Available online at www.sciencedirect.com
SCIENCE DIRECT
Atmospheric Environment 40 (2006) 984–991
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Investigating the mixture of air pollutants associated with adverse health outcomes

Steven Roberts*[,1], Michael A. Martin[2]

Faculty of Economics and Commerce, School of Finance and Applied Statistics, Australian National University, Canberra ACT 0200, Australia

Received 17 May 2005; accepted 12 October 2005

Abstract

Most investigations of the adverse health effects of multiple air pollutants analyse the time series involved by simultaneously entering the multiple pollutants into a Poisson log–linear model. Concerns have been raised about this type of analysis, and it has been stated that new methodologies or models need to be developed for investigating the adverse health effects of multiple air pollutants. To this end, it has recently been stated that it may be more reasonable to assume that there is a mixture of pollutants considered harmful to health and that assessing the adverse health effects of an air pollution mix may be both more meaningful and more tenable than attempting to isolate the effects of individual pollutants. In this paper, a new model is introduced that is able to reveal the mixture of pollutants associated with an adverse health outcome and the effect of this mixture on the adverse health outcome. The model is shown to have a number of advantages over the traditional method of estimating the adverse health effects of multiple air pollutants. In addition, the model is also shown to be an improvement over a previously proposed, and somewhat ad-hoc, method for estimating the mixture of pollutants associated with an adverse health outcome.

Keywords: Air pollution; Mortality; Time series; Multiple pollutants

1. Introduction

Numerous time series studies have investigated the association between daily adverse health outcomes and daily ambient air pollution concentrations (Chock et al., 2000; Cifuentes et al., 2000; Goldberg et al., 2003; Kelsall et al., 1997, 2000; Kwon et al., 2001; Moolgavkar, 2000; Ostro et al., 1999; Roemer and van Wijnen, 2001; Smith et al., 2000a, b; Stieb et al., 2002; Styer et al., 1995). These studies typically fit a Poisson log–linear model to concurrent time series of daily mortality or morbidity, ambient air pollution and meteorological covariates. The fitted models are then used to quantify the adverse health effects of ambient air pollution. Because the US Environmental Protection Agency regulates pollutants independently, most of the current time series research on the adverse health effects of air pollution has focused on estimating the effects of an individual pollutant

*Corresponding author. Tel.: +61 2 6125 3470; fax: +61 2 6125 0087.
E-mail address: steven.roberts@anu.edu.au (S. Roberts).
[1]Senior Lecturer.
[2]Reader.

1352-2310/$ - see front matter
doi:10.1016/j.atmosenv.2005.10.022

(Dominici and Burnett, 2003). However, due to the potentially high correlation between ambient air pollutants, the results from studies that focus on a single pollutant can be difficult to interpret in practice (Vedal et al., 2003). For example, an observed positive association could occur because the single air pollutant is a proxy for another air pollutant or for a mixture of air pollutants.

To overcome the limitations of single-pollutant time series studies, a number of studies have investigated the concurrent adverse health effects of multiple air pollutants (Moolgavkar, 2000; Wong et al., 2002). In the majority of studies of this nature, the multiple air pollutants are simultaneously entered into a single Poisson log–linear model. The results from these studies are used to isolate the adverse health effects of the individual pollutants. However, one important question that these multiple pollutant studies do not answer is whether there is a mixture of pollutants that is associated with the adverse health outcome. Moreover, it has recently been stated that it may be more reasonable to assume that there is a mixture of pollutants that is considered harmful to health (Dominici and Burnett, 2003; Moolgavkar, 2003; Stieb et al., 2002). Assessing the adverse health effects of an air pollution mix may, therefore, be both more interpretable and more feasible than attempting to isolate the effects of individual pollutants. The development of new methodology or models to concurrently estimate the adverse health effects of multiple air pollutants has been identified by statisticians, epidemiologists and policymakers as an important area of future research (Cox, 2000; Dominici and Burnett, 2003).

In this paper, a new model is introduced that reveals the mixture of pollutants associated with an adverse health outcome and the effect of this mixture on that outcome. This new model uses time series data to assign each air pollutant a weight which indicates the pollutant's contribution to the air pollution mixture that is associated with the adverse health outcome under investigation. The model is illustrated by applying it to time series data from nine United States counties for the period 1987–2000.

2. Materials and methods

2.1. Materials

The data used in this paper were obtained from the publicly available National Morbidity, Mortality, and Air Pollution Study (NMMAPS) database. The data extracted consists of concurrent daily time series of mortality, weather and air pollution for nine cities in the United States for the period 1987–2000. The nine cities selected had a relatively large number of days with measurements for all five air pollutants considered. Many of the cities in the NMMAPS database do not collect data on all five air pollutants and/or have a large number of days with missing air pollutant concentrations.

The mortality time series data, aggregated at the level of county, are non-accidental daily deaths of individuals aged 65 and over. Deaths of non-residents were excluded from the mortality counts. The weather time series data are 24 h averages of temperature and dew point temperature, computed from hourly observations.

The five air pollutants considered are particulate matter of less than 10 μm in diameter (PM), ozone (O_3), sulphur dioxide (SO_2), carbon monoxide (CO) and nitrogen dioxide (NO_2). For PM, SO_2, CO and NO_2 average daily concentrations were used. For O_3, the maximum hourly concentration for each day was used. In the analyses that follow, each of the pollutant time series was standardised to have a unit variance.

2.2. Methods

The majority of time series studies that concurrently investigate the adverse health effects of multiple air pollutants simultaneously enter the pollutants into a single Poisson log–linear model. Under this model, the daily adverse health outcome counts are modelled as independent Poisson random variables with a time varying mean μ_t where

$$\log(\mu_t) = \text{confounders}_t + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt}, \tag{1}$$

and where confounders$_t$ represents other time-varying variables which are related to the adverse health outcome, $X_{it}, i = 1, \ldots, k$, represent the k pollutants that are being investigated, and $\beta_i, i = 1, \ldots, k$, measure the adverse health effect of pollutant i. Hereafter, model (1) will be referred to as the "*standard model*".

As discussed above, one important question that the *standard model* cannot answer is whether there is a mixture of pollutants that is associated with the adverse health outcome under investigation. To answer this question, we propose fitting the following model which, like the *standard model*, models the daily adverse health outcome counts as independent Poisson random variables with a time varying mean μ_t, but where

$$\log(\mu_t) = \text{confounders}_t + \beta(w_1 X_{1t} + w_2 X_{2t} + \cdots + w_k X_{kt}), \tag{2}$$

and where $w_i, i = 1, \ldots, k$, is the weight assigned to pollutant i in the relevant air pollutant mixture associated with the adverse health outcome. The weights are constrained to be non-negative, and sum to one. This ensures that $w_1 X_{1t} + w_2 X_{2t} + \cdots + w_k X_{kt}$ can be readily interpreted as an air pollutant mixture. β is the effect of the air pollutant mixture $(w_1 X_{1t} + w_2 X_{2t} + \cdots + w_k X_{kt})$ on the adverse health outcome. The terms confounders$_t$ and $X_{it}, i = 1, \ldots, k$, are as defined in model (1). Model (2) will be referred to as the "*weighted model*".

The *weighted model* is fit by expressing the $w_i, i = 1, .., k$ in model (2) as

$$w_i = \frac{\exp(\lambda_i)}{1 + \sum_{i=1}^{k-1} \exp(\lambda_i)}, \quad i = 1, \ldots, k-1,$$
$$w_k = 1 - \sum_{i=1}^{k-1} w_i.$$

Maximum likelihood estimation is then used to estimate λ_i and hence w_i. Expressing w_i in this form allows the coefficients in model (2) to be estimated using unconstrained rather than constrained optimisation software. The maximum likelihood estimates of $(\beta, w_1, \ldots, w_k)$ are obtained iteratively in two repeated steps. The first step fixes the parameters corresponding to the air pollution terms $(\beta, w_1, \ldots, w_k)$ at their current values and estimates the parameters corresponding to the confounders. This step can be performed using generalised linear modelling software. The second step fixes the parameters corresponding to the confounders at their current values and estimates the parameters corresponding to the air pollution terms. This step can be performed using standard optimization software. The two steps are iterated until convergence to obtain the final parameter estimates. The results in this paper were obtained using the statistical package S-PLUS.

To reiterate, the *standard model* estimates an individual coefficient (β_i) for the effect of each air pollutant on the adverse health outcome, while the *weighted model* produces a (positive) weight for each pollutant (w_i) and an overall estimated coefficient (β) of the effect of the air pollutant mixture defined by the weights $(w_1 X_{1t} + w_2 X_{2t} + \cdots + w_k X_{kt})$ on the adverse health outcome. The output of this new model is therefore twofold: first, the discovery of a relevant air pollutant mixture related to mortality and, second, an indication of the effect of that mixture on the adverse health outcome. The air pollutant mixture found by the *weighted model* is readily interpretable because the weights are constrained to be non-negative and sum to one. It is important to note that there is no re-parameterisation that will allow the *standard model* to return a readily interpretable air pollutant mixture as well as an estimate of the effect of this air pollutant mixture on the adverse health outcome. For instance, the *standard model* may result in different pollutants being assigned coefficients having different signs—this phenomenon will never occur for the weighted model. This means that, unlike the *standard model*, the *weighted model* is able to address the important question of whether there is a biologically relevant pollutant mixture that is related to the adverse health outcome under investigation by forming, as part of the fitting process, a particular linear, or weighted, combination of pollutants. Moreover, through interpretation of the weights, it is also possible to consider the relative importance of individual pollutants to the overall mixture.

3. Simulation study

In this section, a simulation study is conducted to confirm that the *weighted model* is able to recover the air pollutant mixture that is associated with the adverse health outcome of interest.

In order to conduct the simulations, a way of generating realistic mortality time series with known air pollution mortality effects was required. We used a method previously shown to generate a realistic mortality time series, which proceeds by fitting the following Poisson log–linear model similar to those used in previous NMMAPS analyses (Daniels et al., 2000), to the actual Cook County mortality and meteorological time series data

$$\begin{aligned}\log(\mu_t) = \mu &+ S_{t1}(\text{time},\ 7\,\text{df per year}) \\ &+ S_{t2}(\text{temp}_0,\ 6\,\text{df}) + S_{t3}(\text{temp}_{1-3},\ 6\,\text{df}) \\ &+ S_{t4}(\text{dew}_0,\ 3\,\text{df}) + S_{t5}(\text{dew}_{1-3},\ 3\,\text{df}) \\ &+ \gamma \text{DOW}_t\end{aligned} \tag{3}$$

where the subscript t refers to the day of the study, μ_t is the mean number of deaths on day t and μ is an intercept term. The quantities $S_{ti}()$ are smooth functions of time, temperature and dew point temperature, with the indicated degrees of freedom. The smooth functions are represented using natural cubic splines. The quantity temp_0 is the current day's mean 24 h temperature and temp_{1-3} is the average of the previous three days' 24 h mean temperatures. The values dew_0 and dew_{1-3} are similarly defined for the 24 h mean dew point temperature, and DOW_t is a set of indicator variables for the day of the week.

Once model (3) was fit, the estimated mean mortality counts, denoted $\hat{\mu}_t$, were extracted. The effects of the five air pollutants on mortality were then explicitly specified and incorporated into the generated mortality time series, by producing mortality time series of length 2920 days that were Poisson distributed with mean ψ_t on day t where

$$\log(\psi_t) = \log(\hat{\mu}_t) + \theta(\alpha_1 X_{1t} + \alpha_2 X_{2t} + \alpha_3 X_{3t} + \alpha_4 X_{4t} + \alpha_5 X_{5t}), \quad (4)$$

and where, X_{it}, $i = 1, \ldots, 5$ are, respectively, the current day's daily concentrations of PM, NO_2, CO, O_3 and SO_2, α_i, $i = 1, \ldots, 5$ are the corresponding weights for each pollutant and θ is the effect of the air pollutant mixture $(\alpha_1 X_{1t} + \alpha_2 X_{2t} + \alpha_3 X_{3t} + \alpha_4 X_{4t} + \alpha_5 X_{5t})$ on mortality.

In the simulations, nine $(\theta, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ combinations were used. For these nine combinations, the effect of the air pollutant mixture on mortality θ ranged from 0 to 0.1. Since each of the air pollutant time series was standardised to have unit variance, a θ value of 0.1 corresponds to approximately a 10% increase in mortality for a simultaneous 1 standard deviation increment in the concentration of each air pollutant.

Table 1 contains for each of the nine $(\theta, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ combinations, the mean and standard deviation of the β_i and $\sum_{i=1}^{5} \beta_i$ estimates (multiplied by 100) obtained from the *standard model* and the mean and standard deviation of the β and w_i estimates (again scaled by 100) obtained from the *weighted model*. In each case, the mean and standard deviations were based on 200 simulations. These results indicate that the *weighted model* is giving appropriate and interpretable results—the estimates of the effect of the air pollutant mixture on mortality are unbiased and the pollutants that are in the generated air pollutant mixture are, on average, receiving the largest weight. For example, in simulation number 6, where the air pollutant mixture is $\frac{1}{3}\text{PM} + \frac{1}{3}\text{NO}_2 + \frac{1}{3}\text{CO}$ and the effect of this air pollutant mixture is 3.00, the weighted model returns average weights of 33%, 30% and 34% for pollutants PM, NO_2 and CO, respectively, and an air pollutant mixture effect of 3.02 is returned. The *standard model* via summing up the individual pollutant effects $\sum_{i=1}^{5} \beta_i$ also returns unbiased estimates for the effect of the air pollutant mixture on mortality. However, the estimates of the effect of the air pollutant mixture on mortality obtained from the *weighted model* have a smaller estimation variance compared to the estimates obtained from the *standard model*.

In summary, the simulations have shown that the *weighted model* provides an unbiased estimate of the effect of the air pollutant mixture on the adverse health outcome of interest that has smaller estimation variance than the estimate obtained from the *standard model*, and is successful in estimating the mixture of pollutants that is associated with the adverse health outcome. This latter point is important because, although the *standard model* provides an unbiased estimate of the effect of the air pollutant mixture on mortality, it *does not* provide an estimate of the mixture of pollutants that is associated with the adverse health outcome.

4. Application

In this section, the data from the nine cities described above are used to illustrate the use of the *standard model* compared to that of the *weighted model*. As the goal of this study is to introduce the use of the *weighted model*, this section should be viewed as an illustration of the use of the *weighted model* compared to the *standard model*, rather than as explicit reanalysis of the data from each city. For both models, the confounder adjustments used had the same specification described in the previous section.

Table 2 contains the results of fitting the models to the data from each city. This table contains the individual pollutant effects obtained from the *standard model* and the weights assigned to each pollutant and the estimated effect of this air pollutant mixture obtained from the *weighted model*.

The *weighted model* produces more interpretable and parsimonious results than the *standard model* because the *weighted model* assigns weights to each air pollutant and pollutants attracting weight close to 0 can be ignored. For example, in Chicago the *standard model* suggests that increments in the concentration of PM, CO, O_2 and SO_2 will result

Table 1
Results of simulations comparing the *standard* and *weighted models*

	Total	*Weighted model* weights or *standard model* individual pollutant effects				
		PM	NO_2	CO	O_3	SO_2
Actual[a]	***0.00***	***20.00***	***20.00***	***20.00***	***20.00***	***20.00***
s[b]	0.00 (0.79)	0.02 (0.27)	0.01 (0.42)	0.00 (0.31)	0.00 (0.45)	−0.04 (0.29)
w[c]	0.00 (0.57)	18.26 (27.08)	19.28 (31.96)	17.82 (28.00)	22.84 (31.82)	21.80 (30.94)
Actual	***0.75***	***33.33***	***33.33***	***33.33***	***0.00***	***0.00***
s	0.70 (0.81)	0.24 (0.30)	0.23 (0.39)	0.24 (0.33)	−0.05 (0.46)	0.04 (0.31)
w	0.90 (0.36)	28.12 (27.18)	19.01 (27.69)	29.57 (27.33)	8.90 (18.35)	14.40 (21.86)
Actual	***1.25***	***20.00***	***20.00***	***20.00***	***20.00***	***20.00***
s	1.29 (0.79)	0.31 (0.27)	0.19 (0.45)	0.29 (0.33)	0.24 (0.41)	0.27 (0.28)
w	1.31 (0.36)	23.96 (20.79)	19.63 (23.94)	22.38 (18.16)	14.39 (18.70)	19.63 (17.58)
Actual	***1.50***	***33.33***	***33.33***	***33.33***	***0.00***	***0.00***
s	1.52 (0.74)	0.52 (0.29)	0.52 (0.36)	0.48 (0.30)	0.02 (0.40)	−0.01 (0.27)
w	1.63 (0.34)	30.37 (17.39)	27.50 (21.29)	31.62 (18.88)	5.72 (11.37)	4.80 (9.27)
Actual	***2.50***	***20.00***	***20.00***	***20.00***	***20.00***	***20.00***
s	2.57 (0.78)	0.53 (0.28)	0.48 (0.41)	0.50 (0.32)	0.55 (0.42)	0.52 (0.29)
w	2.51 (0.47)	21.4 (12.42)	20.50 (17.45)	20.49 (12.01)	17.14 (14.74)	20.47 (11.70)
Actual	***3.00***	***33.33***	***33.33***	***33.33***	***0.00***	***0.00***
s	2.93 (0.79)	1.00 (0.28)	1.02 (0.41)	0.97 (0.34)	−0.07 (0.43)	0.01 (0.27)
w	3.02 (0.33)	32.99 (9.42)	29.86 (16.63)	34.24 (13.22)	1.05 (4.53)	1.86 (5.09)
Actual	***5.00***	***20.00***	***20.00***	***20.00***	***20.00***	***20.00***
s	4.96 (0.82)	1.00 (0.27)	0.92 (0.41)	1.07 (0.34)	0.99 (0.45)	0.98 (0.32)
w	4.86 (0.56)	20.95 (6.48)	19.36 (10.90)	22.38 (7.40)	17.13 (10.58)	20.17 (6.89)
Actual	***6.00***	***33.33***	***33.33***	***33.33***	***0.00***	***0.00***
s	6.01 (0.80)	1.97 (0.28)	2.03 (0.41)	1.99 (0.32)	0.03 (0.45)	0.00 (0.31)
w	5.99 (0.31)	32.92 (4.69)	32.51 (9.23)	34.11 (7.29)	0.12 (1.22)	0.34 (1.96)
Actual	***10.00***	***20.00***	***20.00***	***20.00***	***20.00***	***20.00***
s	10.01 (0.81)	2.00 (0.29)	1.96 (0.40)	2.02 (0.33)	2.02 (0.46)	2.01 (0.30)
w	9.95 (0.58)	20.14 (3.12)	19.92 (4.74)	20.31 (3.42)	19.32 (5.35)	20.31 (2.96)

[a]100 times the actual values of θ and α_i that were used to generate mortality.
[b]100 times the mean (standard deviation) of the total and individual pollutant effect estimates obtained from the *standard model* over 200 simulations.
[c]100 times the mean (standard deviation) of the total air pollutant effect estimate and weight estimates obtained from the *weighted model* over 200 simulations.

in increased mortality, while increments in the concentration of NO_2 will result in reduced mortality. On the other hand, the *weighted model* predicts that the air pollutant mixture associated with increases in mortality is made up of approximately one part PM to three parts O_3, and that NO_2, CO and SO_2 are not associated with mortality. In this example, the *weighted model* provides three benefits over the *standard model*. Firstly, the *weighted model* avoids the implausible assertion that air pollutant can have a protective effect. Secondly, the *weighted model* assigns three air pollutants weights of 0, indicating that in Chicago these three pollutants are not associated with increases in mortality. Lastly, the *weighted model* provides an estimate of the air pollutant mixture that is associated with increases in mortality. These benefits of the *weighted model* can also be seen in the results for the other eight cities.

The results for the *weighted model* reveal interesting interpretations from the analysis about the effects of the five air pollutants on mortality in the nine cities considered. This insight is unavailable from the results of the *standard model*. In all cities except Houston the *weighted model* suggests that the pollutant mixture associated with daily mortality consists only of one or two pollutants. In both

Table 2
Results of fitting the *standard* and *weighted models* to the data from nine US cities for the period 1987–2000

Model	Total	*Weighted model* weights or *standard model* individual pollutant effects				
		PM	NO_2	CO	O_3	SO_2
Chicago, IL						
Standard[a]		0.59 (0.25)	−1.02 (0.30)	0.42 (0.23)	1.58 (0.43)	0.11 (0.23)
Weighted[b]	1.48 (0.47)	23.89 (26.35)	0.00 (11.72)	0.00 (9.21)	76.11 (30.65)	0.00 (7.59)
Cleveland, OH						
Standard		0.88 (0.62)	0.14 (0.69)	−0.35 (0.55)	0.35 (0.85)	−0.01 (0.54)
Weighted	0.89 (0.77)	100.00 (35.96)	0.00 (26.79)	0.00 (24.88)	0.00 (23.94)	0.00 (21.27)
Denver, CO						
Standard		−1.72 (0.75)	0.60 (0.91)	2.07 (1.00)	−1.10 (1.05)	0.85 (0.66)
Weighted	1.97 (1.05)	0.00 (17.16)	11.95 (22.82)	51.05 (29.37)	0.00 (18.38)	36.99 (27.41)
El Paso, TX						
Standard		−0.95 (0.87)	−0.67 (1.25)	0.77 (1.47)	−1.34 (0.89)	1.06 (0.73)
Weighted	−1.99 (1.24)	31.13 (27.19)	9.08 (21.70)	0.00 (21.98)	59.78 (27.73)	0.00 (17.23)
Houston, TX						
Standard		0.05 (0.42)	0.30 (0.80)	0.33 (0.57)	0.31 (0.69)	0.13 (0.50)
Weighted	1.11 (0.66)	4.84 (18.89)	25.10 (28.19)	29.60 (28.19)	29.34 (23.80)	11.12 (24.24)
Jersey City, NJ						
Standard		0.71 (0.71)	−0.76 (0.94)	−0.12 (0.90)	−0.47 (1.26)	0.95 (1.21)
Weighted	0.88 (1.49)	56.14 (35.64)	0.00 (23.87)	0.00 (27.91)	0.00 (28.95)	43.86 (26.12)
Nashville, TN						
Standard		−0.64 (0.86)	−0.07 (0.80)	−0.47 (0.78)	3.03 (1.40)	−1.38 (0.69)
Weighted	2.04 (1.24)	7.01 (16.56)	0.00 (15.04)	21.96 (19.04)	0.00 (18.96)	71.03 (23.50)
Pittsburgh, PA						
Standard		1.38 (0.62)	−0.01 (0.51)	−0.15 (0.49)	−1.04 (0.85)	−0.30 (0.43)
Weighted	0.88 (0.62)	100.00 (36.10)	0.00 (20.39)	0.00 (20.93)	0.00 (23.85)	0.00 (20.99)
Salt Lake City, UT						
Standard		0.04 (1.42)	−1.79 (2.70)	1.22 (2.62)	0.02 (1.80)	2.77 (1.52)
Weighted	3.61 (1.85)	0.00 (18.11)	0.00 (17.45)	0.00 (16.84)	12.60 (18.97)	87.40 (31.00)

[a] 100 times the mean (standard deviation) of the individual pollutant effect estimates obtained from the *standard model*.
[b] 100 times the mean (standard deviation) of the total air pollutant effect estimate and weight estimates obtained from the *weighted model*.

Cleveland and Pittsburgh the only pollutant that was associated with mortality was PM. In addition, the *weighted model* suggests that NO_2 and CO have little effect on daily mortality. Both these pollutants received weights of 0 in six of the nine cities, and otherwise received small weights.

5. Discussion

Two recent papers have also investigated the mixture of pollutants associated with adverse health outcomes. Roberts (2005) used an indirect method to assign each air pollutant a weight. The weights were obtained by constraining the air pollutant coefficients in the *standard model* (the β_i) to be non-negative. These weights were then scaled to sum to one, and the effect of the air pollutant mixture on the adverse health outcome was given by the sum of the air pollution coefficients $\sum_{i=1}^{k} \beta_i$ Constraining the parameters to be non-negative ensured that the corresponding weights were non-negative. However, an undesirable side effect of constraining the coefficients to be non-negative is that a positive bias is introduced into the estimate of the effect of the air pollutant mixture on the adverse health outcome. The *weighted model* has the advantage of directly estimating the air pollutant mixture associated with the adverse health outcome yet returns an unbiased estimate for the effect of the air pollutant mixture on the adverse health outcome. For these reasons the weighted model is an improvement over the model developed by Roberts. Hong et al. (1999) used a number of air pollution indices to evaluate the combined effects of various air pollutants. The indices used by Hong et al. were selected a priori and gave each pollutant included in the air pollutant index equal weight. This method will perform poorly if the actual air pollutant mixture associated with the adverse health outcome varies from an equal weighting. The approach proposed here avoids this problem, instead estimating the weights directly from the data.

The use of the *weighted model* in a given location may also give rise to other interesting questions that require further investigation such as why in some locations a pollutant attracted a large weight but in other locations it received an estimated weight of 0. For example, why in Cleveland and Pittsburgh was PM given a weight of 100% but in Denver, Houston and Nashville it was given a very small weight? One reason for this result could be that PM has a different chemical composition in Cleveland and Pittsburgh than in Denver, Houston and Nashville. The chemical composition of PM has already been investigated in a few air pollution mortality time series studies (Burnett et al., 2003; Smith et al., 2000a, b). The results from the *weighted model* could assist and inform such investigations, suggesting in this case that the chemical composition of PM in cities where PM is the dominant pollutant in the air pollutant mixture should be contrasted with the chemical composition of PM in cities where PM is a minor component of the air pollutant mixture.

Instead of investigating the unique effects of specific pollutants, it has been suggested that it might be more reasonable to assume that it is a mixture of pollutants that might be considered harmful to health (Dominici and Burnett, 2003; Moolgavkar, 2003; Stieb et al., 2002). For these reasons, the development of new models to concurrently estimate the adverse health effects of multiple air pollutants has been identified by statisticians, epidemiologists and policy makers as an important topic of research (Dominici and Burnett, 2003). The *weighted model* presented here, which provides for identification and estimation of a mixture of pollutants associated with adverse health effects, is an effective step in this direction.

References

Burnett, R.T., Brook, J., Dann, T., et al., 2003. Association between particulate and gas phase components of urban air pollution and daily mortality in eight Canadian cities. Inhalation Toxicology 12 (Suppl. 4), 15–39.

Chock, D.P., Winkler, S.L., Chen, C., 2000. A study of the association between daily mortality and ambient air pollutant concentrations in Pittsburgh, Pennsylvania. Journal of the Air & Waste Management Association 50, 1481–1500.

Cifuentes, L.A., Vega, J., Kopfer, K., Lave, L.B., 2000. Effect of the fine fraction of particulate matter versus the coarse mass and other pollutants on daily mortality in Santiago, Chile. Journal of the Air & Waste Management Association 50, 1287–1298.

Cox, L.H., 2000. Statistical issues in the study of air pollution involving airborne particulate matter. Environmetrics 11, 611–626.

Daniels, M.J., Dominici, F., Samet, J.M., Zeger, S.L., 2000. Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. American Journal of Epidemiology 152, 397–406.

Dominici, F., Burnett, R.T., 2003. Risk models for particulate air pollution. Journal of Toxicology and Environmental Health-Part A 66, 1883–1889.

Goldberg, M.S., Burnett, R.T., Valois, M.F., et al., 2003. Associations between ambient air pollution and daily mortality among persons with congestive heart failure. Environmental Research 91, 8–20.

Hong, Y.C., Leem, J.H., Ha, E.H., Christiani, D.C., 1999. PM_{10} exposure, gaseous pollutants, and daily mortality in Inchon, South Korea. Environmental Health Perspectives 107, 873–878.

Kelsall, J.E., Samet, J.M., Zeger, S.L., Xu, J., 1997. Air pollution and mortality in Philadelphia, 1974–1988. American Journal of Epidemiology 146, 750–762.

Klemm, R.J., Mason Jr., R.M., Heilig, C.M., Neas, L.M., Dockery, D.W., 2000. Is daily mortality associated specifically with fine particles? Data reconstruction and replication analyses. Journal of the Air & Waste Management Association 50, 1215–1222.

Kwon, H.J., Cho, S.H., Nyberg, F., Pershagen, G., 2001. Effects of ambient air pollution on daily mortality in a cohort of patients with congestive heart failure. Epidemiology 12, 413–419.

Moolgavkar, S.H., 2000. Air pollution and daily mortality in three US Counties. Environmental Health Perspectives 108, 777–784.

Moolgavkar, S.H., 2003. Air pollution and daily mortality in two US counties: season-specific analyses and exposure–response relationships. Inhalation Toxicology 15, 877–907.

Ostro, B.D., Hurley, S., Lipsett, M.J., 1999. Air pollution and daily mortality in the Coachella Valley, California: a study of PM10 dominated by coarse particles. Environmental Research 81, 231–238.

Roberts, S., 2005. A new model for investigating the mortality effects of multiple air pollutants in air pollution mortality time series studies, in press.

Roemer, W.H., van Wijnen, J.H., 2001. Daily mortality and air pollution along busy streets in Amsterdam, 1987–1998. Epidemiology 12, 649–653.

Smith, R.L., Davis, J.M., Sacks, J., Speckman, P., Styer, P., 2000a. Regression models for air pollution and daily mortality: analysis of data from Birmingham, Alabama. Environmetrics 11, 719–743.

Smith, R.L., Kim, Y., Fuentes, M., Spitzner, D., 2000b. Threshold dependence of mortality effects for fine and course particles in Phoenix, Arizona. Journal of the Air & Waste Management Association 50, 1367–1379.

Stieb, D.M., Judek, S., Burnett, R.T., 2002. Meta-analysis of time-series studies of air pollution and mortality: effects of gases and particles and the influence of cause of death, age, and season. Journal of the Air & Waste Management Association 52, 470–484.

Styer, P., McMillan, N., Gao, F., Davis, J., Sacks, J., 1995. Effect of outdoor airborne particulate matter on daily death counts. Environmental Health Perspectives 103, 490–497.

Vedal, S., Brauer, M., White, R., Petkau, J., 2003. Air pollution and daily mortality in a city with low levels of pollution. Environmental Health Perspectives 111, 45–51.

Wong, T.W., Tam, W.S., Yu, T.S., Wong, A.H., 2002. Associations between daily mortalities from respiratory and cardiovascular diseases and air pollution in Hong Kong, China. Occupational and Environmental Medicine 59, 30–35.