JOURNAL OF ENVIRONMENTAL SCIENCES 59 (2017) 48–53

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jes

JES JOURNAL OF ENVIRONMENTAL SCIENCES

www.jesc.ac.cn

GIS-based analysis of population exposure to $PM_{2.5}$ air pollution—A case study of Beijing

Zhao Liu, Meihui Xie, Kun Tian, Peichao Gao*

Institute of Geospatial Information, Dep. of Civil Eng., Tsinghua Univ., Beijing 100084, China

ARTICLE INFO

Article history:
Received 12 December 2016
Revised 18 February 2017
Accepted 20 February 2017
Available online 31 March 2017

Keywords:
GIS
Population exposure
$PM_{2.5}$
Spatio-temporal variation

ABSTRACT

$PM_{2.5}$, formally defined as particulate matter with diameter less than 2.5 μm, is one of most harmful air pollutants threatening human health. Numerous epidemiological studies have shown that both short-term and long-term exposures to $PM_{2.5}$ are strongly linked with respiratory diseases. In this study, various types of spatio-temporal data were collected and used to estimate the spatio-temporal variation of $PM_{2.5}$ exposure in Beijing in 2014. The seasonal and daily variation of the population-weighted exposure level (PWEL) in 2014 was estimated and compared. The results show that the population exposure to ambient air pollution differs significantly in the four seasons, and the exposure levels in winter and spring are notably higher than the other seasons; the exposure level changes greatly from North to South, and each sub-district maintains similarity to neighboring sub-districts.

Introduction

Over the past few decades, along with the economic boom and the improvement of living standards, China has also experienced a sharp increase in air pollution, with the middle-eastern region being affected first (Che et al., 2009). Beijing, the capital of China, is also being affected by air pollution, especially fog and haze caused by automobile and industrial exhausts (Wang et al., 2006; Sun et al., 2013). There are 35 automatic air quality monitoring stations established by the government (Zhang et al., 2013), which can keep a close eye on the atmospheric conditions of Beijing and publish hourly monitoring data online for the public. Numerous epidemiological studies have shown that both short-term and long-term exposures to particulate matter with diameter less than 2.5 μm ($PM_{2.5}$) are strongly linked with human respiratory disease, heart disease, lung cancer, and so on (Gamble, 1998; Vallejo et al., 2006). Based on the epidemiological study results, there have also been some studies focused on the spatio-temporal variation of population exposure to $PM_{2.5}$. Li and Li collected $PM_{2.5}$ monitoring data continuously from October 2012 to September 2013 to assess the seasonal and spatial variation of $PM_{2.5}$ in Beijing (Li et al., 2015). Guo, Ma and He produced a new method to study the spatial distribution of regional ambient $PM_{2.5}$ and PM_{10} concentrations (Guo et al., 2009). Zhang and Qi collected 37 days of $PM_{2.5}$ monitoring data in Beijing during autumn 2012 to estimate the short-term population exposure to $PM_{2.5}$. These studies tried to measure the spatial and temporal variations of $PM_{2.5}$ concentrations using different methods from various aspects. Based on the existing research, this paper aimed to study the population exposure to $PM_{2.5}$ air pollution using a variety of spatio-temporal data. The remainder of this paper is structured as follows. Section 1 mainly introduces the different kinds of data used in this research, including the $PM_{2.5}$ monitoring data, the 2010 population census data in Beijing and two kinds of volunteered geographic information (VGI), the street data from the Open Street Map (OSM) and the Point of Interest (POI) data from the GeoHey platform. Section 2 spatializes the $PM_{2.5}$ concentration monitoring data and the population census data in Beijing

* *Corresponding author*. E-mail: xiemeihui1990@163.com (Meihui Xie).

http://dx.doi.org/10.1016/j.jes.2017.02.013

using a spatial interpolation method. The population-weighted exposure level (PWEL) was then used to estimate the potential exposure to $PM_{2.5}$ over the area of Beijing in 2014. Further study on the results of spatial and temporal of population exposure to $PM_{2.5}$ is discussed in Section 3. Section 4 summarizes the calculation results.

1. Data

1.1. $PM_{2.5}$ monitoring data

In order to keep a close eye on the atmospheric conditions in Beijing, the Beijing Municipal Environmental Protection Bureau (BJEPB) established 35 automatic air quality monitoring stations, and has created a public information website (http://www.bjmemc.com.cn/) which can provide hourly monitoring data for the public, from which researchers can obtain historical $PM_{2.5}$ monitoring data. In addition, the Geospatial Data Cloud (http://www.gscloud.cn/) established by the Chinese Academy of Sciences (CAS) has also collected the hourly $PM_{2.5}$ monitoring data for Beijing since 2014. Thus, we obtained the hourly $PM_{2.5}$ monitoring data from the Geospatial Data Cloud for the period January 10 to December 31, 2014. Due to various reasons, there are some missing data from different monitoring stations at different times, and the missing data were all replaced with the value of −9999, which can markedly influence the interpolation results. In these cases, we eliminated all the missing data (Fig. 1).

1.2. The 6th census data and sub-district boundary data in Beijing

The census data of sub-districts in China were collected from the 6th National Population Census in 2010 (National Bureau of Statistics of China, 2002), from which we selected just the area of Beijing City. There are 16 urban areas and 2 counties in Beijing with 326 census units, which are called sub-districts (jiedao). The total population in 2010 reached 19,612,368 in Beijing.

In order to spatialize the census data in different sub-districts, the boundaries of different sub-districts were also collected from websites provided by district and county governments of Beijing. After this, we used the tool ArcGIS for Desktop to delineate the boundary lines of each sub-district. The population of different sub-districts in China is shown in Fig. 2.

1.3. The street data and POI data

With the development of society and the economy, more and more people are not only users of geographic data, but also creators. VGI is a new phenomenon that has arisen in recent years (Goodchild and Li, 2012), which can provide a new method for people to understand the internal mechanisms of the economy, the humanities, and society. The OSM is a typical example of VGI, the information of which is created by participants from throughout the world (Haklay, 2010); all users are creators, but all people can also be users. The OSM data can portray cities in detail, which is often used in different research areas, such as urban planning, transportation planning, and so on.

In this paper, we obtained the road data of Beijing from OSM (http://www.openstreetmap.org/), and the POI data of Beijing from the website maintained by GeoHey (https://geohey.com/).

2. Methods

In order to analyze the population exposure to $PM_{2.5}$ air pollution, there are three main tasks. These include mapping high-resolution $PM_{2.5}$ concentrations in all of Beijing city, spatializing population census data at the sub-district level,

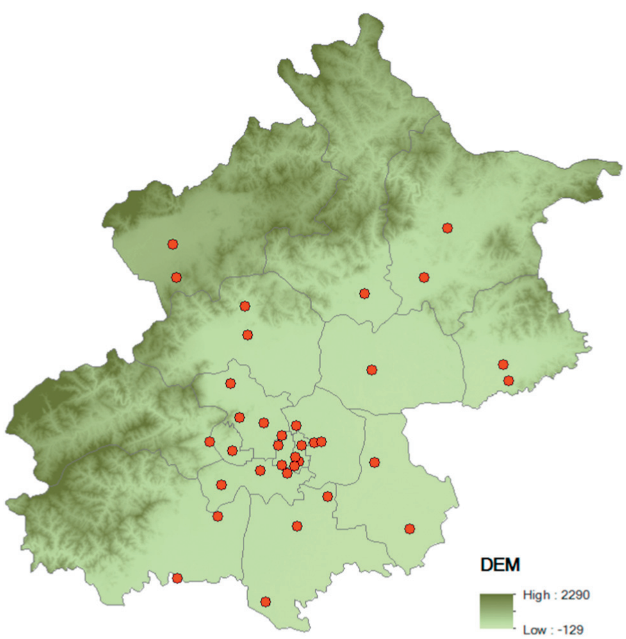

Fig. 1 – $PM_{2.5}$ monitoring stations in Beijing.

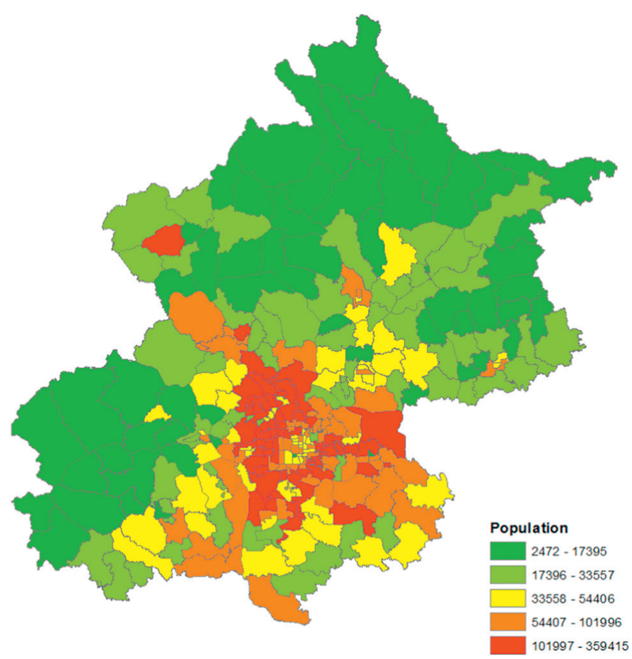

Fig. 2 – Census data in the area of Beijing.

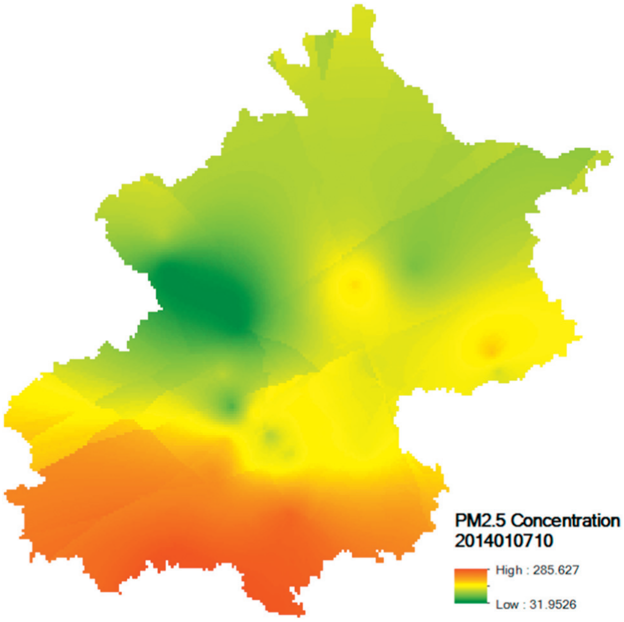

Fig. 3 – A sample of interpolation results of $PM_{2.5}$ concentrations in Beijing.

and finally, estimating the population exposure to $PM_{2.5}$ air pollution.

2.1. *The spatialization of $PM_{2.5}$ concentrations*

Because the $PM_{2.5}$ monitoring stations are distributed sparsely, it is difficult to estimate the $PM_{2.5}$ concentrations using the ground-based monitoring results directly. There are different methods developed to spatialize the $PM_{2.5}$ concentrations over certain areas, such as interpolation methods, regression models, dynamic models, and so on.

Among the interpolation methods, the Kriging interpolation method proposed by Kriging in 1951 is widely used, which is also provided in the toolbox of ArcGIS for Desktop. This paper adopted the Kriging interpolation method to interpolate the monitoring data for the whole area of Beijing with the spatial resolution of 1 km × 1 km. A sample of interpolation results for $PM_{2.5}$ in Beijing at 10:00 on January 7, 2014 is shown in Fig. 3. There were 5645 hourly monitoring data utilized in this paper, and the same interpolation method was applied for these monitoring data. In order to avoid duplication of work, this paper used the Environmental Systems Research Institute (ESRI) model builder to finish the interpolation process.

2.2. *The spatialization of population census data*

Another challenge in this paper was to spatialize the population census data with high spatial resolution. If there is no population, there is no exposure (Hao et al., 2012). Thus, population is an important indicator for measuring the PWEL. However, the population census data of each sub-district cannot reflect the real internal distribution of the population. Various methods have been proposed (Amaral et al., 2006; Ye et al., 2010) to improve the analytical precision of population-related problems. A method using the OSM data and POI data was proposed and verified by Long (Long and Liu, 2013), which can improve the spatial resolution of population data. This paper obtained the road data from the OSM and POI data from the public data part provided by GeoHey.

Four steps are necessary to spatialize the population census data:

1. Extract parcels using road data and sub-district boundary data. Each parcel is surrounded by streets, that is, the polygons within roads should be identified. This paper adopted the tool "polyline to polygon" provided by ArcGIS for Desktop to extract parcels surrounded by streets in each sub-district. Special attention should be paid in this step because some extracted areas are too small when there are overpasses or dual carriageways. After the extraction work was finished, we also removed any areas less than 10 m^2.
2. Calculate the density of POIs in each parcel. POI data contains sites of people's daily activities, such as canteens, banks, shopping malls, living quarters, hospitals, schools, and so on. The distribution of POIs can describe the population density in an area to a certain degree. This paper counted the number of POIs in each parcel, and then calculated the POI density which is defined as d in each parcel.
3. Standardize the parcel according the density of POIs in each sub-district. Assuming there are k sub-districts, and m_k parcels in sub-district k, then the standardization of each parcel in sub-district k is shown as Eq. (1):

$$S_{i,j} = \frac{d_{i,j}}{\sum_{i=1}^{m_k} d_{i,j}} \quad (i = 1, 2, 3 \ ......m_k, j = 1, 2, 3 \ ......k) \tag{1}$$

where, $S_{i,j}$ means the standard value of the ith parcel in jth sub-district, and $d_{i,j}$ means the density of POIs in the ith parcel in jth sub-district.

4. Calculate the population in each parcel. We assumed that the population distribution is similar to the POI distribution in every parcel, and the standardization of population in each parcel is defined as $S_{i,j}$. The population in each parcel is shown as Eq. (2):

$$P_{i,j} = S_{i,j} \cdot P_{i,j} \quad (i = 1, 2, 3 \ ......m_k, j = 1, 2, 3 \ ......k) \tag{2}$$

where, $P_{i,j}$ (person) means the population in the ith parcel in jth sub-district.

2.3. *Estimate the population exposure to $PM_{2.5}$*

In fact, the distribution of population over an area is inconsistent with the distribution of $PM_{2.5}$ concentration (Ivy et al., 2008). Thus, the PWEL was proposed as an important indicator to measure the population exposure to $PM_{2.5}$ air pollution, which requires the population distribution data and $PM_{2.5}$ concentration distribution data. The equation for calculating total PWEL over an area is shown as follows:

$$PWEL = \frac{\sum(P_i \times C_i)}{\sum P_i} \tag{3}$$

where, the parameter i in Eq. (3) represents the index of the computing grid, P_i represents the population in computing grid i, and C_i represents the $PM_{2.5}$ concentrations in the same grid.

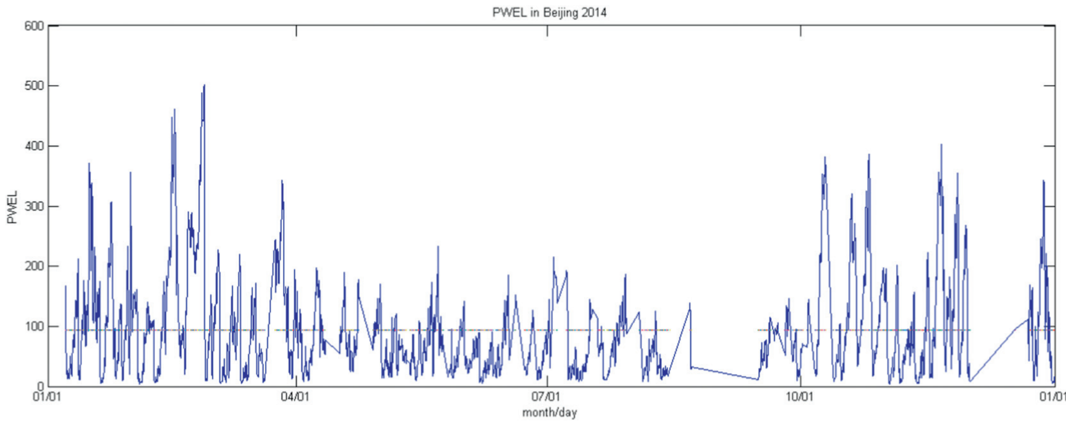

Fig. 4 – The hourly variation of the total population-weighted exposure level (PWEL) during 2014.

In each grid, the PWEL equals the value of $PM_{2.5}$ concentration when the population is larger than zero in the same grid, which is:

$$PWEL_i = \begin{cases} C_i & (P_i > 0) \\ 0 & (P_i = 0) \end{cases} \tag{4}$$

Short-term or long-term exposure to $PM_{2.5}$ concentrations are strongly linked with human health. This paper used 35 $\mu g/m^3$, the 24-hr average standard proposed by WHO (World Health Organization, 2006), as an indicator to calculate the number of exposure days exceeding the standard.

$$D_i = \sum 1 \qquad (PWEL_i > S) \tag{5}$$

The equation shown above calculates the cumulative exposure days during the whole research period, which can indicate the exposure intensity in different parts of the area. Meanwhile, this paper also calculates the cumulative $PM_{2.5}$ concentrations during the whole research period, that is:

$$E_i = \sum PWEL_i \tag{6}$$

This paper uses the cumulative exposure days and cumulative $PM_{2.5}$ concentrations to indicate the level of population exposure to $PM_{2.5}$ ambient concentrations.

3. Spatial and temporal variations of $PM_{2.5}$ concentrations

The spatial distribution of $PM_{2.5}$ concentrations and population with spatial resolution of 1 km × 1 km was produced in this study. Combining the $PM_{2.5}$ concentrations and population spatialization results, the total population exposure to $PM_{2.5}$ concentrations was then calculated, which is shown in Figs. 4 and 5.

Fig. 4 shows that the total population exposure to $PM_{2.5}$ ambient concentrations varies greatly during the whole of 2014. The mean value of total population exposure level in 2014 reached 90 $\mu g/m^3$, far exceeding the standard concentrations provided by WHO. Temporally, the population exposure to $PM_{2.5}$ in winter and spring is more serious than in summer and autumn. Even in winter or autumn, the degree of fluctuation of PWEL in these seasons is also larger than in summer and autumn. The monthly variation of population exposure to $PM_{2.5}$ air pollution was also measured. Fig. 5 shows the monthly variation during 2014. The PWEL in February was much higher than other months, the average of which reached 146 $\mu g/m^3$, and the maximum value in this month also exceeded 250 $\mu g/m^3$. The PWEL in October was also quite serious, with the mean value reaching 136 $\mu g/m^3$. It is worth mentioning that the population exposure to $PM_{2.5}$ in August was lowest compared to other months, with an average PWEL of 44 $\mu g/m^3$, and maximum value of 71 $\mu g/m^3$.

Spatially, the cumulative exposure days exceeding the standard value provided by WHO and average value in 2014 were also calculated. The former results are shown in Fig. 6a; the average number of exposure hours is about 3586.2 hr, equal to 149 days. From Fig. 6a, we can also find that the exposure time decreased from South to North in Beijing, while the exposure hours in the north reached 4400 hr, equal to 188 days. Fig. 4 shows the spatial distribution of exposure days exceeding the mean value of 90 $\mu g/m^3$, with the average exposure time of 1888 hr, or 78.7 days. The exposures time in the area of the Southwest is much more serious than in the Northeast.

Over the duration of data collection, this paper also calculated the cumulative population exposure to $PM_{2.5}$ concentrations, as shown in Fig. 7. The distribution of

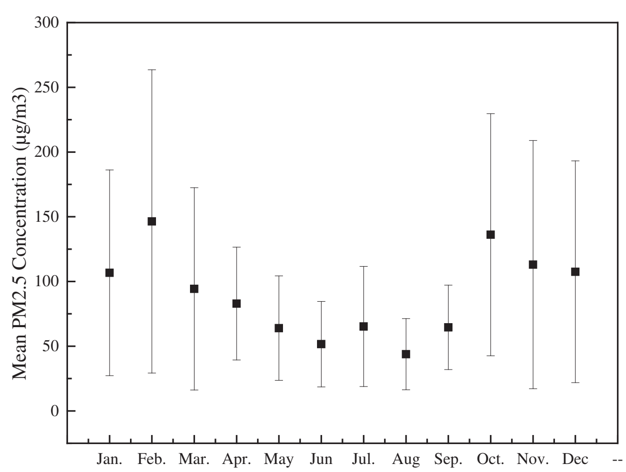

Fig. 5 – The monthly variation of total PWEL during 2014.

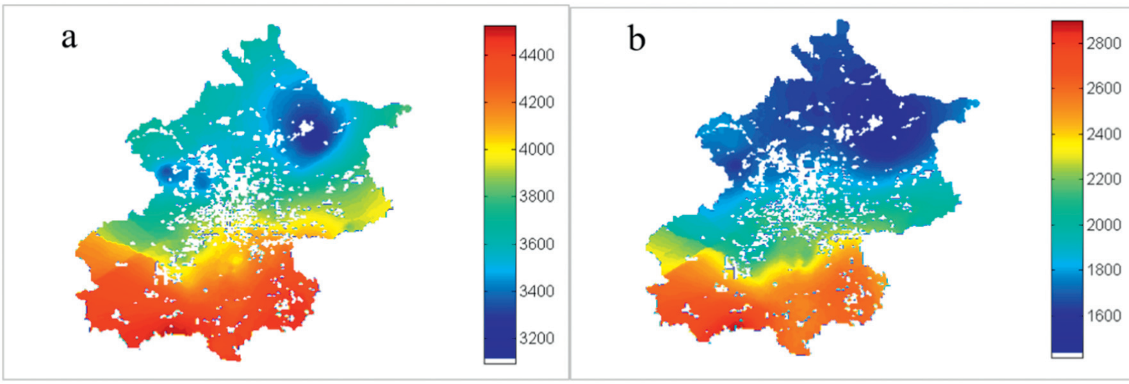

Fig. 6 – (a) Accumulation exposure hours during 2014 with S = 35 μg/m^3 and (b) accumulation exposure hours during 2014 with S = 90 μg/m^3 (Unit: hour).

cumulative population exposure to $PM_{2.5}$ is similar to the distribution of cumulative exposure time.

4. Conclusions

Geographic information system (GIS) is an important and powerful tool to estimate and visualize different natural or artificial phenomena. The GIS-based spatial and temporal variations of population exposure to $PM_{2.5}$ were measured and verified in this study. In order to calculate the population exposure to $PM_{2.5}$, the spatialization of $PM_{2.5}$ concentrations and population with the spatial resolution of 1 km × 1 km was carried out. The total population exposure to $PM_{2.5}$ and cumulative exposure time were then calculated with the spatialization results. Due to the climate characteristics in Beijing, and the heating that generally takes place in November to March next year, the air condition in winter and spring is always more serious than in summer and autumn. The temporal variation of PWEL in 2014 also shows that the air pollution problems in winter and spring are much more serious than in summer and in autumn. Specifically, the PWEL in February was most serious in 2014, even reaching 146 μg/m^3, three times higher than the PWEL in August. Spatially, the severity of $PM_{2.5}$ air pollution decreases from North to South. Considering the standard provided by WHO, there were about 149 days when the average cumulative exposure exceeded the standard, while the figure was 78.7 days considering the mean value. There are also some limitations to our research. Firstly, people do not carry out activities in all places at the same time, so the hypothesis that the population distribution is similar to the POIs distribution is not realistic. Another limitation is that the spatializing of results for the census data in urban areas is more precise than in suburban areas because the streets in urban areas are more intensive. Further work will identify the most common POIs at a certain time, and collect more detailed street data in suburban areas to estimate the population exposure to $PM_{2.5}$ concentrations.

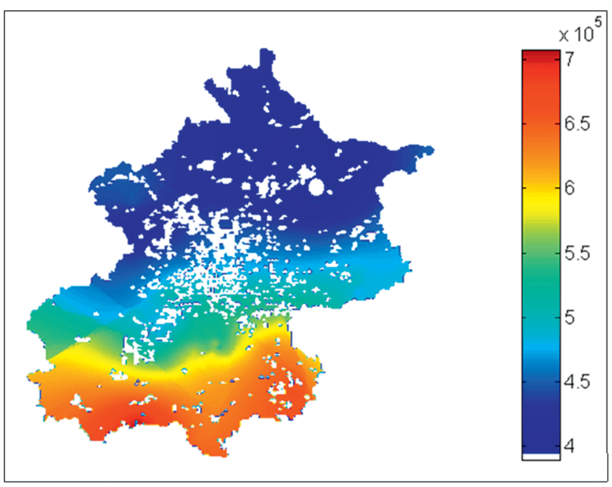

Fig. 7 – Accumulation population exposure to $PM_{2.5}$ concentrations (Units: μg/m^3).

REFERENCES

Amaral, S., Monteiro, A.M., Câmara, G., Quintanilha, J.A., 2006. DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon. Int. J. Remote Sens. 27 (05), 855–870.
Che, H., Zhang, X., Li, Y., Zhou, Z., Qu, J.J., Hao, X., 2009. Haze trends over the capital cities of 31 provinces in China, 1981–2005. Theor. Appl. Climatol. 97 (3–4), 235–242.
Gamble, J.F., 1998. PM2.5 and mortality in long-term prospective cohort studies: cause–effect or statistical associations? Environ. Health Perspect. 106 (9), 535.
Goodchild, M.F., Li, L., 2012. Assuring the quality of volunteered geographic information. Spat. Stat. 1, 110–120.
Guo, T., Ma, Y., He, K., 2009. Study on spatial distributions of PM2.5/PM10 in regional atmospheric environment. Chinese J. Environ. Eng. (01), 147–150.
Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environ. Plann. B. Plann. Des. 37 (4), 682–703.
Hao, Y., Flowers, H., Monti, M.M., Qualters, J.R., 2012. US census unit population exposures to ambient air pollutants. Int. J. Health Geogr. 11 (1), 3.
Ivy, D., Mulholland, J.A., Russell, A.G., 2008. Development of ambient air quality population-weighted metrics for use in time-series health studies. J. Air Waste Manage. Assoc. 58 (5), 711–720.

Li, R., Li, Z., Gao, W., Ding, W., Xu, Q., Song, X., 2015. Diurnal, seasonal, and spatial variation of PM2.5 in Beijing. Sci. Bull. 60 (3), 387–395.

Long, Y., Liu, X., 2013. Automated Identification and Characterization of Parcels (AICP) with OpenStreetMap and Points of Interest (arXiv preprint arXiv:1311.6165).

National Bureau of Statistics of China, 2002. Tabulation on the 2000 Population Census of the People's Republic of China.

Sun, Z., Mu, Y., Liu, Y., Shao, L., 2013. A comparison study on airborne particles during haze days and non-haze days in Beijing. Sci. Total Environ. 456, 1–8.

Vallejo, M., Ruiz, S., Hermosillo, A.G., Borja-Aburto, V.H., Cárdenas, M., 2006. Ambient fine particles modify heart rate variability in young healthy adults. J. Expo. Sci. Environ. Epidemiol. 16 (2), 125–130.

Wang, Y., Zhuang, G., Sun, Y., An, Z., 2006. The variation of characteristics and formation mechanisms of aerosols in dust, haze, and clear days in Beijing. Atmos. Environ. 40 (34), 6579–6591.

World Health Organization, 2006. Air Quality Guidelines: Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide. World Health Organization.

Ye, J., Yang, X., Jiang, D., 2010. The grid scale effect analysis on town leveled population statistical data spatialization. J. Geo-Inf. Sci. 12 (1), 40–47.

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 (5), e63486.