Estimating the acute health effects of coarse particulate matter accounting for exposure measurement error

HOWARD H. CHANG*

Department of Statistical Science, Duke University, Durham, NC 27708, USA
howard.chang@duke.edu

ROGER D. PENG

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health,
Baltimore, MD 21025, USA

FRANCESCA DOMINICI

Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA

SUMMARY
In air pollution epidemiology, there is a growing interest in estimating the health effects of coarse particulate matter (PM) with aerodynamic diameter between 2.5 and 10 μm. Coarse PM concentrations can exhibit considerable spatial heterogeneity because the particles travel shorter distances and do not remain suspended in the atmosphere for an extended period of time. In this paper, we develop a modeling approach for estimating the short-term effects of air pollution in time series analysis when the ambient concentrations vary spatially within the study region. Specifically, our approach quantifies the error in the exposure variable by characterizing, on any given day, the disagreement in ambient concentrations measured across monitoring stations. This is accomplished by viewing monitor-level measurements as error-prone repeated measurements of the unobserved population average exposure. Inference is carried out in a Bayesian framework to fully account for uncertainty in the estimation of model parameters. Finally, by using different exposure indicators, we investigate the sensitivity of the association between coarse PM and daily hospital admissions based on a recent national multisite time series analysis. Among Medicare enrollees from 59 US counties between the period 1999 and 2005, we find a consistent positive association between coarse PM and same-day admission for cardiovascular diseases.

Keywords: Air pollution; Coarse particulate matter; Exposure measurement error; Multisite time series analysis.

1. INTRODUCTION
Ambient particulate matter (PM) is a mixture of solid and liquid particles regulated by the Environmental Protection Agency (EPA) as one of the 6 criteria air pollutants. Under the Clean Air Act, EPA is responsible for establishing national standards for these pollutants to protect public health and the environment.

*To whom correspondence should be addressed.
(Bachmann, 2007). PM can be characterized into the fine and coarse size fractions that represent distinct pollutant mixtures of different sources and properties (Wilson and Suh, 1997). Particle size is an important attribute because it governs the particle’s behaviors in the atmosphere and deposition in the respiratory system. Ambient coarse PM is most often released directly as a primary pollutant through mechanical processes such as dust suspension or physical attrition involving grinding and crushing. Mineral residue resulting from combustion burn out (fly ash) also contributes to the coarse fraction. Biogenic compounds including bacterial endotoxin, pollen, and other animal/plant debris may also be present.

Protecting public health from coarse PM has endured considerable controversy in the regulatory context. EPA’s current National Ambient Air Quality Standards (NAAQS) use ambient PM$_{10}$ concentration (PM with aerodynamic diameter $<$10 $\mu$m) as the pollutant measure to protect public health from coarse PM. Most studies also routinely use PM$_{10}$ to quantify health risks and have consistently found that increased concentrations of ambient (outdoor) PM$_{10}$ are associated with increased risks of various adverse health outcomes (Pope and Dockery, 2006). However, there exists persistent criticism in interpreting the health effects of coarse PM since PM$_{10}$ contains both the coarse and fine fraction.

Recent studies of coarse PM have increasingly focused on exposure to inhalable coarse particles (PM$_{10-2.5}$) of size between 2.5 and 10 $\mu$m aerodynamic diameter. While toxicological evidence supports the potential health effects of coarse PM, current epidemiological evidence is limited and mixed (Brunekreef and Forsberg, 2005). Most time series analysis of ambient PM$_{10-2.5}$ concentrations and short-term mortality showed nonstatistically significant associations except in arid regions such as Mexico City (Castillejos and others, 2000) and Phoenix US (Mar and others, 2004). Results from both the Harvard Six Cities Study (Dockery and others, 1993) and the American Cancer Society cohort (Pope and others, 2002) also found no association between long-term exposure to coarse particles and mortality. However, studies have reported statistically significant short-term effects of ambient PM$_{10-2.5}$ on hospital admissions (Peng and others, 2008) and mortality (Zanobetti and Schwartz, 2009). Particularly, a multisite time series analysis conducted by Peng and others (2008) found that on average across the 108 US counties, ambient PM$_{10-2.5}$ concentration was associated with emergency admissions for cardiovascular diseases, but this association lost statistical significance when adjusted by PM$_{2.5}$.

In a time series design, the health outcome is only available as daily total number of adverse health events in a community, such as a county, a city or a large metropolitan area. Unbiased risk estimates require the exposure measure to coincide with the true average exposure experienced by all at-risk individuals in the community (Zeger and others, 2003; Sheppard, 2005). When the ambient pollutant concentration is spatially smooth, current practice of averaging measurements from outdoor monitors provides a reasonable surrogate measure for the population exposure due to outdoor sources. However, coarse PM concentrations often exhibit higher spatial heterogeneity compared to PM$_{2.5}$ and PM$_{10}$. Therefore, averaging PM$_{10-2.5}$ values from the fixed-location monitors placed in the same community may not capture the true population exposure. Moreover because there is no national monitoring network for PM$_{10-2.5}$, community-level daily PM$_{10-2.5}$ concentrations are calculated based on the limited network of collocated monitor pairs where both PM$_{10}$ and PM$_{2.5}$ are measured at the same location.

One of the main objectives of this paper is to develop a statistical modeling approach and computationally efficient estimation procedures for estimating the health effects of air pollution accounting for exposure measurement error (ME) in multisite time series analyses. We are concerned with the error that results from assigning an incorrect exposure measure to the study population living in an area when (1) pollution concentrations are available from a small number of monitors placed within the community; and (2) the pollution concentrations are highly variable within the community. To incorporate exposure ME in risk estimates, we view monitor-level PM values as error-prone repeated measurements of the true community-level average exposure. Our approach estimates exposure ME by quantifying, on any given day, the disagreement in PM values measured across the monitoring stations located within the same community. Specifically, we develop ME models for a bivariate vector of exposure variables in order
to estimate the effect of PM$_{10-2.5}$ adjusted by PM$_{2.5}$. Joint modeling also addresses the bias where the effect of one pollutant measured with more error is transferred to another pollutant measured with less error (Zidek and others, 1996). Finally, we calculate different measures of county-level ambient daily exposure to PM$_{10-2.5}$ and investigate the sensitivity of the national average effect of PM$_{10-2.5}$ on hospital admissions estimated in Peng and others (2008).

The exposure ME encountered in the analysis of PM$_{10-2.5}$ is related to the statistical problem known as spatial misalignment (Gotway and Young, 1999). Spatial variation in ambient concentrations and exposure ME caused by spatial misalignment have been addressed in several studies on the long-term health effects of air pollution (Zhu and others, 2003; Gyparis and others, 2009). However, few have examined its effects specifically in time series analysis. Through simulation studies, Sheppard and others (2005) find minor effect attenuation when the ambient concentration varies spatially. Peng and Bell (2010) calculate county-level exposure by first interpolating PM$_{2.5}$ chemical constituents concentration via spatial modeling. The authors show that the resulting county-specific short-term health effects are greater in magnitude and have larger standard errors compared to estimates that do not consider spatial variation in pollution concentration. To our knowledge, no study has investigated the effects of exposure ME in the analysis of PM$_{10-2.5}$ and health.

Our ME approach estimates daily county-level PM$_{10-2.5}$ exposure accounting for spatial heterogeneity without explicitly characterizing the spatial gradient. While this approach provides computational advantages, a space–time model for PM$_{10}$ and PM$_{2.5}$ offers an alternative approach for obtaining PM$_{10-2.5}$ exposure measure. However, this increases model complexity significantly when considering daily PM$_{10-2.5}$ concentrations nationally over a long study period. The ability to characterize PM$_{10-2.5}$ spatial variation is also limited by the sparse network of collocated PM$_{10}$ and PM$_{2.5}$ monitors.

The remainder of this paper is organized as follows. Section 2 describes the data sets, the modeling framework and the estimation procedures. In Section 3, we calculate different measures of daily county-level ambient exposure to PM$_{10-2.5}$ and investigate the sensitivity of the national average effect of PM$_{10-2.5}$ on hospital admissions. Section 4 describes a simulation study that examines the impact of PM$_{10-2.5}$ exposure ME and the performance of our proposed method. Finally, discussion and future work appear in Section 5.

2. Methods

2.1 Data

Daily average concentrations of PM$_{10-2.5}$ and PM$_{2.5}$ data for the period 1999–2005 were obtained from the EPA’s National Air Pollution Monitoring Network in the air quality system (AQS). Without a national monitoring network for PM$_{10-2.5}$, the EPA calculates PM$_{10-2.5}$ values indirectly by subtracting PM$_{2.5}$ from PM$_{10}$ measurements at monitors that are physically located at the same place (collocated monitor pairs). We considered PM$_{2.5}$ measurements from all AQS monitors and PM$_{10-2.5}$ measurements from collocated PM$_{2.5}$ and PM$_{10}$ monitor pairs. We restricted our analysis to the 59 US counties with (1) a population greater than 200,000 based on the 2000 census; (2) at least 2 pairs of collocated PM$_{10}$ and PM$_{2.5}$ monitors; and (3) at least 210 daily PM$_{10-2.5}$ measurements over the study period. Locations of the counties in our study are shown in Figure 1.

Time series of daily emergency hospital admissions for cardiovascular and respiratory diseases were assembled for Medicare enrollees aged 65 years or above within each county (Peng and others, 2008). Records were extracted from the National Claim History Files for the period between 1999 and 2005. Based on the International Classification of Diseases, Ninth Revision disease classification, we considered primary diagnosis of admissions due to 2 aggregated causes: (1) cardiovascular disease admissions and (2) respiratory disease admissions.
2.2 Trimmed mean exposure indicator

For county \( c \) on day \( t \), let \( W^c_{1tj} \) and \( W^c_{2tj} \) denote the PM\(_{10} \) and PM\(_{2.5} \) levels at monitor \( j \), where \( j = 1, \ldots, J^c \) and \( J^c \) is the total number of PM\(_{2.5} \) monitors in county \( c \). For some PM\(_{2.5} \) monitors, a collocated PM\(_{10} \) monitor is present and both \( W^c_{1tj} \) and \( W^c_{2tj} \) are observed. Otherwise, only \( W^c_{2tj} \) is observed, and we treat \( W^c_{1tj} \) as missing. In our study, there were 220 collocated PM\(_{2.5} \) and PM\(_{10} \) monitor pairs and an additional 173 PM\(_{2.5} \) monitors that did not have collocated PM\(_{10} \) monitors.

Under the assumption that PM\(_{10} \) and PM\(_{2.5} \) concentrations are spatially homogeneous across each county, the standard daily county-level PM\(_{10} \) exposure \( (X^c_{1t}) \) and PM\(_{2.5} \) exposure \( (X^c_{2t}) \) measures are obtained by averaging all available measurements across monitors on a particular day (Samet et al., 2000). Specifically, the 10% trimmed mean (TM) is used to exclude extreme monitor-level values that may be invalid. In this algorithm, no PM measurements are excluded in computing the mean on days with less than 3 monitor-level PM measurements and only the maximum and minimum measurements are excluded on days between 3 and 9 PM measurements.

2.3 ME model

To illustrate the impact of a spatially heterogeneous pollutant in time series studies of air pollution and health and the importance of defining appropriate population exposure indicators, first consider the following example adopted from Zeger and others (2003) and Sheppard (2005). Denote by \( X_{it} \) the exposure to a pollutant due to outdoor sources for an individual \( i \) on day \( t \). For each individual in the study, we assume that the binary health outcome follows a Bernoulli distribution with probability \( \lambda_0 \exp(\beta X_{it}) \), where \( \lambda_0 \) represents a baseline risk common across individuals and \( \exp(\beta) \) represents the multiplicative change in risk associated with a unit increase in \( X_{it} \).

Let \( N_t \) denote the number of at-risk individuals in a community on day \( t \), and let \( Y_t \) denote the number of hospital admissions on day \( t \) in the population that resides in the community. If the occurrence of outcome is independent across individuals, the community-level outcome count has mean equal to

\[
\mu_t = \lambda_0 \sum_{i=1}^{N_t} \exp(\beta X_{it}). \quad (2.1)
\]
Since the short-term relative risk of air pollution is typically very small, following Zeget and others (2003), a linear approximation to the exponential term in (2.1) gives

$$\log(\mu_t) \simeq \log(\lambda_0 N_t) + \beta \left( \frac{\sum_{i=1}^{N_t} (X_{it} - X_t)}{N_t} \right) + \beta X_t = \log(\lambda_0 N_t) + \beta X_t,$$

where $X_t$ is the average exposure of all at-risk individuals. Under these assumptions, the relative risk obtained by regressing aggregated outcome $Y_t$ on aggregated exposure $X_t$ via log-linear regression is equivalent to the personal risk $\beta$. Therefore, $X_t$ represents the “desired but unobserved” exposure indicator in time series analysis.

Let $X^*_t$ be a surrogate exposure indicator calculated, for example, by averaging measurements from some fixed-location monitoring sites within the community. We can rewrite (2.1) as

$$\mu_t = \lambda_0 \sum_{i=1}^{N_t} \exp(\beta(X_{it} - X_t) + \beta(X_t - X^*_t) + \beta X^*_t). \quad (2.2)$$

Applying a similar linear approximation gives

$$\log(\mu_t) \simeq \log(\lambda_0 N_t) + \beta(X_t - X^*_t) + \beta X^*_t. \quad (2.3)$$

Hence, we wish to have the exposure indicator calculated from monitoring data to coincide with the true average exposure ($X_t = X^*_t$).

For coarse PM, ME can occur when the sparse PM monitoring network does not fully capture the spatial variation in PM$_{10-2.5}$ concentrations. Here, it is also important to account for the spatial distribution of the at-risk population since individuals can be exposed to different levels of ambient PM$_{10-2.5}$. The standard approach of defining community-level exposure by averaging monitor measurements can over- or underestimate the true average exposure. This bias can also vary temporally because the set of PM$_{10-2.5}$ measurements used to calculate average exposure often varies between days due to missing data or different monitoring schedules.

In the ME modeling approach, we view PM measurements from different monitors on the same day within the same county as error-prone repeated measurements of the unobserved population exposure. Let $X^*_t = (X^*_{1t}, X^*_{2t})$ denote the unobserved county-level exposure to ambient PM$_{10-2.5}$ and PM$_{2.5}$ experienced by the at-risk population. Under the classical additive ME model (Fuller, 1987), we assume for $j = 1, \ldots, J^c$,

$$W^c_{ij} = \begin{bmatrix} W^c_{1ij} \\ W^c_{2ij} \end{bmatrix} \sim \text{Normal}\left( \begin{bmatrix} X^c_{1t} \\ X^c_{2t} \end{bmatrix}, \Sigma^c_{w0} = \begin{bmatrix} \sigma^2 \nu^c & \rho^c \sigma_1^c \sigma_2^c \\ \rho^c \sigma_1^c \sigma_2^c & \sigma^2 \nu^c \end{bmatrix} \right), \quad (2.4)$$

where the 2 ME variances ($\sigma^2 \nu^c$ and $\sigma^2 \nu^c$) capture same-day between-monitor variability for PM$_{10-2.5}$ and PM$_{2.5}$. Parameter $\rho^c$ captures the correlation between MEs, which is assumed to be constant across days and monitors within the same county. A positive $\rho^c$ indicates that daily monitor-level PM$_{10-2.5}$ and PM$_{2.5}$ measurement pairs tend to deviate from the true exposures in the same direction. Note that at some monitoring locations and on some days, only $W^c_{2ij}$ is observed and $W^c_{1ij}$ is treated as missing. Through Bayesian inference described in Section 2.5, we simultaneously address the ME problem and the imputation of missing monitor-level PM$_{10-2.5}$ measurements.

Additionally, we assume

$$\log X^c_t = [\log X^c_{1t}, \log X^c_{2t}]^t \sim \text{Normal}( \begin{bmatrix} Z^c_t \eta^1_t, Z^c_t \eta^2_t \end{bmatrix}, \Sigma^c_x ), \quad (2.5)$$
where $Z^c_t$ denotes a covariate vector that includes indicators for month, indicators for day of the week, and calendar dates. Parameters $\eta_1^c$ and $\eta_2^c$ denote the corresponding vectors of regression coefficients. The logarithmic transformation accounts for the strictly positive and right-skewed PM concentration measurements. The diagonal elements of $\Sigma^c_x$ represent the residual variance for log-PM exposure and its off-diagonal element captures the correlation between daily county-level PM$_{10-2.5}$ and PM$_{2.5}$ exposure for county $c$. Note that we choose to model $W^c_{tj}$ in (2.4) without the log-transformation to allow for negative values of observed PM$_{10-2.5}$ concentrations. Also, $X^c_t$ is then interpreted as an average of the observed monitoring data. Alternatively, $W^c_{tj}$ can be modeled on the log-scale where a multiplicative ME is assumed.

However, the error specification in (2.4) assumes that errors between the observed monitor-level PM$_{10-2.5}$ concentrations and the true county-level exposure are distributed identically with the same error variance across monitors. For a spatially heterogeneous PM, true county-level exposure represents a population-weighted average exposure. Hence, measurements taken at monitors located in less populated areas, relative to the total population of the county, may tend to deviate more from the true average exposure experienced by everyone in the county.

We also consider a ME model with heteroskedastic errors between monitors where the ME variance is weighted inversely to the population living around the monitor. Let $P_k^c$ denote the population of city $k$ in county $c$. Following notations from (2.4), we now assume

$$[W^c_{1tj}, W^c_{2tj}]' \sim \text{Normal}\left([X^c_t, X^c_{t+1}]', \frac{1}{\alpha_j^c} \Sigma^c_w\right),$$

(2.6)

where $\alpha_j^c = P_k^c / \sum_k P_k^c$ if monitor $j$ is in city $k^*$. For counties where PM is measured in multiple cities, $\alpha_j^c$ approximates the proportion of at-risk population in county $c$ that are exposed to the PM concentration measured at monitor $j$. City population data were retrieved from the Site Descriptive Data database of EPA's AQS. Under this ME model, monitors in less populated areas contribute less to the estimation of county-level exposure. We refer to the above model as the population-weighted WME approach.

Finally, we assign the following priors for the model parameters: (1) $\eta_1^c$ and $\eta_2^c$ each follows a multivariate Normal distribution with dispersed variances; (2) $\Sigma^c_x$ follows Inv–Wish($v_x, \hat{\Sigma}^c_x$); and (3) $\Sigma^c_w$ follows Inv–Wish($v_w, \hat{\Sigma}^c_w$), where $v_w = 3$ and $\hat{\Sigma}^c_w$ is the estimated covariance of $W^c_{tj} - \hat{W}^c_t$; and (3) $\Sigma^c_x$ follows Inv–Wish($v_x, \hat{\Sigma}^c_x$), where $v_x = 3$ and $\hat{\Sigma}^c_x$ is the estimated residual covariance from regressing log $X^c_t$ on $Z^c_t$. Following Gelman (2006), we also consider assigning a Uniform[0, 100] on the standard deviations (SDs) and a Uniform[−1,1] on the correlation for either $\Sigma^c_w$ or $\Sigma^c_x$, or both. The differences in posterior inference for selected counties with small numbers of days or monitors were negligible.

### 2.4 Health model

We model the expected number of admissions $E(Y^c_t)$ using Poisson regression:

$$\log E(Y^c_t) = \log N^c_t + \beta_0^c + \beta_1^c X^c_{1t} + \beta_2^c X^c_{2t} + \psi^c C^c_t,$$

(2.7)

where $N^c_t$ is the size of the population at risk. Following Dominici and others (2006), confounders ($C^c_t$) include seasonal trends, weather effects, and age-group effects that are modeled via natural cubic splines with degrees of freedom $d$. Specifically, we include: (1) calendar time ($d = 8$ per year), (2) current-day temperature ($d = 6$) and average temperature for the previous 3 days ($d = 6$); (3) current-day dew-point temperature ($d = 3$) and average dew-point temperature for the previous 3 days ($d = 3$); (4) age-group intercept (64–74 vs. 75 and above) and its interaction with a smooth function of calendar time ($d = 1$ per year); and (5) day of the week.
2.5 Risk estimation

For the standard time series analysis without ME modeling, county-level exposures \(X^c\) are obtained using the 10% TM algorithm described in Section 2.2. Relative risks estimates are obtained by fitting the health model via maximum quasi-likelihood separately for each county.

Due to the complex health model and the large data set, we describe a 2-stage approach that involves 2 separate Markov chain Monte Carlo (MCMC) implementations to estimate relative risks with ME modeling. In the first stage, posterior samples of \(PM_{10-2.5}\) and \(PM_{2.5}\) time series \(X^c\) given the observed monitor-level data \(W^c\) are obtained by sampling from the following posterior predictive distribution:

\[
\text{Stage 1: } [X^c|W^c, Z^c] \propto \int [W^c|X^c, \theta_1^c][X^c|Z^c, \theta_2^c] \pi(\theta_1^c, \theta_2^c) d\theta_1^c d\theta_2^c, \tag{2.8}
\]

where \([W^c|X^c, \theta_1^c]\) represents the “measurement model” given by (2.4) with parameters \(\theta_1^c = (\Sigma_w)\) and \([X^c|Z^c, \theta_2^c]\) represents the “exposure model” given by (2.5) with parameters \(\theta_2 = (\eta_1^c, \eta_2^c, \Sigma_z^c)\). Here, the posterior distribution of \(X^c\) does not depend on the health model. Stage 1 computation is carried out by using JAGS version 1.0.3 (Plummer, 2003).

At the second stage, we obtain posterior samples of \(X^c, \beta^c|W^c, Z^c, Y^c\) by using \([X^c|W^c, Z^c]\) from Stage 1 as the prior distribution of \(X^c\). Given health data \(Y^c\) for county \(c\), we assume

\[
\text{Stage 2: } [X^c, \beta^c|W^c, Z^c, Y^c] \propto [Y^c|X^c, \beta^c, \psi^c][X^c|W^c, Z^c] \pi(\beta^c, \psi^c) d\psi^c, \tag{2.9}
\]

where \([Y^c|X^c, \beta^c, \psi^c]\) corresponds to the “health model” from (2.7). To decrease computational burden, we treat \(\psi^c\) as nuisance parameters and carry out a profile sampler approach described in Lee and others (2005). Specifically, we carry out block Metropolis–Hastings updates between \(\beta^c\) and \(X^c\), where the acceptance probabilities are calculated using the profile likelihood. Since both \(X^c\) and \(\beta^c\) are updated by the health data \(Y^c\), we refer to this estimation approach as the “Bayesian” approach. This approach also provides samples of \([X^c|W^c, Z^c, Y^c]\), the posterior distribution of the average PM exposure incorporating the health information. Details of the estimation procedure and an example of validating the 2-stage Bayesian approach are provided in Section 1 of the supplementary material available at Biostatistics online.

We also estimate the county-specific relative risks \(\beta^c\), where we replace \(X^c\) in the health model by the marginal posterior mean \(E[X^c|W^c, Z^c]\) from stage 1. This plug-in method resembles a regression calibration where the unobservable exposure is replaced by its best linear prediction conditional on the covariates measured without error \((Z^c_t)\) and the observed error-prone measurements (Carroll and others, 2006). While this method is computationally simple, it does not fully reflect the uncertainty in the exposure measure when estimating \(\beta^c\). Also, unlike the Bayesian approach, there is no feedback between the health observations and the exposure estimates. However, since the acute health effects of PM are typically small (relative risk of less than 2% per 10 \(\mu g/m^3\) increase in PM), the information in the health model is possibly negligible in determining the posterior distribution of the true county-level PM exposure.

We pool county-specific relative risks \(\hat{\beta}^c = (\hat{\beta}_{1}^c, \hat{\beta}_{2}^c)\) defined in (2.7) by assuming \(\beta^c \sim N(\mu, \Sigma_\beta)\). Here, the parameter of interest, \(\mu\), is interpreted as the pooled (national) \(PM_{10-2.5}\) and \(PM_{2.5}\) effects and \(\Sigma_\beta\) captures the heterogeneity in relative risks between counties. Denote the estimated county-specific relative risks by \(\hat{\beta}^c\) and the corresponding covariance matrix by \(\hat{V}^c\) estimated either with or without exposure ME modeling as described in Section 2.5. For the Bayesian approach with ME modeling, we define \(\hat{\beta}^c\) and \(\hat{V}^c\) as the posterior mean and posterior covariance from the second-stage MCMC. Assuming \(\hat{\beta}^c \sim N(\beta^c, \hat{V}^c)\), \(\mu\) and \(\Sigma_\beta\) are estimated using the 2-level Normal independent sampling estimation algorithm of Everson and Morris (2000).
3. Results

Across counties, the median number of PM$_{2.5}$ monitors in a county was 5; the first quartile (Q$_1$, 25th percentile) was 4; and the third quartile (Q$_3$, 75th percentile) was 8. Similarly, the median number of PM$_{2.5}$ and PM$_{10}$ collocated monitor pairs was 3 (Q$_1 = 2$ and Q$_3 = 5$). Therefore, by restricting our analysis to collocated monitor pairs, the standard TM exposure for PM$_{10-2.5}$ was calculated based on a considerably smaller number of monitors compared to PM$_{2.5}$. In Figure 2, correlation of daily PM measurements at any pair of monitors in the same county is plotted versus the distance between the monitors. There was considerable larger spatial variability in PM$_{10-2.5}$ measurements compared to PM$_{2.5}$.

Our study included approximately 5 million Medicare enrollees between the period 1999 and 2005. There were about 2.6 million admissions for cardiovascular diseases and 1.0 million admissions for respiratory diseases. Across counties, the median daily admission for cardiovascular diseases was 18.7 per 100,000 people (Q$_1 = 15.8$ and Q$_3 = 21.4$) and the median for respiratory diseases was 7.4 per 100,000 people (Q$_1 = 6.3$ and Q$_3 = 8.8$).

We considered 5 exposure measures of daily county-average PM$_{10-2.5}$ level. For example, Figure 3 shows the marginal posterior distributions of PM$_{10-2.5}$ exposure on July 17, 2000 in Harris County, TX. The 4 posterior distributions were obtained under different ME modeling and estimation approaches: (1) constant ME variances across monitors without using the health data (ME, [X|W, Z]); (2) constant ME variances across monitors using the health data (ME, [X|W, Z, Y]); (3) population-WME variances across monitors without using the health data (WME, [X|W, Z]); and (4) population-WME variances across monitors using the health data (WME, [X|W, Z, Y]). For (2) and (4), the relative risks associated with cardiovascular admissions were simultaneously estimated with PM$_{10-2.5}$ exposure. Also, a vertical line is placed at the 10% TM estimate. The differences in exposure estimates reflect which monitor-level observations were used. On this particular day, there were 4 observations of PM$_{10-2.5}$ concentration: 3 observations 18, 61, and 20 were from Houston and one observation 12 was from Deer Park. The TM PM$_{10-2.5}$ measure excluded 12 and 61 in computing the average; the ME measure considered all values.

Fig. 2. Correlations of monitor-level daily PM time series calculated between pairs of PM$_{2.5}$ or PM$_{10-2.5}$ monitoring locations in the same county and plotted versus the distance between monitor pair. For PM$_{2.5}$, we used all available monitors without restricting to those with a collocated PM$_{10}$ monitor.
Measurement error modeling for PM$_{10-2.5}$ exposure

Fig. 3. Posterior distributions of the average exposure to outdoor PM$_{10-2.5}$ concentration on July 17, 2000 in Harris County, TX. The vertical line is placed at the 10% TM estimate. The solid and dotted lines represent 4 different PM$_{10-2.5}$ posterior distributions obtained from ME modeling.

Table 1. Quantiles of county-specific SD of PM$_{2.5}$ and PM$_{10-2.5}$ time series using either TM, ME modeling with constant error variance (ME), or monitor-specific weighted error variance (WME). The median (25th quantile and 75th quantile) of the SD across 59 counties are given

<table>
<thead>
<tr>
<th></th>
<th>TM</th>
<th>ME</th>
<th>WME</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM$_{2.5}$ ($\mu g/ m^3$)</td>
<td>7.6 (6.1, 9.0)</td>
<td>7.1 (5.8, 8.7)</td>
<td>7.3 (5.9, 8.6)</td>
</tr>
<tr>
<td>PM$_{10-2.5}$ ($\mu g/ m^3$)</td>
<td>7.9 (6.2, 12.2)</td>
<td>6.0 (4.4, 10.3)</td>
<td>7.0 (4.8, 11.7)</td>
</tr>
</tbody>
</table>

equally; and the WME measure down weighted the measurement from Deer Park which has a considerably smaller population than Houston.

We calculated the SD of county-level PM$_{2.5}$ and PM$_{10-2.5}$ levels across days and Table 1 gives the median, Q$_1$, and Q$_3$ across 59 counties for different exposure measures. First, daily variation of county-average PM$_{10-2.5}$ levels derived from ME and WME were lower compared to TM and this decrease was less significant for PM$_{2.5}$. PM daily variation decreases when ME is considered because the model assumes that the observed PM concentrations are more noisy than the true exposure. Specifically, a large decrease in time series SD reflects greater disagreement between same-day monitor-level measurements. Moreover, for both PM$_{10-2.5}$ and PM$_{2.5}$, the decrease in daily variation was more significant for the ME measures compared to the WME measures. If PM levels vary across cities of different population sizes, county-average exposure is determined mainly by measurements in cities with large populations. Since the true exposure represents a population-weighted average exposure, the WME approach can result in smaller ME because disagreement between PM measurements in cities with small populations and the true exposure is down-weighted.

Table 2 gives some pairwise correlations between PM$_{10-2.5}$ and PM$_{2.5}$ exposure measures obtained using either TM or WME across counties. Comparing rows 1 and 2, higher correlations are observed
Table 2. Quantiles of correlations between different measures of county-level daily PM\(_{10-2.5}\) and PM\(_{2.5}\) exposures across 59 counties. Exposure measures for PM\(_{2.5}\) and PM\(_{10-2.5}\) are derived using either TM or WME modeling

<table>
<thead>
<tr>
<th>Correlation between</th>
<th>Minimum</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) PM(<em>{10-2.5},)TM \cdot PM(</em>{10-2.5},)WME</td>
<td>0.50</td>
<td>0.89</td>
<td>0.92</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>(2) PM(<em>{2.5},)TM \cdot PM(</em>{2.5},)WME</td>
<td>0.72</td>
<td>0.94</td>
<td>0.97</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>(3) PM(<em>{2.5},)TM \cdot PM(</em>{10-2.5},)TM</td>
<td>-0.20</td>
<td>0.05</td>
<td>0.12</td>
<td>0.23</td>
<td>0.59</td>
</tr>
<tr>
<td>(4) PM(<em>{2.5},)WME \cdot PM(</em>{10-2.5},)WME</td>
<td>-0.15</td>
<td>0.05</td>
<td>0.18</td>
<td>0.32</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Fig. 4. County-specific ME SD (\(\sigma_1^{2,c}\) and \(\sigma_2^{2,c}\) in (2.4) for PM\(_{10-2.5}\) (black) and PM\(_{2.5}\) (gray) plotted versus log county land area (cubic kilometer) for 59 counties. Each bullet denotes the posterior mean and the vertical line indicates the 95% posterior interval.

between different PM\(_{2.5}\) measures compared to PM\(_{10-2.5}\). This is expected since PM\(_{2.5}\) level is less heterogeneous spatially and the ME approach results in less calibration when the between-monitor agreement is strong. Comparing rows 3 and 4, we find that deriving PM\(_{10-2.5}\) and PM\(_{2.5}\) exposures via ME modeling increases the correlation between the 2 pollutants slightly. We also found very high correlation between the average exposure measures derived from the 2 ME models (ME vs. WME) for PM\(_{10-2.5}\) and PM\(_{2.5}\), having minimum correlation of 0.82 and 0.87, respectively, in the 59 counties (not shown in table).

The ME variances, \(\sigma_1^{2,c}\) and \(\sigma_2^{2,c}\) in (2.4) quantify the variability across monitors of the PM values. Figure 4 plots the posterior mean and 95% intervals for the ME SD versus log-transformed county land area (square kilometer). We found greater between-monitor variation for PM\(_{10-2.5}\) (black) measurements compared to PM\(_{2.5}\) (gray) measurements, even though the 2 pollutants had similar average concentration over the study period. The median ME SD across counties for PM\(_{10-2.5}\) is 5.6 (Q\(_1 = 4.4\) and Q\(_3 = 8.8\)) and for PM\(_{2.5}\) is 2.3 (Q\(_1 = 1.7\) and Q\(_3 = 3.2\)). In Figure 4, it also appears that larger counties were associated with greater between-monitor variation in PM measurements. We also found evidence...
Measurement error modeling for PM$_{10-2.5}$ exposure

Fig. 5. Upper panels: scatter plot of county-specific standardized health effect estimates for PM$_{10-2.5}$, $\hat{\beta}_c^1 / \text{SE}(\hat{\beta}_c^1)$, comparing two approaches: (1) including ME modeling with monitor-specific weighted error variance (WME) versus (2) using TM as PM$_{10-2.5}$ exposure. Lower panels: scatter plot of county-specific standardized health-effect estimates for PM$_{10-2.5}$, $\hat{\beta}_c^1 / \text{SE}(\hat{\beta}_c^1)$, using exposure derived from WME comparing Bayesian risk estimation versus regression calibration.

of a weak positive association between PM$_{2.5}$ and PM$_{10-2.5}$ measurement errors at collocated monitors for some counties. The posterior means of $\rho^c$ across 59 counties have a median of 0.10 (min = -0.23, Q$_1$ = -0.1, Q$_3$ = 0.4, max = 0.6).

The 2 upper panels in Figure 5 plot the county-specific standardized coefficients, $\hat{\beta}_c^1 / \text{SE}(\hat{\beta}_c^1)$ to examine the strength and direction of the health effect of PM$_{10-2.5}$ on cardiovascular and respiratory admissions estimated using different exposure measures. Comparing estimates derived from standard TM exposure and WME with the Bayesian risk estimation, we did not observe large changes in the health effects’ direction. However, there is attenuation for large $\hat{\beta}_c^1 / \text{SE}(\hat{\beta}_c^1)$ possibly due to increased uncertainty in risk.
estimates when MEs are accounted for. From the 2 bottom panels in Figure 5, we show that in our application, standard error (SE), $\hat{\beta}_1 / SE(\hat{\beta}_1)$ for cardiovascular and respiratory admissions are very similar between those derived from regression calibration and those estimated through the Bayesian approach. For the Poisson health model, regression calibration will result in some bias in the relative risk estimates; however in the analysis of PM$_{10-2.5}$, the uncertainty in exposure appears to dominate.

Figure 6 gives the pooled estimates of percent increase in cardiovascular and respiratory disease admissions per 10 μg/m$^3$ increase in same-day particulate matter concentration. Exposure measures for PM$_{2.5}$ and PM$_{10-2.5}$ were derived using either TM, ME, or WME, and we considered both regression calibration and Bayesian risk estimations. The original Peng and others (2008) estimates based on 108 counties using TM exposure are also shown. We found consistent positive effects for PM$_{10-2.5}$ and PM$_{2.5}$ with different exposure measures and estimation procedures. For cardiovascular admissions, effects of PM$_{2.5}$ remain statistically significant under different scenarios. The posterior intervals are wider under ME modeling compared to using the standard TM exposure. Also, when ME modeling are used, the confidence intervals are wider for Bayesian risk estimations compared to regression calibration and the bias associated with regression calibration appears negligible.

4. SIMULATION STUDY

This section describes a simulation study that examines the impact of PM$_{10-2.5}$ exposure ME and the performance of our proposed method using data from Clark County, NV. Clark County contains 8 PM$_{10-2.5}$ monitoring locations from 5 cities with population ranging from about 200 to half a million. On each day $t$, we do not observe the complete vector of monitor-level measurements $X_t = (x_{t1}, x_{t2}, \ldots, x_{t8})'$. Across the 1337 days with at least one PM$_{10-2.5}$ measurement, the average number of measurements per day was 4.3.

We generated 100 replicate data sets of the complete monitor-level PM$_{10-2.5}$ values as follows. We assumed $X_t \sim \text{Normal}(\hat{B}Z_t, \hat{\Sigma})$, where $Z_t$ is the $p \times 1$ vector of covariates from (2.5). Parameter $\hat{B}$ is an 8 × $p$ matrix of monitor-specific regression coefficients and parameter $\hat{\Sigma}$ is the 8 × 8 residual variance-covariance matrix.
Table 3. Simulation study results: median bias, average 95% confidence interval (C.I.) width, and root mean squared error for PM$_{10-2.5}$ relative risk in Clark County, Nevada. At-risk population is allocated to each monitor either equally (Scenario 1) or proportional to the city population that the monitor is in (Scenario 2). Exposure measures considered include: TM or ME modeling without (ME) or with (WME) heteroskedastic variances.

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>TM</td>
</tr>
<tr>
<td>Bias ($\times 10^4$)</td>
<td>-0.02</td>
</tr>
<tr>
<td>C.I. width ($\times 10^4$)</td>
<td>1.62</td>
</tr>
</tbody>
</table>

covariance matrix. These parameters were estimated from the observed data and the between-monitor correlation ranges from 0.3 to 0.8. Given the observed PM$_{10-2.5}$ concentrations, missing concentrations were imputed using the corresponding conditional mean and covariance. We used the complete PM data set to generate hospital admissions. However, in constructing the TM exposure measures and in carrying out ME modeling, we followed the observed missing data structure and ignored the imputed PM$_{10-2.5}$ concentrations.

We set the total number of at-risk individuals to be 43,410. We considered 2 exposure scenarios by allocating the population to each monitor either (1) equally or (2) proportional to the city population that the monitor is in. Finally, for each simulated PM$_{10-2.5}$ data set, we generated daily total admission with a baseline risk of $1.8 \times 10^{-4}$ and a relative risk of 0.5% per 10 unit increase PM$_{10-2.5}$ concentration following the Poisson model in (2.7) without additional confounders. Relative risk for PM$_{10-2.5}$ was then estimated using the TM, ME, or WME exposure measures. Here, we only considered the estimation approach without using the health data $(|X|, W, Z)$ due to computational limitation.

Table 3 gives the median bias and average 95% confidence interval length for the relative risk estimates. With the TM exposure measures, we found that attenuation occurs in both scenarios and our ME approaches (ME and WME) reduce this bias. ME modeling increases the confidence interval width when the same number of individuals are exposed to different concentrations (scenario 1). In this case, our method effectively propagates the uncertainty in population exposure when between-monitor disagreement is present. However, when the number of at-risk individuals varies across PM$_{10-2.5}$ concentrations (scenario 2), the WME exposures does not result in wider confidence intervals. This is likely due to recovering some exposure variability that is oversmoothed by simply averaging monitor-level concentrations.

**5. Discussion**

EPA does not regulate PM$_{10-2.5}$ directly but continues to use PM$_{10}$ as a surrogate to protect public health. In the most recent 2006 NAAQS revision for particulate matter, a 24-h PM$_{10-2.5}$ standard was proposed but ultimately not accepted due to insufficient evidence linking short-term PM$_{10-2.5}$ exposure and adverse health outcomes (Environmental Protection Agency, 2006). Time series analysis plays an important role in providing epidemiological evidence for the acute health effects of PM and in establishing regulatory standards (Greenbaum and others, 2001). Its popularity is due to the ability to utilize public databases to estimate the relatively small acute effects with large study populations. However, recent interest in quantifying the health effects of PM$_{10-2.5}$ raises statistical questions regarding the time series design when the pollutant concentration varies spatially.

In this paper, we address the challenge of exposure ME due to spatial misalignment through ME modeling. The goal is to obtain risk estimates that reflect the uncertainty in PM$_{10-2.5}$ exposure in a time series study. This differs from the past work that has focused predominantly on errors due to either (1)
the discrepancy between ambient levels measured outdoors versus total personal exposure (Dominici and others, 2000) or (2) the ecological bias that results from using aggregated outcome and exposure to infer individual-level risk (Sheppard, 2005; Sheppard and others, 2005).

Computing average community-level exposure with monitor-specific weights is a common practice in time series analysis. The TM approach represents a simple way to remove extreme values observed on a particularly day; however, this may oversmooth a spatially varying exposure when the number of measurements per day is small. A similar approach taken by Zanobetti and Schwartz (2009) first removes monitors that are not well correlated with others in the same region to avoid measurements that are due to local pollution sources not reflective of the overall population exposure. However, this approach does not address the scenario when all monitors are poorly correlated with each other. In contrast, our ME method is a parametric approach that provides average exposure estimates and accounts for the same-day-monitor variability using all available data. Moreover, by allowing error variances to be inversely proportional to the population living around each monitor, we automatically specify monitor-specific weights in computing average exposure.

While this paper is motivated by the analysis of PM$_{10-2.5}$ and health, we note that the analysis of PM$_{2.5}$ chemical constituents shares similar challenges in exposure ME. For example, the metal constituents in PM$_{2.5}$ can exhibit high spatial heterogeneity and the minor components are often measured with high instrumental ME. One limitation of the proposed approach is that we need to restrict our analysis to counties with at least 2 pairs of collocated monitors. Future work will borrow information across counties by building regression models for the ME variances to predict the extent of exposure ME for counties that only have a single PM$_{10-2.5}$ measurement each day. Our model also assumes the MEs to be independent between days and does not model the temporal correlation between pollutant concentrations. This is because PM$_{10-2.5}$ measurements are typically only available every sixth day. But at some locations where daily measurements of PM$_{10-2.5}$ are available, additional modeling of these temporal trends should be explored.

SUPPLEMENTARY MATERIAL

Supplementary material is available at http://biostatistics.oxfordjournals.org.

ACKNOWLEDGMENTS

We are grateful to the editor, associate editor, and one anonymous referee for their useful suggestions for improvements. Conflict of Interest: None declared.

FUNDING

United States Environmental Protection Agency (R83622, RD-83241701); National Institute for Environmental Health Sciences (ES012054-03); National Institute for Environmental Health Sciences Center in Urban Environmental Health (P30 ES 03819). It has not been subjected to EPA’s required peer and policy review and therefore does not reflect the views of the Agency and no official endorsement should be inferred.

REFERENCES

Measurement error modeling for PM$_{10-2.5}$ exposure


