Estimating the Association Between Ambient Particulate Matter and Elderly Mortality in Phoenix and Seattle Using Bayesian Model Averaging

Erin M. Sullivan

NRCSE

Technical Report Series

NRCSE-TRS No. 054

August 22, 2000

The **NRCSE** was established in 1996 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.

EPA

Estimating the Association Between Ambient Particulate Matter and Elderly Mortality in Phoenix and Seattle Using Bayesian Model Averaging

Erin M. Sullivan

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Science

University of Washington

2000

Program Authorized to Offer Degree: Statistics

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to her advisors, Peter Guttorp, Tilmann Gneiting, and Merlise Clyde. Their guidance, encouragement, and advice were invaluable to her success and completion of this thesis project.

Although the research described in this article has been funded in part by the United States Environmental Protection Agency through agreement CR825173-01-0 to the University of Washington, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Chapter 1

INTRODUCTION

1.1 Model uncertainty

Statistical modeling is an integral part of many applied data analysis problems. If researchers have a scientific question they wish to answer, and the data to do so, many steps are taken throughout the analysis to create a "final" model, and to make inferences or predictions from this model. One must choose a response variable, or one of several possible transformations. A specific model is selected to describe the relationship between this response and predictors of interest, or their transformations. Potential confounding variables and precision variables may be included or excluded, and again their form must be determined. After this stage of modeling, one usually uses diagnostic techniques to assess the appropriateness and usefulness of the model. Countless decisions must be made by the investigator in this modeling process, such as which link function to use, the form each variable should take, and deciding whether or not to keep unusual observations. Furthermore, each decision is subjective and only lightly governed by arbitrary rules of thumb, such as keeping variables whose coefficients have p-values of less than 0.05.

Often more than one model is considered, and so the researcher must also choose the "best" model, perhaps using a method such as stepwise regression, as well as what aspects of this model will be used to answer the question(s) of interest. This selective reporting of models and results usually ignores the uncertainty resulting from the decisions made by the investigator in building and choosing a model. This can then affect the validity of the p-values and confidence intervals for inferences and predictions made from this model. Selecting a single model and making all subsequent inferences based on this can also lead to policy decisions that are riskier than one may think ([11],[20]). Further discussion of these and other issues of model selection procedures can be found in [26].

As a specific example, we will consider epidemiological studies designed to assess the association between adverse health events and airborne ambient particulate matter (PM). These studies have received great attention in recent years as they help provide a basis for national policy decisions regarding acceptable pollution levels. New standards for levels of PM were proposed by the Environmental Protection Agency (EPA) in 1997, in spite of the remaining high degree of uncertainty regarding the effects of fine and, to a lesser extent, coarse PM on human health.

Most of the PM health-effect studies, current and historical, use model selection methods subject to the criticisms presented above. In this paper we will present and implement an alternative to model selection methods, Bayesian model averaging (BMA). The use of BMA for these particulate matter analyses has been suggested as a possible solution for problems encountered with model selection and exploration ([8], [3]). This method has been presented in detail in [3], [15], [26] and [27] and has been used in the analyses of PM health effects in at least two cities: Birmingham, Alabama and Phoenix, Arizona ([4], [5]). We describe a modified version of the Phoenix, Arizona analysis here, and also use BMA to assess the association between PM and elderly non-accidental mortality in Seattle, Washington. This work can be seen as a continuation of work presented in [3].

1.2 Particulate matter measurements and standards

Before we present the analyses and their motivation, it may be helpful to the reader if we describe the various PM measurements. Regulations set by the EPA concerning PM levels are typically based on the concentration measurement taken at collection sites. Total suspended particulates (TSP) are collected by a device which detects ambient airborne particles with a maximum size of 30 μm ([8]). PM_{10} measurements are made from a device for which 50% of the particles at 10 μm are collected. Therefore, PM_{10} measurements mostly include particles smaller than 10 μm, but also include particles larger than 10 μm ([8]). Similarly, $PM_{2.5}$ measurements are made from a device for which 50% of the particles at 2.5 μm are collected. Other particulate matter variables referred to in this paper are nephelometry, black smoke and British smoke. Nephelometry is a measure of light scattering and is a good measure of particles larger than 0.2 μm and smaller than 0.9 μm ([6]). Black smoke is measured by the amount of blackening on monitor filters. British smoke is measured by the amount of light reflectance from dark particles on monitor filters ([32]).

The most recent particulate matter standards set by the EPA are described in the National Ambient Air Quality Standards (NAAQS) document published in 1997 ([1]). This document states that the annual standard level (based on a three-year average of yearly averages) of $PM_{2.5}$ should not exceed 15 $\mu g/m^3$ and sets the three-year average of the yearly 98th percentile 24-hour level of $PM_{2.5}$ at 65 $\mu g/m^3$ ([1]). Also in this publication, the annual standard level (based on a three-year average of yearly averages) of PM_{10} is retained at 50 $\mu g/m^3$, while the three-year average of the yearly 99th percentile 24-hour level of PM_{10} also remains at 150 $\mu g/m^3$. It should be noted that the standards for PM_{10} have been declared invalid by the Federal Appeals Court in Washington, D.C., on the grounds that the EPA should set separate standards for fine and coarse particles, and PM_{10} measures both fine and coarse particles. The percentiles are calculated nonparametrically, as dictated by the EPA in the NAAQS document, but these could also be estimated parametrically. Appendix A contains an example and comparison of these two methods.

Chapter 2

BAYESIAN MODEL AVERAGING

Bayesian model averaging (BMA) has been suggested by many authors ([3], [15], [26], [27]) as one alternative to the classical model selection process. Model uncertainty is accounted for in BMA, unlike other model selection methods such as stepwise regression or backward/forward elimination, in which it is ignored. This method averages over the best models in the model space to yield useful summary measures, such as the posterior probability of a particular model, the posterior probability that a coefficient is equal to zero and average posterior expected value and variance of parameters ([15]).

2.1 Overview of Bayesian model averaging

Bayesian model averaging provides a coherent alternative for combining inferences from different models and addressing model selection in subsequent inferences ([3], [15], [26], [27]). Under BMA, each model contributes proportionally based on the support it receives from the data, as measured by the posterior probability of each model. Potential models using all possible combinations of covariates are obtained using the leaps and bounds algorithm to find the best models of each size. To provide a baseline reference analysis, we use the Bayes Information Criterion (BIC) or Schwarz Criterion ([29]) for determining posterior model probabilities, where BIC for model M is

$$\text{BIC}(M) \quad = \quad \text{deviance}(M) + \dim(M)\log(n).$$

The quantity deviance(M) is the deviance statistic for model M (-2 log likelihood), $\dim(M)$ is the number of variables in model M, and n is the number of observations. This imposes a heavy penalty on models that contain a large number of parameters and so tends to favor simpler models ([26],[17]).

Suppose we are interested in the posterior probability of a specific model, M, being the correct model out of $m = 1, ..., K$ possible models, $\Pr(M|\text{Data})$. We can find this using Bayes' theorem:

$$\Pr(M|\text{Data}) = \frac{\Pr(\text{Data}|M)\Pr(M)}{\sum_{m=1}^{K}\Pr(\text{Data}|M_m)\Pr(M_m)}.$$

Now assume all models are equally likely a priori, so $\Pr(M) = 1/K$. We can use a BIC approximation result given in [26] to write

$$\Pr(M|\text{Data}) = \frac{\mathrm{e}^{-\mathrm{BIC(M)}/2}}{\sum_{m=1}^{K}\mathrm{e}^{-\mathrm{BIC(m)}/2}}$$

where the sum in the denominator is again over all models $m, m = 1, ..., K$. For some quantity of interest, say β, we can then find the posterior expected value and variance of β, incorporating information from all models considered as

$$\mathrm{E}(\beta|\text{Data}) = \sum_{m=1}^{K}\mathrm{E}(\beta|\text{Data}, m)\Pr(m|\text{Data}),$$

and

$$\begin{aligned}\mathrm{Var}(\beta|\text{Data}) = & \sum_{m=1}^{K}\mathrm{Var}(\beta|\text{Data}, m)\Pr(m|\text{Data}) + \\ & \sum_{m=1}^{K}\{\mathrm{E}(\beta|\text{Data}, m) - \mathrm{E}(\beta|\text{Data})\}^2\Pr(m|\text{Data}).\end{aligned}$$

BMA using BIC has led to improved predictive performance in many situations ([15], [3]), and provides objective probabilities of models. It also has been suggested to be useful when one is interested in the relationship between two variables and there are several other possible covariates that may or may not be necessary to include in the model ([17]). As discussed in chapter 1, this uncertainty about which variables to include in a reported set of models should be accounted for and BMA provides a method for doing this.

2.2 The use of Bayesian model averaging in particulate matter studies

The situation described above is found in several studies, such as [34], [22], and [10], which have failed to account for model uncertainty when trying to assess the relationship between particulate matter and adverse health effects. These studies have used responses such as mortality, asthma cases, and emergency hospital visits, and typically have a small number of predictors of interest. Not all studies have yielded the same results, even among analyses of the same response in the same city (see [19] for many examples). Part of these discrepancies could be due to the model selection process.

Oftentimes, models used in particulate matter analyses include 30 or more variables. This is clearly enough to result in some variables being labeled "statistically significant", using the common $p < 0.05$ criterion, even if they are not truly associated with the response. Also, many models seem to have been explored before a final model is chosen and reported. This concern, that the positive associations between health outcomes and particulate matter are a result of multiple testing and uncertainty inherent in model selection, was raised in the 1998 National Research Council report on "Research Priorities for Airborne Particulate Matter" ([7]). In many studies, results from multiple models may be presented, but no consistent method has been used to combine multiple inferences for the same set of data.

One common reason for the high number of variables included in PM models is the apparent uncertainty in the lag structure of the data. Most PM studies include not only the PM variables of interest, but also selected meteorological variables, other pollutants, variable(s) to account for the baseline trend in mortality, and perhaps an indicator of the day of the week. For the meteorological and PM variables, it is unclear whether today's values affect today's mortality or if the weather and/or PM on some previous day, or an average of several previous days, has a greater effect. Therefore, most studies include these variables for the current day and several previous days, or perhaps a moving average, and then use an automatic model selection procedure to determine which are most significant. Which lags are chosen seems to have a great effect on the outcome. This has been acknowledged as a problem in PM studies by many authors ([16], [18], [36], [12], [19]). See Table 2.1 for examples of several studies in which modeling mortality with various PM lags yields varying results.

Table 2.1: Examples of varying results by lag structure selection in PM studies. Particulate matter measurement definitions are given in section 1.2.

Reference	PM variables	Lag(s)	PM variable with highest reported significance
[24]	TSP and PM_{10}	0-2 days	TSP lags 1 and 2, PM_{10} lag 1
[16]	PM_{10}	5 day mean	mixed, depending on city, monitors and exclusions
[25]	Black smoke and PM_{10}	0 and 1 days, and 3 day mean	mixed, depending on lag, PM_{10} with 3 day mean
[2]	TSP and PM_{10}	many explored	lag 0
[14]	Black smoke	0-2 days	TSP lag 1 and BS lag 2
[35]	PM_{10} and nephelometer	0-2 days	lag 0 and 2-4 day mean
[37]	PM_{10}	0-3 days and 3 day mean	lag 0 and lag 1
[23]	fine particles	means for up to 4 day lag	lag 0 best for summers
[31]	$PM_{2.5}$ and $PM_{10}-PM_{2.5}$	2 day mean	2 day mean significant for both
[10]	$PM_{2.5}$ and PM_{10} and PM_{15}	0 days	lag 0 significant for all
[30]	TSP	0 days	lag 0
[32]	British Smoke	15 hours	significant

In the analyses presented in this report, we consider model uncertainty regarding which meteorological variables should be included, whether there is a coarse or fine particulate matter effect, and which lags of the variables should be included. We calculate the posterior distribution of the relative risk of an increase in non-accidental elderly mortality under BMA associated with simultaneous one interquartile range (IQR) changes in all lags of fine and coarse particulate matter variables included in the models. The use of IQR for calculating relative risk changes is fairly common in PM studies (eg. [34] and [22]). The posterior distribution (approximate) for the relative risk given a model which includes any of the PM variables is a log normal distribution, where the log relative risk has a normal distribution centered at the maximum likelihood estimate of the relative risk under that model and with variance derived from the inverse Fisher information matrix for that model (see [3] for more details on using BMA with Poisson regression models). For models that exclude all PM variables (fine, coarse, or any lags), the relative risk is identically 1. The posterior probability that there is no PM effect is obtained by summing the posterior probabilities of all models that exclude PM.

Chapter 3

PARTICULATE MATTER AND MORTALITY IN PHOENIX, ARIZONA

3.1 Data

Mortality data were obtained from the Arizona Department of Health Services for May 6, 1995−March 30, 1998 and matched to particulate matter and meteorological data for the same time period. Because of concerns about the effect of spatial heterogeneity and potential bias, we constructed four mortality response variables (Figure 3.1). Three, PHXMORT, U2.5MORT and U10MORT, are each a daily count of non-accidental deaths for people age 65 and over in three geographical regions. PHXMORT includes deaths which occurred in the Phoenix metropolitan area using the region defined as Phoenix Division, Arizona by the United States Census Bureau. The second variable, U2.5MORT, includes deaths in a smaller subset of zip codes which are thought to have spatially similar levels of fine particles or $PM_{2.5}$ throughout the region (personal communication Jane Koenig, University of Washington). The third geographical region, is smaller yet and is thought to have fairly homogeneous PM_{10} levels, and therefore spatially homogeneous levels of coarse particles; the corresponding response variable is U10MORT. Zip codes within Phoenix that determine deaths included in U10MORT are 85004, 85006-85009, 85012-85020, 85028-85029, 85031, 85033-85035, 85043, 85051. U2.5MORT is defined using the above zip codes for Phoenix in the U10MORT region plus the following zip codes within Scottsdale, Mesa, Glendale and Tempe: 85201-85205, 85207-85208, 85212-85213, 85234, 85236, 85251, 85253, 85256-85258, 85281-85284, 85296. The fourth response variable, ACCMORT includes all non-traffic related accidental deaths for all age groups which occurred in the Phoenix metropolitan region. See Figure 3.2 for a map of the various study regions. As there is little reason to believe that particulate matter should be associated with non-traffic related accidental mortality, this provides a sensitivity

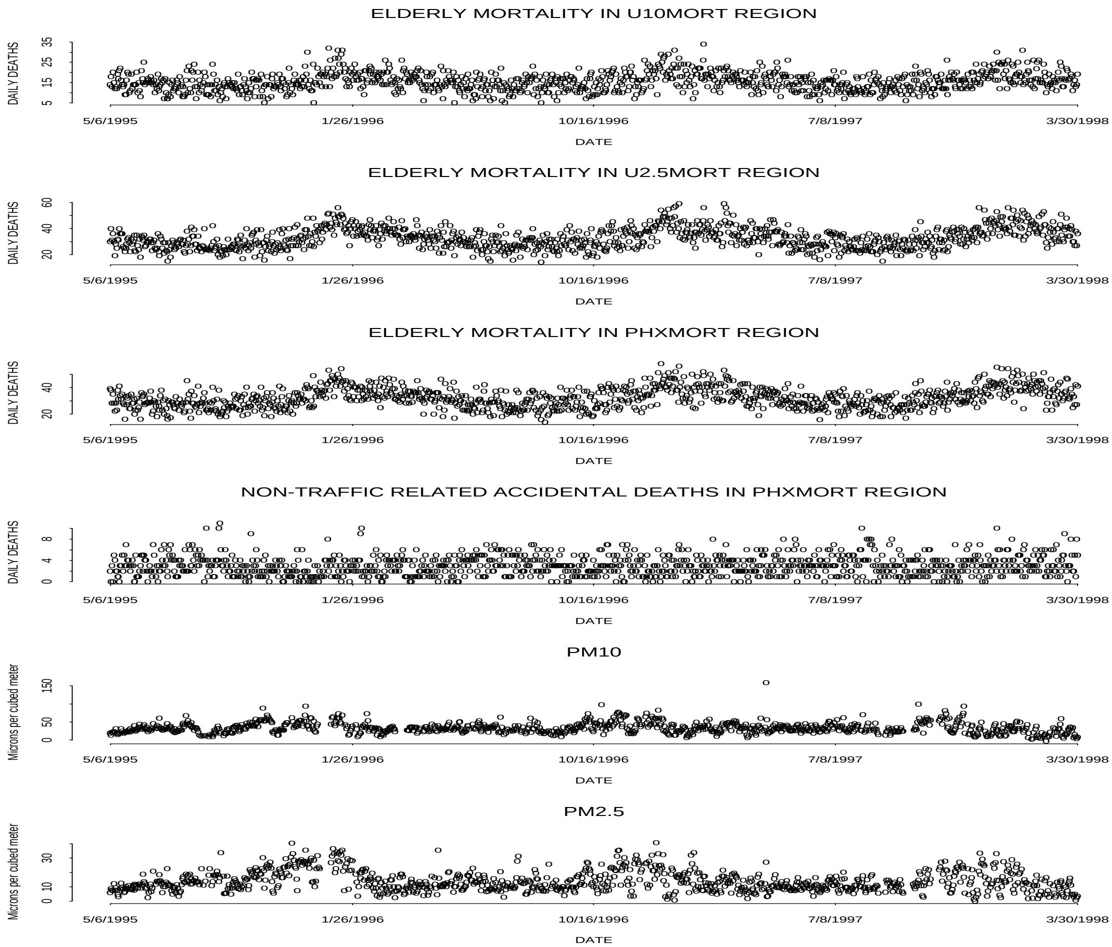

Figure 3.1: Time series of the four mortality responses, coarse PM and fine PM for Phoenix, Arizona from May 6, 1995 to March 30, 1998.

check on the methodology. Because early 1995 was a very mild flu year in contrast to the following years, we have chosen to model mortality using a start date of May 6, 1995.

In light of the upcoming review of the National Ambient Air Quality Standards, our analysis focuses on estimating health effects of fine particles $PM_{2.5}$ and coarse particles (PMC), defined as $PM_{10} - PM_{2.5}$. Daily particulate matter readings (Figure 3.1) from a TEOM monitor (Tapered Element Oscillating Microbalance) were obtained from the EPA's National Exposure Research Laboratory for both PM_{10} and $PM_{2.5}$. In Phoenix, particle mass is typically dominated by the coarse fraction. Table 3.1 gives summaries for PM

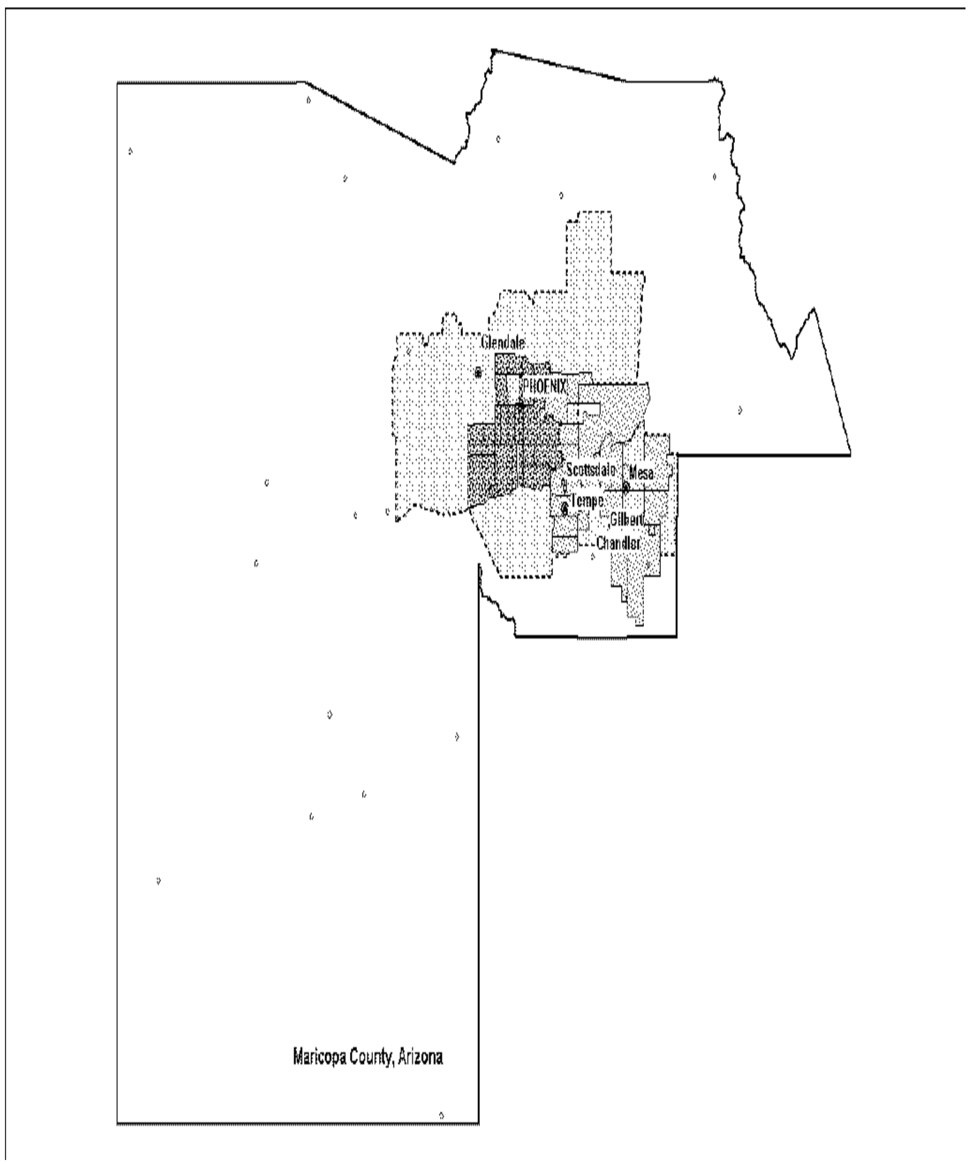

Figure 3.2: Map of Phoenix analysis regions.

variables used in this analysis. The correlation between fine and coarse particle daily levels is 0.65. Over the 3 year period, roughly 13% of the coarse particle data are missing, while approximately 8% of the $PM_{2.5}$ observations are missing. Interquartile ranges used to calculate the relative risk of death are 1.28 and 1.13 for the centered and scaled fine and coarse PM, respectively. Models are based on cases with complete data only (741 days).

Table 3.2 shows the 98th (99th) percentile and yearly averages for $PM_{2.5}$ (PM_{10}) for comparison with the EPA standards outlined in section 1.2. Although it appears that the PM measures did not violate EPA regulations, it should be noted that the EPA standard is based on imputed data (when applicable) for all sites for complete years. The summaries presented in Table 3.2 are based on only one monitor in the Phoenix area, exclude missing observations, and figures for 1995 and 1998 are based only on partial years.

Table 3.1: Summaries of Phoenix particulate matter variables. All measurements are in $\mu g/m^3$.

Variable	Mean	Median	Standard deviation	Maximum	Percent missing	IQR
PMC	33.28	31.16	15.20	158.6	12.8	17.19
$PM_{2.5}$	13.79	12.14	7.16	40.95	8.3	9.17

Additional daily meteorological data were obtained from the U.S. National Climatic Data Center (NCDC) in Ashville, NC, and include minimum and maximum daily temperatures (TMIN and TMAX). Specific humidity (SH) was derived from the NCDC data. To allow for nonlinear effects of temperature and humidity on mortality, we considered squared components of each (TMAXSQ, TMINSQ, and SHSQ). As discussed in chapter 1, there is no consistent agreement on which lags of PM and confounding variables to include. Therefore, we allow for potentially any lag from the present day (lag 0) up to a lag of 3 days (lag 3) to be included.

Table 3.2: Yearly summaries of PM_{10} and $PM_{2.5}$ measurements for the Phoenix site for comparison with EPA standards given in section 1.2. All measurements are in μg/m^3. [a]Averages and percentiles for 1990 are based on measurements for May 6, 1995 through December 31, 1995. [b]Averages and percentiles for 1998 are based on measurements for January 1, 1998 through March 30, 1998.

Year	PM_{10}		$PM_{2.5}$	
	99 %ile	average	98 %ile	average
1995[a]	101.7	49.0	31.4	15.3
1996	110.1	50.6	34.2	14.8
1997	106.9	45.5	27.1	12.5
1998[b]	74.4	33.2	26.5	11.0

3.2 Methods

For each of the four response variables, we model mortality using a Poisson regression model. This model is commonly used for response variables consisting of counts. We assume the mean of the daily mortality counts, $\mathrm{E}(Y) = \mu$, is related to the covariates, X, by

$$g(\mu) = X\beta,$$

where the link function $g(\mu)$ is defined as

$$g(\mu) = \log(\mu).$$

We also assume the variance function

$$V(\mu) = \phi\mu$$

with ϕ estimated within the model, rather than taken to be one. So, we assume that the log of expected mortality has a linear relationship to the predictor variables. Further details about Poisson regression models can be found in Chapter 2 of [21]. There are 70 days with no non-traffic related accidental deaths, so we used log(x + 1) as the response variable for the sensitivity analysis, where x is the daily non-traffic related death count.

We use wavelet trend decomposition with LA(8) wavelet filter ([9], Chapter 6) to estimate the baseline time trend in log(mortality). The trend estimate is based on 6 levels, corresponding to averages 64 days, or roughly two months, apart. Periods of one month or greater have been suggested as the time scale choice when accounting for long-term variation ([3]). This baseline was estimated separately for each of the four response variables and included in all subsequent analyses. We also adjust for the potential confounding meteorological variables discussed above. We consider particle size (fine and coarse) as well as the lag of the effect (0, 1, 2, or 3 days). With the nonlinear BASELINE, there are 29 variables under consideration. All predictor variables were centered and scaled by their mean and standard deviation.

3.3 Results

We will examine several outputs from model averaging to explore the possible association between elderly mortality and particulate matter in Phoenix. Table 3.3 contains the probability the relative risk is one given the data, the posterior mean of the relative risk, and the 95% probability interval for the relative risk estimate for all four analyses. The quantity in column two of this table is calculated as

$$\Pr(RR = 1|\text{Data}) \quad = \quad \sum_{m=1}^{J} \Pr(M_m|\text{Data}),$$

summing over only the $J \leq K$ models which do not contain at least one PM variable. Focusing on these values we see that the data seem to be marginally in support of an association between elderly mortality and particulate matter for the analyses using U2.5MORT and PHXMORT. The probabilities that the relative risk is greater than one given the data are 0.85 and 0.86, respectively, for the two areas. As with most of the output of these analyses, there is no clear-cut rule for what is “large enough” to be considered important, but should be determined by those making decisions based on the data. All four analyses yield relative risk posterior means slightly greater than one and all 95% probability intervals include one, indicating that if any association does indeed exist, the effect is small. Note that the evidence for an association between PM and ACCMORT is very small, as expected.

Table 3.3: Summaries from Bayesian model averaging for the four analyses.

Response Variable	Probability relative risk is one given data	Posterior mean relative risk	95% Probability Interval
U10MORT	0.32	1.02	[1, 1.05)
U2.5MORT	0.15	1.02	[1, 1.04)
PHXMORT	0.14	1.02	[1, 1.04)
ACCMORT	0.83	1.00	(0.99, 1.02)

Figure 3.3 shows, for each analysis, the posterior probability that the coefficient for each particulate matter variable is not equal to zero, given the data. Notably, the posterior probability (in percentage) that the coefficient for lag 1 coarse PM is not equal to zero is 75% for the U2.5MORT analysis, and that for lag 2 coarse PM is 68% in the PHXMORT analysis and 0.41 in the U10MORT analysis. All other coefficients in all analyses had posterior probabilities of not being equal to zero of less than 20%.

Figure 3.4 shows the distribution of relative risks based on the simultaneous change in both coarse and fine PM for each of the four response variables. All plots have a spike at 1.0 reflecting the number of models which do not include any PM variables. In the distribution for ACCMORT, most of the support is on relative risks equal to one, as anticipated. The distributions for the other three responses are more concentrated. As we saw above, there appears to be weak support for a PM association in the U2.5MORT and PHXMORT analyses. The histograms for both have a fair amount of weight dispersed among values greater than one.

Figures 3.5 and 3.6 contain model space plots and 95% probability intervals for relative risks for the top 25 models considered in the analysis for U2.5MORT and that for PHXMORT. The models have been ranked by their log(Bayes Factor) for comparing each model to the worst model. The log(Bayes Factor) does not differ much among these top models, however the conclusions one might draw from each vary widely. First consider the left panel of Figure 3.5, the model space plot for the U2.5MORT analysis. A black square

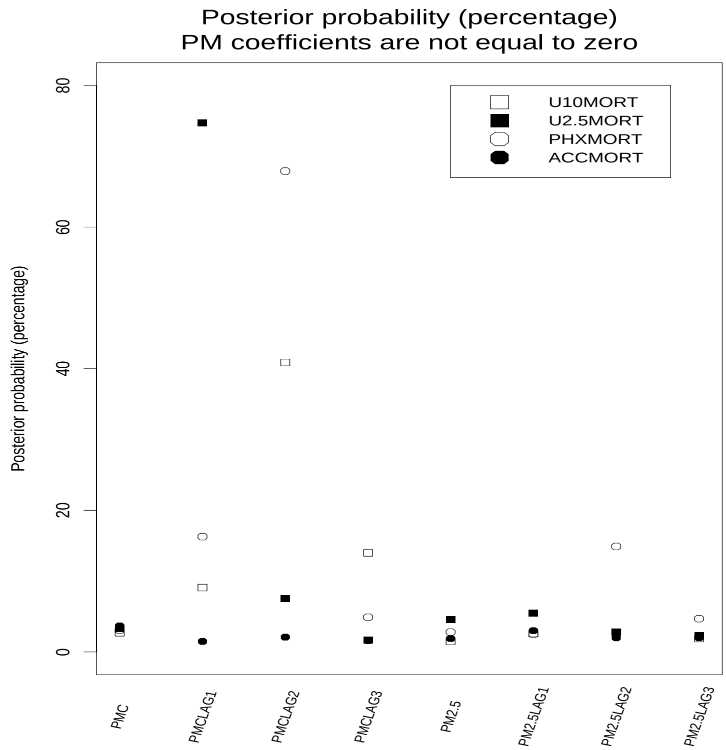

Figure 3.3: Posterior probability (percentage) that each PM coefficient is not equal to zero, given the data, for each of four analyses.

above a variable indicates that it is included in that model. The lag 1 coarse PM variable is present in 20 of the top 25 models. However, only one of the top models contains a $PM_{2.5}$ variable (lag 3 fine PM is included in the 23rd model). It is interesting to note that the best model contains lag 1 coarse PM and the baseline variable only, and the second best model includes only the baseline variable. The maximum and minimum temperature variables appear in many of the top models suggesting an association between extreme temperatures and non-accidental elderly mortality in the U2.5MORT region.

The right panel of Figure 3.5 shows the 95% probability intervals for the relative risk for the top 25 models. Most of the intervals do not include one, suggesting that perhaps increases in coarse PM are associated with a higher relative risk of elderly mortality in the U2.5MORT region. However, the probability interval for the relative risk corresponding to

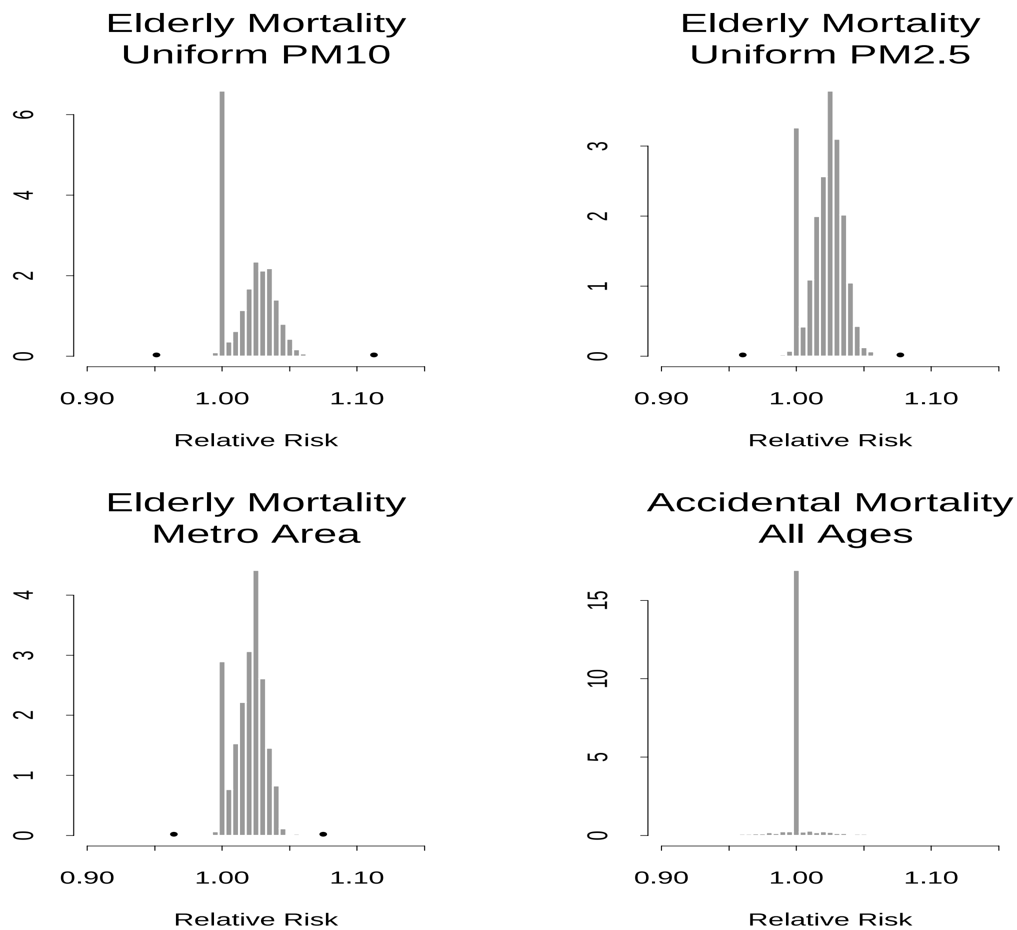

Figure 3.4: Distribution of relative risk for an increase of one IQR in PM_{10} lags 0-3 and one $PM_{2.5}$ variable, incorporating both estimation uncertainty and model uncertainty. The spikes at one correspond to models that do not include any PM variables. The points indicate the range of the distribution.

the only model with the $PM_{2.5}$ variable includes one providing evidence against an association between non-accidental elderly mortality and fine PM for the U2.5MORT region of Phoenix.

Now consider the model space plot for the PHXMORT analysis, shown in Figure 3.6. Now we see that lag 2 coarse PM is present in most (84%) of the top 25 models. Fine PM variables are included in only two of these. In this region, the top model includes only the baseline and the lag 2 coarse PM variable, but the third model contains these two plus the

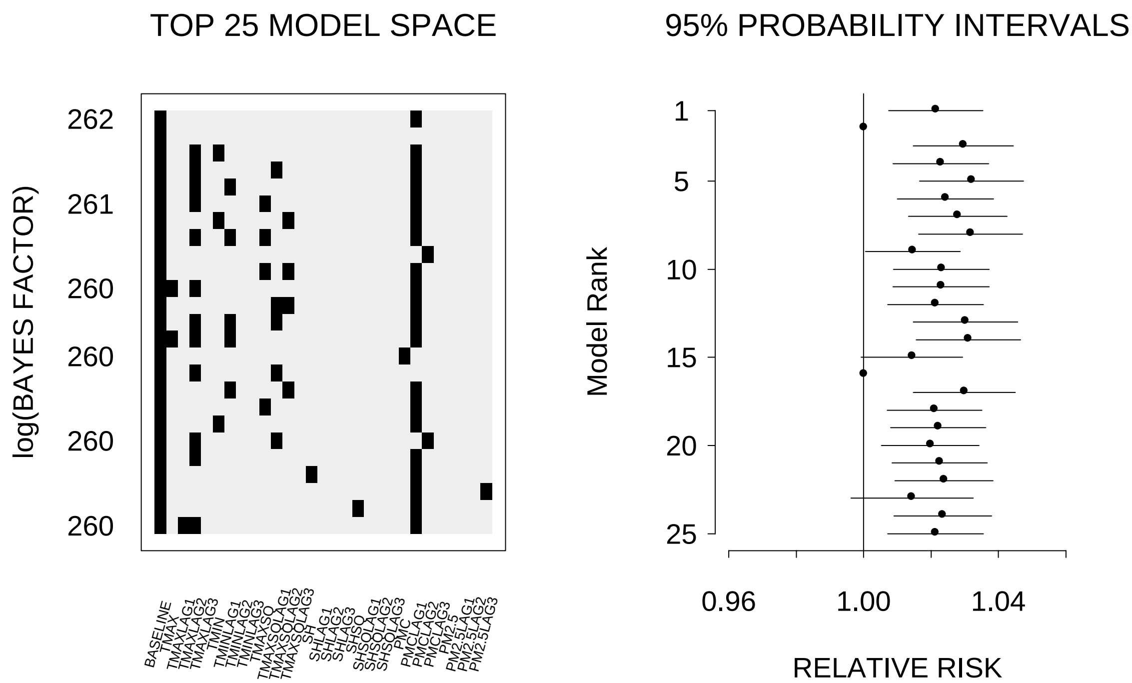

Figure 3.5: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk under each model using elderly mortality for the region with uniform $PM_{2.5}$. Rows in the left figure correspond to models and columns correspond to variables, with black squares indicating that the variable for that column is included for that row. The y-axis for the model space plot is the log(Bayes Factor) for comparing that model to the lowest probability model and is proportional to $-$BIC. Points in the probability intervals are the maximum likelihood estimates of relative risk under that model.

lag 2 fine PM variable. However, the 95% probability interval for the relative risk for this model includes one (see the left panel of Figure 3.6). Most of the other intervals do not include one, again suggesting an association between coarse PM, but not necessarily fine PM, and elderly mortality in the PHXMORT region.

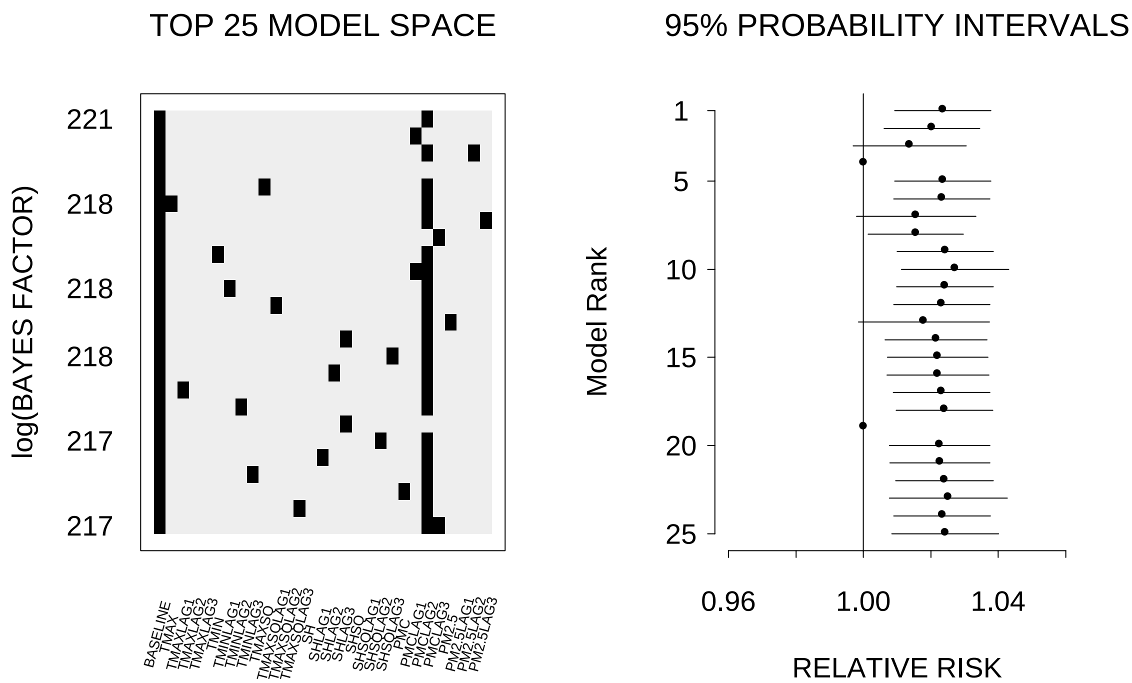

Figure 3.6: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk under each model using elderly mortality for the Phoenix metropolitan area.

3.4 Discussion

For this analysis of the association between non-accidental elderly mortality and particulate matter levels in Phoenix, Arizona, we attempt to take account of some of the uncertainty inherent in model selection methods by using Bayesian model averaging. We used three response variables, based on distinct geographic areas, and found relative risk estimates for each based on simultaneous increases in fine ($PM_{2.5}$) and coarse ($PM_{10}-PM_{2.5}$) particulate matter. We found some suggestion of an association in the area thought to have uniform PM2.5 levels (U2.5MORT) and the Phoenix metropolitan region (PHXMORT). Both had relative risk estimates of 1.02, suggesting a weak effect, if any. This association appears to be due to the coarse fraction PM, as indicated by the posterior probabilities for each coefficient being not equal to zero, and the frequency of inclusion of coarse and fine PM variables in the top 25 models (Figures 3.3, 3.5 and 3.6). We found no association between non-traffic related accidental mortality and particulate matter measurements, as expected.

One important finding from this analysis is that the study region used for the response variable can make a difference in one's findings. How this is typically chosen is not clear in most papers, and can be added as another source of model uncertainty which is not taken account of. Perhaps this area should be chosen based on prior knowledge of the spatial homogeneity of particulate matter levels, as we attempted to do here, rather than population boundaries. If this analysis were to be expanded to include other monitors in the area and spatial variation of PM, then a study region based on population groups might be more appropriate.

There are many modifications that could be made in this analysis to improve the models considered. We have not used all of the information we have about PM and mortality in Phoenix since we ignored days with missing observations. We could attempt to impute this data using a Bayesian approach and account for the uncertainty in the imputing process. This would certainly be a more efficient analysis and may change the results of our study. We also have assumed that the data are missing completely at random which may or may not be reasonable. We have no information about the missingness, and assuming otherwise may change our results and conclusions.

One common issue with Bayesian analyses is the choice of a prior distribution. Here we have used flat priors. However, non-flat priors on variables and lags could be incorporated based on the results of previous studies. One possible problem with using other analyses is that it appears that PM associations and lag structures may vary by city and time-period, so choosing "appropriate" priors would be tricky. In these initial studies of PM using BMA, flat priors may provide the best method for exploring the top models for each data set.

In many PM studies measurements from other pollutants, such as SO_2, CO, NO_2 and O_3, are included as potential confounding variables or variables of interest. If any or all of these pollutants are associated with elderly non-accidental mortality in Phoenix, our relative risk estimates for each model could be biased.

As mentioned in chapter 1, deciding if a variable should be transformed or not is a source of model uncertainty. Here we have assumed a simple log-linear relationship between elderly non-accidental mortality and all particulate matter variables. Exploratory techniques did not reveal a need for more complex relationships, however these could be considered further and included in a BMA process. It has been noted ([3], [8]) that the choice of time scale used to remove long-term variation may influence the results of PM studies. We have chosen our scale here based on a suggestion in [3], but have not explored how changing the scale may change our conclusions. This could certainly be done within the BMA framework and methods for doing so are currently being refined.

Chapter 4

PARTICULATE MATTER AND MORTALITY IN SEATTLE, WASHINGTON

4.1 Previous analyses of particulate matter and adverse health effects in Seattle

We present one recent analysis of PM and mortality in Seattle, Washington for a comparison to that done here. This study done by Samet et al. ([28]) included analyses of the association between PM and elderly mortality and morbidity for the 20 and 90 largest cities in the United States using hierarchical models. For Seattle, results were presented for response variables of total elderly mortality, elderly mortality with cardiorespiratory causes, and elderly mortality caused by other diseases for 1987−1994. The most relevant analysis used a log-linear model for mortality and included a smooth function of time to account for long-term variation, smooth functions for average and dewpoint temperature, smooth functions for the average and dewpoint temperature for the previous three days and an indicator for day of the week. Models included estimates of an overdispersion parameter and excluded days with any missing observations. Pollution variables of interest were PM_{10}, CO, and O_3. Daily measurements were averages of those from up to 8 monitors in the Seattle area. Initial models included only PM_{10}, O_3, or CO. The two and three pollution variable combinations were also included and evaluated. The increase in mortality (in percent) was calculated for a 10 μg/m^3 increase in PM_{10}. Lags 0−2 were primarily explored.

The main analyses from the Samet et al. study reported no significant association between elderly mortality and PM_{10} in Seattle, Washington for any response variable with or without adjustment for other pollution variables. However, another study presented in Appendix B of Samet et al. did find some evidence of an association between PM_{10} and elderly mortality in Seattle, but the focus of the analysis was primarily on exploring the lag structure. The time period studied was 1986−1993. The model used included smooth functions of temperature, humidity, barometric pressure, day of the week and seasonal patterns. The analyses included four different lag models and results varied by which lag structure was used. The estimated percentage of increase in daily deaths in Seattle, and their associated standard errors, for a 10 $\mu g/m^3$ increase in PM_{10} were 0.70(0.23), 0.65(0.31), 1.46(0.31) and 1.46(0.34) for the four lag structures considered. This suggests a possible small association between PM_{10} increases and elderly mortality in Seattle.

We also review two recent studies of asthma hospital admissions and particulate matter levels in the Seattle area. The first, reported by Norris et al. in 1999 ([22]), examines daily emergency department visits for asthma by children aged 18 and younger at Seattle hospitals. The admissions counts were broken into two regions, high and low utilization, based on the number of visits. The data covered the period from September 1, 1995 through December 31, 1996. The two main particulate matter variables of interest were PM_{10} and nephelometer data. Pollution variables SO_2, NO_2, CO and O_3 were also included. Meteorological variables adjusted for were average daily temperature and dewpoint temperature. These two variables, as well as a time-trend, were included as smooth functions estimated by smoothing splines. The smoothing spline chosen for time-trend is approximately equivalent to two-month moving averages. Degrees of freedom for smooth functions of temperature and dewpoint were chosen by minimizing the deviance, adjusted for degrees of freedom. These three smooth functions, and a day-of-week factor variable, were included in a baseline semiparametric Poisson regression model. Pollution variables (lags 0−4) were then added individually to the model and relative rates based on IQR increases (11.6 $\mu g/m^3$ for PM_{10} and 0.3 $m^{-1}/10^{-4}$ for nephelometer data) were calculated.

Norris et al. ([22]) found a significant increase in the relative rate of children's emergency department asthma visits for both nephelometer data and PM_{10} for the low utilization study region and the pooled study region. Only nephelometer reading increases were associated with higher relative rates for the high utilization region. Point estimates for these significant relative rates range from 1.13 to 1.16. Although it is not entirely clear from the article, it seems that these are based on the daily averages (lag 0) of the PM variables only. A few multi-pollutant models were also examined, which contained either PM_{10} or the light scattering variable, and NO_2 and SO_2. Again, the relative rates of childhood asthma emergency department visits for an IQR increase in PM_{10} or nephelometer measurement were significant. In this study the authors also converted the light-scattering IQR to "represent $PM_{2.5}$ gravimetric mass based on colocated nephelometer and $PM_{2.5}$ monitors at the southernmost PM monitoring site" ([22], p. 492), and suggest that an increase of approximately 9.5 $\mu g/m^3$ in $PM_{2.5}$ would also be significantly associated with an increase in emergency department visits for asthma by children.

The second study reviewed here, presented by Sheppard et al. in 1999, focuses on hospital asthma admissions of patients younger than 65 years, in 23 Seattle area hospitals for the period January 1, 1987 through December 31, 1994 ([34]). Primary particulate matter variables of interest were PM_{10}, $PM_{2.5}$, nephelometry data and coarse PM ($PM_{10}-$ $PM_{2.5}$). Lags $1-3$ were the primary focus, but lag 0 and longer lags were also explored for some variables. Other variables included in the model were daily average temperature, SO_2, CO and O_3. The first stage of the analysis included multiple imputation of $PM_{2.5}$ ($72-81\%$ missing depending on site), PM_{10} ($4-40\%$ missing), and SO_2 (6% missing). Semiparametric Poisson regression modeling was used in the second stage to create a base model with a day-of-week factor variable and nonparametric smooth functions of time (64-df smoothing spline) and temperature (4-df smoothing spline). The degrees of freedom for the time trend corresponds to about a 46-day moving average filter. These smooth functions were chosen based on "conceptual simplicity, small Akaike Information Criteria (AIC), and little evidence of over- or underfitting in the residual autocorrelation function" ([34], p.26). Pollution variables were added one at a time and their statistical significance evaluated by Wald statistics and AIC. Relative rate estimates were calculated for a one IQR increase in a pollutant. All variables with a relative rate interval estimate entirely above one were then included in a multi-pollutant model.

In the single pollutant models PM_{10}, $PM_{2.5}$, and coarse PM, all lagged one day, showed significant associations with asthma admissions (relative rate point estimates (RR): 1.05, 1.04, 1.04, respectively; IQR: 19 $\mu g/m^3$, 11.8 $\mu g/m^3$, 9.3 $\mu g/m^3$, respectively). PM_{10} and $PM_{2.5}$, lag 0, also had significant effects (RR: 1.03, 1.03). The multi-pollutant model (including PM_{10}, $PM_{2.5}$ and coarse PM) also showed a significant association between PM increases and asthma admissions ([34]).

4.2 Data

We obtained daily counts of non-accidental deaths for people age 65 and over for a selection of zip codes in King County, Washington from the Washington state death certificate database for January 16, 1990 through December 31, 1995. The study area chosen contains the following King County residential zip codes: 98004-98008, 98011, 98031-98034, 98038, 98040, 98042, 98052, 98055-98056, 98058-98059, 98101-98109, 98112, 98115-98119, 98121-98122, 98125, 98126, 98133-98136, 98144, 98146, 98148, 98155, 98158, 98166, 98168, 98177-98178, 98188 and 98198-98199. We selected this region in part to compare our results with those from the Sheppard et al. study described above ([34]). As a sensitivity analysis we ran the same set of five analyses using daily hospital admissions for appendicitis as the response variable. Since there is no evidence to suggest that PM levels should be associated with appendicitis, this should be a suitable control group. The top panels of Figure 4.1 show these two response variables.

Of the twelve EPA particulate matter monitoring sites in the Puget Sound region, only three sites collect $PM_{2.5}$ data: Kent, Duwamish and Lake Forest Park. None of the sites collect the $PM_{2.5}$ data daily so all have a great number of days with missing measurements. We chose to use data from the Duwamish site as it had the least amount of missing $PM_{2.5}$ values, although the percentage missing is still very high (75.9%). We also obtained PM_{10} data (0% missing), nephelometer data (0.6% missing) and SO_2 measurements (2.2% missing) from this site. Figure 4.1 shows time series plots for each PM variable. Relative risk estimates for analyses which contain PM_{10} and $PM_{2.5}$ variables are based on a simultaneous interquartile range (IQR) increase in centered and scaled PM_{10} (1.07) and centered and scaled $PM_{2.5}$ (1.21).

Table 4.3 gives the 98th percentile and average mean for $PM_{2.5}$ and the 99th percentile and average mean for PM_{10} for each of the study years. It should be noted that although it appears that the $PM_{2.5}$ three-year average levels exceed the NAAQS standard given in section 1.2, the EPA standards are based on all study sites when more than one is present.

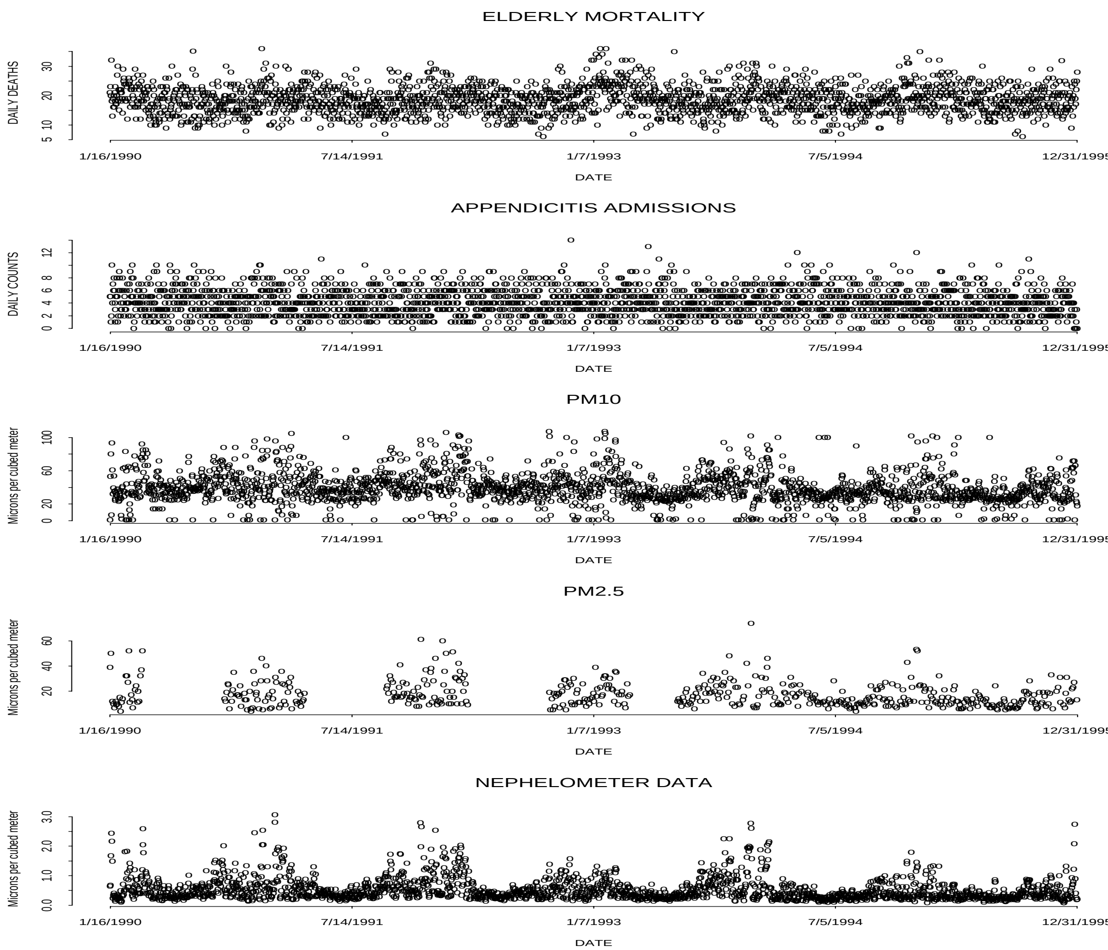

Figure 4.1: Time series of the two response variables, PM_{10}, $PM_{2.5}$ and nephelometer data for Seattle, WA from January 16, 1990 to December 31, 1995.

Table 4.1: Summaries of Seattle pollution variables. All measurements are in $\mu g/m^3$.

Variable	Mean	Median	Standard deviation	Maximum	Percent missing	IQR
PM_{10}	39.16	36	18.62	107	0	20
$PM_{2.5}$	17.44	15	10.12	74	75.9	12.25
Nephelometer	0.53	0.4	0.40	3.07	0.6	0.38
SO_2	0.009	0.004	0.008	0.029	2.2	0.004

Table 4.2: Correlation coefficients for Seattle pollution variables, based on days for which both measurements are complete.

	PM_{10}	$PM_{2.5}$	Nephelometer	SO_2
PM_{10}	1.0	0.61	0.53	0.32
$PM_{2.5}$		1.0	0.91	0.40
Nephelometer			1.0	0.34

Also, the EPA uses an imputation formula for missing data. Since we do not have data for the other monitors in the Seattle region, have not attempted to impute data for missing observations, and only use partial data for 1990, we cannot conclude if any EPA regulations were indeed violated.

We expect different particulate matter levels depending on whether it is a weekend or weekday because of dissimilar traffic and industrial pollution levels during those periods. Indeed, the average PM_{10} measurement on weekends is 34.2 $\mu g/m^3$, while that on weekdays is 41.7 $\mu g/m^3$, and that of $PM_{2.5}$ is 15.5 $\mu g/m^3$ on weekends and 18.2 $\mu g/m^3$ on weekdays. Therefore, we also included an indicator variable for weekends. It should be noted that in [34] a factor variable for day of the week was included. However, since we chose not to impute any $PM_{2.5}$ data and most of these measurements were made on Tuesdays (28.6%) Thursdays (29.4%) or Saturdays (26.0%), an indicator variable for weekend was deemed more appropriate. Daily temperature and dewpoint temperature data at Sea-Tac airport are also included as these types of meteorological variables are typically included in epidemiological PM studies using model selection procedures. See Table 4.1 for summary measures of a selected group of these variables. A centered and scaled version of average daily temperature lag 0−3, average dewpoint temperature lag 0−3 and SO_2 lag 0−3 is included in the model averaging process, along with the weekend indicator and particulate matter variables.

Table 4.3: Yearly summaries of PM_{10} and $PM_{2.5}$ measurements for the Duwamish site for comparison with EPA standards given in section 1.2. All measurements are in $\mu g/m^3$. [a]Averages and percentiles for 1990 are based on measurements for January 16, 1990 through December 31, 1990.

Year	PM_{10}		$PM_{2.5}$	
	99 %ile	average	98 %ile	average
1990[a]	88.0	41.2	51.6	18.9
1991	96.4	42.0	40.6	19.4
1992	101.4	42.5	50.6	20.0
1993	98.4	39.7	43.7	20.5
1994	100.0	34.8	37.0	15.2
1995	97.4	34.8	32.6	14.1

4.3 Methods

We again model the mean of the daily death count ($\mathrm{E}(Y) = \mu$) using Poisson regression assuming

$$g(\mu) = X\beta,$$

and

$$g(\mu) = \log(\mu).$$

As we have varying degrees of overdispersion in these models, we use the variance function

$$V(\mu) = \phi\mu$$

where ϕ is estimated, rather than assumed to be one. There are 39 days with no hospital appendicitis admissions, so we used log(x + 1) as the response variable for the sensitivity analyses, where x is the daily hospital appendicitis admissions count.

As discussed earlier, long-term trends in mortality must be taken account of in these epidemiological studies. For this analysis we use wavelet trend decomposition with LA(8) wavelet filter to estimate the baseline time trend in log(mortality) ([9], Chapter 6). The trend estimate is again based on 6 levels, corresponding to averages 64 days, or roughly two months, apart. This is consistent with degrees of freedom choices made in the studies reviewed above ([28], [22], [34]). This baseline is included in all subsequent analyses.

The Seattle hospital asthma admissions study mentioned above ([34]) included multiple imputation to fill in the missing values for $PM_{2.5}$ from this monitoring site based on the high correlation between $PM_{2.5}$ and other pollutants. However, we have a correlation of only 0.61 between $PM_{2.5}$ and PM_{10} on available days (see Table 4.2). We do not feel this correlation is high enough, nor do we have enough available data, to justify filling in over 75% of the $PM_{2.5}$ data. Because of this limitation, if we include $PM_{2.5}$ and lags $1-3$, we have no days with complete data. Therefore, we did one model-averaging procedure excluding $PM_{2.5}$ and one including $PM_{2.5}$ for each of the lags $0-3$. Ideally we would have used $PMC = PM_{10} - PM_{2.5}$ as our measure of coarse particulate matter, but we would have again been greatly hampered by the amount of missing $PM_{2.5}$ data.

Because of the high correlation between the nephelometer data and $PM_{2.5}$, for days with complete data, we also did one analysis including PM_{10} and the nephelometer data. We then calculated the relative risk of non-accidental elderly mortality for a simultaneous increase of one IQR in PM_{10} and nephelometer measurements. This provided a way to try to quantify the association between mortality and fine PM using more of the available data. All analyses only include days for which all data is complete.

4.4 Results

Table 4.4 gives the posterior probability that the relative risk of non-accidental elderly mortality is equal to one given the data, the posterior mean relative risk and the 95% probability intervals for the five analyses. All confidence intervals include one, suggesting that the relative risk does not significantly change given an interquartile range increase in particulate matter concentration. The posterior probabilities that the relative risk is equal to one range from 0.48 to 0.81, giving very little evidence for an association between particulate matter and elderly mortality.

Table 4.4: Summaries from Bayesian model averaging for six analyses. Each uses elderly non-accidental mortality as the response variable and includes differing measures of fine particulate matter.

PM variables which RR is based on	Probability relative risk is one given data	Posterior mean relative risk	95% Probability interval
PM_{10} lags 0-3	0.81	1.00	[1, 1.01)
PM_{10} lags 0-3 and $PM_{2.5}$ lag 0	0.69	1.00	[1, 1.02)
PM_{10} lags 0-3 and $PM_{2.5}$ lag 1	0.78	1.00	(0.99, 1.01)
PM_{10} lags 0-3 and $PM_{2.5}$ lag 2	0.79	1.00	(0.99, 1.01)
PM_{10} lags 0-3 and $PM_{2.5}$ lag 3	0.48	1.01	[1, 1.04)
PM_{10} lags 0-3 and Neph. lags 0-3	0.73	1.00	[1, 1.01)

Figure 4.2 shows, for each analysis, the posterior probability (in percentage) that the coefficient for each pollution variable is not equal to zero, given the data. All but two of these posterior probabilities are below 20%. One, the PM_{10} lag 0 coefficient from the analysis including $PM_{2.5}$ lag 3 variable, is almost 40%. The probability that the coefficient for SO_2 in the analysis including the $PM_{2.5}$ lag 1 variable is not equal to zero is about 60%. Although these probabilities are much higher than the others presented in this plot, they are still too low to be considered very important.

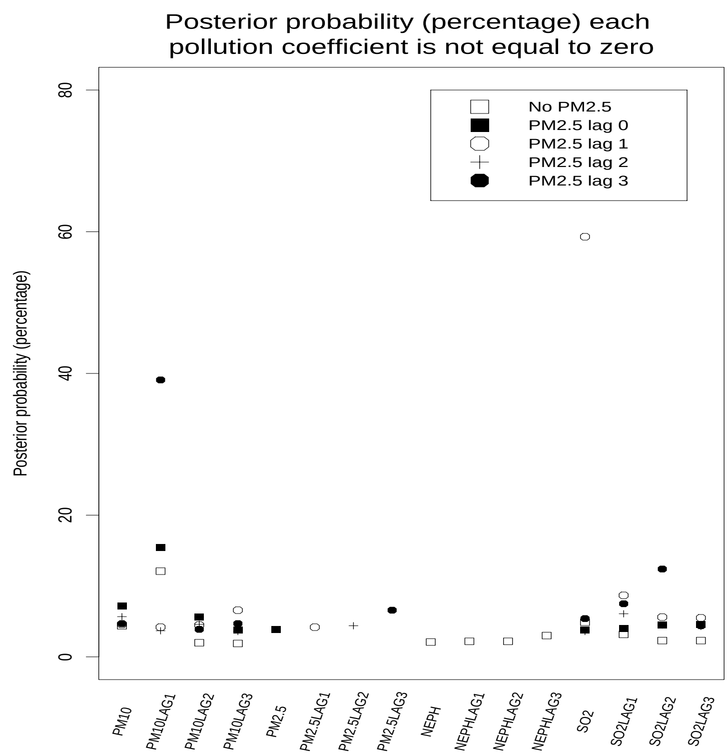

Figure 4.2: Posterior probability (percentage) that the coefficient for each pollution variable is not equal to zero, given the data, for five analyses.

Figure 4.3 shows the relative risk distribution for each of the four analyses. The points indicate the range of the distribution. We can see that because of the small amount of usable data, the range is quite large. The spike at one in each graph is due to the number of models that do not include any PM variables. Most of the weight of the distribution is contained in this point-mass at one, again supporting a conclusion of no association between PM and elderly mortality in Seattle.

Figures 4.4, 4.5, 4.6, 4.7 and 4.8 show the top 25 models by posterior model probability. It is interesting that no variable appears to be consistently in the top models except the baseline variable in any of these plots, with the exception of the SO_2 variable in the plot for the analysis including $PM_{2.5}$ lag 1 (Figure 4.6). Since temperature and dewpoint temperature are typically included in all PM studies, we may have expected them to be in more of the

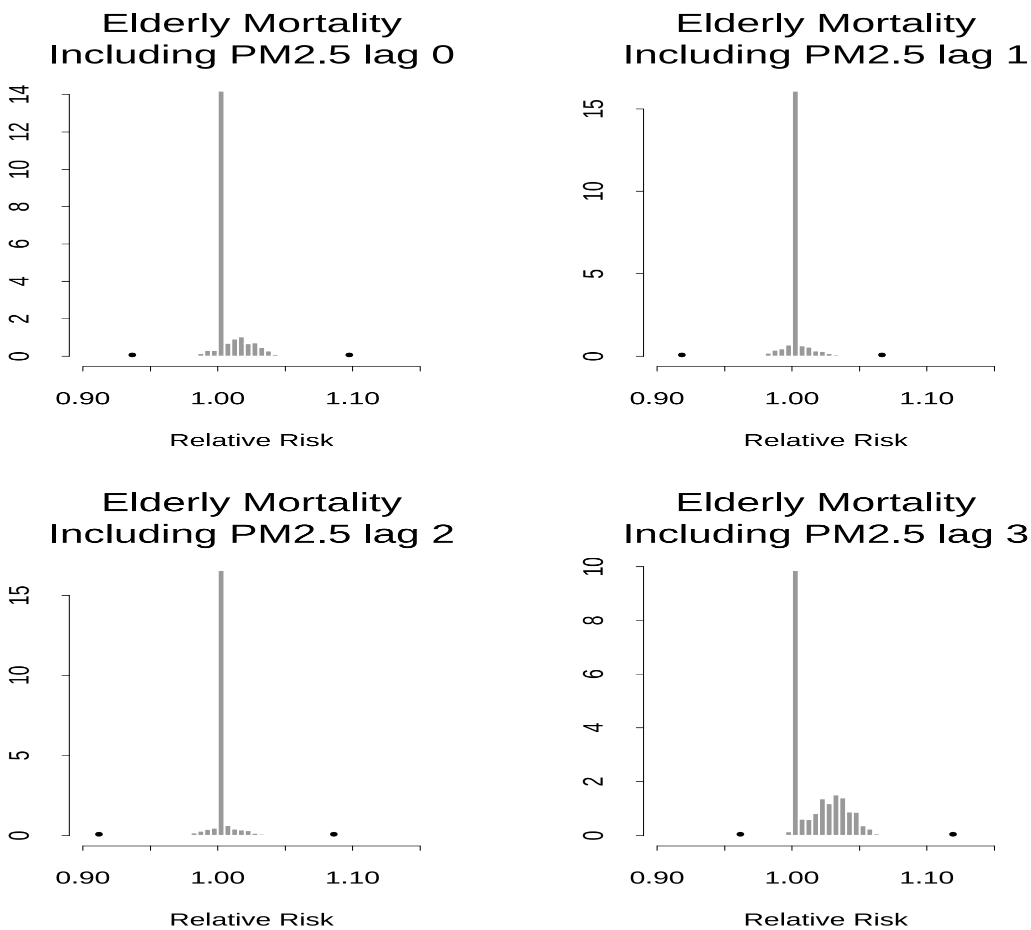

Figure 4.3: Distribution of relative risk for an increase of one IQR in PM_{10} lags 0-3 and one $PM_{2.5}$ variable, incorporating both estimation uncertainty and model uncertainty. The spikes at one correspond to models that do not include any PM variables. The points indicate the range of the distribution.

top models.

Most of the top models for the analyses including $PM_{2.5}$ lag 1 or lag 2 are relatively simple and do not include any PM variable, with some even suggesting a possible protective effect of PM increases. The BMA process which included $PM_{2.5}$ lag 0 shows 12 of the top 25 models including a PM variable, however all but one of the 95% probability intervals for the relative risk estimates in these models include one (Figure 4.5). Now consider the case where PM_{10} lags 0-3 and $PM_{2.5}$ lag 3 are included (Figure 4.8). In that analysis, ten of the top 25 models include a PM_{10} variable and one includes $PM_{2.5}$ lag 3. Seven of the 95% probability intervals for the relative risk do not include one, all of which are for models without $PM_{2.5}$ lag 3. Although this plot seems a bit more interesting than those from the other four analyses, it is consistent with our results above, namely that there is little evidence of an association between mortality and PM in this data set.

As mentioned in section 4.3 we calculated the relative risk of non-accidental elderly mortality for a simultaneous increase of one IQR in PM_{10} and nephelometer data. Table 4.4 contains the results of these calculations. We see that there still appears to be no association between mortality and PM, even though we were able to use almost all of the available data. No significant PM association was found in any of the analyses in which we used hospital appendicitis admissions as our response variable. Since this was done primarily as a sensitivity analysis, specific results are not given here.

4.5 Discussion

In this study we explored the association between non-accidental elderly mortality and particulate matter levels in Seattle, Washington. We used Bayesian model averaging to take account of some of the model uncertainty inherent in model selection processes. All output from this analysis indicated that there is no association between the variables of interest. We found no increase in the estimated relative risk of death for most models, and an estimate slightly above one (1.01) for the analysis including $PM_{2.5}$ lag 3. We also saw no consistent inclusion of variables typically included in PM studies, such as daily temperature and dewpoint temperature. The only variable with a clear relationship to non-accidental elderly mortality is the baseline variable estimated by wavelet trend decomposition.

The conclusions we have drawn for elderly mortality and PM_{10} are consistent with those from the Samet et al. study ([28]). Recall that the main analysis used a model very similar to that presented here and found no significant association between the two. The study done in Appendix B of the same paper included relative humidity and barometric pressure and did not include dewpoint temperature. Results varied by lag structure used, but all significant associations suggested only a small increase in elderly mortality.

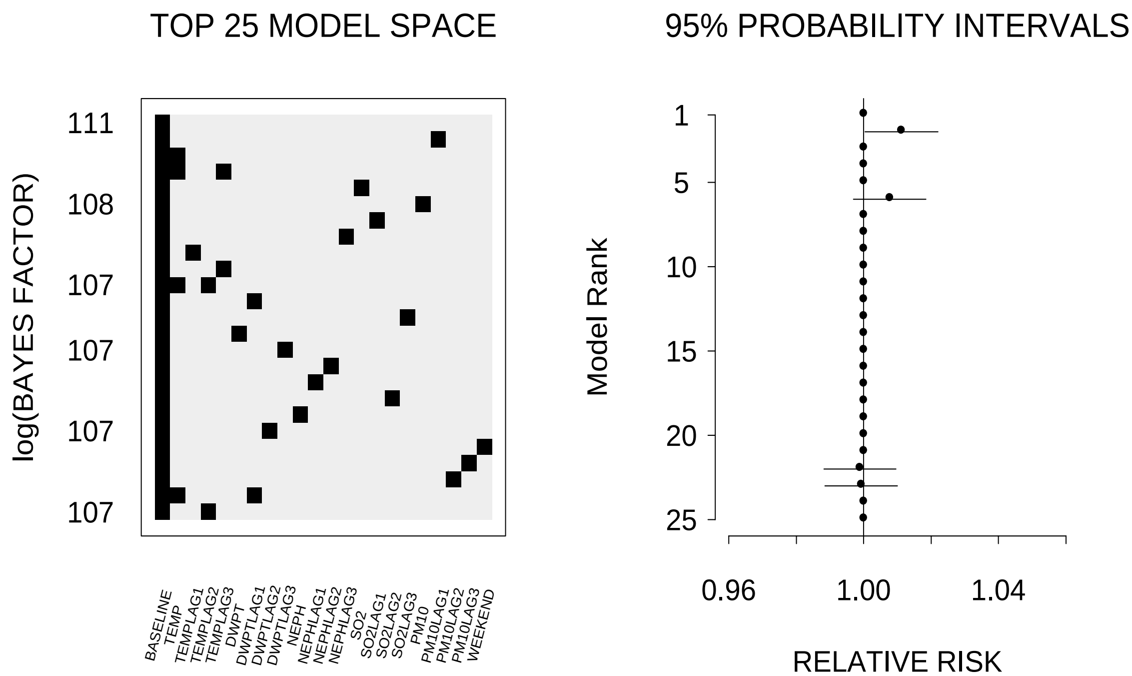

Figure 4.4: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk under each model using elderly mortality and including no $PM_{2.5}$ variables. Relative risk estimates are based on an increase of one IQR in PM_{10}.

Our results are not consistent with those found by Norris et al. ([22]) or Sheppard et al. ([34]). Both of these studies found an increased relative risk of adverse health effects associated with an increase in PM levels. Of course, the comparison between this analysis and those in [22] and [34] is not ideal since we used a different response variable, time period and baseline estimation method. This suggests a few possible explanations for the differing results. Perhaps PM in Seattle is truly associated with asthma-related hospital visits but not with mortality, or at least not during the time period we studied. The baseline estimation method may also be a factor.

It has been noted in [8] that the technique for estimating the long-term variation probably will not make a difference in the results of PM studies, but the time scale may matter. We

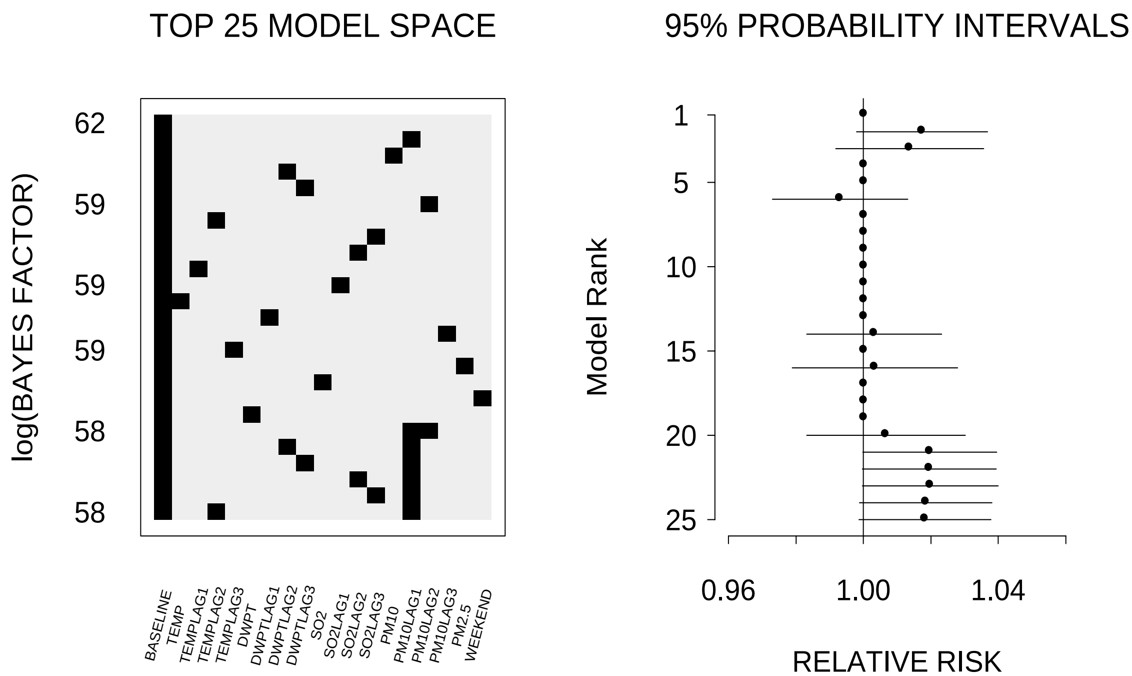

Figure 4.5: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk of elderly mortality, based on a simultaneous increase of one IQR in PM_{10} and $PM_{2.5}$ lag 0.

used a window of 64 days here which is consistent with suggested time scales and with those used in [28],[22], and [34], but we did not explore the effects of changing this baseline. The fact that no other variable except the baseline was consistently in the top models suggests that perhaps the baseline estimate is picking up too much of the small-scale variation. Choosing a higher level, so that we are using averages over a longer time periods, might be appropriate. In this study, we did not incorporate any means to take account of the uncertainty in this baseline, but this could be included in future work.

Besides exploring the effect of the baseline estimation, there are other alterations that could be made to this study which may change the above results and conclusions. One large problem with this data set is the amount of missing data. The data that were readily available from the Seattle monitoring sites was through December 31, 1995. One can see in

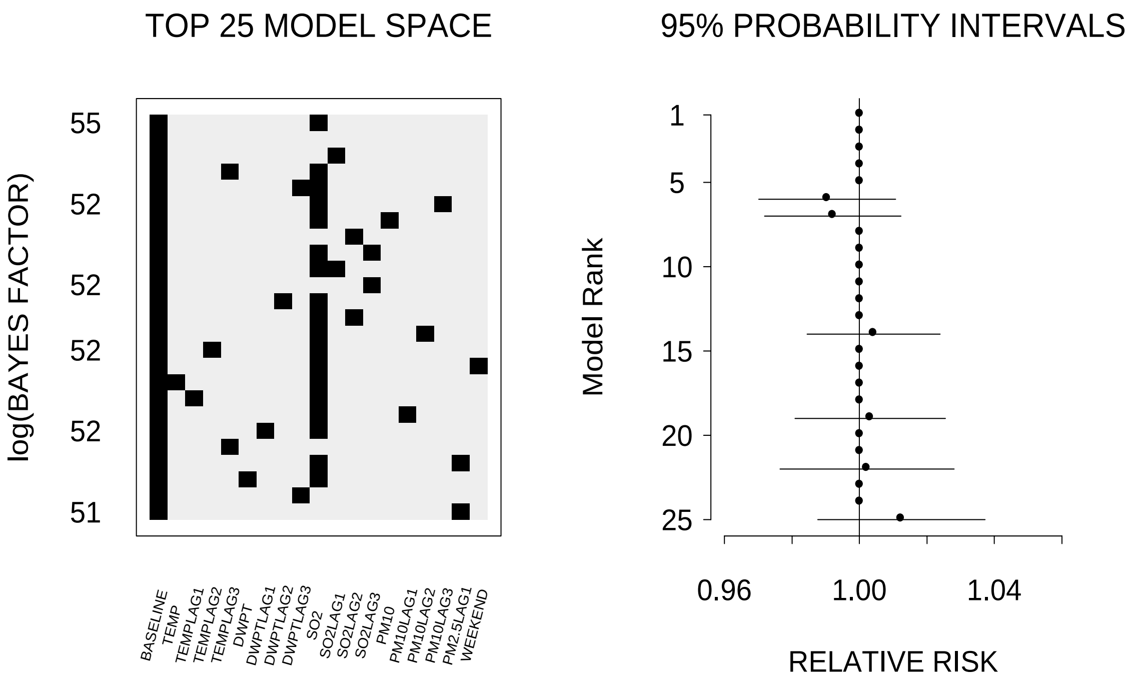

Figure 4.6: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk of elderly mortality, based on a simultaneous increase of one IQR in PM_{10} and $PM_{2.5}$ lag 1.

Figure 4.1 that after about March, 1993 the $PM_{2.5}$ data was collected much more frequently, although still about only every third day. A future analysis may be able to use more data after this date to conduct a more efficient study. However, the nephelometer analysis suggests that the amount of missing $PM_{2.5}$ data did not prevent us from detecting an association between fine PM and mortality. We also treated these data as being missing completely at random. This may be reasonable since most is missing because $PM_{2.5}$ was not measured in the summers of 1990−1993, and because measurements seem to have been taken every three days. If we assume the data are not missing at random, our results and conclusions may change.

We had access to data for a wind stagnation variable measured from the Duwamish site. This variable is defined as the number of hours each day that the wind speed at the

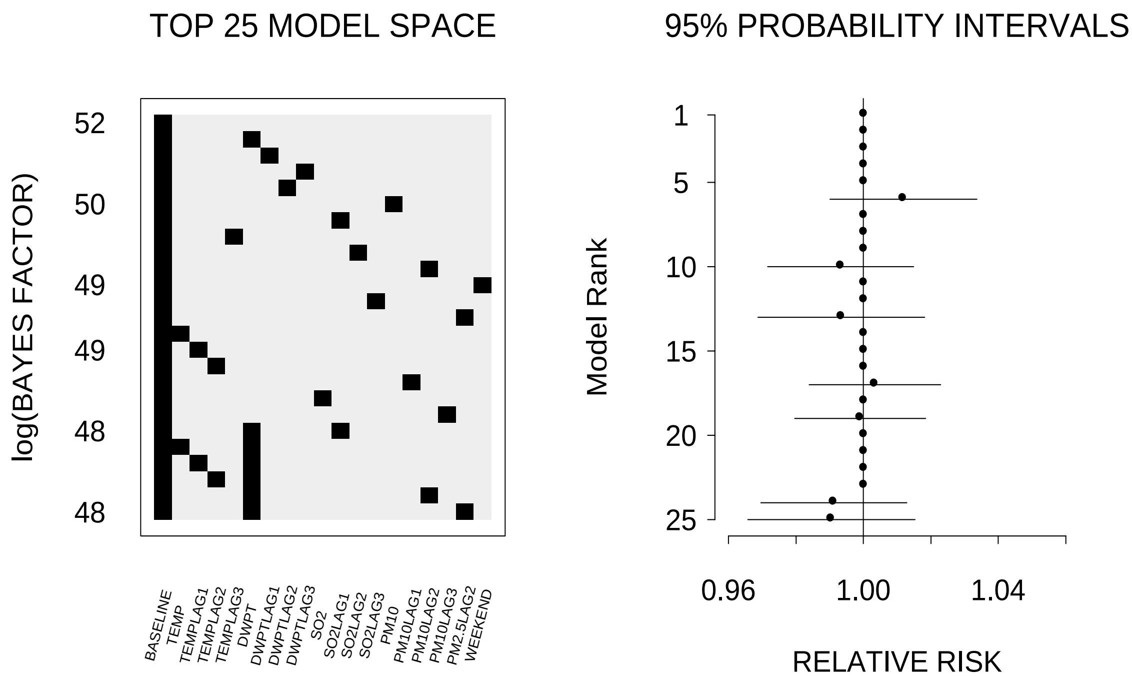

Figure 4.7: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk of elderly mortality, based on a simultaneous increase of one IQR in PM_{10} and $PM_{2.5}$ lag 2.

monitoring site was below the 25th percentile for all hourly wind speed measurements from that monitor. Initial models included this variable as a possible confounder. However, upon closer inspection we decided that there was no scientific basis for an association between mortality and wind stagnation so we chose not to include it. Also, it is not present in the other analyses used for comparison. It should be noted that inclusion of this variable did not change the conclusions of our study.

As in the analysis of the Phoenix, Arizona elderly mortality data, we have used flat priors and assumed a simple log-linear relationship between our response and all predictor variables. With more external information, we could attempt to redo this analysis using appropriate non-flat priors. The analysis could also be extended to include smooth functions of temperature and dewpoint temperature, or other variables, as was done in other studies

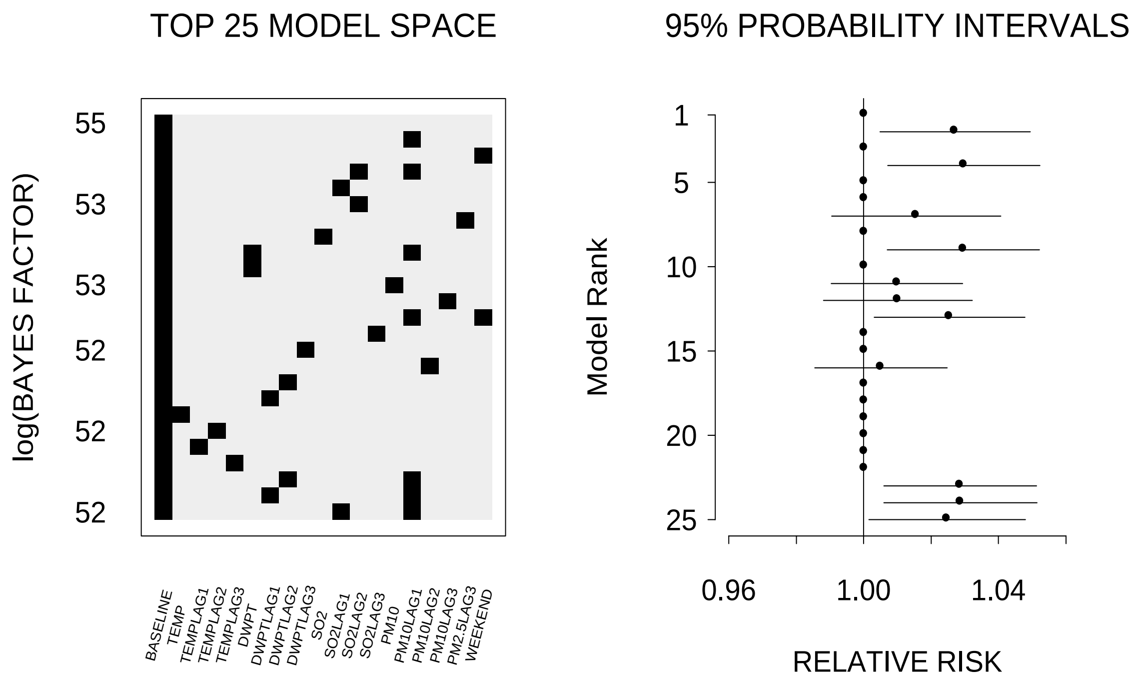

Figure 4.8: The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk of elderly mortality, based on a simultaneous increase of one IQR in PM_{10} and $PM_{2.5}$ lag 3.

([28], [22], [34]).

Chapter 5

DISCUSSION

5.1 Using Bayesian model averaging in particulate matter studies

The use of Bayesian model averaging in particulate matter studies provides a methodological advance by accounting for both parameter uncertainty and model uncertainty in relative risk point estimates and probability interval estimates. Unlike traditional p-values, the Bayesian approach provides posterior probabilities of whether there is a PM effect which incorporate model selection uncertainty. This reduces concerns about finding associations between PM and adverse health effects that do not truly exist, due to model selection methods.

Bayesian model averaging also allows the researcher to look at all the top models, as determined by their posterior probability given the data, in one plot (eg. Figure 3.5). These plots may help readers gain a deeper understanding of the relationships between variables in a dataset than they can by examining the results of only one or two models. Rather than giving one p-value and one confidence interval, they show which variables are consistently in the top models, and therefore which may be important predictors of the response variable.

This method may also be used to learn more about the lag structure of the particulate matter variables in PM studies. If we again consider Figures 3.5 and 3.6 from section 3.3 we see that coarse PM lagged one day is the primary PM variable included in the top models for the smaller region for which $PM_{2.5}$ is thought to be spatially homogeneous (U2.5MORT). However in the larger region, PHXMORT, coarse PM lagged two days is the primary PM variable included in the top models. While we will not attempt to pose any epidemiological or atmospheric hypotheses here, we will suggest that the ability to compare which lags are consistently included in top models may provide evidence to form such hypotheses in the future.

As with any method, Bayesian model averaging does have some drawbacks. First of all, in order to obtain useful results from the method one must start with a good "full" model by including all variables which should be adjusted for. However, there will always be some degree of uncertainty in deciding what to include, and in which form. Of course, this is an issue with traditional model selection methods as well.

There are many potential latent variables in PM mortality studies, such as the presence of flu epidemics and personal particulate matter exposure (indoor and outdoor), which we attempt to estimate with a baseline variable. However, by not accounting for the uncertainty in the baseline estimate, we do not take full advantage of the capabilities of the method. As mentioned in chapter 3, a technique to do this is currently being developed.

We also mentioned in chapters 3 and 4 that we have assumed a simple log-linear relationship between our response and most predictors. We could extend these analyses to include smooth functions of temperature and dewpoint temperature, as in Norris et al. ([22]), or any other variable, and account for the uncertainty in the degrees of freedom using BMA.

Another source of uncertainty in model building mentioned in chapter 1 is the inclusion/exclusion of unusual observations. Unusual or influential observations can be a concern in PM studies as holidays tend to have different PM levels than other days. For example, Norris et al. consider models with and without the observations on December 24 and 25, 1995 because of their unusually high PM measurements. This could easily be accommodated in BMA by adding an indicator variable for these potentially "unusual" or "influential" points.

5.2 Particulate matter and mortality

We have used Bayesian model averaging to analyze data from two cities with vastly different weather patterns and particulate matter composition, Phoenix, Arizona and Seattle, Washington, to attempt to estimate the association between non-accidental elderly mortality and particulate matter concentration measurements. We considered three different regions of mortality for our response variable in Phoenix. For two of these responses, U2.5MORT and PHXMORT, we found some evidence of an association between non-accidental elderly mortality and coarse, but not necessarily fine, particulate matter. At least 80% of the top models for each analysis included some coarse PM variable and 80% of the relative risk probability intervals did not include one. However, some of the top models do not include any PM variables. Also, the point estimates for the relative risk are 1.02, suggesting a weak effect, if any.

Because of the amount of missing $PM_{2.5}$ data for the Seattle area, we considered several different BMA processes. The evidence for any association between non-accidental elderly mortality and coarse or fine PM was underwhelming, even in analyses using most of the data. Relative risk point estimates were not larger than 1.01 and no variable seemed to be clearly related to the response, besides the estimate of long-term variation. These results are consistent with another study which examined PM and mortality in Seattle. However, other studies have found an association between PM and asthma-related morbidity in Seattle.

Several problems and possible improvements with the models considered have been presented in chapters 3, 4 and 5. One other issue with this study, as with most particulate matter studies, is that we looked at acute, rather than chronic, health effects of PM. Although the latter may be a more important response to explore, it would be very difficult to do without long-term data and personal indoor/outdoor particulate matter exposure levels.

BIBLIOGRAPHY

[1] United States Environmental Protection Agency. National Ambient Air Quality Standards for particulate matter; final rule. *Federal Register*, 40 CFR Part 50, 1997.

[2] V.H. Borja-Aburto, D.P. Loomis, S.I Bangdiwala, C.M. Shy, and R.A Rascon-Pacheco. Ozone, suspended particulates, and daily mortality in Mexico City. *American Journal of Epidemiology*, 145(3):258–268, 1997.

[3] M. Clyde. Model uncertainty and health effects studies for particulate matter. *Environmetrics*, in press, 2000.

[4] M. Clyde and H. DeSimone-Sasinowksa. Accounting for model uncertainty in Poisson regression models: Does particulate matter particularly matter? *Institute of Statistics and Decision Sciences, Duke University*, Discussion Paper 97-06, 1998.

[5] M. Clyde, P. Guttorp, and E. Sullivan. Effects of ambient fine and coarse particles on mortality in Phoenix, Arizona. *The National Research Center for Statistics and the Environment*, Technical Report Series 40, 2000.

[6] National Research Council. Airborne particles. *National Academy of Sciences*, Washington, DC, 1977.

[7] National Research Council. Research priorities for airborne particulate matter. *National Academy Press*, Washington, DC, 1998.

[8] L.H. Cox. Statistical issues in the study of air pollution involving airborne particulate matter. *The National Research Center for Statistics and the Environment*, Technical Report Series 41, 2000.

[9] P.F. Craigmile. *Wavelet Based Estimation for Trend Contaminated Long Memory Processes*. PhD thesis, University of Washington, Department of Statistics, 2000. To appear.

[10] D.W. Dockery, C.A. Pope, X. Xu, J.D. Spengler, J.H. Ware, M.E. Fay, B.G. Ferris Jr., and F.E. Speizer. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 329:1753–1759, 1993.

[11] D. Draper. Assessment and propagation of model uncertainty (with discussion). *Journal of the Royal Statistical Society, Series B.*, 56:45–98, 1995.

[12] J.F. Gamble and R.J. Lewis. Health and respirable particulate (PM10) air pollution: a causal or statistical association? *Environmental Health Perspectives*, 104(8):838–850, 1996.

[13] J.S. Hodges. Uncertainty, policy analysis, and statistics. *Statistical Science*, 2:259–291, 1987.

[14] G. Hoek, J.D. Schwartz, B. Groot, and P. Eilers. Effects of ambient particulate matter and ozone on daily mortality in Rotterdam, The Netherlands. *Archives of Environmental Health*, 52(6):455–463, 1997.

[15] J.A. Hoeting, D. Madigan, A.E. Raftery, and C.T. Volinsky. Bayesian model averaging: a tutorial. *Statistical Science*, 14(4):382–417, 1999.

[16] C.A. Pope III, R.W. Hill, and G.M. Villegas. Particulate air pollution and daily mortality on Utah's Wasatch Front. *Environmental Health Perspectives*, 107(7):567–573, 1999.

[17] R.E. Kass and A.E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995.

[18] Y. Kim, D. Spitner, Z. Zhang, R.L. Smith, and M. Fuentes. Accounting for multiple pollutants in pollution-mortality studies. *ASA Proceedings of the Biometrics Section*, 1999.

[19] F.W. Lipfert and R.E. Wyzga. Air pollution and mortality: issues and uncertainties. *Journal of the Air and Waste Management Association*, 45(12):949–966, 1995.

[20] F.W. Lipfert and R.E. Wyzga. Air pollution and mortality: the implications of uncertainties in regression modeling and exposure management. *Journal of the Air and Waste Management Association*, 47(4):517–523, 1997.

[21] P. McCullagh and J.A. Nelder. *Generalized Linear Models.* Chapman and Hall, 1989.

[22] G. Norris, S.N. YoungPong, J.Q. Koenig, T.V. Larson, L. Sheppard, and J.W. Stout. An association between fine particles and asthma emergency department visits for children in Seattle. *Environmental Health Perspectives*, 107(6):489–493, 1999.

[23] B. Ostro. Fine particulate air pollution and mortality in two Southern California counties. *Environmental Research*, 70:98–102, 1995.

[24] A. Peters, J. Skorkovsky, F. Kotesovec, J. Brynda, C. Spix, H.E. Wichmann, and J. Heinrich. Associations between mortality and air pollution in Central Europe. *Environmental Health Perspectives*, 108(4):283–287, 2000.

[25] G.J. Prescott, G.R. Cohen, R.A. Elton, F.G. Fowkes, and R.M. Aguis. Urban air pollution and cardiopulmonary ill health: a 14.5 year time series. *Occupational and Environmental Medicine*, 55(10):697–704, 1998.

[26] A.E. Raftery. Bayesian model selection in social research. *Sociological Methodology*, (25):111–163, 1995.

[27] A.E. Raftery. Approximate Bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika*, 83:251–266, 1996.

[28] J. M. Samet, S.L. Zeger, F. Dominici, F. Curriero, I. Coursac, D.W. Dockery, J. Schwartz, and A. Zanobetti. The national morbidity, mortality, and air pollution study; part II: Morbidity and mortality from air pollution in the United States. *Health Effects Institute Research Report*, 94, Part II, 2000.

[29] G. Schwartz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.

[30] J. Schwartz. Particulate air pollution and daily mortality in Detroit. *Environmental Research*, 102(2):186–189, 1991.

[31] J. Schwartz. Air pollution and hospital admissions for the elderly in Birmingham, Alabama. *American Journal of Epidemiology*, 139(6):589–598, 1994.

[32] J. Schwartz and A. Marcus. Mortality and air pollution in London: a time series analysis. *American Journal of Epidemiology*, 131(1):185–194, 1990.

[33] J. Schwartz, D. Slater, T.V. Larson, W.E. Pierson, and J.Q. Koenig. Particulate air pollution and hospital emergency room visits for asthma in Seattle. *American Review of Respiratory Diseases*, 147:826–831, 1993.

[34] L. Sheppard, D. Levy, G. Norris, T.V. Larson, and J.Q. Koenig. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987-1994. *Epidemiology*, 10(1):23–30, 1999.

[35] R.W. Simpson, G. Williams, A. Petroeschevsky, G. Morgan, and S. Rutherford. Associations between outdoor air-pollution and daily mortality in Brisbane, Australia. *Archives of Environmental Health*, 52(6):442–454, 1997.

[36] R.L. Smith, J.M. Davis, and P. Speckman. Assessing the human health risk of atmospheric particles. *Novartis Foundation Symposium*, 220:59–72, 1999.

[37] J. Wordley, S. Walters, and J.G. Ayres. Short term variations in hospital admissions and mortality and particulate air pollution. *Occupational and Environmental Medicine*, 54(2):108–116, 1997.

Appendix A

PARAMETRIC VS. NON-PARAMETRIC CALCULATION OF EPA STANDARDS

According to the NAAQS document ([1]), the yearly percentiles for PM variables are calculated nonparametrically by the formula

$$P_{q,y} = X_{[i+1]},$$

where $P_{q,y}$ is the qth percentile for year y, $X_{[i+1]}$ is the $(i+1)$th number in the ordered series of measurements, and i is the integer part of the product of q and the number of measurements for that year. One could also estimate the qth percentile for the particulate matter measurements using a parametric method. First, if our variable of interest does not seem to be normally distributed, find a suitable transformation to normality. For example, assume $\log(x)$ is normally distributed with mean μ and standard deviation σ. Then we can estimate the 98th or 99th percentile for x by calculating

$$y_{0.98} = \mu + z_{0.98} * \sigma$$

or

$$y_{0.99} = \mu + z_{0.99} * \sigma$$

where $z_{0.98}$ is the 98th percentile of the standard normal distribution ($z_{0.98} = 2.05$, $z_{0.99} = 2.33$). Then, by raising e to the powers $y_{0.98}$ and $y_{0.99}$, we have parametric estimates of the 98th and 99th percentiles of x.

We apply the parametric and nonparametric methods of percentile calculations in this subsection to provide a comparison between the two. We use PM_{10} and $PM_{2.5}$ data from Seattle, Washington which is described in detail in chapter 4. Using visual comparisons of q-q normal plots and histograms, we decided to calculate the parametric percentile estimates based on both a square-root transformation and a log transformation for PM_{10} and $PM_{2.5}$.

Tables A.1 and A.2 contain the resulting estimates from the two methods for each transformation. Using a square-root transformation results in parametric percentile estimates similar to those from the nonparametric method. The numbers are fairly similar for $PM_{2.5}$. The 99th percentile for PM_{10} for years 1994 and 1995 are much lower using the parametric method. In Table A.2 we see that using a log transformation results in numbers that are very different between parametric and nonparametric methods. The estimates for $PM_{2.5}$ are quite a bit higher using the parametric method, and those for PM_{10} are astronomical. It should be noted that visual assessment deemed the square-root transformation to be more appropriate for PM_{10} and the log transformation to be more appropriate for $PM_{2.5}$.

The results of these calculations suggest that a nonparametric approach may be more suitable than a parametric method for calculating percentiles for EPA particulate matter standards. First of all, the choice of a transformation is subjective and would be difficult to govern by the EPA. Secondly, the "best" transformation to normality seems to be governed by each set of measurements and estimates can be wildly different depending on the choice.

Table A.1: Yearly percentile estimates for PM_{10} and $PM_{2.5}$ measurements from the Duwamish site using parametric and nonparametric methods. The parametric estimates are based on a square-root transformation of the PM variables. All measurements are in $\mu g/m^3$. [a]Averages and percentiles for 1990 are based on measurements for January 16, 1990 through December 31, 1990.

Year	99th percentile of PM_{10}		98th percentile of $PM_{2.5}$	
	Nonparametric	Parametric	Nonparametric	Parametric
1990[a]	88.0	92.3	51.6	52.6
1991	96.4	89.3	40.6	45.2
1992	101.4	94.4	50.6	52.7
1993	98.4	92.4	43.7	46.9
1994	100.0	80.8	37.0	35.4
1995	97.4	78.7	32.6	34.4

Table A.2: Yearly percentile estimates for PM_{10} and $PM_{2.5}$ measurements from the Duwamish site using parametric and nonparametric methods. The parametric estimates are based on a log transformation of the PM variables. All measurements are in $\mu g/m^3$. [a]Averages and percentiles for 1990 are based on measurements for January 16, 1990 through December 31, 1990.

Year	99th percentile of PM_{10}		98th percentile of $PM_{2.5}$	
	Nonparametric	Parametric	Nonparametric	Parametric
1990[a]	88.0	193.1	51.6	70.6
1991	96.4	157.1	40.6	54.0
1992	101.4	184.0	50.6	66.7
1993	98.4	192.2	43.7	54.6
1994	100.0	166.9	37.0	40.4
1995	97.4	145.8	32.6	39.2