



ADDITIONAL MATERIALS AVAILABLE ON THE HEI WEB SITE

Research Report 193

**Particulate Air Pollutants, Brain Structure, and
Neurocognitive Disorders in Older Women**

Chen et al.

Additional Materials 1

Part A: Study Population, Design, and Outcome Ascertainment

Part B: Estimation of Residential Exposures

Part C: Statistical Analyses

These Additional Materials were not formatted or edited by HEI. This document was part of the HEI Review Committee's review process.

Correspondence may be addressed to Dr. Jiu-Chiuan Chen, Department of Preventive Medicine, University of Southern California, Keck School of Medicine, 2001 N. Soto Street, MC 9237, Los Angeles, CA 90089; e-mail: jcchen@usc.edu.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR-83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

© 2017 Health Effects Institute, 75 Federal Street, Suite 1400, Boston, MA 02110-1817

ADDITIONAL MATERIALS

Part A: Study Population, Design, and Outcome Ascertainment

1. Recruitment of Study Participants in WHIMS and WHIMS-MRI

Figure 1: Recruitment of WHIMS Population

(adapted from Shumaker 2004)

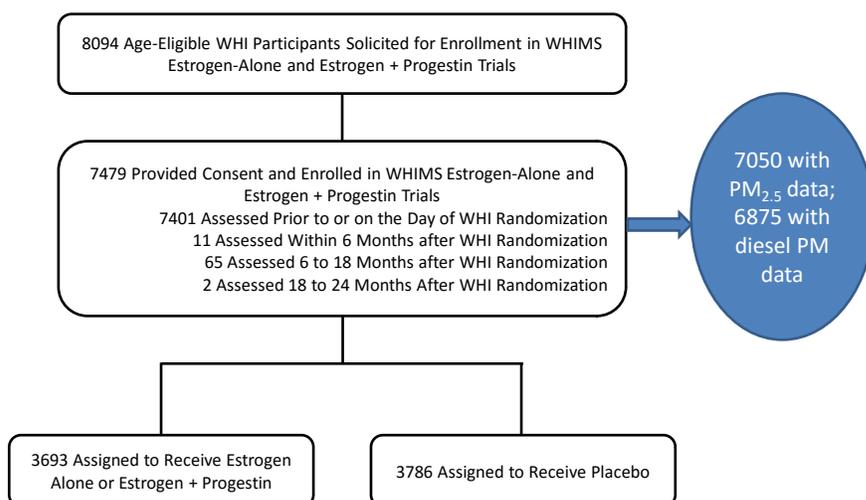
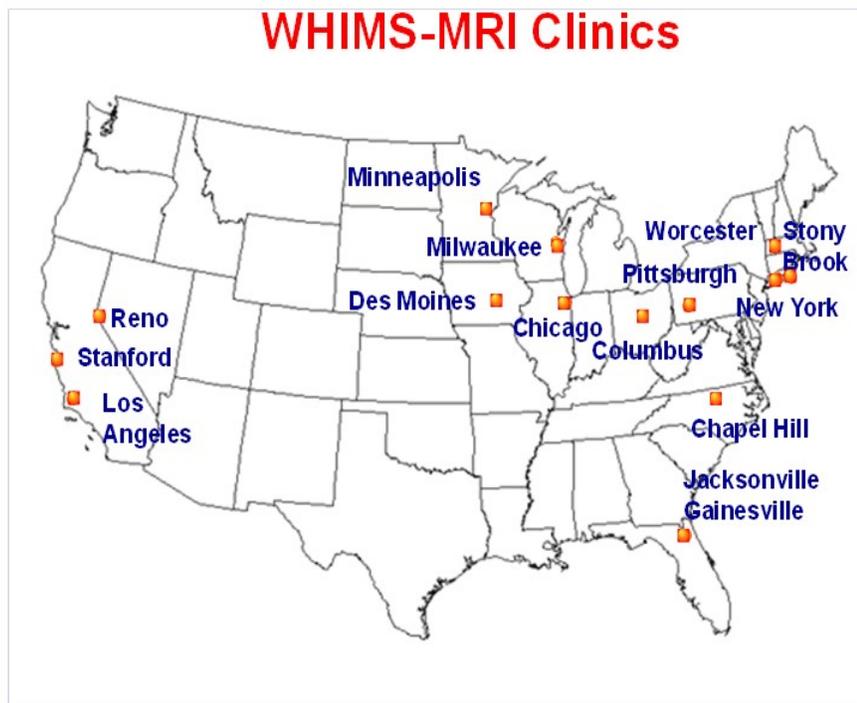
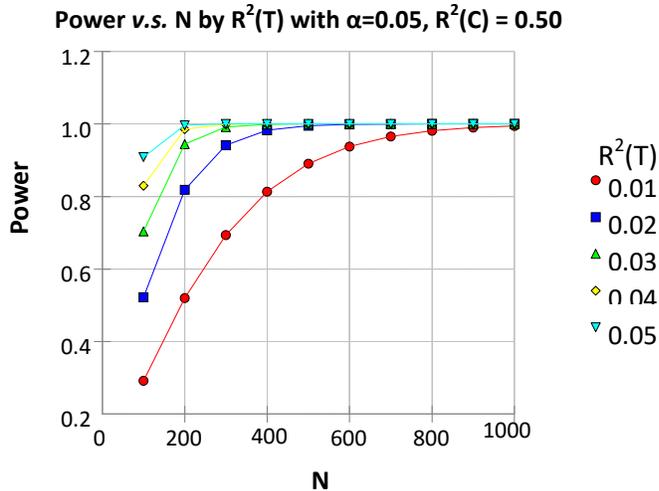


Figure 2: Geographic Distribution of WHIMS MRI Clinic Sites



2. Statistical Power Analyses for Primary Study Aims

We provided the statistical power analyses for the two primary hypotheses (Specific Aims 1–2). In the following Figure, we reported our sample size and statistical power analyses relating MRI-



measured brain volume to ambient air pollutants in multiple regression models (Cohen 1988). With the WHIMS-MRI current sample size (N=1403), the proposed study will have the sufficient statistical power to detect the association between air pollution exposure and MRI-measured brain volume (Aim 1). Our analyses were conditioned on a set of 20 covariates or potential confounders that have accounted for 50% of observed variation (the R²(C)=0.50) in brain volume, and we presented the varying statistical power as a function of sample size (N) and the variation in brain volume as explained by the added continuous variable of air

pollution (the R²(T)=0.01–0.05). Our analyses show that a sample size as small as 1000 will provide 99% power to detect what we see to be clinically important associations with air pollution (i.e. that account for at least 1% or more in the conditional variability of MRI-based measures of neuropathological abnormalities), with adjustment for 20 potential confounders.

The attached table summarized the results from our statistical power analyses (Hsieh and Lavori 2000) for the putative effects on incidence of MCI and all-cause dementia (Aim 2). The proposed study will have a sufficient power to detect a hazard ratio (HR) ≥ 1.75 (per 10- $\mu\text{g}/\text{m}^3$) for the PM_{2.5} effect (power ≥ 0.85) and HR ≥ 1.12 (per inter-quartile change) for the DEP effect (power ≥ 0.83). These statistical power analyses were based on: (1) a presumed final sample size of 7000, which accounts for the possible missing data within the full WHIMS cohort (N=7479); (2) an event rate of 5%, as estimated from the published number of incident events (108 for probable dementia and 202 for MCI) during the on-trials period; (3) relative crude estimates of exposure gradients for annual PM_{2.5} and 1-year average DEP; and (4) a presumed explainable exposure R²=0.40 (i.e., the included covariates account for 40% of variability in air pollution exposure). The exposure gradient for PM_{2.5} was based on reported data in a WHI-OS ancillary study which used the AQS data in 2000 (Miller et al. 2007). For DPM, we use the estimated total diesel PM data in 1996.

Table: Power analyses for the effects of PM_{2.5} and DPM on MCI/probable dementia

PM _{2.5} effect (per 10- $\mu\text{g}/\text{m}^3$ increase)			DPM effect (per inter-quartile [IQR] increase)		
Annual PM _{2.5} (Mean \pm SD) 13.5 \pm 3.7	HR=1.50	0.59	One-year DPM (Mean \pm SD) 1.64 \pm 2.12 (IQR=1.20 $\mu\text{g}/\text{m}^3$)	HR=1.12	0.83
	HR=1.75	0.85		HR=1.14	0.92
	HR=2.00	0.96		HR=1.16	0.97
	HR=2.25	>0.99		HR=1.18	0.99

ADDITIONAL MATERIALS

Part B: Estimation of Residential Exposures

1. Residential location data

Geocoded information of participants' addresses was used to define the residential locations where PM exposures were estimated. Because WHI address data since its inception (in 1993) were collected prospectively at each clinical visit and updated at least biannually, we were able to account for the varying exposure levels due to relocation. For the current study, WHIMS participants would provide more frequent updates of their residential information as they came to the scheduled clinical visit (in 1996-2007), participated in the annual cognitive assessment (1996-2010), or involved in other research activities for the WHIMS Suite of Studies, including the WHIMS extension (2005-2008), the WHIMS Epidemiology of Cognitive Health Outcomes (WHIMS-ECHO; 2008-*onward*), WHI Study on Cognitive Aging (WHISCA; 1999-2010), (Resnick et al. 2004; Espeland et al. 2006; Resnick et al. 2009b) and the WHIMS-MRI-1 (2005-6) (Jaramillo et al. 2007; Espeland et al. 2009) and MRI-2 Studies (2010-11) (Espeland et al. 2013). According to the geocoded information processed for this project, no WHIMS subjects moved out of the geographic regions covered by the exposure estimation described below. The annual exposure (PM_{2.5} in 1999-2007; DPM in 1996-2005) was assigned to each residential location. Given the longitudinal information on residence-year, cumulative exposures were then derived as time-weighted average pertinent to the study outcome measures (e.g., cumulative PM_{2.5} from 1999 to the WHIMS-MRI inception in 2005-6; cumulative DPM from 1996 to the WHIMS-MRI inception in 2005-6).

2. BME estimates of residential exposures to PM_{2.5}

BME Database Development and Spatiotemporal Modeling of Annual PM_{2.5} Exposure

2.1. Obtaining air PM_{2.5} monitoring raw data

We followed the HEI's requirement and had maintained a documentation of the procedure conducted to create and quality assure/quality control this database of the raw data. The data documentation included the description of procedures used to download data from the US Environmental Protection Agency Air Quality System (AQS), as well as documentation of the code used to create results of the comparison between BME estimates and inverse distance weighted estimates (see below in #4).

2.2. Cleaning/Reconstructing annual PM data

Based on the downloaded daily PM_{2.5} data, we calculated the annual PM values, and that we validated these values by achieving a perfect match between an arbitrary subset of our values and those obtained from an independent source.

2.3 Developing BME soft data for annual PM_{2.5} concentrations

In Table 1, we summarized the AQS dataset used for developing the nationwide BME model (#4), which integrated the hard data and soft data, as defined below.). The numbers of hard data in Table 1 are the number of annual PM values obtained at all AQS stations from 1999 to 2007 for which sampling frequency is not missing, and for which at least 75% of intended samples were collected in each quarter.

Table 1: Number of hard data, soft data, and their fraction in the contiguous US and Hawaii from 1999 to 2007.

Year	Hard data (*)	Soft data (**)	Fraction
1999	331	369	0.527
2000	676	315	0.317
2001	858	195	0.185
2002	970	145	0.130
2003	910	177	0.162
2004	1004	127	0.112
2005	1005	141	0.123
2006	989	142	0.125
2007	989	168	0.14
Total	7732	1779	

(*) Number of site/years for which the sampling frequency is not missing,

and for which at least 75% of intended samples were collected in each quarter

(**) Number of site/years for which the sampling frequency is not missing, and for which the number of samples in each quarter was at least equal to one and less than 75% of intended samples

2.4 Developing national-scale BME estimation for yearly PM concentration and comparing them with inverse distance-weighted estimates

We completed this task by implementing of the BME method, and comparing it with inverse distance weighted method. The results of this work are described here, and they were presented in April 2013 at a poster session of the 2013 HEI conference.

Material and Methods

Data

The data used in this work consist in the reconstructed annual average data (see #2), but restricted to AQS stations located in the contiguous US and Hawaii. This removed stations in areas such as Puerto Rico, Alaska, etc., which reduced the total number of hard data to 7732 and soft data to 1779.

Inverse Distance Weighting Method

Inverse Distance Weighting (IDW) method is one of the most commonly used deterministic interpolation methods in exposure assignment and has been applied in many air pollution epidemiological studies. The IDW estimate $\hat{Z}^{IDW}(s)$ of yearly PM_{2.5} at an unmonitored location s is estimated as the weighted average of the values $Z(s_i)$ observed at surrounding data points $s_i, i=1, \dots, n$, which is mathematically expressed as

$$\hat{Z}^{IDW}(s) = \frac{\sum_{i=1}^n w_i(s)Z(s_i)}{\sum_{j=1}^n w_j(s)} \quad (1)$$

$$w_i(s) = \frac{1}{d_i^p}$$

where w_i is a weight based on the distance d_i between the estimation point s and the data point s_i , p is a positive real number called the power parameter and n is the number of data points surrounding s that are used in the estimation.

In order to determine the value of the power parameter p and the number of the data points n , we reviewed previous studies that estimated the concentration of particulate matter using the IDW method. From the 17 relevant studies published that also described nationwide/large-scale exposure interpolation, more than half of these studies did not provide the value of the power parameter p used, however the default value is 2 in the IDW tool of ArcGISTM, the most commonly used GIS software, and four studies clearly stated that the power parameter used was $p=2$. Thus we used $p=2$ in this work. The number of data points n varied by study and ranged from 3 to 15, or was equal to all data points in the study region or within a certain radius (20km - 100km). In this work, we used two implementations of the IDW method, one which used all the data points within a 30 miles radius of the estimation, and the other which used all data points in the study region.

Bayesian Maximum Entropy Method

The BME method introduced by Christakos (1990; 2000) provides a mathematically rigorous framework that integrates a variety of available knowledge bases (e.g., spatial dependency model, empirical relationships, scientific model, physical laws etc.) with data having varying levels of epistemic uncertainty. These data are categorized in hard data corresponding to exact measurements of the process, and soft data, which may have an uncertainty characterized by a PDF of any type (e.g., Gaussian, Uniform). A full description of the epistemic underpinnings of the BME method and its numerical implementation can be found elsewhere (Serre and Christakos; 1999 Christakos et al., 2002). In brief the BME method can be viewed as a two-stage knowledge processing procedure: At the prior stage, maximum entropy theory is used to process the general knowledge base at hand and produces a prior PDF describing the spatial process. Then at the posterior stage, an epistemic Bayesian conditionalization rule is used to update this prior PDF with respect to the site specific hard and soft data available, which produces a BME posterior PDF describing the value of the spatial process at any estimation point of interest.

We denote as Z_i the annual average concentration of $PM_{2.5}$ measured at space/time point $\mathbf{p}_i=(s_i,t_i)$, where s_i and t_i are the spatial coordinate and time of the annual measurement. Let $o(\mathbf{p})$ be a deterministic global space/time offset chosen such that it captures consistent patterns in the space/time distribution of the annual $PM_{2.5}$ data. The word “global” emphasizes that this offset is defined for any space/time coordinate $\mathbf{p}=(s,t)$ within the study domain encompassing all the available data. Let’s define $X(\mathbf{p})$ as a space/time random field (S/TRF) with a homogenous/stationary covariance function, and such that a realization of this S/TRF be the values $Z_i-o(\mathbf{p}_i)$ observed at data points \mathbf{p}_i . The uncertainty and space/time variability of the annual average concentration of $PM_{2.5}$ across the study domain can then be described as in terms of the S/TRF

$$Z(\mathbf{p})= X(\mathbf{p})+o(\mathbf{p}), \quad (2)$$

where $o(\mathbf{p})$ is the global offset chosen deterministically to offset consistent patterns in the annual $PM_{2.5}$ concentration, and $X(\mathbf{p})$ is a residual S/TRF with homogenous/stationary covariance determined by the data Z_i and the choice of the offset $o(\mathbf{p})$. The global offset model used in this work is that presented for similar previous air pollution studies (Serre et al. 2004; Akita et al. 2007, 2012; De Nazelle et al. 2010), which corresponds to an additive space-time model $o(s,t)=o_s(s)+o_t(t)$, where the spatial component $o_s(s)$ is obtained using an exponential kernel smoothing of the time-averaged data, and the temporal component $o_t(t)$ is obtained using an exponential kernel smoothing of the spatially-averaged data. Following the procedure described in these previous works, the parameters of the kernel smoothing are chosen so as to smooth out small area/time variability in the data while capturing regional space/time trends. Operationally this was achieved by selecting smoothing parameters that produced the smallest residual variability while displaying the highest degree of autocorrelation.

In order to describe the BME fundamental equation we will use subscript k (i.e., X_k) to denote the random variable representing the S/TRF at estimation point \mathbf{p}_k (i.e., $X_k = X(\mathbf{p}_k)$). Similarly subscripts h and s are used to represent vectors of random variables corresponding to the S/TRF at the hard data points $\{\mathbf{p}_h\}$ and the soft data points $\{\mathbf{p}_s\}$, respectively. By convention, lower case variables (e.g. x_h, x_s , or x_k) will denote realizations or deterministic values taken by their corresponding upper case random variables (e.g. X_h, X_s or X_k).

In the case that the general knowledge base \mathbf{G} about the S/TRF of the residual concentration $X(\mathbf{p})$ consists in its mean trend $m_X(\mathbf{p}) = E[X(\mathbf{p})]$ and covariance function $c_X(\mathbf{p}, \mathbf{p}')$, then the BME fundamental equation reduces to

$$f_K(x_K) = A^{-1} \int dx_s f_G(x_h, x_s, x_k) f_S(x_s) \quad (3)$$

where A is a normalization constant, the prior PDF f_G obtained from entropy maximization on \mathbf{G} is multivariate normal with mean and covariance given by $m_X(\cdot)$ and $c_X(\cdot)$, respectively, the vector of deterministic values x_h corresponds to the hard data, and f_S is a PDF characterizing the epistemic uncertainty of the soft data. The BME posterior PDF f_K is denoted with a subscript $\mathbf{K} = \mathbf{G} \cup \mathbf{S}$ representing the union of the general knowledge $\mathbf{G} = \{m_X(\cdot), c_X(\cdot)\}$ and site specific knowledge $\mathbf{S} = \{x_h, f_S(\cdot)\}$. The expected value of the BME posterior PDF provides an estimate of the residual concentration $X(\mathbf{p})$ at the estimation point, and the corresponding BME posterior standard deviation provides a useful characterization of the associated estimation uncertainty.

In the limiting case where only hard data are included in the estimation process, the BME estimator is simply the kriging estimator (Christakos, 1990; 2000). Indeed, if we remove the soft data from Equation (3), we obtain $f_K(x_K) = A^{-1} f_G(x_h, x_k)$, where the normalization constant needs to be equal to

$$A^{-1} = \int dx_k f_G(x_h, x_k) = f_G(x_h). \text{ Hence the posterior pdf can be re-written as}$$

$f_K(x_K) = f_G(x_h, x_k) / f_G(x_h) = f_G(x_k | x_h)$ which is nothing more than the conditional pdf under knowledge base \mathbf{G} . If \mathbf{G} is restricted to only including the spatial mean trend $m_X(s)$ and spatial covariance function $c_X(s, s')$, then under maximum entropy maximization $f_G(x_h, x_k)$ is multivariate Gaussian with mean and covariance specified by $m_X(s)$ and $c_X(s, s')$ (Christakos 1990; 2000, Serre and Christakos; 1999, Christakos et al., 2002), and it follows that the posterior pdf is also Gaussian with a conditional mean equal to

$$E[X_k | x_h] = m_k + C_{k,h} C_{h,h}^{-1} (x_h - m_h) \quad (4)$$

where $m_k = m_X(s_k)$, $m_h = m_X(s_h)$, $C_{k,h} = c_X(s_k, s_h)$ and $C_{h,h} = c_X(s_h, s_h)$. This Gaussian conditional mean is a linear combination of the hard data values, with weights equal to those of the classical kriging methods. This demonstrates that BME is consistent with the widely used spatial kriging estimator since it reduces to linear kriging when the mean and covariance is restricted to the spatial domain and the data only include the hard data. This makes BME a consistent extension of the widely used kriging estimator when one needs to (a) extend spatial kriging to the space/time domain and (b) integrate non-Gaussian soft data.

The uncertainty associated with annual PM_{2.5} averages calculated from an incomplete set of daily PM_{2.5} concentrations is characterized by a probability density function (PDF). We assume that an adequate approximation for the PDF at monitoring station i in year t (1999-2007) is a normal distribution with mean $\mu_{s,i}$ and standard deviation $\sigma_{s,i}$ truncated below zero, since the annual average concentrations cannot be negative. The mean $\mu_{s,i}$ is simply set to the sample mean of the n_i daily concentrations measured at station i in year t . The epistemic uncertainty associated with a soft datum arises from the fact that the number of measurements n_i taken over a year may be less than the intended number of measurements n_i^* (i.e. the number of measurement that would have been collected if the station had operated as intended). Hence this uncertainty relates to the difference between the true mean of all n_i^* intended daily measurements, and the sample mean $\mu_{s,i}$ calculated from an incomplete sample of size n_i selected from a finite population of size n_i^* . Therefore, a reasonable value for the standard deviation $\sigma_{s,i}$ is

$$\sigma_{s,i} = \sqrt{\frac{\sum_{j=1}^{n_i} (y_{i,j} - \mu_{s,i})^2 / (n_i - 1)}{n_i}} \times \sqrt{\frac{n_i^* - n_i}{n_i^*}} \quad (5)$$

where the first term of this equation is the standard deviation of the sample mean and the second term is a finite population correction factor to account for the finite population size.

The BME framework presented above was implemented in the MATLAB programming language by using or extending MATLAB functions available in the *BMElib* (version 2.0) numerical library (Serre and Christakos; 1999 Christakos et al., 2002).

Cross-validation Analysis

In order to evaluate the model performance of the proposed and the conventional approaches, a leave-one-out cross-validation analysis was conducted for the following three spatial interpolation and one space/time interpolation methods: (1) the spatial nearest-neighbor approach (NNA), (2) the spatial IDW method, (3) the spatial kriging method, and (4) the space/time BME method. All the annual average concentrations during study period (1999-2007) that met the completeness criterion were used as the validation points. The first three methods are purely spatial interpolation methods, so if annual PM_{2.5} is estimated for a given year of interest, then only the hard data for that year of interest is used. On the other hand the space/time BME estimation differs in two ways. First it accounts for both the hard and soft data available. Second, for a given year of interest, the hard and soft data used comes from that year as well as the adjacent years.

Model performance was evaluated using the following cross-validation statistics: the root mean square error (RMSE), the Pearson correlation coefficient, and the Spearman's rank correlation. In addition to the cross-validation analysis, maps of the estimated PM_{2.5} annual concentration were produced over the contiguous US and California to visually inspect the estimation results.

Results

Global offset and covariance function

The global offset $o(\mathbf{p})=o_s(s)+o_t(t)$ used in this work to model the regional space/time trend in the data (Figure 2) was obtained using the *BMElib* (version 2.0) numerical library. Using *BMElib*, the time-average of annual concentrations are first calculated for each monitoring sites (Fig. 2a), and then a spatial exponential filter is applied to these raw time-averages to obtain the smoothed spatial component $o_s(s)$ shown in Fig 2(b). The spatial offset $o_s(s)$ clearly smoothed out differences in near-stations, while retaining regional trends displaying relatively high concentrations in the East Coast and in California compared to the lower concentrations observed in the Midwest. Likewise, in *BMElib*, the spatial-average of annual concentrations are calculated for each year (dashed line, Fig. 2c), and an temporal exponential filter is used to obtain the smoothed temporal component $o_t(t)$ shown as a plain line in Fig 1(c). This figure clearly exhibits a decreasing trend in the annual PM_{2.5} average concentrations over the study period.

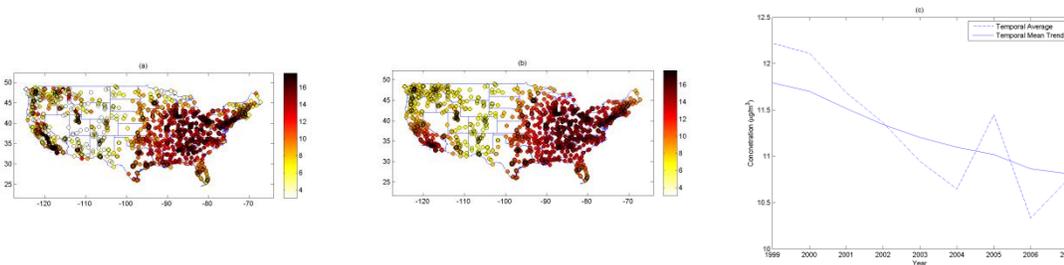


Figure 2: Maps of (a) the time-averaged PM_{2.5} concentration ($\mu\text{g}/\text{m}^3$) calculated for each site as the average of the annual PM_{2.5} concentrations from 1999 to 2007, (b) the corresponding spatial offset $o_s(s)$ ($\mu\text{g}/\text{m}^3$), and (c) time series of annual PM_{2.5} concentrations averaged for each year over all monitoring sites (dotted line) and corresponding temporal offset $o_t(t)$ (solid line).

Based on the global offset $o(\mathbf{p})$ described above, we obtained the covariance function of the residual PM_{2.5} annual concentrations $X(\mathbf{p})$ shown in Figure 3. First we computed the sample covariance values from the residual concentrations, which are shown as red circle in Figure 3. Then, the sample covariance values were used to fit the positive definite covariance model shown with a green line in Figure 3. Since the sample covariance values exhibit both short and long range variability, we used a two components exponential covariance model (Serre et al, 2004), which is given by the following equation.

$$C(r, \tau) = c_1 \exp\left(-\frac{3r}{a_{r1}}\right) \exp\left(-\frac{3\tau}{a_{\tau1}}\right) + c_2 \exp\left(-\frac{3r}{a_{r2}}\right) \exp\left(-\frac{3\tau}{a_{\tau2}}\right) \quad (6)$$

where c_i is the sill, a_{ri} is the spatial range and $a_{\tau i}$ is the temporal range of the i th component ($i = 1$ or 2). The model parameters (Table 2) were estimated using an automated weighted least square procedure. In the case of the spatial kriging method, we only need to model the spatial autocorrelation in the data. This is done by setting

$\tau=0$ in Eq. (6) to obtain the spatial covariance model $C(r) = c_1 \exp(-3r/a_{r1}) + c_2 \exp(-3r/a_{r2})$, where the parameters c_i and a_{ri} ($i = 1$ or 2) take the values listed in Table 2.

Table 2: Estimated covariance model parameters

Parameter		Value
Sill	c_1	$5.71 (\mu\text{g}/\text{m}^3)^2$
Spatial Range	a_{r1}	156 (Km)
Temporal Range	$a_{\tau1}$	148 (Yr)
Sill	c_2	$0.825 (\mu\text{g}/\text{m}^3)^2$
Spatial Range	a_{r2}	1446 (Km)
Temporal Range	$a_{\tau2}$	178 (Yr)

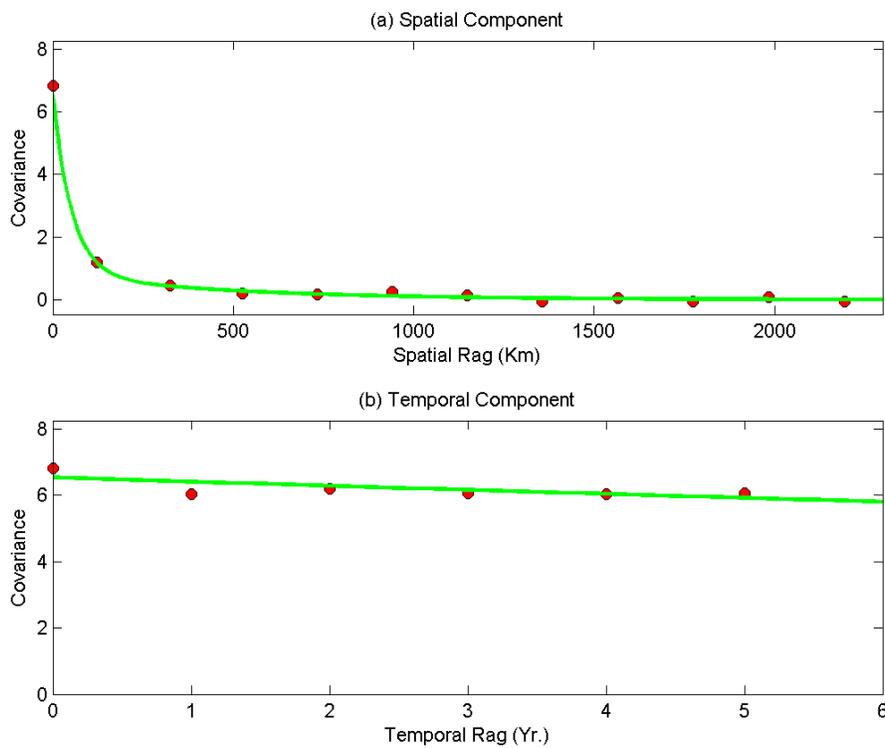


Figure 3: (a) Spatial and (b) temporal component of the covariance function of residual PM_{2.5} annual concentrations

Cross-validation Analysis

Table 3 shows the root mean square error (RMSE), Pearson correlation coefficient (CORR) and Spearman's rank correlation (RANK) obtained from cross validation. First we compared NNA and IDW when these methods are restricted to only using data within a 30 miles radius (upper panel of Table 3). In that case the validation set is reduced to the 5674 PM_{2.5} annual average concentrations observed at monitoring sites in the

contiguous US and Hawaii having at least one monitoring site within a 30 miles radius. In that case NNA (<30 miles) produces a larger RMSE than IDW (<30 miles).

We then compared the four estimation methods when they are not limited to a 30 miles radius (lower panel of Table 3). In this case the cross validation set consists of all 7732 PM_{2.5} annual average concentrations observed at monitoring sites in the contiguous US and Hawaii. We find that NNA produces the largest RMSE among all approaches. The IDW method reduced the RMSE by 9% relative to NNA. Kriging further reduced the RMSE by 18% relative to NNA. Moreover, BME achieved the smallest RMSE amongst all methods, corresponding to a 24% reduction relative to NNA. The improvement of the Pearson correlation coefficient and Spearman’s ranks correlation were also the largest for the BME method. The reduction in RMSE and the corresponding increase in CORR can be seen in the scatter plots (Figure 3).

Table 3: Cross-validation error statistics of four interpolation methods

Method ¹	(1) spatial NNA	(2) spatial IDW	(3) ⁴ spatial kriging	(4) ⁵ space/time BME
RMSE ²	2.20	2.07	N/A	N/A
CORR ²	0.833	0.848	N/A	N/A
RANK ²	0.864	0.879	N/A	N/A
RMSE ³	2.52	2.29	2.01	1.92
CORR ³	0.811	0.845	0.872	0.884
RANK ³	0.846	0.873	0.891	0.902

¹ For a given estimation point, the spatial NNA, IDW and kriging methods only use hard data for the year of estimation, while the space/time BME method uses both hard and soft data for the estimation year as well as adjacent years.

² Cross-validation error statistics based on 5674 PM_{2.5} annual average concentrations observed at monitoring sites in the contiguous US and Hawaii having at least one monitoring site within a 30 miles radius.

³ Cross-validation error statistics based on 7732 PM_{2.5} annual average concentrations observed at the all monitoring sites in the contiguous US and Hawaii.

⁴ Spatial kriging uses the spatial covariance $C(r) = c_1 \exp(-3r/a_{r1}) + c_2 \exp(-3r/a_{r2})$, where the parameters c_i and a_{ri} ($i = 1$ or 2) take the values listed in Table 2.

⁵ Space/time BME uses the space/time covariance model described in equation (6).

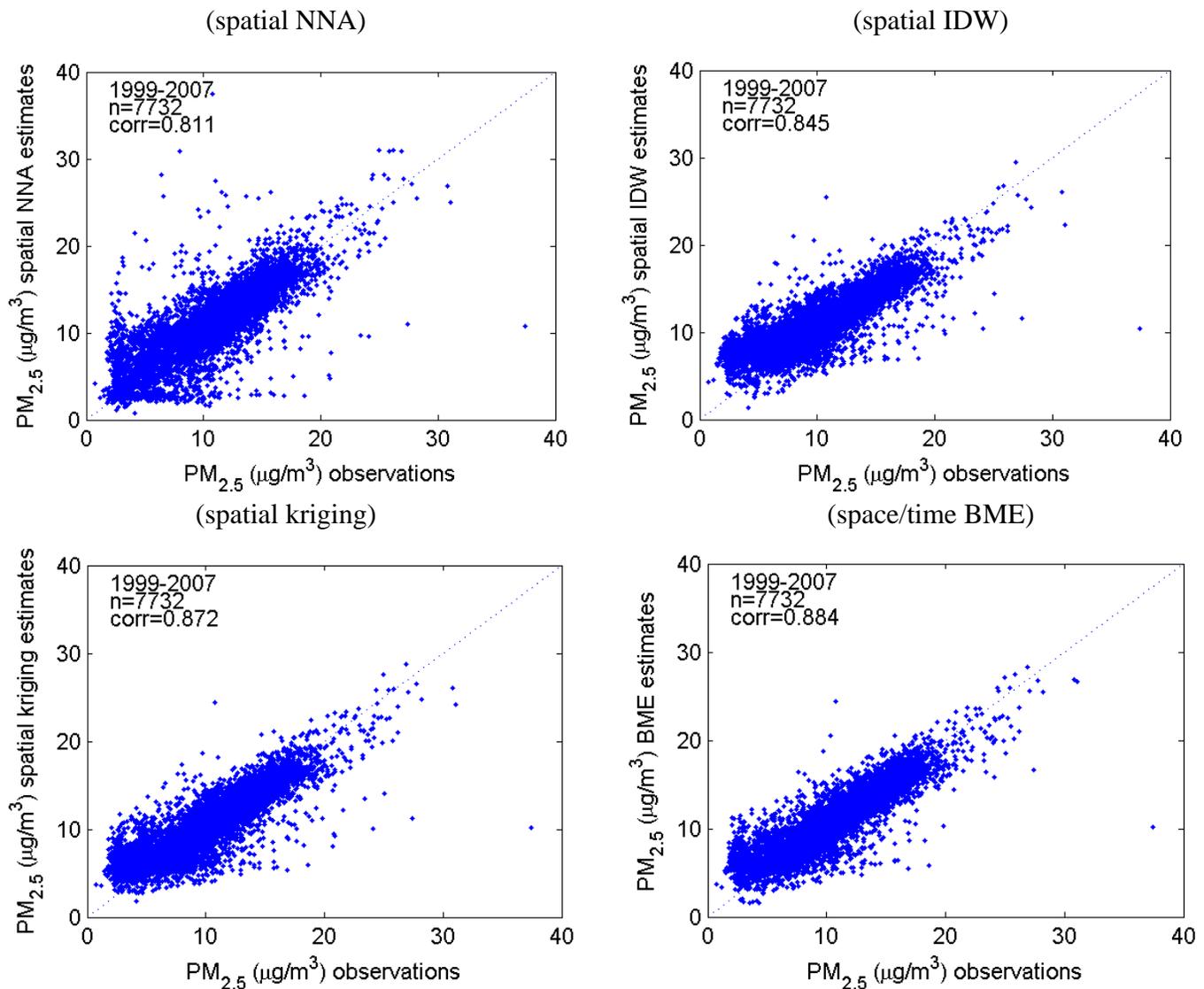


Figure 4: Scatter plots of estimated versus observed annual PM_{2.5} based on 7732 PM_{2.5} annual average concentrations values observed in 1999-2007.

Since PM_{2.5} fluctuates from year to year, we furthermore calculate the performance statistics by years for the RMSE (Table 4) and the CORR (Table 5). We find that for each year the order of performance (from worse to best) is maintained as being spatial NNA, spatial IDW, spatial kriging and space/time BME. The improvement of performance between methods can be seen in the scatter plots for 1999 (Figure 4). The same patterns can be seen in the scatter plots for other years (plots not shown).

Altogether, the cross validation analysis provides evidence supporting that the BME method was correctly implemented since it produced results with lower RMSE and higher CORR than the other methods. The difference in performance statistics is most noticeable when comparing the two non-statistical methods (spatial NNA and spatial IDW) with the two geostatistical methods (spatial kriging and space-time BME). By contrast, the difference between the two geostatistical methods is small. Our cross-validation results showed the performance of BME approach vs. spatial kriging in modeling yearly PM_{2.5} exposure 1999-2007 was only modest, likely reflecting the fact that PM_{2.5} data were fairly complete and we only had 8-9 years of longitudinal data. However, even though the difference is small, the results of this comparison analysis supports that one should use space-time BME, since it produces better performance statistics.

Table 4: RMSE cross-validation error statistics by year

Year	n	RMSE spatial NNA	RMSE spatial IDW	RMSE spatial kriging	RMSE space/time BME
1999	331	3.687	3.363	2.956	2.833
2000	676	2.7	2.367	2.031	1.919
2001	858	2.626	2.419	2.13	1.954
2002	970	2.488	2.276	1.961	1.913
2003	910	2.361	2.193	1.929	1.891
2004	1004	2.427	2.191	1.955	1.859
2005	1005	2.454	2.268	1.95	1.879
2006	989	2.357	2.074	1.866	1.813
2007	989	2.269	2.093	1.846	1.768
1999-2007	7732	2.517	2.289	2.008	1.923

Table 5: CORR cross-validation error statistics by year

Year	n	CORR spatial NNA	CORR spatial IDW	CORR spatial kriging	CORR space/time BME
1999	331	0.767	0.794	0.832	0.847
2000	676	0.806	0.852	0.883	0.896
2001	858	0.813	0.847	0.868	0.892
2002	970	0.811	0.842	0.873	0.882
2003	910	0.824	0.853	0.876	0.885
2004	1004	0.785	0.826	0.851	0.871
2005	1005	0.834	0.866	0.89	0.899
2006	989	0.787	0.835	0.858	0.867
2007	989	0.813	0.841	0.868	0.881
1999-2007	7732	0.811	0.845	0.872	0.884

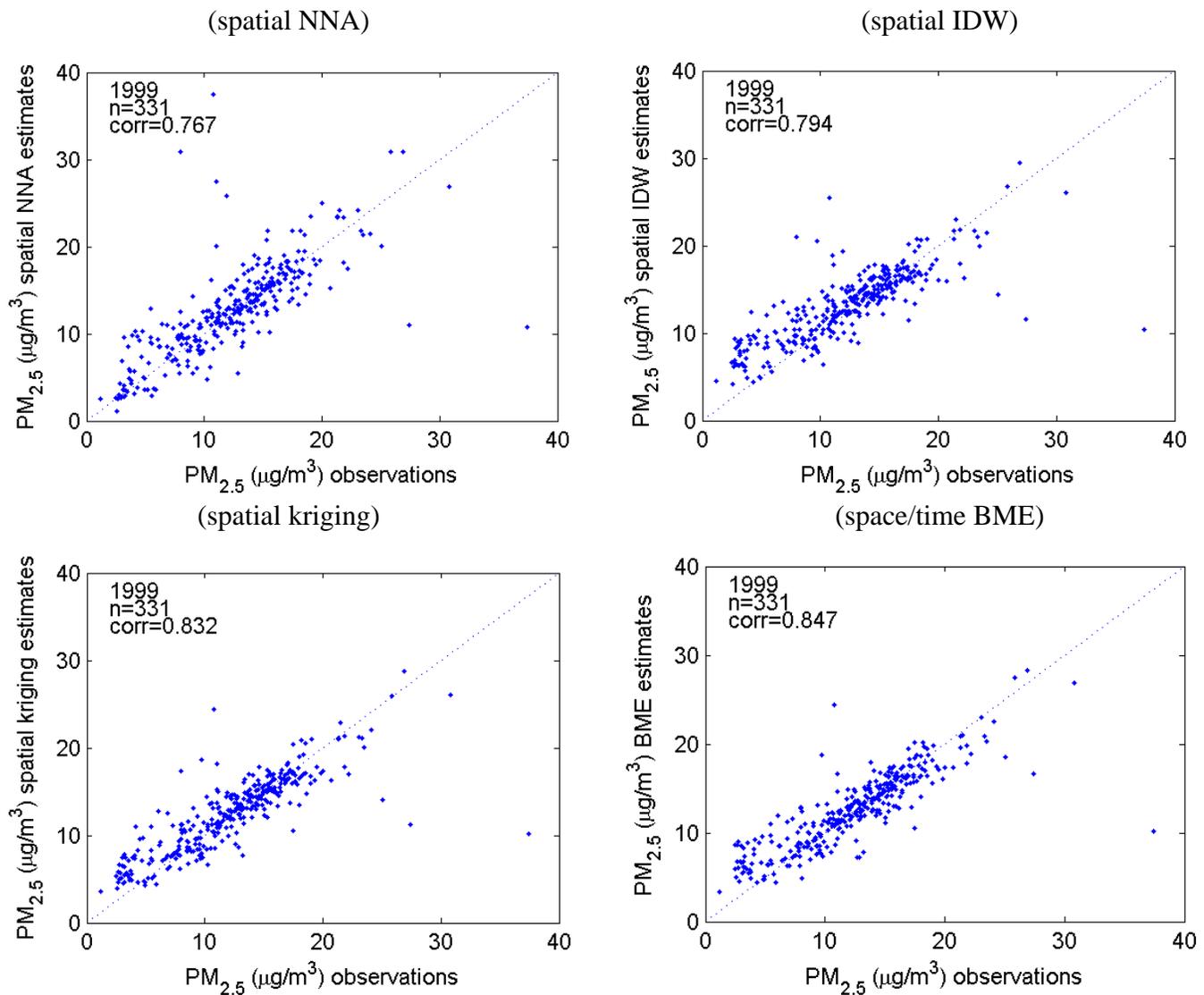


Figure 5: Scatter plots of estimated versus observed annual $PM_{2.5}$ based on 331 $PM_{2.5}$ annual average concentrations values observed in 1999

Estimation map

Figure 6 shows maps of the $PM_{2.5}$ annual average concentration estimated in 2002 over the contiguous US and California by the NNA based on the all the monitoring sites (left), the IDW method based on all the monitoring sites (center), and the BME method (right). The circles show the hard data points, whereas the triangles indicate soft data points used in the BME method. As expected, NNA produced a map showing a patchy distribution of concentrations, since the estimated concentrations were simply assigned the value at the nearest data point. The IDW method resulted in a map with a more realistic and smoother spatial distribution of concentrations. However, since the concentration estimated by the IDW method simply depends on the distance from the data points, the map produced disconnected islands of high and low concentrations centered on data points. The BME method accounts for the autocorrelation of the data and the additional information provided by the soft data shown as triangles, which results in a more realistic map that is less patchy and better describes contiguous areas of high versus low concentrations, such as those displayed along the central valley of California.

(Nearest Neighbor)

(IDW)

(BME)

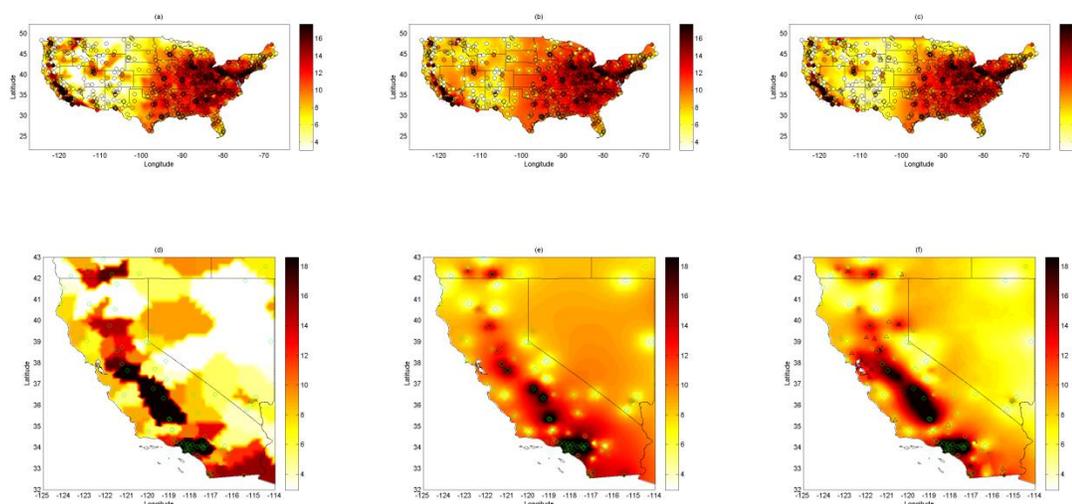


Figure 6: Maps of the PM_{2.5} annual average concentration estimated in 2002 over the contiguous US (top row) and California (bottom row), as obtained by NNA based on the all the monitoring sites (left), the IDW method with all the monitoring sites (middle), and the BME method (right).

References cited:

- Akita, Y.; Carter, G.; Serre, M. L. (2007) Spatiotemporal nonattainment assessment of surface water tetrachloroethylene in New Jersey. *J. Environ. Qual*, 36 (2), 508–20.
- Akita, Y., JC Chen; and M.L. Serre (2012) The moving-window Bayesian Maximum Entropy framework: Estimation of PM_{2.5} yearly average concentration across the contiguous United States. *Journal of Exposure Science and Environmental Epidemiology* . 22(5):496-501
- Christakos G., 1990. "A Bayesian/maximum-entropy view to the spatial estimation problem". *Mathematical Geology*, 22(7): 763-776.
- Christakos G., 2000. *Modern Spatiotemporal Geostatistics*. Oxford University Press, New York, NY (Out of Print.).
- Christakos, G., P. Bogaert, and M.L. Serre (2002) *Temporal GIS: Advanced Functions for Field-Based Applications*, Springer-Verlag, New York, N.Y., 217 p., ISBN: 978-3-540-41476-6.
- De Nazelle, A., S. Arunachalam, M.L. Serre† (2010) Bayesian Maximum Entropy Integration of Ozone Observations and Model Predictions: An Application for Attainment Demonstration in North Carolina, *Environ. Sci. Technol.* Vol. 44, pp. 5707–5713.
- Serre, M. L., and G. Christakos (1999) Modern geostatistics: Computational BME in the light of uncertain physical knowledge--the Equus Beds Study, *Stochastic Environmental Research and Risk Assessment*, Vol. 13, No. 1, pp. 1-26.
- Serre, M.L., G. Christakos, and S-J Lee (2004) Soft Data Space/Time Mapping of Coarse Particulate Matter Annual Arithmetic Average over the U.S., pp. 115-126. In X. Sanchez-Vila, J. Carrera, and J. J. Gomez-Hernandez, editors, *geoENV IV - Geostatistics for Environmental Applications*, Kluwer Academic Publishers, Dordrecht, 560 p., ISBN: 978-1-4020-2114-5.

3. NATA estimates of ambient concentrations at census tract-level of DPM exposures

NATA DPM Estimates: Hierarchical Spatiotemporal Analyses, Census Tract-Specific Interpolation, and Nationwide Mapping in 1996-2005

Introduction

The Environmental Protection Agency's (EPA) National-scale Air Toxics Assessment (NATA) is an ongoing project that involves a comprehensive evaluation of air toxics in the US. Four assessments have been completed (1996, 1999, 2002 and 2005), each providing an annual estimate of outdoor hazardous air pollutants (HAPs) plus diesel particulate matter (DPM) concentrations (in $\mu\text{g}/\text{m}^3$) at the census tract level. Emissions data are the key input to the NATA modeling framework, and the ambient concentrations of DPM and other HAPs were among the primary outputs of NATA models. The output concentrations were split into two on-road and non-road sources. For this study, we chose the on-road DPM exposure as the proxy indicator of exposure to PM from roadway traffics.

The main purpose of conducting the hierarchical spatiotemporal analyses of nationwide DPM estimates was to better understand the potential variability of our assumed exposure variable both over time and within the 3 levels of spatial nesting (census tract within county, county within state). The resulting knowledge would inform us of how best to go about making DPM interpolations for the intervening years (1997; 1998; 2000; 2001; 2003; 2004) when NATA model-based estimates were not generated by the EPA.

Below we detail the procedure taken to investigate the components of variability in NATA DPM and subsequently to interpolate the concentrations 1996-2005 using the 4-year wealth of NATA model-based estimates.

Methods

Assessment of Variability

We used the HLM7 software¹ to model the 4-level data, estimating the variability at each level. The level-1 model provides an estimate of the variability in DPM concentrations attributed to the repeated measures at the tract level. Specifically, $Y_{mijk} = \psi_{0ijk} + \varepsilon_{mijk}$ for $m=1, \dots, 4$ repeated measures (1996, 1999, 2002, 2005) within each census tract, $i=1, \dots, 65141$ census tracts, $j=1, \dots, 3112$ counties and $k=1, \dots, 51$ states (including Washington DC). The DPM concentration is Y_{mijk} , ψ_{0ijk} is the intercept representing the census tract and ε_{mijk} is the residual error representing the repeated measures in time at the census tract level. As $\varepsilon_{mijk} \sim N(0, \sigma^2_{mijk})$, the residual error provides us with an estimate of the variance in the census tracts over the repeated measures. The level-2 model provides an estimate of the variability in the census tracts. Specifically, the intercept from the level-1 model is used to construct $\psi_{0ijk} = \pi_{00jk} + e_{0ijk}$ where $e_{0ijk} \sim N(0, \sigma^2_{ijk})$. Following the same nested structure the level-3 and level-4 models provide estimates of the variance in counties and states. The models are $\pi_{00jk} = \nu_{000k} + r_{00jk}$ where $r_{00jk} \sim N(0, \sigma^2_{jk})$ and $\nu_{000k} = \gamma_{0000} + u_{000k}$ where $u_{000k} \sim N(0, \sigma^2_k)$. No structure was placed on any of the variance-covariance matrices during estimation (i.e. the covariance structure was unrestricted).

Table 4 outlines the specific variance components (as a percent of the total variance) for ambient DPM on-road concentration. We found that time accounted for the largest proportion of the overall variability. For the remaining spatial components, between-census tract accounted for the largest spatial variability (15.48%), with the further partitioning of variance for county and state with almost equal contribution.

Table 6: 4-Level Hierarchical Model Variance Components (% of total variance)

Nesting Level	Ambient On-road
Time (1)	63.87
Between-Census Tract (2)	15.48
Between-County (3)	10.26
Between-State (4)	10.39

These results showed that, for ambient estimates from on-road sources, the remaining spatial variability was largely determined by between-census tract difference as opposed to between-county/within-state or between-state difference in the NATA-modeled annual exposures to diesel PM. The significant proportion (~ 43%) of the spatial variance attributable to the variation at the census-tract level supported the analytical decision using ambient on-road DPM estimate as a traffic-related PM exposure surrogate with reasonable spatial resolution and intra-community variation. In contrast, our earlier analyses (presented at the HEI annual conference 2013; discussed with HEI Research Committee in the 2014 interim review) showed that only 3.65% of the ambient non-road DPM exposure variance would be attributable to the census-tract level, and the overall spatial variance (including all 3 levels) only account for 12% of the overall exposure contrast in 1996-2005.

¹Raudenbush, S.W., Bryk, A.S., & Congdon, R. (2013). HLM 7.01 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.

Interpolation

The first step involved converting the 1996 and 1999 data from 1990 census tracts to 2000 census tracts. We used census tract relationship files available from the US Census Bureau (http://www.census.gov/geo/www/relate/rel_tract.html) for this purpose.

Given the results of the assessment of variability (see Table 6) we proceeded with a strictly temporal interpolation at the census tract level. On examination of the temporal trends in a sample of the census tracts, we found that a linear fit between the years was not ideal. To account for possible non-linearity in concentrations over the 4 years of NATA data, we chose to fit smooth functions using cubic regression splines with a generalized additive model (GAM) for each census tract and DPM concentration type. In addition to being able to account for non-linearity in time, this method allowed for the estimation of standard errors of prediction.

The general model formulation is $y_i = f(x_i) + \varepsilon_i$ where y_i is the DPM on-road concentration in $\mu\text{g}/\text{m}^3$ at time $i=1, \dots, t$ (1996, 1999, 2000, 2005); $f(x_i)$ is function of DPM at year x_i and ε_i are random iid errors. The function $f(x_i)$ is described by basis functions, which for the cubic regression basis are:

$$f(x_i) = \sum_{j=1}^4 b_j(x) \beta_j \text{ where } b_1=1, b_2(x)=x, b_3(x)=x^2, b_4(x)=x^3.$$

The spline essentially connects successive cubic polynomial regressions by knots that are placed at locations where the data are to be connected. The available NATA data for each census tract were fit using this method and then the resultant model was applied to predict concentrations for the missing years (1997, 1998, 2000, 2001, 2003, 2004). These predictions were merged in with the original data, giving a complete yearly time-series of DPM for 1996 to 2005. Using these yearly exposure estimates, we developed nationwide census tract-level maps of diesel PM. The maps for 1996, 1999, 2000, 2005 were NATA estimates, while maps for the intervening years (1997, 1998, 2000, 2001, 2003, 2004) were based on interpolated NATA concentrations from the cubic regression as described before.

ADDITIONAL MATERIALS

Part C: Statistical Analyses

1. Conceptualization of potential confounding and statistical adjustment

When deciding whether a variable is a confounder that should be adjusted for in the statistical analyses, many researchers still use automated variable selection (e.g., backward elimination for covariates with $p > 0.05$ in the final statistical models) or rely on statistical criteria, such as Akaike information criterion (AIC) or Bayesian information criterion (BIC). Analytic approaches based on automated variable selection have known limitations and resulting biases have been reported in the literature (Austin and Tu 2004; Sauer et al. 2013), including case studies for environmental epidemiology (Budtz-Jorgensen et al. 2007). Although the use of AIC and BIC offers practical solutions to select candidate models with optimal statistical fit to the data drawn from population studies on health outcomes with largely known causal structure (e.g., for obtaining the estimates of short air pollution effects on increased mortalities and/or morbidities while controlling for meteorological variables), such approaches often yields analytic results with very little insights about the identification of potential confounders as well as the causal structure where the consequence of (e.g., directionality; magnitude) of resulting confounding occurs.

Following the empirical approach with prior causal knowledge as recommended by modern epidemiologic literature (Greenland and Brumback 2002; Hernan et al. 2002), in this project we used the directed acyclic graph (DAG) to identify a list of important covariates *a priori* as potential confounders for the putative neurotoxic effects of PM exposures. Conceptually, these covariates would have established causal relationships *both* with the studied outcomes (structural

brain volumes; risk of MCI or dementia), *and* also with the PM exposures. Because we relied on residential locations to define the primary exposure variables, any population characteristics or personal attributes that determine where the people live (and thus the estimated exposure levels) in late life could be considered as potential confounders if they are also known to affect the studied outcomes. For epidemiologic studies on brain health and outdoor air pollutants that relied on location-based exposure estimation, rigorous accounting for such potential confounding had been advocated by experts participating in an NIEHS Workshop (Block et al. 2012).

In this regard, age, race/ethnicity, socioeconomic status, and lifestyle factors (including smoking, alcohol consumption, and physical activities) should be conceptualized as potential confounders *a priori* and included in the fully adjusted analyses. Also, to control for possible spatial confounding, our statistical adjustment included the US census-defined geographic region, because there is increasing evidence of geographic differences in dementia incidences (Russ et al. 2012). Finally, we considered other individual-level health-related factors that may contribute to personal decision regarding where to live in late life. In this regard, the possible candidates included physical attributes (e.g., BMI), care seeking behaviors (e.g., use of HT), mental health (e.g., prior depression), and CVD related clinical risk factors (e.g., histories of hypertension, diabetes mellitus, hypercholesterolemia, and CVD). By comparing the results before and after statistical adjustment for CVD related clinical risk factors, one may examine their potential contribution to the associations, if any revealed between neurocognitive outcomes/brains structural volumes and PM exposures, considering the growing evidence for the cardiometabolic abnormalities associated with PM exposure.

2. Operational definition of time-varying long-term exposure variables

The spatiotemporal BME-derived yearly PM_{2.5} exposures were assigned to each time point when WHIMS participants came to the scheduled clinical visits (in 1996-2007) for various assessments, including the annual cognitive screening (1996-2007) during the on-trial and post-trial periods or participate in WHI Study on Cognitive Aging (WHISCA; 1999-2010), (Resnick et al. 2004; Espeland et al. 2006; Resnick et al. 2009) or the WHIMS-MRI-1 (Jaramillo et al. 2007). These exposure estimates were then used to reconstruct the yearly exposure time-series. Given the reconstructed yearly exposure time-series, we derived the time-varying cumulative yearly exposure along the study time scale (and entered as time-varying covariates in the Cox models). This time-varying cumulative exposure variable was defined as the moving average of all yearly exposures aggregated from the WHI inception year to the index year when an incident event (MCI or probable dementia) was classified or the corresponding risk set was defined), assuming all data elements of the reconstructed yearly exposure time-series were missing at random. In order to reduce the statistical ties of event times and also minimize the influence of temporal misalignment in the exposure contrast, all the risk sets were defined on a daily basis whenever an index event occurred along the study time scale.

References

- Austin PC, Tu JV. 2004. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol* 57(11): 1138-1146.
- Sauer BC, Brookhart MA, Roy J, VanderWeele T. 2013. A review of covariate selection for non-experimental comparative effectiveness research. *Pharmacoepidemiol Drug Saf* 22(11): 1139-1145.
- Budtz-Jorgensen E, Keiding N, Grandjean P, Weihe P. 2007. Confounder selection in environmental epidemiology: assessment of health effects of prenatal mercury exposure. *Ann Epidemiol* 17(1): 27-35.
- Greenland S, Brumback B. 2002. An overview of relations among causal modelling methods. *Int J Epidemiol* 31(5): 1030-1037.

Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. 2002. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol* 155(2): 176-184.

Block ML, Elder A, Auten RL, Bilbo SD, Chen H, Chen JC, et al. 2012. The outdoor air pollution and brain health workshop. *Neurotoxicology* 33(5): 972-984.

Russ TC, Batty GD, Hearnshaw GF, Fenton C, Starr JM. 2012. Geographical variation in dementia: systematic review with meta-analysis. *Int J Epidemiol* 41(4): 1012-1032.

Resnick SM, Coker LH, Maki PM, Rapp SR, Espeland MA, Shumaker SA. 2004. The Women's Health Initiative Study of Cognitive Aging (WHISCA): a randomized clinical trial of the effects of hormone therapy on age-associated cognitive decline. *Clin Trials* 1(5): 440-450.

Espeland MA, Coker LH, Wallace R, Rapp SR, Resnick SM, Limacher M, et al. 2006. Association between alcohol intake and domain-specific cognitive function in older women. *Neuroepidemiology* 27(1): 1-12.

Resnick SM, Espeland MA, An Y, Maki PM, Coker LH, Jackson R, et al. 2009. Effects of conjugated equine estrogens on cognition and affect in postmenopausal women with prior hysterectomy. *J Clin Endocrinol Metab* 94(11): 4152-4161.

Jaramillo SA, Felton D, Andrews L, Desiderio L, Hallarn RK, Jackson SD, et al. 2007. Enrollment in a brain magnetic resonance study: results from the Women's Health Initiative Memory Study Magnetic Resonance Imaging Study (WHIMS-MRI). *Acad Radiol* 14(5): 603-612.