COMBINING EVIDENCE ON AIR POLLUTION AND DAILY MORTALITY FROM THE TWENTY LARGEST US CITIES: A HIERARCHICAL MODELING STRATEGY

Francesca Dominici, Jonathan M. Samet and Scott L. Zeger

Francesca Dominici, Department of Biostatistics, Jonathan M. Samet, Department of Epidemiology, and Scott L. Zeger, Department of Biostatistics, all at the Johns Hopkins University School of Hygiene and Public Health.

Correspondence may be addressed to Dr. Francesca Dominici, Department of Biostatistics, School of Hygiene and Public Health, 615 N. Wolfe Street, The Johns Hopkins University, Baltimore, MD 21205-3179, USA. ph: 410-614-5107, fax: 410-955-0958, e-mail: fdominic@jhsph.edu.

Acknowledgments: Research described in this article was conducted under contract to the Health Effects Institute (HEI), an organization jointly funded by the Environmental Protection Institute (EPA R824835) and automotive manufacturers. The contents of this article do not necessarily reflect the views and policies of HEI, nor do they necessarily reflect the views and policies of EPA, or motor vehicles or engine manufacturers. The authors are grateful to Dr. Giovanni Parmigiani and Dr. Curriero for comments on an earlier draft, and to Ivan Coursac and Jing Xu for data base development.

Abstract

Reports over the last decade of association between levels of particles in outdoor air and daily mortality counts have raised concern that air pollution shortens life, even at concentrations within current regulatory limits. Criticisms of these reports have focused on the statistical techniques used to estimate the pollution/mortality relationship and the inconsistency in findings among cities. We have developed analytic methods that address these concerns and combine evidence from multiple locations in order to gain a unified analysis of the data.

This paper presents log-linear regression analyses of daily time-series data from the largest 20 U.S. cities and introduces hierarchical regression models for combining estimates of the pollution-mortality relationship across cities. We illustrate this method by focusing on mortality effects of PM_{10} (particulate matter less than 10 microns in aerodynamic diameter), and performing univariate and bivariate analyses with PM_{10} and O_3 (ozone). In the first stage of the hierarchical model, we estimate the relative mortality rate associated with PM_{10} for each of the 20 cities using semiparametric log-linear models. The second stage of the model describes between-city variation in the true relative rates as a function of selected city-specific covariates. We also fit two variations of a spatial model with the goal of exploring the spatial correlation of the pollutant-specific coefficients among cities. Finally, to explore the results of considering the two pollutants jointly, we fit and compared univariate and bivariate models. All posterior distributions from Stage II are estimated using Markov chain Monte Carlo techniques.

In univariate analyses using concurrent day pollution values to predict mortality, we find that a $10\ \mu g/m^3$ increase of PM_{10} on average in the U.S. is associated with a 0.48% increase in mortality (95% interval $0.05, 0.92$). With adjustment for O_3, the PM_{10} coefficient is slightly

higher.

Results appear to be largely insensitive to the specific choice of vague but proper prior distribution. The models and estimation methods are general and can be used for any number of locations and pollutant measurements and have potential application to other environmental agents.

Key Words: Air Pollution, Longitudinal Data, Hierarchical Models, Markov Chain Monte Carlo, Log-Linear Regression, Mortality, Relative Rate.

1 Introduction

In spite of improvements in measured air quality indicators in many developed countries, the health effects of particulate air pollution remain a regulatory and public health concern. This continued interest is motivated largely by recent epidemiologic studies that have examined both acute and longer-term effects of exposure to particulate air pollution in different cities in the United States and elsewhere in the world (Dockery and Pope, 1994; Schwartz, 1995; American Thoracic Society, 1996a; American Thoracic Society, 1996b; Korrick et al., 1998). Many of these studies have shown a positive association between measures of particulate air pollution – primarily total suspended particles (TSP) or particulate matter less than 10μ in aerodynamic diameter (PM_{10}) – and daily mortality and morbidity rates. Their findings suggest that daily rates of morbidity and mortality from respiratory and cardiovascular diseases increase with levels of particulate air pollution below the current National Ambient Air Quality Standard (NAAQS) for particulate matter in the United States. Critics of these studies have questioned the validity of the data sets used and the statistical techniques applied to them; the critics have noted inconsistencies in findings among studies and even in independent re-analyses of data from the same city (Lipfert and Wyzga, 1993; Li and Roth, 1995). The biologic plausibility of the associations between particulate air pollution and illness and mortality rates has also been questioned (Vedal, 1996).

These controversial associations have been found using Poisson time series regression models fit to the data using the generalized estimating equations (Liang and Zeger, 1986) or generalized additive models (Hastie and Tibshirani, 1990). Following Bradford Hill's criterion of temporality, they have measured the acute health effects, focusing on the shorter-term variations in pollution and mortality by regressing mortality on pollution over the preceding few days. Model approaches

have been questioned (Smith et al., 1997; Clyde, 1998), although analyses of data from Philadelphia (Samet et al., 1997; Kelsall et al., 1997) show that the particle-mortality association is reasonably robust to the particular choice of analytic methods from among reasonable alternatives. Past studies have not used a set of communities; most have used data from single locations selected largely on the basis of the data availability on pollution levels. Thus, the extent to which findings from single cities can be generalized is uncertain and consequently we analyzed data for the 20 largest US locations the population living within the limits of the counties making up the cities. These locations were selected to illustrate the methodology and our findings cannot be generalized to all US with certainty. However, to better represent the nation, in a future application of our methods to the largest 90 cities. Statistical power of analyses within a single city may be limited by the amount of data for any location. Consequently, in comparison to analyses of data from a single site, pooled analyses can be more informative about whether an association exists, controlling for possible counfounders. In addition, a pooled analysis can produce estimates of the parameters at a specific site, which borrow strength from all other locations (DuMouchel and Harris, 1983; DuMouchel, 1990; Breslow and Clayton, 1993).

One additional limitation of epidemiologic studies of the environment and disease risk is the measurement error inherent in many exposure variables. When the target is estimation of the health effects of personal exposure to a pollutant, error is well recognized to be a potential source of bias (Lioy et al., 1990; Mage and Buckley, 1995; Wallace, 1996; Ozkaynak et al., 1996; Janssen et al., 1997; Janssen et al., 1998). The degree of bias depends on the correlation of the personal and ambient pollutant levels. Dominici et al. (2000) have investigated the consequences of exposure measurement error by developing a statistical model that estimates the association

between personal exposure and mortality concentrations, and evaluates the bias likely to occur in the air pollution/mortality relationships from using ambient concentration as a surrogate for personal exposure. Taking into account the heterogeneity across locations in the personal-ambient exposure relationship, we have quantified the degree to which the exposure measurement error biases the results toward the null hypothesis of no effect, and estimated the loss of precision in the estimated health effects due to indirectly estimating personal exposures from ambient measurements. Our approach is an example of regression calibration which is widely used for handling measurement error in non-linear models (Carroll et al., 1995). See also Zidek et al. (1996), Zidek et al. (1998), Fung and Krewski (1999) and Zeger et al. (2000) for measurement error methods in Poisson regression.

The main objective of this paper is to develop a statistical approach that combines information about air pollution/mortality relationships across multiple cities. We illustrated this method with the following two-stage analysis of data from the largest 20 U.S. cities:

1. Given a time series of daily mortality counts in each of three age groups, we used generalized additive models to estimate the relative change in the rate of mortality associated with changes in the air pollution variables (relative rate), controlling for age-specific longer-term trends, weather, and other potential confounding factors, separately for each city;

2. We then combined the pollution-mortality relative rates across the 20 cities using a Bayesian hierarchical model (Lindley and Smith, 1972; Morris and Normand, 1992) to obtain an overall estimate, and to explore whether some of the geographic variation can be explained by site-specific explanatory variables.

This paper considers two hierarchical regression models – with and without modeling possible

spatial correlations – which we referred to as the "baseline" and the "spatial" models.

In both models, we assumed that the vector of the estimated regression coefficients obtained from the first-stage analysis, conditional on the vector of the true relative rates, has a multivariate normal distribution with mean equal to the "true" coefficient and covariance matrix equal to the sample covariance matrix of the estimates. At the second stage of the baseline model, we assume that the city-specific coefficients are independent. In contrast, at the second stage of the spatial model, we allowed for a correlation between all pairs of pollutant and city-specific coefficients; these correlations were assumed to decay toward zero as the distance between the cities increases. Two distance measures were explored.

Section 2, describes the database of air pollution, mortality, and meteorological data from 1987 to 1994 for the 20 U.S. cities in this analysis. In section 3, we fit the log-linear generalized additive models to produce relative-rate estimates for each location. The semiparametric regression is conducted three times for each pollutant: using the concurrent day's (lag 0) pollution values, using the previous day's (lag 1) pollution levels, and using pollution levels from two days before (lag 2).

Section 4, presents the baseline and the spatial hierarchical regression models for combining the estimated regression coefficients and discuss Markov chain Monte Carlo methods for model fitting. In particular, we used the Gibbs sampler (Geman and Geman, 1993; Gelfand and Smith, 1990) for estimating parameters of the baseline model and a Gibbs sampler with a Metropolis step (Hastings, 1970; Tierney, 1994) for estimating parameters of the spatial model. Section 5, summarizes the results, compares between the posterior inferences under the two models, and assesses the sensitivity of the results to the choice of lag structure and prior distributions.

2 Description of the Databases

The analysis database included mortality, weather, and air pollution data for the 20 largest metropolitan areas in the U.S. for the 7-year period 1987-1994 (Figure 1 and Table 1).

Table 1

In several locations, we had an high percentage of days with missing values for PM_{10} because it is generally measured every six days. The cause-specific mortality data, aggregated at the level of county, were obtained from the National Center for Health Statistics. We focused on daily death counts for each site, excluding non-residents who died in the study site and accidental deaths. Because mortality information was available for counties but not smaller geographic units to protect confidentiality, all predictor variables were aggregated to the county level.

Hourly temperature and dew point data for each site were obtained from the EarthInfo CD database. After extensive preliminary analyses that considered various daily summaries of temperature and dew point as predictors, such as daily average, maximum, and eight-hour maximum, we used the 24-hour mean for each day. If a city has more than one weather station, we took the average of the measurements from all available stations. The PM_{10} and O_3 data were also averaged over all monitors in a county. To protect against outliers, a 10% trimmed mean is used to average across monitors, after correction for yearly averages for each monitor. This yearly correction is appropriate since long-term trends in mortality are also adjusted in the log-linear regressions. See Kelsall et al. (1997) for further details. Aggregation strategies based on Bayesian and classical geostatistical models as suggested by Handcock and Stein, (1993), Cressie, (1993), Kaiser and Cressie, (1993) and Cressie at al. (1999) , and Bayesian models for spatial interpolation (Le et al., 1997; Gaudard et al., 1999) are desirable in many contexts because they provide estimates

of the error associated with exposure at any measured or unmeasured locations. However, they were not applicable to our data sets because of the limited number of monitoring stations that are available in the 20 counties.

3 City-Specific Analyses

In this section, we summarize the model used to estimate the air pollution/mortality relative rate separately for each location, accounting for age-specific longer-term trends, weather, and day of the week. The core analysis for each city is a log-linear generalized additive model that accounts for smooth fluctuations in mortality that potentially confound estimates of the pollution effect and/or introduce autocorrelation in mortality series.

This is a study of the acute health effects of air pollution on mortality. Hence, we modeled daily expected deaths as a function of the pollution levels on the same or immediately preceding days not of the average exposure for the preceding month, season, or year as might be done in a study of chronic effects. We built models which include smooth functions of time as predictors as well as the pollution measures to avoid confounding by influenza epidemics which are seasonal and by other longer-term factors.

To specify our approach more completely, let y_{at}^c be the observed mortality for each age group $a = (\leq 65, 65-75, \geq 75$ years) on day t at location c, and $\boldsymbol{x}_{at}^c$ be a $p \times 1$ vector of air pollution variables. Let $\mu_{at}^c = E(y_{at}^c)$ be the expected number of deaths and $v_{at}^c = \text{var}(y_{at}^c)$. We used a log-linear model $\log \mu_{at}^c = \boldsymbol{x}_{at}^{c\prime}\boldsymbol{\beta}^c$ for each city c, allowing the mortality counts to have variances v_{at}^c that may exceed their means (i.e., be overdispersed) with the overdispersion parameter ϕ^c also varying by location so that $v_{at}^c = \phi^c \mu_{at}^c$.

To protect the pollution relative rates $\boldsymbol{\beta}^c$ from confounding by longer-term trends due, for example, to changes in health status, changes in the sizes and characteristics of populations, seasonality, and influenza epidemics, and to account for any additional temporal correlation in the count time-series, we estimated the pollution effect using only shorter-term variations in mortality and air pollution. To do so, we partial out the smooth fluctuations in the mortality over time by including arbitrary smooth functions of calendar time $S^c(\text{time}, \lambda)$ for each city. Here, λ is a smoothness parameter which we pre-specified, based upon prior epidemiologic knowledge of the time scale of the major possible counfounders, to have seven degrees of freedom per year of data so that little information from time-scales longer than approximately two months is included when estimating $\boldsymbol{\beta}^c$. This choice largely eliminates expected confounding from seasonal influenza epidemics and from longer-term trends due to changing medical practice and health behaviors, while retaining as much unconfounded information as possible. We also controlled for age-specific longer-term and seasonal variations in mortality, adding a separate smooth function of time with eight degrees of freedom for each age-group.

To control for weather, we also fit smooth functions of the same day temperature (temp_0), average temperature for the three previous days (temp_{1-3}), each with six degrees of freedom, and the analogous functions for dew point ($\text{dew}_0, \text{dew}_{1-3}$), each with three degrees of freedom. In the US cities, mortality decreases smoothly with increases temperature until reaching a relative minimun and then increases quite sharply at higher temperature. Six degrees of freedom were chosen to capture the highly non linear bend near the relative minimum as best as possible. Since there are missing values of some predictor variables on some days, we restricted analyses to days with no missing values across the full set of predictors.

In summary, we fitted the following log-linear generalized additive model (Hastie and Tibshirani, 1990) to obtain the estimated pollution log-relative rate $\hat{\boldsymbol{\beta}}^c$ and the sample covariance matrix V^c at each location:

$$\begin{aligned} \log \mu_{at}^c &= \boldsymbol{x}_{at}^{c'}\boldsymbol{\beta}^c + \gamma^c \text{DOW} + S_1^c(\text{time}, 7/\text{year}) + \\ &+ S_2^c(\text{temp}_0, 6) + S_3^c(\text{temp}_{1-3}, 6) + S_4^c(\text{dew}_0, 3) + S_5^c(\text{dew}_{1-3}, 3) \\ &+ \text{intercept for age group } a \\ &+ \text{separate smooth functions of time (8 df) for age group } a. \end{aligned} \tag{1}$$

where DOW are indicator variables for day of week. Samet et al. (1995,1997) and Kelsall et al. (1997) give additional details about choices of functions used to control for longer-term trends and weather. Alternative modeling approaches that consider different lag structures of the pollutants and of the meteorological variables have been proposed (Davis et al., 1996; Smith et al., 1997; Smith et al., 1998). More general approaches that consider non-linear modeling of the pollutant variables have been discussed by Smith et al. (1997) .

Because the functions $S^c(x, \lambda)$'s are smoothing splines with fixed λ, the semiparametric model described above has a finite-dimensional representation. Hence, the analytic challenge was to make inferences about the joint distribution of the $\boldsymbol{\beta}^c$s in the presence of finite-dimensional nuisance parameters, which we will refer to as $\boldsymbol{\eta}^c$.

We separately estimated three semi-parametric regressions for each pollutant with the concurrent day (lag 0), prior day (lag 1), and two days prior (lag 2) pollution predicting mortality. The estimates of the coefficients and their 95% confidence intervals for PM_{10} alone and for PM_{10} adjusted by O_3 are shown in Figures 2 and 3. Cities are presented in decreasing order by the size of their populations. The pictures show substantial between-location variability in the estimated relative rates, suggesting that combining evidence across cities would be a natural

approach to explore possible sources of heterogeneity, and to obtain an overall summary of the degree of association between pollution and mortality. To add flexibility in modeling the lagged relationship of air pollution with mortality, we could have used distributed lag models instead of treating the lags separately. While desirable, this is not easily implemented because many cities have PM_{10} data available only every sixth day.

To test if the log-linear generalized additive model (1) has taken appropriate account of the time dependence of the outcome, we calculate, for each city, the autocorrelation function of the standardized residuals. Figure 4 displays the 20 autocorrelation functions; they are centered near zero, ranging between -0.05 and 0.05, confirming that the filtering has removed the serial dependence.

We also examined the sensitivity of the pollution relative rates to the degrees of freedom used in the smooth functions of time, weather and seasonality by halving and doubling each of them. The relative rates changed very little as these parameters are varied over this four-fold range (data not shown).

4 Pooling Results Across Cities

In this section, we present hierarchical regression models designed to pool the city-specific pollution relative rates across cities to obtain summary value for the 20 largest U.S. cities. Hierarchical regression models provide a flexible approach to the analysis of multi-level data. In this context, the hierarchical approach provides a unified framework for making estimates of the city-specific pollution effects, the overall pollution effect, and of the within- and between- cities variation of the city-specific pollution effects.

Results of several applied analyses using hierarchical models have been published. Examples include models for the analysis of longitudinal data (Gilks et al., 1993), spatial data (Breslow and Clayton, 1993), and health care utilization data (Normand et al., 1997). Other modeling strategies for combining information in a Bayesian perspective are provided by Du Mouchel (1990), Skene and Wakefield (1990), Smith and Spiegelhalter (1995), and Silliman (1997) . Recently, spatio-temporal statistical models with applications to environmental epidemiology have been proposed by Wikle et al. (1997), and Wakefield and Morris (1998).

In section 4.1 we present an overview of our modeling strategy. In sections 4.2 and 4.3, we consider two hierarchical regression models with and without modeling of the possible spatial autocorrelation among the $\boldsymbol{\beta}^c$s which we refer to as the "baseline" and "spatial" models, respectively.

4.1 Modeling Approach

The modeling approach comprises two stages. At the first stage, we used the log-linear generalized additive model (1) described in section 3:

$$y_t^c \mid \boldsymbol{\beta}^c, \boldsymbol{\eta}^c \sim \text{Poisson}\left[\mu_t(\boldsymbol{\beta}^c, \boldsymbol{\eta}^c)\right]$$

where $y_t^c = \left(y_{\leq 65t}^c, y_{65-75t}^c, y_{\geq 75t}^c\right)$. The parameters of scientific interest are the mortality relative rates, $\boldsymbol{\beta}^c$s, which for the moment are assumed not to vary across the three age groups within a city. The vector $\boldsymbol{\eta}^c$ of the coefficients for all the adjustment variables, including the splines in the semi-parametric log-linear model, is a finite-dimensional nuisance parameter.

The second stage of the model describes variation among the $\boldsymbol{\beta}^c$s across cities. We regressed the true relative rates on city-specific covariates, $\boldsymbol{z}^c$, to obtain an over-all estimate, and to

explore the extent to which the site-specific explanatory variables explain geographic variation in the relative risks. In epidemiologic terms, the covariates in stage two are possible effect-modifiers. More specifically, we assumed:

$$\boldsymbol{\beta}^c \mid \boldsymbol{\alpha}, \Sigma \sim N_p(\boldsymbol{z}^c\boldsymbol{\alpha}, \Sigma)$$

where p is the number of pollutant variables that enter simultaneously in model (1). Here the parameters of scientific interest are the vector of the regression coefficients, $\boldsymbol{\alpha}$, and the overall covariance matrix, Σ. Unlike the overall air pollution effect $\boldsymbol{\alpha}$, we are not interested in estimating overall non-linear adjustments for trend and weather, therefore we assume that the nuisance parameters $\boldsymbol{\eta}^c$ are independent across cities. Our goal is to make inferences about the parameters of interest, the $\boldsymbol{\beta}^c$s, $\boldsymbol{\alpha}$, and Σ – in the presence of nuisance parameters $\boldsymbol{\eta}^c$s. To estimate an exact Bayesian solution to this pooling problem, we could analyze the joint posterior distributions of the parameters of interest, as well as of the nuisance parameters, and then integrate over the $\boldsymbol{\eta}^c$ dimension to obtain the marginal posterior distributions of the $\boldsymbol{\beta}^c$s. While possible, the computations become extremely laborious and are not practical for either this analysis or a planned model with ninety or more cities.

Given the large sample size at each city, (T ranges from 550 to 2550 days), accurate approximations to the posterior distribution can be obtained using the normal approximation of the likelihood (Le Cam and Yang, 1990). If the likelihood function of $\boldsymbol{\beta}^c$ and $\boldsymbol{\eta}^c$ is approximated by a multivariate normal distribution with mean equal to the maximum likelihood estimates $\hat{\boldsymbol{\beta}}^c$ and $\hat{\boldsymbol{\eta}}^c$ and covariance matrices V_β and V_η, then by definition the marginal likelihood of $\boldsymbol{\beta}^c$ has multivariate normal distribution with mean $\hat{\boldsymbol{\beta}}^c$ and covariance matrix V_β. We then replaced the first stage of the model with a normal distribution with mean and variance equal to the maximum

likelihood estimates of the parameter. Recently it has been shown that, the strategy based on the normal approximation of the likelihood gives an alternative two-stage model that well approximates the original model and leads to more efficient simulation from the posterior (Daniels and Kass, 1998).

To check if inferences based on the normal approximation of the likelihood are proper, we compared our approach with the implementation of the full MCMC approach for a few cities with sample sizes ranging from 2000 in Pittsburgh to 545 in Riverside. Figure 5 shows the histogram of samples for Riverside from $p(\boldsymbol{\beta}^c \mid \text{data})$ – obtained implementing a Gibbs sampler that simulates from $p(\boldsymbol{\beta}^c \mid \boldsymbol{\eta}^c, \text{data})$ and $p(\boldsymbol{\eta}^c \mid \boldsymbol{\beta}^c, \text{data})$ and approximate $p(\boldsymbol{\beta}^c \mid \text{data}) = \int p(\boldsymbol{\beta}^c, \boldsymbol{\eta}^c \mid \text{data}) d\boldsymbol{\eta}^c$ – with samples from $N(\hat{\boldsymbol{\beta}}^c, V^c)$ (solid line). The two distributions are very similar.

4.2 Baseline model

Let $\boldsymbol{\beta}^c = [\beta^c_{PM10}, \beta^c_{O3}]'$ be the log relative rate associated with PM_{10} and O_3 at city c. We considered the following hierarchical model:

$$\begin{aligned} \hat{\boldsymbol{\beta}}^c \mid \boldsymbol{\beta}^c &\sim N_2(\boldsymbol{\beta}^c, V^c) \\ \beta^c_{PM10} &= \boldsymbol{z}^{c'}_{PM10}\boldsymbol{\alpha}_{PM10} + \epsilon^c_{PM10} \\ \beta^c_{O3} &= \boldsymbol{z}^{c'}_{O_3}\boldsymbol{\alpha}_{O_3} + \epsilon^c_{O3} \\ \boldsymbol{\epsilon}^c \mid \Sigma &\sim N_2(0, \Sigma) \end{aligned} \tag{2}$$

where $\boldsymbol{z}^c_{PM10} = [1, P^c_{poverty}, P^c_{>65}, \bar{X}^c_{PM10}]'$, $\boldsymbol{z}^c_{03} = [1, P^c_{poverty}, P^c_{>65}, \bar{X}^c_{03}]'$, $\boldsymbol{\alpha}_{PM10}$ and $\boldsymbol{\alpha}_{O3}$ are 4×1 vectors, and finally $\boldsymbol{\epsilon}^c = [\epsilon^c_{PM10}, \epsilon^c_{O3}]'$, $c = 1, \ldots, 20$. This model specification allowed dependence between the relative rates associated with PM_{10} and O_3, but implied independence between the relative rates of city c and c'.

Under this model, the true PM_{10} and O_3 log-relative rates in city c were regressed on predictor variables including the percentage of people in poverty, $(P^c_{poverty})$ and the percentage of people older than 65 years, $(P^c_{>65})$, and on the average of the daily values of PM_{10} and O_3 over the period 1987-1994 in location c, $\bar{X}^c_{PM10}, \bar{X}^c_{O3}$. If we centered the predictors about their means, the intercepts $\alpha_{0,PM10}, \alpha_{0,O_3}$ can be interpreted as overall effects for a city with mean predictors. A simple pooled estimate of the pollution effect is obtained by setting all covariates to zero. To compare the consequences of considering two pollutants independently and jointly in the model, we fit a baseline-univariate model, — i.e., Σ assumed diagonal — and a baseline-bivariate model, — i.e., Σ assumed to have non-zero off-diagonal elements.

Inference on the parameters $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_{PM10}, \boldsymbol{\alpha}_{O_3}]'$ and Σ represents a synthesis of the information from the 20 cities; for example the parameters $\alpha_{0j},\ [\Sigma]_{jj}\ j = PM_{10}, O_3$, determine the overall level and the variability of the relative change in the rate of mortality associated with changes in the j-th pollutant level on average over all the cities.

The Bayesian formulation was completed by specifying dispersed but proper baseline prior distributions, and then supplementing the baseline analysis with additional sensitivity analysis. A priori, we assumed that the joint prior is the product of the marginals for $\boldsymbol{\alpha}$ and Σ. The following baseline prior specifications for the marginals are used:

$$\begin{aligned} \text{Overall log-relative rates } \boldsymbol{\alpha} &\sim N_{p(k+1)}(\boldsymbol{m}, V_\alpha) \\ \text{Overall covariance matrix } \Sigma &\sim IW_p(df, D) \end{aligned}$$

where $IW_p(df, D)$ denotes the inverse Wishart distribution with df degrees of freedom and scale matrix D, a $p \times p$ positive definite matrix, whose density is proportional to

$$\frac{D^{(df+p-1)/2}}{|\Sigma|^{(df+2p)/2}} \exp\left\{-\frac{1}{2} tr D\Sigma^{-1}\right\}.$$

where p denotes the number of pollutant variables entering simultaneously in the model and k the number of city specific covariates, respectively. We select $\boldsymbol{m}$ equal to a vector of zeros, V_α equal to a diagonal matrix, with diagonal elements equal to 100, $df = 3$ and D a diagonal matrix with diagonal elements equal to 3. In the univariate case we denote Σ by σ^2. These prior hyperparameters lend prior 95% support to the overall effect, the city-specific effects, and the correlation between the PM_{10} and the O_3 log-relative rates equal to $(-15, 15), (-4, 4)$, and $(-0.85, 0.85)$, respectively. This prior specification was selected because it did not impose too much shrinkage of the study-specific parameters toward their overall means, while at the same time specifying a reasonable range for the unknown parameters a priori. A sensitivity analysis is presented in section 4.

Given these prior assumptions, we can draw inferences on the unknown parameters using the posterior distribution

$$p(\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^{20}, \boldsymbol{\alpha}, \Sigma \mid \hat{\boldsymbol{\beta}}^1, \ldots, \hat{\boldsymbol{\beta}}^{20}, V^1, \ldots, V^{20}). \tag{3}$$

To do this, we implemented a Markov chain Monte Carlo algorithm with a block Gibbs Sampler (Gelfand and Smith, 1990) in which the unknowns are partitioned into the following groups: $\boldsymbol{\beta}^c$s, $\boldsymbol{\alpha}$, and Σ. Each group is sampled in turn, given all others. The full conditional distributions were available in closed form. Their derivation was routine (Bernardo and Smith, 1994) and not detailed here. Because of the normality assumptions at the first and second stage of the hierarchical model, computation of the posterior distributions of all the unknowns under an univariate model can be performed via direct simulation following the factorization above:

$$p(\beta^1, \ldots, \beta^{20}, \alpha, \sigma^2 \mid \text{data}) = p(\sigma^2 \mid \text{data}) p(\alpha \mid \sigma^2, \text{data}) \prod_c p(\beta^c \mid \alpha, \sigma^2, \text{data})$$

The first step, simulating σ^2, can be performed numerically (using the inverse cdf method, for example). The second and third steps can be done easily by sampling from normal distributions. This strategy can be conveniently implemented only for the univariate baseline model.

4.3 Spatial model

The assumption of independence of the city-specific coefficients made in the baseline model can be relaxed to a more general model in which the correlation between β^c and $\beta^{c'}$ decays as either a smooth or step function to zero as the distance between the two cities, c and c', increases. In this section, we consider a hierarchical model in which the inferences allow for the possible spatial correlation among the β^cs. We only considered univariate models given the small number of cities; extension to multivariate models is straightforward, but requires a larger data set.

At the second stage of the spatial model, we assumed that there is systematic variation in the air pollution/mortality relationship from pollutant to pollutant as specified in the baseline model (2). We expressed the degree of similarity of the relative rates in locations c and c' as function of an (arbitrary) distance between c and c', by assuming $\rho(c,c') = Corr(\beta^c, \beta^{c'}) = \exp(-\theta\, d(c,c'))$. We considered two distance measures, the Euclidean distance between the cities c and c' in the longitude and latitude coordinates, and a step function such that $d(c,c') = 1$ if location c and c' are within a common "region" and $d(c,c') = \infty$ if not. To make the results of these two models comparable we re-scaled the Euclidean distance such that it ranges between 0 and 4 with median equal to 0.64. The spatial model with $(1,\infty)$ distance can also be specified as a three stage hierarchical model where: the first stage is as the baseline model (2), the second stage describes the heterogeneity of the estimates across regions, and the third stage describes the

heterogeneity of the estimates across regions. For this regional model, we have clustered the 20 cities in the following three regions: north-east, south-east, and west coast. Thus, if we indicate by τ^2 the variability of the estimates across regions and by σ^2 the variability of the estimates within region, then the correlation of the log-relative rates for locations c and c' within a common region is $\tau^2/(\tau^2+\sigma^2)$. Alternate definitions of distance can be incorporated easily into the model as appropriate.

The spatial model specification is completed with the elicitation of the prior distribution. For α and σ^2 we choose the same prior specified in section 4. For the parameter θ under the spatial model with Euclidean distance, we choose a log-normal prior with mean 0.2 and standard deviation 0.5. Let $\tilde{d}$ the median of the distribution of all distances, this specification leads to a prior distribution of the correlation $\exp(-\theta \times \tilde{d})$ having mean 0.45 (95% interval $0.11, 0.74$). For the parameter τ^2 under the spatial model with step distance, we chose an inverse gamma prior $IG(A, B)$ with parameters $A = 5$ and $B = 8.5$. This specification leads to a prior distribution for τ having mean 1.35 (95% prior interval $0.9, 2.2$), and a prior distribution for the correlation $\tau^2/(\tau^2 + \sigma^2)$ having mean 0.45 (95% prior interval $0.13, 0.77$).

In the spatial model, the full conditional for β^cs, α, σ^2 are all available in closed form. In contrast, to sample from the full conditional distribution of θ, we used a Metropolis-Hastings algorithm with a Gamma proposal distribution having mean equal to the current value of θ and fixed variance. The spatial model with a step distance can be more efficiently sampled with a block Gibbs sampler because the full conditional distributions of all the unknown parameters are available in close form.

5 Results

We ran the Gibbs Sampler for 3000 iterations for both the baseline and the spatial models, ignoring the first 100. The autocorrelation, computed from a random sample of the $\alpha_{0,PM10}$, is negligible at lag 5 so we sampled every 5th observations for posterior estimation. The acceptance probabilities for the Metropolis Algorithm averaged between 0.3 and 0.5. Convergence diagnosis has been performed by implementing Raftery and Lewis (1992) methods in `CODA` (Best et al., 1995) which reported the minimum number of iterations, `Nmin`, needed to estimate the variable of interest with an accuracy of plus or minus 0.005 and with probability of attaining this degree of accuracy equal to 0.95. `Nmin` $\simeq 2000$ are proposed.

Figure 6 summarizes results of the pooled analyses under the univariate-baseline model. This figure displays the posterior distributions of city-specific regression coefficients β^c associated with changes in PM_{10} measurements for the 20 cities at the current day, one-day lag, and two-day lag. The marginal posterior distribution of the overall effect ($\alpha_{0,PM10}$) is displayed at far right. Cities are ordered by the decreasing size of their populations. At the current day, the highest relative rate for the PM_{10} variable occurs in New York with 1.05% increase in mortality (95% interval: $0.5, 1.6$) per $10\mu g/m^3$ increase in PM_{10}. Overall, we found that a $10\mu g/m^3$ increase of PM_{10} is associated with an estimated 0.48% increase in mortality (95% interval: $0.05, 0.92$).

Figure 7 summarizes results of the pooled analyses under the bivariate-baseline model. When PM_{10} and O_3 are combined in the same model, we estimated that 10 unit increments in PM_{10} adjusted by O_3 are associated with mortality increases of 0.52% (95% interval: $0.16, 0.85$).

The marginal posterior distribution of the overall regression effect combined and synthesized the information from the 20 locations. Figure 8 shows the marginal posterior distributions of the

overall pollution relative rates at the current day, one-day, and two-day lags obtained from the baseline-univariate,the baseline-bivariate, and the spatial models. At the top right are summarized the posterior probabilities that the overall effects are larger than zero for each lag-specification. In univariate and bivariate analyses, we found significant effects of PM_{10}.

Results of the adjusted analyses under the univariate-baseline model are shown in Table 2. Here we summarize the posterior means and the 95% posterior support intervals for the relationship between the mean of the city-specific coefficients and the percentage in poverty, the percentage of people older than 65, and the mean level of the pollutant. None of these variables are found to predict the PM_{10} relative rate.

An interaction of the pollution effects and age could be detected by the coefficient of the variable $P_{>65}$ in the second-stage regression model. A more direct approach was to estimate a separate pollution relative rate for each age stratum in the first-stage log-linear models and then to pool the trivariate vector $(\hat{\beta}_{<65},\ \hat{\beta}_{65-75}, \hat{\beta}_{>75})$ across cities. When we did so, the estimate of the overall effect of PM_{10} for the three age groups have posterior means 0.63 (95% interval $0.24, 1.05$), 0.26 (95% interval $-0.14, 0.67$), and 0.46 (95% interval $0.04, 0.83$). These results suggest that no a trend was evident in the pollution relative rates with age as is suggested by the second stage regression results in Table 2.

The variability of the regression coefficients, on average, over all the locations was captured by the matrix Σ. Marginal posterior means and 95% posterior support intervals are summarized in Table 4. A large diagonal element signified large variability over cities in the corresponding coefficient, while a large off-diagonal element signifies strong correlation between the PM_{10} and ozone coefficients. Table 4 shows the results. Under the baseline univariate model, the standard

deviation of the true coefficients across cities was estimated to be 0.76 (95% interval $0.41, 1.37$) which is about twice as large as the overall estimate of the pollution effect. Hence, in univariate analyses, the variability in PM_{10} coefficient is non-negligible. The posterior distribution of the off-diagonal elements of Σ indicates a negative mean correlation between the effects of the two pollutants, but with a large standard deviation.

From the posterior samples of θ in the spatial model, we could easily calculate the marginal posterior distributions of the correlation coefficient $\rho(c, c') = \exp(-\theta\ d(c, c'))$ for each distance $d(c, c')$. For the cities having median distance, the posterior mean correlation between β^c and $\beta^{c'}$ was 0.61 (95% interval $0.3, 0.8$). Consider the 25% and 75% quantiles of the distribution of all distances. Each of these quantiles has an associated correlation coefficient. The prior mean of these two correlation coefficients were 0.86 (95% interval $0.68, 0.93$) and 0.3 (95% interval $0.05, 0.58$), all larger than the corresponding prior means.

Under the regional model, with distance equal to a step function, the posterior mean of the within-region correlation of the city-specific relative rates $\tau^2/(tau^2 + \sigma^2)$, was 0.68 (95% interval $0.42, 0.86$). Results for the PM_{10} effects under the two spatial models were similar qualitatively. The posterior means and IQR for the regional effects β^{EAST}, β^{SOUTH} and β^{WEST} are 0.40, $(-0.22, 1.03)$, -0.06, $(-0.96, 0.93)$ and 0.69, $(0.07, 1.35)$ revealing that the adverse health effects of PM_{10} on mortality in the West US is larger than in the East and South US.

We have assessed the robustness of the results with respect to choices of the model (univariate, bivariate, spatial), of the lag-structure (lag 0, lag 1, lag 2) and of the prior distributions. Our sensitivity analysis compared 27 alternative scenarios (three for model-choice, three for lag-structures, and three for prior distributions). For these scenarios we compare the posterior probability that

the overall effect of the PM_{10} is larger than zero. The consequences of these choices are shown in Table 5. Significant effects of PM_{10} on total daily mortality are observed in all three models (weaker under a spatial model with current day pollution predicting mortality). When both pollutants are included in the model, adverse effects of PM_{10} became stronger. Spatial analyses attenuate the effects.

6 Discussion

We have developed a statistical model for obtaining a national estimate of the effect of urban air pollution on daily mortality using data for the 20 largest US cities. The raw data comprised publicly available listings of individual deaths by day and location, and hourly measurements of pollutants and weather variables. Substantial pre-processing of the nearly one gigabyte of information is necessary to create daily time series of mortality, pollutants, and weather for each of the 20 cities.

Because estimation of a national pollution-relative rate is the primary objective of this study, a two-stage approach is developed that allowed the modeling effort to focus on the combining of information across cities. In the first stage, a log-linear regression is used to estimate a pollution-relative rate for each city while controlling for the city-specific longer-term time trends and weather effects. Because we had no specific scientific interest in the time or weather effects, no effort is made to impose modeling assumptions to enable our borrow strength across cities when estimating the effects on mortality of these variables.

In the second stage, we regressed the true relative rates on city-specific covariates to obtain an overall estimate, and to estimate the variation among the coefficients across cities. We then

generated posterior estimates of the overall pollution effect and of the city-specific effects using Markov Chain Monte Carlo methods. Four models for combining relative rates of mortality for PM_{10} across cities are used. In the first, relative rates from different cities are treated as independent of one another. In the second, relative rates from different cities are treated as independent of one another, but are adjusted by O_3. In the third and fourth models the possibility of geographic correlation among the true coefficients is allowed. Results under the four models are similar, bivariate analyses gives slightly higher effects, and spatial analyses slightly attenuate the effects. Results under different models, lag specifications and prior are summarized in Figure 8 and Table 4. Note that the variance of the posterior distribution of the overall relative rate in the spatial models is somewhat sensitive to the prior specification for the between-region variance or equivalently within-region correlation since, with our 20 cities, we have only 3 regions and hence limited information. A similar analyses of the 90 larger cities will provide more precise information about variation across regions.

These analyses demonstrated that there was a consistent association of particulate air pollution PM_{10} with daily mortality across the 20 largest US cities leading to an overall effect, which was positive with high probability. Our overall estimate was that a $10\mu g/m^3$ increase in particulate level is associated with a roughly 0.48$\%$ increase in daily mortality on that day or the next day.

Another multi-city study air pollution and mortality is the multi-center European study, Air Pollution and Health: A European Approach (APHEA), (Katsoyanni et al., 1997; Toulomi et al., 1997). The cities are selected from across Europe, although not on a systematic basis. Data on particulate air pollution and daily mortality are analyzed from 12 cities from western and

central Europe according to a standardized protocol. Model estimates from the individual cities are pooled as the weighted means of the regression coefficients and heterogeneity among cities is explored using a random effects model. For particulate matter, the findings differed between the western and central Europe cities, with a five-old greater effect in the western cities (Katsoyanni et al., 1997). A similar approach is applied to the six selected cities with data available on O_3. A significant effect of O_3 is found, after controlling for levels of black smoke, and index of particulate matter (Toulomi et al., 1997).

While only a first step, the modeling described here establishes a basis for carrying out national surveillance for effects of air pollution and weather on public health. The analyses could be easily extended to studies of cause-specific mortality and other pollutants. Monitoring efforts using models like the one described here would be appropriate given the important public health questions that they can address and the considerable expense to government agencies for collecting the information that forms the basis for this work.

An alternative modeling strategy would have been to use one large Markov chain Monte Carlo method to simultaneously estimate the parameters in the log-linear models within each city, the overall estimate of the pollutant, and all of the nuisance parameters, borrowing strength across cities to obtain more precise estimates of the nuisance functions for each city. This type of approach would be necessary if there were limited information about the nuisance parameters within each city as, for example, in the Neyman and Scott problem (Neyman and Scott, 1960). As this is not the case in our investigation, we focused the modeling and computing effort on combining city-specific relative rate estimates to obtain a national average relative rate.

If the likelihood function for the pollution relative rate and the nuisance parameters is well

approximated by a Gaussian distribution, then our approach will give a close approximation to the posterior distribution from a Markov chain Monte Carlo that simulated both the parameters of interest and the nuisance parameters. We compared the marginal posterior of the $\boldsymbol{\beta}^c$ obtained using a full MCMC with our normal approximation for a few cities; they are indistinguishable.

The approach of taking a weighted average of the city-specific estimates to obtain an estimate of the overall effect, as for example suggested by DerSimonian and Laird, (1986), is a simplified version or approximation to the use of hierarchical models with a Gibbs Sampler. Under the weighted average approach for a random effect model, the weights of the city-specific estimates are modified to take into account of the variability between locations, say σ^2, and an estimate of this variance is included. Rather then including a single estimate of σ^2, the Bayesian method permits incorporating the whole posterior distribution of σ^2. In this way, all the information about the variability among studies is considered. In addition, the Bayesian method provide estimates of the posterior distribution of the city-specific relative rates, of the national estimate and it easily lend itself to generating ranking probabilities as, for example, $P(\text{overall log relative rate} \geq 0 \mid \text{data})$. In addition, the Gibbs Sampler is necessary for approximating the posterior distributions under the spatial model

These analyses alone cannot establish that increased levels of particulate air pollution as measured by PM_{10} cause an increase in mortality. They do, however, establish that there is a consistent association between shorter-term variations in PM_{10} and shorter-term variations in mortality, and that this association is very unlikely explained by the effects of longer-term confounders such as change in medical practice, influenza epidemics, or seasonality, which have been controlled for using city-specific adjustment for longer-term trends. Neither can these

associations be explained by confounding effects of temperature or dew point temperature, which again have been controlled for using city-specific adjustment methods.

References

American Thoracic Society, Bascom, R. (1996a). Health effects of outdoor air pollution, Part 1. *American Journal of Respiratory and Critical Care Medecine* **153**, 3–50.

American Thoracic Society, Bascom, R. (1996b). Health effects of outdoor air pollution, Part 2. *American Journal of Respiratory and Critical Care Medecine* **153**, 477–498.

Bernardo, J. M. and Smith, A. F. M. (1994). *Bayesian Theory*. New York: Wiley.

Best, N. G., Cowles, M. K., and Vines, K. (1995). *CODA: Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output, version 0.30*. Technical report, MRC Biostatistics Unit, University of Cambridge.

Breslow, N. and Clayton, D. (1993). Approximation inference in generalized linear mixed models. *Journal of the American Statistical Association* **88**, 9–25.

Carroll, R. J., Ruppert, D., and Stefanski, L. (1995). *Measurement Error in Nonlinear Models*. Chapman & Hall.

Clyde, M. (1998). Bayesian model averaging and model search strategies. In: *Bayesian Statistics 6*, 83–102.

Cressie, N. (1994). *Statistics for Spatial Data*. Wiley.

Cressie, N., Kaiser, M., Daniels, M., Aldworth, J., Lee, J., Lahiri, S., and Cox, L. (1999). Spatial analysis of particulate matter in an urban environment. In: *GeoEnv 98 Preceedings. Geostatistics for Environmental Applications: Proceedings of the Second European Conference on Geostatistics for Environmental Applications. eds. J. Gomez-Hernandez, A. Soares, R. Froidevaux*.

Daniels, M. and Kass, R. (1998). A note on first-stage approximation in two-stage hierarchical models. *Sankhya, Series B* **60**, 19–30.

Davis, J., Sacks, J., Saltzman, N., Smith, R., and Styer, P. (1996). *Airborne Particulate Matter and daily mortality in Birmingham, Alabama*. Technical report, National Institute of Statistical Science, Research Triangle Park, NC.

DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* **7**, 177–188.

Dockery, D. and Pope, C. (1994). Acute respiratory effects of particulate air pollution. *Annual Review of Public Health* **15**, 107—132.

Dominici, F., Zeger, S., and Samet, J. (2000). A measurement error correction model for time-series studies of air pollution and mortality. *Biostatistics (to appear)* **2**.

DuMouchel, W. (1990). *Bayesian Metaanalysis*. Marcel Dekker.

DuMouchel, W. H. (1990). Bayesian metaanalysis. In: D. A. Berry (ed.), *Statistical Methodology in the Pharmaceutical Sciences*, volume 104 of *Statistics: Textbooks and Monographs*, 509–529, New York, NY, USA. Marcel Dekker.

DuMouchel, W. H. and Harris, J. E. (1983). Bayes methods for combining the results of cancer studies in humans and other species (c/r: P308-315). *Journal of the American Statistical Association* **78**, 293–308.

Fung, Y. and Krewski, D. (1999). On measurement error adjustement methods in poisson regression. *Environmetrics* **10**, 213–224.

Gaudard, M., Karson, M., Linder, E., and Sinha, D. (1999). Bayesian spatial prediction.

Environmental and Ecological Statistics **6**, 147–171.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**(410), 398–409.

Geman, S. and Geman, D. (1993). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Journal of Applied Statistics* **20**, 25–62.

Gilks, W. R., Clayton, D. G., Spiegelhalter, D. J., Best, N. G., McNeil, A. J., Sharples, L. D., and Kirby, A. J. (1993). Modelling complexity: applications of Gibbs sampling in medicine. *Journal of the Royal Statistical Society, Series B, Methodological* **55**, 39–52.

Handcock, M. and Stein, M. (1993). A bayesian analysis of kriging. *Technometrics* **35**, 403–410.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized additive models*. Chapman and Hall, New York.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109.

Janssen, N., Hoek, G., Brunekreef, B., Harssema, H., Mensink, I., and Zuidhof, A. (1998). Personal Sampling of Particles in Adults: Relation among Personal, Indoor, and Outdoor Air Concentrations. *American Journal of Epidemiology* **147**, 537—544.

Janssen, N., Hoek, G., Harssema, H., and Brunekreef, B. (1997). Childhood exposure to PM10: relation between personal, classroom, and outdoor concentrations. *Occupational and Environmental Medicine* **54**, 1—7.

Kaiser, M. and Cressie, N. (1993). The construction of multivariate distributions from markov

random fields. *Journal of Multivariate Analysis (to appear)* **2**, 35–54.

Katsoyanni, K., Toulomi, G., Spix, C., Balducci, F., Medina, S., Rossi, G., Wojtyniak, B., Sunyer, J., Bacharova, L., Schouten, J., Ponka, A., and Anderson, H. R. (1997). Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 european cities: results from time series data from the aphea project. *British Medical Journal* **314**, 1658–1663.

Kelsall, J., Samet, J., and Zeger, S. (1997). Air pollution, and mortality in Philadelphia, 1974-1988. *American Journal of Epidemiology* **146**, 750—762.

Korrick, S., Neas, L., Dockery, D., and Gold, D. (1998). Effects of ozone and other pollutants on the pulmonary function of adult hikers. *Environmental Health Perspectives* **106**, 93–99.

Le, N., Sun, W., and Zidek, J. (1997). Bayesian multivariate spatial interpolation with data missing by design. *Journal of the Royal Statistical Society, Series B* **59**, 501–510.

Le Cam, L. and Yang, G. L. (1990). *Asymptotics in Statistics. Some Basic Concepts*. Springer-Verlag.

Li, Y. and Roth, H. (1995). Daily mortality analysis by using different regression models in Philadelphia county, 1973-1990. *Inhalation Toxicology* **7**, 45–58.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear model (with discussion). *Journal of the Royal Statistical Society, Series B, Methodological* **34**, 1–41.

Lioy, P., Waldman, J., Buckley, T., Butler, J., and Pietarinen, T. (1990). The personal, indoor, and outdoor concentrations of PM10 measured in an industrial community during the winter. *Atmospheric Environment* **24**, 57—66.

Lipfert, F. and Wyzga, R. (1993). Air pollution and mortality: Issues and uncertainty. *Journal Air Waste Manage. Assoc* **45**, 949—966.

Mage, T. and Buckley, T. J. (1995). The Relationship between Personal Exposures and Ambient Concentrations of Particulate Matter. In: *Air and Waste Management, 88-th Annual Meeting*, 1–16.

Morris, C. N. and Normand, S.-L. (1992). Hierarchical models for combining information and for meta-analysis. In: J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.), *Bayesian Statistics 4*, 321–344, Oxford. Oxford University Press.

Neyman, J. and Scott, E. L. (1960). Correction for bias introduced by a transformation of variables (com: P656-661). *The Annals of Mathematical Statistics* **31**, 643–655.

Normand, S., Glickman, M., and Gatsonis, C. (1997). Statistical methods for profiling providers of medical care: Issues and applications. *Journal of the American Statistical Association* **92**, 803–814.

Ozkaynak, H., Xue, J., Spengler, J., Wallace, L., Pellizzari, E., and Jenkins, P. (1996). Personal Exposure to Airborne Particles and Metals: Results from the Particle Team Study in Riverside, California. *Journal of Exposure Analysis and Environmental Epidemiology* **6**, 57—78.

Raftery, A. and Lewis, S. (1992). How many iterations in the gibbs sampler? In: *Bayesian*

Statistics 4. Proceedings of the Fourth Valencia International Meeting, 763–774.

Samet, J., Zeger, S., Kelsall, J., Xu, J., and Kalkstein, L. (1997). *Air pollution, weather and mortality in Philadelphia, In Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants. The Phase IB report of the Particle Epidemiology Evaluation Project*. Health Effect Institute, Cambridge, MA.

Schwartz, J. (1995). Air pollution and daily mortality in Birmingham, alabama. *American Journal of Epidemiology* **137**, 1136–1147.

Silliman, N. (1997). Hierarchical selection models with applications in meta-analysis. *Journal of the American Statistical Association* **92**, 926–936.

Skene, A. M. and Wakefield, J. C. (1990). Hierarchical models for multicentre binary response studies. *Statistics in Medicine* **9**, 919–929.

Smith, L., Davis, J., Sacks, J., Speckman, P., and Styer, P. (1997). Assessing the human health risk of atmospheric particle. *ASA Proceedings, Env Stat* .

Smith, R., Davis, J., Sacks, J., Speckman, P., and Styer, P. (1998). *Air pollution and mortality in Birmingham, Alabama: A reappraisal*. Technical report, National Institute of Statistical Science, Research Triangle Park, NC.

Smith, T. C., Spiegelhalter, D. J., and Thomas, A. (1995). Bayesian approaches to random-effects meta-analysis: A comparative study. *Statistics in Medicine* **14**, 2685–2699.

Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics* **22**(4), 1701–1762.

Toulomi, G., Katsouyanni, K., Zmirou, D., and Schwartz, J. (1997). Short-term effects of

ambient oxidant exposure on mortality: A combined analysis within the APHEA project. *American Journal of Epidemiology* **146**, 177–183.

Vedal, S. (1996). Ambient particles and health: Lines that divide. *J. Air and Waste Manage. Assoc.* **47**, 551—–581.

Wakefield, J. and Morris, S. (1998). Spatial Dependence and Errors-in-Variables in Environmental Epidemiology: Investigating the Relationship Between Magnesium and Myocardial Infarction. In: *Bayesian Statistics 6*, 510–516.

Wallace, L. (1996). Indoor Particles: A Review. *J. Air and Waste Manage. Assoc.* **46**, 98—126.

Wikle, C., Berliner, M., and Cressie, N. (1997). *Hierarchical Bayesian Space-Time Models*. Technical report, Iowa State University.

Zeger, S., Thomas, D., Dominici, F., Cohen, A., Schwartz, J., Dockery, D., and Samet, J. (2000). Measurement Error in Time-Series Studies of Air Pollution: Concepts and Consequences. *Environmental Health Perspectives* **108**(5).

Zidek, J., Wong, H., Le, N., and Burnett, R. (1996). Causality, measurement error and multicollinearity in epidemiology. *Environmetrics* **7**, 441–451.

Tables Captions

- **Table 1.** Summary by location of the county population (Pop), percentage days with missing values ($P_{missO3}, P_{missPM10}$), percentage of people in poverty ($P_{poverty}$), percentage of people older than 65 years ($P_{>65}$), average of pollutant levels for O_3 and PM_{10} ($\bar{X}_{O3}, \bar{X}_{PM10}$), and average daily of deaths ($\bar{Y}$).

- **Table 2.** Results of the adjusted second-stage analyses under the baseline-univariate model. (PM_{10} entered independently in the model). Posterior means and 95% posterior support intervals $(,)$ of the coefficients for the relationship between the true relative rate, β^c, the percentage in poverty, $P_{poverty}$, the percentage of people older than 65, $P_{>65}$, and the mean level of the pollutant $\bar{X}_{PM10}$. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from two days before (lag 2).

- **Table 3.** Posterior means and 95% support intervals of the elements of Σ (*std of* PM_{10} *effects* = standard deviation across locations of the β^c_{PM10}, *std of* O_3 *effects* = standard deviation across locations of the $\beta^c_{O_3}$, *corr of* PM_{10} *and* O_3 *effects* = correlation between the β^c_{PM10} and $\beta^c_{O_3}$) under the three models (univariate, bivariate, spatial).

- **Table 4.** Posterior probabilities that the overall effects of PM_{10} are larger than 0 by lag, by three prior distributions under the three models (univariate, bivariate, spatial). The three prior specifications consist in the following 95% support intervals of the overall effects, the city-specific effects, and of the spatial correlation for the relative rates of the two closets cities having median distance: a: $(-15, 15), (-4, 4), (0.11, 0.74)$; b:

$(-4, 4), (-4, 4), (0.11, 0.74)$; c: $(-4, 4), (-7, 7), (0, 0.9)$.

Figure Captions

Figure 1. Map of the 20 cities with largest population including the surrounding country. The dimensions of the circles are proportional to the county populations. The cities are numbered from 1 to 20 following the order of Table 1.

Figure 2. Results of regression models for the twenty cities by selected lag: $\hat{\beta}^c$, and 95 $\%$ confidence intervals of $\hat{\beta}^c \times 1000$ for PM_{10}. Cities are presented in decreasing order by population living within their county limits. The vertical scale can be interpreted as the percentage increase in mortality per 10 $\mu g/m^3$ increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from two days before (lag 2).

Figure 3. Results of Regression Models for the twenty cities by selected lag: $\hat{\beta}^c$, and 95 $\%$ Confidence Intervals of $\hat{\beta}^c \times 1000$ for PM_{10} adjusted by O_3. Cities are presented in decreasing order by population living within their county limits. The empty circle placed at Minneapolis represents the missingness of the ozone data in this city. The vertical scale can be interpreted as the percentage increase in mortality per 10 $\mu g/m^3$ increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from two days before (lag 2).

Figure 4. Plots of city-specific autocorrelation functions of standardized residuals r_t, where $r_t = (Y_t - \hat{Y}_t)/\sqrt{(\hat{Y}_t)}$ and $\hat{Y}_t$ are the fitted values from log-linear generalized additive model (1).

Figure 5. Comparison between the normal approximation of the likelihood of $\boldsymbol{\beta}^c$ and the marginal posterior distribution of $\boldsymbol{\beta}^c$. The solid line represents the normal density $N(\hat{\boldsymbol{\beta}}^c, V^c)$ where $\hat{\boldsymbol{\beta}}^c$ and V^c are the MLEs of a semiparametric Poisson regression model. The histogram represents the marginal posterior distribution of $\boldsymbol{\beta}^c$ obtained by implementing a full Gibbs Sampler for the parameter of interest $\boldsymbol{\beta}^c$ and for the coefficients of the natural cubic splines $\boldsymbol{\eta}^c$.

Figure 6. Results of pooled analyses under the univariate baseline model (PM_{10} entered independently in the model). Boxplots of samples from the posterior distributions of city-specific regression coefficients, β^c, associated with changes in PM_{10} measurements. For comparison, samples from the marginal posterior distribution of the corresponding overall effects are displayed at far right. The vertical scale can be interpreted as the percentage increase in mortality per 10 $\mu g/m^3$ increase in PM_{10}. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from two days before (lag 2).

Figure 7. Results of pooled analyses under the bivariate baseline model (PM_{10} and O_3 entered simultaneously in the model). Boxplots of samples from the posterior distributions of city-specific regression coefficients, β^c, associated with changes in PM_{10} adjusted by O_3 measurements. For comparison, samples from the marginal posterior distribution of the corresponding overall effects, are displayed at far right. The vertical scale can be interpreted as the percentage increase in mortality per 10 $\mu g/m^3$ increase in PM_{10} respectively. The results are reported using the concurrent day (lag 0) pollution values to predict mortality, using the previous day's (lag 1) pollution levels, and using pollution levels from two days

before (lag 2).

Figure 8. Results of pooled analyses under the univariate baseline, bivariate-baseline, and spatial models. Marginal posterior distributions of the overall effects, $\alpha_{0,pm10}$ for different lags. At the top right are specified the posterior probabilities that the overall effects are larger than zero.

Table 1:

Locations (State)	Label	Pop	P_{missO3}	P_{missPM}	$P_{poverty}(\%)$	$P_{>65}(\%)$	$\bar{X}_{O3}(ppb)$	$\bar{X}_{PM}(\mu g/m^3)$	$\bar{Y}$
LosAngeles	la	8863164	0	80.2	14.8	9.7	22.84	45.98	148
NewYork	ny	7510646	0	83.3	17.6	13.2	19.64	28.84	191
Chicago	chic	5105067	0	8.2	14.0	12.5	18.61	35.55	114
Dallas-Fortworth	dlft	3312553	0	78.6	11.7	8.0	25.25	23.84	49
Houston	hous	2818199	0	72.9	15.5	7.0	20.47	29.96	40
SanDiego	sand	2498016	0	82.2	10.9	10.9	31.64	33.63	42
SantaAna-Ana.	staa	2410556	0	83.6	8.3	9.1	22.97	37.37	32
Phoenix	phoe	2122101	0.1	85.1	12.1	12.5	22.86	39.75	38
Detroit	det	2111687	36.3	53.9	19.8	12.5	22.62	40.90	47
Miami	miam	1937094	1.4	83.4	17.6	14.0	25.93	25.65	44
Philadelphia	phil	1585577	0.7	83.1	19.8	15.2	20.49	35.41	42
Minneapolis	minn	1518196	100	5.4	9.7	11.6	NA	26.86	26
Seattle	seat	1507319	37.3	24.5	7.8	11.1	19.37	25.25	26
SanJose	sanj	1497577	0	67.7	7.3	8.6	17.87	30.35	20
Cleveland	clev	1412141	41.4	55.6	13.5	15.6	27.45	45.15	36
SanBernardino	sanb	1412140	0	81.6	12.3	8.7	35.88	36.96	20
Pittsburgh	pitt	1336449	1.3	0.8	11.3	17.4	20.73	31.61	38
Oakland	oakl	1279182	0	82.6	10.3	10.6	17.24	26.31	22
SanAntonio	sana	1185394	0.1	77.1	19.4	9.8	22.16	23.83	20
Riverside	river	1170413	0	81.3	14.8	11.3	33.41	51.99	20

Table 2:

City-Specific covariates	Lag 0	Lag 1	Lag 2
Overall-PM_{10}	0.48 (0.02,0.91)	0.55 (0.11,1)	0.44 (0.00,0.9)
$P_{poverty}$ (%)	-0.005 (-0.17,0.07)	0.00 (-0.11,0.13)	0.03 (-0.10,0.14)
$P_{>65}$ (%)	-0.03 (-0.20,0.15)	0.02 (-0.15,0.2)	0.00 (-0.17,0.17)
$\bar{X}^c_{PM10}$ ($\mu g/m^3$)	0.01 (-0.05,0.07)	-0.01 (-0.07,0.05)	0.00 (-0.07,0.06)

Table 3:

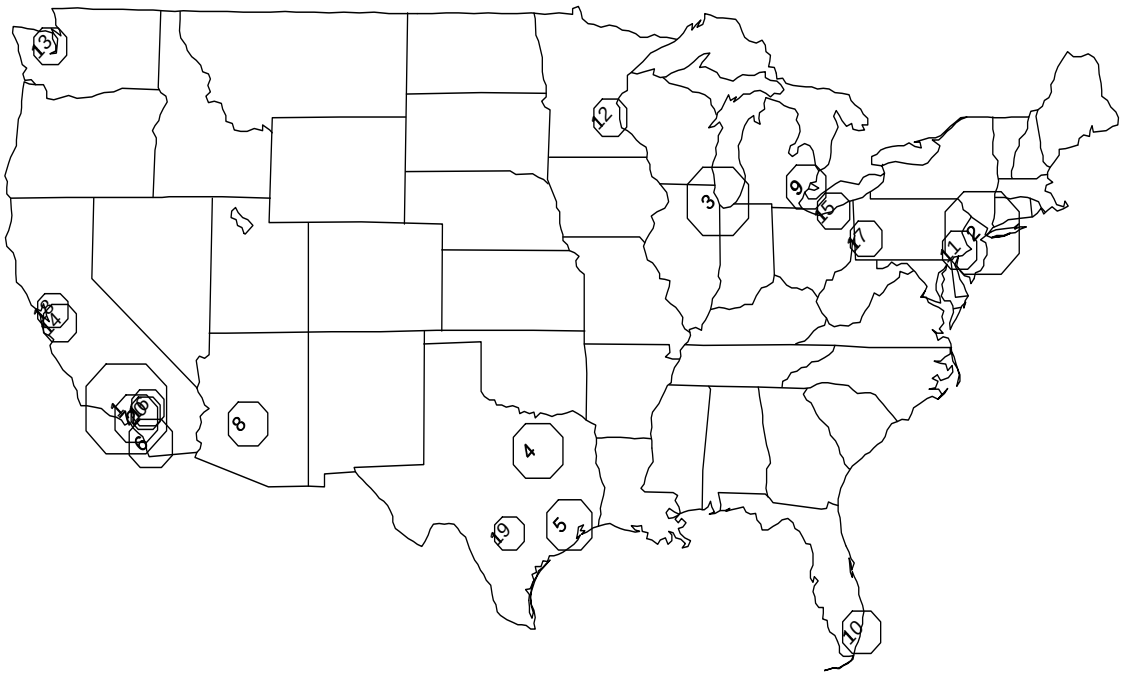
euclidean distance model			
	25 %	*50 %*	*75%*
distance quantiles	0.2	0.67	1.57
prior correlation $e^{-\theta d}$	0.8 (0.53,0.93)	0.45 (0.11,0.74)	0.16 (0.06,0.5)
posterior correlation $e^{-\theta d}$	0.86 (0.68,0.93)	0.61 (0.30,0.80)	0.30 (0.05,0.58)

Table 4:

Model	*std of PM_{10} effects*	*std of O_3 effects*	*corr.of PM_{10} and O_3 effects*
Baseline-Bivariate	0.36 (0.17,0.75)	0.91 (0.33,2.01)	-0.09 (-0.5,0.22)
Baseline-Univariate	0.76 (0.41,1.37)	1.28 (0.69,2.28)	
Spatial	0.71 (0.38,1.27)	1.21 (0.61,2.32)	

Table 5:

PM_{10} $(\mu g/m^3)$									
Model	Prior 1[a]			Prior 2[b]			Prior 3[c]		
	Lag 0	Lag 1	Lag 2	Lag 0	Lag 1	Lag 2	Lag 0	Lag 1	Lag 2
Baseline-Univariate	0.98	0.98	0.99	0.98	0.96	0.98	0.95	0.96	0.93
Baseline-Bivariate	1	1	0.97	1	0.99	0.99	0.98	1	0.93
Spatial	0.83	0.95	0.92	0.83	0.93	0.91	0.78	0.89	0.85


13
12
3
9
15
17
2
14
16
18
8
4
19
5
10

Univariate Model for PM_{10}

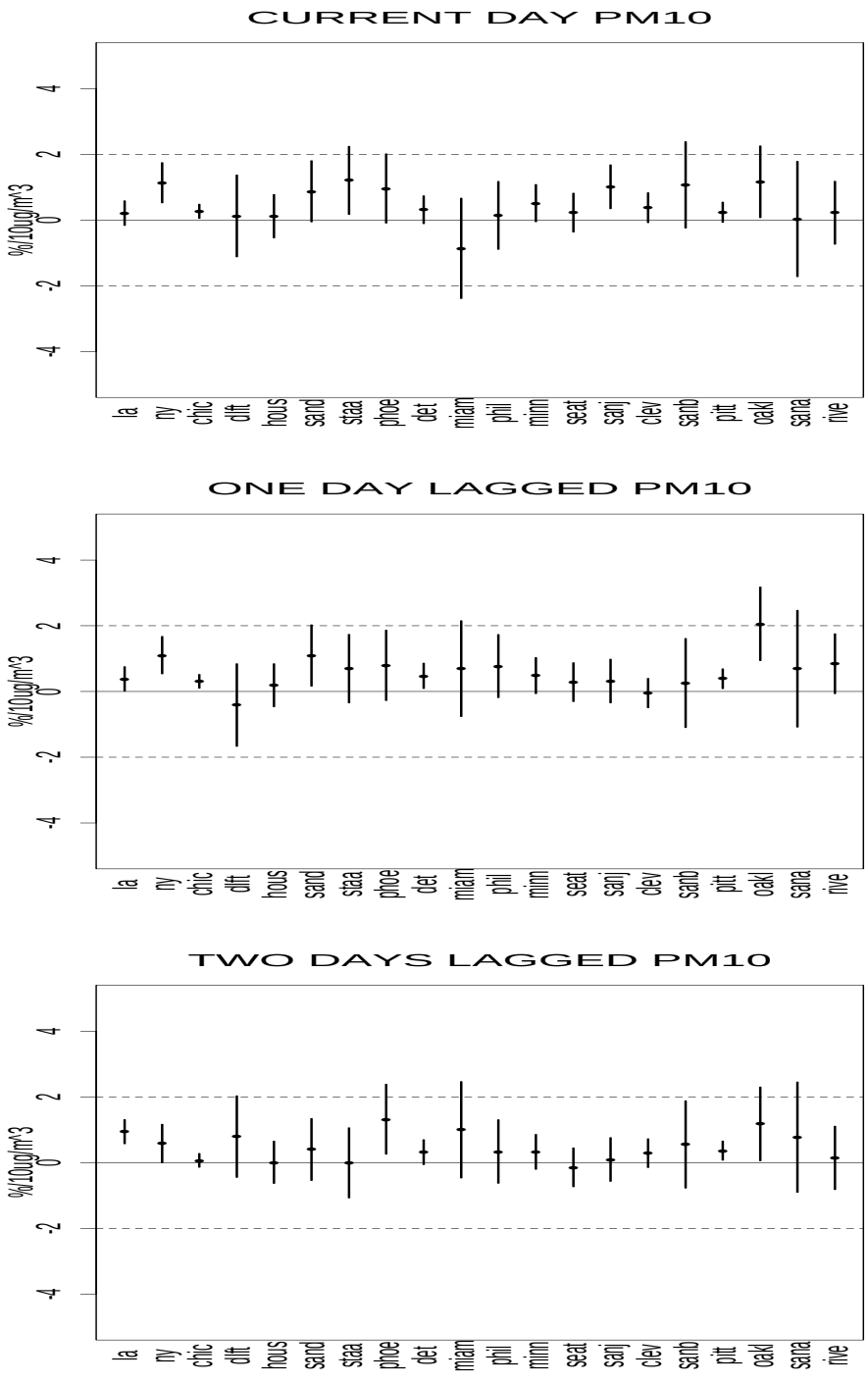


Figure 2:

Bivariate Model for PM_{10} adjusted by O_3

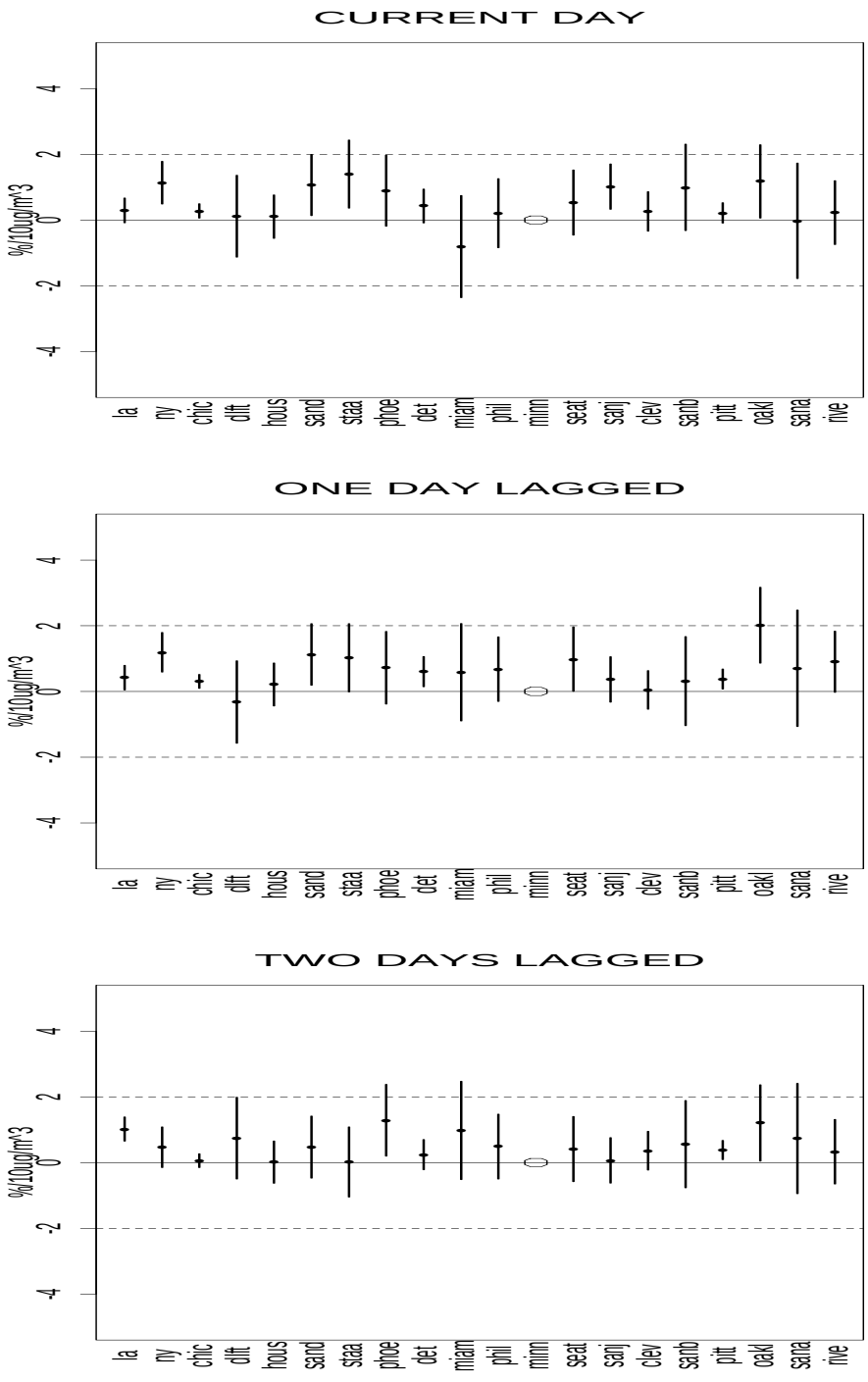


Figure 3:

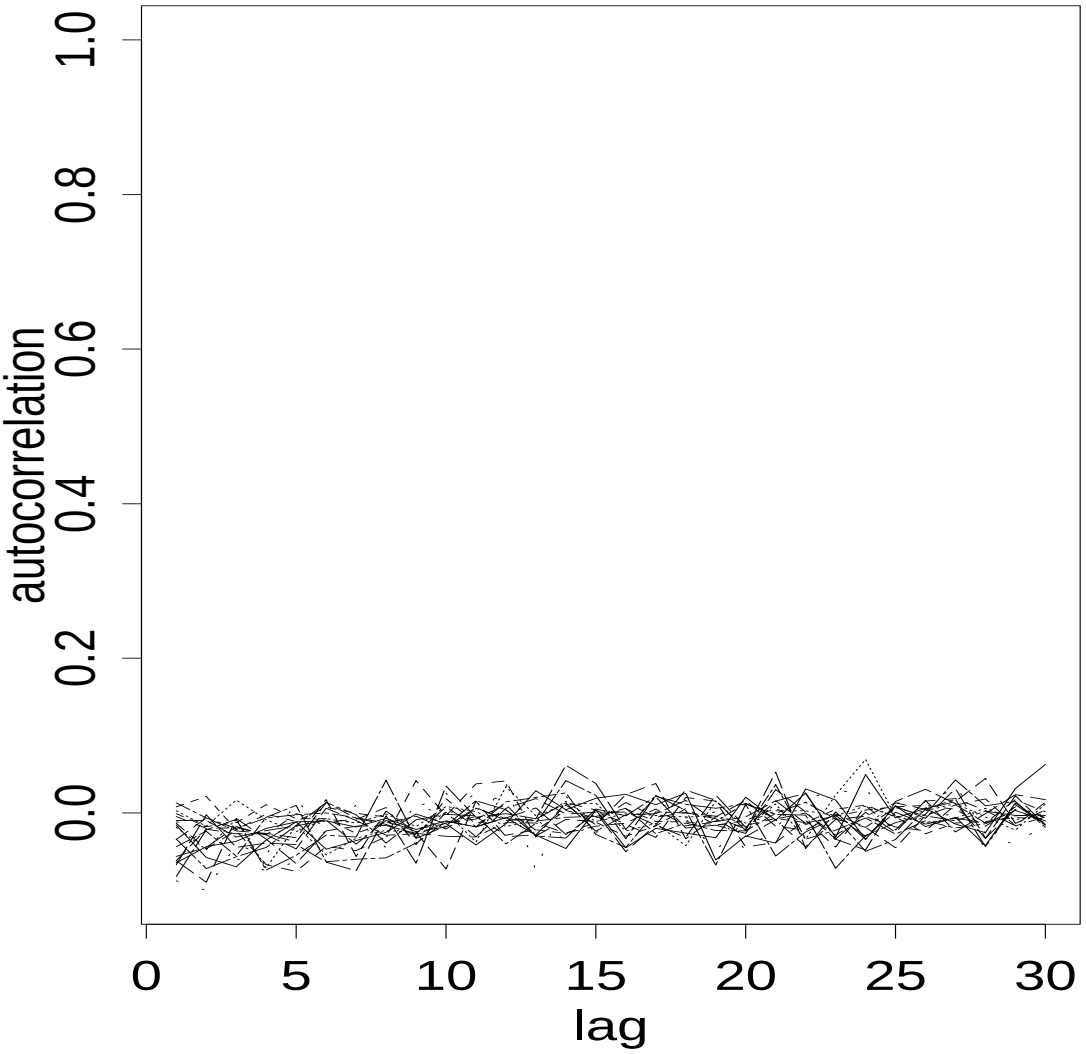


Figure 4:

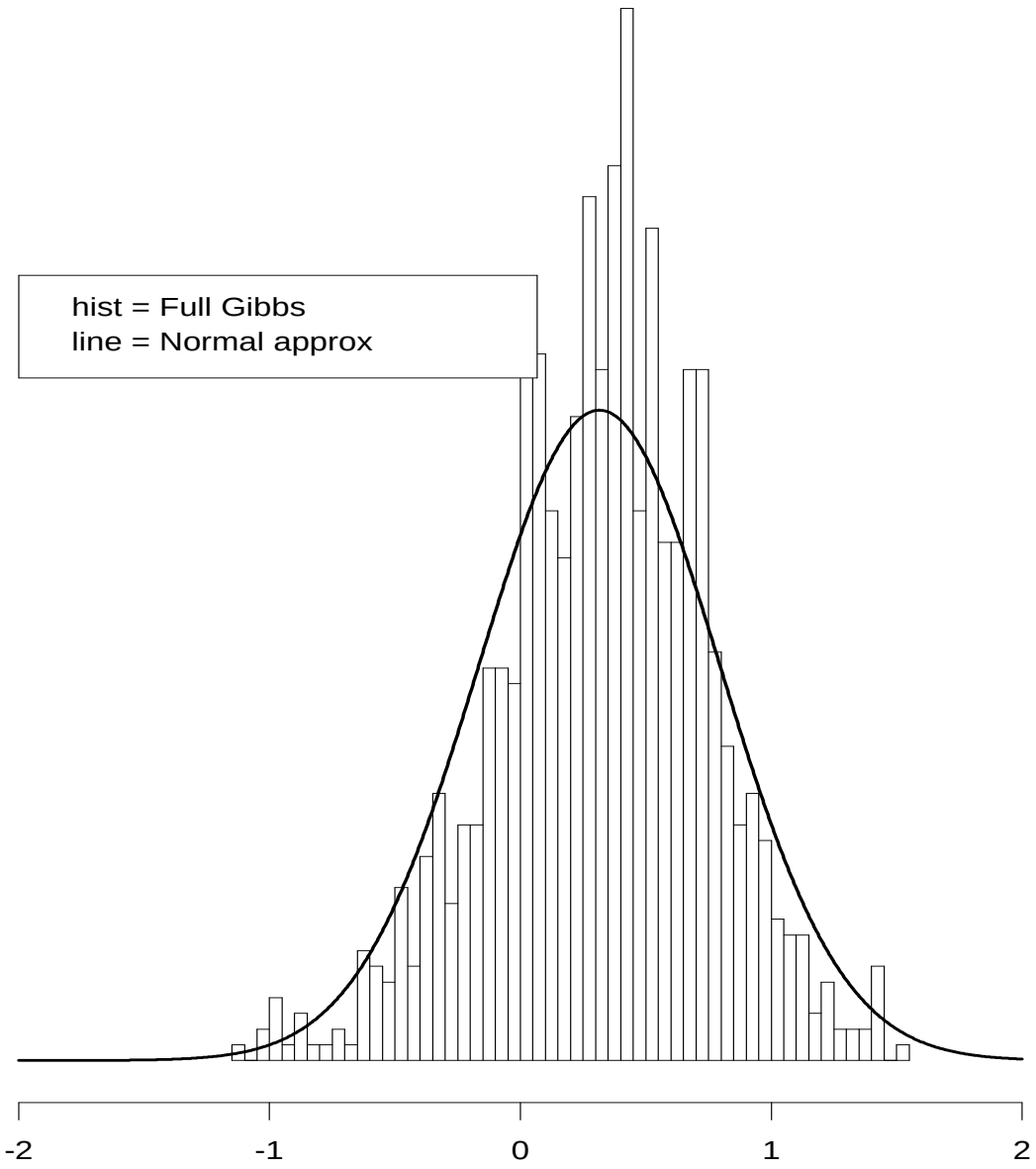


Figure 5:

Univariate Model for PM_{10}

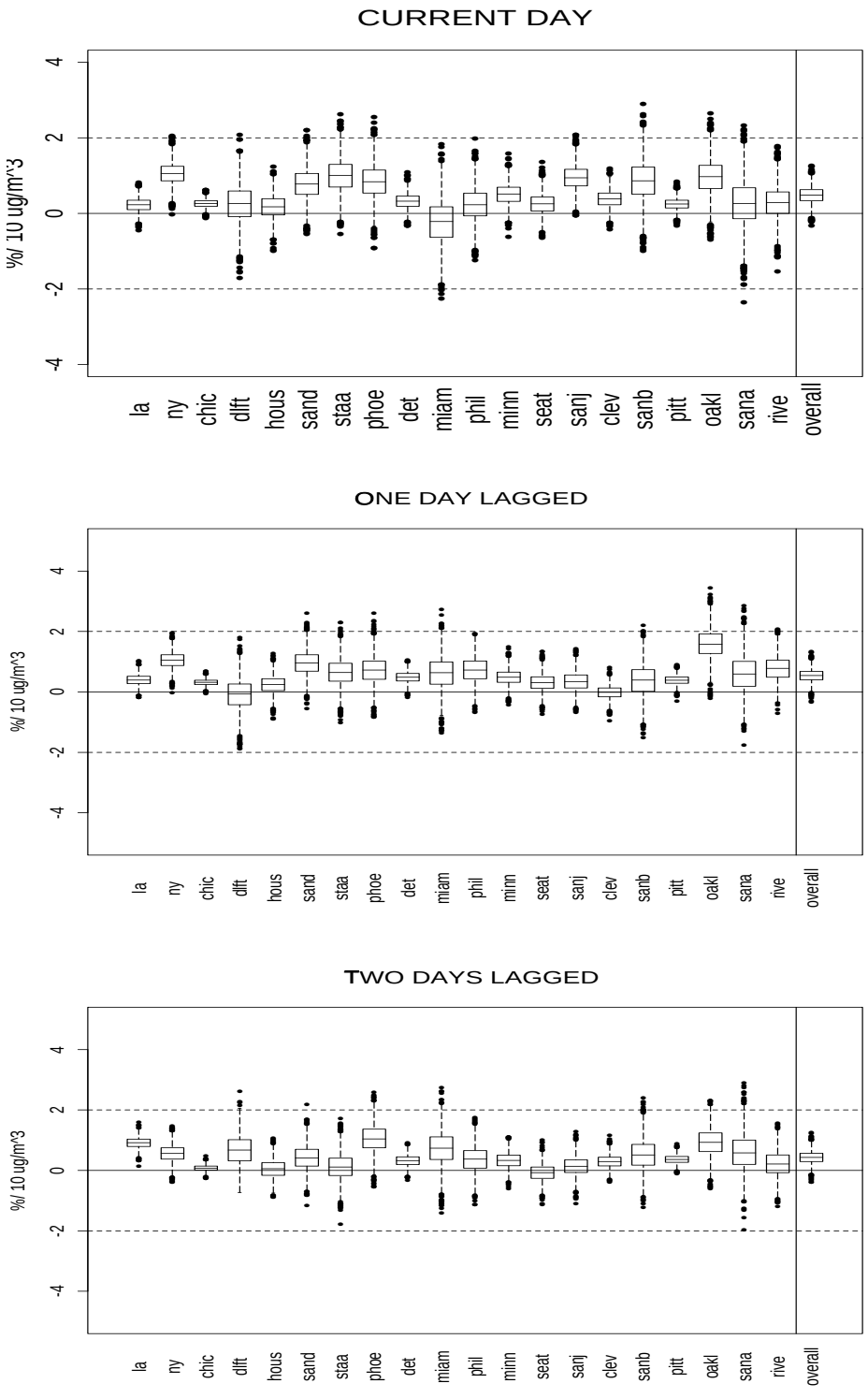


Figure 6:

Bivariate Model for PM_{10} adjusted by O_3

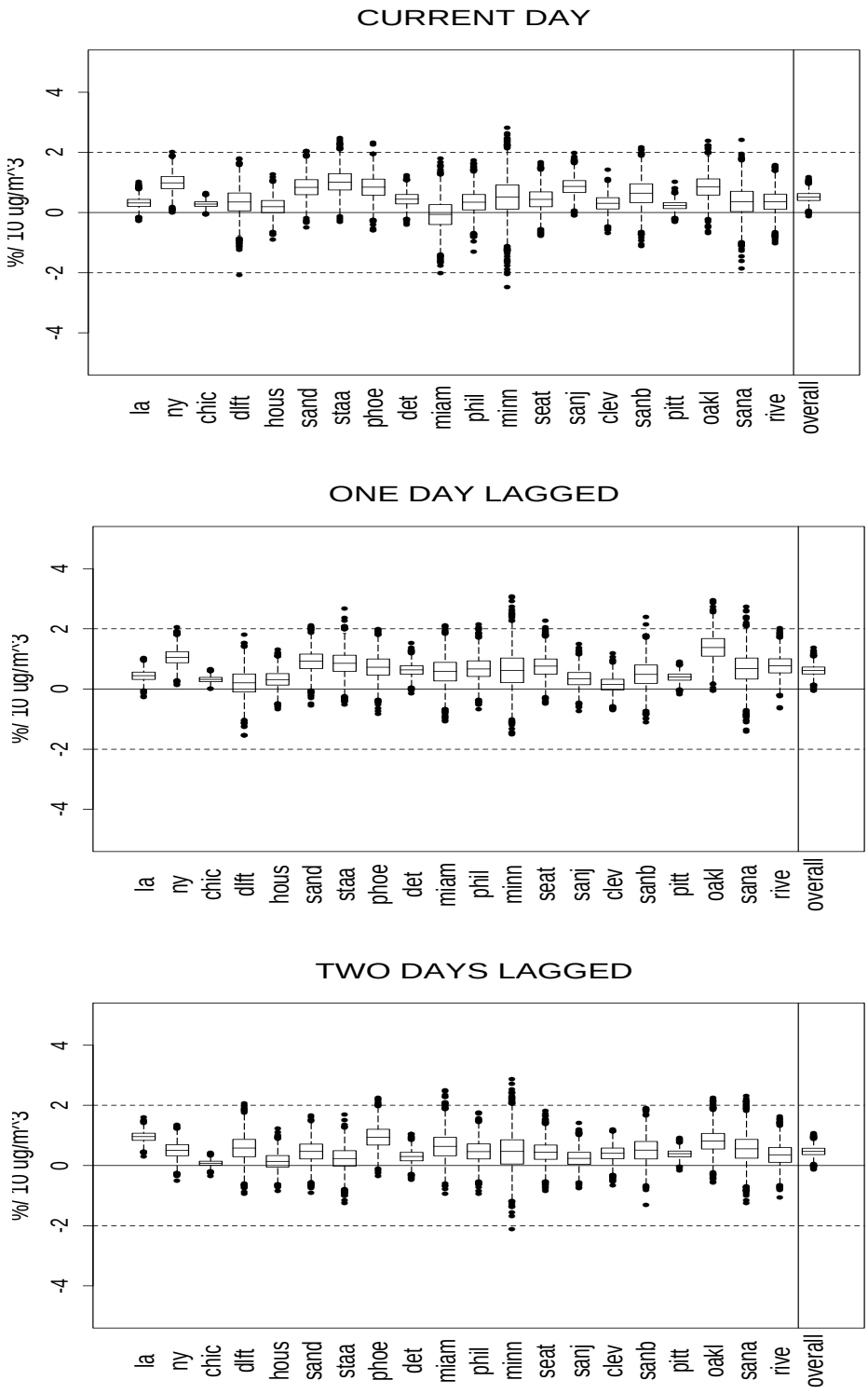
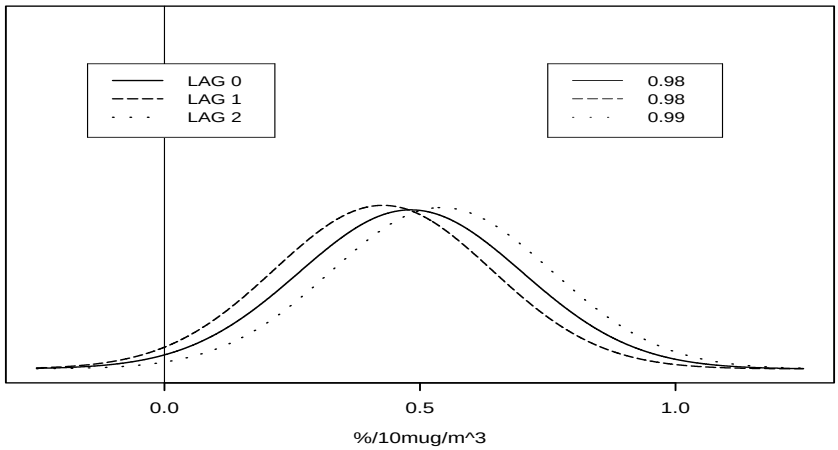
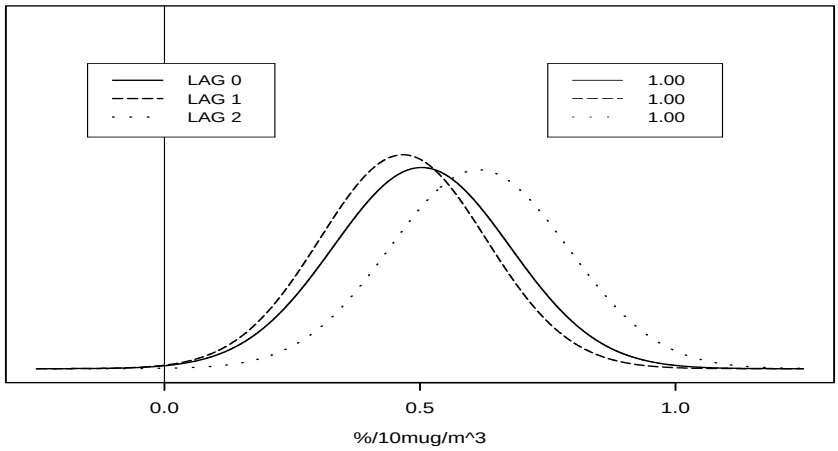


Figure 7:

Univariate Model for PM_{10}



Bivariate Model for PM_{10} adjusted for O_3



Spatial Model for PM_{10} with Euclidean distance

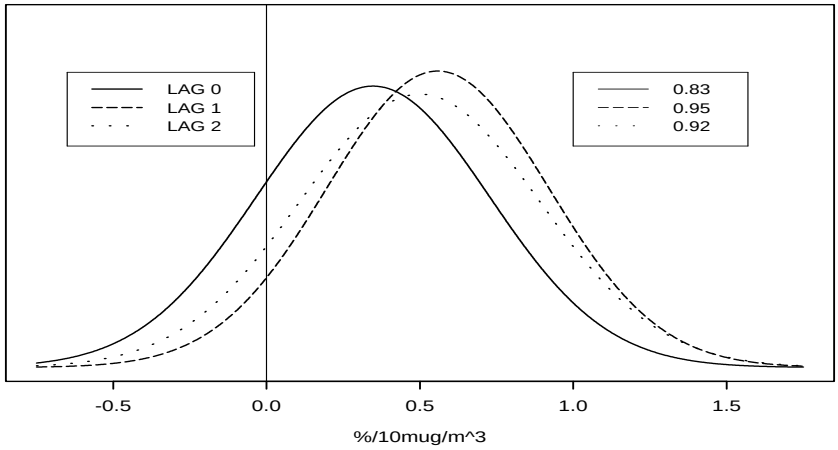


Figure 8: