortality. To pool the results from each city, a two stage hierarchical model and a three-stage regional model with non-informative priors on the variance components was used. This investigation will not discuss the models for estimating a multi-site coefficient.

1.2.2.3 Other Time Series Studies.

There might be differences between Europe and the United States, and within Europe Itself, which might influence the health effects of air pollution; these include emission sources, pollution mixtures, climate, lifestyle, and the underlying health of the population. The APHEA project included data from 15 cities including 5 in central-eastern Europe (4 Polish and 1 Slovakian). In this case, the summary estimates from the city specific estimates were weighted means of the regression coefficients, with weights inversely proportional to local variances.

In Mexico time series studies have also been conducted, however, these have not been multi-site but performed only in MCMA. The total population of MCMA at the moment of the studies was of at least 16 million people; again, emission sources, pollution mixtures, climate, lifestyle, and the underlying health of the population are factors that could be influential for different health effects outcomes in this setting.

Table A2 shows some results from different time series analysis performed in different locations, once more, the reader must be aware that the estimates presented in the table might again be difficult to compare because of the methods employed, the quality of the information available (for mortality cause and environmental estimates of pollutants' concentrations). One more thing to take into consideration is that although there are several reasonable alternatives for obtaining adequate control for confounding, an optimal method has not been identified.

1.2.3 Uncertainty.

In previous sections a brief discussion on the kind of studies and results available in air pollution health effects literature was set forward. However, there are large sources of uncertainty with respect to the different studies conducted. For time series studies, some sources of uncertainty include the differences in the regression model design used across studies, the uncertainty introduced by pollution exposure estimates, the differences in lags between exposure and response used by different studies and the population characteristics; for example considering more sensitive subgroups as the elderly should increase the magnitude of the coefficients but also the random variation due to the smaller number of

deaths. Studies such as NMMAPS have tried to deal with some of these and other issues that introduce uncertainty in time series models.

In the case of cohort studies the major criticisms have been in the sense that ambient pollution concentrations are poor surrogates for personal exposures because of population's mobility and time spent outdoors vs. indoors for example. Also, because of the characteristics of the Cox proportional hazards model, the population studied is assumed to be homogenous in other risk factors affecting survival times such as economic status or education level and this assumption is generally not satisfied across the studies available. More recent findings (Pope et.al., 2002) have tried to deal to some extent with these issues, but more work in this respect is still needed.

An alternative approach to get a picture of how the scientific community perceives these issues is with expert judgment. The classical model for expert judgment has been introduced in (Cooke, 1991) and applied in many risk and reliability studies. This model for combining expert judgments bears its name because of a strong analogy with classical hypothesis testing. Expert judgment is performed for multiple reasons, for example, in many cases, there might not be sufficient data and hence a fair amount of uncertainty. The scientific foundation for subjective probability comes from the theory of rational decision-making; hence, the main aim of the method is to provide the basis for achieving rational consensus.

The reader is referred to chapter 4 and the original references for details about the method. It is good however to have in mind that for the case at hand, structured expert judgment could be a useful tool for characterizing the uncertainty regarding coefficients estimated by time series studies and cohort studies and hence for decision making.

1.3 About the rest of the thesis.

So far a brief introduction to the mathematical models used for exploring air pollution health effects and main findings of these has been presented. In the rest of the thesis the models will be explored in more details and examples shown when possible. In chapter 2 the reader can find a discussion on the Cox proportional hazards model as used by epidemiologists. Since conducting a cohort study represents a very large availability of human, economic and time related resources that this investigation lacks at the moment, the analysis of the model will use data available in different sources and Monte Carlo methods.

Chapter 3 will present the theory behind GLMs and GAMs not only in the mathematics but also as they are used by epidemiologists to investigate acute effects of air pollution in health. Examples will be presented for the city of Toluca (in central Mexico) using mortality, air pollution and atmospheric data for 1999-2000.

In chapter 4 the theory for expert judgment will be sketched and an exercise recently conducted with Mexican experts on air pollution health effects will be discussed. Finally in chapter 5, comparisons with results from previous chapters will be performed and conclusions and recommendations based on the findings of this investigation will be presented for the reader's consideration. Since each chapter keeps a certain amount of independence from other chapters, appendixes are presented at the end of each chapter that will help the reader through each chapter.

APPENDIX 1

A1.1 Review of Criteria Pollutants

1. Nitrogen Dioxide (NO_2)

Nature and sources. This is a reddish brown gas formed in the ambient air through the oxidation of Nitric Oxide (NO). Nitrogen Oxides (NO_x) contribute to the formation of ozone, particulate matter, haze, and acid rain. The major source of man made NO_x's are high temperature processes such as those that occur in power plants and automobiles. Home heaters and gas stoves are big sources of indoors NO_2.

Health and Environmental Effects. Short-term exposures (e.g., less than 3 hours) to low levels of NO_2 may lead to changes in airway responsiveness and lung function in individuals with preexisting respiratory illnesses. These exposures may also increase respiratory illnesses in children. Long-term exposures to NO_2 may lead to increased susceptibility to respiratory infection and may cause irreversible alterations in lung structure. NO_x react in the air to form ground-level ozone and fine particle pollution, which are associated with adverse health effects.

NO_x contribute to a wide range of environmental effects directly and when combined with other precursors in acid rain and ozone. Increased nitrogen inputs to terrestrial and wetland systems can lead to changes in plant species composition and diversity. Similarly, direct nitrogen inputs to aquatic ecosystems such as those found in estuarine and coastal waters can lead to eutrophication (a condition that promotes excessive algae growth, which can lead to a severe depletion of dissolved oxygen and increased levels of toxins harmful to aquatic life).

Nitrogen, alone or in acid rain, also can acidify soils and surface waters. Acidification of soils causes the loss of essential plant nutrients and increased levels of soluble aluminum that are toxic to plants. Acidification of surface waters creates conditions of low pH and levels of aluminum that are toxic to fish and other aquatic organisms. NO_x also contribute to visibility impairment.

2. Ozone (O_3)

Nature and Sources. Ground-level ozone is the primary constituent of smog. Ozone is not emitted directly into the air but is formed by the reaction of Volatile Organic Coumpounds and NO_x in the presence of heat and sunlight. Ground-level ozone forms readily in the atmosphere, usually during hot summer weather. VOCs are emitted from a variety of sources, including motor vehicles, chemical plants, refineries, factories, consumer and commercial products, and other industrial sources. NO_x is emitted from motor vehicles, power plants, and other sources of combustion. Changing weather patterns contribute to yearly differences in ozone concentrations from region to region. Ozone and the pollutants that form ozone also can be transported into an area from pollution sources found hundreds of miles upwind.

Health and Environmental Effects. Short-term (1 to 3 hours) and prolonged (6 to 8 hours) exposures to ambient ozone have been linked to a number of health effects of concern. For example, health effects attributed to ozone exposure include significant decreases in lung

function and increased respiratory symptoms such as chest pain and cough. Exposures to ozone can make people more susceptible to respiratory infection, result in lung inflammation, and aggravate preexisting respiratory diseases such as asthma. Also, increased hospital admissions and emergency room visits for respiratory problems have been associated with ambient ozone exposures. These effects generally occur while individuals are actively exercising, working, or playing outdoors. Children that are active outdoors during the summer when ozone levels are at their highest are most at risk of experiencing such effects. Other at-risk groups include adults who are active outdoors (e.g., some outdoor workers) and individuals with preexisting respiratory disease such as asthma and chronic obstructive pulmonary disease. In addition, longer-term exposures to moderate levels of ozone present the possibility of irreversible changes in the lung structure, which could lead to premature aging of the lungs and worsening of chronic respiratory illnesses.

Ozone also affects vegetation and ecosystems, leading to reductions in agricultural crop and commercial forest yields, reduced growth and survivability of tree seedlings, and increased plant susceptibility to disease, pests, and other environmental stresses (e.g., harsh weather). In long-lived species, these effects may become evident only after several years or even decades, thus having the potential for long-term effects on forest ecosystems. Ground-level ozone damage to the foliage of trees and other plants can also decrease the aesthetic value of ornamental species as well as the natural beauty of our national parks and recreation areas.

3. Sulfur Dioxide (SO_2)

Nature and Sources. Sulfur dioxide belongs to the family of SO_x gases. These gases are formed when fuel containing sulfur (mainly coal and oil) is burned at power plants and during metal smelting and other industrial processes. Most SO_2 monitoring stations are located in urban areas. The highest monitored concentrations of SO_2 are recorded near large industrial facilities. Fuel combustion, largely from electricity generation, accounts for most of the total SO_2 emissions.

Health and Environmental Effects. High concentrations of SO_2 can result in temporary breathing impairment for asthmatic children and adults who are active outdoors. Short-term exposures of asthmatic individuals to elevated SO_2 levels during moderate activity may result in breathing difficulties that can be accompanied by symptoms such as wheezing, chest tightness, or shortness of breath. Other effects that have been associated with longer-term exposures to high concentrations of SO_2, in conjunction with high levels of PM, include aggravation of existing cardiovascular disease, respiratory illness, and alterations in the lungs' defenses. The subgroups of the population that may be affected under these conditions include individuals with heart or lung disease, as well as the elderly and children.

Together, SO_2 and NO_x are the major precursors to acidic deposition (acid rain), which is associated with the acidification of soils, lakes, and streams and accelerated corrosion of buildings and monuments. SO_2 also is a major precursor to $PM_{2.5}$, which is a significant health concern, and a main contributor to poor visibility.

4. Particulate Matter (PM).

Particulate matter is the general term used for a mixture of solid particles and liquid droplets found in the air. Some particles are large enough to be seen as dust or dirt. Others are so small they can be detected only with an electron microscope. $PM_{2.5}$ describes the "fine" particles

that are less than or equal to 2.5 µm in aerodynamic diameter. "Coarse fraction" particles are greater than 2.5 µm, but less than or equal to 10 µm in diameter. PM_{10} refers to all particles less than or equal to 10 µm in aerodynamic diameter (about one-seventh the diameter of a human hair). PM can be emitted directly or formed in the atmosphere. "Primary" particles, such as dust from roads or black carbon (soot) from combustion sources, are emitted directly into the atmosphere.

Health and Environmental Effects. Particles that are small enough to get into the lungs (those less than or equal to 10 µm in diameter) can cause numerous health problems and have been linked with illness and death from heart and lung disease. Various health problems have been associated with long-term (e.g., multiyear) exposures as well as daily and, potentially, peak (e.g., 1 hour) exposures to particles. Particles can aggravate respiratory conditions such as asthma and bronchitis and have been associated with cardiac arrhythmias (heartbeat irregularities) and heart attacks.

Particles of concern can include both fine and coarse-fraction particles, although fine particles have been more clearly linked to the most serious health effects. People with heart or lung disease, the elderly, and children are at highest risk from exposure to particles.

In addition to health problems, PM is the major cause of reduced visibility in many parts of the United States. Airborne particles also can impact vegetation and ecosystems and can cause damage to paints and building materials. (See sections on NO2, and SO2.)

5. Carbon Monoxide (CO)

Nature and Sources. Carbon monoxide is a colorless and odorless gas, formed when carbon in fuel is not burned completely. It is a component of motor vehicle exhaust. Non-road vehicles also account for the total CO emissions from transportation sources. High concentrations of CO generally occur in areas with heavy traffic congestion. In some cities, as much as 95 percent of all CO emissions may come from automobile exhaust. Other sources of CO emissions include industrial processes, non-transportation fuel combustion, and natural sources such as wildfires. Peak CO concentrations typically occur during the colder months of the year when CO automotive emissions are greater and nighttime inversion conditions (where air 14 pollutants are trapped near the ground beneath a layer of warm air) are more frequent.

Health Effects. CO enters the bloodstream through the lungs and reduces oxygen delivery to the body's organs and tissues. The health threat from levels of CO sometimes found in the ambient air is most serious for those who suffer from cardiovascular disease such as angina pectoris. At much higher levels of exposure not commonly found in ambient air, CO can be poisonous, and even healthy individuals may be affected. Visual impairment, reduced work capacity, reduced manual dexterity, poor learning ability, and difficulty in performing complex tasks are all associated with exposure to elevated CO levels.

6. Lead (Pb)

Nature and Sources. In the past, automotive sources were the major contributor of lead emissions to the atmosphere. As a result of EPA's regulatory efforts to reduce the content of lead in gasoline, however, the contribution of air emissions of lead from the transportation sector, and particularly the automotive sector, has greatly declined over the past two decades.

Today, industrial processes, primarily metals processing, are the major source of lead emissions to the atmosphere. The highest air concentrations of lead are usually found in the vicinity of smelters and battery manufacturers.

Health and Environmental Effects. Exposure to lead occurs mainly through inhalation of air and ingestion of lead in food, water, soil, or dust. It accumulates in the blood, bones, and soft tissues and can adversely affect the kidneys, liver, nervous system, and other organs. Excessive exposure to lead may cause neurological impairments such as seizures, mental retardation, and behavioral disorders. Even at low doses, lead exposure is associated with damage to the nervous systems of fetuses and young children, resulting in learning deficits and lowered IQ. Recent studies also show that lead may be a factor in high blood pressure and subsequent heart disease. Lead can also be deposited on the leaves of plants, presenting a hazard to grazing animals and humans through ingestion.

Table A1 Summary of Findings from Different Cohort Studies

Study	Mortality Cause (pollutant)	Relative Risk for a 10 units increase (C.I.)	Units
Original SCS[1]	All cause mortality (PM_{10})	1.09 (1.03, 1.15)	$\mu g/m^3$
	All cause mortality ($PM_{2.5}$)	1.13 (1.04, 1.23)	$\mu g/m^3$
	Lung Cancer ($PM_{2.5}$)	1.18 (0.89, 1.57)	$\mu g/m^3$
	Cardio Pulmonary ($PM_{2.5}$)	1.18 (1.06, 1.32)	$\mu g/m^3$
	All cause mortality (Sulfates)	1.33 (1.10, 1.62)	$\mu g/m^3$
Original ACS[2]	All cause mortality ($PM_{2.5}$)	1.07 (1.04, 1.10)	$\mu g/m^3$
	Cardiopulmonary ($PM_{2.5}$)	1.12 (1.07, 1.17)	$\mu g/m^3$
	All cause mortality (Sulfates)	1.07 (1.04, 1.11)	$\mu g/m^3$
	Lung Cancer (Sulfates)	1.17 (1.05, 1.29)	$\mu g/m^3$
	Cardio Pulmonary (Sulfates)	1.12 (1.08, 1.17)	$\mu g/m^3$
Extended AHSMOG	All cause mortality in males (PM_{10})[3]	1.04 (0.99, 1.10)	$\mu g/m^3$
	Nonmalignant Respiratory (PM_{10})[4]	1.06 (0.99, 1.15)	$\mu g/m^3$
	Lung Cancer in males (SO_2)[3]	1.14 (0.65, 1.56)	ppb
	Lung Cancer in females (SO_2)[5]	1.00 (0.78, 1.29)	ppb
	Lung Cancer in females (NO_2)[5]	0.99 (0.94, 1.05)	ppb
Reanalysis SCS (Extended model)[6]	All cause mortality ($PM_{2.5}$)	1.14 (1.05, 1.24)	$\mu g/m^3$
	Cardiopulmonary ($PM_{2.5}$)	1.16 (1.04, 1.30)	$\mu g/m^3$
	Lung Cancer ($PM_{2.5}$)	1.15 (0.86, 1.54)	$\mu g/m^3$
Reanalysis ACS (Extended model)[6]	All cause mortality ($PM_{2.5}$)	1.07 (1.04, 1.10)	$\mu g/m^3$
	Cardiopulmonary ($PM_{2.5}$)	1.11 (1.07, 1.16)	$\mu g/m^3$
	Lung Cancer ($PM_{2.5}$)	1.00 (0.91, 1.11)	$\mu g/m^3$
Extended ACS[7]	All cause mortality ($PM_{2.5}$)	1.06 (1.02, 1.11)	$\mu g/m^3$
	Cardiopulmonary ($PM_{2.5}$)	1.09 (1.03, 1.16)	$\mu g/m^3$
	Lung Cancer ($PM_{2.5}$)	1.14 (1.04, 1.23)	$\mu g/m^3$

1. Mean concentration of each pollutant included individually in the CPHM
2. For $PM_{2.5}$ mean concentration, for sulfates, median concentration taken as air pollution measure. Adjusted for age, sex, race, cigarette smoking, exposure to cigarette smoke, BMI, drinks per day of alcohol, education, occupational exposure.
3. Including age, years of education, pack years of past smoking, history of high blood pressure, years lived with a smoker and total exercise level.
4. Including past years of pack smoking, BMI, total physical exercise, age within age strata.
5. Adjusting for age, years of education, pack years of past smoking, years lived with a smoker, years worked with a smoker, occupational exposure for more than 10 years, BMI.
6. Controlling for tobacco consumption, education level, occupational exposure, BMI, marital status, alcohol consumption and sex.
7. Estimated and adjusted based on the baseline random effects CPHM, controlling for age, sex, race, smoking, history of high blood pressure, years lived with a smoker and total exercise level.

Table A2 Summary of Findings from Different Time Series Studies

Regression coefficients for effect of PM					
Study	**Location**	**Health effect**	**Model characteristics**	**% change per 10 $\mu g/m^3$**	**Statistic for significance**
NMMAPS	Houston	Total mortality	PM_{10} Unadjusted for other pollutants	0.19 (-0.47, 0.84)	Na
			PM_{10} Adjusted for O_3, NO_2	0.01 (-0.67, 0.69)	Na
			PM_{10} Adjusted for O_3, SO_2	0.13 (-0.54, 0.79)	Na
			PM_{10} Adjusted for O_3, CO	0.04 (-0.66, 0.75)	Na
		Cardiorespiratory	PM_{10} Unadjusted for other pollutants	0.55 (-0.35, 1.45)	Na
		Other disease	PM_{10} Unadjusted for other pollutants	-0.64 (-2.51, 1.23)	Na
	Phoenix	Total mortality	PM_{10} Unadjusted for other pollutants	0.65 (-0.42, 1.72)	Na
			PM_{10} Adjusted for O_3, NO_2	1.95 (-0.39, 4.29)	Na
			PM_{10} Adjusted for O_3, SO_2	1.72 (0.14, 3.31)	Na
			PM_{10} Adjusted for O_3, CO	-0.07 (-1.42, 1.29)	Na
		Cardiorespiratory	PM_{10} Unadjusted for other pollutants	1.12 (-0.22, 2.63)	Na
		Other disease	PM_{10} Unadjusted for other pollutants	0.16 (-1.40, 1.72)	Na
	Atlanta	Total mortality	PM_{10} Unadjusted for other pollutants	0.05 (-1.60, 1.70)	Na
			PM_{10} Adjusted for O_3, NO_2	-1.12 (-3.36, 1.11)	Na
			PM_{10} Adjusted for O_3, SO_2	-0.95 (-3.15, 1.24)	Na
			PM_{10} Adjusted for O_3, CO	-0.70 (-3.17, 1.17)	Na
		Cardiorespiratory	PM_{10} Unadjusted for other pollutants	-0.32 (-2.69, 2.04)	Na
		Other disease	PM_{10} Unadjusted for other pollutants	0.44 (-1.86, 2.74)	Na
	National	Total mortality (using a non informative prior)	$PM_{10} + O_3$	0.27 (0.12, 0.44)	Na
			$PM_{10} + O_3 + NO_2$	0.21 (-0.01, 0.44)	Na
			$PM_{10} + O_3 + SO_2$	0.21 (0.02, 0.42)	Na
			$PM_{10} + O_3 + CO$	0.24 (0.05, 0.43)	Na
APHEA	All	Total mortality aggregated using fixed effects	Black smoke coefficient using old sinusoidal terms to control for seasonality (also controlling for temperature, relative humidity, day of the week, epidemic periods and holidays)	0.26 (0.18, 0.34)	*p* value = <0.0001
	Western			0.57 (0.42, 0.73)	*p* value = <0.001
	Central-eastern			0.12 (0.02, 0.22)	*p* value = 0.25
	All	Total mortality aggregated using fixed effects	Black smoke coefficient using GAM methodology (also controlling for temperature, relative humidity, day of the week, epidemic periods and holidays)	0.44 (0.36, 0.51)	*p* value = <0.0001
	Western			0.61 (0.47, 0.77)	*p* value = <0.001
	Central-eastern			0.44 (0.28, 0.45)	*p* value = 0.04

Table A2 Summary of Findings from Different Time Series Studies (continue)

Regression coefficients for effect of PM					
Study	**Location**	**Health effect**	**Model characteristics**	**% change per 10 μg/m^3**	**Statistic for significance**
Borja Aburto, et.al (1998)	Southwestern MCMA	Total mortality	Three pollutant model ($PM_{2.5}$, O_3 and NO_2) controlling for temperature on the 3 days before death and smoothed periodic cycles	1.68 (0.23, 3.14)	t = 2.27
		Respiratory		1.68 (-2.88, 6.42)	t = 0.72
		Cardiiovascular		3.42 (0.67, 6.18)	t = 2.43
Castillejos et.al (2000)	Southwestern MCMA	Total mortality	PM_{10} coefficient controlled for ozone, NO_2, temperature on 3 days before death and non parametrically smoothed cycles	2.47 (1.14, 3.81)	t = 3.63
		Respiratory		6.40 (2.16, 10.64)	t = 2.96
		Cardiiovascular		1.96 (-0.56, 448)	t = 1.53

A1.2 Commonly Used Definitions in Epidemiology.

Comments by Roger Cooke

1. Risk

$$Risk = R = \frac{\#\ \text{of new cases}}{\#\ \text{persons exposed}} \quad \text{for a period of time}$$

Risk is dimensionless, intended to measure probability of occurrence between t_0 and t_1.

2. Incidence rate

$$Incidence\ rate = I = \frac{\#\ \text{new cases}}{\#\ \text{persons exposed} \times \text{exposure time}}$$

$$R = I \times T$$

3. Mortality rate

Mortality rate = incidence rate of death

4. Prevalence proportion

$$P(t) = prevalence\ \ proportion(t) = \frac{\#\ \text{people with effect (t)}}{\#\ \text{people in the population (t)}}$$

At equilibrium $\frac{P}{(1-P)} = I \times D_{av}$; D_{av} = average duration of disease.

P_w = Probability per unit time of recovery, $D_{av} = \frac{1}{P_w}$

P_d = Probability per unit time of getting sick = I

$N_d(t)$ = # of people sick at t

$$N_d(t+1) = N_d(t)(1-P_w) + (N - N_d(t))P_d$$

$$N_d(t+1) = N_d(t) \Leftrightarrow \frac{N_d(t)}{N - N_d(t)} = \frac{P_d}{P_w}$$

5. Risk Ratio and Rate Ratio

R_x = Risk of exposed group, R_u = Risk of unexposed group

$$RR = Risk\ Ratio = \frac{R_x}{R_u} \cong \frac{I_x T}{I_u T} = \frac{I_x}{I_u} = Rate\ Ratio$$

Ratios of probabilities of occurrence in a time interval ~ ratio of failure rates if interval is small

6. Relative Risk

Relative Risk = "Instantaneous Risk Ratios" = Rate Ratio.

7. Attributable Fraction

$$Atributable\ Fraction = \frac{R_x - R_u}{R_x} = \frac{RR-1}{RR}$$

8. Risk Ratio for Cox

$$RR_{Cox}(X_1, X_0) = \frac{P(\text{death in } t+\delta \mid \text{alive at } t)(X_1, \mathrm{Y})}{P(\text{death in } t+\delta \mid \text{alive at } t)(X_0, \mathrm{Y})} = \frac{\gamma(t)e^{bX_1 + \mathrm{Y}\bullet\mathrm{C}}}{\gamma(t)e^{bX_0 + \mathrm{Y}\bullet\mathrm{C}}} = e^{b(X_1 - X_0)}$$

Where **Y** is vector of other covariates and **C** is vector of coefficients for other covariates, $\gamma(t)$ is baseline hazard.

*Though measurable in principle, in fact the coefficients are found by maximizing the partial likelihood. It is interpreted as a risk ratio for two populations having same covariates **Y**, differing only in X. However under the assumptions of the Cox model, this ratio does not depend on the particular values of **Y**.*

9. Compare risk in populations P_1 and P_0:

Covariates X,**Y**, in P_1 $X = x_1$, in P_0 $X = x_0$; **y**(i) = covariate value for individual i.

$$R_1 = \frac{\#\text{ of new cases in } P_1}{\#\text{ persons exposed in } P_1} \text{ in } (t, t+\delta);$$

$$R_1 = \sum_{i \in P_1} \frac{P(\text{event in } t+\delta \mid \text{no event at } t)}{\#\text{ persons exposed in } P_1} = \frac{\gamma_1(t) e^{bx_1} \sum_{i \in P_1} e^{y(i)\bullet C}}{\#\text{ persons exposed in } P_1}$$

$$RR = \frac{R_1}{R_0} = \frac{\gamma_1(t) e^{bx_1} \sum_{i \in P_1} e^{y(i)\bullet C} \Big/ \#\text{ persons exposed in } P_1}{\gamma_0(t) e^{bx_0} \sum_{i \in P_0} e^{y(i)\bullet C} \Big/ \#\text{ persons exposed in } P_0} = RR_{Cox} \times \frac{\gamma_1(t) \sum_{i \in P_1} e^{y(i)\bullet C} \Big/ \#\text{ persons exposed in } P_1}{\gamma_0(t) \sum_{i \in P_0} e^{y(i)\bullet C} \Big/ \#\text{ persons exposed in } P_0}$$

*$RR = RR_{Cox}$ if the distribution of **y** and $\gamma(t)$ is same in both populations. This is true if we compare the effect of "exposure" in ONE population. If $x_1 = x_0$ but $P_1 \neq P_0$, then the ratio of risks depend on P_1, P_2. But risk ratios R_1/R_0 in each population are the same.*

Chapter 2. Long Term Effects of Air Pollution in Mortality.

The first model that will be discussed is the Cox Proportional Hazards Model (CPHM) that is used in survival analysis for investigating long term effects of air pollution (particulate matter specifically) on health. To gather the data needed to fit this model requires considerable economic, human and time resources (see chapter one) that unfortunately are not available for the present setting. However some effort in assessing model adequacy will be encountered along the way.

2.1. Cox Proportional Hazards Model.

In 1972 Cox proposed a regression model for survival time in which the "partial likelihood" would be studied instead of the full likelihood function. This model assumes a hazard function with the following form:

$$h(t,z,\beta) = \Lambda_0(t) r(z,\beta) \tag{1}$$

In equation (1), the hazard is a function of the base line hazard Λ_0 that depends on survival time and another function r that characterizes how the hazard changes as a function of covariates; β is the vector of covariates to be estimated from data. In practice the familiar choice for r is $e^{Z\beta}$ where z and β are the vectors of covariates and parameters respectively. To estimate parameters form this model the ln of the partial likelihood function (*PL* below) is maximized. For covariates that are constant in time, all that maters for the Cox model is the rank of events.

$$PL\ (\beta) = \prod_{i=1}^{n} \left[\frac{e^{Z_i \beta}}{\sum_{j \in R(t_i)} e^{Z_j \beta}} \right]^{c_i} \tag{2}$$

In the above expression c is an indicator that is one when an event is observed and zero otherwise, n is the total population and the summation in the denominator runs over all individuals in the risk set[1] at time t_i. Usually confidence intervals are obtained with a normal approximation in the usual way by taking the second derivative of the log likelihood function, computing the information matrix and estimating the standard deviation for the kth covariate coefficient estimate as the square root of the kth element in the diagonal of the inverse information matrix. For more details the reader is referred to (Allison, P.D., 2003) and (Homser, D.W. and Lemeshow, S.).

The intuitive explanation for the model above is that given that the first death of the population occurs at t_1 the probability that it happened to individual 1 is the hazard of this individual divided by the sum of the hazards of all other individuals in the population (3).

[1] The risk set is the set of survival times that have not been observed or censored at t_i

$$\frac{h_1}{\sum_{j=1}^{n} h_j} = \frac{e^{Z_1\beta}}{\sum_{j=1}^{n} e^{Z_j\beta}} \tag{3}$$

After individual 1 is removed from the population, the same reasoning applies to the surviving population; given that a second event occurred at time t_2, the probability that it happened to the second individual is as (4) and so on until n.

$$\frac{h_2}{\sum_{j=2}^{n} h_j} = \frac{e^{Z_2\beta}}{\sum_{j=2}^{n} e^{Z_j\beta}} \tag{4}$$

In many important studies, model adequacy is not examined, and only estimates of the coefficients for some covariates are given with Wald confidence intervals for example (Pope, et al., 2002) and (Dockery, et al., 1993). The coefficients are used then for different risk assessments for regulation purposes. Next the issue of model adequacy will be discussed.

2.2. Model Adequacy.

For simplicity, consider a model with time invariant covariates *X*, *Y*, *Z* and parameters *A*, *B*, *C* the hazard rate has then the following form:

$$h(X, Y, Z) = \Lambda_0(t) e^{XA+YB+ZC} \tag{5}$$

If this hazard rate holds, then for an individual with covariate values (*x*, *y*, *z*) the survivor function is:

$$e^{-h(x,y,z)} \tag{6}$$

With (*x*, *y*, *z*) fixed, the time to death *T* random and with a constant baseline hazard scaled to one, then (6) considered as a function of the random variable *T* is uniformly distributed on [0,1], that is:

$$T \sim \frac{-\ln(U)}{h} \tag{7}$$

Where *U* is a uniform random variable on [0,1]. This holds for each individual in the population, so if the values form (8) are ordered and plot against their number the points should lay along the diagonal if the proportional hazards model is true with coefficients *A*, *B*, *C* and constant baseline hazard equal to one.

$$e^{-t_i (e^{x_i A + y_i B + z_i C})} \tag{8}$$

However if the baseline hazard is also estimated from the data, (9) should plot as uniform. In that equation "^" denotes the estimators that naturally do not correspond to the values that generated the data.

$$e^{-\Lambda_0(t_i)(e^{x_i\hat{A}+y_i\hat{B}+z_i\hat{C}})} \tag{9}$$

The next steps illustrate a numerical experiment that will help the reader understand the issues relative to model adequacy. All calculations assume no ties and no censored observations.

1. Choose coefficients (A, B, C), choose a constant baseline hazard scaled to one, and choose a distribution for (X, Y, Z)
2. Sample independently 100 values for (X, Y, Z) an 100 values for U; and compute one hundred t_i using (7)
3. Find estimators for (A, B, C) and confidence bounds as described in the previous section and estimate the baseline hazard.

Model (5) will be denoted as h_{XYZ}. In practice, information is available for some of the covariates, but many others may not be represented in the model. For instance in air pollution studies, variables that measure disease prevalence, stress, medical care, home environment, genetic disposition are usually not included in the model and these could also have an influence in the event together with all other variables which are actually included in the model. To study the effects of model incompleteness a model using covariates X and Y and another one using only covariate X (denoted by h_{XY} and h_X respectively) will estimate the coefficient A and these will be compared with the "true" model h_{XYZ}. This procedure is repeated 100 times with the same values for (A, B, C) and (X, Y, Z) sampled independently from a uniform distribution in [0,3]

Figure 2.1 One hundred Ordered Estimates of A for h_{XYZ}, h_{XY} and h_X X,Y,Z~U[0.3], (A,B,C) = (1.006, 1, 1).

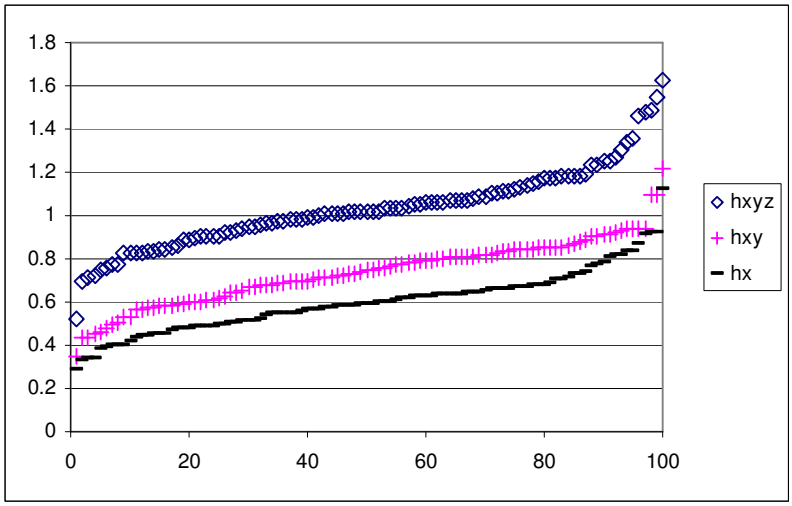
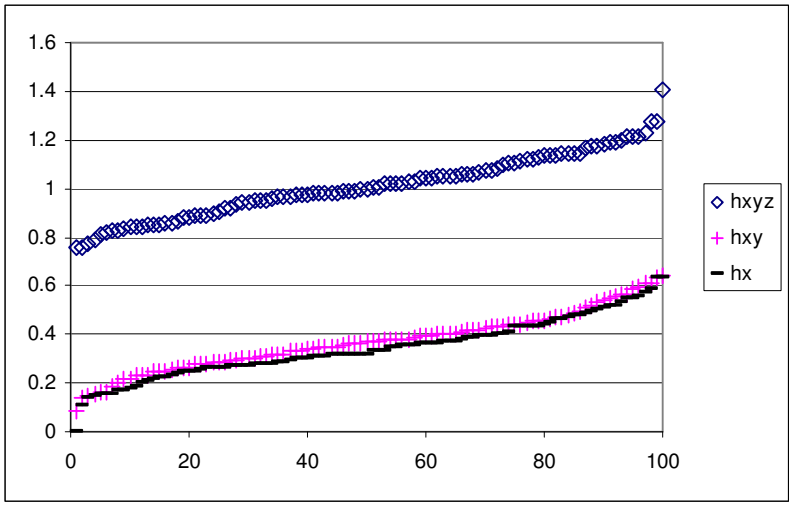


Figure 4.1 shows that the two incomplete models tend to under estimate the coefficient A. This tendency becomes more pronounced in figure 4.2 where Z takes its maximum value, 3.

Figure 2.2 One hundred Ordered Estimates of A for h_{XYZ}, h_{XY} and h_X
***X,Y,Z*~U[0, 3], (*A,B,C*) = (1.006, 1, 3).**



The statement that equation (9) should plot as uniform is checked below in figure 2.3; following (Kalbfleisch and Prentice, 2002) for a model without ties, the estimator of the cumulative hazard function is:

$$\hat{\Lambda}_0(t_i) = \sum_{j \le i} (1 - \hat{\alpha}_j) \tag{10}$$

$$\hat{\alpha}_i = \left\{ 1 - \frac{e^{Z_i \hat{\beta}}}{\sum_{j \ge i} e^{Z_j \hat{\beta}}} \right\}^{e^{-Z_i \hat{\beta}}} \tag{11}$$

The estimator above is obtained in a similar manner as the Nelson-Aalen and Kaplan-Meier estimators for the hazard function. For the simulation presented in this chapter $\hat{\beta} = (\hat{A}, \hat{B}, \hat{C})$ is the vector of estimated coefficients and $Z_i = (X, Y, Z) \sim U[0, 3]$ is the vector of covariates. Figure 2.3 show that when the base line hazard function is estimated, the three models, even the two incomplete models that clearly underestimate the coefficient A, lay along the diagonal, however when the additional restriction that $\hat{\Lambda}_0 \equiv 1$ is imposed, then uniformity is lost for the two incomplete models as shown in figure 2.4.

Interval estimators are very important in any statistical analysis. If 95% confidence bands for the estimator of A are computed as described in the previous section, i e. assuming that the Wald statistic that is the estimated coefficient divided by the estimator of its standard deviation is asymptotically standard normal, then the "true" model h_{XYZ} would miss the true value only 1 time out of the 100 repetitions. Figure 2.5 shows the confidence bounds with the model h_{XY} again computed as described in the previous section. The experiment followed

through this chapter shows that the interval estimators do not capture the true value of the coefficient at all in 100 repetitions and the picture does not improve for the model h_X.

Figure 2.3 One hundred Ordered Values of (9) for h_{XYZ}, h_{XY} and h_X
***X*,*Y*,*Z*~U[0, 3], (*A*,*B*,*C*) = (1.006, 1, 3).**

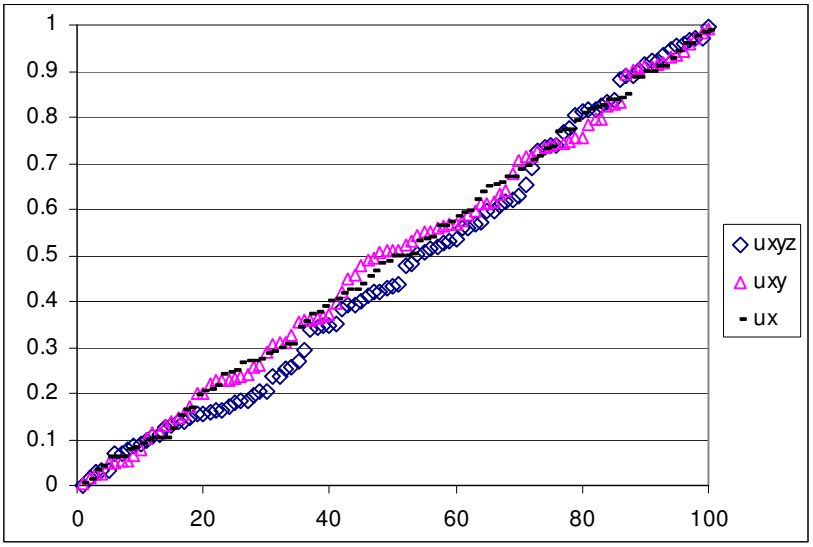
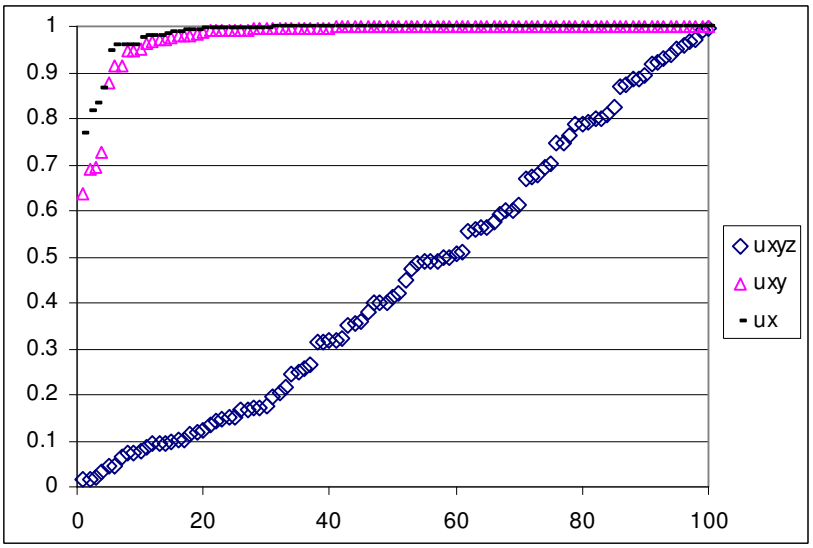


Figure 2.4 One hundred Ordered Values of (9) for h_{XYZ}, h_{XY} and h_X
***X*,*Y*,*Z*~U[0, 3], (*A*,*B*,*C*) = (1.006, 1, 3), $\hat{\Lambda}_0 \equiv 1$**



The models h_{XY} and h_X are clearly incorrect and misestimate the covariate A. Relative risk coefficients based on incomplete models would be biased. The conclusion is that the lack of fit in the incomplete models is compensated by the estimated baseline hazard function. The last observation suggests that lack of fit in the covariates might be detected by comparing the estimated baseline hazard function with the population cumulative hazard function.

The Nelson-Aalen estimator (12) is used for the population cumulative hazard function, where d_i are the number of events at t_i an n_i is the population at risk at the same time. Figures 2.6 and 2.7 show that the cumulative baseline hazard function for models h_{XY} and h_X have moved closer to the population cumulative hazard reflecting the heavier load for the missing covariate Z.

$$\Lambda(t_i) = \sum_{j \le i} d_j / n_j \tag{12}$$

Figure 2.5 One Hundred 95% Confidence Bands for *A* with Model h_{XY} of Figure 2.3

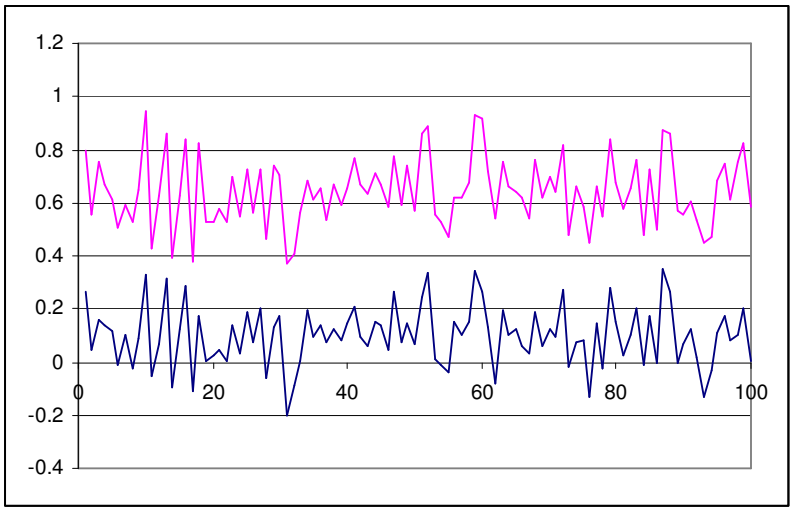


Figure 2.6 Cumulative Population and Base Line Hazard Functions for h_{XYZ}, h_{XY}, h_X, (*X*, *Y*, *Z*)~U[0,3], (*A*, *B*, *C*)=(1.006, 1, 1)

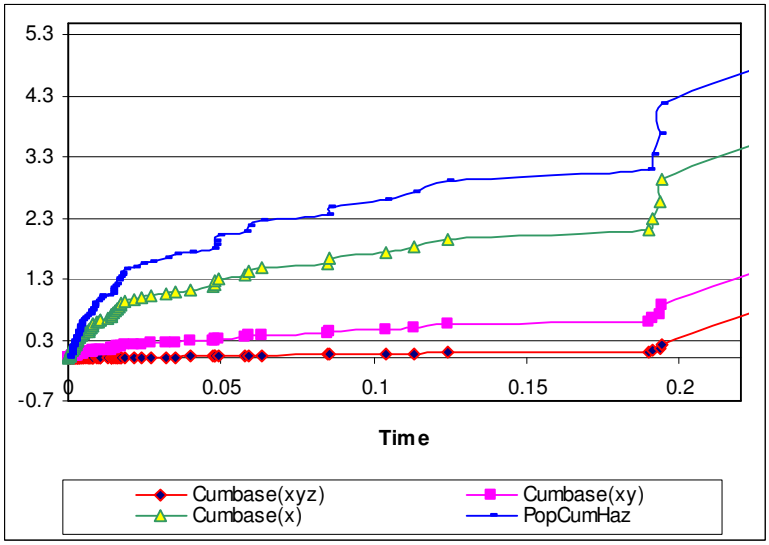
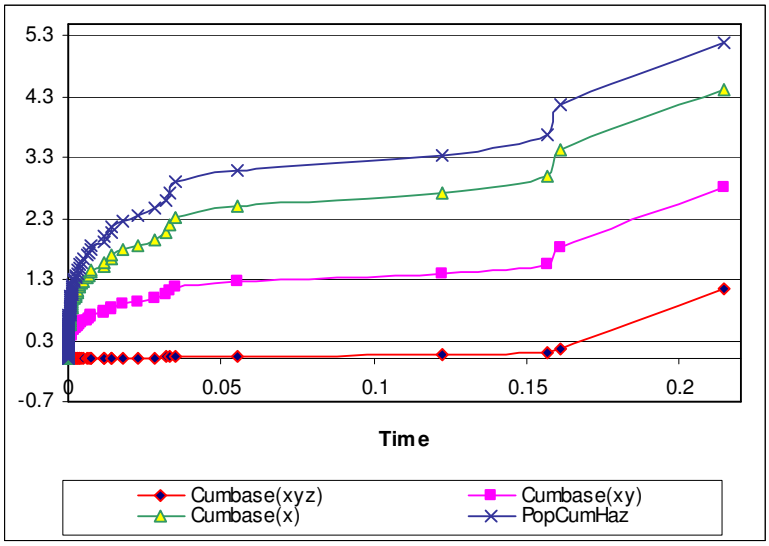


Figure 2.7 Cumulative Population and Base Line Hazard Functions for h_{XYZ}, h_{XY}, h_X, (X, Y, Z)~U[0,3], (A, B, C)=(1.006, 1, 3)



A possibility to assess model adequacy would be to test the null hypothesis that the cumulative hazard function is equal to the population cumulative hazard function. If the null hypothesis cannot be rejected, then using the Cox model would not be indicated. With the same notation as before, the variance estimator of the Nelson-Aalen cumulative hazard estimator (12) is computed as in (13).

$$\hat{V}(t_i) = \sum_{j \le i} d_j (n_j - d_j) / n_j^3 \qquad (13)$$

Figures 2.8 and 2.9 show the same picture as 2.6 and 2.7 respectively plus the "2-σ" bands for the population cumulative hazard function. Sigma is of course, the square root of the variance estimator (13); the reader may see in figure 2.8 that even when A is underestimated in the case where $C = 1$ the test proposed is unable to detect the lack of fit of the models h_{XY} and h_X; however in the case where $C = 3$, the model h_X would fail a test for model adequacy as the cumulative baseline hazard is not easily distinguished from the population cumulative hazard function as observed in figure 2.9. If the assumptions from appendix A1.2 were true and the $RR = RR_{Cox}$ then a change of k units in a continuous scale covariate would translate in a change of $k\beta$ in the log hazard. For the case described above this change would be of k if $C = 1$ and 3 times larger when $C = 3$.

2.3. Competing Risks.

Another way to look at survival data, is the competing risks approach, the idea is that the event may be one of distinct "failure" types. Right censoring is a form of competing risk. In the competing risks approach the data is modeled as a sequence of identical independent distributed pairs (T_i, δ_i), $i = 1, 2, \ldots$. Each T is the minimum of two or more variables,

corresponding to the competing risks. It will be assume that there are only two competing risks, described by two random variables D and C such that $T = \min(D, C)$. D will be time of death which is of primary interest, while C is a censoring time corresponding to termination of observation by other causes.

Figure 2.8 Cumulative Population and Base Line Hazard Functions for h_{XYZ}, h_{XY}, h_X, (X, Y, Z)~U[0,3], (A, B, C)=(1.006, 1, 1)

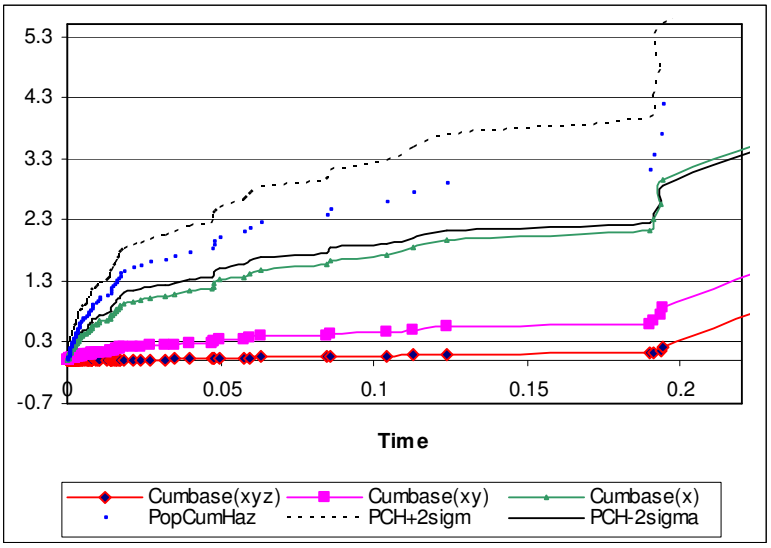
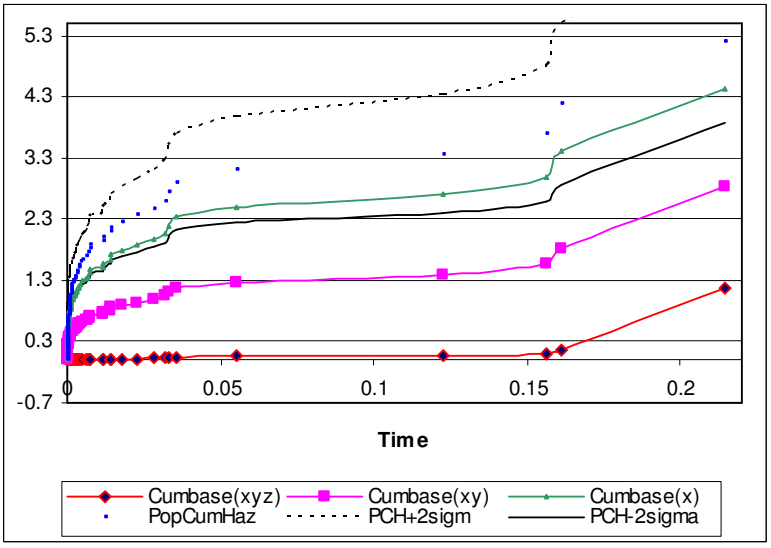


Figure 2.9 Cumulative Population and Base Line Hazard Functions for h_{XYZ}, h_{XY}, h_X, (X, Y, Z)~U[0,3], (A, B, C)=(1.006, 1, 3)



In addition to the time T one observes the indicator variable $\delta = I\,(D < C)$ which describes the cause of the termination of observation. For simplicity we assume that $P(D = C) = 0$. It is

known (Tsiatis, 1975) that from observation of (T, δ) only the subsurvivor functions of C and D may be identified:

$$S_D^*(t) = P(D > t, D < C) = P(T > t, \delta = 1) \tag{14}$$

$$S_C^*(t) = P(C > t, C < D) = P(T > t, \delta = 0) \tag{15}$$

but not in general the true survivor functions of D and C, $S_D(t)$ and $S_C(t)$. Note that $S_D^*(t)$ depends on C, though this fact is suppressed in the notation. Note also that $S_D^*(0)=P(D<C)=P(\delta=1)$ and $S_C^*(0)=P(C<D)=P(\delta=0)$, so that $S_D^*(0)+S_C^*(0)=1$.

The conditional subsurvivor functions are defined as the survivor functions conditioned on the occurrence of the corresponding type of event. Assuming continuity of $S_D^*(t)$ and $S_C^*(t)$ at zero, these functions are given by

$$CS_D^*(t) = P(D > t \mid D < C) = P(T > t \mid \delta = 1) = S_D^*(t) / S_D^*(0) \tag{16}$$

$$CS_C^*(t) = P(C > t \mid C < D) = P(T > t \mid \delta = 0) = S_C^*(t) / S_C^*(0) \tag{17}$$

Closely related to the notion of subsurvivor functions is the probability of censoring beyond time t,

$$\Phi(t) = P(C < D \mid T > t) = P(\delta = 0 \mid T > t) = \frac{S_C^*(t)}{S_D^*(t) + S_C^*(t)} \tag{18}$$

As mentioned above, without any additional assumptions on the joint distribution of D and C, it is impossible to identify the marginal survivor functions $S_D(t)$ and $S_C(t)$. However, by making extra assumptions, one may restrict to a class of models in which the survivor functions are identifiable. A classical result on competing risks (Tsi, et al.,_____) states that, assuming independence of D and C, we can determine uniquely the survivor functions of D and C from the joint distribution of (T, δ), where at most one of the survivor functions has an atom at infinity. In this case the survivor functions of D and C are said to be identifiable from the censored data (T, δ). Hence, an independent model is always consistent with data.

If the censoring is assumed to be independent then the survivor function for T, the minimum of D and C, can be written as

$$S_T(t) = S_D(t)\, S_C(t) \tag{19}$$

If it is assumed that D obeys a proportional hazard model, and that the censoring is independent, then the coefficients may be estimated by maximizing the partial likelihood function adapted to account for censoring:

$$PL\ (\beta) = \prod_{i \in D_N} \left[\frac{e^{Z_i \beta}}{\sum_{j \in R(t_i)} e^{Z_j \beta}} \right] \tag{20}$$

where D_N is the subset of observed times $t_1,\ldots, t_N$ at which death is observed to occur, and j runs over all times corresponding to death or censoring.

If the estimated coefficients are substituted in the survivor function (19), and use the familiar Kaplan Meier estimator for S_C, then the ideas of the previous section may be applied to assess model adequacy. In appendix A2.1 some results from the theory of competing risks are summarized. Those results lead to heuristics for model selection.

The probability $\Phi(t)$ of censoring after time t, yields a diagnostic for model selection, together with the conditional subsurvivor functions $CS_D^*(t)$ and $CS_C^*(t)$. Statistical tests are developed in (Bunea, et al. 2002). The following statements, which follow from the results shown in appendix A1, may guide in model selection.

Table 2.1 Heuristics for Model Selection in Competing Risks.

Model	Conditional Subsurvivor functions	Probability of censoring beyond t
Independent Exponential Competing Risks	$CS_D^*(t) = CS_C^*(t)$ and exponential	$\Phi(t)$ is constant.
Random Signs Censoring	$CS_D^*(t) > CS_C^*(t)$ for all $t > 0$.	$\Phi(0) > \Phi(t)$ for all $t > 0$.
Conditional Independence Model(V, U and W exponential)	$CS_D^*(t) = CS_C^*(t)$	$\Phi(t)$ is constant.
Mixture of Exponentials Model	$CS_D^*(t) \leq CS_C^*(t)$ for all $t > 0$	$\Phi(0) < \Phi(t)$ for all $t > 0$.

Next, an example with a subset of the Mayo Clinic Lung Cancer Data (Loprinzi, et al., 1994) will be presented. This data is available in several software applications (S-Plus, SAS) and in the website http://www.mayo.edu/hsr/people/therneau/book/data/lung.html.

2.4. Example: The Mayo Clinic Lung Cancer Data.

The data consists of 228 observations in total, from which 165 are observed times of death and 63 correspond to censoring times. The censoring is assumed to be independent. 8 covariates are used to construct a proportional hazards model: Enrolling institution, age, sex, Eastern Cooperative Oncology Group (ECOG) performance score (as judged by physician), Karnofsky performance score (as judged by physician), Karnofsky performance score (as judged by the patient), daily calories consumed at meals and weight loss in the last 30 days. The ECOG and Karnofsky scores measure performance status of the patient and have been used as prognostic factors for patients in clinical trials.

First, the partial likelihood (20) is maximized with respect to β to obtain the vector of estimated coefficients. Following (Kalbfleish & Prentice, 2002) the population cumulative hazard is estimated as:

$$\hat{\Lambda}_{pop}(t) = -\ln(\hat{S}(t)) \tag{21}$$

Where $\hat{S}$ is the Kaplan-Meier estimator of the survivor function. The baseline cumulative hazard is computed in the manner explained previously (see equations (10)-(11)). Figure 2.10 shows the Cox cumulative baseline hazard function and the population cumulative hazard function. The baseline hazard is nearly linear up to 883 days, indicating a nearly constant baseline hazard rate. The last observations are censored; the fact that the baseline hazard is estimated only at times of death explains the flat shape after 883 days.

Figure 2.10 Cumulative Baseline Hazard (Cox Model) and Population Cumulative Hazard for Mayo Clinic Lung Cancer Data

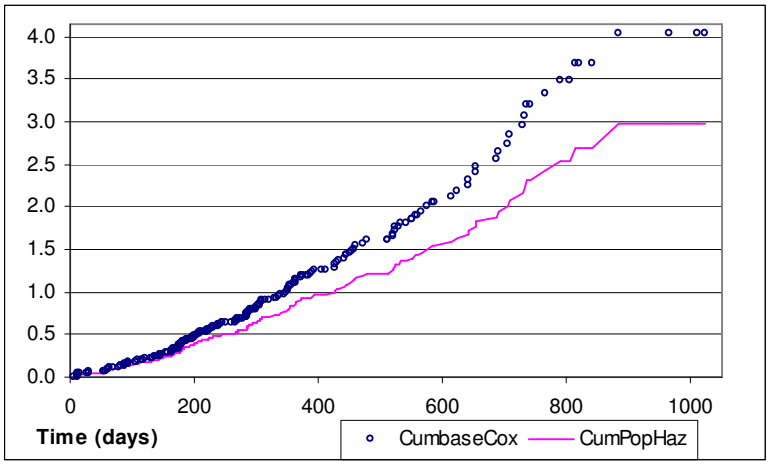


Figure 2.11 Cumulative Baseline Hazard (Cox Model) and Population Cumulative Hazard for Mayo Clinic Lung Cancer Data with 2 Sigma Confidence Bounds.

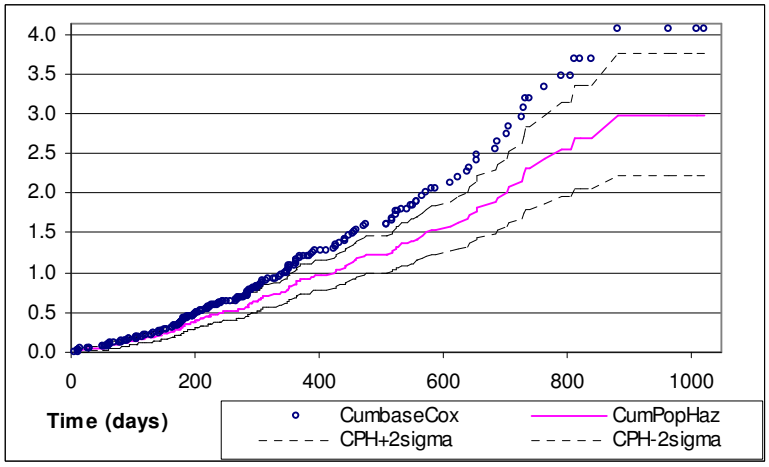
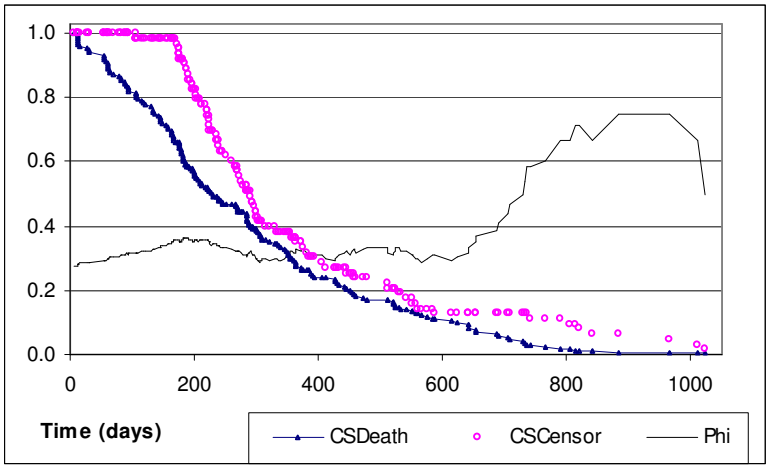


Figure 2.11 adds the 2-sigma bounds computed as in (13). The Cox baseline hazard function is close to the upper bound for the cumulative population hazard. This means that the Cox model is barely able to distinguish the cumulative baseline hazard and population cumulative hazard functions.

Figure 2.12 presents the empirical conditional subsurvivor functions for death and censoring and shows the probability of censoring after time t i. e. the function $\Phi(t)$ They are computed as shown in the appendix A2.2. The reader may see that the conditional subsurvivor function of censoring dominates that of death. The $\Phi(t)$ function is increasing up to the last observed event at day 883 where the conditional subsurvivor function for death becomes constant and $\Phi(t)$ decreases. From table 2.1 it could be inferred that this pattern correspond to a mixture of exponentials model (see appendix A2.1).

Figure 2.12 Cumulative Baseline Hazard (Cox Model) and Population Cumulative Hazard for Mayo Clinic Lung Cancer Data with 2 Sigma Confidence Bounds.



Since the test proposed in the previous section for this specific data set gives the result that the population cumulative hazard and the Cox baseline cumulative hazard are not easily distinguishable; and the censoring mechanism appears to come from an independent exponential censoring a mixture of exponential life variables, the suggestion would be to add more covariates to the Cox model in order to take load from the cumulative baseline hazard and make the coefficients more easily interpretable.

2.5. Final Remarks.

As stated before, tests like the one performed in the last part of this chapter to the Mayo Clinic Lung Cancer Data are not common in air pollution health effects studies. While trying to control the model for certain covariates, the power of the whole model after controlling is evaluated with likelihood related tests, however the methods presented in this chapter could also serve as a guide for model performance evaluation. According to the results presented above the Cox proportional hazard model with constant covariates entails a mixed exponential life distribution.

In the next chapter models that investigate short term effects of air pollution will be discussed and examples will be drawn with data from the central Mexico capital city of the State of Mexico, Toluca.

APPENDIX 2

A2.1. Independent Competing Risks Models.

Independent Exponential Competing Risks

A model in which *D* and *C* are independent is always consistent with the data, but an independent *exponential* model is not in general consistent with the data. One can derive a sharp criterion for independence and exponentiality in terms of the subsurvivor functions (Cooke, 1996):

Theorem. Let *D* and *C* be independent life variables. Then any two of the following conditions imply the others:

$$S_D(t) = e^{-\lambda t} \tag{8}$$

$$S_C(t) = e^{-\gamma t} \tag{9}$$

$$S_D^*(t) = \frac{\lambda}{\lambda+\gamma} e^{-(\lambda+\gamma)t} \tag{10}$$

$$S_C^*(t) = \frac{\gamma}{\lambda+\gamma} e^{-(\lambda+\gamma)t} \tag{11}$$

Thus if *D* and *C* are independent exponential life variables with failure rates λ and γ, then the conditional subsurvivor functions of *D* and *C* are equal and correspond to exponential distributions with failure rate $\lambda+\gamma$.
Moreover, the probability of censoring beyond time *t* is constant. Thus

$$CS_D^*(t) = CS_C^*(t) = e^{-(\lambda+\gamma)t} \tag{12}$$

$$\Phi(t) = \frac{\gamma}{\lambda+\gamma} \tag{13}$$

Random Signs Censoring

Suppose that the event that the time of death of a subject is censored is independent of the age *D* at which the subject would die, but given that the subject's time of death is censored, the time at which it is censored may depend on *D*. This situation is captured in the following definition:

Let *D* and *C* be life variables with $C = D - W\delta$, where $0 < W < D$ is a random variable and δ is a random variable taking values [1,-1], with *D* and δ independent. The variable $T \equiv [\min(D,C), I(D<C)]$ is called a random sign censoring of *D* by *C*. Note that in this case

$$S_D^*(t) = P\{D > t, \delta = -1\} = P\{D > t\}P\{\delta = -1\} = S_D(t)P\{C > D\} = S_D(t)S_D^*(0) \tag{14}$$

Hence $S_D(t)=CS_D^*(t)$ and it follows that the distribution of *D* is identifiable under random signs censoring.

A joint distribution of (D, C) which satisfies the random signs requirement, exists if and only if $CS_D^*(t) > CS_C^*(t)$ for all $t>0$. In this case the probability of censoring beyond time t, $\Phi(t)$, is maximum at the origin.

Conditional Independence Model

Another model from which there is identifiability of marginal distributions is the conditional independence model introduced by Hokstad. This model considers the competing risk variables D and C to be sharing a common quantity, V, and to be independent given V. More precisely, the assumption is that $D = V + W$, $C = V + U$, where V, U and W are mutually independent. Hokstad and Jensen derived explicit expressions for the case when V, U and W are exponentially distributed:

Theorem. Let V, U, W be independent with $S_V(t) = e^{-\lambda t}$, $S_U(t) = e^{-\gamma t}$ and $S_W(t) = e^{-\theta t}$. Then

$$S_D^*(t) = \frac{\lambda\theta e^{-(\lambda+\theta)t}}{(\gamma+\theta)(\lambda-\gamma-\theta)} - \frac{\theta e^{-(\lambda)t}}{(\lambda-\gamma-\theta)} \tag{15}$$

$$S_D^*(t) = \frac{\lambda\gamma e^{-(\lambda+\theta)t}}{(\gamma+\theta)(\lambda-\gamma-\theta)} - \frac{\gamma e^{-(\lambda)t}}{(\lambda-\gamma-\theta)} \tag{16}$$

$$CS_D^*(t) = CS_C^*(t) = S_D^*(t) + S_C^*(t) \tag{17}$$

$$\Phi(t) = \frac{\lambda}{\lambda+\theta} \tag{18}$$

Moreover, if V has an arbitrary distribution such that $P(V \geq 0) = 1$, and V is independent of U and W then still $CS_D^*(t) = CS_C^*(t)$.

Thus, as in the case of independent exponential competing risks we have equal conditional subsurvivor functions, and the probability of censoring beyond time t, $\Phi(t)$, is constant. However, the conditional subsurvivor functions need not be exponential. Nothing is known about their general form.

Mixture of Exponentials Model

Suppose that $S_D(t)$ is a mixture of two exponential distributions with parameters λ_1, λ_2 and mixing coefficient p, and that the censoring survivor distribution $S_C(t)$ is exponential with parameter λ_y:

$$S_D(t) = pe^{-\lambda_1 t} + (1-p)e^{-\lambda_2 t} \tag{18}$$

$$S_C(t) = pe^{-\lambda_y t} \tag{19}$$

The properties of the corresponding competing risk model is given by (Bunea, et al., 2003)

Theorem. Let D and C be independent life variables with the above distributions. Then,

$$S_D^*(t) = p\frac{\lambda_1}{\lambda_y+\lambda_1}e^{-(\lambda_1+\lambda_y)t} + (1-p)\frac{\lambda_2}{\lambda_y+\lambda_2}e^{-\lambda_2 t} \tag{20}$$

$$S_C^*(t) = p\frac{\lambda_y}{\lambda_y+\lambda_1}e^{-(\lambda_1+\lambda_y)t} + (1-p)\frac{\lambda_y}{\lambda_y+\lambda_2}e^{-\lambda_2 t} \tag{21}$$

$$CS_D^*(t) = \frac{e^{-(\lambda_1+\lambda_y)t} + \frac{(1-p)}{p}\frac{\lambda_2}{\lambda_1}\frac{(\lambda_y+\lambda_1)}{(\lambda_y+\lambda_2)}e^{-(\lambda_2+\lambda_y)t}}{1+\frac{(1-p)}{p}\frac{\lambda_2}{\lambda_1}\frac{(\lambda_y+\lambda_1)}{(\lambda_y+\lambda_2)}} \tag{22}$$

$$CS_C^*(t) = \frac{e^{-(\lambda_1+\lambda_y)t} + \frac{(1-p)}{p}\frac{(\lambda_y+\lambda_1)}{(\lambda_y+\lambda_2)}e^{-(\lambda_2+\lambda_y)t}}{1+\frac{(1-p)}{p}\frac{(\lambda_y+\lambda_1)}{(\lambda_y+\lambda_2)}} \tag{23}$$

$$CS_D^*(t) \le CS_C^*(t) \tag{24}$$

Moreover, $\Phi(t)$ is minimal at the origin, and is strictly increasing when $\lambda_1, \neq \lambda_2$.

A2.2. Empirical Subsurvivor and Conditional Subsurvivor functions (Bunea, 2003).

$$\hat{S}_X^*(t) = \frac{\text{number of } X \text{ events after time t}}{\text{total number of events}} \tag{25}$$

$$\hat{S}_Y^*(t) = \frac{\text{number of } Y \text{ events after time t}}{\text{total number of events}} \tag{26}$$

$$C\hat{S}_X^*(t) = \frac{\text{number of } X \text{ events after time t}}{\text{total number of } X \text{ events}} \tag{27}$$

$$C\hat{S}_Y^*(t) = \frac{\text{number of } Y \text{ events after time t}}{\text{total number of } Y \text{ events}} \tag{28}$$

Chapter 3. Short Term Effects of Air Pollution in Mortality: an Example with Toluca City.

As explained in chapter one, the short term effects of PM in mortality are described by regression models. In this chapter, the theoretical foundation of these kinds of regressions is explored and examples are shown with data from Toluca's Metropolitan Area (ZMT). Toluca was chosen to exemplify the theory presented here because it is a relatively important city in Mexico that still lacks of a formal study on air pollution effects on health. One remark that the reader should remember through this chapter is that the mortality data available corresponds to two years, which is a smaller set of observations than in other studies hence explaining possible differences with these as for example wider confidence bands.

3.1. Description of Toluca City.

Toluca is the capital city of the central Mexican state, State of Mexico. Air pollution data for Toluca city is recorded in RAMA-T by 7 monitoring stations in 3 municipalities in the State of Mexico. The codes for the stations, the municipalities to which they belong and the names of each station are summarized in Table and Figure 3.1

Table 3.1 RAMA-T Description.

Abbreviation	Name	Municipality
OX	Nueva Oxtotitlan	Toluca
CE	Toluca Centro	Toluca
MT	Metepec	Metepec
SM	San Mateo Atenco	San Mateo Atenco
AP	Aeropuerto	Toluca
SL	San Lorenzo Tepaltitlan	Toluca
SC	San Cristobal Huichochitlan	Toluca

Data for daily number of deaths from the country was available from the mortality database, INEGI/SSA 1990 and 2000 delivered personally by Raydel Valdés from Instituto Nacional de Salud Pública, Centro de Investigación en Salud Poblacional in Cuernavaca Mexico. From this data base, daily counts were extracted for the municipalities of the State of Mexico coded by INEGI (National Institute of Geography Statistics and Informatics) as 106 (Toluca), 076 (San Mateo Atenco) and 054 (Metepec).

The data base for pollution is "Bases de datos históricas de los contaminantes criterio de la ZMVT" provided by Instituto Nacional de Ecología-SEMARNAT (Nacional Institute of Ecology). Air pollution data is corrected for temperature and barometric pressure according to the Mexican official norm "Normas Oficiales Mexicanas (NOM-034-ECOL-1993 a NOM-038-ECOL-1993)" with the following criteria: temperature of 298 K(25°C) and barometric pressure of 101 KPa (760 mm de Hg). The pollutants are given with the following units: Ozone ppm, NO2 ppm, PM_{10} $\mu g/m^3$. A summary for the data is presented in the next tables 3.2 and 3.3.

Figure 3.1 RAMA-T in Toluca City.

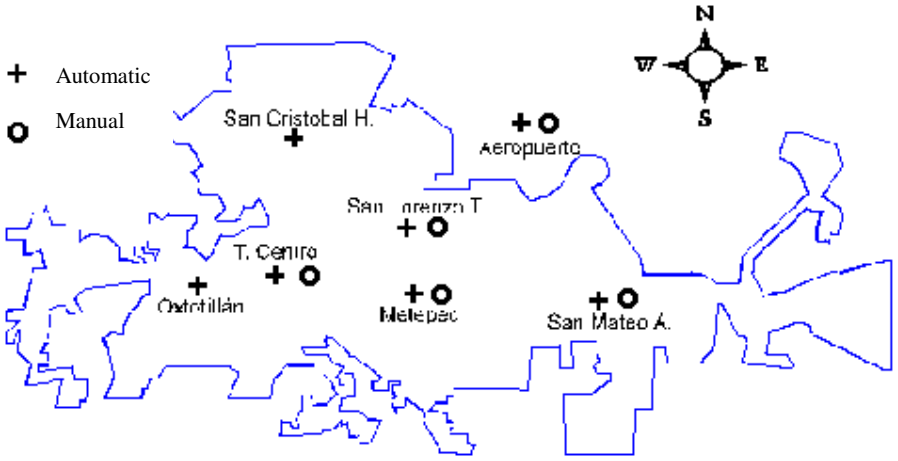


Table 3.2 Variables Description

nad64-(circ64+res64)	onad64
Daily circulatory system related number of deaths in people less than 65	circ64
Daily respiratory system related number of deaths in people less than 65	res64
Daily non accidental number of deaths in people less than 65	nad64
nad65-(circ65+res65)	onad65
Daily circulatory system related number of deaths in people 65 or older	circ65
Daily respiratory system related number of deaths in people 65 or older	res65
Daily non accidental number of deaths in people 65 or older	nad65
Daily respiratory system related number of deaths	respt
Daily circulatory system related number of deaths	circt
namt-(circt+respt)	onamt
Daily non accidental number of deaths	namt
24 hour PM10 average concentration (µg/m^3)	pm10
24 hour O3 maximum concentration (ppm)	O3max
24 hour NO2 average concentration (ppm)	NO2av
24 hour minimum temperature (°C)	TMPmin..
3 previous days average temperature (°C)	pdtmin.3
24 hour average relative humidity (%)	RH
3 previous days average relative humidity (%)	pdrh.3
24 hour PM10 average concentration of the third day prior to death (µg/m^3)	pm10.3
24 hour O3 maximum concentration of the third day prior to death (ppm)	O3max.3
24 hour NO2 average concentration of the third day prior to death (ppm)	NO2av.3

Table 3.3 Statistical Summary of Variables of Interest

	onad64	circ64	res64	nad64	onad65	circ65	res65	nad65	respt	circt	onamt	namt
Min:	0	0	0	0	0	0	0	0	0	0	1.0	2.0
Mean:	6.1	0.7	0.6	7.4	3.0	1.6	0.6	5.2	1.2	2.4	9.0	12.6
Median:	6	1	0	7	3	1	0	5	1	2	9	12
Max:	15.0	5.0	6.0	17.0	9.0	7.0	5.0	14.0	9.0	9.0	19.0	24.0
Variance:	6.1	0.8	0.7	8.1	3.0	1.6	0.8	5.3	1.7	2.3	9.5	14.3
Sum:	4,434	544	449	5,427	2,161	1,185	463	3,809	912	1,729	6,595	9,236
	82%	10%	8%	100%	57%	31%	12%	100%	10%	19%	71%	100%

Units as in Table 3.2.

Table 3.3 Statistical Summary of Variables of Interest (continue)

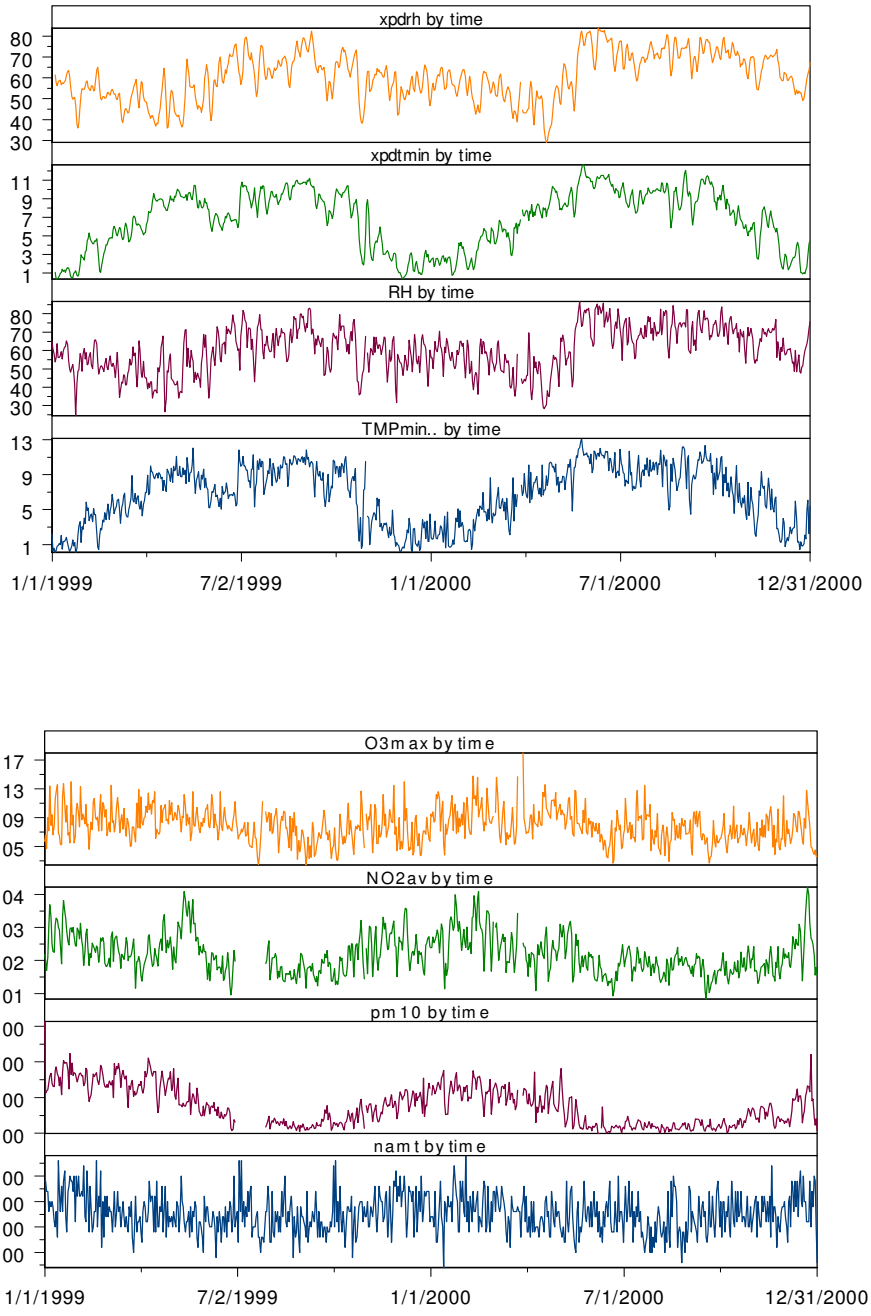
	pm10	O3max	NO2av	TMPmin..	pdtmin.3	RH
Min:	9.717	0.024	0.008	0.100	0.333	24.483
Mean:	47.604	0.079	0.022	6.701	6.715	60.581
Median:	43.548	0.078	0.021	7.214	7.400	60.948
Max:	166.725	0.180	0.042	13.157	12.662	86.607
NA's :	35.000	8.000	34.000	4.000	4.000	4.000
Variance:	740.892	0.001	0.000	10.986	10.075	145.953

Units as in Table 3.2. NA's means number of missing observation values.
Data is from 731 daily averages (2 years).

From table 3.3 it is observed that there are 9,236 non accidental deaths in two years (731 days) in Toluca, roughly 19% of these are related to circulatory system diseases and 10% are respiratory system related deaths; also, 41% of the total number of deaths is among people aged 65 or more. The proportion of circulatory and respiratory related mortality is higher in the older age group (31% and 12% respectively) than in the younger. Some words of caution that the reader must bare in mind is that previous studies of "Mexico City death certificates for the years 1991 and 1994 suggested that 12% of deaths might be misclassified among broad categories such as respiratory and cardiovascular causes" (Borja-Aburto, 1998) this problem could be also observed in the case of Toluca.

The Mexican norm for ozone (NOM-020-SSA1-1993) is 0.11 ppm 1 hour average, for nitrogen dioxide (NOM-023-SSA1-1993) 0.21 ppm 1 hour average and 50 $\mu g/m^3$ one year average for PM_{10} (NOM-025-SSA1-1993). The average PM_{10} concentration for the two years is below the Mexican allowable level for 1 year, and so are the average concentrations of Ozone and nitrogen dioxide for their respective norms (table 3.3). A better picture of the daily variations for these variables can be observed in figure 3.2; for example, in some days the PM_{10} average concentration is above a 100 $\mu g/m^3$, ozone also exceeds in some days the maximum allowable while the NO_2 daily average concentration is always bellow the maximum 1 hour tolerance level.

Figure 3.2 Time series for selected variables.



namt = non accidental mortality, pm10 = PM_{10} ($\mu g/m^3$), NO2av = NO_2 (ppm) daily average concentration, O3max = Daily maximum concentration of ozone (ppm), TMPmin.. = minimum 24hour temperature (C°), RH = Relative Humidity (%), xpdtmp = 3 previous days average temperature, xpdrh = 3 previous days average relative humidity.

3.2. Generalized Linear Models.

The first classes of models to be considered are generalized linear models (GLM), these are an extension of classical linear models of the form

$$E(Y_i) = \mu_i = x_i^T \beta + \varepsilon; \qquad Y_i \sim N(\mu_i, \sigma^2) \tag{3.1}$$

Where the random variables Y_i are independent and form the basis of most analysis of continuous data. However, methods analogous to those developed for linear models are now used in the more general situations where the response variable follows other distributions than normal (they might be discrete) and where the relationship between expected response and explanatory variable is not necessarily linear. The exponential family of distributions (Appendix) offers some properties that make it desirable for these kinds of models.

A generalized linear model has three components:

1. Response variables $Y_1,\ldots,Y_N$ which are assumed to share the same distribution from the exponential family in the canonical form.
2. A set of parameters $\beta_1, \ldots, \beta_p$ and explanatory variables

$$X = \begin{matrix} x_{11} & . & . & . & x_{1p} \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ x_{N1} & . & . & . & x_{Np} \end{matrix}$$

3. A monotone link function g such that $g(\mu_i) = \mathrm{x_i}^{\mathrm{T}} \beta + \varepsilon$ with $\mu_i = \mathrm{E}(Y_i)$

Models with data showing the auto-regressions of time series are excluded from a GLM analysis because they assume independent observations are being studied. A second assumption is that there is a single error term in the model. In classical linear models, the error term is assumed to be normal with constant variance. Hence, whether choosing Y or $\ln(Y)$ as scale has to take into consideration the restrictions of constant variance and normality for the error components together with additivity of the systematic components. In GLM this scaling problem is removed because normality and constant variance is no longer a requirement for the error component.

Statistical packages such as S-Plus already support a routine for fitting GLM as described in the appendix. For example consider the data for Toluca city presented in previous sections, if we assume that $Y_1,\ldots,Y_N$ are independent random variables with Y_i denoting the number of deaths observed in a given day, then $\mathrm{E}(Y_i) = \mu_i = \theta_i$ and the dependence of θ_i on explanatory variables is usually modeled as

$$\theta_i = e^{x_i^T \beta} \tag{3.2}$$

Therefore, the generalized linear model with the logarithmic link function is

$$E(Y_i) = \mu_i = e^{x_i^T \beta}; \qquad Y_i \sim Poisson(\mu_i) \tag{3.3}$$

$$\ln(\mu_i) = x_i^T \beta \tag{3.4}$$

In some kind of models θ_i is multiplied by a constant n_i for example if Y is the number of insurance claims for a particular model of car then n_i would denote the number of cars of a particular kind. The amount $\ln(n_i)$ will be added to (3.4); this amount is called the offset and it is a known constant which can be incorporated into the estimation procedure.

One of the simplest models, is to explain impact on mortality in Toluca from PM_{10} using a model just like the one described by (3.2) – (3.4) in health effect studies the offset is normally not considered; the two vectors (x_0 and x_1) incorporated for the estimation of a model containing PM_{10} as the only predictor would be a vector of ones (that estimates the parameter β_0 and a vector containing the PM_{10} measurements); the results from fitting this very simple model in S-Plus are shown in table 3.4; the value for each coefficient was checked in Microsoft Excel and Matlab with the algorithm and the program presented in the appendix.

Table 3.4 Example of Poisson Regression for Mortality and PM_{10} Concentration in Toluca City.

```
glm(formula = namt ~ pm10, family = poisson(link = log), data =
      MORTALITYSMALLCALCzmvt1, na.action = na.exclude, control = list(
      epsilon = 0.0001, maxit = 50, trace = F))

Coefficients:
                          Value      Std. Error      t value
Intercept   (β0)     2.424638586    0.0217589095    111.43199
     pm10   (β1)     0.002369523    0.0003842811      6.16612
```

The table above presents the estimated values for the coefficients (β_0, β_1), its standard error (square root of the diagonal element corresponding to covariate i in the inverse of the information matrix) and the Wald test statistic i. e. the ratio of the coefficient value to its standard error, which is asymptotically standard normal. In practice, whenever the p value (left tale of the standard normal distribution evaluated at the Wald test statistic) is less than 0.05 the hypothesis that the true value of the coefficient is zero is rejected. For the example above, the p value of the Wald statistic would be less than 0.000 showing a high probability that the coefficient is significantly different than zero.

The usual interpretation or risk ratio applies to the situation at hand, hence, for a binary explanatory variable with $x_j = 0$ if the factor is absent and $x_j = 1$ if it is present the risk ratio will look as formula (3.4)

$$RR = \frac{E(Y_i | present)}{E(Y_i | absent)} = e^{\beta_j} \tag{3.5}$$

For a continuous explanatory variable, a one unit increase will result in a multiplicative effect of e^{β} on μ, if all other factors remain the same. If the regression above were correct then there would be a multiplicative effect of $e^{pm10} = 1.002372336$ on the mortality rate for a one $\mu g/m^3$ increase in PM_{10} exposure; this actually means a 0.23% increase in total mortality per $\mu g/m^3$ increase in PM_{10} average concentration or a 2.3 % increase in total mortality per 10 $\mu g/m^3$ increase in PM_{10} average concentration. Often times, epidemiologists take the estimated coefficient (β_1 in table 3.4) and multiply it by 1000 (= 2.369523) to express the % change in mortality due to a 10 $\mu g/m^3$ increase in PM_{10} average concentration which for small β is a good approximation.

To make inferences about the estimator, one common measure is the analysis of deviance; this estimator is derived from the notions of saturated, maximal or full model. If there are N observations all with potentially different values for the linear component $x_i^T \beta$ then the full model can have N parameters. However if some of the observations have the same linear component or covariate pattern, they are called replicates. In this case the maximum number of parameters that may be estimated from the saturated model m are less than or equal to N.

Call β_{max} the vector of parameters for the saturated model and $\mathbf{b}_{max}$ it's vector of maximum likelihood estimators. The likelihood function evaluated at $\mathbf{b}_{max}$ will be larger than any other likelihood function for these observations if $L(\mathbf{b};\mathbf{y})$ denotes the maximum value of the likelihood function for the model of interest, then the likelihood ratio is:

$$\lambda = \frac{L(b_{max}; y)}{L(b; y)} \tag{3.6}$$

$$\ln(\lambda) = l(b_{max}; y) - l(b; y) \tag{3.7}$$

The quantity $2\ln(\lambda)$ has a chi-squared distribution with degrees of freedom equal to the number of paramteres in the full model minus the number of parameters to be estimated (see appendix) and can be used for inference in a similar way as an anova table.

3.3. Generalized Additive Models.

In Generalized Additive Models (GAM) the elements to be considered are,

1. Response variables $Y_1,\ldots,Y_N$ which are assumed to share the same distribution from the exponential family in the canonical form.
2. A set of explanatory variables

$$X = \begin{matrix} x_{11} & . & . & . & x_{1d} \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ x_{N1} & . & . & . & x_{Nd} \end{matrix}$$

3. A link function g such that $\eta_i = g(\mu_i) = f_0 + \sum_{j=1}^{d} f_j(X_j)$ with $\mu_i = \mathrm{E}(Y_i|X_i)$

f_j is a nonparametric function estimated from data using smoothing operations, describing the relationship between each of the j predictors and the transformed mean response. The estimation of these functions is accomplished by the algorithm described below (Hastie and Tibshirani, 1990) and (Schimek, 1999).

Table 3.5 Local Scoring Algorithm.

1. Initialization of the outer loop: $f_0^0 = g(n^{-1}\sum_{j=1}^{p} f_j(X_j)), f_1^0 = f_2^0 \ldots = f_d^0 = 0$

2. Scoring steps: For k = 0, 1, 2, . . .

   Compute working observations $z_i^k = \eta_i^k + (y_i - \mu_i^k)\left(\frac{\partial \eta_i}{\partial \mu_i}\right)$

   And weights $w_i^k = \left(\frac{\partial \mu_i^k}{\partial \eta_i^k}\right)^2 (V_i^k)^{-1}$

   Recall that for the Poisson regression $\eta_i = \ln(\mu_i) = f_0^k + \sum_{j=1}^{d} f_j^k(X_j)$

   $z_i^k = \eta_i^k + y_i / e^{\eta_i^k} - 1$ and $w_i^k = e^{\eta_i^k}$

3. Backfitting steps: Solve for f^{k+1} for $m = 0, 1, 2, \ldots$

   a) *Initialization of inner loop*:

   $f_0^m = n^{-1}\sum_{j=1}^{n} z_i^k, f_j^0 = f_j^k \quad j, p, c = 1, \ldots, d$

   b) *Compute updates for f_j*:

   $f_j^{m+1} = S_j\left[z - f_0 - \sum_{c<j} f_c^{m+1} - \sum_{p>j} f_p^{m+1}\right]$ index c for current and p for previous.

   $f_j^m = f_j^{m+1}$

   c) *Stop backfitting when*:

   $\left\| f_j^m - f_j^{m+1} \right\|$ is very small

4. Compute termination criteria:

   Set $f_j^{k+1} = f_j^{m+1}$, and compute $\frac{\sum_{j=1}^{d} \left\| f_j^{k+1} - f_j^k \right\|}{\sum_{j=1}^{d} \left\| f_j^k \right\|} \le \varepsilon$

In the above algorithm, S_j is a weighted cubic smoothing-spline operator; from (Schimek, 1999) and (Green and Silverman, 1994) if $h_i = x_{i+1} - x_i$ for $i = 1, 2, \ldots, n-1$ then $K = \Delta^T C^{-1} \Delta$ is a quadratic penalty matrix where Δ is a tri-diagonal (n-2) × n matrix with $\Delta_{i,i} = 1/h_i$, $\Delta_{i,i+1} = -(1/h_i + 1/h_{i+1})$, $\Delta_{i,i+2} = 1/h_{i+1}$, and a symmetric tri-diagonal (n-2) × (n-2) matrix C with $c_{i-1,i} = c_{i,i-1} = h_i / 6$, $c_{i,i} = (h_i + h_{i+1})/3$. Then $S_j = (W + \lambda_j K_j)^{-1} W$, where W is a diagonal matrix with elements w_{ii} the reader may find the motivation for this approach in the appendix.

Inference in this model is similar to that of GLM and the reader is referred to the references to go further in this respect.

3.4 Results for Toluca City.

The models to be studied and compared in the present chapter for Toluca city when using GLM or GAM for mortality counts Y_t, will take the form (3.8) and (3.9) respectively. A model very similar to (3.9) was used in the NMMAPS study (Dominici et. al., 2000) and is also comparable to previous studies in Mexico (Borja Aburto, 1998) and (Castillejos, 2000).

$$\ln(E(Y_t)) = \beta_0 + \beta_1 X_{t-l} + \beta_2 DOW + \beta_3 time + \beta_4 temp + \beta_5 hum + \beta_6 prevtemp + \beta_7 prevhum \quad (3.8)$$

$$\ln(E(Y_t)) = \beta_0 + \beta_1 X_{t-l} + \beta_2 DOW + S(time, \lambda_1) + S(temp, \lambda_2) + S(prevtemp, \lambda_3) + S(hum, \lambda_2) + S(prevhum, \lambda_3) \quad (3.9)$$

In the expression above X is the matrix of air pollution measures, *time* is the calendar time, *temp* is a measure of temperature (minimum in this case), *prevtemp* is the average minimum temperature of 3 previous days, *hum* is a measure of humidity or dew point (relative humidity), *prevhum* is the average relative humidity of 3 previous days, *DOW* are indicator variables for day of the week, l is the lag of the pollution exposure which is generally restricted to 0 to 7 days. The functions $S(\cdot, \lambda)$ denote smooth functions of the covariates with smoothing parameter λ. For the models presented below λ_1 = 7/year, λ_2 = 6 and λ_3 = 3, this choice is again motivated by the model used in NMMAPS.

The model (3.9) is semi-parametric in the sense that besides the parametric part of the model that enters in the usual linear relationship of GLM, smooth functions of certain covariates are included to take into consideration possible non linear relationship of these to daily mortality counts. The model is fit using the *local scoring* algorithm, described in previous sections and separating the parametric from the nonparametric part of the fit; the parametric part is fit using weighted linear least squares (Appendix) within the backfitting algorithm. Next the results of fitting models (3.8) and (3.9) for Toluca city will be presented.

Table 3.6 a) has the same interpretation as table 3.4; the first two columns show the name and estimated coefficients of each variable obtained by solving equation (28) in the appendix, the calculations were checked in Matlab with the code presented in appendix 3.3; the algorithm used by S-plus identifies a singularity in the estimate of the information matrix shown in the left hand side of (28) when the 7 indicator variables for day of the week are used and hence eliminates the last indicator (sun) to solve the system described by (28) with all but this covariate. The next column shows standard errors computed as the square root of the main diagonal elements of the inverse of the estimated information matrix after the last iteration; and the last column shows the Wald statistic; the p value for this statistic for pm10 in the example below would be 0.2207. As before, whenever the p value (left tale of the standard normal distribution evaluated at the Wald test statistic) is less than 0.05 the hypothesis that the true value of the coefficient is zero can be rejected; table 3.6 a) shows that it is very likely that the 0.0006 estimator of the coefficient for PM_{10} is not significantly different than zero.

Table 3.6 b) is to be interpreted in the same way as table 3.6) except that the estimations are obtained with GAMs; as in a) the indicator variable for Sunday is removed from the calculations as singularities are found when using this covariate. The variables that are smoothed are denoted by s(variable, λ_i). The reader should notice that the estimated coefficient shows a positive influence of PM_{10} 24 hour average concentrations on health however this coefficient is statistically not significantly different than zero (p value = 0.276).

Table3.6 Results for the Relation of PM_{10} in Total non Accidental Mortality in Toluca from Models 3.8 and 3.9.

a) GLM Model (3.8)

	Value	Std. Error	t value
(Intercept)	2.570793e+000	0.13303444897	19.32426868
pm10	6.127692e-004	0.00079600471	0.76980605
O3max	-2.561561e-001	0.58421245327	-0.43846398
NO2av	3.141127e+000	2.69499757224	1.16553982
mon	-8.101378e-003	0.04052810624	-0.19989531
tus	-2.638994e-002	0.04110338911	-0.64203814
wed	-4.079026e-002	0.04145779026	-0.98389857
thr	-9.821575e-003	0.04131613149	-0.23771768
fri	-3.157690e-002	0.04176243050	-0.75610770
sat	-1.866582e-002	0.04091670076	-0.45619076
sun	NA	NA	NA
day	8.120721e-007	0.00006142607	0.01322032
TMPmin..	-2.349803e-003	0.00794113366	-0.29590270
pdtmin.3	-2.094585e-002	0.00847319387	-2.47201351
RH	-9.081254e-004	0.00156641054	-0.57974932
pdrh.3	1.863330e-003	0.00176058157	1.05836046

b) GAM Model (3.9)

	Value	Std. Error	t value
(Intercept)	2.69236718192	0.1319031336	20.4116999
pm10	-0.00047217204	0.0007955877	-0.5934884
O3max	-0.13939559296	0.5848676198	-0.2383370
NO2av	5.00340947549	2.7031572276	1.8509502
mon	-0.00488470905	0.0405263250	-0.1205318
tus	-0.02584708725	0.0410963289	-0.6289391
wed	-0.03964965077	0.0414670037	-0.9561735
thr	-0.00752704590	0.0413266194	-0.1821355
fri	-0.03062093956	0.0417558846	-0.7333323
sat	-0.01076352800	0.0409239990	-0.2630126
s(day, 14)	-0.00004388885	0.0000614229	-0.7145356
s(TMPmin.., 6)	0.00155339734	0.0079387640	0.1956724
s(pdtmin.3, 3)	-0.01066454682	0.0084476575	-1.2624265
s(RH, 6)	-0.00183729973	0.0015626216	-1.1757804
s(pdrh.3, 3)	-0.00043561371	0.0017526721	-0.2485426

Table3.7 Analysis of Deviance for the Models Described in table 3.6.

a) GLM Model (3.8)

	Df	Delta D.	Df	Resid. Dev	Pr(Chi)
NULL			690	787.1315	
mon	1	0.17390	689	786.9576	0.6766723
tus	1	0.00400	688	786.9536	0.9495826
wed	1	0.42047	687	786.5331	0.5167041
thr	1	0.19920	686	786.3339	0.6553645
fri	1	0.08428	685	786.2496	0.7715771
sat	1	0.03262	684	786.2170	0.8566695
sun	0	0.00000	684	786.2170	
day	1	3.73514	683	782.4819	0.0532792
TMPmin..	1	56.72469	682	725.7572	0.0000000
pdtmin.3	1	7.58442	681	718.1728	0.0058875
RH	1	0.20787	680	717.9649	0.6484395
pdrh.3	1	0.60426	679	717.3606	0.4369579
NO2av	1	2.18471	678	715.1759	0.1393871
O3max	1	0.14690	677	715.0290	0.7015182
pm10	1	0.59172	676	714.4373	0.4417533

b) GAM Model (3.9)

	Df	Delta D	Df	Resid. Dev	Pr(Chi)
NULL			690	787.1315	
Mon	1	0.17390	689	786.9576	0.6766723
tus	1	0.00400	688	786.9536	0.9495826
wed	1	0.42047	687	786.5331	0.5167041
thr	1	0.19920	686	786.3339	0.6553645
fri	1	0.08428	685	786.2496	0.7715771
sat	1	0.03262	684	786.2170	0.8566695
sun	0	0.00000	684	786.2170	
s(day, 14)	1	3.73514	683	782.4819	0.0532792
s(TMPmin.., 6)	1	56.72469	682	725.7572	0.0000000
s(pdtmin.3, 3)	1	7.58442	681	718.1728	0.0058875
s(RH, 6)	1	0.20787	680	717.9649	0.6484395
s(pdrh.3, 3)	1	0.60426	679	717.3606	0.4369579
NO2av	1	2.18471	678	715.1759	0.1393871
O3max	1	0.14690	677	715.0290	0.7015182
pm10	28	60.92454	648	654.1045	0.0003109

Table 3.7 a) shows the analysis of deviance for GLM, this table consists of 6 columns; the reader should look first at column 5 that shows the value of the deviance computed with the coefficients that a fit with all variables up to a given row would find, for instance a model that includes all variables up to O3max would have a deviance of approximately 715.03. This is approximately a chi squared variable with 677 degrees of freedom as shown in column 4 (the degrees of freedom are equal to the number of non missing observations in the data minus the number of parameters in the model that are 13 for this case); the deviance of the model including all variables is shown at the last row, column five, of the table (714.4373 with 676 degrees of freedom) if this value is subtracted from the value found previously with all variables except PM concentrations a chi squared variable with 1 degree of freedom will be found. This is shown in the column labeled Delta D.

In practice, whenever the *p* value (p(Chi) in the table above) of this number is less than 0.05 the null hypothesis that the model including PM_{10} is equal to the model including all other variables but PM_{10} can be rejected, showing that a model including PM_{10} would describe better the data than a model excluding this variable. For the example above it is observed that it is likely (44% probability) that a model containing all variables up to O3 is not significantly different from a model containing also PM_{10}.

The reader may see that the only difference between tables 3.7 a) and b) is the last row. This is because the hypothesis to be tested is whether a model like 3.9 that contains smooth functions of calendar time, temperature and relative humidity is significantly different from a GLM that contains all variables except PM_{10}. In this case, Delta D has 28 degrees of freedom this is because the deviance up to O3max has 677 degrees of freedom and the GAM model contains 690 degrees of freedom from the data minus one degree of freedom for the parameter for the intercept and each indicator variable except Sunday, 14 for the smoothing parameter λ_1, 6 for each of the two λ_2 parameters, 3 each of the two λ_3 and one for each of the three pollutants considered giving a total of 648 degrees of freedom. It can be observed that the null hypothesis that a GAM containing all variables of interest with the smooth functions of selected variables is equal to a GLM containing all variables except PM_{10} can be rejected with some confidence.

Further analysis will present risk ratios for 10 $\mu g/m^3$ increase in PM_{10} using the coefficients computed by models (3.8) and (3.9), normal approximations of 95% confidence intervals and *p* values for the Wald statistic and the Delta D statistic that for GLM test the null hypothesis that a model containing all variables except PM_{10} is equal to a model containing also this variable; and for GAMs the null hypothesis that a GAM with all variables is equal to a GLM with all variables except PM_{10}, this *p* values will be labeled p(t) for the Wald statistic and p(chi) for the Delta D statistic. The analysis breaks total non accidental mortality (ICD-10 codes A to R) in circulatory system related mortality (ICD-10 code I), respiratory system related mortality (ICD-10 code J) and all other non accidental mortality. The results are presented in table 3.7 for same day pollution and in table 3.8 for a 3 day lag pollution; the choice for the air pollution lag follows the methodology employed in NMMAPS.

Table 3.7 Risk Ratio for 10 μg/m^3 increase in PM_{10} average concentration in the same day.

Cause of death and pollutants included in the model		GLM (Model 3.8) Risk Ratio and (95% C.I.)	p(t)	p(Chi)	GAM (Model 3.9) Risk Ratio and (95% C.I.)	p(t)	p(Chi)
Total non accidental mortality	PM_{10}	1.010 (0.995, 1.024)	0.096	0.192	1.003 (0.989, 1.018)	0.336	0.000
	$PM_{10}+O_3max$	1.009 (0.994, 1.024)	0.115	0.230	1.001 (0.987, 1.016)	0.424	0.000
	$PM_{10}+O_3max+NO_2av$	1.006 (0.991, 1.022)	0.221	0.442	0.995 (0.980, 1.011)	0.276	0.000
Total circulatory system related mortality	PM_{10}	0.999 (0.967, 1.033)	0.478	0.957	0.990 (0.958, 1.023)	0.275	0.205
	$PM_{10}+O_3max$	0.993 (0.960, 1.027)	0.343	0.687	0.983 (0.950, 1.017)	0.166	0.211
	$PM_{10}+O_3max+NO_2av$	0.983 (0.948, 1.019)	0.175	0.843	0.965 (0.931, 1.001)	0.028	0.151
Total respiratory system related mortality	PM_{10}	1.017 (0.973, 1.064)	0.224	0.448	1.004 (0.961, 1.050)	0.426	0.105
	$PM_{10}+O_3max$	1.026 (0.980, 1.074)	0.139	0.279	1.009 (0.963, 1.056)	0.356	0.125
	$PM_{10}+O_3max+NO_2av$	1.026 (0.977, 1.077)	0.153	0.306	1.005 (0.957, 1.055)	0.417	0.127
Other non accidental mortality total	PM_{10}	1.011 (0.994, 1.028)	0.104	0.209	1.005 (0.988, 1.022)	0.291	0.009
	$PM_{10}+O_3max$	1.011 (0.993, 1.029)	0.115	0.231	1.003 (0.986, 1.021)	0.349	0.009
	$PM_{10}+O_3max+NO_2av$	1.010 (0.991, 1.028)	0.157	0.315	1.000 (0.982, 1.019)	0.499	0.008
Total non accidental mortality (people < 65)	PM_{10}	1.003 (0.984, 1.022)	0.377	0.753	1.002 (0.984, 1.021)	0.406	0.001
	$PM_{10}+O_3max$	1.006 (0.987, 1.026)	0.266	0.532	1.004 (0.985 1.023)	0.344	0.001
	$PM_{10}+O_3max+NO_2av$	1.003 (0.983, 1.023)	0.394	0.787	0.997 (0.977, 1.017)	0.389	0.001
Total circulatory system related mortality (people < 65)	PM_{10}	**1.067** (1.006,1.131)	**0.015**	**0.031**	**1.062** (1.002, 1.125)	**0.021**	**0.010**
	$PM_{10}+O_3max$	**1.067** (1.005, 1.133)	**0.017**	**0.034**	**1.065** (1.004, 1.130)	**0.019**	**0.010**
	$PM_{10}+O_3max+NO_2av$	1.047 (0.983, 1.115)	0.078	0.158	1.027 (0.964, 1.094)	0.203	0.013
Total respiratory system related mortality (people < 65)	PM_{10}	0.952 (0.893, 1.014)	0.063	0.127	0.961 (0.903, 1.022)	0.103	0.0201
	$PM_{10}+O_3max$	0.969 (0.907, 1.034)	0.170	0.341	0.974 (0.914, 1.039)	0.211	0.0453
	$PM_{10}+O_3max+NO_2av$	0.969 (0.903, 1.039)	0.185	0.369	0.970 (0.906, 1.039)	0.194	0.0458
Other non accidental mortality total (people < 65)	PM_{10}	1.000 (0.980, 1.021)	0.483	0.967	0.998 (0.978, 1.019)	0.439	0.0167
	$PM_{10}+O_3max$	1.002 (0.981, 1.024)	0.411	0.823	0.998 (0.977, 1.020)	0.441	0.0169
	$PM_{10}+O_3max+NO_2av$	1.001 (0.979, 1.024)	0.458	0.917	0.995 (0.973, 1.018)	0.342	0.0155
Total non accidental mortality (people > 65)	PM_{10}	1.019 (0.996, 1.042)	0.049	0.099	1.002 (0.980, 1.025)	0.421	0.0488
	$PM_{10}+O_3max$	1.013 (0.990, 1.037)	0.131	0.263	0.996 (0.973, 1.019)	0.371	0.0672
	$PM_{10}+O_3max+NO_2av$	1.011 (0.987, 0.879)	0.190	0.380	0.991 (0.967, 1.015)	0.227	0.0662
Total circulatory system related mortality (people > 65)	PM_{10}	0.970 (0.932, 1.009)	0.063	0.127	0.957 (0.920, 0.996)	0.016	0.1017
	$PM_{10}+O_3max$	**0.960 (0.921, 1.000)**	**0.026**	**0.053**	**0.946 (0.907, 0.986)**	**0.004**	**0.0709**
	$PM_{10}+O_3max+NO_2av$	**0.955 (0.914, 0.998)**	**0.020**	**0.040**	**0.937 (0.897, 0.980)**	**0.002**	**0.0603**
Total respiratory system related mortality (people > 65)	PM_{10}	**1.086 (1.020, 1.155)**	**0.005**	**0.010**	1.052 (0.988, 1.121)	0.058	0.1126
	$PM_{10}+O_3max$	**1.085 (1.017, 1.157)**	**0.007**	**0.014**	1.047 (0.980, 1.118)	0.086	0.1316
	$PM_{10}+O_3max+NO_2av$	**1.085 (1.014, 1.161)**	**0.009**	**0.020**	1.043 (0.973, 1.119)	0.115	0.1467
Other non accidental mortality total (people > 65)	PM_{10}	**1.033 (1.003, 1.065)**	**0.016**	**0.032**	1.019 (0.989, 1.050)	0.107	0.2753
	$PM_{10}+O_3max$	1.029 (0.997, 1.061)	0.036	0.073	1.015 (0.984, 1.047)	0.170	0.3656
	$PM_{10}+O_3max+NO_2av$	1.027 (0.995, 1.061)	0.052	0.104	1.010 (0.978, 1.044)	0.273	0.3733

Table 3.8 Risk Ratio for 10 μg/m^3 increase in PM_{10} average concentration at lag 3.

Cause of death and pollutants included in the model		GLM (Model 3.8)			GAM (Model 3.9)		
		Risk Ratio and (95% C.I.)	p(t)	p(Chi)	Risk Ratio and (95% C.I.)	p(t)	p(Chi)
Total non accidental mortality	PM_{10}	**1.014** (1.001 1.026)	**0.016**	**0.031**	**1.011** (0.998, 1.023)	**0.043**	**0.001**
	$PM_{10}+O_3max$	**1.014** (1.001 1.027)	**0.017**	**0.035**	**1.010** (0.997, 1.023)	**0.064**	**0.001**
	$PM_{10}+O_3max+NO_2av$	**1.014** (1.001 1.028)	**0.021**	**0.043**	**1.011** (0.998, 1.025)	**0.052**	**0.001**
Total circulatory system related mortality	PM_{10}	1.012 (0.983, 1.041)	0.209	0.418	1.012 (0.983, 1.041)	0.215	0.241
	$PM_{10}+O_3max$	1.008 (0.978, 1.038)	0.304	0.609	1.018 (0.986, 1.051)	0.137	0.135
	$PM_{10}+O_3max+NO_2av$	1.004 (0.973, 1.036)	0.402	0.804	1.004 (0.972, 1.036)	0.408	0.283
Total respiratory system related mortality	PM_{10}	1.032 (0.993, 1.071)	0.055	0.111	1.022 (0.984, 1.062)	0.130	0.110
	$PM_{10}+O_3max$	**1.036** (0.997, 1.078)	**0.036**	**0.075**	1.025 (0.985, 1.066)	0.112	0.113
	$PM_{10}+O_3max+NO_2av$	1.031 (0.988, 1.031)	0.079	0.162	1.018 (0.975, 1.062)	0.211	0.150
Other non accidental mortality total	PM_{10}	1.012 (0.997, 1.027)	0.060	0.121	1.007 (0.992, 1.022)	0.184	0.012
	$PM_{10}+O_3max$	1.012 (0.997, 1.028)	0.059	0.118	1.006 (0.991, 1.022)	0.221	0.014
	$PM_{10}+O_3max+NO_2av$	1.015 (0.998, 1.031)	**0.039**	**0.078**	1.010 (0.994, 1.027)	0.113	0.010
Total non accidental mortality (people < 65)	PM_{10}	1.008 (0.992, 1.025)	0.155	0.311	**1.011** (0.995, 1.028)	**0.088**	**0.002**
	$PM_{10}+O_3max$	1.011 (0.994, 1.028)	0.100	0.201	**1.012** (0.995, 1.029)	**0.085**	**0.002**
	$PM_{10}+O_3max+NO_2av$	1.011 (0.993, 1.030)	0.109	0.220	**1.011** (0.995, 1.031)	**0.073**	**0.002**
Total circulatory system related mortality (people < 65)	PM_{10}	1.042 (0.991, 1.097)	0.055	0.111	1.028 (0.977, 1.082)	0.145	0.023
	$PM_{10}+O_3max$	1.025 (0.971, 1.082)	0.183	0.368	1.003 (0.951, 1.059)	0.451	0.029
	$PM_{10}+O_3max+NO_2av$	1.024 (0.968, 1.084)	0.204	0.409	1.003 (0.947, 1.061)	0.462	0.029
Total respiratory system related mortality (people < 65)	PM_{10}	0.981 (0.928, 1.036)	0.242	0.483	0.980 (0.927, 1.035)	0.229	0.028
	$PM_{10}+O_3max$	0.997 (0.943, 1.055)	0.460	0.921	0.992 (0.938, 1.049)	0.388	0.061
	$PM_{10}+O_3max+NO_2av$	0.981 (0.922, 1.042)	0.265	0.529	0.976 (0.918, 1.037)	0.212	0.073
Other non accidental mortality total (people < 65)	PM_{10}	1.007 (0.989, 1.026)	0.209	0.418	1.011 (0.993, 1.029)	0.122	0.029
	$PM_{10}+O_3max$	1.011 (0.992, 1.030)	0.127	0.256	1.013 (0.994, 1.032)	0.095	0.040
	$PM_{10}+O_3max+NO_2av$	1.013 (0.993, 1.033)	0.103	0.207	**1.016** (0.996, 1.036)	**0.056**	**0.034**
Total non accidental mortality (people > 65)	PM_{10}	**1.021** (1.002, 1.041)	**0.016**	**0.032**	1.009 (0.990, 1.029)	0.172	0.038
	$PM_{10}+O_3max$	**1.018** (0.998, 1.039)	**0.039**	**0.079**	1.006 (0.986, 1.027)	0.267	0.055
	$PM_{10}+O_3max+NO_2av$	**1.019** (0.997, 1.041)	**0.044**	**0.090**	1.008 (0.986, 1.029)	0.241	0.056
Total circulatory system related mortality (people > 65)	PM_{10}	0.998 (0.965, 1.033)	0.464	0.927	1.003 (0.969, 1.039)	0.424	0.178
	$PM_{10}+O_3max$	1.000 (0.965, 1.037)	0.495	0.989	1.008 (0.973, 1.045)	0.329	0.162
	$PM_{10}+O_3max+NO_2av$	0.995 (0.958, 1.034)	0.399	0.797	1.002 (0.965, 1.041)	0.455	0.162
Total respiratory system related mortality (people > 65)	PM_{10}	**1.082** (1.027, 1.140)	**0.002**	**0.004**	**1.064** (1.008, 1.122)	**0.012**	**0.097**
	$PM_{10}+O_3max$	**1.076** (1.019, 1.136)	**0.004**	**0.009**	1.057 (0.999, 1.118)	0.027	0.134
	$PM_{10}+O_3max+NO_2av$	**1.079** (1.019, 1.143)	**0.005**	**0.011**	1.058 (0.997, 1.123)	0.032	0.142
Other non accidental mortality total (people > 65)	PM_{10}	1.021 (0.995, 1.047)	0.060	0.121	0.999 (0.974, 1.025)	0.475	0.247
	$PM_{10}+O_3max$	1.015 (0.988, 1.043)	0.135	0.270	0.993 (0.966, 1.020)	0.300	0.298
	$PM_{10}+O_3max+NO_2av$	1.019 (0.990, 1.048)	0.101	0.204	0.998 (0.970, 1.027)	0.442	0.274

From table 3.7 the reader may observe that for the aggregated data, GLM and GAM show central estimate that indicate a positive association between daily mortality and PM_{10} concentration levels except for circulatory system related mortality where the relationship is opposite, however in most cases (except for a GAM with total circulatory system related mortality including 3 pollutants and perhaps the GLM for total non accidental mortality including only the pollutant of interest) the hypothesis that the coefficient is different than zero would not be rejected. The hypothesis that a model including PM and a model including all other variables except this one are equal, would not be rejected in any case for the GLM. The null hypothesis for GAMs however would be rejected except for circulatory and respiratory mortality, showing that for total non accidental mortality and for mortality causes other than respiratory and circulatory, a model including smooth terms and PM_{10} is preferable to a GLM excluding PM as an explanatory variable.

In the age group of people less than 65 years old, a consistent association of PM_{10} and circulatory system related mortality is observed in models including only this pollutant and ozone. It is to be notice that the *p* values for both hypothesis tests are less than 0.05 and the central estimates using either GAM or GLM range from a 6.2% to a 6.7% increase in circulatory system related mortality per 10 $\mu g/m^3$ increase in PM_{10} levels. In all cases the GAM for this age group would be a preferable model than a linear model as measured by the Delta D statistic, however, only in the two cases previously mentioned the hypothesis that the coefficient is equal to zero would be rejected ($p(t)< 0.02$).

In the older age group, in 50% of the cases (total non accidental mortality and less so in circulatory system related mortality), GAMs including PM_{10} seem not to be significantly different to a GLM not including this pollutant as for the hypothesis that the coefficient found by GAMs is significantly different than zero, it would be rejected in circulatory system related mortality that surprisingly shows a central estimate with a negative influence on mortality by this cause, this finding is consistent when considering GLM. On the other hand, a very significant linear influence of PM_{10} same day average concentrations is observed on respiratory system related mortality with at least an 8.5% increment per 10 $\mu g/m^3$ increase in PM_{10}, the reader should observe that a linear model better fits the data than a model considering smooth terms ($p(Chi)>0.10$ for GAMs). When fitting a model not considering ozone or nitrogen dioxide, a significant association of this pollutant with total non accidental mortality and mortality from causes other that circulatory or respiratory system diseases was observed; the central estimates show a 1.9% and 3.3% increase in baseline mortality per 10 unit increase of air pollution respectively.

For a 3 day lag effect of air pollution in total mortality, the GAMs seem to better describe the data than a GLM not including PM_{10} as measured by p(Chi), except in the case of non accidental deaths not related to respiratory or circulatory system diseases. A consistent statistical association of 3 day lag PM concentration was found when using GLM and GAM to explain total non accidental mortality ($p(t)<0.064$ and $p(chi)<0.043$ in all cases) showing an approximate 1.0 to 1.4% increase in mortality for 10 units increase in PM_{10} average concentration in the third day previous to death. Strong association was observed in the older age group for total non accidental mortality (% change in mortality for 10 unit increase in air pollution between 1.8 and 2.1%) and total respiratory mortality (% change in mortality for 10 unit increase in air pollution between 7.6 and 8.2%) where the prediction power appears to be stronger as measured by the two test statistics. Some association, though less clear, is

observed in the models including 3 pollutants for other non accidental mortality and 2 pollutants for total respiratory mortality.

When looking at GAMs, the smallest *p* value for rejecting the null hypothesis that the coefficient for PM is different than zero is observed in total respiratory system related mortality in people 65 or older with the model considering one pollutant with 6.4% increase in mortality for 10 unit increase in 3 day lag PM_{10} concentration. In the younger age group, the most important association between 3 day lag air pollution and mortality was observed with the 3 pollutant model in other non accidental mortality, and a consistent 1.1% increase in total non accidental mortality per 10 unit increase in air pollution is to be noted.

3.5 Partial Correlation.

The coefficients found previously are those that maximize the log likelihood function (9) in the appendix, assuming a Poisson distribution for mortality counts. Other question that is of interest is to investigate how strong is the association between the daily number of deaths from different causes with the variables included in models 3.8 and 3.9; to begin with, simple product moment correlation coefficients[2] could be observed. To keep consistency with the previous models, mortality variables are in the log scale, the indicator variables for day of the week and the dummy variable for the time of study were excluded from the analysis.

Table 3.9 Correlation Coefficients for Total non Accidental Mortality and Selected Variables.

	ln(namt)	pm10	O3max	NO2av	TMPmin	Pdtmin.3	RH	3pdrh.3
ln(namt)	1.0000							
pm10	0.1959	1.0000						
O3max	0.0766	0.4533	1.0000					
NO2av	0.1803	0.6452	0.5544	1.0000				
TMPmin	-0.2529	-0.5641	-0.1942	-0.4819	1.0000			
pdtmin.3	-0.2716	-0.5963	-0.1767	-0.4897	0.8953	1.0000		
RH	-0.1443	-0.6875	-0.4163	-0.4406	0.4447	0.4065	1.0000	
3pdrh.3	-0.1405	-0.7169	-0.3936	-0.4516	0.3813	0.4648	0.7670	1.0000

From the table above it is observed that all correlation coefficients for the log of total non accidental mortality with a given variable are smaller than 0.28 in absolute value, all three pollutant measures are positively correlated with the log of daily number of deaths (PM_{10} showing the "strongest" linear dependence with a coefficient equal to 0.19), while the atmospheric variables present a negative correlation with this variable. All three pollutant measures are positively correlated with each other with a coefficient between 0.45 and 0.64, and negatively correlated with atmospheric variables the strongest association in the matrix above is found between pdtmin.3 with TMPmin and 3pdrh.3 with RH which is natural because pdtmin.3 is the 3 previous day moving average of minimum temperature and 3pdrh.3 is the 3 previous day moving average of relative humidity.

[2] The reader should remember that the product moment correlation for variable X and Y is the ratio of the covariance between the two variables to the product of the standard deviation for each variable. The coefficient is zero if the two variables are independent but it is not necessarily the case that when the coefficient is zero, the two variables are independent. The product moment correlation coefficient measures linear dependence.

To have a better picture of the dependence of daily mortality with PM_{10} partial correlation coefficients can be computed. Given n variables one may be interested in the regression of any one, upon any subset of the others. Partial correlation of $\{i, j\}$ based on $\{k, l, ..\}$ is related to the regression of i on j, in the presence of variables k, l, …,n as:

$$\rho_{ij;k,l,\ldots,n} = \mathrm{sgn}(b_{ij;k,l,\ldots,n})(b_{ij;k,l,\ldots,n} b_{ji;k,l,\ldots,n})^{1/2} \tag{3.8}$$

Where $b_{ij;k,l,\ldots,n}$ are the numbers that minimize (3.9)

$$E\left[(X_i - b_{ij;k,\ldots,n} X_j - \ldots - b_{in;j,\ldots,n-1} X_n)^2\right] \tag{3.8}$$

For a joint normal distribution, partial and conditional correlations are equal. In this case the interpretation is that if holding other variables fixed reduces the correlation between two variables, it will be inferred that their interdependence arises in part because of the influence of these other variables; if the partial correlation is close to zero, the interdependence of two given variables could be due to the influence of all other variables (those that are held fixed). When the partial correlation of two variables is larger than the original correlation, the conclusion is that the other variables were obscuring the true stronger association between two given variables.

In this study, following (Kendall, 1961) partial correlation coefficients are computed as (3.9) where C_{ij} is the cofactor of the (i, j)th element in the symmetric correlation matrix. All correlation and partial correlation coefficients were computed in Matlab with the program presented in the appendix.

$$\rho_{ij;k,l,\ldots,n} = -\frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} \tag{3.9}$$

Table 3.10 Partial Correlation Coefficients for Total non Accidental Mortality and Selected Variables.

	ln(namt)	pm10	O3max	NO2av	TMPmin	pdtmin.3	RH	3pdrh.3
ln(namt)	-1.000							
pm10	0.014	-1.000						
O3max	-0.006	0.073	-1.000					
NO2av	0.037	0.337	0.424	-1.000				
TMPmin	-0.006	-0.036	-0.005	-0.066	-1.000			
pdtmin.3	-0.091	-0.130	0.119	-0.068	0.851	-1.000		
RH	-0.023	-0.227	-0.112	0.058	0.340	-0.303	-1.000	
3pdrh.3	0.020	-0.306	-0.058	0.031	-0.338	0.338	0.585	-1.000

The partial correlation of the log of non accidental mortality with all other variables is smaller than 3.7×10^{-2} suggesting that any correlation that could be observed between these variables in table 3.9 could be induced by the influence of all other variables. It is also observed that the correlation between ozone and PM with NO2 does not drop dramatically when keeping the values of all other variables fixed suggesting that this correlation could be important. As it is

expected the 3 previous days moving averages for minimum temperature and relative humidity show the highest values for partial correlation with the single day measures of these variables. Since PM_{10} is the variable that this thesis is mainly concerned with, further analysis will present correlation and partial correlation coefficients for this variable and the measures of daily number of deaths considered through this chapter computed in the same way as in tables 3.9 and 3.10.

Table 3.11 Correlation Coefficients and Partial Correlation Coefficients for Mortality and PM_{10} in Toluca.

		PM_{10}		PM_{10} (3 day lag structure)	
	ln(*Y*)	CC	PCC	CC	PCC
Total	non accidental deaths	0.196	0.014	0.209	0.067
	circulatory system related	0.124	−0.056	0.132	−0.019
	respiratory system related	0.186	0.025	0.204	0.028
	other	0.102	0.030	0.105	0.066
< 65 years old	non accidental deaths	0.123	0.044	0.135	0.059
	circulatory system related	0.083	0.031	0.069	0.016
	respiratory system related	0.040	−0.043	0.063	−0.073
	other	0.082	0.052	0.102	0.082
> 65 years old	non accidental deaths	0.180	0.013	0.176	0.044
	circulatory system related	0.099	−0.071	0.075	−0.086
	respiratory system related	0.140	0.012	0.156	0.047
	other	0.064	0.049	0.027	0.018

CC = Pearson Correlation Coefficient. PCC = Partial Correlation Coefficient.

From the table above it is observed that all correlation coefficients for PM_{10} and different causes of mortality in different age groups are smaller than 0.196 for the same day average concentration and 0.209 for a 3 day lag structure, suggesting that a linear association between PM_{10} and mortality, if any, is very weak. Moreover, for the same day average PM_{10} concentration the largest absolute value of the partial correlation coefficient is 5.6×10^{-2} observed in total circulatory system related mortality, and for the 3 day lag structure this number is 8.6×10^{-2}. These numbers being so small suggest that any possible linear association between two given variables could be due to the influence of the other variables in the model.

3.6 Final Remarks.

As measured by Pearson's and partial correlation coefficients, it would be difficult to observe any dependence of daily mortality counts on PM_{10} average concentration in the same day or with a 3 day lag structure with traditional regression techniques due to the assumption of joint normality. GAM and GLM however give some evidence of the influence of PM_{10} on daily mortality because they assume a Poisson distribution for mortality counts. It is to be noted that the estimated effects of air pollution are generally larger when fitting a simple linear model than when taking into account possible non-linear effects of confounding covariates, however GAMs generally fit better the data than GLM except in the case of the older age group where the Deviance statistic is smaller when using the more simple linear model. Taking these considerations into account it is also observable, that PM_{10} levels are more

strongly associated with respiratory disease related mortality, especially in the older age group for both, the model that investigates same day pollution levels and the 3 day lag model. These results are not far from previously reported studies (chapter one).

There are many other possibilities to work with air pollution data and daily mortality counts, for instance, information for more than one city could be available and a combination of these information could be possible or Fourier analysis could be used in the air pollution time series for investigating different time scales. All possibilities would introduce some uncertainty, however, time series analysis of air pollution still are valuable for some reasons, including the fact that they provide important information in identifying whether particles acutely cause illness or death, presumably because persons with underlying heart and lung disease are more at risk. Also, by comparing mortality from day to day within the same population, time series studies are less subject to "ecologic bias" than cohort studies. Time series studies may also provide evidence relevant to scientific questions that support a causal relationship of particles with mortality including: the effects of co-pollutants, cause-of-death-specific pollution effects, and geographic variations in the pollution effects.

The next chapter will present an alternative approach for getting a picture of the uncertainty surrounding air pollution-mortality relationship estimates as perceived by the scientific community and an example for Mexico City will be presented.

APPENDIX 3

A3.1. Exponential Family of distributions

Consider a random variable Y whose probability distribution depends on a single parameter θ. The distribution belongs to the exponential family if it can be written as,

$$f(y;\theta) = s(y)t(\theta)e^{a(y)b(\theta)} \tag{1}$$

Take $a(y) = y$, $s(y) = e^{d(y)}$ and $t(\theta) = e^{c(\theta)}$ then the distribution is said to be in canonical or standard form and $b(\theta)$ is called the natural parameter

$$f(y;\theta) = e^{yb(\theta)+c(\theta)+d(y)} \tag{2}$$

If there are other parameters, in addition to the parameter of interest θ they are regarded as nuisance parameters forming parts of the functions a, b, c and d and they are treated as though they are known. Take for example the Poisson distribution with discrete probability function

$$f(y;\theta) = \frac{\theta^y e^{-\theta}}{y!} \qquad \text{or,} \tag{3}$$

$$f(y;\theta) = e^{y\ln\theta-\theta-\ln y!} \tag{4}$$

in the canonical form. Some properties of the exponential family will be illustrated using the Poisson distribution. First observe that for any exponential distribution with parameter θ:

$$\int f(y;\theta)dy = 1 \qquad \frac{d}{d\theta}\int f(y;\theta)dy = \frac{d}{d\theta}1 = 0 \qquad \int \frac{df(y;\theta)}{d\theta}dy = \int \frac{d^2 f(y;\theta)}{d\theta^2}dy = 0 \tag{5}$$

Now for distributions in the exponential family:

$$f(y;\theta) = e^{a(y)b(\theta)+c(\theta)+d(y)} \tag{6}$$

Applying the results from (5) to (6) we have for the distributions in the exponential family and for the Poisson distribution in particular:

$$\int \frac{df(y;\theta)}{d\theta}dy = 0$$

$$E(a(y)) = E(y) = -c'(\theta)/b'(\theta) = \theta \tag{7}$$

$$\text{var}(a(y)) = \text{var}(y) = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{[b'(\theta)]^3} = \theta \tag{8}$$

For the log-likelihood function

$$l(y;\theta) = a(y)b(\theta) + c(\theta) + d(y) \tag{9}$$

The score vector statistic is

$$U = \frac{\partial l(y;\theta)}{\partial \theta} = a(y)b'(\theta) + c'(\theta) \tag{10}$$

And if this quantity is treated as a random variable, then from (7),

$$E(U) = E(a(y))b'(\theta) + c'(\theta) = 0 \tag{11}$$

The information will be denoted by $\Im$; so applying the formula of the variance of linear transformations of random variables and from 8

$$\Im = \text{var}(U) = \text{var}(a(y))b'(\theta)^2 = \frac{b''(\theta)c'(\theta)}{b'(\theta)} - c''(\theta) \tag{12}$$

Observe that

$$\Im = \text{var}(U) = E(U^2) = -E(U') \tag{13}$$

Consider independent random variables $Y_{1, \ldots,} Y_N$ satisfying the properties of a generalized linear model, we want to estimate parameters β which are related to the Y_i's through $E(Y_i) = \mu_i$ and $g(\mu_i) = \mathbf{x}^T \beta = \eta_i$. Each Y_i will have a log likelihood function as (9) and since the distribution is in the canonical form, the log likelihood function for all the Y_i's is

$$l = \sum_{i=1}^{N} l_i = \sum_{i=1}^{N} y_i b(\theta_i) + \sum_{i=1}^{N} c(\theta_i) + \sum_{i=1}^{N} d(y_i) \tag{14}$$

To obtain the maximum likelihood estimator for the parameter β_j we need

$$\frac{\partial l}{\partial \beta_j} = U_j = \sum_{i=1}^{N} \frac{\partial l_i}{\partial \beta_j} = \sum_{i=1}^{N} \frac{\partial l_i}{\partial \theta_i} \cdot \frac{\partial \theta_i}{\partial \mu_i} \cdot \frac{\partial \mu_i}{\partial \beta_j} \tag{15}$$

Consider each term in the right hand side of (15) separately, then from (9), (7), (8) and the fact that $E(Y_i) = \mu_i$

$$\frac{\partial l_i}{\partial \theta_i} = y_i b'(\theta_i) + c'(\theta_i) = b'(\theta_i)(y_i - \mu_i) \tag{16}$$

$$\frac{\partial \theta_i}{\partial \mu_i} = 1 / \frac{\partial \mu_i}{\partial \theta_i} \tag{17}$$

$$\frac{\partial \mu_i}{\partial \theta_i} = \frac{-c''(\theta_i)}{b'(\theta_i)} + \frac{c'(\theta_i)b''(\theta_i)}{[b'(\theta_i)]^2} = b'(\theta_i)\,\text{var}(Y_i) \tag{18}$$

$$\frac{\partial \mu_i}{\partial \beta_i} = \frac{\partial \mu_i}{\partial \eta_i} \cdot \frac{\partial \eta_i}{\partial \beta_i} = \frac{\partial \mu_i}{\partial \eta_i} x_{ij} \tag{19}$$

Hence,

$$U_j = \sum_{i=1}^{N} \frac{y_i - \mu_i}{\text{var}(y_i)} x_{ij} \left(\frac{\partial \mu_i}{\partial \eta_i} \right) \tag{20}$$

And,

$$\Im_{jk} = E(U_j U_k) = E\left\{ \sum_{i=1}^{N} \left[\frac{y_i - \mu_i}{\text{var}(y_i)} x_{ij} \left(\frac{\partial \mu_i}{\partial \eta_i} \right) \right] \sum_{l=1}^{N} \left[\frac{y_l - \mu_l}{\text{var}(y_l)} x_{lk} \left(\frac{\partial \mu_l}{\partial \eta_l} \right) \right] \right\} \tag{21}$$

Because $E[(y_i - \mu_i)(y_l - \mu_l)] = 0$ for $i \neq l$ since the Y_i's are independent and using $E[(y_i - \mu_i)]^2 = \text{var}(Y_i)$ then,

$$\Im_{jk} = \sum_{i=1}^{N} \frac{E\left[(y_i - \mu_i)^2\right] x_{ij} x_{ik}}{\text{var}(y_i)^2} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 = \sum_{i=1}^{N} \frac{x_{ij} x_{ik}}{\text{var}(y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \tag{22}$$

A3.2. Scoring Procedure for Parameter Estimation.

For the method of **scoring** the vector of estimates of the parameters **b** become

$$\begin{aligned} b^{(m)} &= b^{(m-1)} + \left[\Im^{(m-1)} \right]^{-1} U^{(m-1)} \\ \Im^{(m-1)} b^{(m)} &= \Im^{(m-1)} b^{(m-1)} + U^{(m-1)} \end{aligned} \tag{23}$$

$$\Im_{jk} = \sum_{i=1}^{N} \frac{E\left[(y_i - \mu_i)^2\right] x_{ij} x_{ik}}{\text{var}(y_i)^2} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 = \sum_{i=1}^{N} \frac{x_{ij} x_{ik}}{\text{var}(y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \tag{24}$$

The right hand side of (23) is

$$\sum_{k=1}^{p} \sum_{i=1}^{N} \frac{x_{ij} x_{ik}}{\text{var}(y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 b_k^{(m-1)} + \sum_{i=1}^{N} \frac{(y_i - \mu_i) x_{ij}}{\text{var}(y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right) \tag{25}$$

If W is the N × N diagonal matrix with elements

$$w_{ii} = \frac{1}{\text{var}(y_i)^2} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \tag{26}$$

Then, (24) and (25) become respectively

$$\Im_{jk} = X^T W X \tag{27}$$

$$X^T W z$$

$$z_i = \sum_{k=1}^{p} x_{ik} b_k^{(m-1)} + (y_i - \mu_i)\left(\frac{\partial \eta_i}{\partial \mu_i}\right)$$

With μ_i and $\partial \eta_i / \partial \mu_i$ evaluated at $b^{(m-1)}$ hence the iterative equation can be written as

$$X^T W X b^{(m)} = X^T W z \tag{28}$$

This equation has to be solved iteratively because **z** and **W** depend on **b**. Thus for generalized Linear Models, maximum likelihood estimators are obtained by an iterative weighted least squares procedure. For the Poisson regression with the logarithmic link function equations (26) and (27) become,

$$w_{ii} = e^{x_i^T \hat{\beta}} \tag{29}$$

$$z_i = \hat{\eta} + (y_i / e^{x_i^T \hat{\beta}}) - 1 \tag{30}$$

A3.3 Matlab code for GLM.

```
clear;clc;load namt_20.txt;load pm10_20.txt;n=length(pm10_20)
load O3_20.txt;load no2av_20.txt;load mon_20.txt;load tues_20.txt;
load wed_20.txt;load thr_20.txt;load fri_20.txt;load sat_20.txt;%load sun_20.txt;
;load day_20.txt;load tmp_20.txt;load tmp3_20.txt;
load rh_20.txt;load rh3_20.txt;
xo=ones(n,1);
x1=zeros(n,1);
bo=zeros(15,1);k=0;
X=[xo pm10_20 O3_20 no2av_20 mon_20 tues_20 wed_20 thr_20 fri_20 sat_20 day_20
tmp_20 tmp3_20 rh_20 rh3_20];
eps=1
while eps>0.0001
    eta=X*bo;
    W=diag(exp(eta),0);
    z=zeros(n,1);
    for i=1 : n
        z(i)=eta(i)+(namt_20(i)/exp(eta(i)))-1;
    end
    A=X'*W*X;
    c=X'*W*z;
    %R = chol(A);
    %b = R\(R'\c)
    %[L,U] = lu(A);
    %b= U\(L\c)
    b=A^-1*c
    eps=norm(bo-b)
    bo=b;
    k=k+1
end
```

A3.4 The Sampling Distribution for the Deviance.

From a Taylor expansion of the log likelihood function for an estimated value ***b*** of ***β*** and the fact that if ***b*** is the maximum likelihood estimator of ***β*** (so that U(b)=0)

$$l(\beta) - l(b) = -\frac{1}{2}(\beta - b)^T \Im(b)(\beta - b) \tag{31}$$

Then

$$2[l(b) - l(\beta)] = (\beta - b)^T \Im(b)(\beta - b) \sim \chi_p^2 \tag{32}$$

Where p the degrees of freedom, is given by the number of parameters. Specifically, for the Deviance we have

$$\begin{aligned} D &= 2[l(b_{\max}; y) - l(b; y)] \\ &= 2[l(b_{\max}; y) - l(\beta_{\max}; y)] - 2[l(b; y) - l(\beta; y)] - 2[l(\beta_{\max}; y) - l(\beta; y)] \end{aligned} \tag{33}$$

The first term in square brackets in (33) has the χ_m^2 distribution, the second one χ_p^2 and the last one (v), is a positive constant near zero if the model of interest fits the data almost as well as the saturated model. Hence $D \sim \chi_{m-p,v}^2$ where v is the non-centrality parameter.

For the Poisson model we have

$$D = 2\left\{\sum_{i=1}^{n} y_i \ln(y_i/\mu) - (y - \mu)\right\} \tag{34}$$

A3.5. Derivation of the Local Scoring Algorithm for Generalized Additive Models.

Let $\eta_i = f_0 + \sum_{j=1}^{d} f_j(X_j)$ and consider the log likelihood as a function of $\boldsymbol{\eta}$. Let H be the space of all functions f that have two continuous derivatives and call a function smooth if it is in H. consider the following optimization problem. Find $f_j \in H_j$ to maximize:

$$l(\eta, y) - 1/2 \sum_{j=1}^{d} \lambda_j \int \{f_j''(x)\}^2 dx \tag{35}$$

Where $\lambda_j \geq 0$ are smoothing parameters. The solution is a natural cubic spline interpolant (see for example Green & Silverman, 1994), that is, each coordinate function is a cubic spline[3], if the expression above is parametrized by the evaluation of the cubic splines $f_j(x)$ at the observed points $x_{ij}, \ldots, x_{nj}$ then the problem (3.4) may be stated as

[3] Suppose we are given real numbers $t_1, \ldots, t_n$ on some interval $[a, b]$, such that $a < t_1 < t_2 < \ldots < t_n < b$. A function f defined on $[a, b]$ ia a cubic spline if:

1. in each interval (a, t_1), (t_1, t_2), (t_2, t_3), . . ., (t_n, b) f is a cubic polynomial;
2. the polynomial pieces fit together at the points t_i in such a way that g itself and its first and second derivatives are continuous at each t_i and hence in the whole $[a, b]$

$$l(\eta, y) - \tfrac{1}{2}\sum_{j=1}^{d} \lambda f_j^T K_j f_j, \tag{36}$$

Where K is a penalty matrix defined as $K = \Delta^T C^{-1} \Delta$, such that $h_i = x_{i+1} - x_i$ for $i = 1, 2, \ldots,$ n-1. Δ is a tri-diagonal (n-2) × n matrix with $\Delta_{i,i} = 1/ h_i$, $\Delta_{i,i+1} = -(1/ h_i + 1/ h_{i+1})$, $\Delta_{i,i+2} = 1/ h_{i+1}$, and a symmetric tri-diagonal (n-2) × (n-2) matrix C with $c_{i-1,i} = c_{i,i-1} = h_i / 6$, $c_{i,i} = (h_i + h_{i+1})/3$.

If $u = \partial l/\partial \eta$ and $W = -\partial l/\partial \eta \eta^T$ is a diagonal matrix with elements $w_{ii} = \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 (V_i)^{-1}$ then the Fisher scoring iterations become

$$\begin{pmatrix} W^k + \lambda_1 K_1 & W^k & \ldots & W^k \\ W^k & W^k + \lambda_2 K_2 & \ldots & W^k \\ \ldots & \ldots & \ldots & \ldots \\ W^k & W^k & \ldots & W^k + \lambda_d K_d \end{pmatrix} \begin{pmatrix} f_1^{k+1} - f_1^k \\ f_2^{k+1} - f_2^k \\ \ldots \\ f_d^{k+1} - f_2^k \end{pmatrix} = \begin{pmatrix} u^k - \lambda_1 K_1 f_1^k \\ u^k - \lambda_2 K_2 f_2^k \\ \ldots \\ u^k - \lambda_2 K_2 f_2^k \end{pmatrix} \tag{37}$$

By taking $S_j = (W + \lambda_j K_j)^{-1} W$ and $z = \eta + W^{-1} u$ then (36) can be written as

$$\begin{pmatrix} I & S_1 & \ldots & S_1 \\ S_2 & I & \ldots & S_2 \\ \ldots & \ldots & \ldots & \ldots \\ S_d & S_d & \ldots & I \end{pmatrix} \begin{pmatrix} f_1^{k+1} \\ f_2^{k+1} \\ \ldots \\ f_d^{k+1} \end{pmatrix} = \begin{pmatrix} S_1 z^k \\ S_2 z^k \\ \ldots \\ S_d z^k \end{pmatrix} \tag{38}$$

And finally

$$\begin{pmatrix} f_1^{k+1} \\ f_2^{k+1} \\ \ldots \\ f_d^{k+1} \end{pmatrix} = \begin{pmatrix} S_1 \left(z^k - \sum_{j \neq 1} f_j^{k+1} \right) \\ S_2 \left(z^k - \sum_{j \neq 2} f_j^{k+1} \right) \\ \ldots \\ S_d \left(z^k - \sum_{j \neq d} f_j^{k+1} \right) \end{pmatrix} \tag{39}$$

A3.6. Matlab Code for Correlation and Partial Correlation Coefficients Calculation.

```
clear;clc;load lnnamt.txt;load pm10.txt;n=length(pm10)
load O3.txt;load no2av.txt;load mon.txt;load tues.txt;
load wed.txt;load thr.txt;load fri.txt;load sat.txt;load sun.txt;
;load day.txt;load tmp.txt;load tmp3.txt;
load rh.txt;load rh3.txt;
X=[lnnamt pm10 O3 no2av  tmp tmp3 rh rh3];

Ro = corrcoef(X);
a=length(Ro)
for i=1:a
    for j=1:a
        R=Ro;
        R(i,:)=[];
        R(:,j)=[];
        CofR(i,j)=det(R)*(-1)^(i+j);
    end
end
for i=1:a
    for j=1:a
        pro(i,j)=-CofR(i,j)/(CofR(i,i)*CofR(j,j))^0.5;
    end
end
```

Chapter 4. Expert Judgment as a Tool for Assessing Health Effects of Air Pollution.

Long term effects of PM on human health are investigated by epidemiologists with the Cox Proportional Hazards Model (chapter two); some statistical techniques available for investigating short term effects of air pollution in mortality were exemplified for Toluca City in the previous chapter. In chapter four the classical method for expert judgment will be set forward as a tool for investigating both long term and short term health effects of air pollution and for characterizing the uncertainty surrounding published health effect estimates.

4.1. An Introduction to Expert Judgment.

Expert judgment is performed for multiple reasons. For example, in many cases, there might not be sufficient data available to investigate a certain phenomenon, but experts could have an "idea" of the behavior of it; or there could be different opinions in the scientific community regarding the "true" outcome of a certain experiment; in both cases decision makers would like to reach an agreement regarding the behavior of the phenomenon that is being investigated. One thing that could be done is take the different points of view of the experts and combine them in a certain way to get an outcome that possesses enough scientific validity to be used for taking decisions.

According to Cooke (1991), any methodology of science (and structured expert judgment is not an exception) must aim at rational consensus. There are some methodological principles that structured expert judgment must meet to achieve rational consensus.

a. Reproducibility. It must be possible for other scientists to review and if necessary reproduce an experiment (calculations).
b. Scrutability/Accountability. In a scientific report, sources of data and instruments for measuring and performing calculations must be identified. In the case of expert judgment, experts' names and assessments, and all processing tools should be subject to empirical quality controls.
c. Empirical control. Scientific statements and scientific theories should be feasible in principle. It is recognized that theories can never be conclusively verified but at least it should be possible in principle to discover a reproducible conflict with observations, if the theory is in fact false. So, it must be possible (in principle) to evaluate expert probabilistic opinion on the basis of possible observation.
d. Neutrality. The method for combining/evaluating expert opinion should encourage experts to state their true opinion.
e. Fairness. All experts are treated equally, prior to processing the results of observation.

The classical model for expert judgment has been introduced in (Cooke, 1991) and applied in many risk and reliability studies. This model for combining expert judgments bears its name because of a strong analogy with classical hypothesis testing. In the following section it will be described the basic model for the case where experts asses their uncertainty for quantities taking values in a continuous range. The scientific foundation for subjective probability comes from the theory of rational decision-making; hence, the main aim of the method is to provide the basis for achieving rational consensus.

4.2. The Classical Method for Expert Judgment.

The first step in the classical model for expert judgment is to elicit questions on variables for which the experts provide a number of quantiles from their uncertainty distribution; the experts are asked to assess physical quantities which could be hypothetically measured in experiments. The classical model takes the data from each expert and constructs a weighted combination of expert's probability assessments. These weights are based on two key performance measures, calibration and information which are assessed on variables whose true values are known post hoc (though not known to the experts at the time of assessment). Calibration corresponds to statistical likelihood. In the language of statistics, this is the "p-value" at which we would reject the hypothesis that a given experts' probabilistic statements are true. Thus, low values for the calibration score (near zero) indicate low support for the hypothesis that the experts' probability statements are accurate; high values (near one) indicate high support for this hypothesis. Information or informativeness measures the degree to which the experts' distributions are concentrated.

The weights are based on the theory of proper scoring rules[4] and satisfy a proper scoring rule constraint. This means that an expert receives his/her maximal expected score by, and only by, stating his/her true beliefs. There is thus no advantage in trying to mask one's assessments so as to achieve maximal score. The weights are proportional to the product of the calibration and information scores, if the calibration scores exceeds a "significance level" cutoff, which may be found by optimization.

The classical model computes "performance based" weighted combinations, but also uses the performance measures to assess the quality of other combinations. In particular, the performance of the equal weight combination is assessed. Generally, the combination exhibiting the best calibration and informativeness is recommended. A detailed explanation of these notions is found in Appendix B

4.3. Application of Expert Judgment for Air Pollution Health Effects in Mexico City.

A group from researchers from Delft University of Technology, Harvard Center for Risk Analysis and the Department of Environmental Health of the National Public Health Institute of Finland was put together as part of the Harvard-MIT Mexico city project. An elicitation of Mexico City air pollution experts was carried on during March 15th to 21st 2004, the purpose of this elicitation was to get a picture of the uncertainty, as perceived by Mexican experts, in the mortality response to air pollution by PM. Six Mexican experts participated in the experiment, they are presented in table 4.1. It is important to point out that the order in which they are presented does not correspond to the code they were assigned for the analysis of the data.

[4] S**coring** is assigning a numerical value to probability assessments on the basis of observation. A scoring rule is called strictly proper if a subject receives his best expected score if and only if his stated assessment corresponds to his true opinion. See for example Cooke 1991, chapter 9.

Table 4.1 Mexico City Air Pollution Experts.

Msc. Matiana Ramírez Aguilar

Surgery Physician from the National Polytechnic Institute, Mexico (ESM-IPN). Msc. Scholar Medicine (ESM-IPN). Master in Environmental Health (University of California Los Angeles). PhD Candidate in Epidemiology Sciences (INSP). Associate Researcher level A form CISP - SPHM.

Dr. Carlos Santos-Burgoa

General Director of Equity and Development in Health, Sub secretary of Quality and Innovation, Health Department, Mexico, Medical Doctor graduated from the National Autonomous University of México. Masters in Public Health and PhD, in Environmental and Occupational Epidemiology from the Johns Hopkins School of Hygiene and Public Health. He was Chair of Epidemiology and Biostatistics and later on for six years Dean of The School of Public Health of Mexico. He was the General Director of Environmental Health and currently is the General Director for Equity and Health Director at the Secretariat of Health in Mexico. He was full professor of epidemiology and Director of the Global Health Program at the University of Michigan School of Public Health, where he is currently associate professor.

Dr. Mauricio Hernandez Avila

Recently appointed General Director of the National Institute of Public Health in Mexico (INSP), former Director of the centre for Research in Public Health (INSP), Titular Researcher level C in INSP, National Researcher level II. Professor and Associated Researcher at The Rollins School of Public Health at Emory University, he is Surgery Physician from the National Autonomous University of Mexico (UNAM), Specialty in Applied Mathematics. (IIMASS-UNAM), Master degree in Epidemiology (Harvard School of Public Health), PhD in Epidemiology (Harvard School of Public Health).

Msc. Castillejos Salazar Margarita B.

Msc. In Medical Demographics from the University of London. Since 1975 she is research professors, Level "C" in the Department of Health Attention, at Universidad Autónoma Metropolitana, Unidad Xochimilco. She has conducted since 1985, several research projects on air pollution, particularly in health effects on children, with financial support from Harvard University, El Colegio de México, el National Council for Science and Technology, el CONSERVA, la Comisión Ambiental Metropolitana (CAM), and others, author of 19 scientific documents for national and international journals. Awarded with the *Prize "Matilde M. de Santos"* by the Mexican Foundation for Health for her research paper *"Effects of Ambient Ozone on Respiratory Function and Symptoms in Schoolchildren in Mexico, City",* recipient of the *Fullbrigth* schoolarship from Harvard University and member of the National System of Researchers. She was a member of the work group that performed the Revision to the Official Norms NOM-020-SSA1-1993, NOM-025-SSA1-1993, and in the project for the Official Mexican Norm for Particles Less than 2.5 microns (PM2.5), from an invitation of the General Direction of Environmental Health. From 2000 to 2001 directed the System for Integral Development of the Family (DIF-D.F.), nowadays she collaborates as assessor of the Government of the Federal District, in environmental issues.

Dr. Álvaro R. Osornio Vargas

A researcher at the *Instituto Nacional de Cancerología* (National Cancerology Institute) and a research associate with the Environmental Health Department at UNAM's Environmental Sciences Program, Dr. Osornio is a medical doctor who graduated from the Faculty of Medicine at UNAM. He has experience as a pathologist and has a Ph.D. in Basic Biomedical Research. His main area of research concerns the mechanisms of damage interceding in the effects produced by environmental contaminants such as particulates. This focus enabled a collaboration of over 10 years with the NIH's National Institute of Environmental Health Sciences, which was partially supported by the Fogarty Center. Dr. Osornio has recently published his results on the toxic effects of particulate contaminants in Mexico City and the impact of their size and composition. In addition, he is an enthusiastic popularizer on environmental issues, particularly with children and adolescents. He belongs to the *Sistema Nacional de Investigadores* (National System of Scientific Researchers), and is a member of the Mexican Academy of Sciences and the American Thoracic Society.

Dr. Victor Hugo Borja Aburto

M.D. Universidad Autonoma Metropolitana Xochimilco, Ph.D., Epidemiology. University North Carolina at Chapel Hill, Coordinator at Coordination for health at work, Social Security Mexican Institute, Mexico. Author and coauthor of numerous papers on air pollution and mortality in the MCMA, director of 12 students in epidemiology and environmental health at master's level and one doctoral researcher. He is interested in the areas of Reproductive and Environmental Epidemiology and Health Effects of Air Pollution. He has directed 12 students in epidemiology and environmental health at the master level and one doctoral student. He has published 18 papers in different national and international journals in the past 3 years.

During the interviews, the protocol previously developed by the research team was used; a summary of this protocol is presented in appendix A4.2. The tables and figures presented in this report are the output generated by EXCALIBUR, the software developed at the department of mathematics of Delft University of Technology for performance based combination of expert judgments.

The part of the protocol that will be used in this chapter includes 9 questions of interest (1 to 9 in appendix A4.2) that are variables for which a distribution is desired to be obtained from the experts' combined opinion and 12 calibration questions, that are used for measuring expert's performance (10-21 in appendix A4.2).

Dr. Leonora Rojas from the National Institute of Ecology in Mexico made individual appointments with each expert and a maximum of 2 interviews per day were conducted during the week of the 15th to 21st of March 2004 with no interviews performed on Wednesday 17th and Saturday 20th. All interviews, except those of Dr. Álvaro R. Osornio Vargas and Dr. Mauricio Hernandez Avila were performed at the office provided by the National Institute of Ecology in Periférico 5000, Col. Insurgentes Cuicuilco, C.P. 04530, Delegación Coyoacán, México D.F.

During the interviews a maximum of 3 persons plus the expert were present. All the interviews were conducted by Prof. Cooke from Delft University of Technology who asked the questions. After the expert provided the 5th, 25th, 50th, 75th and 95th percentile of their distribution, one other member of the team (Andrew Wilson from Harvard School of Public Health, Jouni Tuomistu from the Department of Environmental Health in the National Public Health Institute, Finland or Oswaldo Morales Napoles from Delft University of Technology) asked questions regarding the experts' rationale to get to his/her estimates. The purpose of doing so was on one hand, to gain insight in the experts' true opinion and on the other, to help the expert think about all the factors that to his/her opinions were relevant for a given answer. The results presented here are preliminary in the sense that the answers to all of the calibration questions are not currently available. Thus it is of course possible that the results will change when the full set of calibration data becomes available.

4.3.1 Calibration and Information.

Table 4.2 below, shows the calibration and information scores for the six experts in this study. The first column gives the expert number; the second column gives the calibration score. The ratio of highest to lowest score is about 3×10^5. It will be noted that only expert 3 had a score corresponding to a p-value above 5%. Expert 4's score is marginal, and the others are quite low. Calibration scores in the order 0.001 would fail to confer the requisite level of confidence in the results. The information scores for all items and for calibrations items are shown in columns 3 and 4 respectively. It will be noted that the overall information scores are quite similar, within a factor 1.5.

For the variables with realizations, the differences are a bit larger, but the overall pattern is similar. Note that the expert with the best calibration score (nr 3) also has the lowest information score for the calibration variables. This is a recurrent pattern. This expert also has one of the highest scores for overall informativeness. This might lead one to question the representativeness of the calibration questions; however the differences are not large.

Table 4.2 Calibration and Information Scores.

```
Case name : MEXICANEXPERTSforthesis120-7-2004              CLASS version W4.0
_____________________________________________________________________________
Results of scoring experts
_____________________________________________________________________________
   Nr.|Calibr.   |Mean relat|Mean relat|Numb|UnNormaliz|
      |          | total    |realizatii|real|weight    |
______|__________|__________|__________|____|__________|
     1|  0.007064|     0.971|     1.156|  12|  0.008169|
     2|8.336E-007|      1.44|     1.376|  12|1.147E-006|
     3|   0.08608|     1.131|    0.8659|  12|   0.07453|
     4|   0.01767|     1.278|     0.885|  12|   0.01564|
     5| 0.0009732|     1.207|     1.077|  12|  0.001048|
     6|3.754E-007|     1.015|     1.652|  12|  6.2E-007|
_____________________________________________________________________________
_____________________________________________________________________________
                                                        (c) 1999 TU Delft
```

The fifth column gives the number of calibration variables and the last column gives the "un-normlaized weight"; this is the product of columns 2 and 4. If this column were normalized and used to form weighted combinations, experts 1, 3, 4 and 5 would be influential with (8.2, 75, 15.7 and 1.1 per cent respectively).

To get a picture of the degree of homogeneity within the expert group, range graphs showing all assessments per item are useful. Figure 4.1 shows all expert assessments for the first 9 variables. For each expert, the upper and lower quantiles are given as " [---] "; the 25% and 75% quantiles are given as " < -- > " , and the median is given as " | " . The intrinsic range is shown below the expert assessments. In the appendix the range graphs for all items, including the equal weight decision maker are given.

Figure 4.1 Range Graphs for Variables of Interest in the Mexican Air Pollution Experts Elicitation.

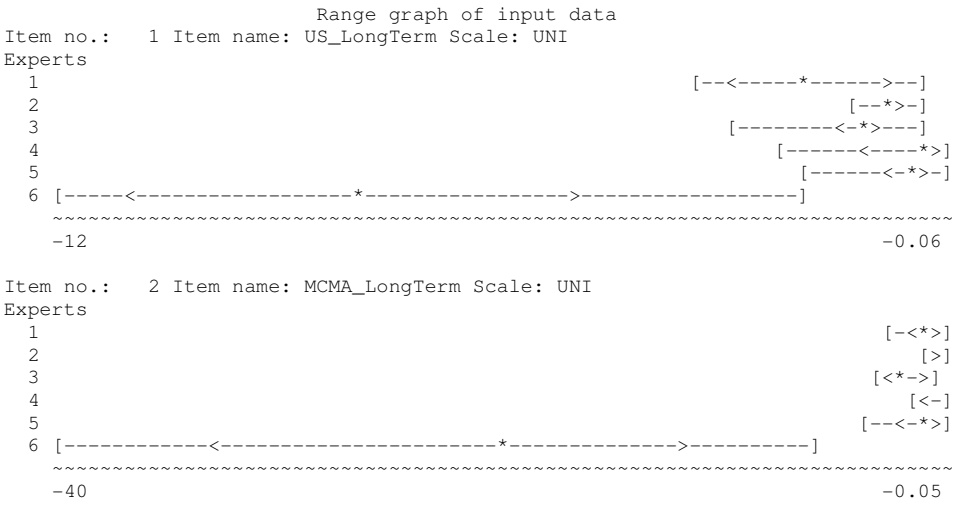


Figure 4.1 Range Graphs for Variables of Interest in the Mexican Air Pollution Experts Elicitation (Continue).

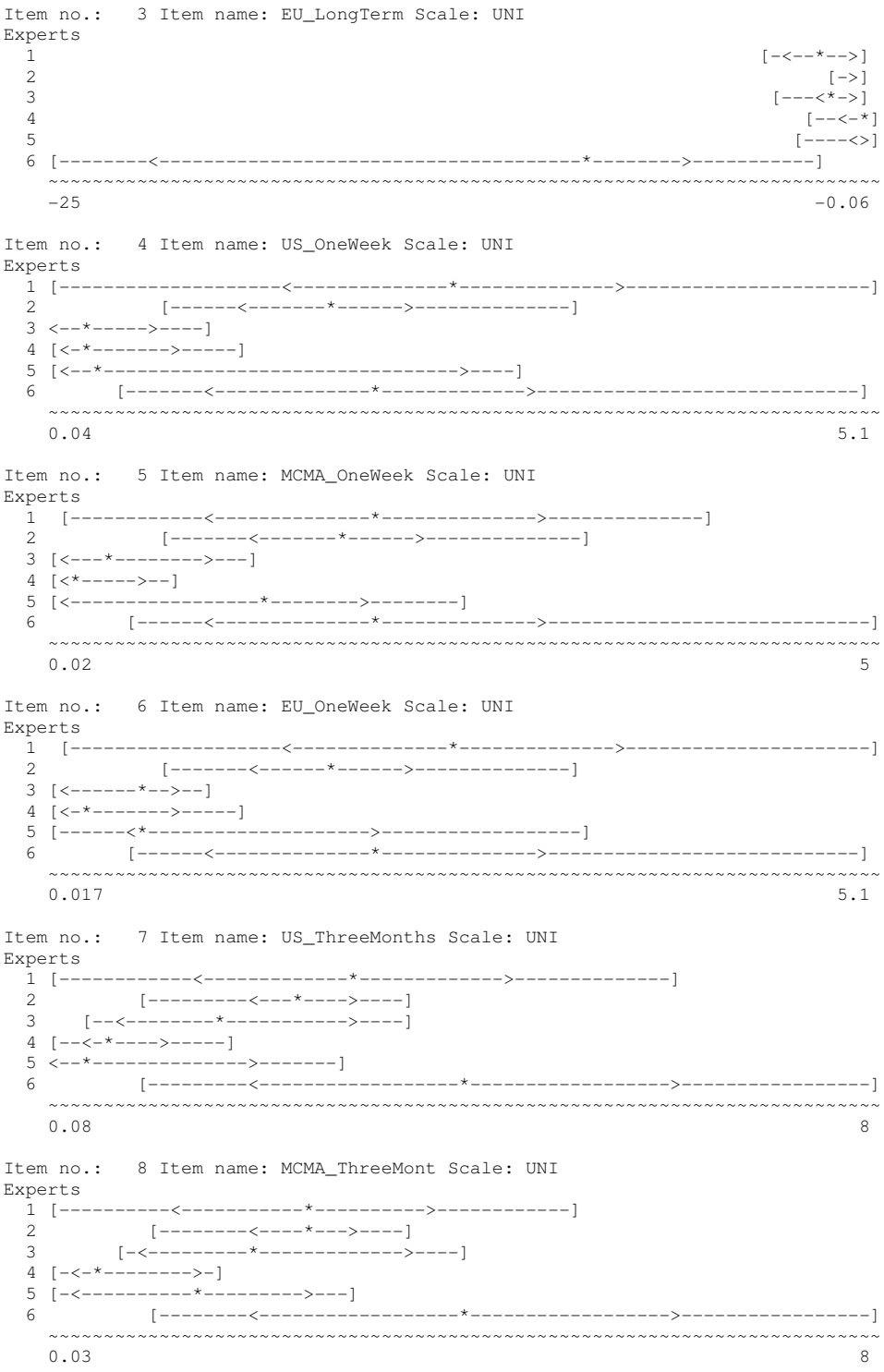
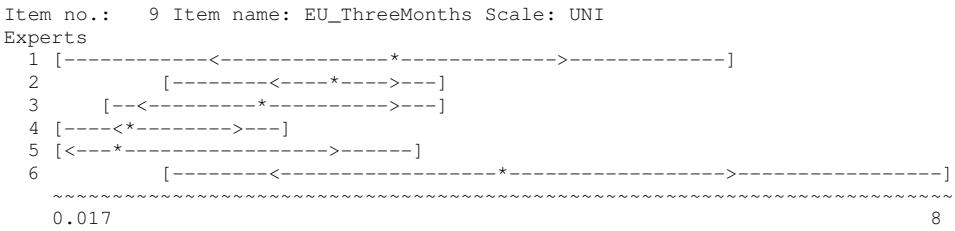


Figure 4.1 Range Graphs for Variables of Interest in the Mexican Air Pollution Experts Elicitation (Continue)



The predominant picture is that the experts' central 90% confidence, generally show considerable overlap. To appreciate Figure 1 numerical table 4.3 compares each expert to the "EWDM"; that is the equal weight decision maker. In discrepancy analysis, the relative information of each expert with respect to the Decision Maker is computed per item. These scores are averaged over all variables (column 6) and calibration variables only (column 7).

Table 4.3 Discrepancy Analysis Relative to Equal Weight Decision Maker.

Case name : MEXICANEXPERTSforthesis120-7-2004 CLASS version W4.0

Results of scoring experts and Relative Information to the DM
Bayesian Updates: no Weights: equal DM Optimisation: no
Significance Level: 0 Calibration Power: 1

Nr.	Mean relat total	Mean relat realizatii	Numb real	UnNormaliz weight	Rel.Inf to total	Rel.Inf to realiz.
1	0.971	1.156	12	0.008169	0.4867	0.5673
2	1.44	1.376	12	1.147E-006	0.9532	1.165
3	1.131	0.8659	12	0.07453	0.6174	0.6134
4	1.278	0.885	12	0.01564	0.648	0.5829
5	1.207	1.077	12	0.001048	0.5236	0.5581
6	1.015	1.652	12	6.2E-007	0.9696	1.223

(c) 1999 TU Delft

It may be seen that the relative information with respect to the equal weight combination is generally one half of the information in the experts' individual assessments.

4.3.2 Combination Schemes.

In this exercise, experts give their uncertainty assessments on calibration variables in the form of 5%, 25%, 50%, 75% and 95% quantiles. To combine all experts' assessments into one uncertainty assessment on each calibration variable there are three combination schemes. The combined distributions are weighted sums of the individual experts' distributions, with non-negative weights adding to one. Different combination schemes are distinguished by the method according to which the weights are assigned to densities. These schemes are designated "decision makers". Two kinds of decision makers are described below.

4.3.2.1 Equal Weight Decision Maker.

The equal weight decision maker (table 4.4) results by assigning equal weight to each density. If E experts have assessed a given set of variables, the weights for each density are 1/E; hence for variable i in this set the decision maker's density is given by:

$$f_{ewdm,i} = \left(\frac{1}{E}\right) \sum_{j=1\cdots E} f_{j,i} \tag{1}$$

where $f_{j,i}$ is the density associated with expert j's assessment for variable i.

Table 4.4 Equal Weight Decision Maker.

```
Case name : MEXICANEXPERTSforthesis120-7-2004              CLASS version W4.0
_____________________________________________________________________________
Results of scoring experts
 Bayesian Updates: no      Weights:  equal      DM Optimisation:  no
 Significance Level:          0    Calibration Power:          1
_____________________________________________________________________________
   Nr.| Calibr.  |Mean relat|Mean relat|Numb|UnNormaliz|Normaliz.w|Normaliz.w
      |          | total    |realizatii|real|weight    |without DM|with DM
______|__________|__________|__________|____|__________|__________|_________
     1|  0.007064|     0.971|     1.156|  12|  0.008169|    0.1667|   0.04149
     2|8.336E-007|      1.44|     1.376|  12|1.147E-006|    0.1667|5.827E-006
     3|   0.08608|     1.131|    0.8659|  12|   0.07453|    0.1667|    0.3785
     4|   0.01767|     1.278|     0.885|  12|   0.01564|    0.1667|   0.07942
     5| 0.0009732|     1.207|     1.077|  12|  0.001048|    0.1667|   0.00532
     6|3.754E-007|     1.015|     1.652|  12|  6.2E-007|    0.1667|3.149E-006
  EWDM|    0.2541|     0.474|    0.3838|  12|   0.09753|          |    0.4953
_____________________________________________________________________________
_____________________________________________________________________________
                                                        (c) 1999 TU Delft
```

Table 4.4 shows 8 columns in total, the first 6 are the same as in table 4.2, the decision maker is the 7th expert identified as "EWDM". As it may be seen in the table above, the third column shows that the combined opinion of the six experts, i.e. EWDM is better calibrated than every expert individually. Weights for each expert and for the decision maker are shown in the last three columns. Since the equal weight combination has been chosen in this case, column 7 displays the same weight for all experts except the decision maker, and finally column 8 shows that when normalizing the numbers in column 6 the decision maker and expert number 3 would share almost all the weight.

4.3.2.1 Global Weight Decision Maker.

This kind of decision maker is from the class of performance based decision makers that are those where the weights are based on the experts' performance. Two performance based decision makers are supported in the software EXCALIBUR. The "global weight" decision maker uses average information over all calibration variables and computes one set of weights for all items. The "item weight" decision maker constructs weights for each item separately, using the experts' information scores for the given item, rather than the average information score.

In this study the global and items weights do not differ, and we focus on the former. The global weight decision maker (Table 5) uses performance based weights which are defined, per expert, by the product of expert's calibration score and his(her) overall information score on calibration variables, and by an optimization[5] procedure. For expert *j*, the same weight is used for all variables assessed. Hence, for variable *i* the global weight decision maker's density is:

$$f_{gwdm,i} = \frac{\sum_{j=1\cdots E} w_j f_{j,i}}{\sum_{j=1\cdots E} w_j} \qquad (2)$$

Table 4.5 Global Weight Decision Maker.

Case name : MEXICANEXPERTSforthesis120-7-2004 CLASS version W4.0

Results of scoring experts
Bayesian Updates: no Weights: global DM Optimization: yes
Significance Level: 0.08608 Calibration Power: 1

Nr.	Calibr.	Mean relat total	Mean relat realizatii	Numb real	UnNormaliz weight	Normaliz.w without DM	Normaliz.w with DM
1	0.007064	0.971	1.156	12	0	0	0
2	8.336E-007	1.44	1.376	12	0	0	0
3	0.08608	1.131	0.8659	12	0.07453	1	0.5
4	0.01767	1.278	0.885	12	0	0	0
5	0.0009732	1.207	1.077	12	0	0	0
6	3.754E-007	1.015	1.652	12	0	0	0
GWDM	0.08608	1.131	0.8659	12	0.07453		0.5

(c) 1999 TU Delft

In this case, all experts with a calibration score less than the significance level found by the optimization procedure are unweighted as reflected by the zeros in column 5.

4.3.3 Robustness

Robustness analysis addresses the question, to what extent the results of the study would be affected by loss of a single expert or calibration variable in each case the total relative information with respect to the background measure, calibration and the total relative information with respect to the original decision maker are computed. Robustness is an issue whenever we optimize performance (robustness analysis for the equal weight decision maker is omitted). Tables 4.6 and 4.7 show the robustness analysis for the global weight decision maker; that is, they show how the scores would change if calibration variables (table 4.6) or experts (table 4.7) were removed from the analysis one at a time.

[5] For each value of α it is defined a decision maker dm_α computed as a weighted linear combination of the experts whose calibration score exceeds α. dm_α is scored with respect to calibration and information. The weight which this dm_α would receive if he were added as a "virtual expert" is called the "virtual weight" of dm_α. The value of α for which the virtual weight of dm_α is the greatest is chosen as the cut-off value for determining the unweighted expert.

Columns 3 and 4 from table 4.7 show that there is no significant gain or loss in the decision maker's information scores with respect to the background measure when excluding a given item one at the time. In most of the cases there would be a gain in the calibration score when

Table 4.6 Robustness Analysis by Items.

Case name : MEXICANEXPERTSforthesis120-7-2004 CLASS version W4.0

Robustness analysis on seed Items
Bayesian Updates: no Weights: global DM Optimization: yes
Significance Level: 0.0861 Calibration Power: 1.0000

Nr.	Id of excl. item	Rel.info/b total	Rel.info/b realizatii	Calibr.	Rel.info/o total	Rel.info/o realizati
1	Days2003exceed	1.14	0.8587	0.117	0	0
2	Days1995exceed	1.143	0.865	0.117	0	0
3	Days2003below	1.165	0.9044	0.2428	0	0
4	Days1995below	1.164	0.9022	0.2428	0	0
5	Hours2003excee	1.116	0.8161	0.06145	0	0
6	Hours1995excee	1.115	0.8145	0.117	0	0
7	Hours2003below	0.9397	0.6181	0.4742	0.4987	0.4773
8	Hours1995below	0.7877	0.6589	0.1982	0.7669	0.6736
9	NAD2000highFD	1.121	0.825	0.2428	0	0
10	NAD2000lowFD	0.9073	0.5978	0.05177	0.3705	0.3632
11	NAD2000highSM	0.9883	0.684	0.2428	0.1848	0.1702
12	NAD2000lowSM	1.164	0.9022	0.117	0	0
13	None	1.131	0.8659	0.08608		

(c) 1999 TU Delft

Table 4.7 Robustness Analysis by Experts.

Case name : MEXICANEXPERTSforthesis120-7-2004 CLASS version W4.0

Robustness analysis on Experts
Bayesian Updates: no Weights: global DM Optimisation: yes
Significance Level: 0.08608 Calibration Power: 1

Nr.	Id excl.exp	Rel.info/b total	Rel.info/b realizatii	Calibr.	Rel.info/o total	Rel.info/o realizati
1	1	1.129	0.8659	0.08608	0	0
2	2	1.006	0.648	0.08608	0	0
3	3	0.8029	0.6421	0.06288	0.871	0.8629
4	4	1.122	0.8517	0.08608	0	0
5	5	1.113	0.8363	0.08608	0	0
6	6	0.8032	0.7902	0.08608	0	0
7	None	1.131	0.8659	0.08608	0	0

(c) 1999 TU Delft

removing a given item. The largest gain in calibration would be achieved by removing question 16 in appendix A4.1.

A similar picture as in table 4.6 is observed when analyzing columns two and three from table 4.7; for the calibration score the most relevant comment is that there would be a loss of calibration in the decision maker by removing expert 3 from the analysis.

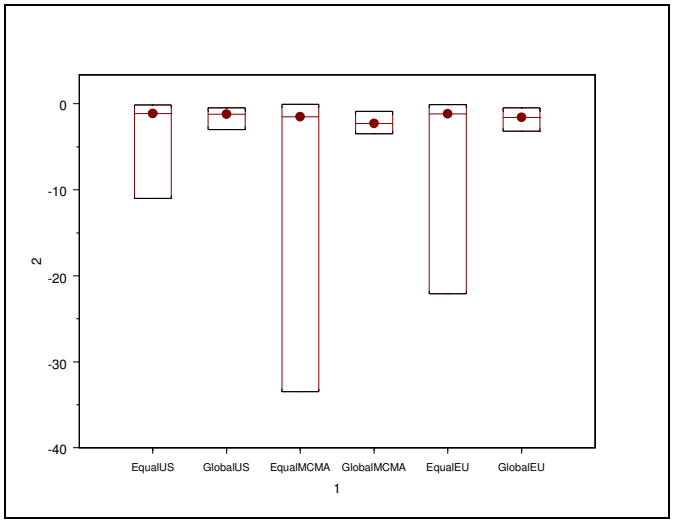
A relevant comparison is between the last two columns of table 4.3 and tables 4.6 and 4.7. We see that the effect of "perturbing" the model by removing an expert or an item, is small

relative to the differences among the experts themselves. In case of removing expert 3, there is a significant change, but even this is within the inter-expert differences.

4.3.4 Further Issues on Combination Schemes.

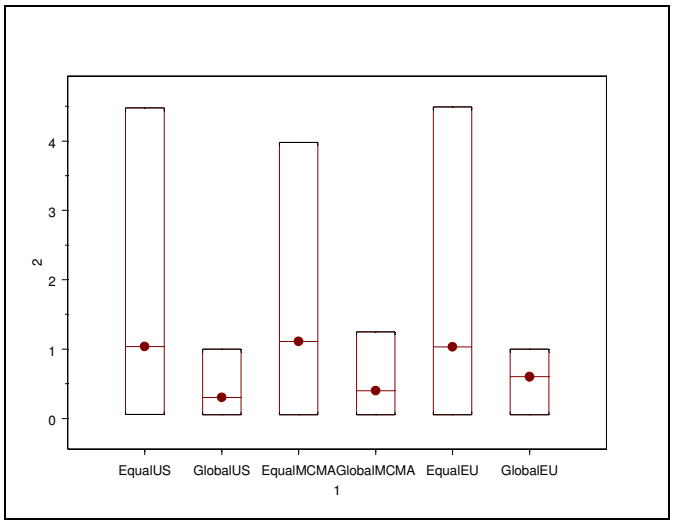
Figure 4.2 shows the median, 5th and 95th percentile of the estimates for long term effects in mortality as assessed in questions 1, 2 and 3 of the elicitation protocol using two different combination schemes (see also Table A4.4), i.e. Global and Equal weight decision makers. The distribution generated with the equal weight decision maker, expresses the large disagreements among experts as to the lower percentile of the distribution (largest health effect observable), than that generated by optimization. The plot also shows that the differences in the median estimate generated with both combination schemes are not as large as those for the lower percentile estimate.

Figure 4.2 Median, 5th and 95th Quantiles for the Long Term Effects of Air Pollution in Mortality



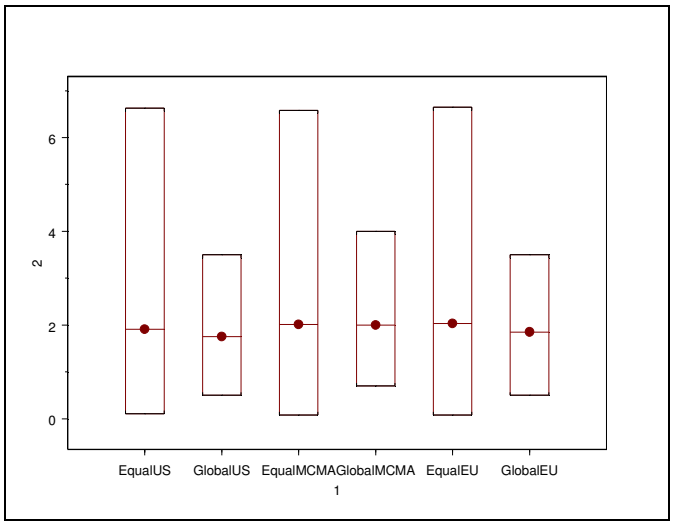
Data generated in Excalibur with results from the Mexico City Elicitation.

Figure 4.3 Median, 5th and 95th Quantiles for One Week Effects of Air Pollution in Mortality



Data generated in Excalibur with results from the Mexico City Elicitation.

Figure 4.4 Median, 5th and 95th Quantiles for Three Month Effects of Air Pollution in Mortality



Data generated in Excalibur with results from the Mexico City Elicitation.

Figures 4.3 and 4.4 show approximately the same trend as figure 4.2 i.e. consensus about the median but larger differences concerning the upper 95th percentile of the uncertainty distribution that correspond to the larger health effect observable.

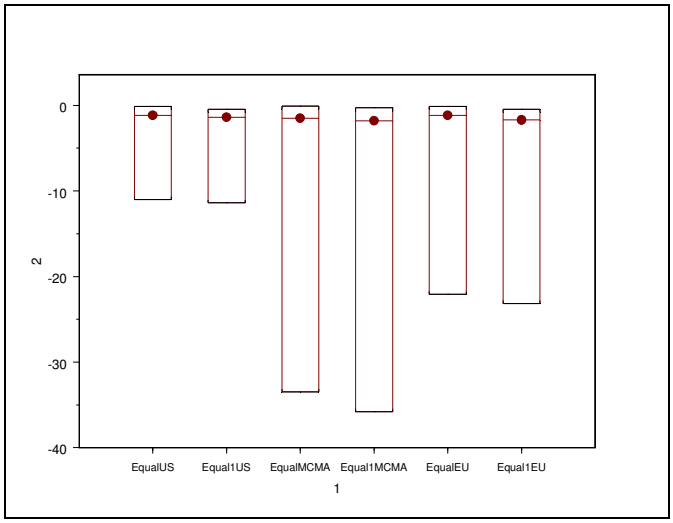
Another option to be explored is the exclusion of experts 4 and 5 (The elicitation staff think that estimates given by these, experts might not be independent of the elicitors' opinion in the subject) .The results for Global and Equal combination are shown in tables 4.8 and A4.5. The calibration score for the two new decision makers EWDM_1 and GWDM_1 is the same as in the previous case, while the information scores show negligible variation.

Table 4.8 Decision Makers Excluding Experts 4 and 5.

```
Case name : MEXICANEXPERTSforthesis 21-7-2004              CLASS version W4.0
_____________________________________________________________________________
Results of scoring experts
 Bayesian Updates: no      Weights:  equal      DM Optimisation:  no
 Significance Level:          0    Calibration Power:          1
_____________________________________________________________________________
   Nr.| Calibr.    |Mean relat|Mean relat|Numb|UnNormaliz|Normaliz.w|Normaliz.w
      |            | total    |realizatii|real|weight    |without DM|with DM
______|____________|__________|__________|____|__________|__________|_________
     1|   0.007064|    0.9335|     1.094|  12|   0.00773|      0.25|   0.04246
     2| 8.336E-007|     1.402|     1.313|  12|1.095E-006|      0.25|6.015E-006
     3|    0.08608|     1.094|    0.8053|  12|   0.06932|      0.25|    0.3808
     6| 3.754E-007|    0.9796|     1.595|  12|5.986E-007|      0.25|3.288E-006
EWDM_1|     0.2541|    0.4322|    0.4131|  12|     0.105|          |    0.5767
GWDM_1|    0.08608|     1.094|    0.8053|  12|   0.06932|          |       0.5
_____________________________________________________________________________
_____________________________________________________________________________
                                                        (c) 1999 TU Delft
```

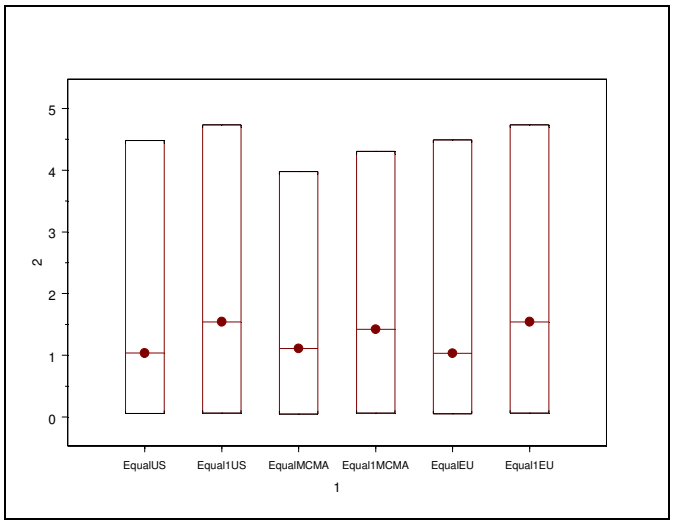
In table 4.2 it was observed that if normalizing the last column experts 4 and 5 would still have some influence in a distribution constructed with these weights, in this sense, one question that arises is how different the distributions constructed using the 6 experts would be to that excluding experts 4 and 5 from the analysis and that is done in figures 5, 6 and 7 below.

Figure 4.5 Median, 5th and 95th Quantiles for Long Term Effects of Air Pollution in Mortality



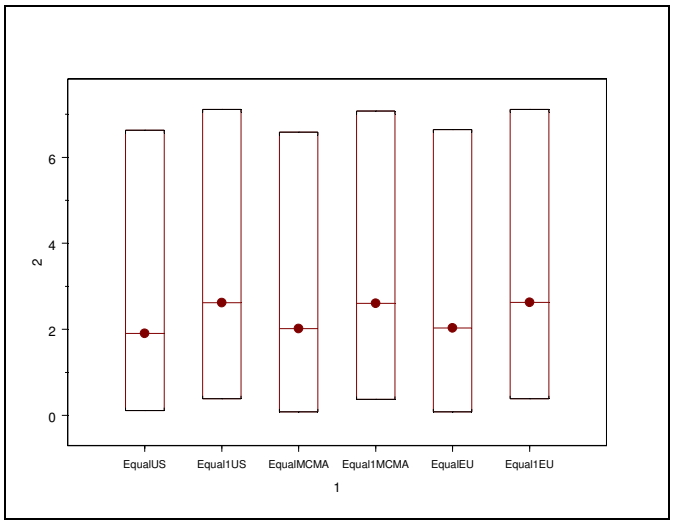
Data generated in Excalibur with results from the Mexico City Elicitation.
Equal1 excludes experts 4 and 5.

Figure 4.6 Median, 5th and 95th Quantiles for One Week Effects of Air Pollution in Mortality



Data generated in Excalibur with results from the Mexico City Elicitation.
Equal1 excludes experts 4 and 5.

Figure 4.7 Median, 5th and 95th Quantiles for Three Months Effects of Air Pollution in Mortality



Data generated in Excalibur with results from the Mexico City Elicitation.
Equal1 excludes experts 4 and 5.

Figures 4.5, 4.6 and 4.7 show the uncertainty distribution for the variables of interest constructed with the equal weight combination scheme with all experts (Equal) and excluding experts 4 and 5 (Global). It is observed that there is little difference in the distributions by excluding these two experts from the analysis.

4.3.5 Choice of Combination.

In choosing between the two decision makers considered in this analysis, the following should be borne in mind:

1) Both the equal weight and the performance based decision makers show acceptable statistical performance
2) The performance based decision maker is significantly more informative than the equal weight decision maker
3) The overall scores between these two are rather close,
4) The robustness of the performance based decision maker is quite satisfactory.

Thus the conclusion is that there are no compelling scientific arguments driving the choice between the performance based and equal weight decision maker.

Two things are to be noticed by comparing tables 4.4 and 4.5: The “GWDM” decision maker is less well calibrated than the “EWDM” decision maker by approximately a factor of 3, though is more informative. The second thing to be noticed is that when choosing an optimization procedure for determining weights, expert number three dominates the scene and the decision maker considers only this opinion in the combination.

APENDIX 4

A4.1. Calibration and Information Scores in the Classical Model for Expert Judgment.

Calibration.

We have asked for experts' uncertainty over a number of calibration variables; these variables are chosen to resemble the quantities of interest, and to tell us something about the expertise of the people from whom the assessments of the variables of interest will be required. An expert states *n* fixed quantiles for his/her subjective distribution for each of several uncertain quantities taking values in a continuous range. There are *n*+1 'inter-quantile intervals' into which the realizations (actual values) may fall. Let

$$P = (p_1,\ldots,p_{n+1}) \quad (1)$$

denote the theoretical probability vector associated with these intervals. Thus, if the expert assesses the 5%, 25%, 50%, 75% and 95% quantiles for the uncertain quantities, then n = 5 and p = (5%, 20%, 25%, 25%, 20%, 5%). The expert believes there is 5% probability that the realization falls between his/her 0% and 5% quantiles, a 20% probability that the realization falls between his/her 5% and 25% quantiles, and so on.

Suppose we have such quantile assessments for N seed variables. Let

$$s = (s_1,\ldots s_{n+1}) \quad (2)$$

denote the empirical probability vector of relative frequencies with which the realizations fall in the inter quantile intervals. Thus

$$s_1 = \frac{(\#\text{ realizations less than or equal to the 5\% quantile})}{N}$$

$$s_2 = \frac{(\#\text{ realizations strictly above the 5\% quantile and less than or equal to the 25\% quantile})}{N}$$

$$s_3 = \frac{(\#\text{realizations strictly above the 25\% quantile and less than or equal to the 50\% quantile})}{N}$$ and, so on.

If the expert is well calibrated, he/she should give intervals such that – in a statistical sense- 5% of the realizations of the calibration variables fall into the corresponding 0% to 5% intervals, 20% fall into the 5% to 25% intervals, etc.

Under the hypothesis that the uncertain quantities may be viewed as independent samples from the probability vector p, the quantity:[6]

$$2NI(s,p) = 2N\sum_{i=1}^{n+1} s_i \ln(\frac{s_i}{p_i}) \quad (3)$$

[6] I(s,p) is the Shannon relative information of s with respect to p. For all s,p with $p_i > 0$, i = 1,…, n+1, we have $I(s,p) \geq 0$ and I(s,p) = 0 if and only if s=p (see Kullback 1959).

is asymptotically Chi-square distributed with n degrees[7] of freedom. Thus, if χ_n^2 is the cumulative distribution function for a Chi-square variable with n degrees of freedom, then

$$CAL = 1 - \chi_n^2 (2NI(s,p)) \quad (4)$$

is the upper tail probability, and is asymptotically equal to the probability of seeing a disagreement no larger than I(s,p) on N realizations, under the hypothesis that the realizations are drawn independently from p.

CAL is a measure of the expert's calibration. Low values (near zero) correspond to poor calibration. This arises when the difference between s and p cannot be plausibly explained as the result of mere statistical fluctuation. In the language of hypothesis testing, CAL is the "p-value" at which the hypothesis that the expert's probability values, as given by p, are true would be rejected. For example, if N = 10, and we find that 8 of the realizations fall below their respective 5% quantile or above their respective 95% quantile, then we could not plausibly believe that the probability for such events was really 5%. This phenomenon is sometimes called "overconfidence". Similarly, if 8 of the 10 realizations fell below their 50% quantiles, then this would indicate a "median bias". In both cases, the value of CAL would be low. High values of CAL indicate good calibration.

Information.

Information shall be measured as Shannon's relative information with respect to a user-selected background measure. The background measure will be taken as the uniform (or loguniform) measure over a finite "intrinsic range" for each variable. For a given uncertain quantity and a given set of expert assessments, the intrinsic range is defined as the smallest interval containing all the experts' quantiles and the realization, if available, augmented above and below by K%.

The relative information of expert e on a given variable is:

$$I(e) = \sum_{i=1}^{n+1} p_i \ln(\frac{p_i}{r_i}) \quad (5)$$

Where r_i are the background measures of the corresponding intervals and n the number of quantiles assessed. Overall informativeness per expert is the average of the information scores over all variables. For each expert, an information score for all variables is obtained by summing the information scores for each variable[8]. Roughly speaking, with the uniform background measure, more informative distributions are gotten by choosing quantiles, which are closer together, whereas less informative distributions result when the quantiles are farther apart.

The calibration score is a "fast" function; that is, differences of several orders of magnitude are observed in a relatively small group of experts with, say 12 calibration variables. On the other hand, information is a "slow" function; differences are typically within a factor three. In

[7] $P(2NI(s;p) \leq x) \approx \chi_n^2 (2NI(s;p))$

[8] This corresponds to the information in the expert's joint distribution relative to the product of the background measures under the assumption that the expert's distributions are independent.

combining expert judgments, these scores are multiplied and normalized; hence in combining experts, the calibration score dominates over information score. Information serves to modulate between more or less equally well-calibrated experts.

A4.2 Summary of the Elicitation Protocol for Mexico City Air Pollution Experts.

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
1	US	Long-term	1 $\mu g/m^3$	$PM_{2.5}$	Ambient	18 ug/m3

What is your estimate of the true, but unknown, percent change in the total annual, non-accidental mortality rate in the adult U.S. population resulting from a permanent 1 $\mu g/m^3$ reduction in long-term annual average $PM_{2.5}$ (from a population-weighted baseline concentration of 18 $\mu g/m^3$) throughout the U.S.? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
2	MCMA	Long-term	1 $\mu g/m^3$	$PM_{2.5}$	Ambient	35 ug/m3

What is your estimate of the true, but unknown, percent change in the total annual, non-accidental mortality rate in the adult MCMA population resulting from a permanent 1 $\mu g/m^3$ reduction in long-term annual average $PM_{2.5}$ (from a population-weighted baseline concentration of 35 $\mu g/m^3$) throughout the MCMA? To express the uncertainty associated with the concentration-response relationship, please provide the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
3	EU	Long-term	1 $\mu g/m^3$	$PM_{2.5}$	Ambient	20 ug/m3

What is your estimate of the true, but unknown, percent change in the total annual, non-accidental mortality rate in the adult European population resulting from a permanent 1 $\mu g/m^3$ reduction in long-term annual average $PM_{2.5}$ (from a population-weighted baseline concentration of 20 $\mu g/m^3$) throughout the EU? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
4	US	Short-term (one week)	10 μg/m^3	$PM_{2.5}$	Ambient	Current

What is your estimate of the true, but unknown, percent change in non-accidental mortality in the adult U.S. population over the one week following a 10 μg/m^3 increase in $PM_{2.5}$ levels on a single day throughout the U.S.? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
5	MCMA	Short-term (one week)	10 μg/m^3	$PM_{2.5}$	Ambient	Current

What is your estimate of the true, but unknown, percent change in non-accidental mortality in the adult MCMA population over the one week following a 10 μg/m^3 increase in $PM_{2.5}$ levels on a single day throughout the MCMA? To express the uncertainty associated with the concentration-response relationship, please provide the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
6	EU	Short-term (one week)	10 μg/m^3	$PM_{2.5}$	Ambient	Current

What is your estimate of the true, but unknown, percent change in non-accidental mortality in the adult Eurpoean population over the one week following a 10 μg/m^3 increase in $PM_{2.5}$ levels on a single day throughout the EU? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
7	US	Short-term (three months)	10 μg/m^3	$PM_{2.5}$	Ambient	Current

What is your estimate of the true, but unknown, percent change in non-accidental mortality in the adult U.S. population over the three months following a 10 μg/m^3 increase in $PM_{2.5}$ levels on a single day throughout the U.S.? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
8	MCMA	Short-term (three months)	10 μg/m^3	$PM_{2.5}$	Ambient	Current

What is your estimate of the true, but unknown, percent change in non-accidental mortality in the adult MCMA population over the three months following a 10 μg/m^3 increase in $PM_{2.5}$ levels on a single day throughout the MCMA? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

Question	*Setting*	*Exposure (Effect Interval)*	*Change*	*Pollutant*	*Composition*	*Baseline*
9	EU	Short-term (three months)	10 μg/m^3	$PM_{2.5}$	Ambient	Current

What is your estimate of the true, but unknown, percent change in non-accidental mortality in the adult European population over the three months following a 10 μg/m^3 increase in $PM_{2.5}$ levels on a single day throughout the EU? To express the uncertainty associated with the concentration-response relationship, please provide the 5th, 25th, 50th, 75th, and 95th percentiles of your estimate.

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

10. On how many **days in 2003** does the daily average PM_{10} concentration **exceed the 24 hr limit** at least one of the above RAMA stations (max 365)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

11. On how many **days in 1995** did the daily average PM_{10} concentration **exceed the 24 hr limit** at least one of the above RAMA stations (max 365)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

12. On how many **days in 2003** does the daily average PM_{10} concentration lie **below 50 µg/m^3** in at least one of the above RAMA stations (max 365)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

13. On how many **days in 1995** did the daily average PM_{10} concentration lie **below 50 µg/m^3** in at least one of the above RAMA stations (max 365)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

14. On how many **hours in 2003** does the hourly average O_3 concentration **exceed the 1hr limit** in at least one of the above RAMA stations (max 8736)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

15. On how many **hours in 1995** did the hourly average O_3 concentration **exceed the 1hr limit** in at least one of the above RAMA stations (max 8736)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

16. On how many **hours in 2003** does the hourly average O_3 concentration lie **below 0.05 ppm** in at least one of the above RAMA stations (max 8736)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

17. On how many **hours in 1995** did the hourly average O_3 concentration lie **below 0.05 ppm** in at least one of the above RAMA stations (max 8736)?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

18. What is the number of non-accidental deaths in the week (7 days starting from January 1st) of 2000 with the **highest** average PM_{10} concentration (92 $\mu g/m^3$) averaged over these three RAMA stations?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

19. What is the number of non-accidental deaths in the week (7 days starting from January 1st) of 2000 with the **lowest** average PM_{10} concentration (23 $\mu g/m^3$) averaged over these three RAMA stations ?

5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

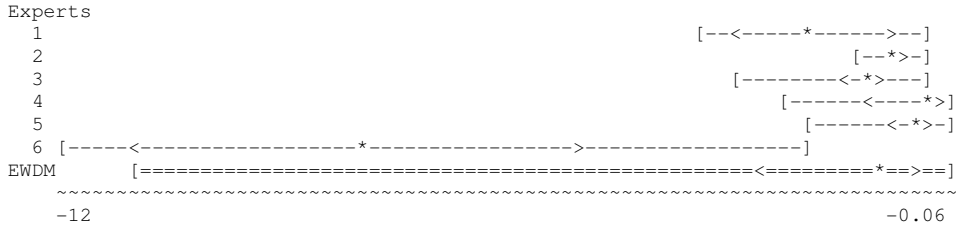
20. What is the number of non-accidental deaths in the week (7 days starting from January 1st) of 2000 with the **highest** average PM_{10} concentration (120 $\mu g/m^3$) averaged over these two RAMA stations ?

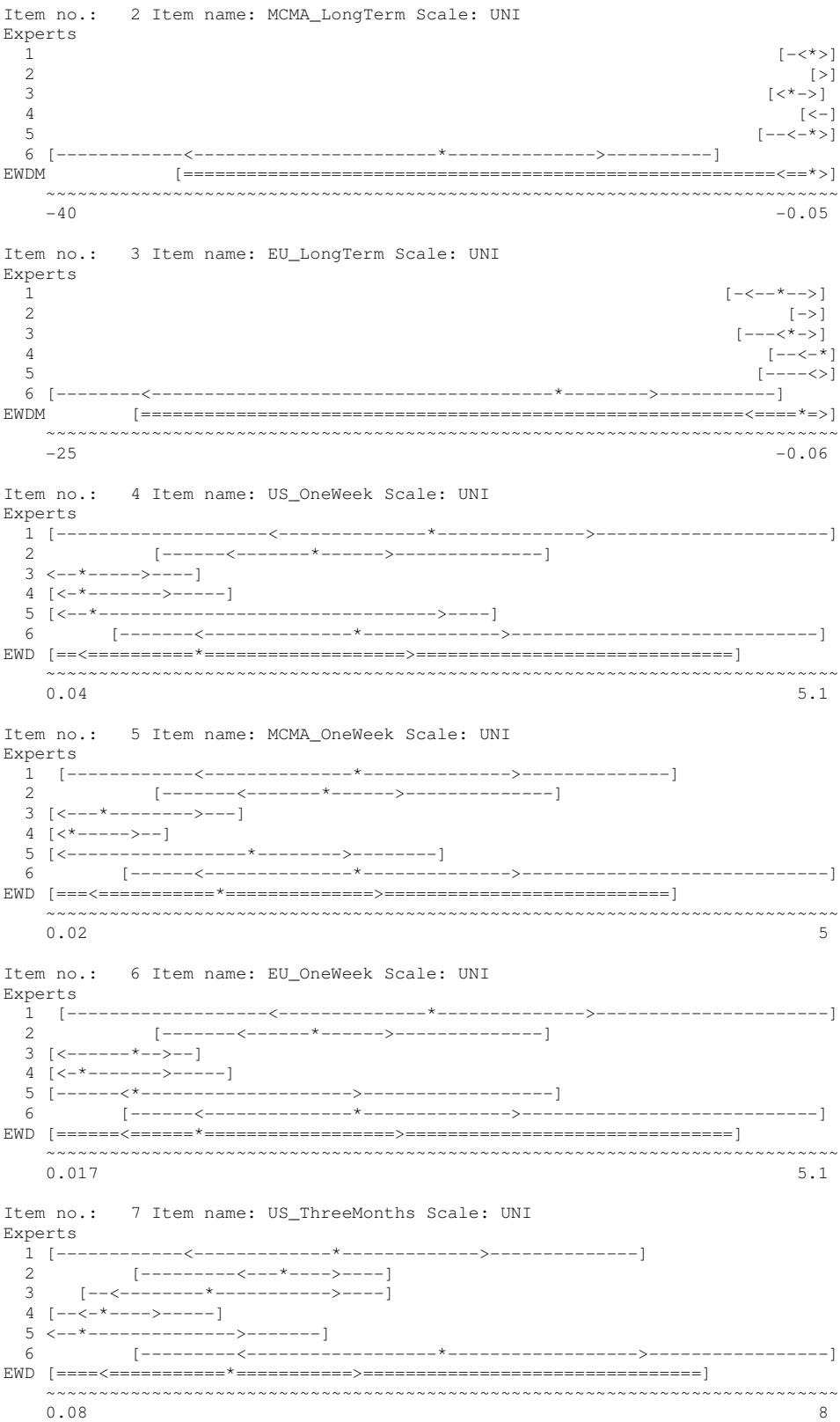
5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

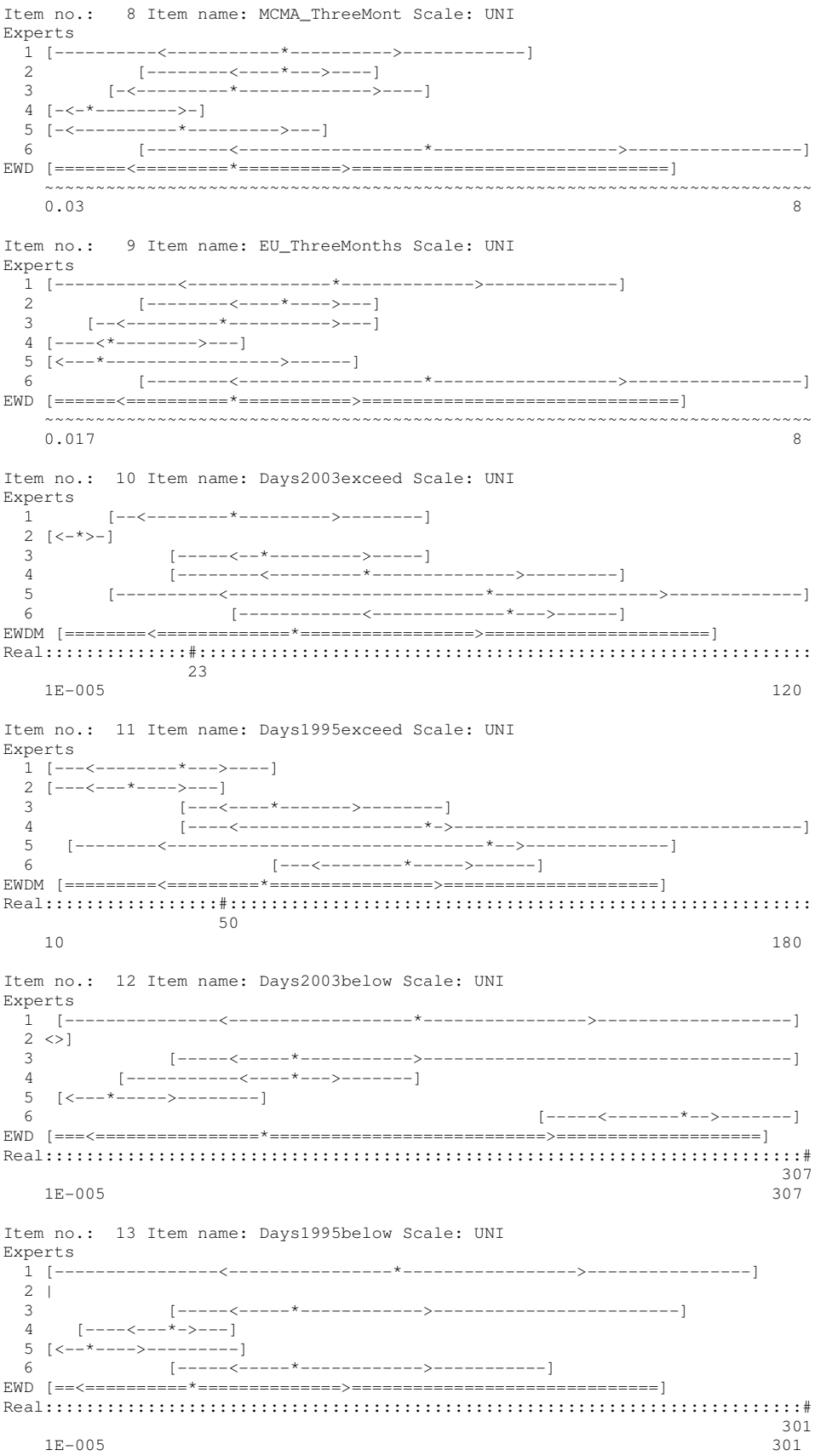
21. What is the number of non-accidental deaths in the week (7 days starting from January 1st) of 2000 with the **lowest** average PM_{10} concentration (23 $\mu g/m^3$) averaged over these two RAMA stations ?

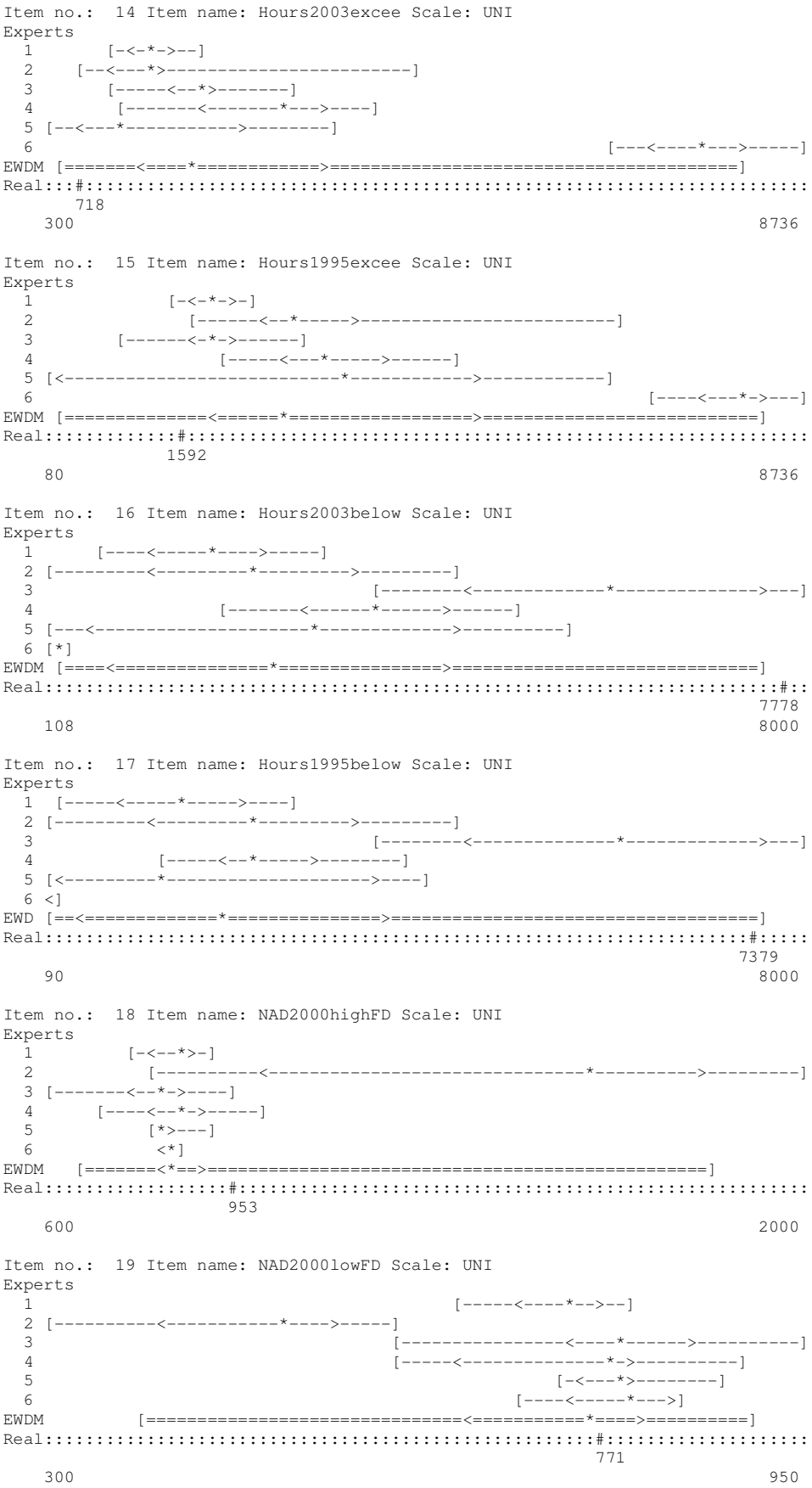
5% :_____ 25%:_____ 50% :_____ 75%:_____ 95%:_____

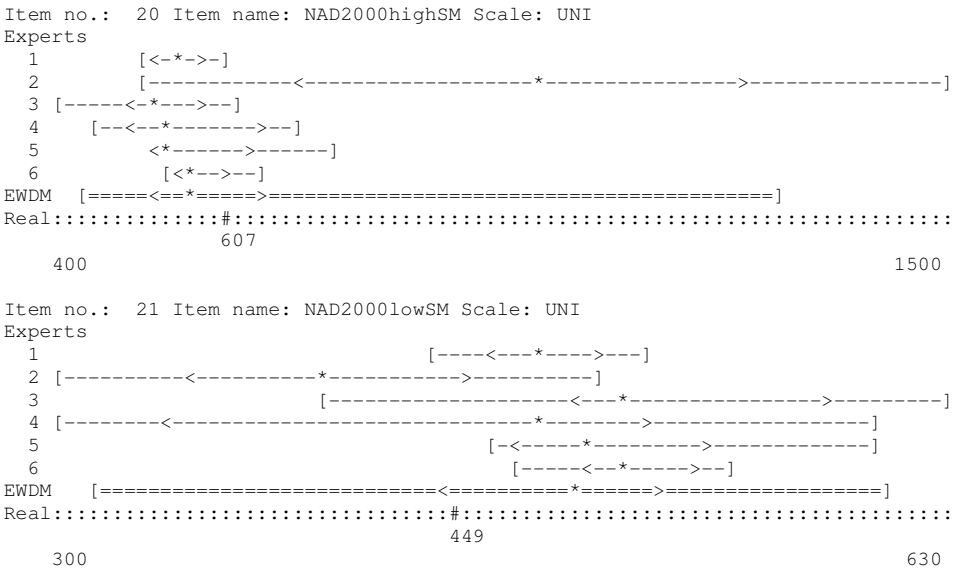
A4.3 Range Graphs for Calibration Questions in the Mexico City Air Pollution Experts Elicitation.




Item no.: 2 Item name: MCMA_LongTerm Scale: UNI
Experts
1
2
3
4
5
6
EWDM
-40
-0.05
Item no.: 3 Item name: EU_LongTerm Scale: UNI
Experts
1
2
3
4
5
6
EWDM
-25
-0.06
Item no.: 4 Item name: US_OneWeek Scale: UNI
Experts
1
2
3
4
5
6
EWD
0.04
5.1
Item no.: 5 Item name: MCMA_OneWeek Scale: UNI
Experts
1
2
3
4
5
6
EWD
0.02
5
Item no.: 6 Item name: EU_OneWeek Scale: UNI
Experts
1
2
3
4
5
6
EWD
0.017
5.1
Item no.: 7 Item name: US_ThreeMonths Scale: UNI
Experts
1
2
3
4
5
6
EWD
0.08
8


Item no.: 8 Item name: MCMA_ThreeMont Scale: UNI
Experts
1
2
3
4
5
6
EWD
0.03
8
Item no.: 9 Item name: EU_ThreeMonths Scale: UNI
Experts
1
2
3
4
5
6
EWD
0.017
8
Item no.: 10 Item name: Days2003exceed Scale: UNI
Experts
1
2
3
4
5
6
EWDM
Real
23
1E-005
120
Item no.: 11 Item name: Days1995exceed Scale: UNI
Experts
1
2
3
4
5
6
EWDM
Real
50
10
180
Item no.: 12 Item name: Days2003below Scale: UNI
Experts
1
2
3
4
5
6
EWD
Real
307
1E-005
307
Item no.: 13 Item name: Days1995below Scale: UNI
Experts
1
2
3
4
5
6
EWD
Real
301
1E-005
301


Item no.: 14 Item name: Hours2003excee Scale: UNI
Experts
EWDM
Real
718
300
8736
Item no.: 15 Item name: Hours1995excee Scale: UNI
Experts
EWDM
Real
1592
80
8736
Item no.: 16 Item name: Hours2003below Scale: UNI
Experts
EWDM
Real
7778
108
8000
Item no.: 17 Item name: Hours1995below Scale: UNI
Experts
EWD
Real
7379
90
8000
Item no.: 18 Item name: NAD2000highFD Scale: UNI
Experts
EWDM
Real
953
600
2000
Item no.: 19 Item name: NAD2000lowFD Scale: UNI
Experts
EWDM
Real
771
300
950



A4.4 Uncertainty Distribution for Variables of Interest (all experts).

VARIABLE	QUANTILE				
	5%	25%	50%	75%	95%
EqualUS_LongTerm	-11.01	-2.61	-1.16	-0.67	-0.13
GlobalUS_LongTerm	-3.00	-1.50	-1.24	-1.10	-0.50
EqualMCMA_LongTerm	-33.48	-3.02	-1.51	-0.82	-0.08
GlobalMCMA_LongTerm	-3.50	-2.80	-2.30	-1.50	-0.90
EqualEU_LongTerm	-22.07	-2.74	-1.19	-0.68	-0.11
GlobalEU_LongTerm	-3.20	-2.00	-1.60	-1.00	-0.50
EqualUS_OneWeek	0.06	0.27	1.04	2.39	4.48
GlobalUS_OneWeek	0.05	0.10	0.30	0.70	1.00
EqualMCMA_OneWeek	0.05	0.33	1.11	2.13	3.98
GlobalMCMA_OneWeek	0.05	0.10	0.40	1.00	1.25
EqualEU_OneWeek	0.05	0.50	1.03	2.29	4.49
GlobalEU_OneWeek	0.05	0.10	0.60	0.80	1.00
EqualUS_ThreeMonths	0.11	0.70	1.91	3.20	6.63
GlobalUS_ThreeMonths	0.50	0.75	1.75	3.00	3.50
EqualMCMA_ThreeMonths	0.08	0.89	2.02	3.21	6.58
GlobalMCMA_ThreeMonths	0.70	0.90	2.00	3.50	4.00
EqualEU_ThreeMonths	0.08	0.77	2.03	3.22	6.65
GlobalEU_ThreeMonths	0.50	0.80	1.85	3.00	3.50

A4.5 Uncertainty Distribution for Variables of Interest (excluding experts 4 and 5).

VARIABLE	QUANTILE				
	5%	25%	50%	75%	95%
Equal1US_LongTerm	-11.38	-3.48	-1.41	-0.92	-0.46
Global1US_LongTerm	-3.00	-1.50	-1.24	-1.10	-0.50
Equal1MCMA_LongTerm	-35.79	-8.37	-1.81	-0.93	-0.27
Global1MCMA_LongTerm	-3.50	-2.80	-2.30	-1.50	-0.90
Equal1EU_LongTerm	-23.14	-3.86	-1.72	-0.91	-0.44
Global1EU_LongTerm	-3.20	-2.00	-1.60	-1.00	-0.50
Equal1US_OneWeek	0.06	0.65	1.54	2.65	4.74
Global1US_OneWeek	0.05	0.10	0.30	0.70	1.00
Equal1MCMA_OneWeek	0.06	0.71	1.42	2.48	4.30
Global1MCMA_OneWeek	0.05	0.10	0.40	1.00	1.25
Equal1EU_OneWeek	0.06	0.73	1.54	2.65	4.74
Global1EU_OneWeek	0.05	0.10	0.60	0.80	1.00
Equal1US_ThreeMonths	0.39	1.52	2.62	3.76	7.12
Global1US_ThreeMonths	0.50	0.75	1.75	3.00	3.50
Equal1MCMA_ThreeMonths	0.37	1.51	2.60	3.71	7.07
Global1MCMA_ThreeMonths	0.70	0.90	2.00	3.50	4.00
Equal1EU_ThreeMonths	0.39	1.54	2.63	3.76	7.12
Global1EU_ThreeMonths	0.50	0.80	1.85	3.00	3.50

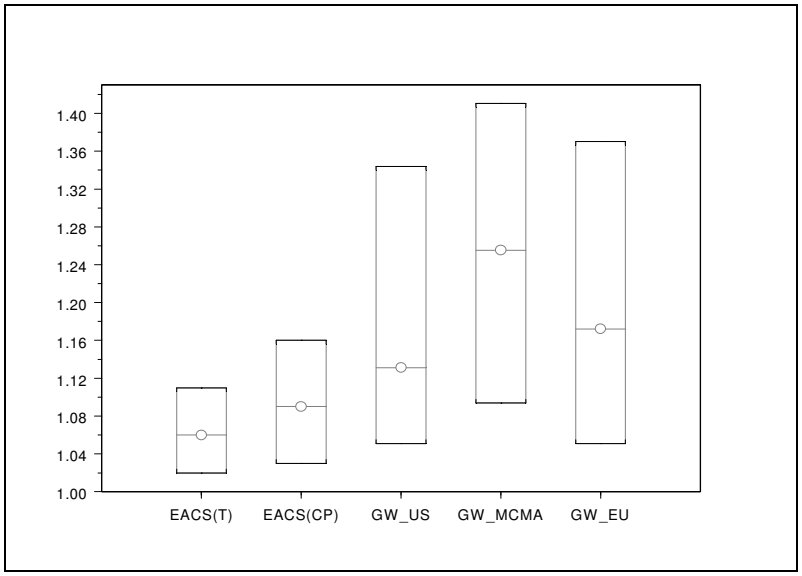
Chapter 5. Conclusions and Recommendations.

Several conclusions have been presented throughout this research. This section will summarize those findings considered more relevant so that recommendations may be put forward in a more comprehensible way for the reader. In this chapter all relevant quantities were transformed to relative risks for 10 units increase in air pollution to make comparison easier. The reader should refer to original sources or the chapter of interest for more details on how the original quantities are reported.

5.1 Long Term Effects of Air Pollution in Mortality.

In figure 5.1 the estimates for long term mortality (5^{th}, 50^{th} and 95^{th} percentiles) found with the expert judgment exercise described in chapter 4 (Global Weights Decision Maker) transformed to risk ratios are plotted against one of the most recent studies that use a variant of the Cox proportional hazards model described in chapter 1. The first two boxes correspond to the risk ratio estimates from (Pope, et. al., 2002) for total and cardiopulmonary mortality respectively[9]. The next three boxes correspond to equivalent estimates from the expert judgment study for application in the United States, Mexico City, and Europe in that order.

Figure 5.1 Relative Risk from Cox Model per 10 $\mu g/m^3$ Increase in $PM_{2.5}$ (Central and Interval Estimates).



EACS(T)= Extended American Cancer Society Study (Total Mortality); EACS(T)= Extended American Cancer Society Study (Cardiopulmonary Mortality); Estimated and adjusted based on the baseline random effects CPHM, controlling for age, sex, race, smoking, history of high blood pressure, years lived with a smoker and total exercise level.
GW_US = Global Weight Decision Maker United States Setting; GW_US = Global Weight Decision Maker for United States; GW_MCMA = Global Weight Decision Maker for Mexico City Metropolitan Area; GW_EU = Global Weight Decision Maker for European Union.

[9] The confidence intervals shown for this study are only thought to reflect the sampling error in estimating the coefficients. The authors don't make any suggestion that these are the true uncertainty intervals of the estimates.

From the picture two conclusions may be drawn. First, that the experts (or at least Mexican experts) believe that the "true" effect of $PM_{2.5}$ in mortality is larger than that reported in the ACS study as expressed by the central estimate: 50^{th} percentile of global weights decision maker compared to the central estimate from the ACS study. For the 3 geographic locations, the experts believe that the central estimate for total mortality could be even larger than the central estimate found in the study of reference for cardiopulmonary mortality. This observation is consistent with the results from chapter 1 where an incomplete Cox model considerably underestimated the "true" value of the coefficient.

The second conclusion is that experts express their uncertainty in the estimates found across the literature by providing a range for the 90% confidence much wider than the one normally reported in the literature, this result is expected as the uncertainty surrounding a given estimate is influenced by other factors than just statistical fluctuation reflected by confidence intervals. This expert opinion combination does not consider the possibility of no effect of particulate matter on mortality.

The results from the investigation show that more emphasis should be placed in assessing model adequacy. The reports that investigate long term exposure to air pollution should include not only central and interval risk estimates but also measures of model performance and analysis of the whole survival experience of the population studied so that the decision maker can include these measures in their analysis. Another conclusion extracted from chapter four is that expert judgment could also be a useful tool for characterizing and monitoring the uncertainty as perceived by the scientific community regarding such kind of studies.

5.2 Short Term Effects of Air Pollution in Mortality.

Figure 5.2 is the equivalent of 5.1 for short term effects of air pollution in total mortality. The first two boxes in the plot correspond to estimates for the United States; the next two correspond to estimates for Europe and the last 4 correspond to estimates for Mexico. From the estimates for the United States it may be observed that experts believe that $PM_{2.5}$ has a larger influence than that documented for PM_{10} in the total daily number of deaths, and that they express larger uncertainty regarding the true effect by giving a larger range for their 90% confidence than the one found in previous studies.

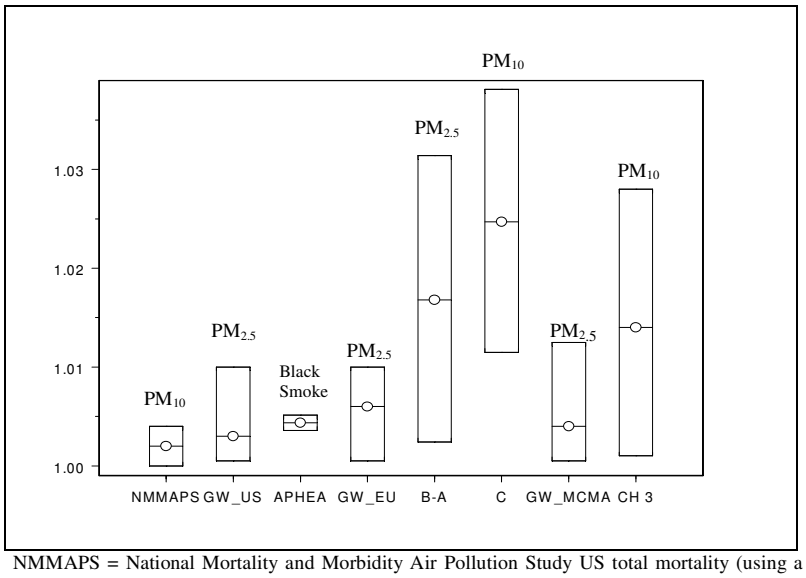
The smaller confidence bounds estimate for the risk ratio in the APHEA study is possibly obtained because of the use of black smoke as air pollution measurement, when these estimates are compared to the Mexican expert's "optimized" opinion it is observed that experts' believe that short term effects of air pollution in mortality could be different across geographical location, possibly because of baseline health status or differences in the pollution mix. Their best estimate for Europe is larger than the one for USA whereas their uncertainty is in general expressed in the same manner as for the previous setting.

The estimate found with GLM for effects of PM_{10} in mortality in Toluca city is roughly the same as the one found in (Borja-Aburto, 1998) for $PM_{2.5}$ in Mexico City when considering smooth functions of some covariates in the regression analysis and is also not far from the one found later by (Castillejos, et. al., 2000) for PM_{10}. This leads to the conclusion that a regression that does not control for possible non linear effects of certain covariates is also

capable to capture air pollution health effects (the estimates presented where found statistically significant with hypothesis testing).

From figure 5.2 it is observed that the experts believe that the estimates available for Mexico City (and hence possibly the one found for Toluca in this investigation) might be over estimated. Taking these considerations together a recommendation would be to conduct a multi-site air pollution study in Mexico. Furthermore, from the results of chapter three a multi-site study incorporating methodologies previously used (Fourier decomposition of time series, hierarchical models for combining site-specific estimates, etc.) is desirable but not sufficient. To go further in investigating short term effects of air pollution, different sensitivity and uncertainty measures (like the attempt made in chapter three of using also partial correlation coefficients) should be applied to the models and incorporated in the reports, and an expert judgment exercise similar to the one sketched in chapter four would be desirable after new multi site information becomes available.

Figure 5.2 Short Term Air Pollution Effects on Mortality (Central and Interval Estimates).

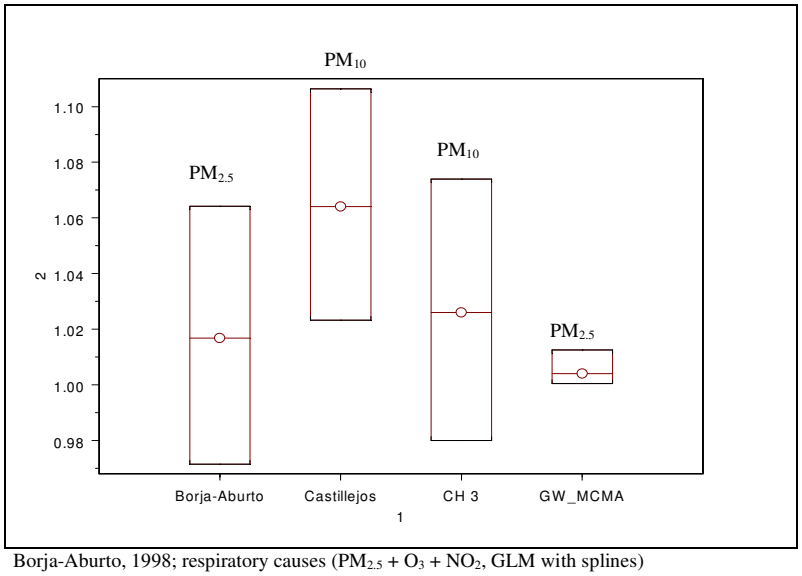


NMMAPS = National Mortality and Morbidity Air Pollution Study US total mortality (using a non informative prior, $PM_{10} + O_3 + NO_2$);
GW_US = Global Weight Decision Maker for United States, adult population ($PM_{2.5}$);
APHEA = Air Pollution and Health: a European Approach (Black smoke, using GAM)
GW_EU = Global Weight Decision Maker for European Union, adult population ($PM_{2.5}$).
B-A = Borja-Aburto, 1998 ($PM_{2.5} + O_3 + NO_2$, GLM with splines)
C = Castillejos,2000 ($PM_{10} + O_3 + NO_2$, GLM with splines)
GW_MCMA = Global Weight Decision Maker for Mexico City Metropolitan Are, adult population ($PM_{2.5}$);
CH 3 = Estimate found in chapter 3 for Toluca ($PM_{2.5}$, 3 day lag, GLM)

In figure 5.3 the reader may observe that the effects of $PM_{10,2.5}$ in mortality related to the respiratory system observed in Mexico City and Toluca are likely to be larger than the effect that experts' believe $PM_{2.5}$ could have in adult mortality in Mexico city. However, the variation of the central indicator is so large that the possibility of no effect of air pollution in respiratory system related mortality is also observed in the first box in figure 5.3.

In chapter four, the classical method for expert judgment proved to be a useful tool for dealing with the uncertainty surrounding air pollution health effect estimates, and it could also be a useful tool for investigating cause specific mortality that could provide information about the relationship of the pollutants with frail groups of people. The multi-site aggregation should include also cause specific mortality analysis.

Figure 5.3 Respiratory System Related Mortality Central and Interval Estimates.



Borja-Aburto, 1998; respiratory causes ($PM_{2.5} + O_3 + NO_2$, GLM with splines)
Castillejos,2000; respiratory causes ($PM_{10} + O_3 + NO_2$, GLM with splines)
CH 3 = Estimate found in chapter 3 for Toluca, respiratory causes (3 day lag, GLM)
GW_MCMA = Global Weight Decision Maker for Mexico City Metropolitan Area, adult population ($PM_{2.5}$);

This investigation suggests that there is enough evidence derived from the mathematical models available at the moment to support the hypothesis that air pollution has negative effects in health; however it also indicates that not everything is settled with these tools, especially regarding model performance and sensitivity and uncertainty analysis. These issues should be address more extensively in future research.

References

Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Beeson WL, Yang JX. "*Long-term inhalable particles and other air pollutants related to mortality in nonsmokers*". Am J Resp Crit Care Med. 1999; 159: 373-382.

Allison, P.D. "*Survival Analysis Using SAS a Practical Guide*". SAS Publishing, Caray. 2003

Borja-Aburto VH, Castillejos M, Gold DR, Bierzwinski S, Loomis D. "*Mortality and ambient fine particles in southwest Mexico City*", 1993-1995. Environ Health Perspect 106: 849-856 (1998).

Bunea, Cornel "*Mathematical Models for Reliability Data*" ISBN 90-9016926-1, 2003

Castillejos M, Borja-Aburto VH, Dockery DW, et al. 2000. "*Airborne coarse particles and mortality*". Inhalable Toxicology 12: 61-72.

Cooke, R.M. and Goossens, L.H.J. "*Procedures guide for structured expert judgement in accident consequence modelling*" {COS 02}

Cooke, R.M. and Slijkhuis, K.A. "*Expert Judgment in the Uncertainty Analysis of Dike Ring Failure Frequency*" {COS 02}

Cooke, R.M., (1991). "Experts in Uncertainty." Oxford University Press.

Cooke, R.M., Bedford, T. "*Mathematical tools for Probability Risk Analysis*".

Cooke, R.M., Bedford, T. "*The Design of Reliability Data BasesPart I and II* " Reliability Engineering and system safety, 1996.

Cox, D.R. "*Regression Models and Life Tables*" Royal Statistical Society, Series B, Vol. 34, No2. 1972, pp. 187-220

Dobson, Annette J. "*An introduction to Generalized linear models*" 2nd ed. Boca Raton: Chapman and Hall/CRC, 2002

Dockery, D., III, C. A. P., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E. 1993. "*An association between air pollution and mortality in six U.S. cities.*" New England Journal of Medicine, 329, 1753-1759.

Dominici F. and McDermott A. and Zeger S.L. and Samet J.M "*On the Use of Generalized Additive Models in Time Series of Air Pollution and Health American Journal of Epidemiology*" vol 156, no 3, 193-203.

Gamble, John F. "*$PM_{2.5}$ and mortality in long-term prospective cohort studies: cause-effect or statistical associations?*" Environmental Health Perspectives Vol. 106, Number 9, September 1998.

Green, P.J. Silverman, B.W. "*Nonparametric regression and generalized linear models: a roughness penalty approach*" London: Chapman and Hall, 1994

Hastie, T.J. Tibshirani, R.J. "*Generalized additive models*" London: Chapman and Hall, 1991

Hastie, Trevor Tibshirani, Robert Friedman, Jerome "*The elements of statistical learning data*

mining, inference, and prediction". New York: Springer, 2001

Health Effects Institute. 2000. "*Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality*." Health Effects Institute, Cambridge, MA. PART I

Health Effects Institute. 2000. "*Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality*." Health Effects Institute, Cambridge, MA. PART II

Homser, D.W. and Lemeshow, S. "*Applied Survival Analysis*" Wiley, New York.

Kalbfleish, J.D. and Prentice, R.L. "*The Statistical Analysis of Failure Time Data*", second edition, Wiley New York. 2002

Katsouyanni K, Touloumi G, Spix C. "*Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from times series data from the APHEA project*". BMJ 314:1658–1663 (1997).

Lipfert, F. W. and R. E. Wyzga (1995). "Air-Pollution and Mortality - Issues and Uncertainties." Journal of the Air & Waste Management Association Elect Power Res Inst,Palo Alto,Ca 45(12): 949-966.

McCullagh, P. Nelder, J.A. "*Generalized linear models*" 2nd ed. London: Chapman and Hall, 1989-1990

Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. "*Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution*". JAMA, March 6, 2002 – Vol. 287, No 9.

Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D., Evans, J. S., Speizer, F. E., and Heath, C. W. 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.

Samet, J. M., Dominici, F., Zeger, S. L., Schwartz, J., and Dockery, D. W. 2000. "*The National Morbidity, Mortality, and Air Pollution Study*." Health Effects Institute, Cambridge, MA. PART I

Samet, J. M., Dominici, F., Zeger, S. L., Schwartz, J., and Dockery, D. W. 2000. "*The National Morbidity, Mortality, and Air Pollution Study*." Health Effects Institute, Cambridge, MA. PART II

Samoli E, Schwartz J, Wojtyniak B, Touloumi G, Spix C, Balducci F, Medina S, Rossi G, Sunyer J, Bacharova L, Anderson HR, Katsouyanni K. "*Investigating regional differences in short-term effects of air pollution on daily mortality in the APHEA project: a sensitivity analysis for controlling long-term trends and seasonality*". Environ Health Perspect. 2001 Apr;109(4):349-53.

Schimek, Michael G. "*Smoothing and regression approaches, computation, and application*" New York: Wiley,2000

Venables, W.N. Ripley, B.D. "*Modern applied statistics with S-Plus*" 2nd ed. New York: Springer, 1997