Applying land use regression model to estimate spatial variation of PM_{2.5} in Beijing, China

Jiansheng Wu · Jiacheng Li · Jian Peng · Weifeng Li ·
Guang Xu · Chengcheng Dong

Received: 5 July 2014 / Accepted: 20 November 2014 / Published online: 10 December 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Fine particulate matter (PM_{2.5}) is the major air pollutant in Beijing, posing serious threats to human health. Land use regression (LUR) has been widely used in predicting spatiotemporal variation of ambient air-pollutant concentrations, though restricted to the European and North American context. We aimed to estimate spatiotemporal variations of PM_{2.5} by building separate LUR models in Beijing. Hourly routine PM_{2.5} measurements were collected at 35 sites from 4th March 2013 to 5th March 2014. Seventy-seven predictor variables were generated in GIS, including street network, land cover, population density, catering services distribution, bus stop density, intersection density, and others. Eight LUR models were developed on annual, seasonal, peak/non-peak,

and incremental concentration subsets. The annual mean concentration across all sites is 90.7 $\mu\text{g}/\text{m}^3$ (SD=13.7). PM_{2.5} shows more temporal variation than spatial variation, indicating the necessity of building different models to capture spatiotemporal trends. The adjusted R^2 of these models range between 0.43 and 0.65. Most LUR models are driven by significant predictors including major road length, vegetation, and water land use. Annual outdoor exposure in Beijing is as high as 96.5 $\mu\text{g}/\text{m}^3$. This is among the first LUR studies implemented in a seriously air-polluted Chinese context, which generally produce acceptable results and reliable spatial air-pollution maps. Apart from the models for winter and incremental concentration, LUR models are driven by similar variables, suggesting that the spatial variations of PM_{2.5} remain steady for most of the time. Temporal variations are explained by the intercepts, and spatial variations in the measurements determine the strength of variable coefficients in our models.

Responsible editor: Gerhard Lammel

Electronic supplementary material The online version of this article (doi:10.1007/s11356-014-3893-5) contains supplementary material, which is available to authorized users.

J. Wu · J. Li (✉) · J. Peng · C. Dong
The Key Laboratory for Environmental and Urban Sciences, School of Urban Planning and Design, Shenzhen Graduate School, Peking University, Shenzhen 518055, China
e-mail: billylee17@pku.edu.cn

J. Wu · J. Peng
College of Urban and Environmental Sciences, Laboratory for Earth Surface Processes, Ministry of Education, Peking University, Beijing 100871, China

W. Li (✉)
Department of Urban Planning and Design, University of Hong Kong, Hong Kong, SAR, China
e-mail: wfli@hku.hk

G. Xu
Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

G. Xu
University of Chinese Academy of Sciences, Beijing 100049, China

Keywords Land use regression · Fine particulate matter · PM_{2.5} · Spatiotemporal variation · Outdoor exposure · Air pollution · Beijing

Introduction

Due to rapid urban growth, air pollution has been one of the most serious environmental problems in Beijing, the capital of China. The annual average ambient fine particulate matter (PM_{2.5}, or particles up to 2.5 μm in diameter) concentrations in this city were measured at about 70–100 $\mu\text{g}/\text{m}^3$ (Cheng et al. 2013a; Wang et al. 2013), which were about two to three times higher than the WHO Level 1 Interim Target of 35 $\mu\text{g}/\text{m}^3$ (WHO Press 2006). Globally, the city has been listed among the most PM_{2.5}-polluted areas (Brauer et al. 2012; Van Donkelaar et al. 2010). Overexposure to PM_{2.5} has been

proved to be associated with a wide variety of health effects such as aggravation of existing heart and lung disease and premature mortality (Anderson et al. 2012; Dockery et al. 1993; Jerrett et al. 2005b; Pope and Dockery 2006).

Several epidemiological studies (Breitner et al. 2011; Hou et al. 2010; Zhang et al. 2013; Zhang et al. 2012) have evaluated the association of exposure to particulate matter with adverse health impacts in Beijing. However, these studies are often limited to both rough spatial variability by using interpolation methods and poor temporal variability by using mean concentration throughout the sampling period. Yet, PM_{2.5} pollution is understood to vary over space and time (Hoek et al. 2002; Zhang et al. 2013; Zhao et al. 2009). To better measure the intra-urban heterogeneity in PM_{2.5} concentrations and provide evidence for epidemiological studies, many approaches have been put forward such as exposure indicator variables, interpolation methods, dispersion models, and land use regression (LUR) models (Briggs 2005; Hoek et al. 2008; Holmes and Morawska 2006; Jerrett et al. 2005a; Ryan and LeMasters 2007).

First developed by Briggs et al. (1997), LUR has been proved to have more advantages than other conventional approaches (Hoek et al. 2008). By means of establishing a statistical relationship between pollutant concentrations measured at 20–100 sites and potential influencing factors such as land use, traffic, and population density, LUR is able to predict concentrations at unsampled locations throughout a given domain within the framework of GIS (Henderson et al. 2007; Hoek et al. 2008). LUR has achieved great success in predicting concentrations of major air pollutants, mainly in Europe and North America (Hoek et al. 2008), yet it has seldom been applied to China (Li et al. 2010a, b), one of the most seriously air-polluted regions in the world.

Although LUR has a strong ability to describe spatial variation of pollutants, most LUR models fail to reflect temporal variation, since they consider long-term (i.e., annual) average concentration (Dons et al. 2013; Hoek et al. 2008). Several recent LUR studies have attempted to capture temporal variation, mainly in three ways: (1) by modifying the intercept of LUR models based on a temporal variation observed at background monitoring sites (Gan et al. 2011; Nethery et al. 2008; Saraswat et al. 2013; Slama et al. 2007); (2) by adding dummy variables to reflect different periods (Dons et al. 2013; Noth et al. 2011); and (3) by building LUR models for different periods with dynamic or static variables (Dons et al. 2013; Gulliver et al. 2011). Other studies even integrated the above methods (Adam-Poupart et al. 2014; Patton et al. 2014). The first two approaches are limited to producing the same spatial pattern of pollutants over time, because they are based on a typical LUR model. Only the intercept varies to reflect the corresponding periods, yet other variables remain the same. Moreover, certain explaining

variables may lose their statistical significance or contradict with the direction of effect in a prior assumptions (Dons et al. 2013). The last method is a more rigorous way but may increase workload.

The heavy air pollution in China is likely to have more diverse sources than that in European and North American cities: not only traditional sources such as traffic and industry emission but also coal burning and biomass combustion for cooking and winter heating (Carter et al. 2014; Yu et al. 2013; Zhao et al. 2013b). Moreover, other specific activities like setting off fireworks during the Spring Festival (Wang et al. 2007; Zhang et al. 2010) and open straw burning over the harvest season (Cheng et al. 2013b; Li et al. 2014; Qu et al. 2012) may also aggravate pollution concentrations in some periods, especially in winter. Wide varieties of air pollution sources are likely to increase the difficulties to build LUR models.

In this paper, we aim to capture both spatial and temporal variations (seasonal and diurnal) of PM_{2.5} in Beijing by building separate LUR models based on different periods. By means of model comparison, factors dominating the distribution of PM_{2.5} were analyzed on annual, seasonal, and peak/non-peak subsets. In particular, we built an LUR model for incremental concentration during peak hours in order to better explore diurnal variation of PM_{2.5} in Beijing.

Materials and methods

PM_{2.5} measurements

Routine monitoring data were collected in real time at 35 air-quality monitoring stations, which are available on the Website of the Beijing Environmental Monitoring Center (<http://zx.bjmemc.com.cn/>). Continuous hourly PM_{2.5} concentrations were gathered for approximately a whole year from 4th March 2013 to 5th March 2014. Fine particulate matter was measured using the 1405-F TEOM Monitor (Thermo Fisher Scientific Inc., Waltham, MA, USA) concurrently at all sites. The monitor measured particle mass concentrations based on the tapered-element oscillating microbalance continuous monitoring method (TEOM) at a flow rate of 3 L/min. All samplers were placed approximately 3 m to the ground, and the TEOM sensors were housed in a single-cabinet compact enclosure conditioned within temperature of 20±2 °C. Quality assurance/quality control (QA/QC) procedures at each site were followed according to the environmental protection standards of China (HJ 618-2011; MEPCN 2011). Data loss was inevitable during the measurement, and different approaches were adopted to deal with the missing data in existing LUR models (Amini et al. 2014; Gulliver et al. 2011; Ross et al. 2013). Since there were few missing data in our measurements (Supplementary Table SII),

we did not discard any sites or any periods, and the original dataset was used for analysis.

The 35 stations were divided into four categories to guarantee adequate spatial variation in measured concentrations, including 12 urban environmental evaluation sites, 16 suburban environmental evaluation sites, 5 traffic pollution monitoring sites, and 2 regional background control sites. A detailed description of the 35 sites in their four categories can be seen in Supplementary Table S12. The distribution of all sites is illustrated in Fig. 1.

In order to explore seasonal variation of $PM_{2.5}$ in Beijing, the measured concentrations were classified and integrated for four seasons. Seasonal average $PM_{2.5}$ concentrations were calculated and served as dependent variables of seasonal LUR models. For diurnal variations, a previous study (Zhao et al. 2009) demonstrated that urban and rural $PM_{2.5}$ concentrations display different diurnal patterns, but both show a significant peak between 5:00 p.m. and 11:00 p.m. Thus, the period between 5:00 p.m. and 11:00 p.m. is defined as peak hours, and the rest as non-peak hours. Mean concentrations for these two periods were used to build diurnal LUR models.

To further investigate factors that influence the spatial variation of increased $PM_{2.5}$ concentrations during peak hours, an LUR model that used the change in $PM_{2.5}$ concentrations during peak hours as dependent variable was built, which was termed incremental concentration model. The change in $PM_{2.5}$ concentrations during peak hours was calculated as the difference between the maximum hourly mean concentration during peak hours and the average concentration at 5:00 p.m. Previous LUR studies focused on directly measured concentrations, and this may be a novel approach that used the change in pollution concentrations as dependent variable in LUR studies. In summary, eight LUR models were developed on annual, seasonal, peak/non-peak, and incremental concentration subsets.

Development of predictor variables

Predictor variables considered to be related to air pollutant concentrations were calculated in ArcGIS (ESRI China Inc., version 10.0). A total of 77 variables in seven categories were created to characterize the street network, land cover,

Fig. 1 Classification and distribution of air-quality monitoring sites in Beijing

population density, catering services distribution, bus stop density, intersection density, and other characteristics at or around the sampling sites. Then we sub-divided each category by variation of distance at different buffer radii (Table 1). For each category, data source and detailed description are available in Supplementary Table S13.

In the absence of traffic intensity data, we used the length of specific road types as a reasonable proxy (Henderson et al. 2007). Some studies (Dong et al. 2014; Feng et al. 2013) have shown that average traffic volume on major roads in Beijing is ~97 vehicle/km, which is close to volume in the study implemented in New York (Ross et al. 2007). The major type of fuel used in this area is gasoline; some other types of fuel are also used in different kinds of vehicles, according to Beijing Municipal Transportation Commission.

Particularly, we added catering service variables in our models, and they were not commonly used in previous LUR studies (Abernethy et al. 2013). The locations of restaurants were obtained from Sina Microblog (<http://weibo.com/>), a social networking Website similar to Twitter. Big data of social sensing can easily collect more massive and distributed social information compared with traditional methods. Totally, 13,553 catering service shops were collected, including 440,449 records of check-ins, then the scale of different restaurants was divided into five grades based on check-in numbers.

As for the spatial distribution of winter heating, the major winter heating stations that belonged to the Beijing District

Heating Group (<http://www.bdhg.com.cn>, in Chinese) were included in industrial land use. Most of the winter heating stations in six downtown districts used gas-fired boilers instead of coal-fired boilers, and the stations in the suburbs used both types of boilers. There were also some small-scale winter heating sources, which used coal as fuel generally. These small-scale sources were often scattered and difficult to characterize. Therefore, we only gathered major winter heating stations in the type of industrial land use.

Model development and evaluation

We used the model-building algorithm developed by Henderson et al. (2007). A priori assumptions of a required sign of regression coefficients for specific variables (e.g., the coefficient of road length is positive to PM_{2.5} concentrations) were made so as to increase the applicability of LUR models beyond the monitoring sites (Henderson et al. 2007; Hoek et al. 2008), which is described in Table 1.

In order to reduce the potential for collinearity among variables belonging to the same category, and ensuring interpretability of model parameters, the following model-building algorithm was used:

1. Remove variables with less than five nonzero values
2. In each sub-category, rank all variables by the absolute strength of their correlation with the measured pollutant and identify the highest-ranking variable

Table 1 Classification, descriptions, and assumed effect of spatial predictor variables

Category (N)	Description	Variable sub-categories	Circular buffer radii (m)	Variables names and assumed effect
Road length (12)	Total length of 2 road types within each buffer (m)	Major roads (mr) Secondary roads (sr)	100, 200, 300, 500, 750, 1000 ^a	mr_xx (+) sr_xx (+)
Land cover (36)	Total area of 6 land cover types within each buffer (m ²)	Impervious (bud) Natural vegetation (vege) Water body (wat) Bare land (bare) Cropland (crop) Industrial land (ind)	100, 300, 500, 1000, 2000, 3000	bud_xx (+) vege_xx (-) wat_xx (-) bare_xx (+) corp_xx (NA) ind_xx (+)
Population density (3)	Number of persons within each buffer (units)	Population density (pop)	1000, 2000, 5000	pop_xx (+)
Catering services (8)	Number of restaurants within each buffer (units)	Catering services (cat)	100, 200, 300, 500, 750, 1000, 2000, 3000	cat_xx (+)
Bus stop density (8)	Number of bus stops within each buffer (units)	Bus stops (bus)	100, 200, 300, 500, 750, 1000, 2000, 3000	bus_xx (+)
Intersection density (8)	Number of intersections within each buffer (units)	Intersections (inter)	100, 200, 300, 500, 750, 1000, 2000, 3000	inter_xx (+)
Others (2)	Other variables describing attributes of site	Normalized Difference Vegetation Index (NDVI) Elevation (DEM)	N/A	NDVI (+) DEM (NA)

xx the circular buffer radii (meters), NA not assigned because no effect could be assumed

^a The maximum traffic-related buffer distance was set as 1000 m, in accordance with the principle described by Hoek et al. (2008)

3. Remove other variables in each sub-category that are correlated (Pearson's $r > 0.6$) with the highest-ranking variable
4. Enter all remaining variables into a stepwise linear regression with a 95 % confidence interval in IBM SPSS Software (SPSS China Inc., version 20.0)
5. Remove the variables that have insignificant t statistics ($\alpha = 0.05$) or are inconsistent with a priori assumptions
6. Repeat steps 4 and 5 until there are no more variables that can contribute more than 1 % to the adjusted R^2 .

As linear regression assumes independence of the residuals, a standard diagnostic test was carried out to check spatial autocorrelation of the residuals in ArcGIS using Global Moran's I (Hoek et al. 2008). We evaluated models by Leave-One-Out (LOO) cross validation, where models were developed for $N-1$ sites (N is the total number of sampling sites). The predicted concentrations were compared with the measured concentrations at the left-out site. The above procedure was repeated N times. Then, the root mean squared error (RMSE) and normalized mean squared error (NMSE) were calculated to describe the accuracy of the models. Generally, lower RMSE and NMSE values mean more stable and accurate models.

Regression mapping and population exposure

After obtaining the final valid LUR models, regression equations were applied to a regular 30×30 m grid covering all of the study region to generate a continuous surface of exposure for visualization and computation of population-weighted exposure. In order to limit the estimated concentrations to a reasonable range, predicted

concentrations at the low and high ends were truncated at a certain percentage of measured maximum/minimum values in some LUR studies (Abernethy et al. 2013; Amini et al. 2014; Henderson et al. 2007). There is no rigorous standard for cutting the range of predicted concentrations. We made similar assumptions described by Amini et al. (2014) that predicted concentrations were truncated at 120 % of the maximum and 60 % of the minimum measured values.

The spatial distribution of modeled $PM_{2.5}$ concentrations was then integrated with population density for analysis, which aimed to calculate the outdoor population exposure to $PM_{2.5}$ in Beijing. The outdoor population exposure was calculated as below:

$$C_{pe} = \frac{\sum \text{pop}(i, j) \cdot C(i, j)}{\sum \text{pop}(i, j)}$$

Where, C_{pe} means the average outdoor population exposure across the whole city, $\text{pop}(i, j)$ means the population density of the raster in column i and row j , and $C(i, j)$ is the concentration of $PM_{2.5}$ of the raster in column i and row j .

Results

Descriptive statistics for $PM_{2.5}$ measurements

The hourly $PM_{2.5}$ measurements were analyzed on annual, seasonal, peak/non-peak, and incremental concentration subsets, as described above. Descriptive statistics of $PM_{2.5}$ measurements for these periods are shown in Table 2. The measured concentrations of all these periods were normally

Table 2 Descriptive statistics of $PM_{2.5}$ measurements in eight periods

Period	Measured-mean concentration of all sites (SD ^a)	Measured-mean concentration of traffic sites (SD ^a)	Measured-mean concentration of non-traffic sites (SD ^a)	Range (min–max)	Coefficient of spatial variation ^b of all sites	Moran's I (p value) of all sites
Annual	90.7 (13.7)	99.6 (5.4)	89.2 (14.1)	62.1–115.9	0.15	0.17 (0.01)
Spring	85.9 (9.9)	95.4 (5.4)	84.3 (9.7)	62.1–107.8	0.12	0.19 (0.00)
Summer	78.2 (11.1)	91.8 (5.9)	75.9 (10.1)	58.2–100.7	0.14	0.40 (0.00)
Autumn	84.0 (15.3)	95.5 (5.1)	82.0 (15.7)	57.7–109.5	0.18	0.23 (0.00)
Winter	114.9 (23.0)	116.2 (8.4)	114.6 (24.7)	70.7–163.3	0.20	0.08 (0.16)
Peak hours	96.8 (13.5)	105.6 (6.5)	95.3 (13.8)	68.6–120.2	0.14	0.17 (0.01)
Non-peak hours	88.2 (14.0)	97.2 (5.3)	86.7 (14.5)	59.1–114.1	0.16	0.17 (0.01)
Incremental concentration	19.4 (9.0)	20.6 (2.7)	19.2 (9.7)	0.6–43.9	0.47	0.03 (0.44)

^a Between-site standard deviation of measured-mean concentration for each period

^b Coefficient of between-site spatial variation is defined as SD/mean

distributed (Supplementary Fig. SI 1), which was tested by Shapiro-Wilk test. Apart from the winter period and the incremental concentration during peak hours, PM_{2.5} concentrations presented significant spatial autocorrelation with positive Globe Moran's *I* and *p* values of less than 0.05 during the remaining periods. Moreover, differences between the five traffic sites and the remaining sites were also listed in Table 2. The measured concentrations in traffic sites were generally $\sim 10 \mu\text{g}/\text{m}^3$ higher than those in non-traffic sites, and the value might be larger in warmer seasons. The standard deviation in traffic sites was also smaller.

As is shown in Supplementary Fig. SI2a, different categories of monitoring sites showed diverse PM_{2.5} diurnal patterns but had a significant peak between 5:00 p.m. and 11:00 p.m. in common. This was quite similar to a previous study (Zhao et al. 2009) and provided evidence for distinguishing peak hours from non-peak hours of the diurnal period.

PM_{2.5} shows more temporal variation than spatial variation in Beijing. Coefficient of within-site hourly temporal variation for all sites ranged between 0.81 and 1.06 (Supplementary Table SI2), while coefficient of between-site spatial variation in annual mean measurements was merely 0.15 (Table 2). Evident seasonal and diurnal variations of ambient PM_{2.5} concentrations are illustrated in Table 2 and Supplementary Fig. SI2a, which demonstrates the necessity of building different models to capture spatiotemporal trends (Dons et al. 2013).

LUR models and model evaluation

The annual model (Table 3; Fig. 2) explained 58 % of the variability in measured concentrations, and was driven by the area of natural vegetation land use within the buffer 3000 m (vege_3000), the length of major roads within the buffer 1000 m (mr_1000) and the area of water body within the buffer 500 m (wat_500). The model coefficients of these variables were -2.25×10^{-6} , 4.87×10^{-3} , and -1.87×10^{-4} , respectively.

In regard to seasonal models (Table 3; Fig. 3), the models of spring, summer, and autumn had acceptable adjusted R^2 values ranging between 0.53 and 0.65. The predictor variables of these seasonal models were quite similar to the variables used in the annual model, which resulted in similar patterns of PM_{2.5} pollution maps (Figs. 2 and 3). The length of major roads within the buffer 100 m (mr_100) appeared in the spring and summer models, reflecting small-scale variability of traffic emissions. Model coefficients of the variables were generally smaller in spring and summer, and it was likely due to smaller coefficients of spatial variation in the measurements (Table 2). The winter model was poor fit, yielding an adjusted R^2 of 0.43. The area of natural vegetation land use within the buffer 3000 m (vege_3000) and the area of farmland within the buffer 3000 m (crop_3000) were included as statistically

significant predictor variables, with large model coefficients of -3.24×10^{-6} and 1.27×10^{-6} , respectively.

The model of peak and non-peak hours (Table 3; Fig. SI3) explained 64 and 54 % of the variability in measured PM_{2.5} concentrations, severally. The explanatory variables were the same with the annual model, and the model coefficients were also close to these in the annual model. Unlike the model of peak and non-peak hours, the model of incremental concentration (Table 3, Fig. SI4) was explained by the area of natural vegetation land use within the buffer 3000 m (vege_3000) and the area of farmland within the buffer 300 m (crop_300); road length and water body land use variables were not included in the result. The model coefficients were -1.32×10^{-6} and -3.98×10^{-5} , respectively, which were smaller than those in other models as the incremental concentration was lower than the directly measured values. The result of the incremental concentration model was also unsatisfactory, with an adjusted R^2 value of 0.44.

Spatial autocorrelation of model residuals was checked by Globe Moran's *I* in ArcGIS. The results of *p* value for these models were greater than or equal to 0.05 (Table 3), which indicates weak spatial autocorrelation of the residuals. All these spatiotemporal LUR models were evaluated by LOO cross validation, yielding RMSE values ranging between 6.5 and $19.1 \mu\text{g}/\text{m}^3$, and NMSE values ranging from 0.57 to 10.19 %.

We compared our model with pre-existing urban PM_{2.5} LUR studies carried out mainly in Europe and North America (Table 4). In general, our models produced lower adjusted R^2 values (mean adjusted $R^2=0.55$) than those studies (mean adjusted $R^2=0.68$). RMSE values of our model validation were larger mainly due to the high PM_{2.5} concentrations measured in Beijing, but the NMSE values were relatively small for most models, and relative errors (RMSE/mean) were nearly the same as previous studies. More similarities and differences between this study and previous urban PM_{2.5} LUR models are listed in Table 5.

Regression mapping and population exposure

The estimated concentrations of PM_{2.5} in unsampled locations were calculated on the basis of eight LUR equations. The results of regression mapping across Beijing for annual model and seasonal models are shown in Figs. 2 and 3. The maps of concentrations for diurnal models and incremental concentration during peak hours are presented in Supplementary Figs. SI2 and SI3, respectively.

Truncating predicted concentrations would certainly have effects on regression mapping results. Overall, 3.99–48.21 % out of 18,235,037 grid cells in the modeling domain were affected by truncating among all periods, which primarily distributed in mountainous areas. Our LUR models trended to produce very low estimated concentrations in these regions,

Table 3 Results of eight spatiotemporal LUR models for PM_{2.5} of different periods

Models	Model terms and coefficients					Adjusted R^2	LOO-RMSE ^b (NMSE) ^c	Moran's I (p value)
	Constant	vege_3000	mr_1000	wat_500	mr_100			
Annual	94.39 (2.49) ^a	-2.25×10^{-6} (3.63×10^{-7})	4.87×10^{-3} (1.48×10^{-3})	-1.87×10^{-4} (6.09×10^{-5})		0.58	9.3 (1.03 %)	0.05 (0.30)
Spring	88.74 (1.73)	-1.76×10^{-6} (2.50×10^{-7})	2.67×10^{-3} (1.18×10^{-3})	-1.47×10^{-4} (4.15×10^{-5})	4.18×10^{-2} (1.95×10^{-2})	0.65	6.5 (0.57 %)	0.01 (0.57)
Summer	79.53 (2.08)	-1.76×10^{-6} (3.01×10^{-7})	3.77×10^{-3} (1.43×10^{-3})	-1.37×10^{-4} (5.00×10^{-5})	4.87×10^{-2} (2.35×10^{-2})	0.59	8.1 (1.05 %)	0.12 (0.05)
Autumn	86.44 (2.99)	-2.35×10^{-6} (4.35×10^{-7})	6.08×10^{-3} (1.77×10^{-3})	-1.83×10^{-4} (7.29×10^{-5})		0.53	11.2 (1.72 %)	0.07 (0.20)
Winter	110.41 (5.11)	-3.24×10^{-6} (7.12×10^{-7})			1.27×10^{-6} (4.90×10^{-7})	0.43	19.1 (2.66 %)	0.08 (0.16)
Peak hours	101.21 (2.28)	-2.36×10^{-6} (3.32×10^{-7})	4.44×10^{-3} (1.35×10^{-3})	-1.92×10^{-4} (5.57×10^{-5})		0.64	8.6 (0.77 %)	0.05 (0.27)
Non-peak hours	91.57 (2.70)	-2.21×10^{-6} (3.93×10^{-7})	5.04×10^{-3} (1.60×10^{-3})	-1.85×10^{-4} (6.59×10^{-5})		0.54	10.1 (1.27 %)	0.05 (0.29)
Incremental concentration	24.45 (1.67)	-1.32×10^{-6} (2.78×10^{-7})			-3.98×10^{-5} (1.62×10^{-5})	0.44	6.8 (10.19 %)	0.11 (0.06)

All included variables are statistically significant at the 95 % confidence level

vege_3000 area of natural vegetation land use within the buffer 3000 m, wat_500 area of water body within the buffer 500 m, mr_1000 (mr_100) length of major roads within the buffer 1000 m (100 m), crop_3000 (crop_300) area of farmland within the buffer 3000 m (300 m)

^a Standard errors for the regression coefficients

^b Root mean square error of Leave-One-Out cross validation

^c Normalized mean square error of Leave-One-Out cross validation

Fig. 2 Spatial distribution of modeled PM_{2.5} annual concentrations in Beijing

and some predicted values were even negative. But the predicted high values fit well with measured concentrations. Detailed effects of truncating predicted concentrations are listed in Supplementary Table SI4.

In order to better demonstrate the spatial variation of intra-urban PM_{2.5} concentrations during different periods, the Central Axis of Beijing (displayed in Fig. 2) was selected for cross-section analysis. The predicted PM_{2.5} concentrations for these eight periods along the transect are shown in Fig. 4.

According to the spatial surfaces of PM_{2.5} concentrations (Figs. 2 and 3; Supplementary Figs. SI2 and SI3) and the profile of predicted concentrations along the Central Axis of Beijing (Fig. 4), PM_{2.5} concentrations exhibit similar geographic patterns among different periods apart from the winter model and the model of incremental concentration. PM_{2.5} concentrations are relatively high in plain areas distributed in the east and

south part of Beijing, while the pollution is lighter in mountainous areas with large percentage of vegetation cover. Areas close to the major roads exhibit serious PM_{2.5} pollution as the peak values in Fig. 4 spatially correspond to the major roads. In winter, PM_{2.5} concentrations are higher and more homogeneous in plain areas. In addition, the regression mapping result of incremental concentration shows a more dispersed pattern.

Combining the results of regression mapping with spatial distribution of population, outdoor population exposure to PM_{2.5} was generated. Truncating predicted concentrations had little effect on outdoor population exposure results, and the maximal percentage of population affected by truncating was merely 2.80 % among all periods, since very few population were distributed in mountainous areas. Per capita annual outdoor population exposure to PM_{2.5} was 96.5 μg/m³ across the whole study

Fig. 3 Spatial distribution of modeled PM_{2.5} seasonal concentrations in Beijing

area. More detailed outdoor population exposure to PM_{2.5} for different periods is presented in Table 6.

Discussion

By combining hourly routine monitoring data with spatial predictor variables, we have successfully applied LUR models of different periods to demonstrate the spatiotemporal variability of PM_{2.5} in Beijing. These eight models could explain 43 to 65 % of variability in PM_{2.5} concentrations and were able to predict

concentrations generally within 10 % normalized mean square error (NMSE).

Predictor variables

Regarding the predictor variables included in LUR models, these models were only driven by some traditional variables involving land use and road length, even though we attempted to build LUR models considering various potential sources of PM_{2.5}. Apart from the models for winter and incremental concentration, the final variables entered into LUR models remained almost consistent. Besides, if two models share the

Table 4 Comparison among results of urban PM_{2.5} LUR models

Paper	Study area	Size of the study area	No. of monitoring sites	Sampling period	Sampling campaign	Measured mean/median PM _{2.5} concentrations	Variables entered into the model	Adjust R^2	RMSE of model validation	Relative errors ^b
Brauer et al. (2003)	Stockholm, Sweden Munich, Germany	~86 km ^{2a} 310 km ²	42 40	1999.03–2000.04 1999.03–2000.07	Four 2-week sampling campaign	10.3 µg/m ³ 13.6 µg/m ³	Traffic intensity nearest road; population 1000–5000 m Traffic intensity 50 m; traffic intensity 50–250 m; address density 300 m	0.50 0.56	1.1 µg/m ³ 1.4 µg/m ³	11 % 10 %
Henderson et al. (2007)	Netherlands Vancouver, Canada	41,864 km ² 2200 km ²	40 25	1999.02–2000.03 2003.03–2003.05	High traffic roads 250 m; address density 300 m; region Five 1-week sampling campaign	17.5 µg/m ³ 4.0 µg/m ³	Commercial area 300 m; residential area 750 m; industrial area 300 m; altitude Traffic intensity; distance urban green 3000 m	0.73 0.52	1.6 µg/m ³ 1.5 µg/m ³	9 % 38 %
Hoek et al. (2011)	Amsterdam, Netherlands	219 km ²	50	2002.10–2004.03	–	21.6 µg/m ³	–	0.54	–	–
Eeftens et al. (2012)	20 European cities	–	20 sites at each city	2008.10–2011.04	Three 2-week sampling campaign	8.3–29.3 µg/m ³	–	0.35–0.94	0.8–2.6 µg/m ³	5–15 %
Ross et al. (2013)	New York, USA	789 km ²	150	2008.12–2010.12	Continuous measurements	11.3 µg/m ³	Industrial area 500 m; number of boilers burning residual oil 1000 m; truck traffic density 1600 m; road density 100 m; vegetative area 100 m	0.79	–	<6 %
Saraswat et al. (2013)	New Delhi, India	~551 km ^{2a}	39	2010.02–2010.05	Totally 136 h during the 13-week sampling	133 µg/m ³	Urban background value; population 5000 m; distance major road	0.85, 0.73	–	25 and 29 %
Wu et al. (2014a)	Kaohsiung, Taiwan	325 km ²	29	2011.03–2011.12	Two 2-week sampling campaign	38.3 µg/m ³	Mid-level site; high-level site; major road length 500 m; total road length 25 m; distance gravel plant 5000 m; crop area 100 m	0.65	11.0 µg/m ³	29 %
Wang et al. (2014)	17 European cities	–	356	2008.10–2011.05	Three 2-week sampling campaign	16.2 µg/m ³	Urban background value; traffic intensity 50–1000 m; traffic intensity 50 m; road length 100 m	0.48	2.4 µg/m ³	15 %
This study	Beijing, China	16,411 km ²	35	2013.03–2014.03	Continuous measurements	78.2–114.9 µg/m ³	Major road length 100/1000 m; water area 500 m; natural vegetation area 3000 m; crop area 3000 m	0.43–0.65	6.5–19.1 µg/m ³	8–17 %

^a The area of the study region was estimated from the figures in the manuscript^b The percentage of RMSE to measured-mean concentrations

Table 5 Similarities and differences between this study and previous urban PM_{2.5} LUR researches

	Items	This study	Previous urban PM _{2.5} LUR studies
Differences	PM _{2.5} measurements	Continuous measurements from routine networks	Principally purpose-designed monitoring, composed by several intermittent campaign throughout the sampling period
	Measured PM _{2.5} concentrations	High (78.2–114.9 µg/m ³)	Low, especially in Europe and North America (average measured, 20.0 µg/m ³)
	Adjust <i>R</i> ²	Slightly lower with a scope of 0.43–0.65, mean Adj. <i>R</i> ² =0.55	Typically ranging between 60 and 70 %, mean Adj. <i>R</i> ² =0.68
	RMSE of model validation	High RMSE values between 6.5 µg/m ³ and 19.1 µg/m ³	Low RMSE values with an average of 1.9 µg/m ³
Similarities	No. of monitoring sites	Around 40, few studies include more than 100 sites in a city	
	Explanatory variables	Some traditional variables involving land use and traffic, there are also some variables concerning particular sources, e.g., gravel plant, catering services	
	Relative errors	Generally around 12 %	

same predictor variables, the measured concentrations of the corresponding periods trend to be more correlative (Supplementary Table SI4). This indicates that spatial variations of PM_{2.5} remain steady for most of the time, which is noticeable from the almost overlapping profiles in Fig. 4, and similar patterns of PM_{2.5} concentrations among Figs. 2 and 3 and Supplementary Fig. SI3.

Generally, vehicle emissions played an important role in contributing to PM_{2.5} pollution in Beijing (Cheng et al. 2013a; Song et al. 2006; Sun et al. 2004). LUR models for most of the periods contained road length variables, and road length variables were the only variables with positive model coefficients. This leads to the distinguishable major road network in most regression mapping results and peak values in Fig. 4 that correspond to the major roads spatially. Although the length of major road was a strong predictor in most periods, it was principally with 1000 m buffers and less so with the smaller buffer of 100 m. This is likely due to the fact that PM_{2.5} is a secondary pollutant, and it is unable to reflect small-scale variability of traffic emissions as much as primary pollutants (e.g., black carbon, nitric oxide). Secondary pollutants are generally less sensitive to variables with smaller buffers. Similar conclusion was also drawn by Henderson et al. (2007) as the nitric oxide (NO) map showed more pronounced small-scale spatial contrast than the nitrogen dioxide (NO₂). The second reason for not picking up small-scale variability of traffic emissions much is possibly due to the small number of traffic locations in the sample. Traffic pollution monitoring sites only accounted for a small proportion. The SD values among traffic sites were relatively smaller compared with the remaining sites (Table 2). This might also indicate that the existing traffic sites caught little small-scale spatial variation at

the roadside, while there were both high and low emissions in the remaining sites. PM_{2.5} was lower in areas with more natural vegetation and water bodies. On average, the predicted PM_{2.5} concentration in natural vegetation and water body land cover areas were approximately 42 µg/m³ lower than other areas (effects of truncating were included). This is largely because natural vegetation and water bodies are distributed in areas where there are fewer pollution sources. Moreover, trees can remove fine particles via dry deposition (Nowak et al. 2006), but the effect is very limited as the values in major American cities were lower than 0.1 µg/m³ according to Nowak et al. (2013). A similar study (Baumgardner et al. 2012) in Mexico City showed that the annual removal rate of peri-urban forests for PM₁₀ was approximately 2 %. King et al. (2014) has also suggested an absence of pollution sources in vegetation-covered areas played a larger role than the effect of pollution removal from vegetation in LUR models. The presence of water body land use variable (wat_500) in many models was largely due to a regional background control site next to Miyun reservoir. If we dropped this site from the samples, no variable concerning water body was included in these models.

Spatiotemporal variations explained by LUR models

To further explore the spatiotemporal variations of ambient PM_{2.5} concentrations, LUR models were compared among different periods. Generally, the temporal variations are explained by the intercepts, and the spatial variations in the measurements determine the strength of variable coefficients in our models.

Fig. 4 Cross-section analysis for modeled $PM_{2.5}$ concentrations along the central axis of Beijing

As for temporal variations, previous spatiotemporal LUR studies (Dons et al. 2013; Noth et al. 2011) often used time dummy variables to characterize pollutant concentrations in different periods. In other words, time dummy variables were employed to adjust intercepts of LUR models to reflect temporal variations. Since $PM_{2.5}$ showed more temporal variation than spatial variation in this study, predicted concentrations along the Central Axis presented almost overlapping pattern in Fig. 4, and the relative height of these profiles was determined by the intercepts of LUR models to represent temporal variations among periods. Remarkably, similar conclusion could be seen in an existing spatiotemporal LUR study carried out in New Delhi (Saraswat et al. 2013). Citywide temporal variation of $PM_{2.5}$ and BC was largely on account of the variability monitored from a background site.

For spatial variations, periods with larger coefficient of spatial variation trend to have stronger variable coefficients. There is less vegetation in winter (leaf-off) as deciduous forest plays the dominant role in Beijing (He et al. 2013; Xie et al. 2010), but the *vege_3000* variable (area of natural vegetation land use in a 3000-m buffer) had a stronger coefficient than other seasons (leaf-on). As is concluded above, $PM_{2.5}$ was lower in areas with more natural vegetation mainly because there were fewer pollution sources, rather than the effect of pollution removal from vegetation. The larger effect of vegetation variable in winter might be principally on account of the larger spatial variation of $PM_{2.5}$ measurements in winter (Table 2), which is due to more pollution sources (e.g., winter heating, setting off fireworks) distributed in plain areas. The larger coefficient of vegetation variable was a response to the larger spatial variation, and it could generate enough spatial

Table 6 Results of outdoor population exposure to PM_{2.5} for different periods

Period	Whole Beijing (<i>N</i> =24,403)			Six downtown districts (<i>N</i> =2040)		
	Mean exposed concentration (μg/m ³)	SD ^a	Range ^b (μg/m ³)	Mean exposed concentration (μg/m ³)	SD ^a	Range ^b (μg/m ³)
Annual	96.5	26.4	37.2–139.1	99.1	16.2	37.2–139.1
Spring	89.1	21.3	37.3–129.3	90.6	12.3	37.3–126.9
Summer	82.1	21.2	34.9–120.9	84.4	14.0	34.9–120.9
Autumn	90.6	25.9	34.6–131.4	94.1	18.2	34.6–131.4
Winter	117.3	38.5	42.4–141.4	116.4	14.4	42.4–132.8
Peak hours	102.3	27.4	41.1–144.3	104.8	15.9	41.1–144.3
Non-peak hours	94.0	26.2	35.5–136.9	96.8	16.3	35.5–136.9
Incremental concentration ^c	19.9	8.5	0.4–24.4	22.2	6.0	0.4–24.4

N number of grids estimated; each grid has a spatial resolution of 0.5' (816.3 m approximately)

^a Standard deviation of population-weighted exposure concentration for each period

^b The truncated minimum and maximum value of population-weighted exposure concentration for each period

^c The average outdoor population exposure PM_{2.5} concentration increased during peak hours

contrast in regression mapping results. This is also suggested by King et al. (2014). Similarly, the coefficient of mr_1000 (length of major roads within the buffer 1000 m) was smaller in peak hours than in non-peak hours, which was mainly due to a stronger spatial variation in non-peak hours.

Seasonal variations

Seasonal variations of PM_{2.5} concentrations are similar in spring, summer, and autumn, but show a different pattern in winter. The similar patterns of these three seasons are more likely due to meteorological factors, since specific sources (e.g., winter heating) play a weaker role during these periods. The winter model with only two variables produced a lower *R*², which was mainly on account of a larger spatial variation in the measured PM_{2.5} concentrations. This might be explained by specific PM_{2.5} sources that are distributed in a small scale, such as fossil fuel combustion, biomass burning for cooking and winter heating (Carter et al. 2014; Cheng et al. 2013b; Yu et al. 2013; Zhao et al. 2013b), and setting off fireworks in the winter (Wang et al. 2007; Zhang et al. 2010). Road length variables were no longer significant in winter compared with these in other three seasons, as the *p* values were 0.54 and 0.59 for the most predictive road-length variables in categories of major and common roads in the winter LUR model.

Seasonal variations of PM_{2.5} can appear on account of particular variables. In the winter LUR model, a robust positive coefficient was generated between PM_{2.5} concentrations and areas of cropland within 3000 m. Nevertheless, cropland of various buffer radii showed a negative correlation with PM_{2.5} concentrations in summer and uncertain directions in

spring and autumn. This is apparent because cropland can be regarded as vegetation in summer, while it serves as bare land in winter. Bare land can be a potential source of PM_{2.5} mainly in two ways: the major one is soil or sand dust caused by wind erosion (Chen et al. 2014; Wang et al. 2006), especially in winter; another is soil or sand dust from construction activities (Chen et al. 2014; Wang et al. 2006; Wu et al. 2014b), and construction sites were classified as bare land type in our study. Besides, small-scale distributed combustion of crop residuals also contributed to the PM_{2.5} pollution in winter (Cheng et al. 2013b; Wang et al. 2009), despite open straw burning was prohibited in Beijing. Even though all LUR models were based on the same static variables, rather than dynamic variables that changed over different time periods (Dons et al. 2013), specific static variables such as cropland could play various roles in LUR models over different periods (McCarty 2011).

Diurnal variations

Diurnal variations of PM_{2.5} concentrations were explored primarily concerning the peak hours. Our initial purpose was to explore factors that influence the spatial variation of increased PM_{2.5} concentrations during peak hours by means of model comparison between peak and non-peak hours and building the model for incremental concentration. Only a few LUR studies (Dons et al. 2013; Hong and Bae 2012; Patton et al. 2014) have investigated the issue regarding traffic rush hours.

As shown in Supplementary Fig. SI2b, pronounced increase of PM_{2.5} during traffic peak hours did not occur in all seasons, though they were significant among different

categories of sites when involving the annual mean concentration. This phenomenon could be influenced by meteorological conditions and were more apparent in colder seasons (autumn and winter). Some previous studies (Hu et al. 2014; Zhao et al. 2009) outside of the LUR framework have demonstrated that meteorological factors have a significant impact on PM_{2.5} diurnal cycle in Beijing. Our models of peak and non-peak hours were driven by the same explanatory variables and produced similar patterns of PM_{2.5} concentrations, which is more likely to be interpreted by meteorological factors and boundary layer height (Du et al. 2013) instead of specific sources such as traffic emission and cooking. In addition, the models of peak hours did not include more traffic variables, which might also be explained by the secondary nature of PM_{2.5}. Another pre-existing hourly regression model for a primary pollutant of BC (Dons et al. 2013) had shown that traffic variables were more significant on peak hours.

The LUR model for incremental concentration was more poorly fit than other models, yielding an adjusted R^2 of 0.43 and an RMSE accounting for 75 % of standard deviation of the measurements. It was the same case with the winter model; the poor fit was largely due to a larger spatial variation in the incremental concentration during peak hours. The existing variables could not match the spatial distribution of PM_{2.5} measurements or specific PM_{2.5} sources such as cooking during peak hours well. Categorical or time variables (e.g., peak cooking hour) could be used as a proxy of these sources to improve spatiotemporal LUR models. The spatial pattern was remarkably higher in six downtown districts, even though no traffic-related variables were involved in the model, which indicates human activities contribute a lot to the increment in PM_{2.5} concentrations diurnally (Dons et al. 2013; Zhao et al. 2013a). Employing incremental concentration to build LUR models may be a novel approach and calls for further research.

Outdoor exposure

The results of outdoor exposure in Beijing were alarming, since the predicted long-term outdoor exposure of PM_{2.5} was more than twice the ambient air quality standards of 35 $\mu\text{g}/\text{m}^3$ in China. Even the minimum measured annual PM_{2.5} concentration at all 35 sites (62.1 $\mu\text{g}/\text{m}^3$ at Miyun Reservoir) exceeded the standards. The situations in six downtown districts (Xicheng, Dongcheng, Haidian, Chaoyang, Shijingshan, and Fengtai) were more terrible.

A previous study (Zhang et al. 2013) conducted in six downtown districts in autumn, 2012 produced a similar result (102.5 $\mu\text{g}/\text{m}^3$) to our model, but the kriging method might decrease spatial variation in PM_{2.5} exposures (Lee et al. 2014; Mulholland et al. 1998). Spatiotemporal LUR models are supposed to generate results with fine resolution and high variations for both short- and long-term outdoor exposures

(Dons et al. 2014), which are aimed at promoting epidemiological and risk assessment studies in Beijing.

Limitations

This is among the first LUR studies implemented in a seriously air-polluted Chinese context. Li et al. (2010a, b) have made the first attempt to apply LUR models in Chinese cities with limited number of routine monitoring sites. Their models introduced meteorological variables and yielded moderate adjusted R^2 values. Our models have made large improvements on them with more precise model-building algorithm and regression mapping results and have exhibited spatial variation of PM_{2.5} in more detailed periods.

However, a few limitations of our LUR models should be recognized. First, the models are based on 35 observations, which is fewer than the 40–80 recommended by Hoek et al. (2008). The limited number of monitoring sites is a common restriction in LUR models implemented outside of Europe and North America (Supplementary Table SI6). Second, our models are limited by the availability of more refined predictor variables, including uncharacterized PM_{2.5} sources (e.g., biomass burning), meteorological data, and dynamic variables. The inaccessibility of these predictors is also a universal limitation for LUR models in seriously polluted developing countries (Supplementary Table SI6). The absence of variables characterizing detailed pollution sources often leads to a smaller proportion of variability explained by LUR models, which was concluded by Allen et al. (2013) and Saraswat et al. (2013). Moreover, recently published spatiotemporal LUR models (Patton et al. 2014; Smargiassi et al. 2012) usually incorporated meteorological conditions, especially wind speed and direction, as these factors play a crucial role in the spatiotemporal pattern of air pollutants (Gupta and Christopher 2009; McKendry 2000). As for dynamic variables that change over time, they can generally improve the performance of spatiotemporal LUR models (Dons et al. 2013). Yet, our models simply used different seasons and hours of the day as a reasonable proxy for more refined meteorology and dynamic variables (e.g., dynamic traffic variables). Further research should be done with both meteorological data and dynamic variables to build more precise spatiotemporal LUR models. Finally, the adjusted R^2 values of our models were generally smaller than the existing urban PM_{2.5} LUR models in Table 4 and the models outside of Europe and North America in Supplementary Table SI6, which was largely due to the lack of more detailed predictor variables and large spatial variations in the measurements.

As for the methods of building spatiotemporal LUR models, one previous research (Dons et al. 2013) compared different ways and summarized them in detail. However, our study used the most basic method by developing spatiotemporal LUR models for each period independently. Based on

the results above, adding dummy variables to represent temporal variation may be more convenient, as most of the periods show a similar spatial variation. However, this is not suitable for winter since a different temporal variation is presented. Dons et al. (2013) have also suggested another approach to build spatiotemporal LUR models, which was termed grouping of periods. In summary, different periods are grouped together if they have similar statistical characteristics of pollution measurements. Then, a spatiotemporal LUR model with time dummy variables is built to measure the spatial variation of pollutants throughout the grouped period, and temporal variations are explained by dummy variables within the grouped period. If dynamic predictor variables corresponding to period are available, they can be used instead of the static variables. We would try this approach in our future study to test its feasibility and performance.

Conclusion

We have constructed eight LUR models to explore spatial variations of annual, seasonal, diurnal, and incremental $PM_{2.5}$ concentrations in Beijing. These LUR models generally produce convincing results and exhibit reliable spatial variation of $PM_{2.5}$ with routine monitoring of air quality data. Most LUR models, apart from the models for winter and incremental concentration, share similar significant predictors including major road length, vegetation, and water land use, which indicates that the spatial variations of $PM_{2.5}$ keep steady for most of the time. In these LUR models, temporal and spatial variations are respectively explained by the intercepts and the absolute value of variable coefficients. Annual outdoor exposure for $PM_{2.5}$ is as high as $96.5 \mu\text{g}/\text{m}^3$ in Beijing, which is about two to three times higher than the WHO Level 1 Interim Target (WHO Press 2006) of $35 \mu\text{g}/\text{m}^3$. Spatiotemporal LUR models can provide promising references for both short- and long-term epidemiological studies and pollution abatement in Beijing. Air pollution is a serious public health concern in China, and it calls for a wider application of LUR models in the future.

Acknowledgments This study was financially supported by the National Science Foundation of China (No. 41130534). We thank anonymous reviewers and the editor (Prof. Dr. Gerhard Lammel) for their constructive comments.

Conflict of interest None declared.

References

- Abernethy RC, Allen RW, McKendry IG, Brauer M (2013) A land use regression model for ultrafine particles in Vancouver, Canada. *Environ Sci Technol* 47:5217–5225
- Adam-Poupart A, Brand A, Fournier M, Jerrett M, Smargiassi A (2014) Spatiotemporal modeling of ozone levels in Quebec (Canada): a comparison of kriging, land-use regression (LUR), and combined Bayesian maximum entropy-LUR approaches. *Environ Health Perspect* 122:970–976
- Allen RW, Gombojav E, Barkhasragchaa B, Byambaa T, Lkhasuren O, Amram O, Takaro TK, Janes CR (2013) An assessment of air pollution and its attributable mortality in Ulaanbaatar, Mongolia. *Air Qual Atmos Health* 6:137–150
- Amini H, Taghavi-Shahri SM, Henderson SB, Naddafi K, Nabizadeh R, Yunesian M (2014) Land use regression models to estimate the annual and seasonal spatial variability of sulfur dioxide and particulate matter in Tehran, Iran. *Sci Total Environ* 488:343–353
- Anderson JO, Thundiyil JG, Stolbach A (2012) Clearing the air: a review of the effects of particulate matter air pollution on human health. *J Med Toxicol* 8:166–175
- Baumgardner D, Varela S, Escobedo FJ, Chacalo A, Ochoa C (2012) The role of a peri-urban forest on air quality improvement in the Mexico City megalopolis. *Environ Pollut* 163:174–183
- Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Gehring U, Heinrich J, Cyrys J, Bellander T, Lewne M (2003) Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* 14:228–239
- Brauer M, Amann M, Burnett RT, Cohen A, Dentener F, Ezzati M, Henderson SB, Krzyzanowski M, Martin RV, Van Dingenen R (2012) Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. *Environ Sci Technol* 46:652–660
- Breitner S, Liu L, Cyrys J, Bröske I, Franck U, Schlink U, Leitte AM, Herbarth O, Wiedensohler A, Wehner B (2011) Sub-micrometer particulate air pollution and cardiovascular mortality in Beijing, China. *Sci Total Environ* 409:5196–5204
- Briggs D (2005) The role of GIS: coping with space (and time) in air pollution exposure assessment. *J Toxicol Environ Health Part A* 68:1243–1261
- Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebreton E, Pryl K, van Reeuwijk H, Smallbone K, Van Der Veen A (1997) Mapping urban air pollution using GIS: a regression-based approach. *Int J Geogr Inf Sci* 11:699–718
- Carter EM, Shan M, Yang XD, Li JR, Baumgartner J (2014) Pollutant emissions and energy efficiency of Chinese gasifier cooking stoves and implications for future intervention studies. *Environ Sci Technol* 48:6461–6467
- Chen Y, Schleicher N, Chen Y, Chai F, Norra S (2014) The influence of governmental mitigation measures on contamination characteristics of $PM_{2.5}$ in Beijing. *Sci Total Environ* 490:647–658
- Cheng SY, Lang JL, Zhou Y, Han LH, Wang G, Chen DS (2013a) A new monitoring-simulation-source apportionment approach for investigating the vehicular emission contribution to the $PM_{2.5}$ pollution in Beijing, China. *Atmos Environ* 79:308–316
- Cheng Y, Engling G, He KB, Duan FK, Ma YL, Du ZY, Liu JM, Zheng M, Weber RJ (2013b) Biomass burning contribution to Beijing aerosol. *Atmos Chem Phys* 13:7765–7781
- Dockery DW, Pope CA 3rd, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE (1993) An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 329:1753–1759
- Dong C, Shao C, Richards SH, Han LD (2014) Flow rate and time mean speed predictions for the urban freeway network using state space models. *Transp Res C Emerg Technol* 43(1):20–32

- Dons E, Van Poppel M, Kochan B, Wets G, Int Panis L (2013) Modeling temporal and spatial variability of traffic-related air pollution: hourly land use regression models for black carbon. *Atmos Environ* 74: 237–246
- Dons E, Van Poppel M, Int Panis L, De Prins S, Berghmans P, Koppen G, Matheeußen C (2014) Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. *Sci Total Environ* 476:378–386
- Du C, Liu S, Yu X, Li X, Chen C, Peng Y, Dong Y, Dong Z, Wang F (2013) Urban boundary layer height characteristics and relationship with particulate matter mass concentrations in Xi'an, central China. *Aerosol Air Qual Res* 13:1598–1607. doi:10.4209/aaqr.2012.10.0274
- Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, Declercq C, Dedele A, Dons E, de Nazelle A (2012) Development of land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* 46:11195–11205
- Feng C, Qi Z, Yuanhua J, Jian L (2013) Research of traffic flow multi-objectives intelligent control method for junction network. *Telecommun Syst* 53:77–84
- Gan WQ, Koehoorn M, Davies HW, Demers PA, Tamburic L, Brauer M (2011) Long-term exposure to traffic-related air pollution and the risk of coronary heart disease hospitalization and mortality. *Environ Health Perspect* 119:501–507
- Gulliver J, Morris C, Lee K, Vienneau D, Briggs D, Hansell A (2011) Land use regression modeling to estimate historic (1962–1991) concentrations of black smoke and sulfur dioxide for Great Britain. *Environ Sci Technol* 45:3526–3532
- Gupta P, Christopher SA (2009) Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. A neural network approach. *J Geophys Res: Atmos* (1984–2012) 114 doi:10.1029/2008jd011497
- He C, Convertino M, Feng ZK, Zhang SY (2013) Using LiDAR data to measure the 3D green biomass of Beijing urban forest in China. *PLoS One* 8:e75920
- Henderson SB, Beckerman B, Jerrett M, Brauer M (2007) Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environ Sci Technol* 41:2422–2428
- Hoek G, Meliefste K, Cyrys J, Lewné M, Bellander T, Brauer M, Fischer P, Gehring U, Heinrich J, van Vliet P (2002) Spatial variability of fine particle concentrations in three European areas. *Atmos Environ* 36:4077–4088
- Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, Briggs D (2008) A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos Environ* 42:7561–7578
- Hoek G, Beelen R, Kos G, Dijkema M, van der Zee SC, Fischer PH, Brunekreef B (2011) Land use regression model for ultrafine particles in Amsterdam. *Environ Sci Technol* 45:622–628
- Holmes NS, Morawska L (2006) A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available. *Atmos Environ* 40:5902–5928
- Hong ESA, Bae CHC (2012) Exposure of bicyclists to air pollution in Seattle, Washington hybrid analysis using personal monitoring and land use regression. *Transp Res Record*:59–66
- Hou Q, An X, Wang Y, Guo J (2010) An evaluation of resident exposure to respirable particulate matter and health economic loss in Beijing during Beijing 2008 Olympic Games. *Sci Total Environ* 408:4026–4032
- Hu GY, Zhang YM, Sun JY, Zhang LM, Shen XJ, Lin WL, Yang Y (2014) Variability, formation and acidity of water-soluble ions in PM_{2.5} in Beijing based on the semi-continuous observations. *Atmos Res* 145:1–11
- Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahuvaroglu T, Morrison J, Giovis C (2005a) A review and evaluation of intraurban air pollution exposure models. *J Expo Sci Environ Epidemiol* 15:185–204
- Jerrett M, Burnett RT, Ma R, Pope CA 3rd, Krewski D, Newbold KB, Thurston G, Shi Y, Finkelstein N, Calle EE, Thun MJ (2005b) Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* 16:727–736
- King KL, Johnson S, Kheirbek I, Lu JWT, Matte T (2014) Differences in magnitude and spatial distribution of urban forest pollution deposition rates, air pollution emissions, and ambient neighborhood air quality in New York City. *Landsc Urban Plan* 128:14–22
- Lee JH, Wu CF, Hoek G, de Hoogh K, Beelen R, Brunekreef B, Chan CC (2014) Land use regression models for estimating individual NO_x and NO₂ exposures in a metropolis with a high density of traffic roads and population. *Sci Total Environ* 472:1163–1171
- Li C, Bai Z, Kong S, Han B, You Y, Ding X, Du S, Liu A (2010a) A land use regression for predicting NO₂ and PM₁₀ concentrations in different seasons in Tianjin region, China. *J Environ Sci-China* 22: 1364–1373
- Li C, Du SY, Bai ZP, Kong SF, You Y, Bin H, Han DW, Li ZY (2010b) Application of land use regression for estimating concentrations of major outdoor air pollutants in Jinan, China. *J Zhejiang Univ Sci* 11: 857–867. doi:10.1631/jzus.A1000092
- Li J, Song Y, Mao Y, Mao Z, Wu Y, Li M, Huang X, He Q, Hu M (2014) Chemical characteristics and source apportionment of PM_{2.5} during the harvest season in eastern China's agricultural regions. *Atmos Environ* 92:442–448
- McCarty JL (2011) Remote sensing-based estimates of annual and seasonal emissions from crop residue burning in the contiguous United States. *J Air Waste Manage Assoc* 61:22–34
- McKendry IG (2000) PM₁₀ levels in the Lower Fraser Valley, British Columbia, Canada: an overview of spatiotemporal variations and meteorological controls. *J Air Waste Manage Assoc* 50:443–452
- MEPCN (2011) Determination of atmospheric particles PM₁₀ and PM_{2.5} in ambient air by gravimetric method. http://english.mep.gov.cn/standards_reports/standards/Air_Environment/air_method/201111/t20111101_219390.htm. Accessed 07 Nov 2014
- Mulholland JA, Butler AJ, Wilkinson JG, Russell AG, Tolbert PE (1998) Temporal and spatial distributions of ozone in Atlanta: regulatory and epidemiologic implications. *J Air Waste Manage Assoc* 48:418–426
- Nethery E, Teschke K, Brauer M (2008) Predicting personal exposure of pregnant women to traffic-related air pollutants. *Sci Total Environ* 395:11–22
- Noth EM, Hammond SK, Biging GS, Tager IB (2011) A spatial-temporal regression model to predict daily outdoor residential PAH concentrations in an epidemiologic study in Fresno, CA. *Atmos Environ* 45:2394–2403
- Nowak DJ, Crane DE, Stevens JC (2006) Air pollution removal by urban trees and shrubs in the United States. *Urban Urban Green* 4:115–123
- Nowak DJ, Hirabayashi S, Bodine A, Hoehn R (2013) Modeled PM_{2.5} removal by trees in ten US cities and associated health effects. *Environ Pollut* 178:395–402
- Patton AP, Collins C, Naumova EN, Zamore W, Brugge D, Durant JL (2014) An hourly regression model for ultrafine particles in a near-highway urban area. *Environ Sci Technol* 48:3272–3280
- Pope CA III, Dockery DW (2006) Health effects of fine particulate air pollution: lines that connect. *J Air Waste Manage Assoc* 56:709–742
- WHO Press (2006) WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. World Health Organization. http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf. Accessed 25 Oct 2014
- Qu CS, Li B, Wu HS, Giesy JP (2012) Controlling Air pollution from straw burning in China calls for efficient recycling. *Environ Sci Technol* 46:7934–7936
- Ross Z, Jerrett M, Ito K, Tempalski B, Thurston GD (2007) A land use regression for predicting fine particulate matter concentrations in the New York City region. *Atmos Environ* 41:2255–2269
- Ross Z, Ito K, Johnson S, Yee M, Pezeshki G, Clougherty JE, Savitz D, Matte T (2013) Spatial and temporal estimation of air pollutants in

- New York City: exposure assignment for use in a birth outcomes study. *Environ Health* 12:51
- Ryan PH, LeMasters GK (2007) A review of land-use regression for characterizing intraurban air models pollution exposure. *Inhal Toxicol* 19:127–133
- Saraswat A, Apte JS, Kandlikar M, Brauer M, Henderson SB, Marshall JD (2013) Spatiotemporal land use regression models of fine, ultra-fine, and black carbon particulate matter in New Delhi, India. *Environ Sci Technol* 47:12903–12911
- Slama R, Morgenstern V, Cyrus J, Zutavern A, Herbarth O, Wichmann H-E, Heinrich J, LISA Study Group (2007) Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model. *Environ Health Perspect* 115:1283
- Smargiassi A, Brand A, Fournier M, Tessier F, Goudreau S, Rousseau J, Benjamin M (2012) A spatiotemporal land-use regression model of winter fine particulate levels in residential neighbourhoods. *J Expo Sci Environ Epidemiol* 22:331–338
- Song Y, Zhang Y, Xie S, Zeng L, Zheng M, Salmon LG, Shao M, Slanina S (2006) Source apportionment of PM_{2.5} in Beijing by positive matrix factorization. *Atmos Environ* 40:1526–1537
- Sun Y, Zhuang G, Wang Y, Han L, Guo J, Dan M, Zhang W, Wang Z, Hao Z (2004) The air-borne particulate pollution in Beijing—concentration, composition, distribution and sources. *Atmos Environ* 38:5991–6004
- Van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, Villeneuve PJ (2010) Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ Health Perspect* 118:847
- Wang Y, Zhuang GS, Sun YL, An ZS (2006) The variation of characteristics and formation mechanisms of aerosols in dust, haze, and clear days in Beijing. *Atmos Environ* 40:6579–6591
- Wang Y, Zhuang GS, Xu C, An ZS (2007) The air pollution caused by the burning of fireworks during the lantern festival in Beijing. *Atmos Environ* 41:417–431
- Wang Q, Shao M, Zhang Y, Wei Y, Hu M, Guo S (2009) Source apportionment of fine organic aerosols in Beijing. *Atmos Chem Phys* 9:8573–8585
- Wang JF, Hu MG, Xu CD, Christakos G, Zhao Y (2013) Estimation of citywide air pollution in Beijing. *PLoS One* 8:e53400
- Wang M, Beelen R, Bellander T, Birk M, Cesaroni G, Cirach M, Cyrus J, de Hoogh K, Declercq C, Dimakopoulou K, Eeftens M, Eriksen KT, Forastiere F, Galassi C, Grivas G, Heinrich J, Hoffmann B, Ineichen A, Korek M, Lanki T, Lindley S, Modig L, Molter A, Nafstad P, Nieuwenhuijsen MJ, Nystad W, Olsson D, Raaschou-Nielsen O, Ragettli M, Ranzi A, Stempfelet M, Sugiri D, Tsai MY, Udvardy O, Varro MJ, Vienneau D, Weinmayr G, Wolf K, Yli-Tuomi T, Hoek G, Brunekreef B (2014) Performance of multi-city land use regression models for nitrogen dioxide and fine particles. *Environ Health Perspect* 122:843–849
- Wu CF, Lin HI, Ho CC, Yang TH, Chen CC, Chan CC (2014a) Modeling horizontal and vertical variation in intraurban exposure to PM_{2.5} concentrations and compositions. *Environ Res* 133:96–102
- Wu SW, Deng FR, Wei HY, Huang J, Wang X, Hao Y, Zheng CJ, Qin Y, Lv HB, Shima M, Guo XB (2014b) Association of cardiopulmonary health effects with source-apportioned ambient fine particulate in Beijing, China: a combined analysis from the Healthy Volunteer Natural Relocation (HVNRL) study. *Environ Sci Technol* 48:3438–3448
- Xie GD, Li WH, Xiao Y, Zhang BA, Lu CX, An K, Wang JX, Xu K, Wang JZ (2010) Forest ecosystem services and their values in Beijing. *Chin Geogr Sci* 20:51–58. doi:10.1007/s11769-010-0051-y
- Yu L, Wang G, Zhang R, Zhang L, Song Y, Wu B, Li X, An K, Chu J (2013) Characterization and source apportionment of PM_{2.5} in an urban environment in Beijing. *Aerosol Air Qual Res* 13:574–583. doi:10.4209/aaqr.2012.07.0192
- Zhang M, Wang X, Chen J, Cheng T, Wang T, Yang X, Gong Y, Geng F, Chen C (2010) Physical characterization of aerosol particles during the Chinese New Year's firework events. *Atmos Environ* 44:5191–5198
- Zhang Y, Guo Y, Li G, Zhou J, Jin X, Wang W, Pan X (2012) The spatial characteristics of ambient particulate matter and daily mortality in the urban area of Beijing, China. *Sci Total Environ* 435:14–20
- Zhang A, Qi Q, Jiang L, Zhou F, Wang J (2013) Population exposure to PM_{2.5} in the urban area of Beijing. *PLoS One* 8:e63486
- Zhao X, Zhang X, Xu X, Xu J, Meng W, Pu W (2009) Seasonal and diurnal variations of ambient PM_{2.5} concentration in urban and rural environments in Beijing. *Atmos Environ* 43:2893–2900
- Zhao H, Che H, Zhang X, Ma Y, Wang Y, Wang H, Wang Y (2013a) Characteristics of visibility and particulate matter (PM) in an urban area of Northeast China. *Atmos Pollut Res* 4:427–434. doi:10.5094/apr.2013.049
- Zhao PS, Dong F, Yang YD, He D, Zhao XJ, Zhang WZ, Yao Q, Liu HY (2013b) Characteristics of carbonaceous aerosol in the region of Beijing, Tianjin, and Hebei, China. *Atmos Environ* 71:389–398