Environmental Science & Technology

Article

pubs.acs.org/est

Spatiotemporal Land Use Regression Models of Fine, Ultrafine, and Black Carbon Particulate Matter in New Delhi, India

Arvind Saraswat

Institute for Resources Environment and Sustainability The University of British Columbia, Rm 411, 2202 Main Mall, Vancouver, BC V6T 4T1, Canada

Joshua S. Apte

Energy and Resources Group, University of California, Berkeley, 310 Barrows Hall, Berkeley, California 94720, United States

Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 90R3022, Berkeley, California 94720, United States

Milind Kandlikar*

Liu Institute for Global Issues & Institute for Resources Environment and Sustainability, The University of British Columbia, Room 101B, 6476 NW Marine Drive, Vancouver, BC V6T 1Z2, Canada

Michael Brauer

School of Population and Public Health, The University of British Columbia, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada

Sarah B. Henderson

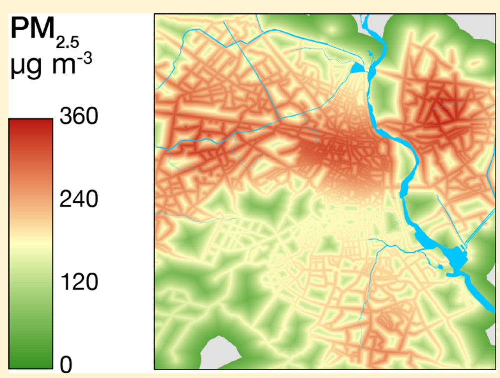
British Columbia Centre for Disease Control, 655 West 12th Avenue, Vancouver, BC V5Z 4R4, Canada

Julian D. Marshall

Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, Minnesota 55455, United States

Supporting Information

ABSTRACT: Air pollution in New Delhi, India, is a significant environmental and health concern. To assess determinants of variability in air pollutant concentrations, we develop land use regression (LUR) models for fine particulate matter ($PM_{2.5}$), black carbon (BC), and ultrafine particle number concentrations (UFPN). We used 136 h (39 sites), 112 h (26 sites), 147 h (39 sites) of $PM_{2.5}$, BC, and UFPN data respectively, to develop separate morning (0800–1200) and afternoon (1200–1800) models. Continuous measurements of $PM_{2.5}$ and BC were also made at a single fixed rooftop site located in a high-income residential neighborhood. No continuous measurements of UFPN were available. In addition to spatial variables, measurements from the fixed continuous monitoring site were used as independent variables in the $PM_{2.5}$ and BC models. The median concentrations (and interquartile range) of $PM_{2.5}$, BC, and UFPN at LUR sites were 133 (96–232) $\mu g\ m^{-3}$, 11 (6–21) $\mu g\ m^{-3}$, and 40 (27–72) $\times 10^3\ cm^{-3}$ respectively. In addition (a) for $PM_{2.5}$ and BC, the temporal variability was higher than the spatial variability; (b) the magnitude and spatial variability in pollutant concentrations was higher during morning than during afternoon hours. Further, model R^2 values were higher for morning (for $PM_{2.5}$, BC, and UFPN, respectively: 0.85, 0.86, and 0.28) than for afternoon models (0.73, 0.69, and 0.23); (c) the $PM_{2.5}$ and BC concentrations measured at LUR sites all over the city were strongly correlated with measured concentrations at a fixed rooftop site; (d) spatial patterns were similar for $PM_{2.5}$ and BC but different for UFPN; (e) population density and road variables were statistically significant predictors of pollutant concentrations; and (f) available geographic predictors explained a much lower proportion of variability in measured $PM_{2.5}$, BC, and UFPN than observed in other LUR studies, indicating the importance of temporal variability and suggesting the existence of uncharacterized sources.



Received: April 5, 2013
Revised: September 29, 2013
Accepted: October 2, 2013

1. INTRODUCTION

Air pollution has been an environmental and health concern in New Delhi, the Indian capital, for decades. Fine particles ($PM_{2.5}$ or particulate matter with aerodynamic diameter $d_a \leq 2.5$ μm) have been consistently associated with adverse health outcomes.[1−4] Ambient $PM_{2.5}$ frequently exceeds regulatory standards in New Delhi, with annual average concentrations of 123 ± 87 μg m^{-3} between 2008 and 2011 across seven regulatory monitoring stations.[5] Those concentrations are an order of magnitude higher than the guideline value of 10 μg m^{-3} (annual mean) set by the World Health Organization.[6] For comparison, annual average $PM_{2.5}$ concentrations in New York City were around 11 μg m^{-3} in 2009[7] and ∼100 μg m^{-3} or in Beijing in 2010.[8] Black carbon (BC) and ultrafine particles (UFP, < 0.1 μm diameter) are constituents of $PM_{2.5}$. BC is an indicator of incomplete combustion, e.g., from biomass burning and motor vehicles.[9] Apart from adverse health effects due to inhalation, BC absorbs sunlight and is a climate forcing agent.[10] Elevated BC levels have been associated with regional climate effects across India and China.[11] Recent research also indicates that BC may be a more specific marker than $PM_{2.5}$ for health effects.[12] Sources of UFP include emissions from internal combustion engines, power plants, incinerators, forest fires, and cooking.[13] UFP have also been associated with adverse health effects.[13−15]

Ambient $PM_{2.5}$ pollution is understood to be one of the leading risk factors for premature mortality in South Asia, resulting in more than 750 000 premature deaths in 2010.[16] Poor air quality in New Delhi is an important public health issue for the city's 16.7 million people,[17−24] yet epidemiologic studies[25,26] have been limited by the quality, duration and spatial coverage of the urban air quality monitoring network. New Delhi has six continuous monitoring stations for $PM_{2.5}$. Better coverage may be required to assess the spatial variability of $PM_{2.5}$ and, hence, the variation in human exposure. In the absence of more detailed monitoring data, land use regression (LUR) models can be an effective tool for assessing within-city variability of air pollution.[27] LUR and results are useful for epidemiologic studies, risk assessments, and prioritizing air quality management.[28]

LUR was developed as an alternative to dispersion models and as a means to assess small-scale spatial variation of air pollutants within urban areas. LUR models have been developed for many cities in North America and Europe.[28,29] Models may be used to assign population exposure for epidemiologic studies. In brief, a statistical relationship is established between land use characteristics and pollutant concentrations measured by targeted sampling at a limited number of sites, and then, the relationship is used to predict pollutant concentrations at unmeasured locations throughout a given domain. Generally, LUR has been used to estimate mean annual concentrations of a pollutant, using one to two week sampling at 20−100 sites.[28] Land use, road network, population density, and traffic flow variables are typically used as inputs to the models, though LUR models have been reported to perform well even in the absence of traffic flow data as an input.[30] The site selection processes vary,[30−33] but in general a minimum number of spatially dispersed sites characterizing different land uses that are able to capture the spatial variation of a pollutant for a given domain are required. A small number of sites and a large number of predictor variables can lead to inflated R^2 values.[34]

Air pollution in Indian cities likely has a wider range of sources than air pollution in the European and North American cities where LUR has been previously applied: not just traffic and industry, but also small-scale, distributed sources such as biomass burning for cooking and heat, open burning of solid waste, and diesel generators for backup power.[35−37] In the case of New Delhi, there is also substantial seasonal variation in absolute concentrations of $PM_{2.5}$ and percentage contributions from different sources.[24,38,39]

There is also substantial variation in $PM_{2.5}$ concentrations during daylight hours (diurnal variation).[5] To capture this trend, we sought to develop separate LUR models for morning and afternoon periods. Specifically, $PM_{2.5}$ and BC models were developed to predict the spatial distribution of pollutants over time, using data from a fixed continuous monitoring site. The models are thus spatiotemporal rather than simply spatial. These models could be used to obtain estimates for any given time interval within the study period. For example, hourly concentration estimates can be obtained if we input hourly pollutant concentration in a model equation. As no continuous UFPN measurements were available, the UFPN models only describe spatial variability in the morning and afternoon hours for the duration of the study.

2. MATERIALS AND METHODS

2.1. Site Selection. We undertook field measurements of air quality in New Delhi during February−May 2010, encompassing the local spring and summer seasons. We employed a site selection approach similar to what Brauer et al.[40] employed for the TRAPCA study in three European cities with about 40 sites in each city. We used local knowledge, Google Earth, and city maps (1:5500 to 1:12 500) to classify neighborhoods based on the following criteria: population density, distance to the city center (Connaught Place), residential or commercial type, density of the road network, and green spaces. Sites were allocated to neighborhoods that captured maximum variation in these variables. We made minor adjustments to site locations during the monitoring campaign on the basis of preliminary site visits; adjustments were less than 10 m in magnitude and were made to ensure that a spot was available to leave monitoring devices undisturbed at normal breathing height for the sampling period.

2.2. Instrumentation and Field Measurements. Data collection occurred in parallel with fieldwork for a companion study of urban and in-vehicle exposure to particulate matter in New Delhi.[41] The air quality sampling instruments, protocols, and postprocessing techniques have been described previously.[41] Continuous monitoring of $PM_{2.5}$ and BC was conducted at a fixed rooftop site located in a high-income residential neighborhood in southern New Delhi (Figure 1). The rooftop provided a background site that was relatively free from the influence of local traffic and point sources, so that data from this site could be used to characterize citywide diurnal and seasonal trends. Sampling at LUR sites was divided into morning and afternoon sampling periods. At each site, measurements were collected for 1−3 h during the morning (0800−1200) and/or afternoon (1200−1800). Given equipment limitations, only one LUR site (plus the one fixed-location) was sampled at a time. Sampling at LUR sites was conducted close to normal breathing height, in contrast to regulatory monitoring which is often on rooftops.

2.2.1. GPS and Meteorological Data. We used a GPS device (GPSMap 60CSx, Garmin Inc.) with an accuracy of ±3−5 m to record the spatial coordinates of all sites at the time of data collection. We recorded meteorological data (temperature, relative humidity, wind speed, wind direction, and rainfall; all recorded at 5 min intervals) via a weather station (Model PWS-1000TD,

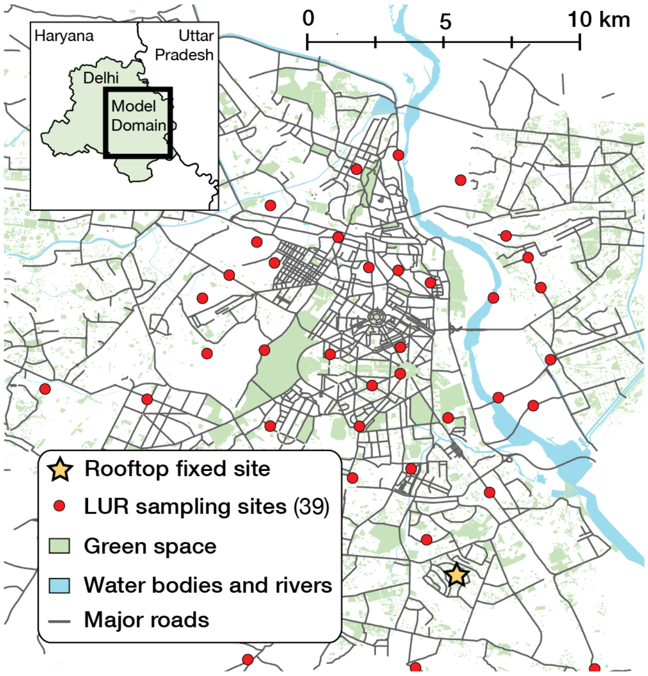


Figure 1. Land use regression sites and rooftop site overlaid on major road network.

Zephyr Instruments, East Granby, CT) at the fixed (i.e., central rooftop) location.

2.2.2. $PM_{2.5}$ Data Collection. Fine particulate matter was measured using two TSI DustTrak 8250 aerosol monitors (TSI Inc., Shoreview, MN, USA) fitted with $PM_{2.5}$ impactor inlets. The DustTrak infers particle mass concentrations based on 90° light scattering measured by a laser photometer. This detection method is subject to error because relative humidity (RH) and particle properties (i.e., density, shape, size, refractive index) influence particle light scattering.[41−43] To account for the RH effect, we applied an empirical correction equation[42] to the raw DustTrak measurements using 5 min average RH measurements from the rooftop fixed site. The regular sampling program of DustTrak measurements was supplemented with ∼35 colocated, time-integrated (∼1−4 h) gravimetric measurements of $PM_{2.5}$ collected with a single-stage impactor, from which a nonlinear gravimetric calibration curve was developed.[41] The final $PM_{2.5}$ mass determination was obtained by applying this gravimetric calibration to the time-resolved RH-corrected DustTrak observations.[44]

During each measurement session, $PM_{2.5}$ data collection was carried out simultaneously at the central rooftop location and at one of the LUR sites. $PM_{2.5}$ data were collected at 48 LUR sites (44 afternoon sessions, 22 morning sessions; Figure 1). Nine afternoon sites and three morning sites were later dropped from the analysis because population and road network data were not available in geographic information system (GIS) format at those locations. This could have been avoided if the availability of these data (in GIS format) was verified before the sampling. We employed a time resolution for measurements of 30 s (rooftop site) and 1 s (LUR sites).

2.2.3. Black Carbon Data Collection. Similar to $PM_{2.5}$, BC data collection was carried out simultaneously at the central rooftop location and at one of the LUR sites. Concentrations were measured using portable aethalometers (model AE-51 micro-Aeth, Magee Scientific, Berkeley, CA). A previously developed empirical correction factor[45] was validated for use in New Delhi[41] and then applied to raw BC data to correct for underestimation of BC concentrations with increasing aethalometer filter loadings.[45,46] Five-minute moving averages were used to remove sharp, consecutive negative, and positive concentrations peaks that can reflect measurement artifacts.[41,47] Overall, BC concentrations were measured at 30 LUR sites (29 afternoon sessions and 20 morning sessions). Four afternoon sessions and three morning sessions were dropped from the analysis because population and road network data were unavailable in GIS format at those locations.

2.2.4. UFPN Data Collection. Ultrafine particle concentrations were measured using a Condensation Particle Counter (CPC, model CPC 3007, TSI Inc., Shoreview, MN). Although this instrument provides the total number count of all particles in the size range 10 nm < d_p < 1 μm, this result closely approximates UFPN (d_p < 100 nm) under the conditions encountered in this study. Concentrations were measured at 1 Hz. Because UFPN concentrations often exceeded the upper measurement limit (10^5 particles cm^{-3}) of the CPC,[41] we employed a custom-built dilutor, which reduced inflow concentrations by a factor of 5.5. In addition, we applied the empirical correction factor of Westerdahl et al. (2005)[48] to account for particle coincidence errors when diluted UFPN concentrations exceeded 10^5 particles cm^{-3}.

Sampling for UFPN was conducted at LUR sites only, because a second CPC was not available. We measured UFPN concentrations at 48 LUR sites (46 afternoon sessions and 21 morning sessions). UFPN data for nine afternoon sites and two morning sites were dropped from the analysis because population and road network data were unavailable in GIS format.

2.2.5. Data Reduction and Quality Control. Processed data for all pollutants were used to obtain hourly medians for the LUR sites. Any hour with less than 15 min of data was discarded. The arithmetic mean of these medians for morning or afternoon hours was assigned to an LUR site as its concentration for morning or afternoon hours, respectively.

For the rooftop site, 10th percentile concentrations were computed for each hour, and any hour with less than 15 min of data was discarded. The 10th percentile was selected to approximate the urban background diurnal profile and to be free from the influence of short-duration peaks due to local sources. The arithmetic mean of the hourly 10th percentile concentrations corresponding to the sampling period at an LUR site was computed for each LUR site. The natural logarithm of this value was used as an independent variable in the model building processes for $PM_{2.5}$ and BC and is hereafter called ln(ROOF).

2.3. Spatial and Socioeconomic Variables. We used maps in a GIS to generate 14 spatial variables related to land use and demographics (Table 1). Shape files were converted into rasters of 5 m cells using ESRI ArcGIS 9.3. Separate rasters were created for major roads, minor roads, and green spaces. The road rasters were used to estimate road length around each LUR site for chosen radii and to estimate the shortest distance from each LUR site to the nearest major and minor roads. A green-space raster was used to obtain area of green-space around each LUR site for chosen radii.

We use Indian Census data to generate independent socioeconomic variables (Table 1) as possible surrogates for sources such as domestic wood and waste burning. Census data were based on the 2001 census by the Government of India, available at the ward level for New Delhi. A ward is an electoral unit for the local municipal government, with a total of 156 wards in New Delhi. Population attributes per ward were assigned to the ward centroids and the Spatial Analyst feature in ArcGIS was

dx.doi.org/10.1021/es401489h | Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Potentially Predictive Independent Variables

variable	description	median (IQR[a])
DIST.J1	shortest distance to the nearest major road from a cell (m)	65 (24–130)
DIST.J2	shortest distance to the nearest minor road from a cell (m)	37 (14–77)
RD1.100	major road length in a buffer of 100, 250, and 500 m radii, respectively (m)	200 (0–271)
RD1.250		672 (467–10^3)
RD1.500		2.2×10^3 (1.3×10^3–3.5×10^3)
RD2.100	minor road length in a buffer of 100, 250, and 500 m radii respectively (m)	247 (42–459)
RD2.250		1.3×10^3 (586–2.7×10^3)
RD2.500		5.6×10^3 (3.1×10^3–11.3×10^3)
GREEN.100	area of green space in a buffer of 100, 250, and 500 m radii, respectively (m^2)	0 (0–2.1×10^3)
GREEN.250		2.4×10^3 (0– 18.4×10^3)
GREEN.500		31.7×10^3 (981–98.2×10^3)
SC.5000	density of Scheduled Caste and Scheduled Tribe population (persons/hectare) obtained using density analysis with 5000 m radius	25 (17–44)
ILL.5000	density of illiterate population (persons/hectare) obtained using density analysis with 5000 m radius	50 (27–86)
POP.5000	density of population (persons/hectare) obtained using density analysis with 5000 m radius	177 (97–289)
ln(ROOF)	log of mean hourly 10th percentile concentrations from the rooftop site	4.4 (4–5) $PM_{2.5}$
		1.6 (1.1–2.7) BC

[a]Interquartile range.

used to obtain smoothed density surfaces[30] for the total population and for two indicators of low socioeconomic status: illiterate population and Scheduled Caste and Scheduled Tribe population. Scheduled Caste and Scheduled Tribe are defined in Articles 341 and 342, respectively, of the Indian Constitution and are generally considered to be socioeconomically disadvantaged groups.[49] In total, 14 variables were considered in LUR model development.

2.4. Model Building. Air pollutant concentrations are typically log-normally distributed.[50] We used the natural logarithm of the mean of measured hourly medians at each LUR site as the dependent variable in each model, following verification with the Kolmogorov–Smirnov test[51] in the R statistical software package.[52] Eligible independent variables included the land use and socioeconomic variables (Table 1) for all models and ln(ROOF) for the $PM_{2.5}$ and BC models.

For $PM_{2.5}$ and BC models, we assumed that the pollutant concentrations were associated with a multiplicative combination of a background temporal component (from the rooftop site) and the spatial components. We also assumed that (a) the temporal component was spatially invariant and (b) the spatial components were temporally invariant. The first assumption is supported by the existence of a strong temporal correlation between measurements at the LUR sites and the rooftop site. The second assumption is reasonable given that changes in the spatial variables would be small over the modeling period. For example, we do not expect that the patterns of population density changed over the 13-weeks of sampling.

We generated two temporal models per pollutant, each based on log hourly median concentrations: morning models (observations 0800–1200) and afternoon models (1200–1800). We used a model-building algorithm similar to the one developed by Henderson et al.,[30] which was designed to produce a parsimonious model in which the influence of individual variables is interpretable and consistent with a priori assumptions about determinants of spatial variability in air pollution. A statistical relationship was established between the observed pollutant concentrations at the LUR sites, the potentially predictive land use and socioeconomic variables, and the rooftop concentration variables (for $PM_{2.5}$ and BC). These relationships were used to predict pollutant concentrations throughout the spatial domain. The model-building algorithm is as follows:

1. Each variable was ranked based on its correlation with the log of hourly median concentration.
2. The top-ranking variable in each subcategory (e.g., major roads) was identified.
3. Variables in each subcategory that had a correlation of 0.6 or higher with the top-ranking variable (step 2) were dropped.
4. Remaining variables were included in robust linear regression models.
5. Variables that were not significant at a 90% confidence level or that had a coefficient with a counterintuitive sign were dropped.
6. Repeat steps 4 and 5 to convergence.

Steps 1 and 2 of the model-building algorithm provide ordered sets of the most predictive variables, and step 3 avoids collinearity between independent variables from the same subcategory. Diagnostic plots in step 4 showed a few high leverage points in almost all models. Robust linear regression was used to prevent the undue influence of the high leverage points on coefficient estimates.[53] Steps 4 and 5 were repeated to convergence, and the resulting models had those variables only that were statistically significant at 90% confidence level. Here, we present the regular linear regression coefficients and statistics for the independent variables selected using this model-building algorithm. Robust linear regression coefficients and statistics are presented as Supporting Information (Table S1). For models with more than one independent variable, we also report the square of partial correlation for each independent variable. Squared partial correlation reflects the contribution of an independent variable as a fraction of the model R^2 and can be used as an estimate of effect size for each individual variable.[54,55]

2.5. Model Evaluation. Models were evaluated with diagnostic plots and leave-one-out cross validation.[56] We used ESRI ArcGIS9.3 to obtain Moran's I statistic to check for presence of spatial autocorrelation[57] in the LUR models. For the $PM_{2.5}$ and BC models, Moran's I was calculated for every two week period, in the morning and afternoon models. For UFPN models, Moran's I was calculated for the morning and afternoon models for the complete duration of the study. All models were

dx.doi.org/10.1021/es401489h | Environ. Sci. Technol. XXXX, XXX, XXX–XXX

Table 2. Descriptive Statistics for Hourly Median Concentrations of Fine Particulate Matter ($PM_{2.5}$), Black Carbon (BC), and Ultrafine Particle Number (UFPN) at the Land Use Regression Sites

species	GM (GSD)	median	min	P_{10}	P_{25}	P_{75}	P_{90}	max	N (sites, h)	
$PM_{2.5}$ (μg m^{-3})	140 (1.8)	133	40	61	96	232	335	680	39	136
BC (μg m^{-3})	12 (2.6)	11	2	4	6	21	43	140	26	112
UFPN (10^3 cm^{-3})	43 (2.0)	40	7	18	27	72	113	190	39	147

Table 3. Final Spatiotemporal LUR Model Specifications and Results for Fine Particulate Matter ($PM_{2.5}$), Black Carbon (BC), and Ultrafine Particle Number (UFPN)

	model terms and coefficients: log-median concentration[a]						
model	constant	ln(ROOF)	POP.5000	DIST.J1	RD2.500	$R^{2\ b}$	LOOE[c]
$PM_{2.5}$ (morning)	1.22**	0.77**	2.0×10^{-3}**	-8.2×10^{-4}**		0.85	0.25
N = 19[d]		*(0.83)*[e]	*(0.53)*	*(0.32)*			
$PM_{2.5}$ (afternoon)	0.01	0.98**	1.5×10^{-3}**			0.73	0.29
N = 35		*(0.72)*	*(0.22)*				
BC (morning)	0.80**	0.73**	2.2×10^{-3}**	-7.2×10^{-4}		0.86	0.34
N = 17		*(0.85)*	*(0.41)*	*(0.15)*			
BC (afternoon)	0.37	0.83**	1.5×10^{-3}**	-1.0×10^{-3}		0.69	0.44
N = 25		*(0.67)*	*(0.14)*	*(0.11)*			
UFPN (morning)	10.35**		3.3×10^{-3}**			0.28	0.59
N = 18							
UFPN (afternoon)	9.79**		2.3×10^{-3}*		4.7×10^{-5}*	0.23	0.72
N = 37			*(0.10)*		*(0.09)*		

[a]Spatiotemporal LUR models predict the natural log of the hourly median concentration at each fixed site. [b]Model coefficient of determination (R^2). [c]Leave-one-out cross-validation error. [d]*N* is the number of sites. [e]Squared partial correlation. **Statistically significant at 95% confidence level. *Statistically significant at 90% confidence level.

also checked for temporal autocorrelation using Durbin−Watson test.[58]

2.6. Regression Mapping. We rendered the regression models as concentration maps for the study domain using the Spatial Analyst feature in ArcGIS. We used coefficient estimates obtained in the model building process to predict pollutant concentrations for each cell.

3. RESULTS

Median (interquartile range) concentrations for $PM_{2.5}$, BC, and UFPN across all LUR sites were, respectively, 133 (96−232) μg m^{-3} $PM_{2.5}$, 11 (6−21) μg m^{-3} BC, and 40 (27−72) × 10^3 particles cm^{-3} (Table 2). These pollutant concentrations are substantially higher than those reported in other LUR studies.[28] The correlation between the hourly median concentrations of $PM_{2.5}$ and the corresponding BC concentrations at LUR sites was 0.79. At the LUR sites the coefficient of correlation between the hourly median concentrations of $PM_{2.5}$ and UFPN was 0.42, and for BC and UFPN the correlation was 0.51. The coefficient of correlation between hourly medians at the fixed rooftop site and the corresponding hourly medians at the LUR sites was 0.84 for $PM_{2.5}$ and 0.85 for BC. Average temperatures recorded during the sampling campaign were as follows: 20 °C for February, 28 °C for March, 36 °C for April, and 37 °C for the first week of May.

Values of Moran's I confirmed the presence of spatial autocorrelation, such that sites with higher concentrations were more likely to be closer to other sites with higher concentrations. No evidence of temporal autocorrelation was found using the Durbin−Watson test. Monthly diurnal variation plots for $PM_{2.5}$ and BC have been included in the Supporting Information (Figure S1 and S2).

3.1. $PM_{2.5}$ Models. The morning $PM_{2.5}$ model (Table 3, Figure 2a) explained 85% of the variability in measured concentrations. The explanatory variables were population density, distance to nearest major road, and ln(ROOF), which accounts for citywide day-to-day and diurnal variation. Coefficient estimates indicate that the predicted concentrations increased with population density and decreased as the distance to nearest major road increased. The squared partial correlation indicated that the ln(ROOF) explained more variability than the land use variables.

The afternoon $PM_{2.5}$ model (Table 3, Figure 2a) was more poorly fit than the morning model. The explanatory variables were population density and ln(ROOF). To understand the relative importance of temporal and spatial variation, we dropped the ln(ROOF) term from both the morning and afternoon models. No useful models were obtained, suggesting a dominant role of temporal variation in predicting the $PM_{2.5}$ measurements.

3.2. Black Carbon Models. The morning BC model (Table 3, Figure 2b) explained 86% of the variability in measured BC concentrations. The explanatory variables were population density and ln(ROOF). The afternoon model (Table 3, Figure 2b) explained 69% of the variability in the measurements. For both morning and afternoon BC LUR models, distance from the nearest major road was chosen as model variable using the robust linear regression selection algorithm (Table S1), but its coefficients were not statistically significant in the regular linear regression models. That finding suggests that robust regression can improve the performance of LUR models. The morning BC LUR model was a better fit than the afternoon model. To understand the relative importance of temporal and spatial variation, we tried dropping the ln(ROOF) term from both BC models; as with $PM_{2.5}$, no useful models were obtained.

3.3. UFPN Models. Unlike the $PM_{2.5}$ and BC models, the UFPN models did not include measurements from the fixed rooftop site as an independent variable. The morning model (Table 3, Figure 2c) had population density as the only statistically significant predictor. The afternoon model (Table 3, Figure 2c)

dx.doi.org/10.1021/es401489h | Environ. Sci. Technol. XXXX, XXX, XXX−XXX

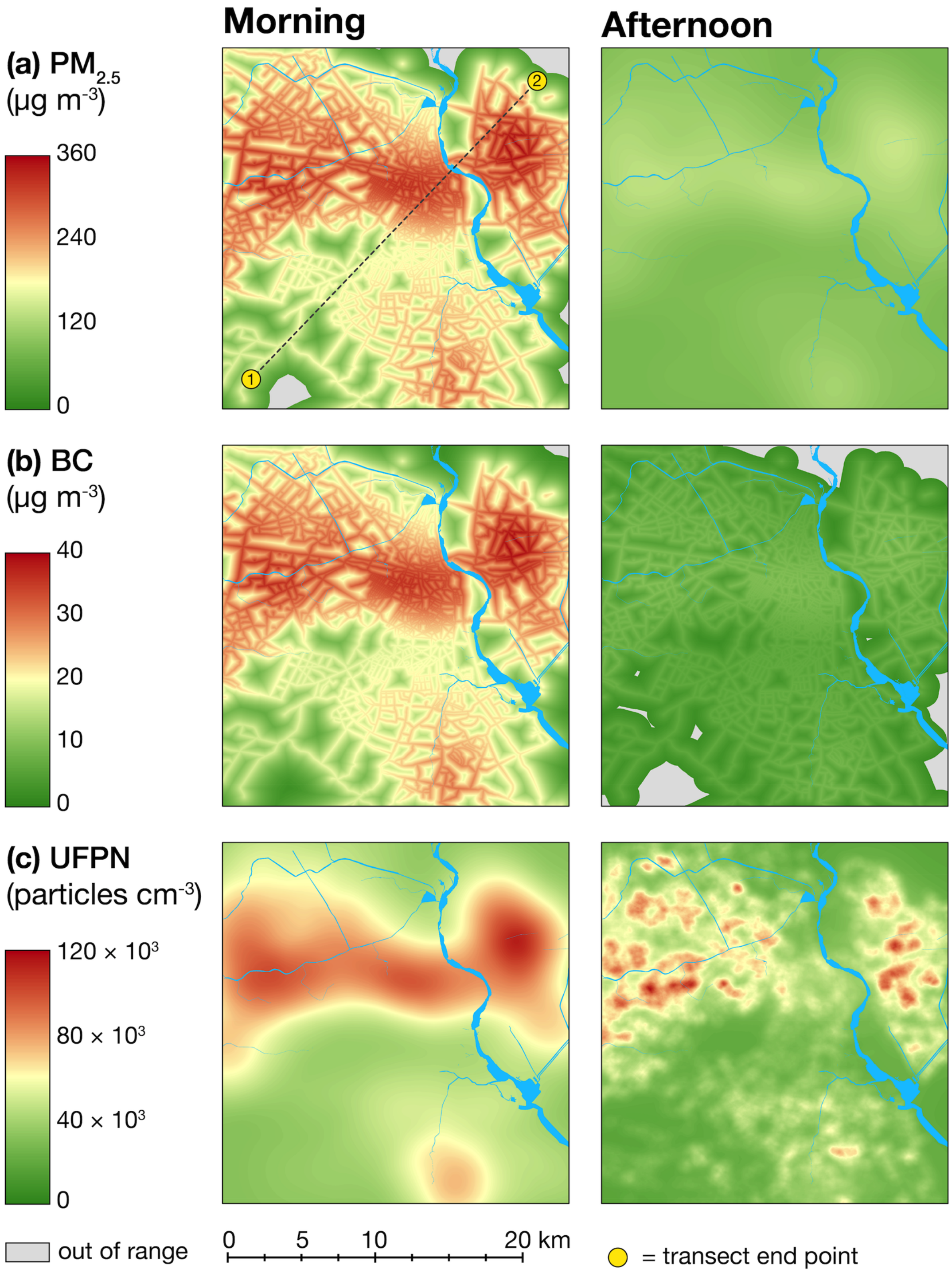


Figure 2. Predicted average fine particulate matter ($PM_{2.5}$), black carbon (BC), and ultrafine particle number concentration (UFPN) spatial variation for the duration of the study. Transect for Figure 3 displayed in part (a).

included population density and minor road length in a buffer of 500 m as statistically significant predictor variables.

4. DISCUSSION

This study documents the first $PM_{2.5}$, BC, and UFPN LUR models developed for New Delhi, India, and among the first developed for highly polluted locations in rapidly developing economies.[59,60] The $PM_{2.5}$ and BC models are spatiotemporal[61] in nature, leveraging data from a fixed continuous monitoring site. As such, the $PM_{2.5}$ and BC models can be used to predict pollutant concentrations for smaller time periods (~1 h periods, which is in contrast to the annual averages predicted by most conventional LUR models). We developed separate models for morning and afternoon hours because of the strong diurnal patterns. For all three pollutants, the spatial patterns were widely different for morning and afternoon hours. Spatial patterns of $PM_{2.5}$ and BC were similar at both times of day, and both were different from UFPN. Our LUR models predicted higher concentrations for, and high variability during, morning hours. Predicted pollutant levels (Figure 3) along an arbitrary transect

dx.doi.org/10.1021/es401489h | Environ. Sci. Technol. XXXX, XXX, XXX−XXX

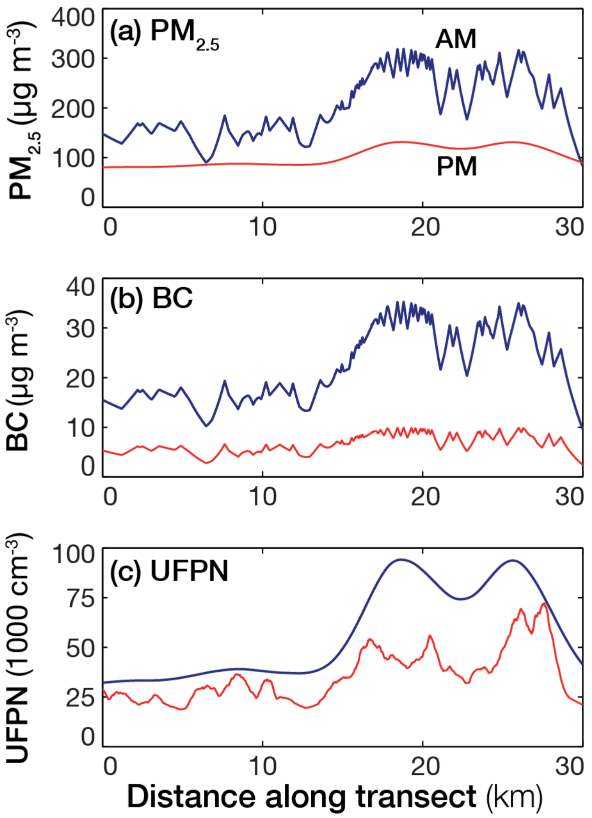


Figure 3. Predicted average fine particulate matter ($PM_{2.5}$), black carbon (BC), and ultrafine particle number concentration (UFPN) variation along an arbitrary transect (from point 1 to point 2, Figure 2a). The red lines indicate results for the afternoon, and the blue line results for the morning.

for morning and afternoon models illustrate this point. This observation (concentrations are more uniform during afternoon than during morning) is likely attributable to more rapid meteorological dispersion in the afternoon.[5,39,44,62] Concentrations of BC, and spatial autocorrelation in modeled BC concentrations, are higher in morning than in afternoon. The UFPN model also showed greater spatial autocorrelation in morning than in afternoon.

High correlation between measurements at the LUR sites and measurements at the fixed site indicate the important role of temporal variability and urban background concentrations. The high correlation between BC and $PM_{2.5}$ measurements at LUR sites and the similar spatial patterns suggest that similar factors (sources and/or meteorology) may be driving both concentrations. Different spatial patterns and weaker correlation between UFPN and the other pollutants indicate that the factors influencing UFPN may be different from those for $PM_{2.5}$ and BC. The observed correlation between $PM_{2.5}$ measurements at LUR sites and $PM_{2.5}$ measurements at the fixed site is higher than the reported correlation in a similar study in Massachusetts that used a similar modeling approach.[63]

Mapping the spatial variability of air pollutants is useful to understand population exposures, and the results can complement information obtained from the regulatory monitoring networks used for air quality management and planning. LUR models typically use one or two weeks of data collected simultaneously at a number of sites.[28] Such monitoring is resource intensive, with the cost of a traditional $PM_{2.5}$ LUR study (including ~40 sites) being approximately 30 000 Euros, assuming that the air-monitoring equipment is available to use.[28] Purchasing the monitoring equipment would roughly double the cost. These levels of funding are often unavailable in developing countries. The spatiotemporal approach used here is applicable in more resource-constrained situations, and our results demonstrate that a reasonable understanding of spatiotemporal patterns of $PM_{2.5}$ and BC can be developed. Furthermore, we expect that performance of UFPN LUR models will improve with the availability of continuous monitoring data from a fixed site. This approach is similar to the development of LUR models with mobile monitoring, which has been explored elsewhere.[64,65]

There are other challenges to developing LUR models for cities in developing countries. Socioeconomic data related to the independent variables in LUR are often not available, or are available at a coarse spatial resolution. For example, the models for New Delhi show that higher levels of $PM_{2.5}$, BC, and UFPN are associated with higher population density. However, population data were available at ward level only; the average ward size is ~100 000 people (Delhi: 16 M people, 156 wards). Availability of population data at a finer scale may improve model performance, as might data on other sources such as biomass burning. The available geographic predictors explained a much smaller proportion of variability in measured $PM_{2.5}$, BC, and UFPN than observed in other studies,[28] indicating the importance of temporal variability and the likelihood of uncharacterized sources and sources that do not correlate with land-use. Performance of our LUR models is in part limited by the influence of uncharacterized sources such as biomass burning (sources that are not necessarily correlated with land use) and the likely impact of a far greater number of distributed sources compared with developed country settings. We recognize the presence of spatial autocorrelation as a limitation of our models. We did not include spatial clustering in the predictions as it decreases interpretability by replacing important but unknown predictors.

Our modeling was limited by the availability of geographic data for constructing predictor variables. Other LUR models have considered more than 100 variables to describe spatial contrasts,[66] compared with the 14 available for this analysis. It is recognized that a larger number of sampling sites can benefit validation of LUR models and is recommended for future projects.[67] Finally, our models are only applicable from February through May, which is when the measurements used here were conducted. Separate models could be developed for other months using the approach employed here.

■ ASSOCIATED CONTENT

Supporting Information

Figures S1 and S2 and Table S1. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 604 822 5918. E-mail: mkandlikar@ires.ubc.ca.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

1. We gratefully acknowledge assistance provided by the following field staff: A. H. Reich, S. J. Deshpande, G. Kaushik, and A. Chel.
2. The following colleagues provided technical/logistical assistance or useful feedback: M. G. Apte, J. Blair, L. A.

Gundel, S. K. Guttikunda, O. T. Hadley, A. D. A. Hansen, T. W. Kirchstetter, M. M. Lunden, W. W. Nazaroff, E. Novotny, C. C. O. Reynolds, K. R. Smith, T. Green, R. Chan, and R. Wang.

3. Partial funding for this work was provided by the Auto-21 NCE (AS; MK); ORSIL no. F07-0010: Leaders Opportunity Fund grant; AS was also supported by a Four-Year Fellowship from the University of British Columbia. Efforts by JSA were partially supported by a Fulbright-Nehru grant from US-India Educational Foundation and by US Environmental Protection Agency under EPA STAR Graduate Fellowship Program. The funding agencies neither reject nor endorse the conclusions and the views expressed herein.

REFERENCES

(1) Dockery, D. W.; Pope, C. A.; Xu, X. P.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. G.; Speizer, F. E. An association between air pollution and mortality in 6 United States cities. *N. Engl. J. Med.* **1993**, *329*, 1753−1759.

(2) Pope, C. A.; Dockery, D. W. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manage. Assoc.* **2006**, *56*, 709−742.

(3) Pope, C. A.; Ezzati, M.; Dockery, D. W. Fine-particulate air pollution and life expectancy in the United States. *N. Engl. J. Med.* **2009**, *360*, 376−386.

(4) Krewski, D.; Jerrett, M.; Burnett, R. T.; Ma, R.; Hughes, E.; Shi, Y.; Turner, M. C.; Pope, C. A.; Thurston, G.; Calle, E. E.; Thun, M. J.; Beckerman, B.; DeLuca, P.; Finkelstein, N.; Ito, K.; Moore, D. K.; Newbold, K. B.; Ramsay, T.; Ross, Z.; Shin, H.; Tempalski, B. *Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality*; Health Effects Institute: Boston, MA, Jun, 2009; research report 140; http://pubs.healtheffects.org/view.php?id=315.

(5) Guttikunda, S. K.; Calori, G. A GIS based emissions inventory at 1 km × 1 km spatial resolution for air pollution analysis in Delhi, India. *Atmos. Environ.* **2013**, *67*, 101−111.

(6) WHO. *Air Quality Guidelines: Global Update 2005*. World Health Organization Europe Regional Office: Copenhagen, Denmark, 2006.

(7) Ross, Z.; Ito, K.; Johnson, S.; Yee, M.; Pezeshki, G.; Clougherty, J. E.; Savitz, D.; Matte, T. Spatial and temporal estimation of air pollutants in New York City: exposure assignment for use in a birth outcomes study. *Environ. Health* **2013**, *12*, 51.

(8) Wang, J.-F.; Hu, M.-G.; Xu, C.-D.; Christakos, G.; Zhao, Y. Estimation of citywide air pollution in Beijing. *PLoS ONE* **2013**, *8*, e53400.

(9) Novakov, T.; Andreae, M. O.; Gabriel, R.; Kirchstetter, T. W.; Mayol-Bracero, O. L.; Ramanathan, V. Origin of carbonaceous aerosols over the tropical Indian Ocean: Biomass burning or fossil fuels? *Geophys. Res. Lett.* **2000**, *27*, 4061−4064.

(10) Grieshop, A. P.; Reynolds, C. C. O.; Kandlikar, M.; Dowlatabadi, H. A black-carbon mitigation wedge. *Nature Geosci.* **2009**, *2*, 533−534.

(11) Menon, S.; Hansen, J.; Nazarenko, L.; Luo, Y. Climate effects of black carbon aerosols in China and India. *Science* **2002**, *297*, 2250−2253.

(12) Janssen, N. A.; Hoek, G.; Simic-Lawson, M.; Fischer, P.; Bree, L. V.; Brink, H. T.; Keuken, M.; Atkinson, R.; Anderson, H. R.; Brunekreef, B.; Cassee, F. R. Black carbon as an additional indicator of the adverse health effects of airborne particles compared to PM_{10} and $PM_{2.5}$. *Environ. Health Perspect.* **2011**, *119*, 1691−1699.

(13) Oberdörster, G.; Oberdörster, E.; Oberdörster, J. Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles. *Environ. Health Perspect.* **2005**, *113*, 823−839.

(14) Seaton, A.; Macnee, W.; Donaldson, K.; Godden, D. Particulate air pollution and acute health effects. *Lancet* **1995**, *345*, 176−178.

(15) Peters, A.; Wichmann, H.; Tuch, T.; Heinrich, J.; Heyder, J. Respiratory effects are associated with the number of ultrafine particles. *Am. J. Respir. Crit. Care Med.* **1997**, *155*, 1376−1383.

(16) Lim, S. S.; Vos, T.; Flaxman, A. D.; Danaei, G.; Shibuya, K.; Adair-Rohani, H.; AlMazroa, M. A.; Amann, M.; Anderson, H. R.; Andrews, K. G.; Aryee, M.; Atkinson, C.; Bacchus, L. J.; Bahalim, A. N.; Balakrishnan, K.; Balmes, J.; Barker-Collo, S.; Baxter, A.; Bell, M. L.; Blore, J. D.; Blyth, F.; Bonner, C.; Borges, G.; Bourne, R.; Boussinesq, M.; Brauer, M.; Brooks, P.; Bruce, N. G.; Brunekreef, B.; Bryan-Hancock, C.; Bucello, C.; Buchbinder, R.; Bull, F.; Burnett, R. T.; Byers, T. E.; Calabria, B.; Carapetis, J.; Carnahan, E.; Chafe, Z.; Charlson, F.; Chen, H.; Chen, J. S.; Cheng, A. T.-A.; Child, J. C.; Cohen, A.; Colson, K. E.; Cowie, B. C.; Darby, S.; Darling, S.; Davis, A.; Degenhardt, L.; Dentener, F.; Des Jarlais, D. C.; Devries, K.; Dherani, M.; Ding, E. L.; Dorsey, E. R.; Driscoll, T.; Edmond, K.; Ali, S. E.; Engell, R. E.; Erwin, P. J.; Fahimi, S.; Falder, G.; Farzadfar, F.; Ferrari, A.; Finucane, M. M.; Flaxman, S.; Fowkes, F. G. R.; Freedman, G.; Freeman, M. K.; Gakidou, E.; Ghosh, S.; Giovannucci, E.; Gmel, G.; Graham, K.; Grainger, R.; Grant, B.; Gunnell, D.; Gutierrez, H. R.; Hall, W.; Hoek, H. W.; Hogan, A.; Hosgood Iii, H. D.; Hoy, D.; Hu, H.; Hubbell, B. J.; Hutchings, S. J.; Ibeanusi, S. E.; Jacklyn, G. L.; Jasrasaria, R.; Jonas, J. B.; Kan, H.; Kanis, J. A.; Kassebaum, N.; Kawakami, N.; Khang, Y.-H.; Khatibzadeh, S.; Khoo, J.-P.; Kok, C.; Laden, F.; Lalloo, R.; Lan, Q.; Lathlean, T.; Leasher, J. L.; Leigh, J.; Li, Y.; Lin, J. K.; Lipshultz, S. E.; London, S.; Lozano, R.; Lu, Y.; Mak, J.; Malekzadeh, R.; Mallinger, L.; Marcenes, W.; March, L.; Marks, R.; Martin, R.; McGale, P.; McGrath, J.; Mehta, S.; Memish, Z. A.; Mensah, G. A.; Merriman, T. R.; Micha, R.; Michaud, C.; Mishra, V.; Hanafiah, K. M.; Mokdad, A. A.; Morawska, L.; Mozaffarian, D.; Murphy, T.; Naghavi, M.; Neal, B.; Nelson, P. K.; Nolla, J. M.; Norman, R.; Olives, C.; Omer, S. B.; Orchard, J.; Osborne, R.; Ostro, B.; Page, A.; Pandey, K. D.; Parry, C. D. H.; Passmore, E.; Patra, J.; Pearce, N.; Pelizzari, P. M.; Petzold, M.; Phillips, M. R.; Pope, D.; Pope Iii, C. A.; Powles, J.; Rao, M.; Razavi, H.; Rehfuess, E. A.; Rehm, J. T.; Ritz, B.; Rivara, F. P.; Roberts, T.; Robinson, C.; Rodriguez-Portales, J. A.; Romieu, I.; Room, R.; Rosenfeld, L. C.; Roy, A.; Rushton, L.; Salomon, J. A.; Sampson, U.; Sanchez-Riera, L.; Sanman, E.; Sapkota, A.; Seedat, S.; Shi, P.; Shield, K.; Shivakoti, R.; Singh, G. M.; Sleet, D. A.; Smith, E.; Smith, K. R.; Stapelberg, N. J. C.; Steenland, K.; Stöckl, H.; Stovner, L. J.; Straif, K.; Straney, L.; Thurston, G. D.; Tran, J. H.; Van Dingenen, R.; van Donkelaar, A.; Veerman, J. L.; Vijayakumar, L.; Weintraub, R.; Weissman, M. M.; White, R. A.; Whiteford, H.; Wiersma, S. T.; Wilkinson, J. D.; Williams, H. C.; Williams, W.; Wilson, N.; Woolf, A. D.; Yip, P.; Zielinski, J. M.; Lopez, A. D.; Murray, C. J. L.; Ezzati, M. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990−2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **2012**, *380*, 2224−2260.

(17) Kandlikar, M. Air pollution at a hotspot location in Delhi: Detecting trends, Seasonal cycles and oscillations. *Atmos. Environ.* **2007**, *41*, 5934−5947.

(18) Census Census of India. http://censusindia.gov.in/2011-prov-results/prov_data_products_delhi.html (accessed November 7, 2011).

(19) Kathuria, V. Impact of CNG on vehicular pollution in Delhi: a note. *Transp. Res. D.* **2004**, *9*, 409−417.

(20) Jayaraman, G. Nidhi. Air pollution and associated respiratory morbidity in Delhi. *Health Care Manage. Sci.* **2008**, *11*, 132−138.

(21) Saksena, S.; Joshi, V.; Patil, R. S. Cluster analysis of Delhi's ambient air quality data. *J. Environ. Monit.* **2003**, *5*, 491−499.

(22) Shandilya, K. K.; Khare, M.; Gupta, A. B. Suspended particulate matter distribution in rural-industrial Satna and in urban-industrial South Delhi. *Environ. Monit. Assess.* **2007**, *128*, 431−445.

(23) Srivastava, A.; Jain, V. K. Seasonal trends in coarse and fine particle sources in Delhi by the chemical mass balance receptor model. *J. Hazard. Mater.* **2007**, *144*, 283−291.

(24) Guttikunda, S. K.; Gurjar, B. Role of meteorology in seasonality of air pollution in megacity Delhi, India. *Environ. Monit. Assess.* **2011**, *184*, 3199−3211.

(25) Cropper, M. L.; Simon, N. B.; Alberini, A.; Arora, S.; Sharma, P. K. The health benefits of air pollution control in Delhi. *Am. J. Agric. Econ.* **1997**, *79*, 1625−1629.

(26) Rajarathnam, U.; Sehgal, M.; Nairy, S.; Patnayak, R. C.; Chhabra, S. K.; Kilnani; Ragavan, K. V. Time-series study on air pollution and mortality in Delhi. *Res Rep Health Eff Inst* **2011**, 47−74.

(27) Marshall, J. D.; Nethery, E.; Brauer, M. Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmos. Environ.* **2008**, *42*, 1359−1369.

(28) Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* **2008**, *42*, 7561−7578.

(29) Novotny, E. V.; Bechle, M. J.; Millet, D. B.; Marshall, J. D. National satellite-based land-use regression: NO_2 in the United States. *Environ. Sci. Technol.* **2011**, *45*, 4407.

(30) Henderson, S. B.; Beckerman, B.; Jerrett, M.; Brauer, M. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environ. Sci. Technol.* **2007**, *41*, 2422−2428.

(31) Gilbert, N. L.; Goldberg, M. S.; Beckerman, B.; Brook, J. R.; Jerrett, M. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. *J. Air Waste Manage. Assoc.* **2005**, *55*, 1059−1063.

(32) Jerrett, M.; Arain, M. A.; Kanaroglou, P.; Beckerman, B.; Crouse, D.; Gilbert, N. L.; Brook, J. R.; Finkelstein, N.; Finkelstein, M. M. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. *J. Toxicol. Env. Health Part A* **2007**, *70*, 200−212.

(33) Kanaroglou, P. S.; Jerrett, M.; Morrison, J.; Beckerman, B.; Arain, M. A.; Gilbert, N. L.; Brook, J. R. Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. *Atmos. Environ.* **2005**, *39*, 2399−2409.

(34) Basagaña, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; Foraster, M.; de Nazelle, A.; Nieuwenhuijsen, M.; Vila, J.; Künzli, N. Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmos. Environ.* **2012**, *54*, 634−642.

(35) Guttikunda, S. Air pollution in Delhi. *Econ. Polit. Wkly.* **2012**, *47*, 24−27.

(36) Guttikunda, S. K.; Jawahar, P. Application of SIM-air modeling tools to assess air quality in Indian cities. *Atmos. Environ.* **2012**, *62*, 551−561.

(37) Kandlikar, M.; Ramachandran, G. The causes and consequences of particulate air pollution in urban India: a synthesis of the science. *Annu. Rev. Energy Environ.* **2000**, *25*, 629−684.

(38) Chowdhury, Z.; Zheng, M.; Schauer, J. J.; Sheesley, R. J.; Salmon, L. G.; Cass, G. R.; Russell, A. G. Speciation of ambient fine organic carbon particles and source apportionment of $PM_{2.5}$ in Indian cities. *J. Geophys. Res., Atmos.* **2007**, *112*.

(39) Tiwari, S.; Srivastava, A. K.; Bisht, D. S.; Parmita, P.; Srivastava, M. K.; Attri, S. D. Diurnal and seasonal variations of black carbon and $PM_{2.5}$ over New Delhi, India: Influence of meteorology. *Atmos. Res.* **2013**, *125−126*, 50−62.

(40) Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P.; Gehring, U.; Heinrich, J.; Cyrys, J.; Bellander, T.; Lewne, M.; Brunekreef, B. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. *Epidemiology* **2003**, *14*, 228−239.

(41) Apte, J. S.; Kirchstetter, T. W.; Reich, A. H.; Deshpande, S. J.; Kaushik, G.; Chel, A.; Marshall, J. D.; Nazaroff, W. W. Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Delhi, India. *Atmos. Environ.* **2011**, *45*, 4470−4480.

(42) Ramachandran, G.; Adgate, J. L.; Pratt, G. C.; Sexton, K. Characterizing indoor and outdoor 15 minute average PM2.5 concentrations in urban neighborhoods. *Aerosol Sci. Technol.* **2003**, *37*, 33−45.

(43) Arku, R. E.; Vallarino, J.; Dionisio, K. L.; Willis, R.; Choi, H.; Wilson, J. G.; Hemphill, C.; Agyei-Mensah, S.; Spengler, J. D.; Ezzati, M. Characterizing air pollution in two low-income neighborhoods in Accra, Ghana. *Sci. Total Environ.* **2008**, *402*, 217−231.

(44) Both, A. F.; Balakrishnan, A.; Joseph, B.; Marshall, J. D. Spatiotemporal aspects of real-time $PM_{2.5}$: Low and middle-income neighborhoods in Bangalore, India. *Environ. Sci. Technol.* **2011**, *45*, 5629−5636.

(45) Kirchstetter, T. W.; Novakov, T. Controlled generation of black carbon particles from a diffusion flame and applications in evaluating black carbon measurement methods. *Atmos. Environ.* **2007**, *41*, 1874−1888.

(46) Jimenez, J.; Claiborn, C.; Larson, T.; Gould, T.; Kirchstetter, T. W.; Gundel, L. Loading effect correction for real-time aethalometer measurements of fresh diesel soot. *J. Air Waste Manage. Assoc.* **2007**, *57*, 868−873.

(47) Hansen, A. D. A.; Rosen, H.; Novakov, T. The aethalometer – An instrument for the real-time measurement of optical absorption by aerosol particles. *Sci. Total Environ.* **1984**, *36*, 191−196.

(48) Westerdahl, D.; Fruin, S.; Sax, T.; Fine, P. M.; Sioutas, C. Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. *Atmos. Environ.* **2005**, *39*, 3597−3610.

(49) Dushkin, L. Scheduled caste policy in India: history, problems, prospects. *Asian Surv.* **1967**, *7*, 626−636.

(50) Limpert, E.; Stahel, W. A.; Abbt, M. Log-normal distributions across the sciences: Keys and clues. *Bioscience* **2001**, *51*, 341−352.

(51) Massey, F. J. The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68−78.

(52) *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011.

(53) Huber, P. J. Regression. In *Robust Statistics*; John Wiley & Sons, Inc.: Hoboken, NJ, 2005; pp 153−198.

(54) Onwuegbuzie, A.; Leech, N.; Whitcome, J. A framework for making quantitative educational research articles more reader-friendly for practitioners. *Qual. Quan.* **2008**, *42*, 75−87.

(55) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2ed.; L. Erlbaum Associates: Hillsdale, N.J., 1988.

(56) Picard, R. R.; Cook, R. D. Cross-validation of regression models. *J. Am. Stat. Assoc.* **1984**, *79*, 575−583.

(57) Getis, A. Reflections on spatial autocorrelation. *Reg. Sci. Urban Econ.* **2007**, *37*, 491−496.

(58) Durbin, J.; Watson, G. S. Testing for serial correlation in least squares regression. II. *Biometrika* **1951**, *38*, 159−177.

(59) Chen, L.; Bai, Z.; Kong, S.; Han, B.; You, Y.; Ding, X.; Du, S.; Liu, A. A land use regression for predicting NO_2 and PM_{10} concentrations in different seasons in Tianjin region, China. *J. Environ. Sci.* **2010**, *22*, 1364−1373.

(60) Allen, R. W.; Gombojav, E.; Barkhasragchaa, B.; Byambaa, T.; Lkhasuren, O.; Amram, O.; Takaro, T. K.; Janes, C. R. An assessment of air pollution and its attributable mortality in Ulaanbaatar, Mongolia. *Air Quality, Atmos. Health* **2011**, 1−14.

(61) Rose, N.; Cowie, C.; Gillett, R.; Marks, G. B. Validation of a spatiotemporal land use regression model incorporating fixed site monitors. *Environ. Sci. Technol.* **2010**, *45*, 294−299.

(62) Tripathi, S. N.; Dey, S.; Tare, V.; Satheesh, S. K. Aerosol black carbon radiative forcing at an industrial city in northern India. *Geophys. Res. Lett.* **2005**, *32*, L08802.

(63) Clougherty, J.; Wright, R.; Baxter, L.; Levy, J. Land use regression modeling of intra-urban residential variability in multiple traffic-related air pollutants. *Environ. Health* **2008**, *7*, 17.

(64) Larson, T.; Su, J.; Baribeau, A.-M.; Buzzelli, M.; Setton, E.; Brauer, M. A spatial model of urban winter woodsmoke concentrations. *Environ. Sci. Technol.* **2007**, *41*, 2429−2436.

(65) Larson, T.; Henderson, S. B.; Brauer, M. Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environ. Sci. Technol.* **2009**, *43*, 4672−4678.

(66) Wang, R.; Henderson, S. B.; Sbihi, H.; Allen, R. W.; Brauer, M. Temporal stability of land use regression models for traffic-related air pollution. *Atmos. Environ.* **2013**, *64*, 312−319.

(67) Johnson, M.; Isakov, V.; Touma, J. S.; Mukerjee, S.; Özkaynak, H. Evaluation of land-use regression models used to predict air quality concentrations in an urban area. *Atmos. Environ.* **2010**, *44*, 3660−3668.

Supplementary Information

Spatiotemporal land use regression models of fine, ultrafine and black carbon particulate matter in New Delhi, India

Arvind Saraswat[a], Joshua S. Apte[b], Milind Kandlikar[a,c]*, Michael Brauer[d], Sarah Henderson[d,e] & Julian Marshall[f]

[a]Institute for Resources Environment and Sustainability, The University of British Columbia, Vancouver, BC Canada

[b]Energy and Resources Group, University of California, Berkeley CA USA

[c]Liu Institute for Global Issues, The University of British Columbia, Vancouver, BC Canada

[d]School of Population and Public Health, Faculty of Medicine, The University of British Columbia, Vancouver, BC Canada

[e]BC Centre for Disease Control, Vancouver, BC Canada

[f]Department of Civil Engineering, University of Minnesota, Minneapolis MN USA

*Corresponding author.

Figure S.1: Diurnal variation of $PM_{2.5}$ concentrations at the rooftop site by month

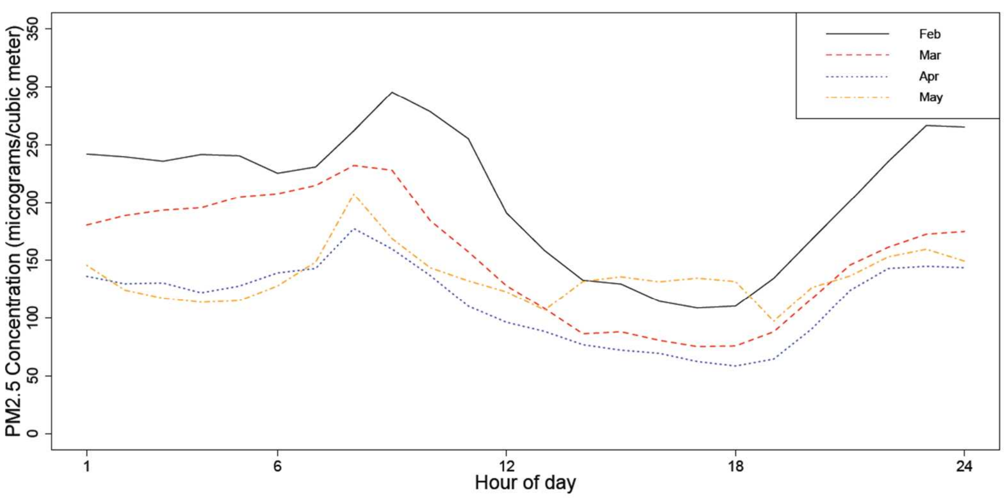


Figure S.2: Diurnal variation of BC concentrations at the rooftop site by month

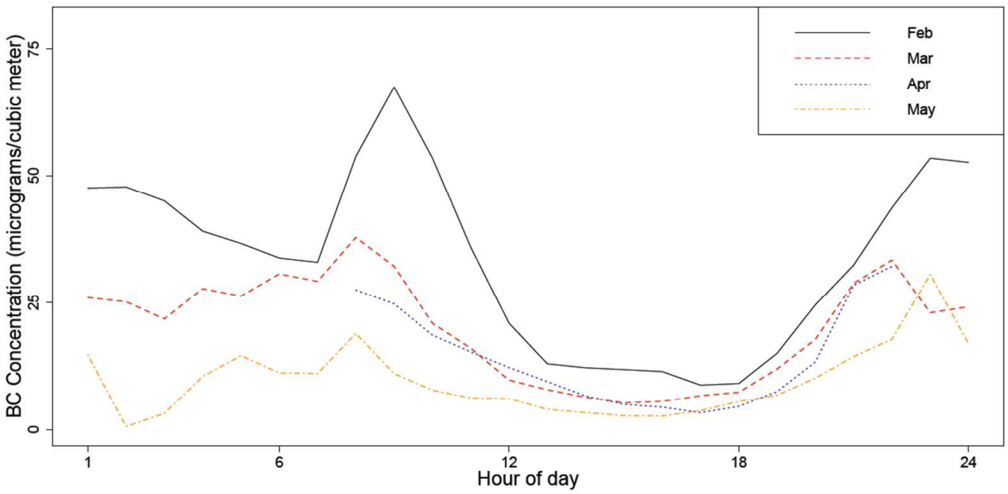


Table S1: LUR models using robust regression

Model	Model terms and coefficients: log-median concentration [a]					LOOE [b]
	Constant	ln(*ROOF*)	*POP.5000*	*DIST.J1*	*RD2.500*	
$PM_{2.5}$ (morning) *N = 19*[c]	1.07	0.80 **	2.0×10^{-3} **	-7.5×10^{-4} **		0.25
$PM_{2.5}$ (afternoon) *N = 35*	-0.29	*1.05* **	1.4×10^{-3} **			0.29
BC (morning) *N = 17*	0.64*	0.75 **	2.5×10^{-3} **	-6.7×10^{-4} *		0.34
BC (afternoon) *N = 25*	-0.03	0.97 **	2.2×10^{-3} **	-1.1×10^{-3} **		0.45
PN (morning) *N = 18*	10.38**		3.1×10^{-3} **			0.59
PN (afternoon) *N = 37*	9.76**		2.3×10^{-3} **		4.9×10^{-5} **	0.72

[a] Spatiotemporal LUR models predict the natural log of the hourly median concentration at each fixed site.
[b] Leave-one-out cross-validation error.
[c] *N* - number of sites