Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Application of land use regression for estimating concentrations of major outdoor air pollutants in Jinan, China*

Li CHEN[†1,2,3], Shi-yong DU[4], Zhi-peng BAI[†‡1,2], Shao-fei KONG[1,2],
Yan YOU[1,2], Bin HAN[1,2], Dao-wen HAN[4], Zhi-yong LI[1,2]

([1]*College of Environmental Science and Engineering, Nankai University, Tianjin 300071, China*)
([2]*State Environmental Protection Key Laboratory of Urban Ambient Air Particulate Matter Pollution and Control, Tianjin 300071, China*)
([3]*College of Urban and Environmental Science, Tianjin Normal University, Tianjin 300387, China*)
([4]*Jinan Institute of Environmental Sciences, Jinan 250014, China*)
[†]E-mail: amychenli1981@126.com; zbai@nankai.edu.cn

Received Mar. 10, 2010; Revision accepted Aug. 11, 2010; Crosschecked Oct. 12, 2010

Abstract: SO_2, NO_2, and PM_{10} are the major outdoor air pollutants in China, and most of the cities in China have regulatory monitoring sites for these three air pollutants. In this study, we developed a land use regression (LUR) model using regulatory monitoring data to predict the spatial distribution of air pollutant concentrations in Jinan, China. Traffic, land use and census data, and meteorological and physical conditions were included as candidate independent variables, and were tabulated for buffers of varying radii. SO_2, NO_2, and PM_{10} concentrations were most highly correlated with the area of industrial land within a buffer of 0.5 km (R^2=0.34), the distance from an expressway (R^2=0.45), and the area of residential land within a buffer of 1.5 km (R^2=0.25), respectively. Three multiple linear regression (MLR) equations were established based on the most significant variables (p<0.05) for SO_2, NO_2, and PM_{10}, and R^2 values obtained were 0.617, 0.640, and 0.600, respectively. An LUR model can be applied to an area with complex terrain. The buffer radii of independent variables for SO_2, NO_2, and PM_{10} were chosen to be 0.5, 2, and 1.5 km, respectively based on univariate models. Intercepts of MLR equations can reflect the background concentrations in a certain area, but in this study the intercept values seemed to be higher than background concentrations. Most of the cities in China have a network of regulatory monitoring sites. However, the number of sites has been limited by the level of financial support available. The results of this study could be helpful in promoting the application of LUR models for monitoring pollutants in Chinese cities.

Key words: Land use regression (LUR), Air pollution, Background concentration, Geographic information system (GIS)
doi:10.1631/jzus.A1000092 **Document code:** A **CLC number:** X51

1 Introduction

Numerous studies have shown that land use regression (LUR) models can be applied to obtain accurate, small scale air pollutant concentrations without a detailed pollutant emission inventory (Briggs *et al*., 2000; Brauer *et al*., 2003). LUR models aim to illustrate the spatial distribution of air pollutant concentrations based on nearby traffic, land use and other variables. Provided that related geographic data are obtainable multiple linear regression (MLR) models developed using existing monitoring sites can then in most instances be applied to unmonitored sites (Ross *et al*., 2007).

Briggs *et al*. (1997) first applied an LUR model to air pollution mapping in small area variations in air quality and health (SAVIAH). In the past ten years, numerous epidemiological studies have adopted LUR models (Briggs, 2005; Slama *et al*., 2007; Ryan *et al*.,

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (No. 20677030), the Development Plan of Key National Fundamental Research (No. 2011CB503801), and the Special Research Funds for Science Development in Jinan (No. 200904015), China

2008). Studies on the application of LUR models were carried out initially in Europe, then later in North America and Japan (Gilbert *et al*., 2005; Ross *et al*., 2007; Wheeler *et al*., 2008; Kashima *et al*., 2009; Mukerjee *et al*., 2009). Several recent studies have indicated that LUR models can be used to predict air pollutant concentrations (Briggs *et al*., 2000; Brauer *et al*., 2003; Henderson *et al*., 2007; Hoek *et al*., 2008; Mavko *et al*., 2008). Furthermore, other studies have considered the impact of meteorological and physical conditions while predicting air pollutant concentrations (Arain *et al*., 2007; Su *et al*., 2007; Mavko *et al*., 2008; Ryan *et al*., 2008).

The regional background concentration is a very important index for environmental air quality management and atmospheric chemistry. Hoek *et al*. (2008) stated that it can generally be assumed that the constant in the LUR model represents the (invariant) regional background concentration. High background concentrations and intercepts can be similar if only source strength variables are in the MLR equations for soot and even for $PM_{2.5}$ (Hoek *et al*., 2008). Three approaches were used to predict the fine spatial distribution of background air pollution across the European Union for NO_2, PM_{10}, O_3, SO_2, and CO (Beelen *et al*., 2009).

Transportation and the patterns of land use are complicated in Chinese cities, and the applicability of LUR to predict SO_2, NO_2, and PM_{10} is unknown. This application could contribute to the wide use of LUR models in Asia. The aim of our study is to enlarge the application extension and to assess the improvement in accuracy of our LUR model. We have five component tasks: (1) to predict the spatial distributions of SO_2, NO_2, and PM_{10} in the study area; (2) to test whether it is possible to construct an LUR model with regulatory monitoring data and to identify which air pollutant concentrations can be predicted most accurately; (3) to determine appropriate buffer radii for the independent variables such as traffic, land use and census data for SO_2, NO_2, and PM_{10}; (4) to examine whether the LUR model is applicable in the study area with complicated meteorological and physical conditions, and (5) to discuss the relationship between the values of the background concentrations of regional air pollutants and the intercept values of the MLR equations derived from LUR.

2 Method

2.1 Study area and overview

Located at 36°40′ N and 117º00′ E, Jinan is in the northwest of Shangdong Province, China. It borders Mount Tai to the south and the Yellow River to the north. The land slopes downwards gradually from south to north. Jinan is the capital of Shandong Province and has six districts and three counties. The study area included Lixia and part of the Tianqiao, Huaiyin, Shizhong, and Liyu districts. The boundary of study area was an expressway. Jinan covers an area of 8177 km^2 including 3257 km^2 of urban settlement. Typically, there are four distinct seasons in Jinan with a long winter and summer and a relatively short spring and autumn. Average annual temperature is 13–14 °C and average annual precipitation is 654 mm. The total population of Jinan was 60.485 million at the end of 2007, and the population density was 740 person/km^2. The major industries in Jinan include traffic and mechanical equipment manufacturing, electronic information, metallurgy, iron steel, petrochemical, food and medicine, construction, and service industries. By the end of 2007, there were 0.996 million automobiles in Jinan. The increasing fuel consumption and number of construction projects, and the continuous growth of the population have brought serious air pollution problems to the city where PM_{10} has been found to be the principal atmospheric pollutant.

In this study, air pollutant concentrations from monitoring sites were used as dependent variables in LUR, while the surrounding land use, transportation and other data obtained from the geographic information system (GIS) and included in the regression equation were taken as independent variables. We have assembled a database of information on traffic, land use, and census data, and information on local meteorological and physical conditions from around the monitoring sites in Jinan, China (Fig. 1).

2.2 Dependent variables: ambient SO_2, NO_2, and PM_{10} monitoring data

The purpose of establishing a regional environmental air monitoring network is to confirm the possible high levels of air pollutants, the influence of the most important sources of air pollutions on environmental air quality, the background level of air

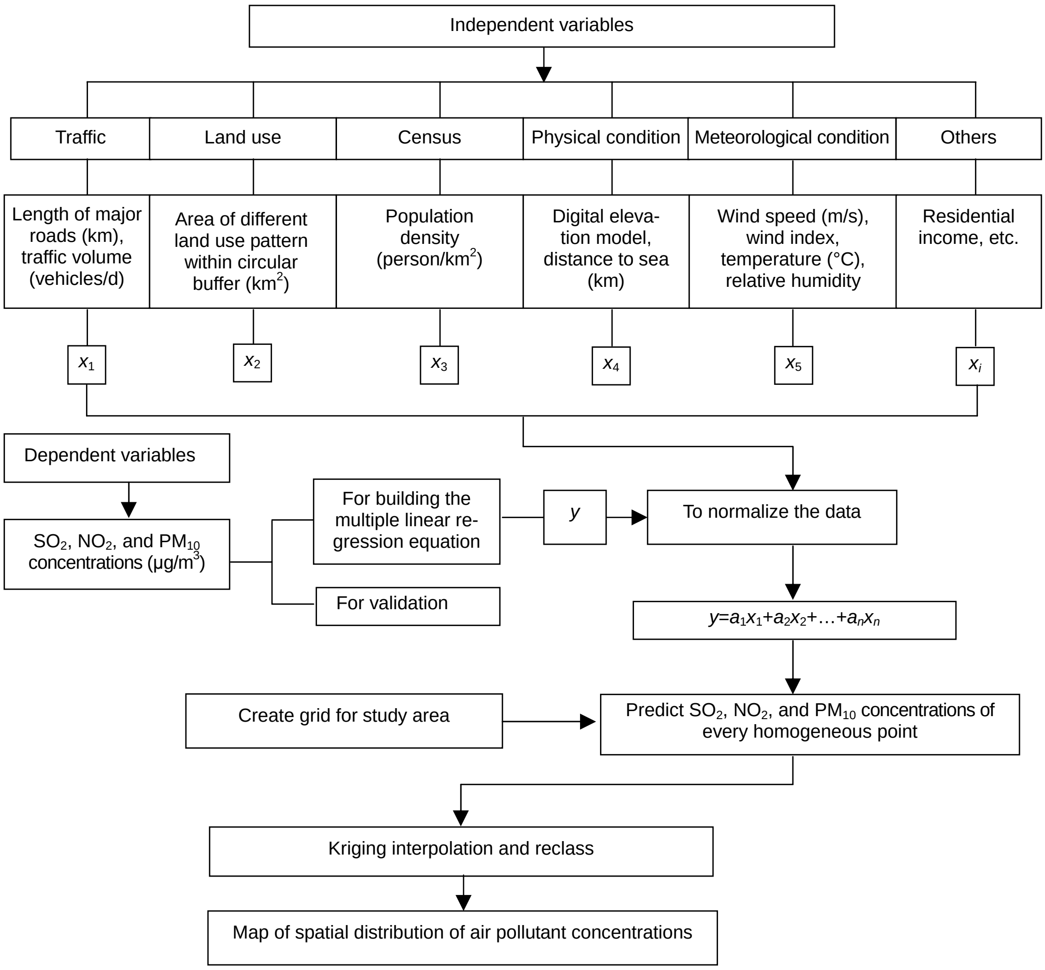

Fig. 1 Flowchart for the land use regression model

pollutants, the variation in trends for air pollutants, and to provide support for organizing the deployment of local air pollution control systems.

Monitoring SO_2 (API-110, ultraviolet fluorescence), NO_2 (API-210, chemiluminescence), and PM_{10} (BAM-1020, β ray absorbing method) data from regulatory monitoring sites operated and managed by the Jinan Institute of Environmental Sciences, China were used. It is essential to ensure that the monitoring data are scientific and reliable in QA/QC (quality assurance/quality control) procedures. QA/QC procedures at each site were followed according to the automated methods for ambient air quality monitoring issued by the Ministry of Environmental Protection, China. SO_2 and NO_2 concentrations were recorded every 1 min, and PM_{10} concentration was recorded every 1 h. The automated monitoring instrument took samples throughout each day (24 h).

The spatial distribution of air pollutants concentrations was measured from Aug. 1, 2008 to July 31, 2009 at 14 monitoring sites in the study area. The monitoring sites were representative of areas with greater population density, different land use types and heavy traffic. For each monitoring site, the mean, median, minimum, maximum and standard deviation of each pollutant concentration were calculated.

Autocorrelations and the distribution of air pollutant concentrations were explored using Geoda095i and SPSS13.0.

2.3 Independent variables: traffic, land use, census data, meteorological, and physical variables

All independent variables were assembled to form a database (Fig. 2). Circular buffers around each monitoring site with varying radii (0.5, 1, 1.5, and 2 km) were generated. All GIS layers and buffer layers were intersected. Traffic, land use and census data in each buffer surrounding each monitoring site were calculated based on Arcgis 9.2.

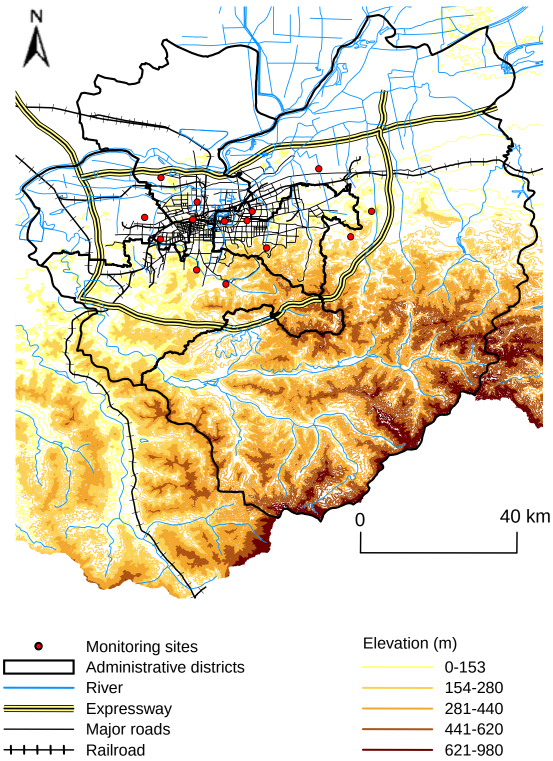

Fig. 2 Study area, air quality monitoring sites

2.3.1 Traffic data

The traffic data was obtained from the Jinan Institute of Environmental Sciences, China. The mobile source variables considered included the total length of major roads, the traffic volume on major roads, and the distance from each monitoring site to the expressway. The traffic volume data was incomplete, as minor roads and local street volumes were not included.

2.3.2 Land use data

The land use data was obtained from the Institute of Surveying and Mapping Survey, China. In this study, land use patterns included residential land, industrial land, agricultural land, commercial land, water area, and woodland.

2.3.3 Census population data

The population density in the buffer area of all monitoring sites was obtained based on the census data from the Jinan Statistical Yearbook 2008. For buffers that spanned more than one district, the population density was weighted according to the area of each administrative district.

2.3.4 Meteorological conditions

Wind direction is an important factor affecting the air pollutant concentration at a monitoring site. A wind index was calculated according to Ryan (2008). The Euclidean direction from the nearest source was derived to illustrate the relationship between the location of each monitoring site, the wind direction and the nearest traffic source. The derived wind index was the derived based on a cosine transformation (McCune and Keon, 2009):

$$\text{wind index}=[1-\cos(\theta-a)]/2,$$

where θ is the Euclidean direction from the nearest traffic source to the monitoring site in degrees east of north, α is the dominant wind direction. The derived wind index ranges from 0 to 1. When a monitoring site is located upwind of the nearest traffic source, the wind index equals to 0. However, if a monitoring site is directly downwind of the nearest traffic source, the wind index equals 1.

The temperature, humidity, and wind speed of each monitoring site were also considered in this study. The annual averages of these variables at each monitoring site were used.

2.3.5 Physical conditions

The shortest distance to the sea from each monitoring site was calculated. This study area included plains, hills, and mountains, and elevation was taken into account in the LUR model.

2.4 Land use regression model development

A univariate regression was made between the dependent variable and each independent variable, and the most significant were included in the MLR equations (Table 1). Since variables relating to traffic, land use and census data surrounding the monitoring sites within different buffer radii (e.g., length of major roads within 1.5 and 2 km of a monitoring site) might be colinear, the same traffic, land use and census variables within one buffer radius were included in the final MLR model. Five monitoring sites were selected randomly based on land use patterns (one on industrial land, one on agricultural land, one on commercial land and two on residential land), and then removed from the initial LUR model for later model validation. Thus, the LUR models initially developed were composed of 9 monitoring sites. Comparison was made between the estimated data and the measured concentrations at these five monitoring sites. Finally, data from all 14 monitoring sites in the LUR model were recalculated again after data from the five monitoring sites were returned to the full dataset.

Table 1 Independent variables considered for inclusion in LUR model

Air pollutant	Variable (buffer radii)	R^2	p
SO_2	Traffic volume (0.5 km)	0.29	0.01
	Area of industrial land (0.5 km)	0.34	<0.01
	Distance to sea	0.27	0.01
NO_2	Length of major roads (2 km)	0.39	<0.01
	Distance to expressway	0.45	<0.01
	Area of residential land (2 km)	0.28	0.01
PM_{10}	Area of residential land (1.5 km)	0.25	0.01
	Area of industrial land (1.5 km)	0.23	0.02
	Distance to sea	0.19	0.03

The spatial autocorrelations for SO_2, NO_2, and PM_{10} values were tested using Geoda095i. A four-nearest neighbor approach was used in this study. Statistical significance was derived with 999 repetitions of a permutation test.

Using a bootstrap, the sensitivities of model parameters were analyzed. In each iteration, three randomly selected monitoring sites were excluded and the coefficients were recorded. Data from the three sites were then returned, and all of the above steps were repeated 10000 times to find the stability of the independent variable coefficients.

A smooth surface of interpolated predications was created based on the MLR equation and the Kriging method for visualization.

3 Results

3.1 Descriptive statistics for air pollutants

Annual average concentrations of SO_2, NO_2, and PM_{10} were normally distributed (Figs. 3a–3c) with a mean for SO_2 of 68 μg/m^3 (standard deviation=15, median=67, min=46, max=96), for NO_2 of 41 μg/m^3 (standard deviation=19, median=34, min=20, max=70), and for PM_{10} of 144 μg/m^3 (standard deviation=35, median=137, min=103, max=237). With data from all validation monitoring sites included, all statistics were based on the complete dataset.

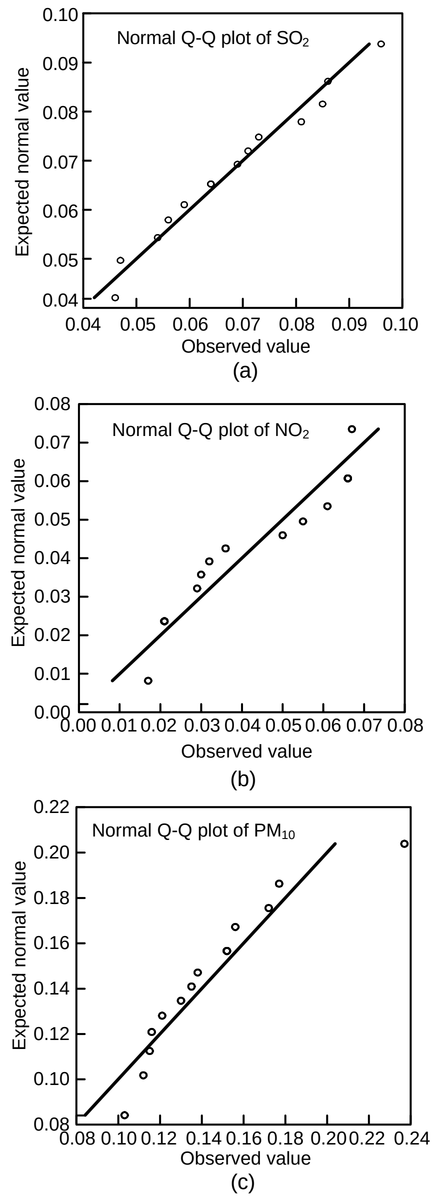

Fig. 3 Distributions of air pollutants concentrations. (a) SO_2; (b) NO_2; (c) PM_{10}

3.2 Regression model building and results

Pearson correlation coefficients were calculated between the length of major roads within different buffer radii (from 100 m to 5 km) and NO_2 concentration. The length of major roads within a buffer of 2 km was most highly correlated with NO_2 concentration based on the *p* value in this study. Pearson correlation coefficients were also calculated between the area of industrial land within different buffer radii (from 100 m to 5 km) and SO_2 concentration. The area of industrial land within a buffer of 0.5 km was most highly correlated with SO_2 concentration based on the *p* value in this study.

The independent variables most significantly correlated with SO_2, NO_2, and PM_{10} concentrations were area of industrial land (0.5 km), distance to expressway, and area of residential land (1.5 km), which explained 34%, 45%, and 25%, respectively of the variation (Table 1). Independent variables significantly correlated ($p<0.05$) with air pollutant concentrations were considered for inclusion in the final MLR equations.

The regression models built with independent variables are summarized in Table 2. In this study, Moran's *I* values were −0.265 (p=0.052), −0.097 (p=0.461), 0.012 (p=0.198) for SO_2, NO_2, and PM_{10}, respectively. The predictive performance was moderate, with R^2=0.617, R^2=0.640, and R^2=0.600 for SO_2, NO_2, and PM_{10}, respectively. The wind index, as an important variable for NO_2 in this study, when included in the LUR model led to a 0.06 increase in the R^2 value.

Measured annual average concentrations at nine monitoring sites ranged from 46 to 96 μg/m^3, 17 to 67 μg/m^3, and 103 to 237 μg/m^3 for SO_2, NO_2, and PM_{10}, respectively. Fig. 4 compares predicted SO_2,

Table 2 Multiple linear regression results for the association between dependent and independent variables

	β	t	p	VIF
Model: SO_2				
Intercept	68	3.473	0.006	1.03
Traffic volume (0.5 km)	12	2.908	0.016	1.32
Area of industrial land (0.5 km)*	9	2.864	0.017	1.21
Distance to sea	−8	−2.745	0.021	1.34
Final LUR model including five validation sites (R^2=0.617)				
Model: NO_2				
Intercept	41	7.760	0.000	1.04
Length of major roads (2 km)	16	4.578	0.003	1.21
Distance to expressway	−8	−2.456	0.018	1.33
Area of residential (2 km)	−22	−2.807	0.019	1.24
Final LUR model including five validation sites (R^2=0.640)				
Model: PM_{10}				
Intercept	142	5.453	0.002	1.12
Area of residential (1.5 km)	−14	−2.612	0.016	1.32
Area of industrial land (1.5km)	19	2.628	0.025	1.41
Distance to sea	−18	−6.705	0.007	1.27
Final LUR model including 5 validation sites (R^2=0.600)				

* Buffer radii; β is the coefficient of independent variable. These are the t-statistics and their associated 2-tailed p-values used in testing whether a given coefficient is significantly different from zero. α=0.05. VIF: variance inflation factor

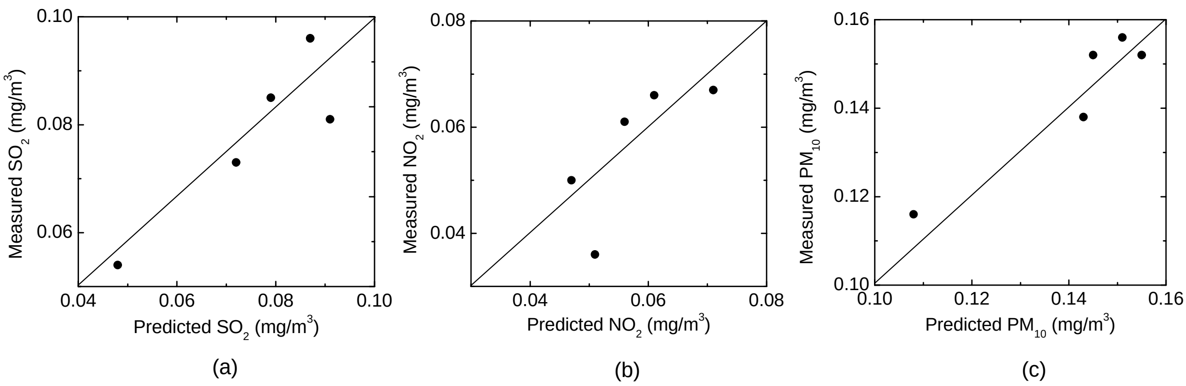

Fig. 4 Measured versus predicted air pollutants concentrations. (a) SO_2; (b) NO_2; (c) PM_{10}

NO_2, and PM_{10} concentrations to measured concentrations at five of the 14 validation sites. The predicted and measured concentrations were well correlated.

In leave-one-out evaluation, only 13 monitoring sites were used for prediction in each run. This procedure was repeated continuously until all 14 sites had been selected. Mean errors for the SO_2, NO_2 and PM_{10} data were 0, but their standard deviations were 0.35, 0.40, and 0.44, respectively, of the monitoring sites' data means.

3.3 Bootstrap

The results of the LUR model based on bootstrap were relatively stable in respect of the selected samples (Fig. 5). A bimodal distribution could be observed, with one peak representing the traffic volume (0.5 km) variables in the SO_2 model and the other representing the length of major roads (2 km) variable in the NO_2 model. In the SO_2 model, deviation from normality was caused by a single sample that had the highest SO_2 concentration (96 μg/m^3) of all samples. In the NO_2 model, deviation from normality was caused by a single sample that had the lowest NO_2 concentration (17 μg/m^3) of all samples. As it was unreasonable to believe that the underlying data for the two monitoring sites were inaccurate, they were not excluded.

3.4 Kriging of predicted SO_2, NO_2, and PM_{10} concentrations

The origin of coordinates at latitude 36°32′27″ N and longitude 116°50′56″ E was assigned, and a grid of 1 km×1 km in the study area was drawn (Figs. 6a–6c). SO_2, NO_2, and PM_{10} concentrations were calculated at each intersection point using MLR equations. The annual spatial concentration distributions of SO_2, NO_2, and PM_{10} were interpolated using the Kriging approach. The interpolation maps were

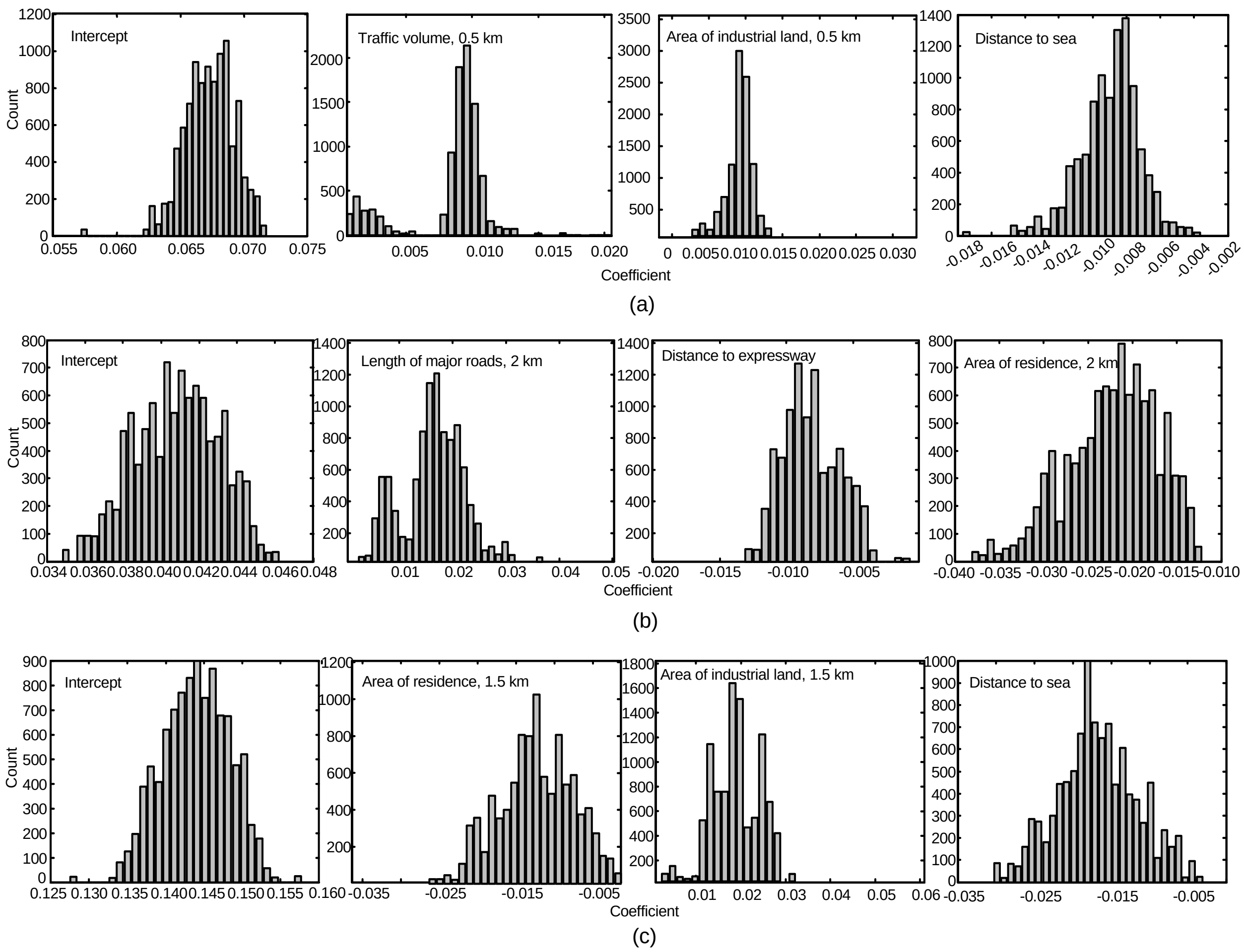

Fig. 5 Bootstrap of final parameters. (a) SO_2; (b) NO_2; (c) PM_{10}

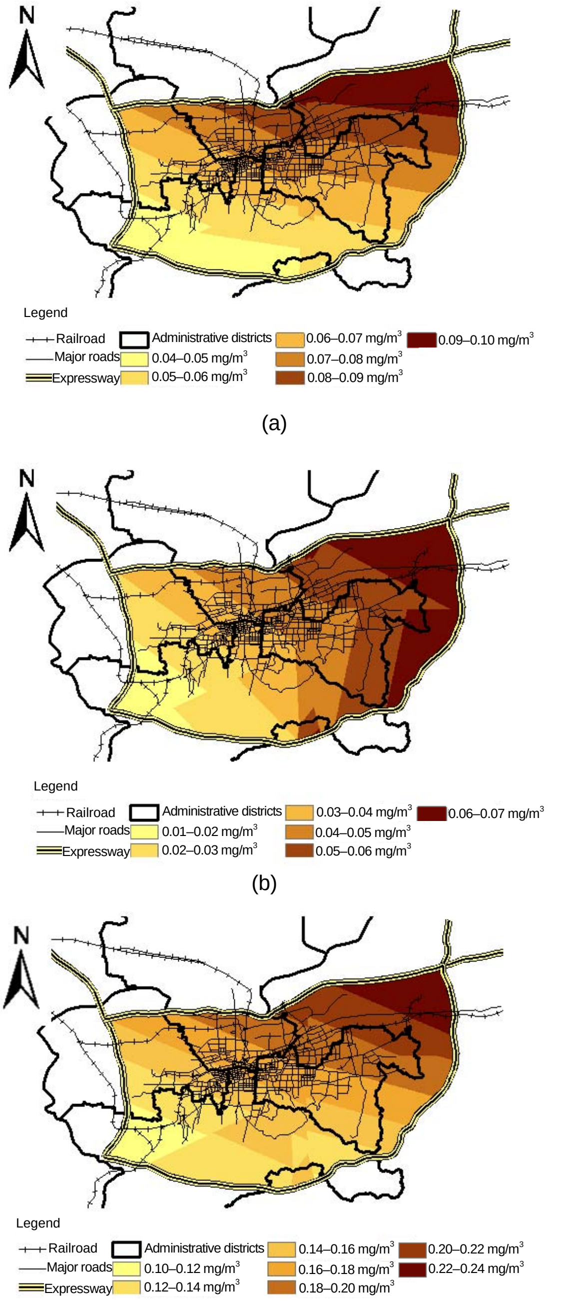

Fig. 6 LUR model applied in Jinan, China. Resolution was 10 km. (a) SO_2; (b) NO_2; (c) PM_{10}

reclassified using a manual method based on the concentration gradients of SO_2, NO_2, and PM_{10}.

There was a gradual decline in predicted concentrations from northeast to southwest in the study area for SO_2, NO_2, and PM_{10}. Most industrial areas are located in the northeast of the study area, and many green areas are located in the west. This suggests that industrial areas are still a major pollutant sources in Jinan, while woodland areas may play a role in decreasing air pollutant concentrations.

3.5 Background concentration versus intercepts of MLR equations

The intercepts of the MLR equations reflected air pollutant background concentrations in the study area (Hoek *et al*., 2008). Table 2 shows that the intercepts of MLR equations for SO_2, NO_2, and PM_{10} were 68, 41, and 142, respectively. SO_2 concentrations at seven monitoring sites were lower than 68 μg/m^3, NO_2 concentrations at eight monitoring sites were lower than 41 μg/m^3, and PM_{10} concentrations at eight monitoring sites were lower than 142 μg/m^3. It is likely that the intercept values of MLR equations are higher than the values of background concentrations.

4 Discussion

LUR models with a combination of traffic, land use, census, meteorological and physical variables were developed for predicting SO_2, NO_2, and PM_{10} concentrations in Jinan. More than 60% of the variability in SO_2, NO_2, and PM_{10} concentrations could be explained by these models, and the concentrations were generally within 20% of the actual values.

With regulatory air quality data, LUR models could successfully predict air pollutant concentrations. The predictions of traffic-related pollutant concentrations in Chinese cities were closely related to regulatory air quality data. Concentrations of SO_2, NO_2, and PM_{10} could be modeled moderately well from traffic, industrial land and census variables, respectively.

An LUR model applied to the prediction of SO_2 concentrations by Wheeler *et al*. (2008) gave an R^2 value of 0.69. In our study, the R^2 of the LUR model for SO_2 was 0.617. Hoek *et al*. (2001) used an LUR model to predict the NO_2 concentration in the Netherlands, and obtained an R^2 of 0.85 (the highest R^2 in previous studies of NO_2). The lowest R^2 for NO_2 from an LUR model was 0.54 (Gilbert *et al*., 2005). The R^2 for NO_2 from our model was 0.640. An LUR model applied to the prediction of PM_{10} concentrations in London gave R^2 values from 0.45–0.60 (Briggs *et al*., 2010). In our research, the R^2 for PM_{10} was 0.600.

According to R^2 values, the performance of LUR models has not been ideal, but was acceptable in our study. To improve the performance of the MLR equations, data from more monitoring sites should be included in the LUR models and more independent variables should be considered.

With a mountain in southern Jinan, the topography of our study area was complicated, and this reflects the general topography of Shandong Province. There are more than 3000 km^2 of mountains and hills and 5000 km^2 of plains in Jinan, and the altitude ranges from 8.7 to 989 m in Jinan. Nevertheless, the LUR model could still be applied successfully in such a complicated study area.

Most previous LUR models have not included a wind index, wind speed, temperature, and relative humidity due to lack of data at the monitoring sites. Jerrett *et al*. (2005) mentioned the application of an LUR model including wind direction in a developed model. Arain *et al*. (2007) described the use of high-resolution interpolated wind direction data in an LUR model in Canada. Furthermore, we noted that meteorological variables are required to capture local scale atmospheric patterns, so in our study, the meteorological data of each monitoring site were used. Although in LUR models meteorological variables are very important, when included, the R^2 of the overall models did not increase sharply. When a wind index was included in this study, the R^2 value increased from 0.640 to 0.700 in the LUR model for NO_2. A 0.03 increase resulted from the inclusion of wind speed in the LUR model since it was an important variable for PM_{10} in this study.

The buffer radii of independent variables for SO_2, NO_2, and PM_{10} were chosen to be 0.5, 2, and 1.5 km, respectively based on the univariate models. For SO_2, inclusion of independent variables within a 1, 1.5, and 2 km buffer and for PM_{10}, inclusion of independent variables within a 2 km buffer did not result in significant variability in the univariate models. This suggests that the SO_2 concentration was influenced by local sources, such as industrial areas, which have no significant impact on the distant areas.

An insufficient number of the monitoring sites have been one of the limitations of this study. There were only 14 regulatory monitoring sites in the study area, and this may have resulted in errors in the model. To estimate the stability of the LUR model, bootstrap was used. The results showed that the LUR model was relatively stable. The distance to major roads and expressway may be evaluated as a non-linear function, but it is still considered to be a linear function because the final model is a multiple linear regression.

In most models residential land use has a positive slope, but in our model it is negative. We think there are three reasons. Firstly, the air quality of residential land is better than that of other place because of the air pollution controls implemented by the government. Secondly, the major air pollution source of PM_{10} is open source (bare soil and cropland) while in residential areas very little dust is raised because of the sealed surfaces. Finally, the major air pollution source of NO_2 is traffic, and the traffic volume is transferred from residential land to surrounding areas, especially during rush hour.

Background concentrations of air pollutants are important parameters in air quality management in China, and so far there is no good method estimation. The multi-year annual averages of the “clean” remote monitoring sites provide the data for the background concentration of pollutants in North America. Since the data reflect the effects of anthropogenic emissions from within North America as well as background, they might overestimate the actual background concentrations (USEPA, 2003). The performance of an urban atmospheric dispersion model used to estimate NO_x background concentration in Copenhagen, Denmark has been evaluated (Venegas and Mazzeo, 2002). Maps of background air pollution using universal Kriging, ordinary Kriging, and regression have been drawn by Beelen *et al*. (2009). Li *et al*. (2007) selected a control station for clean air located far away from the city to estimate the background concentrations of PM_{10} and SO_2 over the Beijing city region. The PM_{10} background concentration was 58–67 μg/m^3. Values of intercepts of MLR equations of PM_{10} and NO_2 in our study were found to be higher than the values of background concentrations. The reasons for overestimating background concentrations were as follows. Firstly, the number of monitoring sites in the study area was insufficient. Secondly, the independent variables included in this study did not include all the variables that can affect the concentrations of air pollutants. Thus, the intercept values of MLR equations may include the effect of some other independent variables. The intercepts were calculated considering only the source strength variables. For SO_2 model, the intercept value was 62, while the value from the MLR equation was 68. For the NO_2 model, the intercept value was 40, while the value from the MLR equation was 41. The background concentrations and intercept values could be similar if only the source strength variable was considered for SO_2 and NO_2 models, because the intercept values of MLR equations seemed to be higher than the actual values of background concentrations, and the intercept values considering only the source strength variables were lower than the intercept values from MLR equations in this study. However, LUR is still a promising method for evaluating air pollutant background concentrations.

Air pollutant sources are very complicated in China. LUR is superior to the other models in that it is able to predict the air pollutant concentration without depending on outdated or inaccurate source emission data. Finally, more monitoring sites should be used to confirm the importance of LUR models in epidemiological studies.

References

Arain, M.A., Blair, R., Finkelstein, N., Brook, J.R., Sahsuvaroglu, T., Beckerman, B., Zhang, L., Jerrett, M., 2007. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. *Atmospheric Environment*, **41**(16): 3453-3464. [doi:10.1016/j.atmosenv.2006.11.063]

Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., Hoogh, K.D., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. *Science of the Total Environment*, **407**(6):1852-1867. [doi:10.1016/j.scitotenv.2008.11.048]

Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., *et al*., 2003. Spatial variation in nitrogen dioxide in three European areas. *Science of the Total Environment*, **332**(1-3):217-230. [doi:10.1016/j.scitotenv.2004.04.014]

Briggs, D., 2005. The role of GIS: coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health, Part A*, **68**(13-14): 1243-1261. [doi:10.1080/15287390590936094]

Briggs, D., Hoogh, C.D., Gulliver, J., 2010. Comparative assessment of GIS-based methods and metrics for modeling exposure to air pollution. *Journal of Toxicology and Environmental Health, Part A*, in press

Briggs, D.J., Collins, S., Elliot, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., van Reeuwijk, H., Smallborne, K., van der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science*, **11**(7):699-718. [doi:10.1080/136588197242158]

Briggs, D.J., de Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., Samllbone, K., 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment*, **253**(1-3):151-167. [doi:10.1016/S0048-9697(00)00429-0]

Gilbert, N.L., Goldberg, M.S., Beckerman, B., Brook, J.R., Jerrett, M., 2005. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. *Journal of the Air Waste Management Association*, **55**:1059-1063.

Henderson, S.B., 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. *Environmental Science & Technology*, **41**(7):2422-2428. [doi:10.1021/es0606780]

Hoek, G., Meliefste, K., Brauer, M., Vliet, P.V., Brunekreef, B., Fischer, I., Lebret, E., Cyrys, J., Gehring, U., Heinrich, J., 2001. Risk Assessment of Exposure to Traffic-related Air Pollution for the Development of Inhalant Allergy, Asthma and Other Chronic Respiratory Conditions in Children (TRAPCA). Final Report, IRAS, University Utrecht, Utrecht.

Hoek, G., Beelen, R., Hoogh, K.D., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, **42**(33):7561-7578. [doi:10.1016/j.atmosenv.2008.05.057]

Jerrett, S., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, **15**(2):185-204. [doi:10.1038/sj.jea.7500388]

Kashima, S., Yarifuji, T., Tsuda, T., Doi, H., 2009. Application of land use regression to regulatory air quality data in Japan. *Science of the Total Environment*, **407**(8): 3055-3062. [doi:10.1016/j.scitotenv.2008.12.038]

Li, J.X., Yu, T., Zhao, Y., Li, H.J., Li, L.J., Liu, W.Q., 2007. Estimation of Beijing's atmospheric PM_{10} background concentration based on analysis of synoptic types. *Acta Scientiae Circumstantiae*, **27**(9):1525-1533 (in Chinese).

Mavko, M.E., Tang, B., George, L.A., 2008. A subneighborhood scale land use regression model for predicting NO_2. *Science of the Total Environment*, **398**(1-3):68-75.

McCune, B., Keon, D., 2009. Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science*, **13**(4):603-606. [doi:10.1111/j.1654-1103.2002.tb02087.x]

Mukerjee, S., Smith, L.A., Johnson, M.M., Neas, L.M., Stallings, C., 2009. Spatial analysis and land use regression of VOCs and NO_2 from school-based urban air monitoring in Detroit/Dearborn, USA. *Science of the*

Total Environment, **407**(16):4642-4651 [doi:10.1016/j.scitotenv.2009.04.030]

Ross, Z., Jerrett, M., Ito, K., Tempalski, B., Thurston, G.D., 2007. A land use regression for predicting fine particulate matter concentrations in the New York City region. *Atmospheric Environment*, **41**(11):2255-2268. [doi:10.1016/j.atmosenv.2006.11.012]

Ryan, P.H., LeMasters, G.K., Levin, L., Burkle, J., Biswas, P., Hu, S., Grinshpun, S., Reponen, T., 2008. A land-use regression model for estimating microenvironmental diesel exposure given multiple addresses from birth through childhood. *Science of the Total Environment*, **404**(1):139-147. [doi:10.1016/j.scitotenv.2008.05.051]

Slama, R., Morgenstern, V., Cyrys, J., Zutavern, A., Herbarth, O., Wichmann, H.E., Heinrich, J., 2007. Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model. *Environment Health Perspect*, **115**(9):1283-1292. [doi:10.1289/ehp.10047]

Su, J.G., Brauer, M., Ainslie, B., Steyn, D., Larson, T., Buzzelli, M., 2007. An innovative land use regression model incorporating meteorology for exposure analysis. *Science of the Total Environment*, **390**(2-3):520-529. [doi:10.1016/j.scitotenv.2007.10.032]

USEPA (US Environmental Protection Agency), 2003. Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. Office of Air Quality Planning and Standards Staff Paper, First Draft 2-53.

Venegas, L., Mazzeo, N., 2002. An evaluation of DAUMOD model in estimating urban background concentrations. *Water, Air, and Soil Pollution Focus*, **2**(5-6):433-443. [doi:10.1023/A:1021376510644]

Wheeler, A.J., Smith-Doiron, M., Xu, X., Gilbert, N.L., Brook, J.R., 2008. Intra-urban variability of air pollution in Windsor, Ontario measurement and modeling for human exposure assessment. *Environmental Research*, **106**(1):7-16. [doi:10.1016/j.envres.2007.09.004]