Environmental Research 158 (2017) 542–552

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Development of land use regression models for $PM_{2.5}$, SO_2, NO_2 and O_3 in Nanjing, China

Lei Huang[a,b,*], Can Zhang[a], Jun Bi[a,**]

[a] State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Xianlin Campus, Box 624, 163 Xianlin Avenue, Nanjing 210023, China

[b] Lamont-Doherty Earth Observatory, Columbia University, P.O. Box 1000, 61 Rt. 9W, Palisades, NY 10964, USA

ARTICLE INFO

Keywords:
Ambient pollution
Land use regression
Simulation
Air quality monitoring networks
Spatial analysis

ABSTRACT

Ambient air pollution has been a global problem, especially in China. Comparing with other methods, Land Use Regression (LUR) models can obtain air pollutant concentration distribution at finer scale without the air pollution source data based on a few monitoring sites and predictors. However, limited LUR studies have been conducted on the basis of regular monitoring networks. Thus, we explored the applicability of conducting LUR models for four key air pollutants: $PM_{2.5}$, SO_2, NO_2 and O_3, on the basis of national monitoring networks which have good representation of areas with different characteristics in Nanjing, China. Fifty-nine potential predictor variables were considered, including land use type, population density, traffic emission, industrial emission, geographical coordinates, meteorology and topography. LUR models of these four air pollutants were with good explained variance for four key air pollutants. Adjusted explained variance of the LUR models was highest for NO_2 (87%), followed by SO_2 (83%), and was lower for $PM_{2.5}$ (72%) and O_3 (65%). Annual average distributions of pollutants in 2013 were obtained based on predicted values, which revealed that O_3 in Nanjing was more heavily impacted by regional influences. This study would not only contribute to the wider use of LUR studies in China but also offer important reference for the application of regular monitoring network with high representativeness in LUR studies. These results would also support for air epidemiological studies in the future.

1. Introduction

Ambient air pollution has been a global problem and a large number of epidemiological studies have shown adverse effects caused by ambient air pollution on public health, especially the influence of the fine particulate matter on the morbidity and mortality of cardiopulmonary diseases (Adar and Kaufman, 2007; Englert, 2004; West et al., 2016). China now is under the rapid development of industrialization with extensive production and high emission load, the air quality is difficult to be substantially improved. Meanwhile, with the constantly accelerating urbanization, population and city size are enlarging and also the aging population problem is becoming increasingly acute. All of these above make China face with exceptionally serious air problem (Huang et al., 2013a). The report of the Global Burden of Disease shows that fine particulate matter ($PM_{2.5}$) is the seventh largest important death risk factor in the world and the fourth largest important death risk factor in China (Cohen et al., 2005; Stephen S Lim et al., 2012). In China, the air pollution in 2010 resulted in 25.23 million loss of disability adjusted life years (Yang et al., 2013). Many epidemiologic studies have been conducted in China to characterize the adverse health effects of air pollution using different health outcomes. However, previous epidemiologic studies have largely used measured pollutant concentrations at fixed monitors to represent the exposure of large populations which ignored the spatial variations of air pollutants (Gehring et al., 2013; West et al., 2016). Various studies have demonstrated that large within-city contrasts in ambient air pollution in urban regions may result in the error of the estimate health effect and even the misclassification of exposure risk in the traditional epidemiologic studies (Punger and West, 2013). There is also evidence from epidemiologic studies that within-city contrasts of particulate matter air pollution are larger than between-city (Miller et al., 2007). So it is of great significance to assess the within-city concentrations of key air

Abbreviations: $PM_{2.5}$, particles with aerodynamic diameter less than 2.5 μm; SO_2, sulfur dioxide; NO_2, nitrogen dioxide; O_3, ozone

* Corresponding author at: State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Xianlin Campus, Box 624, 163 Xianlin Avenue, Nanjing 210023, China.

** Corresponding author at: State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Xianlin Campus, Box 624, 163 Xianlin Avenue, Nanjing 210023, China.

E-mail addresses: huanglei@nju.edu.cn (L. Huang), jbi@nju.edu.cn (J. Bi).

http://dx.doi.org/10.1016/j.envres.2017.07.010
Received 22 May 2017; Received in revised form 4 July 2017; Accepted 5 July 2017
0013-9351/

pollutants, especially in China.

Methods for air pollutant concentration simulation mainly include interpolation methods such as kriging and inverse distance weighing (IDW) interpolation, conventional dispersion models, chemical transport models, inversion of satellite remote sensing image and land use regression (LUR) models (Anenberg et al., 2010; Bravo et al., 2012; Briggs, 2005; Hoek et al., 2008; Hu et al., 2017, 2014b, 2014c, 2015; Jerrett et al., 2005; Laurent et al., 2014; Ma et al., 2016; Ostro et al., 2015; Sarnat et al., 2011). Interpolation methods are mechanistic and rely on the monitoring network with high density, so the application and also the accuracy of interpolation methods are limited (Jerrett et al., 2005). Conventional dispersion models and chemical transport models require many kinds of data including the air pollutant source data, topography data and meteorology data and all of them should be with high precision which make the simulation process complicated and high-cost (Solomos et al., 2015). Method of satellite remote sensing image inversion are mainly based the association between aerosol optical depth (AOD) and air pollutants concentrations and usually used to obtain the widespread concentration distribution, but this method is limited by the imaging time and spatial resolution of remote sensing image (Levy et al., 2007; Ma et al., 2016). Comparing with above methods, LUR models can obtain air pollutants concentration distributions at finer scale without the air pollution source data based on a few monitoring sites and predictors mainly extracted by geographic information system (GIS) technology, which is convenient and efficient (de Hoogh et al., 2014; Zou et al., 2009).

With the development of GIS technology, the application of LUR model becomes more and more extensive, especially in Europe and the North America and the simulation of ultrafine particles (UFP), NO_2, nitrogen oxide (NO_X) and volatile organic compounds (VOC) on the basis of LUR model have already been successfully conducted in Europe and the North America (Beelen et al., 2013; Cattani et al., 2017; Hoek et al., 2008; Olvera et al., 2012; Ross et al., 2007; van Nunen et al., 2017). While only a few LUR studies have been done in China due to the limited data accessibility and also the lack of data since a nationwide regulatory $PM_{2.5}$ monitoring network did not exist until the end of 2012 (Liu et al., 2016; Ma et al., 2016; Meng et al., 2015). Several LUR studies have been performed in Beijing, Tianjin, Shanghai, Jinan and Hong Kong, China (Chen et al., 2017; Hu et al., 2016; Li et al., 2010; Liu et al., 2016; Shi et al., 2016; Wu et al., 2015) while there are no relevant studies in the city of Nanjing. Considering that different cities differ in land use pattern, population distribution, terrain condition, meteorology condition and also the concentration distribution of air pollutants, it would be very meaningful to conduct LUR studies in the city of Nanjing to contribute to the wider use of LUR models in China.

Briggs et al. first introduced LUR models in Small Area Variations In Air quality and Health (SAVIAH) study and LUR models were called as "regression mapping" at first (Briggs et al., 1997). This air pollutant concentration simulation method develops linear regression models based on a small number of air pollution monitoring sites and potential predictor variables obtained by GIS technology which derive from different kinds of land use types. Then the model is applied to a large number of unsampled points in the study area which can be called "virtual monitoring site" and the whole air pollutant concentration surface can be obtained with fine spatial resolution. Thus for LUR models, the quality of monitoring data and the chose of predictors are of great importance for the model performance.

Except for the land use factors, other local specific variables, e.g., the altitude and meteorology, were also included in some studies to improve the explanation of the LUR models (Abernethy et al., 2013; Hoek et al., 2008; Tang et al., 2013). Methods of monitoring data acquisition in current studies mainly include purposed-designed monitoring and routine network, however, limited studies have made use of air pollution concentration data from routine monitoring network (Hoek et al., 2008). Measurements of air pollutants from routine networks are of low trust for the reason that routine networks in most urban areas are often focus on the potential hotspots such as heavily trafficked street and industrial areas, or preferentially placed away from hotspots, which introduce biased estimate of the public exposure to air pollution (Marshall et al., 2008). While in China, national monitoring networks were established in accordance with the principle of representing areas of different characteristics which would avoid the above problems of routine networks at a certain extent. The national monitoring network is regular monitoring network, so its measurements are continuous in temporal coverage (Meng et al., 2015). And the national monitoring network data also enable the public to obtain free, which is cost-effective. Meanwhile, in China, air pollutant concentration data obtained from the national monitoring sites are often used to represent the air pollution level of the whole city which would lead to estimate error in public air pollution exposure. So it would meaningful to conduct LUR studies on the basis of national monitoring network to explore the model performance, then an economical and efficient air pollutant concentration simulation method can be obtained.

This paper attempted to explore the applicability of conducting LUR models on the basis of national monitoring network which is routine network in the city of Nanjing, China. This would not only enrich the rare evidence of developing LUR models for air pollutants in China, but also offer important reference for the application of regular monitoring network with high representativeness in LUR studies. Meanwhile, on the basis of previous LUR studies, we also considered sufficient potential predictor variables including land use type, population density, traffic emission, industrial emission, geographical coordinates, meteorology and topography to explore the most efficient influencing factors of air pollutants. This paper would offer a basic systematic LUR method to promote the wider use in Chinese other cities. In addition, the spatial variations of four key air pollutants including $PM_{2.5}$, SO_2, NO_2 and O_3 in Nanjing at annual average level in 2013 were also explored, which would also contribute to the long-term epidemiological studies in air pollution.

2. Materials and methods

In this section, we first introduced the study area and the data used to train the LUR model. Then we elaborated the LUR model development and also the methods to evaluate the model performance. In LUR models, for each air pollutant, annual average concentration at each national monitoring site was selected as dependent variable. Fifty-nine potential independent variables included GIS predictor variables of different buffers (road length, residential land, agricultural land, green space, water body and population density) and local specific variables (location including latitude and longitude, temperature, 24r precipitation, relative humidity, wind speed, wind index, altitude and slope). LUR models were built and analyzed using statistical package (SPSS). ArcGIS 10.1 was used to analyze and visualize the result data. The detailed procedures for LUR model development are shown in Fig. 1.

2.1. Study area

Nanjing, the capital of Jiangsu province in China, is one of the important regional central cities in the core economic area of Yangtze River Delta (YRD) owning high population density and well-developed industry (Li et al., 2011). With the rapid economic development and continuous energy consumption, Nanjing has been suffering severe air pollution in recent years, which has been one of the most important environmental issues in Nanjing and a long-lasting haze episode has occurred in Nanjing and its surrounding areas from October 15–31, 2009 (Kang et al., 2013; Zu et al., 2017). Fine particles were reported to be the serious air pollutants in China and $PM_{2.5}$ has been demonstrated to be the key factor for air pollution improvement in the YRD region where Nanjing city lies in (Cheng et al., 2013; Hu et al., 2014a; van Donkelaar et al., 2010; Wang et al., 2014). Meanwhile O_3 pollution was also found to be serious in YRD region with the occurrence frequency of