

ORIGINAL ARTICLE

Prediction of fine particulate matter chemical components with a spatio-temporal model for the Multi-Ethnic Study of Atherosclerosis cohort

Sun-Young Kim^{1,2}, Lianne Sheppard^{2,3}, Silas Bergen^{3,4}, Adam A. Szpiro³, Paul D. Sampson⁵, Joel D. Kaufman^{2,6,7} and Sverre Vedal²

Although cohort studies of the health effects of PM_{2.5} have developed exposure prediction models to represent spatial variability across participant residences, few models exist for PM_{2.5} components. We aimed to develop a city-specific spatio-temporal prediction approach to estimate long-term average concentrations of four PM_{2.5} components including sulfur, silicon, and elemental and organic carbon for the Multi-Ethnic Study of Atherosclerosis cohort, and to compare predictions to those from a national spatial model. Using 2-week average measurements from a cohort-focused monitoring campaign, the spatio-temporal model employed selected geographic covariates in a universal kriging framework with the data-driven temporal trend. Relying on long-term means of daily measurements from regulatory monitoring networks, the national spatial model employed dimension-reduced predictors using universal kriging. For the spatio-temporal model, the cross-validated and temporally-adjusted R^2 was relatively higher for EC and OC, and in the Los Angeles and Baltimore areas. The cross-validated R^2 s for both models across the six areas were reasonably high for all components except silicon. Predicted long-term concentrations at participant homes from the two models were generally highly correlated across cities but poorly correlated within cities. The spatio-temporal model may be preferred for city-specific health analyses, whereas both models could be used for multi-city studies.

Journal of Exposure Science and Environmental Epidemiology (2016) **26**, 520–528; doi:10.1038/jes.2016.29; published online 18 May 2016

Keywords: empirical/statistical models; epidemiology; exposure modeling; particulate matter

INTRODUCTION

Modern epidemiological studies focusing on the association between long-term exposure to fine particulate matter (PM_{2.5}) and health rely on predictions of PM_{2.5}, because PM_{2.5} measurements at each individual's location are infeasible. One of the common prediction approaches includes assigning average PM_{2.5} concentrations of monitors within an administrative unit to participants residing in that area.¹ Other approaches assign measurements on the basis of the monitor nearest to a participant's home or apply weighted averages on the basis of inverse distances.^{2,3} These approaches, however, do not represent well all the spatial variability in the underlying exposure surface; this in turn results in exposure measurement error in the health effect analysis. Recent advances in prediction models have led to better representation of variability in PM_{2.5} concentrations across cohort locations than relatively simple and commonly used approaches. For instance, land use regression models used geographic variables that affect the spatial variation of long-term average PM_{2.5} concentrations.^{4,5} More sophisticated spatio-temporal models, based on shorter-term average concentrations over 2 weeks or a month, characterize spatial and temporal variability using regression and smoothing techniques.^{6–9}

Study of the health effects of long-term concentrations of PM_{2.5} chemical components has been limited. Most studies of PM_{2.5} components have investigated associations of short-term concentrations.^{10–12} Thus, few prediction models for PM_{2.5} components have been developed. Ostro *et al.*¹³ investigated long-term associations of eight PM_{2.5} components and mortality in the California Teachers Study based on the nearest-monitor approach. Sun *et al.*¹⁴ adopted area-averaging, nearest-monitor, and inverse-distance-weighting methods to predict four PM_{2.5} components to examine the associations with subclinical atherosclerosis in the Multi-Ethnic Study of Atherosclerosis (MESA). Because some PM_{2.5} components, such as elemental and organic carbon (EC and OC) are affected largely by local sources such as traffic, it is likely that these simple prediction approaches provide poor predictions, particularly when distant monitors were used. A recent study of eight trace elements from PM_{2.5} in 20 European cities for the European Study of Cohorts for Air Pollution Effects demonstrated good capacity to represent local-scale spatial variability based on land use regression.¹⁵ The National Particle Component and Toxicity (NPACT) study at the University of Washington focused on PM_{2.5} components and investigated the association with cardiovascular outcomes in the MESA cohort.¹⁶ This study aimed to develop two distinct exposure models based

¹Institute of Health and Environment, Seoul National University, Seoul, Korea; ²Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA; ³Department of Biostatistics, University of Washington, Seattle, Washington, USA; ⁴Department of Mathematics and Statistics, Winona State University, Winona, Minnesota, USA; ⁵Department of Statistics, University of Washington, Seattle, Washington, USA; ⁶Department of Medicine, University of Washington, Seattle, Washington, USA and ⁷Department of Epidemiology, University of Washington, Seattle, Washington, USA. Correspondence: Dr. Sun-Young Kim, Institute of Health and Environment, Seoul National University, Room 717, Building 221, 1, Gwanak-ro, Gwanak-gu, Seoul 08826, Korea. Tel: +82 2 880 2759. Fax: +82 2 762 9105.

E-mail: puha0@snu.ac.kr

Received 12 November 2015; accepted 2 April 2016; published online 18 May 2016

on two different data sources to predict PM_{2.5} component concentrations at participant homes: the spatio-temporal and national spatial models. The spatio-temporal model was constructed based on 2-week average concentrations in each city area, whereas the national spatial model was constructed relying on long-term average concentrations in the continental U.S.

Here, we examine the prediction ability of the spatio-temporal modeling approach for long-term concentrations of PM_{2.5} components for the MESA cohort and compares its performance to that of the national spatial model previously developed for the same cohort.¹⁷ We focused on four PM_{2.5} components: EC, OC, sulfur, and silicon as roughly reflecting combustion-related traffic emissions, primary and secondary organic aerosol, secondary inorganic aerosol, and airborne crustal matter, respectively.

MATERIALS AND METHODS

Data

NPACT/MESA Air monitoring data. The NPACT study obtained 2-week integrated samples of PM_{2.5} chemical component measurements from the MESA and Air Pollution (MESA Air) study monitoring campaign.^{18,19} This campaign concentrated on the geographic areas covered by the MESA subject residences in each of the six MESA city regions: Los Angeles, Chicago, Minneapolis-St. Paul, Baltimore, New York, and Winston-Salem (Supplementary Figure S1). Three to seven fixed sites operated for the entire study period, whereas ~50 rotating home-outdoor sites were sampled in each of two seasons. Although the NPACT/MESA Air monitoring sites are located where most MESA participants live, there were very few regulatory monitoring sites near these subjects (Supplementary Figure S1). Although NPACT sampled for trace elements, including sulfur and silicon, between August 2005 and August 2009, sampling for EC and OC was limited to March 2007 through August 2008. Supplementary Figure S2 shows the sampling design of fixed and home-outdoor sites in Los Angeles; similar patterns hold for all MESA cities.

We summarize the sampling and analysis methods here; details can be found in Vedal *et al.*¹⁶ The NPACT/MESA Air monitoring campaign collected 2-week samples of PM_{2.5} components using the Harvard Personal Environmental Monitors with a 2.5 μm cut size when operated with pump-flow rate of 1.8 l/min. Sulfur and silicon were quantified by the X-ray fluorescence analysis of Teflon filters. EC and OC were determined by the IMPROVE_A thermal optical reflectance method from quartz filters. All data used in this analysis passed strict data cleaning and quality assurance criteria. In addition, we excluded seven measurements based on technician flags and evidences of filter contamination. We additionally excluded 5–15 outlying measurements, depending on the component, because they exceeded 2.5 times the inter-quartile range distance from temporally and spatially defined quartiles in each city.¹⁶ Those measurements dramatically affected model fitting and evaluation in our preliminary analysis. Then we added 1 and log-transformed the 2-week average measurements to meet normality assumption. Silicon was modeled in nanograms per cubic meter, whereas other components were in micrograms per cubic meter.

Regulatory monitoring data. There are two nation-wide regulatory monitoring programs for PM_{2.5} components: the U.S. Environmental Protection Agency (EPA) Chemical Speciation Network (CSN) and the Interagency Monitoring of Protected Visual Environments (IMPROVE). CSN monitoring sites are located mostly in urban areas and have collected PM_{2.5} components on an every 3rd or 6th day schedule since 1999.²⁰ The IMPROVE program deployed most monitoring sites in national parks and rural areas, and have sampled every 3rd day since 1987.²¹ The sampling and analysis protocols of these two networks were described elsewhere.^{20,21} We initially planned to combine the CSN and IMPROVE data with the NPACT/MESA Air monitoring data to develop our spatio-temporal models. However, we found that there were important differences between the two sets of networks due to the different sampling protocols and filter analysis methods for carbon. For example, we previously showed fair or poor correlations (0.27–0.62) of co-located EC measurements between the two networks in six cities.²² We concluded that the data were not sufficiently consistent between the two networks to justify combining all available data into one unified model. Instead, we used the NPACT/MESA Air monitoring data for the spatio-temporal model and the CSN and IMPROVE monitoring data for the national spatial model. For the national spatial model, daily measurement data was downloaded from the EPA Air

Quality System database for both CSN and IMPROVE for 2009 and 2010. We computed annual averages at sites where the data were at least two-thirds complete for a year and consecutive measurements were available for no >45 days. Then, we square-root transformed these to reduce skewness.

Geocoding and geographic variables. Residential addresses of 7014 MESA and MESA Air participants who consented to use of their addresses were geocoded using TeleAtlas 2000 according to standardized procedures. Geocoded locations of the NPACT/MESA Air and agency monitoring sites were obtained from geocoding and hand-held GPS devices, and EPA sources, respectively.

We created more than 800 candidate geographic variables at monitoring and cohort locations (Supplementary Table S1) to be used for both models. These variables included population density, vegetative index, impervious surface, types of land use, elevation, emissions of primary pollutants, and proximity to and density of road networks. We preprocessed these covariates, eliminating those that did not vary across locations, log transforming distance variables, and recoding distance variables by truncating at 25 km to avoid implausible extreme values. The preprocessing was applied by city area for the spatio-temporal model and nationally for the national spatial model. After this area-specific data processing, the number of candidate geographic variables in each area ranged between 52 and 116.

Exposure Prediction Model

Spatio-temporal model framework. We developed separate models for 2-week average log concentration measurements in each region and for each component. Our spatio-temporal modeling approach was based on the MESA Air study framework, previously described for PM_{2.5} and NO_x.^{7,8,23,24} The MESA Air spatio-temporal model assumed that log 2-week averages of PM_{2.5} are composed of three features: spatially varying long-term means, spatially varying temporal trends, and spatially varying and temporally independent spatio-temporal residuals. Because the component models relied on much less monitoring data not supplemented with the regulatory monitoring data, we used a simplified version of the MESA Air spatio-temporal model with a single temporal trend characterized by a simple spatial structure. Although we aimed to estimate long-term concentrations, we developed the spatio-temporal model on the 2-week scale in order to most effectively take advantage of the spatially rich but temporally sparse NPACT/MESA Air monitoring data based on long time-series for 4 years at 3–5 fixed sites and 1–3 temporal measurements at about 100 home-outdoor sites.

The spatio-temporal model for PM_{2.5} components represents the log 2-week average component concentration ($C(s,t)$) in terms of a long-term mean ($\beta_0(s)$), a temporal trend ($\beta_1(s)f(t)$), and spatio-temporal residuals ($\epsilon(s,t)$), shown in the equation below.

$$C(s, t) = \beta_0(s) + \beta_1(s)f(t) + \epsilon(s, t)$$

$$\beta_0(s) \sim \left(\alpha_{00} + \sum_{j=1}^{j=m} \alpha_{j0} X_{j0}(s), \Sigma(\phi_0, \sigma_0^2, \tau_0^2) \right)$$

$$\beta_1(s) \sim (\alpha_{01} + \alpha_{11} X_{11}(s), \tau_1^2)$$

$$\epsilon(s, t) \sim (0, \Sigma(\phi_\epsilon, \sigma_\epsilon^2, \tau_\epsilon^2))$$

The long-term mean and temporal trend vary spatially with a trend coefficient ($\beta_1(s)$) scaling the spatially-constant temporal basis function ($f(t)$). The temporal basis function was estimated by smoothing the first temporal component of a singular value decomposition of the space-time monitoring data matrix. The long-term mean was characterized by a universal kriging model with a land-use regression mean model and spatial correlation modeled with an exponential covariance function.²⁵ The covariance function had parameters for the range (ϕ), partial sill (σ^2), and nugget (τ^2), which represent the spatial correlation distance, spatial variability, and non-spatial variability, respectively. Geographic covariates for the long-term mean ($X_{j0}(s)$) were selected from a subset identified by the least absolute shrinkage and selection operator (lasso) followed by an exhaustive search.²⁶ The spatially varying trend coefficient was modeled by the one geographic variable most associated with the trend coefficient; its variance model had no spatial structure (i.e., zero range and partial sill). The spatio-temporal residual field was assumed to be temporally independent with mean zero and spatially correlated with an exponential covariance model.

Spatio-temporal model fitting and prediction procedure. Estimation of the temporal basis function was restricted to the PM_{2.5} component data at fixed sites. To determine the set of geographic variables to be included in the long-term mean, we performed the variable selection using data from home-outdoor sites for provisionally-computed long-term averages after removing a temporal trend. The geographic variables were rescaled to have common mean and unit variance. We selected twelve candidate variables from the lasso and then chose the final set of up to five (for sulfur and silicon) or four (for EC and OC) based on fivefold cross-validated R² in an exhaustive search. This selection approach aimed to maximize prediction ability for PM_{2.5} components rather than to identify the associations of the geographic variables with the PM_{2.5} components. Given the estimated temporal basis function, selected geographic variables and monitoring data, we estimated regression and covariance parameters. For the model evaluation, we performed 10-fold cross-validation for 2-week average measurements across home-outdoor sites and computed summary statistics such as root mean square error (RMSE) and R-squared statistic (R²). To focus on the spatial prediction ability of our spatio-temporal models, we computed temporally-adjusted R² statistic adjusting for temporal variability in addition to the usual (unadjusted) R² as shown in the equation below.

$$\text{Traditional } R^2 = \max \left(0, \frac{E \left[(C(s, t) - \hat{C}(s, t))^2 \right]}{E \left[(C(s, t) - \bar{C}(\cdot, \cdot))^2 \right]} \right)$$

$$\text{Temporally-adjusted } R^2 = \max \left(0, \frac{E \left[(C(s, t) - \hat{C}(s, t))^2 \right]}{E \left[(C(s, t) - \bar{C}_{\text{fixed}}(\cdot, t))^2 \right]} \right)$$

The temporally-adjusted R² accounted for temporal variability using either an estimated trend based on fixed sites or spatial averages of fixed sites at each time.^{16,24}

We predicted log 2-week average concentrations at participant addresses conditional on the estimated spatio-temporal model parameters and geographic covariates. These were exponentiated and 1 was subtracted to obtain 2-week predictions on the native scale. We also computed the unit of silicon back to the original microgram per cubic meter units. We restricted the prediction area to participants living within 10 km of any NPACT/MESA Air monitors to avoid extrapolation. In addition, we excluded a few extremely high or low predictions at addresses where covariate values for a particular geographic variable were far outside the range across the entire monitoring locations. Finally, we averaged the 2-week average predicted concentrations for 1 year from May 2007 to April 2008 when all four component data are available. Our spatio-temporal models were implemented in the R package SpatioTemporal on the Comprehensive R Archive Network.²⁷

National spatial model. We briefly summarize the previously developed national spatial modeling approach based on annual averages of PM_{2.5} component concentrations from the CSN and IMPROVE monitoring

networks; for more detail see Bergen et al.¹⁷ Instead of variable selection, this model adopted partial least squares (PLS) to estimate reduced numbers of spatial predictors, called PLS scores, from a large set of geographic variables.^{7,28} The first few PLS scores were used to characterize the mean structure in a universal kriging model. Two PLS scores were selected for all components except for EC with three based on 10-fold cross-validation. Given selected PLS scores, we estimated regression coefficients and covariance parameters. Although the national spatial model was developed over the entire U.S., we evaluated it in the MESA areas using CSN and IMPROVE monitoring sites within 200 km of the centers of the six MESA cities to be comparable to the spatio-temporal model. Finally, we predicted annual average concentrations for the PM_{2.5} components at MESA participant addresses in the same prediction area (i.e., within 10 km of any NPACT/MESA Air monitors) and back transformed these to the original microgram per cubic meter units.

RESULTS

NPACT/MESA Air Monitoring Data

Table 1 shows the summary statistics of 2-week concentrations for four PM_{2.5} components in each of the six MESA regions from the NPACT/MESA Air monitoring network. Sulfur concentrations were high in the cities on the East Coast, although those of EC were high in highly-populated cities such as Los Angeles and New York. Silicon concentrations were high in Los Angeles as expected given the dry climate contributing to lifting dust.

Spatio-Temporal Model Fitting

Trend estimation. Figure 1 shows the computed singular value decomposition and trend function for log-transformed PM_{2.5} components in Los Angeles. The results for the other five cities are shown in Supplementary Figure S3. Sulfur generally showed a clear seasonal pattern in all six cities. The seasonal pattern of EC and OC was seen in Los Angeles with higher EC in summer but higher OC in winter.

Variable selection. Supplementary Tables S2 and S3 give the classes of geographic variables included in the final selected models for each component and area from the potential variables described in Supplementary Table S1. For most pollutants and areas, the final models included traffic variables and urban and rural land use characteristics; inclusion of geographic coordinates, distances to sources, emission variables, vegetation, imperviousness, and elevation varied across PM_{2.5} components and areas. The variable selection cross-validated R²s using selected variables for the regression of “long-term average” PM_{2.5} component concentrations are also shown in Supplementary Table S3. They

Table 1. Summary statistics of 2-week concentrations of four PM_{2.5} components in the NPACT/MESA Air monitoring network.

City	Type	Sulfur			Silicon			EC			OC		
		N of sites	N of samples	Mean (SD) (µg/m ³)	N of sites	N of samples	Mean (SD) (µg/m ³)	N of sites	N of samples	Mean (SD) (µg/m ³)	N of sites	N of samples	Mean (SD) (µg/m ³)
Los Angeles	Fixed	7	535	1.15 (0.59)	7	536	0.16 (0.08)	7	200	1.81 (0.79)	7	200	2.15 (1.06)
	Home	89	153	1.08 (0.62)	108	172	0.15 (0.08)	70	88	1.79 (0.87)	70	87	2.24 (1.03)
Chicago	Fixed	5	375	1.12 (0.44)	5	374	0.11 (0.04)	5	138	1.38 (0.39)	5	138	1.81 (0.63)
	Home	104	187	1.09 (0.36)	89	152	0.10 (0.06)	50	80	1.27 (0.32)	50	82	1.88 (0.62)
St. Paul	Fixed	3	257	0.73 (0.23)	3	256	0.11 (0.05)	3	93	0.87 (0.23)	3	95	1.71 (0.37)
	Home	104	187	0.70 (0.23)	104	187	0.11 (0.04)	54	89	0.79 (0.21)	54	90	1.70 (0.40)
Baltimore	Fixed	4	331	1.53 (0.62)	4	329	0.09 (0.04)	4	133	1.45 (0.52)	4	133	2.18 (0.71)
	Home	85	156	1.73 (0.67)	85	156	0.09 (0.05)	61	99	1.23 (0.35)	61	99	2.19 (0.89)
New York	Fixed	3	191	1.34 (0.56)	3	191	0.11 (0.05)	3	80	2.22 (0.93)	3	81	1.84 (0.74)
	Home	107	190	1.38 (0.57)	105	186	0.10 (0.05)	49	78	1.83 (0.77)	49	81	2.09 (0.71)
Winston-Salem	Fixed	4	352	1.51 (0.75)	4	352	0.09 (0.05)	4	105	1.07 (0.24)	4	105	2.55 (0.69)
	Home	92	177	1.71 (0.72)	92	177	0.11 (0.05)	47	84	1.05 (0.27)	48	86	2.75 (0.79)

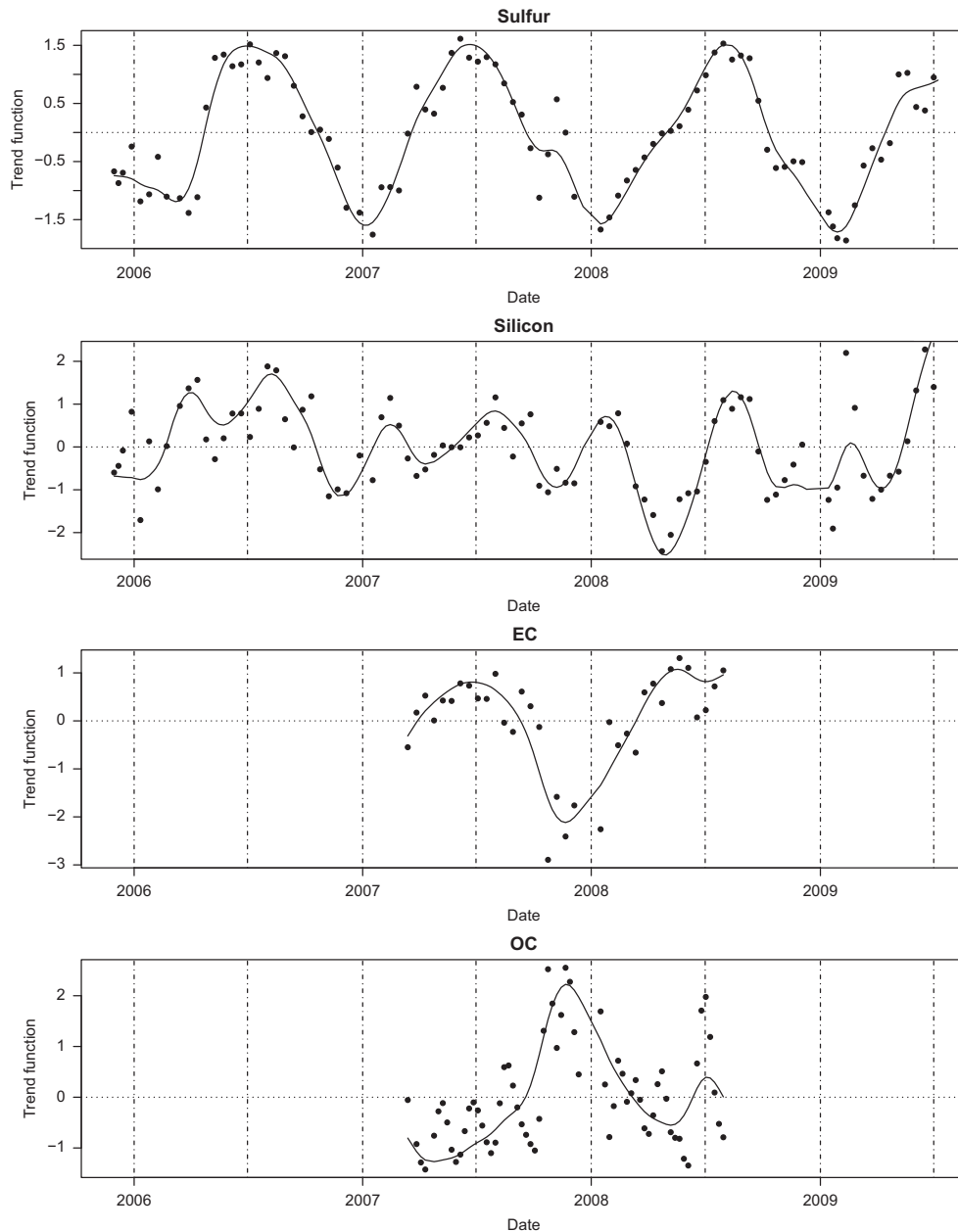


Figure 1. Estimated smooth temporal trends for four log-transformed PM_{2.5} components in Los Angeles.

were generally higher in all areas for EC and OC than for sulfur and silicon. Sulfur and silicon in St. Paul as well as New York and sulfur in Baltimore showed R^2 lower than 0.2. These low R^2 are possibly due to our conservative approach computing R^2 statistics, less spatial variability of sulfur and silicon, or absence of important geographic variables.

Parameter estimation. The estimates for the regression coefficients and variance model parameters in the six MESA cities are shown in Supplementary Figure S4. Los Angeles and Chicago tended to show larger range and partial sill representing stronger spatial correlation structure than other areas. In general, the estimated regression coefficients for EC and OC were significantly different from zero, whereas those for silicon and sulfur were not.

Features of Spatio-Temporal and National Spatial Models

Model evaluation. Table 2, Figure 2 and Supplementary Figure S5 show statistics and scatter plots for cross-validated predictions of 2-week concentrations from the city-specific spatio-temporal model across MESA home-outdoor sites and cross-validated predictions of annual averages from the national spatial model across the CSN/IMPROVE sites in the MESA areas. Not surprisingly, in the spatio-temporal predictions many of the temporally-adjusted R^2 s were much lower than the unadjusted R^2 s. Across all areas, the temporally-adjusted R^2 s, particularly when spatial averages were used, were generally higher for EC and OC than for sulfur and silicon. Los Angeles and Baltimore gave higher temporally-adjusted R^2 than other cities. Temporally-adjusted R^2 s for sulfur, silicon, EC and OC across all six cities combined were 0.84, 0.38, 0.79, and 0.59, respectively. These MESA-wide statistics

Table 2. Cross-validation statistics of predicted concentrations of four PM_{2.5} components between spatio-temporal and national spatial models in six MESA Air areas.

City	Pollutant	Spatio-temporal model ^a				National spatial model ^b	
		RMSE (µg/m ³)	R ²	Temporally-adjusted R ² ^c		RMSE (µg/m ³)	R ²
				Estimated trend	Average		
Overall ^d	Sulfur	0.19	0.92	0.84	0.82	0.05	0.94
	Silicon	0.03	0.61	0.38	0.28	0.04	0.45
	EC	0.32	0.75	0.79	0.79	0.21	0.70
	OC	0.42	0.75	0.59	0.55	0.28	0.79
LA	Sulfur	0.11	0.97	0.77	0.35		
	Silicon	0.03	0.68	0.66	0.49		
	EC	0.50	0.73	0.54	0.51		
	OC	0.59	0.67	0.49	0.37		
Chicago	Sulfur	0.19	0.74	0.54	0.15		
	Silicon	0.04	0.35	0.07	0.00		
	EC	0.18	0.69	0.51	0.49		
	OC	0.44	0.48	0.20	0.20		
Minneapolis-St. Paul	Sulfur	0.05	0.94	0.78	0.59		
	Silicon	0.03	0.65	0.39	0.19		
	EC	0.14	0.57	0.33	0.32		
	OC	0.15	0.85	0.47	0.46		
Baltimore	Sulfur	0.13	0.96	0.77	0.48		
	Silicon	0.00	0.82	0.58	0.35		
	EC	0.22	0.62	0.56	0.59		
	OC	0.33	0.86	0.36	0.35		
NY	Sulfur	0.30	0.71	0.12	0.00		
	Silicon	0.03	0.33	0.27	0.36		
	EC	0.65	0.15	0.58	0.52		
	OC	0.48	0.46	0.63	0.57		
Winston-Salem	Sulfur	0.24	0.89	0.41	0.09		
	Silicon	0.03	0.77	0.14	0.04		
	EC	0.19	0.48	0.19	0.18		
	OC	0.42	0.72	0.13	0.12		

^aCity-specific model using 2-week concentrations on log scale. ^bNation-wide model using annual average concentrations on square-root scale. ^cAdjusted-temporal trend was defined by two approaches, which are unsmoothed-temporal trend estimated using measurements across NAPCT/MESA Air fixed sites and average of measurements across fixed sites at each time; higher temporally adjusted R²s than traditional R²s indicate the over-adjustment of the temporal trend. ^dEvaluation of the national spatial model was restricted to regulatory monitoring sites within 200 km from the centers of six MESA cities; city-specific evaluation was not carried out given limited numbers of regulatory monitoring sites in each city area.

were generally higher than the city-specific temporally-adjusted R²s. R²s for sulfur, silicon, EC and OC in the national spatial model were 0.94, 0.45, 0.70, and 0.79, respectively.

Predicted long-term PM_{2.5} component concentrations. The city-specific summaries and spatial distributions of predicted long-term PM_{2.5} component concentrations varied by component and city (Supplementary Table S4, Figure 3 and Supplementary Figures S6–S8). Figure 3 and Supplementary Figures S6–S8 display maps of predicted long-term average concentrations of four PM_{2.5} components from the spatio-temporal model in each city region. Prediction maps for EC show local spatial heterogeneity with higher concentrations close to large roads and in largely populated or commercial areas, although this pattern somewhat varied depending on the city. In contrast, concentration surfaces for sulfur were smoothed at a large spatial scale. Predicted concentrations at MESA Air participant homes were generally higher from the spatio-temporal model than from the national spatial model, particularly for sulfur and EC. The two sets of predictions for all components except OC varied more between cities than within each city, with both between- and within-city variability being larger from the spatio-temporal model than from the national spatial model (Figure 4). In addition, predictions from the two sets of models were positively correlated across cities for all components, but with much lower correlation for OC, despite

comparable overall means, than for other components (correlation coefficients are 0.91, 0.55, 0.82, and 0.19 for sulfur, silicon, EC, and OC, respectively). These correlations across cities were higher than those within cities.

DISCUSSION

This study developed the spatio-temporal exposure prediction model to obtain long-term average residential concentrations of four PM_{2.5} chemical components at participant addresses, specifically for application in an epidemiological study. The model performance of four components in each of six cities was good with some exceptions for a few cities and pollutants. The spatial prediction ability was generally better for EC and OC than for sulfur and silicon. The spatio-temporal model was based on different monitoring data and modeling approach from those in our national spatial model previously developed for PM_{2.5} components. We, however, found generally consistent model performance across the six MESA cities driven by the large between-city variability of PM_{2.5} components; predicted long-term concentrations of PM_{2.5} components from the two models were fairly or highly correlated across cities. In contrast, the predictions are less highly correlated within each city.

We developed rich exposure prediction models in order to reduce measurement error in predicted individual-level

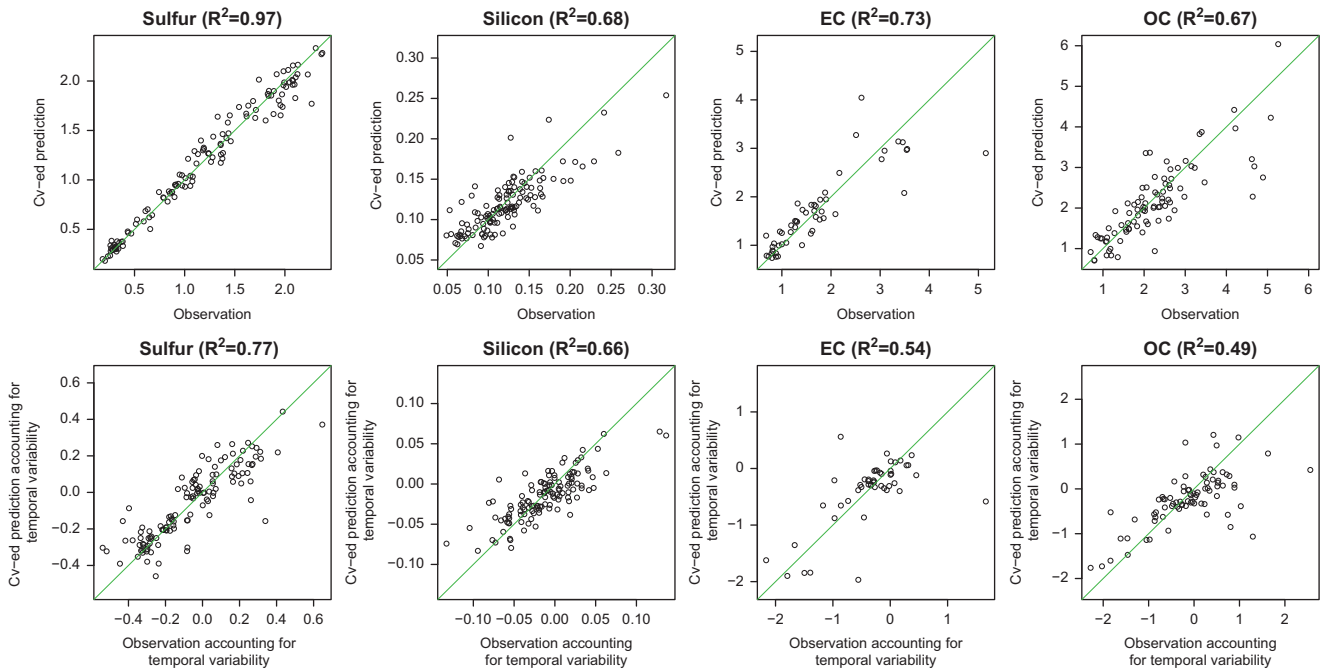


Figure 2. Component-specific scatter plots of observations and cross-validated predictions from the spatio-temporal model for 2-week average concentrations (top) and for 2-week average concentrations after accounting for temporal variability (bottom) across home-outdoor sites in Los Angeles.

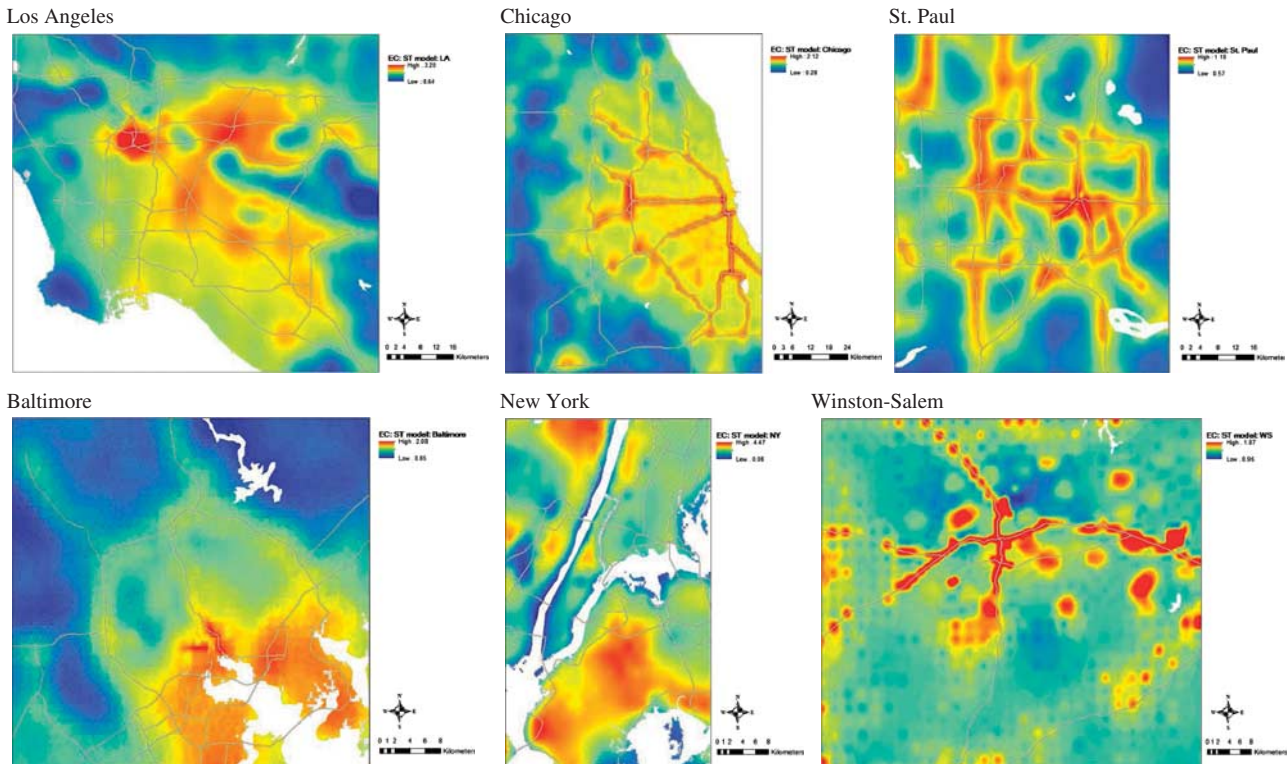


Figure 3. Maps of predicted long-term concentrations of EC ($\mu\text{g}/\text{m}^3$) from the spatio-temporal model in the six MESA city areas.

concentrations and then to provide more valid and precise health effect estimates. To our knowledge, this study is the one of a few studies focusing on the development of exposure prediction approaches for $\text{PM}_{2.5}$ components. Most previous cohort studies

assessed health effects of long-term $\text{PM}_{2.5}$ component concentrations using relatively simple prediction approaches such as area-averaging and nearest-monitor methods in representing spatial distribution.^{1,13,29} These approaches, however, could have

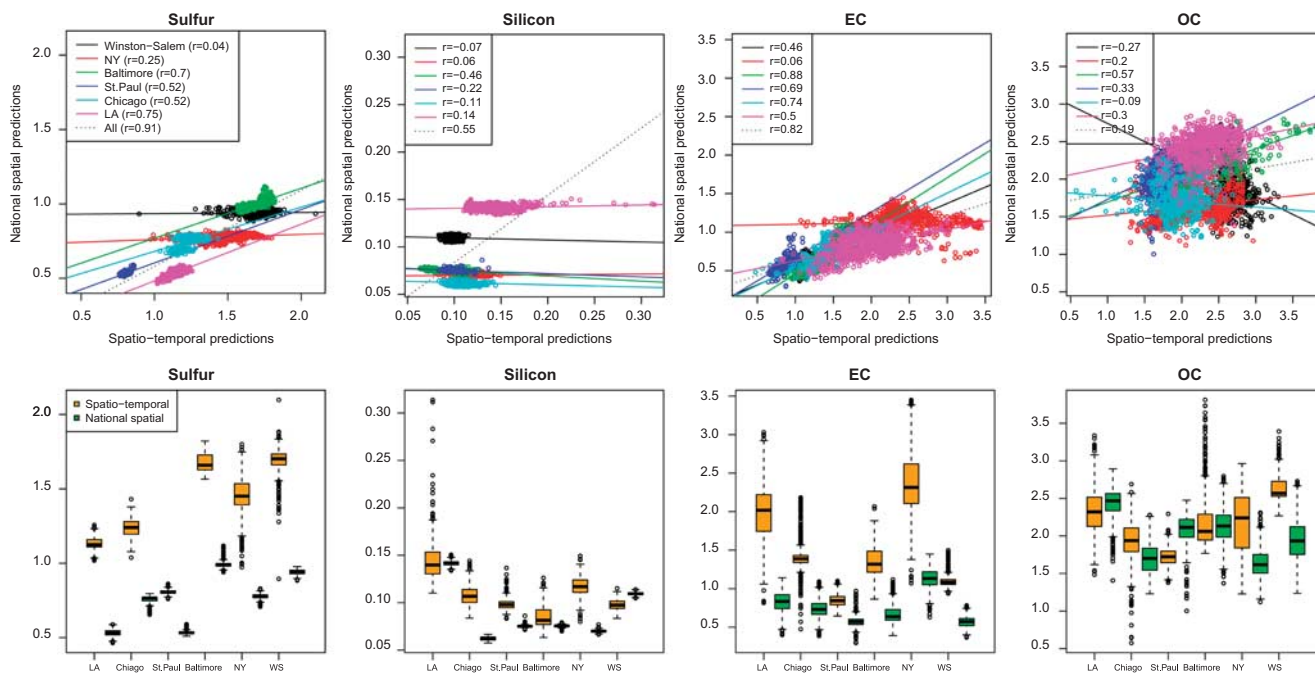


Figure 4. Component-specific scatter plots and box plots for spatio-temporal and national spatial model predictions of long-term average concentrations across the six MESA city areas.

high-exposure measurement error given spatially-limited regulatory monitoring networks, which do not represent fine-scale spatial heterogeneity of $PM_{2.5}$ components. This measurement error could then affect inference in the health effect analysis. We have shown by simulation that nearest-monitor predictions give more biased health effect estimates than kriging when the underlying pollution field has spatial structure.³⁰ Sun *et al.*¹⁴ investigated the association with subclinical atherosclerotic outcomes using simple prediction approaches based on the same NPACT/MESA Air monitoring data used in our spatio-temporal model. Supplementary Figure S9 showed that these predictions were highly correlated with predictions from the spatio-temporal model across cities but present little or no within-city variability. Recently, de Hoogh *et al.*¹⁵ adopted city-specific land use regression on long-term concentrations of eight trace elements of $PM_{2.5}$ across 20 monitoring sites in each of 20 European cities. Their approach is similar to our provisional approach for variable selection in the spatio-temporal modeling procedure. However, their cross-validated R^2 s for sulfur and silicon were generally higher than ours. This difference may be due to their approach to compute R^2 statistic. Their R^2 was computed based on the leave-one-out cross-validation, which can overestimate model performance particularly given a small number of sites.³¹

We chose highly conservative approaches using cross-validated and MSE-based R^2 in evaluating our two exposure prediction models to avoid overestimating model performance. One of the common evaluation approaches in land-use regression studies is the leave-one-out cross-validation.⁵ However, this approach was overly optimistic for model performance particularly when studies include limited numbers of monitoring sites, which make it difficult to adopt other evaluation methods.^{31,32} Given about a hundred home-outdoor sites, our cross-validation was based on the 5 or 10 group cross-validation. In addition, we computed MSE-based R^2 by subtracting mean square prediction error relative to data variability from 1 as opposed to model-based R^2 calculated by the squared correlation coefficient. The model-based R^2 tended to overestimate prediction ability because observations are compared to predictions based on the regression line instead of

the identity line as in MSE-based R^2 .³¹ Our evaluation approach using MSE-based R^2 in the 5- or 10-fold cross-validation was likely to provide more reasonable but relatively lower R^2 s than those reported in other studies.

Within-city predicted concentrations and variability of $PM_{2.5}$ components were generally higher from the spatio-temporal model than those from the national spatial model. This high within-city variability, assuming it reflects true exposure variation, can provide health effect estimates with increased precision and allows better assessment of the health effects across people residing within a single city. Features contributing to these within-city differences between the two exposure prediction models include the data sources, modeling approaches, and evaluation methods. Whereas the spatio-temporal model was developed based on the NPACT monitoring data, the national spatial model relied on the CSN/IMPROVE monitoring data. Higher predictions for sulfur, silicon, and EC in the spatio-temporal model compared to those in the national spatial model correspond to higher concentrations measured at NPACT/MESA Air monitoring sites relative to those at CSN/IMPROVE sites (Supplementary Figure S10). These differences could be due to there being many more NPACT/MESA Air sites located in central urban areas with higher concentrations. The number of monitoring sites in NPACT/MESA Air within 200 km of six MESA city areas was 92–123 compared to 7–32 in CSN/IMPROVE.²² In addition, operationally the land use information was incorporated differently in the two models. The spatio-temporal model relied on variable selection techniques to choose a subset of geographic variables. We focused on predictive performance in our variable selection instead of scientific interpretability because we considered that our limited monitoring data were unlikely to support identification of scientifically meaningful associations with geographic variables. Using all available geographic information with a dimension reduction technique as used in the national spatial model could be an alternative option. However, we note that the spatio-temporal model variable selection was based on detrended provisional “long-term averages”; these were quite uncertain and thus limited our confidence in applying the PLS approach. Lastly,

we devised a temporally-adjusted R^2 to evaluate spatial prediction ability for the spatio-temporal model. In the national spatial model our use of long-term averages removed all temporal variability so the traditional R^2 only represents spatial performance. All these fundamental differences between the two prediction models make it difficult to directly compare their performance statistics and conclude that one model is preferable to the other.

Despite different data sources and modeling characteristics, the two exposure models also showed consistent features across pollutants and a similar ordering of low to high predicted concentrations across cities. The two prediction models presented relatively strong mean structures for EC and OC and prominent spatial dependence structure for sulfur and silicon. The proportion of the spatial variability represented by the long-term mean was larger for EC and OC in the spatio-temporal model, whereas temporal variability represented by the temporal trend was larger for sulfur and silicon (Supplementary Table S5). Similarly, the regression part of the national spatial model explained most of the variability for EC and OC.¹⁷ Predicted concentrations from the spatio-temporal model were higher in some cities than others; similar patterns applied for predictions from the national spatial models for all components except for OC. The inconsistent pattern of OC may reflect complex features of OC produced by local sources such as traffic as well as atmospheric processes driven by meteorology. The value of adding meteorological variables as spatio-temporal covariates could be assessed in future work.

This study includes some limitations and implications for future studies. In the NPACT study, we originally intended to characterize within-city distribution of PM_{2.5} components based on the dedicated monitoring campaign for the target cohort combined with additional regulatory data. However, our preliminary exploratory analysis led us to limit our analysis to only the NPACT data given its incompatibility with CSN/IMPROVE data.²² The simplified spatio-temporal model based on limited monitoring data, though the NPACT/MESA Air monitoring campaign provided much richer spatial data for PM_{2.5} components in urban areas than any other previous studies, did not allow us to represent all the spatial variation within each city and may affect the ensuing health effect analyses. In addition, we focused on the four PM_{2.5} components, which are considered as least ambiguous markers for pollution sources of our interest. We plan to expand our modeling approaches to other components treated as being strongly related to specific pollution sources.

We described two modeling approaches for predicting long-term concentrations of PM_{2.5} components. The city-specific spatio-temporal model performed well for components largely affected by local sources and in metropolitan areas. Both performed reasonably well across cities. Predictions were generally consistent across the six study areas except for organic carbon; consistency was relatively weak within each city. Relatively large within-city variability of predictions from the spatio-temporal model and similar patterns of between-city contrasts for the two models suggest that the spatio-temporal model may be preferable for epidemiological studies focusing on within-area associations whereas studies combining multiple areas could use either modeling approach. These predictions of PM_{2.5} components allow us to assess associations of long-term exposure to PM_{2.5} components and health.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was primarily supported by the National Particle Component Toxicity (NPACT) initiative funded by the Health Effects Institute (HEI) (Health Effects Institute 4749-RFA05), along with the Multi-Ethnic Study of Atherosclerosis and Air Pollution

by the U.S. Environmental Protection Agency (EPA) Science to Achieve Results program (STAR) research assistance agreement (RD 831697). This publication has not been formally reviewed by the EPA. The views expressed in this document are solely those of the University of Washington and the EPA does not endorse any products or commercial services mentioned in this publication. Additional support was provided by the National Institute of Environmental Health Sciences (NIEHS) (T32ES015459 and P50 ES015915), the U.S. EPA (RD 83479601 and CR-834077101-0), and the National Research Foundation of Korea (Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education: 2013R1A6A3A04059017).

REFERENCES

- Laden F, Schwartz J, Speizer FE, Dockery DW. Reduction in fine particulate air pollution and mortality: extended follow-up of the Harvard Six Cities study. *Am J Respir Crit Care Med* 2006; **173**: 667–672.
- Lipsett MJ, Ostro BD, Reynolds P, Goldberg D, Hertz A, Jerrett M *et al*. Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort. *Am J Respir Crit Care Med* 2011; **184**: 828–835.
- Miller KA, Siscovick DS, Sheppard L, Shepherd K, Sullivan JH, Anderson GL *et al*. Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* 2007; **356**: 447–458.
- Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M *et al*. Development of land use regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* 2012; **46**: 11195–11205.
- Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P *et al*. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos Environ* 2008; **42**: 7561–7578.
- Paciorek CJ, Yanosky JD, Puett RC, Laden F, Suh HH. Practical large-scale spatio-temporal modeling of particulate matter concentrations. *Ann Appl Stat* 2009; **3**: 370–397.
- Sampson PD, Szpiro AA, Sheppard L, Lindström J, Kaufman JD. Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmos Environ* 2011; **45**: 6593–6606.
- Szpiro AA, Sampson PD, Sheppard L, Lumley T, Adar SD, Kaufman JD. Predicting intraurban variation in air pollution concentrations with complex spatio-temporal interactions. *Environmetrics* 2010; **21**: 606–631.
- Yanosky JD, Paciorek CJ, Suh HH. Predicting chronic fine and coarse particulate exposures using spatiotemporal models for the Northeastern and Midwestern United States. *Environ Health Perspect* 2009; **117**: 522–529.
- Bell ML, Dominici F, Ebisu K, Zeger SL, Samet JM. Spatial and temporal variation in PM_{2.5} chemical composition in the United States for health effects studies. *Environ Health Perspect* 2007; **115**: 989–995.
- Ostro B, Feng WY, Broadwin R, Green S, Lipsett N. The effects of components of fine particulate air pollution on mortality in California: results from CALFINE. *Environ Health Perspect* 2007; **115**: 13–19.
- Peng RD, Bell ML, Geyh AS, McDermott A, Zeger SL, Samet JM *et al*. Emergency admissions of cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. *Environ Health Perspect* 2009; **117**: 957–963.
- Ostro B, Lipsett M, Reynolds P, Goldberg D, Hertz A, Garcia C *et al*. Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. *Environ Health Perspect* 2010; **118**: 363–369.
- Sun M, Kaufman JD, Kim SY, Larson TV, Gould TR, Polak JF *et al*. Particulate matter components and subclinical atherosclerosis: common approaches to estimating exposure in a Multi-Ethnic Study of Atherosclerosis cross-sectional study. *Environ Health* 2013; **12**: 39.
- de Hoogh K, Wang M, Adam M, Badaloni C, Beelen R, Birk M *et al*. Development of land use regression models for particle composition in twenty study areas in Europe. *Environ Sci Technol* 2013; **47**: 5778–5786.
- Vedal S, Kim SY, Miller KA, Fox JR, Bergen S, Gould T *et al*. NPACT epidemiologic study of components of fine particulate matter and cardiovascular disease in the MESA and WHI-OS cohorts. Research Report 178. Health Effects Institute: Boston, MA, 2013.
- Bergen S, Sheppard L, Sampson PD, Kim SY, Richards M, Vedal S *et al*. A national prediction model for components of PM_{2.5} and measurement error corrected health effect inference. *Environ Health Perspect* 2013; **121**: 1017–1025.
- Cohen MA, Adar SD, Allen RW, Avol E, Curl CL, Gould T *et al*. Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environ Sci Technol* 2009; **43**: 4687–4693.
- Kaufman JD, Adar SD, Allen RW, Barr RG, Budoff MJ, Burke GL *et al*. Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease: the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Am J Epidemiol* 2012; **176**: 825–837.

- 20 U.S. EPA. *Air Quality Criteria for Particulate Matter (Report No. EPA 600/P-99/002aF-bF)*, vol. 1. U.S. Environmental Protection Agency: Washington, DC, 2004.
- 21 Hand JL, Copeland SA, Day DE, Dillner AM, Indresand H, Malm WC *et al*. Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States, Report V. Colorado State University, Fort Collins CO., 2011.
- 22 Kim SY, Sheppard L, Larson TV, Vedal S. Combining PM_{2.5} component data from multiple sources: data consistency and characteristics relevant to epidemiological analyses of predicted long-term exposures. *Environ Health Perspect* 2015; **123**: 651–658.
- 23 Keller JP, Olives C, Kim SY, Sheppard L, Sampson PD, Szpiro AA *et al*. A unified spatiotemporal modeling approach for prediction of multiple air pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. *Environ Health Perspect* 2015; **123**: 301–309.
- 24 Lindstrom J, Szpiro AA, Sampson PD, Oron AP, Richards M, Larson TV *et al*. A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environ Ecol Stat* 2013a; **21**: 411–433.
- 25 Banerjee S, Carlin BP, Gelfand AE. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC Press: Boca Raton, FL, 2004.
- 26 Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc Series B* 1996; **58**: 267–288.
- 27 Lindström J, Szpiro AA, Sampson PD, Bergen S, Oron AP. SpatioTemporal: Spatio-Temporal Model Estimation. R package version 1.1.1. 2012.
- 28 Sampson PD, Richards M, Szpiro AA, Bergen S, Sheppard L, Larson TV *et al*. A regionalized national universal kriging model using partial least squares regression for estimating annual PM_{2.5} concentrations in epidemiology. *Atmos Environ* 2013; **75**: 383–392.
- 29 Pope CA III, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K *et al*. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002; **287**: 1132–1141.
- 30 Kim SY, Sheppard L, Kim H. Health effects of long-term air pollution: influence of exposure prediction methods. *Epidemiology* 2009; **20**: 442–450.
- 31 Wang M, Beelen R, Eeftens M, Meliefste K, Hoek G, Brunekreef B. Systematic evaluation of land use regression models for NO₂. *Environ Sci and Technol* 2012; **46**: 4481–4489.
- 32 Wang M, Beelen R, Basagana X, Becker T, Cesaroni G, de Hoogh K *et al*. Evaluation of land use regression models for NO₂ and particulate matter in 20 European study areas: the ESCAPE project. *Environ Sci Technol* 2013; **47**: 4357–4364.

Supplementary Information accompanies the paper on the *Journal of Exposure Science and Environmental Epidemiology* website (<http://www.nature.com/jes>)