



Contents lists available at ScienceDirect

Environmental Pollution

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

Dynamic assessment of PM_{2.5} exposure and health risk using remote sensing and geo-spatial big data[☆]



Yimeng Song^a, Bo Huang^{a,*}, Qingqing He^b, Bin Chen^c, Jing Wei^d, Rashed Mahmood^e

^a Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, Hong Kong

^b School of Resource and Environmental Engineering, Wuhan University of Technology, Wuhan, 430070, China

^c Department of Land, Air and Water Resources, University of California, Davis, CA, 95616, USA

^d State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

^e Department of Atmospheric Science, School of Environmental Studies, China University of Geosciences, Wuhan, Hubei, China

ARTICLE INFO

Article history:

Received 22 January 2019

Received in revised form

12 June 2019

Accepted 13 June 2019

Available online 5 July 2019

Keywords:

Human mobility

Spatiotemporal heterogeneity

Remote sensing

Big data

Environmental health

ABSTRACT

In the past few decades, extensive epidemiological studies have focused on exploring the adverse effects of PM_{2.5} (particulate matters with aerodynamic diameters less than 2.5 μm) on public health. However, most of them failed to consider the dynamic changes of population distribution adequately and were limited by the accuracy of PM_{2.5} estimations. Therefore, in this study, location-based service (LBS) data from social media and satellite-derived high-quality PM_{2.5} concentrations were collected to perform highly spatiotemporal exposure assessments for thirteen cities in the Beijing-Tianjin-Hebei (BTH) region, China. The city-scale exposure levels and the corresponding health outcomes were first estimated. Then the uncertainties in exposure risk assessments were quantified based on in-situ PM_{2.5} observations and static population data. The results showed that approximately half of the population living in the BTH region were exposed to monthly mean PM_{2.5} concentration greater than 80 μg/m³ in 2015, and the highest risk was observed in December. In terms of all-cause, cardiovascular, and respiratory disease, the premature deaths attributed to PM_{2.5} were estimated to be 138,150, 80,945, and 18,752, respectively. A comparative analysis between five different exposure models further illustrated that the dynamic population distribution and accurate PM_{2.5} estimations showed great influence on environmental exposure and health assessments and need be carefully considered. Otherwise, the results would be considerably over- or under-estimated.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

In the past few decades, PM_{2.5} (fine particulate matter with aerodynamic diameters of less than 2.5 μm) has become a serious health concern (Cohen et al., 2004; Anenberg et al., 2010; Shang et al., 2013). The fine particulate matter, composed of various chemical compounds and aerosols (e.g., combustion particles, organic compounds, and metals), can penetrate human bronchi and lungs, causing great damage to human health (Pope and Dockery, 2006). Nowadays, numerous epidemiologic studies have explored and established associations between PM_{2.5} exposure and premature mortalities, such as cardiopulmonary death and respiratory death (Dockery et al., 1993; Lelieveld et al., 2015; Nel, 2005). Such

findings are especially concerning in China. Due to unprecedented economic development and urbanization in recent decades, China has experienced a tremendous increase in energy consumption and is highly polluted with PM_{2.5} (Kan et al., 2012; Sun et al., 2016). In 2013, a previous study reported that more than 96% of the Chinese population were exposed to PM_{2.5} concentrations exceeding the World Health Organization (WHO) Air Quality PM_{2.5} Interim Target-1 (IT-1, annual mean of 35 μg/m³) level (Ma et al., 2014). In 2014, the annual mean PM_{2.5} concentration was about 61 μg/m³ (Fang et al., 2016) in China, which exceeded the WHO's recommended air quality standard (10 μg/m³/year) by six times. Recent estimations suggest that PM_{2.5} pollution is associated with 1.2 million premature deaths per year from 1999 to 2010 (Peng et al., 2016). Therefore, it is critically important to accurately estimate the population exposure to PM_{2.5} in order to implement countermeasures to reduce air pollution.

The accurate estimation of PM_{2.5} concentrations is one of the

[☆] This paper has been recommended for acceptance by Admir Créso Targino.

* Corresponding author.

E-mail address: bohuang@cuhk.edu.hk (B. Huang).

most critical prerequisites in PM_{2.5}-related environmental pollution and epidemiologic studies (Chen et al., 2018a). Station-based and satellite-station-hybrid models are two widely used approaches in estimating the surface PM_{2.5} concentrations. Station-based models use the station or monitor-based records and interpolation methods to obtain the pollution levels with high update frequencies (Nyhan et al., 2016; Park and Kwan, 2017). However, the sparsely distributed surface monitor stations cannot fully capture the spatial variability of PM_{2.5} concentrations, thus leading to relatively low accuracy. Compared to station-based models, satellite-station-hybrid models can more effectively estimate the ground-level PM_{2.5} concentrations at high spatial resolutions by integrating satellite remote sensing data and in-situ monitoring records. Satellite-derived aerosol optical depth (AOD) product is an important optical property and has been proved to be highly positively correlated with PM_{2.5} (He and Huang, 2018a). Satellite-station-hybrid models can effectively present the spatial heterogeneity of air pollution (Lv et al., 2017), and a large number of statistical regression models have been developed to quantify the PM_{2.5}-AOD relationships and improve the accuracy of ground-level PM_{2.5} estimations. These models mainly include the simple linear regression model (Wang and Christopher, 2003), the linear mixed effect (LME) model (Ma et al., 2015; Xie et al., 2015), the geographically weighted regression (GWR) model (Song et al., 2014), and the geographically and temporally weighted regression (GTWR) model (He and Huang, 2018a, b). However, most of the current PM_{2.5} exposure studies are based on station-based models rather than satellite-station-hybrid models, and the uncertainties caused by the above-mentioned issue have not been comprehensively explored, especially the comparison with satellite-station-hybrid models.

Estimating population exposure to ambient PM_{2.5} is challenging. Demographic data based on administration cells (census-tract) have been widely used to provide reliable population information (Fleischer et al., 2014; Gray et al., 2014). However, these spatially-aggregated census data ignore the spatial heterogeneity of population distribution and have a low update frequency (5–10 years), leading to low assessment accuracy. Gridded population data, such as LandScan Global Population (Dobson et al., 2000), can provide detailed spatial distributions of population and have been used in numerous exposure assessment studies (Lelieveld et al., 2015; Van Donkelaar et al., 2014). Nevertheless, population census data are also the source data for these pixel-based population distribution maps, significantly limiting their spatiotemporal accuracy (Chen et al., 2018b). Approaches like questionnaire surveys or on-site investigations may remedy above limitations by offering more exact and timely location information (Pope et al., 2002), but the limited samples (i.e., interviewees or participants) prevent generalizing the results beyond a small area. More importantly, given the dynamic changes in population distribution, these aforementioned static data are less likely to capture the ever-changing interactions between humans and air pollution.

To address this issue, researchers have collected and involved human mobility information to improve the accuracy of exposure estimations. For example, by using GPS-based wearable air monitors, researchers can collect the real-time locations (GPS trajectories) of participants and the ambient air pollution levels (Wang et al., 2018). However, privacy issues and the amount of manpower required to collect such information create roadblocks to large-scale, long-term, and continuous studies. Recently, the rapid development and popularization of mobile internet technology, diverse computing platforms (e.g., smartphones, tablets, and PCs), and location-based services have promoted the generation of location-based service (LBS) data to record people's spatiotemporal activities (Liu et al., 2015). Compared to

demographic data or GPS-based records, LBS data are representative indicators with a much better spatiotemporal scale and sufficient sample size (Hawelka et al., 2014). Despite increasing studies involving the LBS data (Chen et al., 2018a; Nyhan et al., 2016; Song et al., 2018), the contribution of integrating LBS data and satellite-station-based PM_{2.5} concentrations in assessing PM_{2.5} exposure and health impacts remains unclear.

Therefore, this objective of this study is to improve the model for PM_{2.5} exposure and health risk assessments and addresses the uncertainties caused by station-based PM_{2.5} concentration and static population data. For this, a new exposure assessment model is first developed using dynamic population distribution data obtained from social media and PM_{2.5} concentrations derived from an improved satellite-station-hybrid model. Then the PM_{2.5} exposure levels and the number of premature deaths caused by PM_{2.5} are estimated in 2015 for thirteen cities in the Beijing-Tianjin-Hebei (BTH) region, China. Lastly, the comparative analysis between five different exposure models are conducted to evaluate the importance of time-resolved population distribution and improved PM_{2.5} concentration maps in environmental exposure/health-related studies.

2. Materials and methods

2.1. Study area

The PM_{2.5} exposure levels and the number of premature deaths in the BTH metropolitan region were assessed for the year 2015. As the capital region of China, the BTH region is the biggest urbanized and economically developed region in northern China. The region consists of thirteen cities (Fig. 1): two municipalities, Beijing (BJ) and Tianjin (TJ); one provincial capital city, Shijiazhuang (SJ); and 10 prefecture-level cities including Tangshan (TS), Qinhuangdao (QH), Handan (HD), Xingtai (XT), Baoding (BD), Zhangjiakou (ZJ), Chengde (CD), Cangzhou (CZ), Langfang (LF), and Hengshui (HS). In 2015, the total population of the BTH region was 111 million (China, 2016). Heavy air pollution episodes frequently occur in this densely populated region. The annual mean PM_{2.5} concentration exceeded 90 µg/m³ in 2014 (He and Huang, 2018a), which was about nine times higher than the WHO's recommended air quality standard.

2.2. Location-based service data

In this study, the LBS data (i.e., geo-tagged messages) were obtained from Weibo (<http://weibo.com>) to reflect the dynamic change in population distribution. Weibo is one of the most popular social media platforms and microblogging services in China. In September 2015, Weibo had 222 million monthly active users (MAUs), and mobile MAUs represented 83% of the total MAUs. The ratio of male users to female users was 1:1, and young people (aged 17–35 years old) comprised the largest segment, accounting for 79% of active users (Weibo-Corporation, 2015). Similar to Twitter (<http://twitter.com>), Weibo users can post messages (e.g., short texts, images, voice, video, etc.) and share their real-time locations via this online social network. Using a series of probability models and indicators of mobility patterns, Jurdak et al. (2015) conducted a comprehensive analysis and demonstrated that the geo-tagged data from social media could capture several features of human mobility and effectively represent real-time population distribution at the metropolitan scale. Jurdak et al. (2015) also concluded that the three inevitable issues (i.e., potential sampling bias, location bias, and communication modality) in using geo-tagged social media data (e.g., Twitter) did not strongly influence the performance of LBS data in characterizing dynamic population distributions at city scale. Moreover, LBS data from social media have been

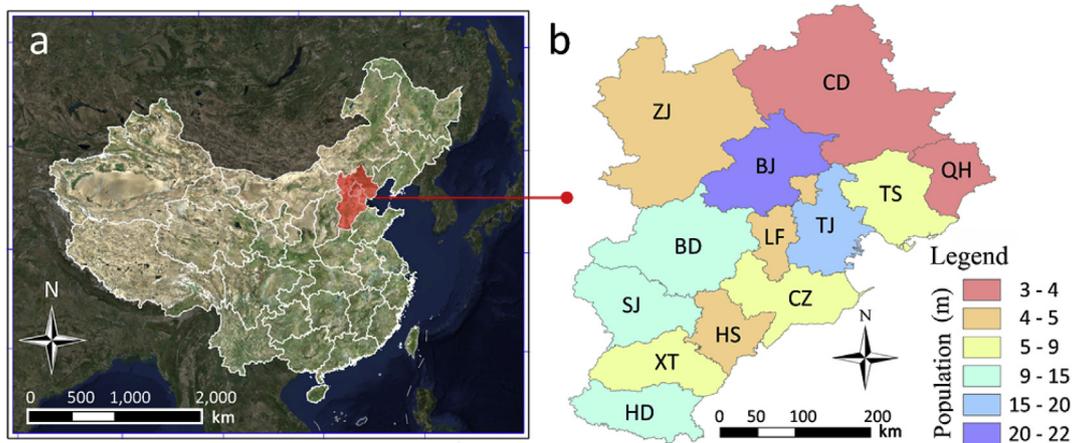


Fig. 1. (a) Location and (b) population of the Beijing-Tianjin-Hebei region in China.

successfully used in pioneering studies on human mobility (Huang and Li, 2016; Luo et al., 2016), travel behavior (Rashidi et al., 2017; Zhang et al., 2017), environmental exposure (Chen et al., 2018b; Zheng et al., 2019), land-use classification (Chen et al., 2017b; Liu et al., 2017), and urban planning (Cai et al., 2017; Zhang and Zhou, 2018). These studies illustrated that LBS data from social media could be a useful proxy for describing dynamic population distribution.

This study collected 7.8 million geo-tagged records through the Weibo application programming interface (API) (<http://open.weibo.com>) from 1 January 2015 to 31 December 2015. All information about Weibo users' identities and other private information was deleted before the data were released. Information obtained from geo-tagged records (i.e., LBS data) included user ID (anonymous), gender, message, release time, and location. Only the "release time" and "location" data were used to generate the dynamic population distribution maps.

2.3. Demographic data

The demographic data in the BTH region for the year 2015 was obtained from the National Scientific Data Sharing Platform for Population and Health (<http://www.ncmi.cn>). This dataset has recorded annual resident population information at the county level since 2004 and is maintained by the Chinese Infectious Disease Network Reporting System.

2.4. PM_{2.5} concentrations

The ground-level PM_{2.5} concentrations used in this study were estimated using our recently developed satellite-ground-hybrid model (He and Huang, 2018a). The model includes two spatio-temporal models, i.e., the GTWR model and the improved GTWR model by incorporating seasonal characteristics. The GTWR model has been widely applied in environmental studies (Bai et al., 2016; Chu et al., 2015; Guo et al., 2017) by carefully considering the spatial and temporal variabilities. The inputs variables mainly included the Moderate Resolution Imaging Spectroradiometer (MODIS) 3-km AOD product, ground-level PM_{2.5} measurements, and ancillary variables (i.e., the relative humidity, temperature, wind speed, and normalized difference vegetation index). The 3-km AOD product was fused based on the newly released 3-km Dark Target and the 10-km Deep Blue AOD products to improve the spatial coverage and data quality compared to the official AOD products (Wei et al., 2019). Then the model was employed to generate the daily 3-km

PM_{2.5} concentration dataset. The PM_{2.5} estimations were evaluated against the ground-level PM_{2.5} measurements using the 10-fold cross-validation method, and the results showed that GTWR models can capture more than 80% of the daily PM_{2.5} variations in the BTH region (He and Huang, 2018a).

2.5. Mortality database

The total mortality and the incidences of mortality data sets for different health endpoints were collected for the health impact assessments. The total mortality data were collected from the Statistical Yearbook (2015) published by the governments of the thirteen cities. The incidences of mortality for different health endpoints (i.e., cardiovascular mortality and respiratory mortality) were collected from the Statistical Yearbook (2015) and National Disease Surveillance Points System (2015).

2.6. Quantifying dynamic population distribution

The LBS data were used as indicators to quantify the spatio-temporal pattern of population distribution. Due to the differences in socioeconomic development and mobile internet popularity among different cities, estimates for dynamic population distribution were conducted separately for each city. The monthly density maps of LBS data were generated and used in this study by aggregating all geo-tagged records for each grid because the Weibo geo-tagged messages are always sparse and insufficient during a relatively short period (Cai et al., 2017; Liu et al., 2015; Steiger et al., 2015). The monthly density maps can more comprehensively quantify the population distribution.

Then, the monthly density maps of LBS data were used to redistribute the demographic data for each city (Eqs. (1) and (2)), assuming that the inter-city human mobility did not dramatically affect the total population of a city within the monthly time window. That is, population changes caused by human movement across different cities were neglected in this study.

$$W_{ij} = \frac{p_{ij}}{\sum_{i=1}^n p_{ij}} \quad (1)$$

$$Pop_{ij} = TP \times W_{ij} \quad (2)$$

where p_{ij} is the number of Weibo geo-tagged records at the i th pixel in a given month j , n is the total number of the pixels in a city, W_{ij} is the weight for redistributing population, TP is the total population

in a city, and Pop_{ij} denotes the population approximation in the i th pixel in a given month j .

2.7. $PM_{2.5}$ exposure assessment

Due to the dynamic changes in $PM_{2.5}$ concentrations and population distribution, a population-weighted metric was adopted to estimate the dynamic population exposure to $PM_{2.5}$. The daily $PM_{2.5}$ concentrations were first averaged into monthly maps to make the temporal resolution consistent with the LBS-based population maps. $PM_{2.5}$ exposure levels were obtained by integrating population distribution and $PM_{2.5}$ concentrations. For each city, the monthly exposures were assessed via a pixel-based method (Eq. (3)), which can effectively reduce the potential zoning effect of the modifiable areal unit problem (MAUP) (Ho et al., 2015; Ho et al., 2018). Finally, the annual mean exposure levels for different cities can be calculated using Eq. (4).

$$Et_j = \frac{Pop_{ij} \times PM_{ij}}{\sum_{i=1}^n Pop_{ij}} \quad (3)$$

$$Ey = \sum_{j=1}^{12} Et_j / 12, (j = 1, 2, \dots, 12) \quad (4)$$

where PM_{ij} and Pop_{ij} are the average $PM_{2.5}$ concentration and the estimated population in i th pixel in a given month j , respectively; n is the total number of pixels in a city; Et_j is the monthly mean population-weighted exposure in a city, and Ey is the annual mean population-weighted exposure of $PM_{2.5}$ in a city.

2.8. Health impact assessment

The premature deaths associated with $PM_{2.5}$ exposure were estimated using the concentration-response (C-R) function that relates changes in fine particle concentrations to changes in mortality. The C-R function was developed based on a log-linear relationship between relative risk and concentrations (Anenberg et al., 2010). Detailed information regarding the C-R function can be found in Section 1 of the Supplementary Materials.

In this study, we estimated premature deaths for three types of health endpoints (i.e., all-cause mortality, cardiovascular mortality, and respiratory mortality) under the assumption that the entire population was exposed to a level same as their ambient outdoor $PM_{2.5}$ concentration. Notably, the relative risk (RR) from $PM_{2.5}$ concentration varied significantly at the city level (Lu et al., 2015; Shang et al., 2013). However, because large amounts of clinical or mortality data are unavailable, the city-level RR values regarding short-term $PM_{2.5}$ exposure were only estimated for less than 25 cities in the Greater China Region (including mainland China,

Taiwan, Hong Kong, and Macao), according to a systematic review based on both English- and Chinese-language literature databases from 1990 to 2013 (Lu et al., 2015). Therefore, the meta-analysis-based RR derived from a large number of epidemiological studies has become an acceptable solution for this issue for regional-scale studies. The short-term RR used in this study was derived from a meta-analysis of nine related studies concerning Chinese cities (Shang et al., 2013). It has also been widely used in regional studies concerning the adverse health effects caused by $PM_{2.5}$ exposure in parts of China, including the BTH region (Yang et al., 2019; Zhao et al., 2019), the Pearl River Delta region (Lu et al., 2016), the Yangtze River Delta region (Liao et al., 2017; Wang et al., 2015), the Henan Province (Shen et al., 2017), and across mainland China (Chen et al., 2017a). Specifically, the RR suggested that an increase of $10 \mu\text{g}/\text{m}^3$ in the annual mean $PM_{2.5}$ would cause 0.38% (95% CI: 0.31%, 0.45%), 0.44% (95% CI: 0.33%, 0.55%), and 0.51% (95% CI: 0.30%, 0.75%) increase in all-cause mortality, cardiovascular mortality, and respiratory mortality, respectively. In addition, the baseline incidence rates (y_0) of different health endpoints and the exposed population (Pop) in 2015 were collected. The baseline of $PM_{2.5}$ concentration (C_0) was set as the WHO's recommended air quality level (i.e., $10 \mu\text{g}/\text{m}^3/\text{year}$). The actual $PM_{2.5}$ concentration (C) was the calculated annual mean population-weighted $PM_{2.5}$ exposure level for each city.

2.9. Comparison of five exposure models

To better understand the uncertainties caused by station-based $PM_{2.5}$ concentration and static population data, we compared five exposure models based on different datasets (Table 1). Station-based $PM_{2.5}$ concentrations and pixel-based population data were used in the first four models. Station-based $PM_{2.5}$ concentrations were interpolated into continuous surface data at 3-km spatial resolution using the inverse distance-weighted technique and then averaged to monthly maps (He and Huang, 2018a). The 1-km LandScan Global Population data (2015) were downloaded from the Oak Ridge National Laboratory (<http://web.ornl.gov/sci/landscan>) and were aggregated into 3-km cell grids. The exposure and health impact assessments of these models were conducted using the above-introduced methods, respectively. For Model-1/3, the county-level demographic data were equally distributed into each 3-km pixel.

3. Results

3.1. Monthly population distribution

The monthly population distributions in the BTH region were estimated by integrating density maps of LBS data and city-scale

Table 1
Exposure assessment models used for comparison.

Model	Datasets		Notes
	Population	$PM_{2.5}$	
Model-1	County-level demographic data	Station-based interpolated $PM_{2.5}$ concentration	(1) Low spatial accuracy in terms of $PM_{2.5}$ concentration; (2) Lower spatiotemporal variability in terms of population distribution.
Model-2	Pixel-based Landscan population data	Station-based interpolated $PM_{2.5}$ concentration	(1) Low spatial accuracy in terms of $PM_{2.5}$ concentration; (2) Lower temporal variability in terms of population distribution.
Model-3	County-level demographic data	Satellite-station-based $PM_{2.5}$ concentration	(1) Lower spatiotemporal variability in terms of population distribution.
Model-4	Pixel-based Landscan population data	Satellite-station-based $PM_{2.5}$ concentration	(1) Lower temporal variability in terms of population distribution.
Model-5	LBS-based dynamic population data	Satellite-station-based $PM_{2.5}$ concentration	(1) Improved spatial accuracy in terms of $PM_{2.5}$; (2) Considering the spatiotemporal variability in terms of population distribution

demographic data. Fig. 2a shows the result in November 2015, in which the intensity represents the specific amount of the population within the area covered by each pixel, with stretched colors from yellow to dark blue denoting the varied population densities. The results show that the pixel-based population map could more appropriately characterize the spatial distribution of people. In the BTH region, a vast majority of the population is centered in the urban core of each city, and the remaining proportion is spread over the entire region with several concentrated hot spots. The zoomed-in Beijing map (Fig. 2b) suggests that the LBS-based population map could provide a spatially explicit visualization of population density, especially for urban cores. The correlation coefficient matrix (Fig. 2c) of the monthly population maps in Beijing further indicated the dynamics of population distributions over space and time. These monthly dynamic changes could be explained by several factors that directly or indirectly affect the population distribution and transport, such as weather conditions, commuting

routines, and special holidays.

3.2. Ground-level PM_{2.5} concentrations

Fig. 3a–d shows the monthly mean PM_{2.5} concentration maps in January, April, September, and October, which were estimated using the GTWR satellite-station-hybrid model. Overall, the PM_{2.5} concentration level in the southeast plain was higher than that of the northwest mountainous region. Besides, the spatial patterns of PM_{2.5} concentrations changed significantly at the month scale in the year. In winter (e.g., January, Fig. 3a), for instance, increased coal burning for heat and unfavorable meteorological conditions jointly contributed to the most severe PM_{2.5} pollution in the BTH region (Zhang et al., 2009). In spring (e.g., in April, Fig. 3b), the high PM_{2.5} concentrations in the plain were mainly caused by sand dust raised by strong winds (Lv et al., 2017). In late summer and early autumn (i.e., September, Fig. 3c) the PM_{2.5} pollution is slightest, but large-

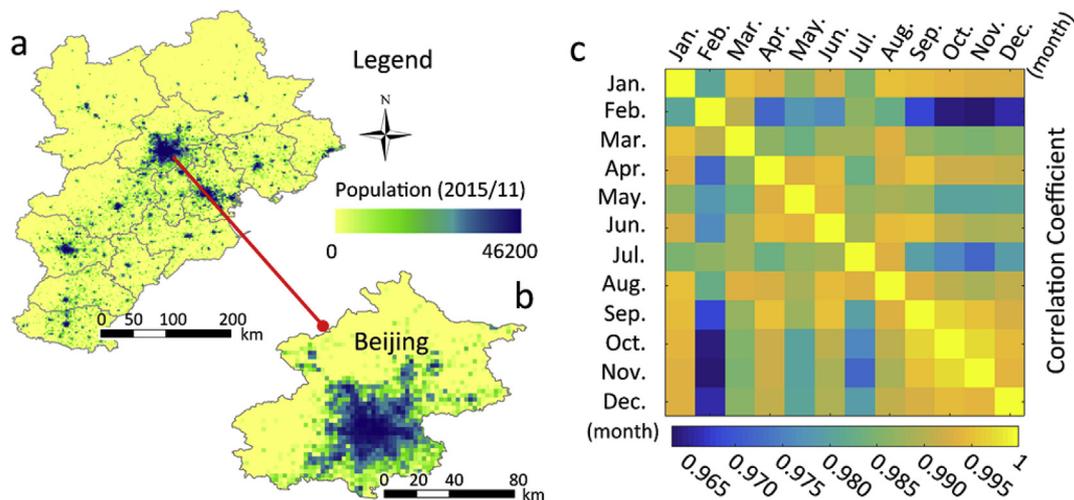


Fig. 2. (a–b) Maps of LBS-based population distribution (2015/11) in the BTH region and Beijing, respectively. (c) Correlation coefficient matrix of the monthly LBS-based population distribution in Beijing in 2015.

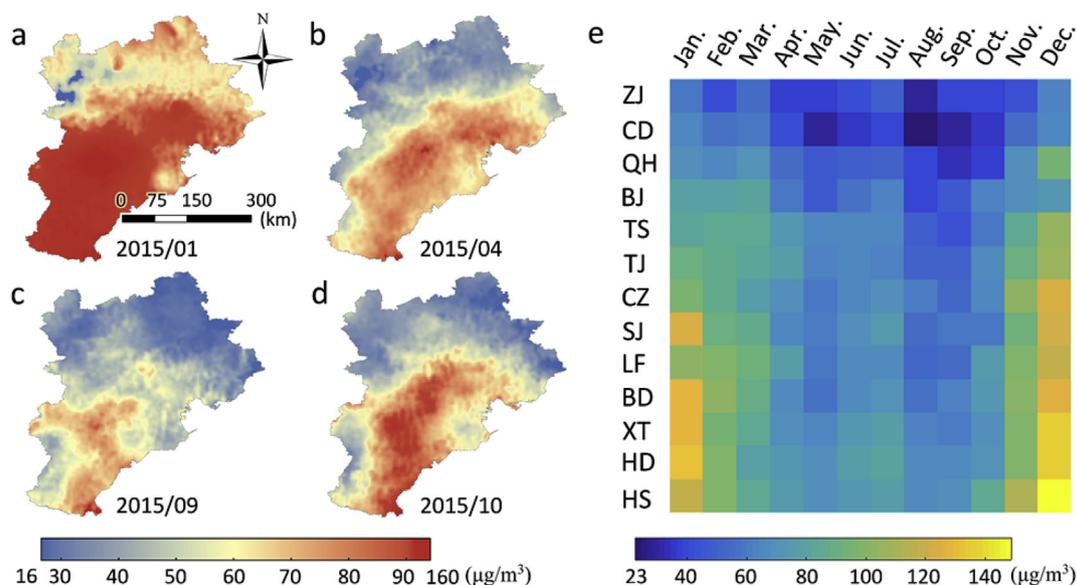


Fig. 3. (a–d) Monthly mean PM_{2.5} concentration maps in January, April, September, and October 2015, respectively. (e) City-scale monthly mean PM_{2.5} concentrations from 2015/01 to 2015/12. The x-axis represents January to December 2015, and the y-axis represents the thirteen selected cities in the BTH region.

scale straw burning in late autumn (i.e., October, Fig. 3d) in the northwest rural areas raised the PM_{2.5} level again (Duan et al., 2004). The monthly mean PM_{2.5} concentrations at the city scale (Fig. 3e) suggest that the highest (lowest) PM_{2.5} pollution always occurred in December (August) at most of the selected cities.

3.3. Temporal dynamics of population exposure to PM_{2.5}

The temporal dynamics of population exposure to PM_{2.5} were evaluated using the satellite-station-hybrid model by incorporating dynamic population distribution and PM_{2.5} concentrations. Fig. 4a shows the temporal variation of monthly exposure for each city from January to December 2015. Except for Zhangjiakou, Chengde, and Hengshui, most cities presented an obvious unimodal structure in the time-series profiles, with peak PM_{2.5} exposure levels in December. Meanwhile, a slight rise of PM_{2.5} exposure levels in July could be found in some cities (e.g., Beijing, Handan, Zhangjiakou). The annual mean PM_{2.5} exposure in 2015 shows that a cluster of cities in the southwest of the BTH region (i.e., Handan, Xingtai, Shijiazhuang, Hengshui, and Baoding) have experienced relatively higher exposure levels than other cities (Fig. 4b). Baoding and Hengshui were the two most polluted cities with annual mean PM_{2.5} exposures reaching up to 93.16 μg/m³/year and 90.63 μg/m³/year, respectively.

To better present the monthly variation in PM_{2.5} exposure, we aggregated the pixel-based assessments to obtain the monthly cumulative percentage of PM_{2.5} exposure in the BTH region (Fig. 5). The results show that January to February and November to December are the two worst periods with high PM_{2.5} pollution in 2015. More than 50% (20%) of the population in the BTH region were exposed to an average PM_{2.5} concentration higher than 80 (120) μg/m³. In December, a sudden deterioration appeared and 50% of the population were exposed to an average PM_{2.5} level higher than 110 μg/m³. This phenomenon could be explained by increasing heating demands and coal consumption during the cold season. Notably, most of the coal combustion and emissions

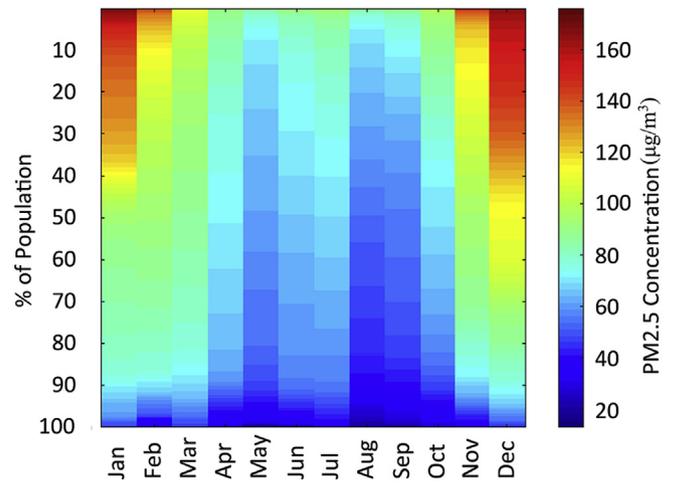


Fig. 5. Cumulative percentage of PM_{2.5} exposure in the BTH region. The x-axis represents January to December 2015, and the y-axis represents the percentage of the population in the entire BTH region.

occurred inside areas of human settlement, which directly deteriorates ambient air quality and human health. The period with the lowest PM_{2.5} exposure level was around August, but still, 50% of the population were exposed to PM_{2.5} concentrations higher than 50 μg/m³.

3.4. Total attributive deaths for different causes

Three health endpoints were considered in this study and coded according to the International Classification of Diseases, Revision 10: all causes (A00–R99), cardiovascular diseases (I00–I99), and respiratory diseases (J00–J98). Table 2 presents the annual premature deaths estimated by the C-R function under the WHO AQG PM_{2.5} scenario (10 μg/m³/year). The results showed that the total

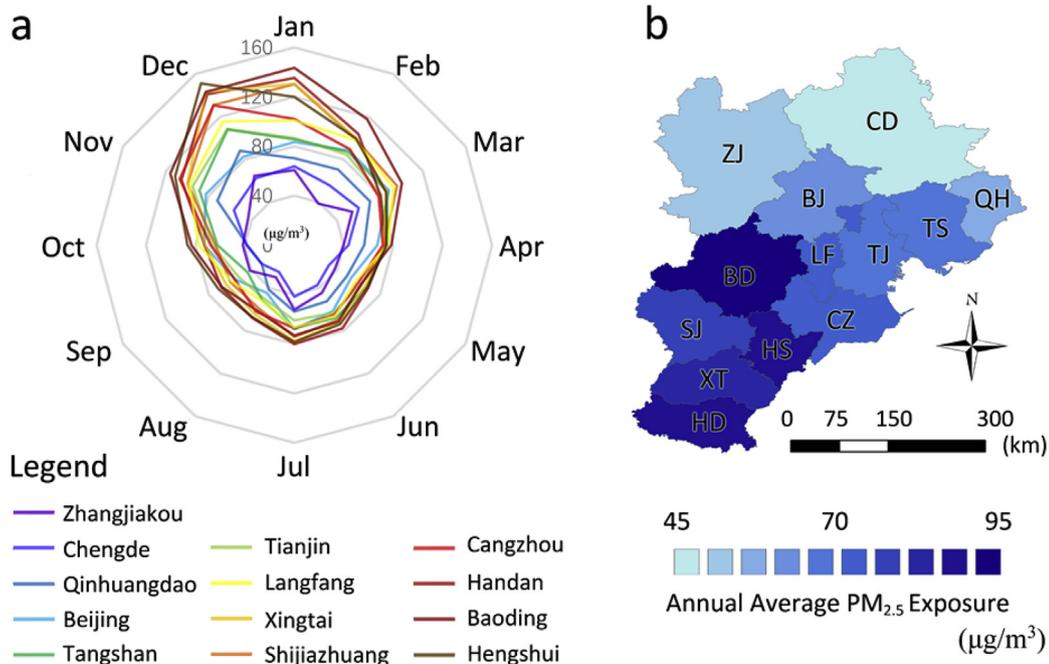


Fig. 4. (a) Temporal variation of the monthly PM_{2.5} exposure levels of the selected cities. (b) Annual mean PM_{2.5} exposure of each city in the BTH region.

Table 2
Premature deaths due to PM_{2.5} for each city in the BTH region.

City	All cause (A00–R99)		Cardiovascular (I00–I99)		Respiratory (J00–J98)	
	Premature Mortality	95% CI	Premature Mortality	95% CI	Premature Mortality	95% CI
Beijing	22,153	(18,447, 25,706)	12,965	(8,107, 17,432)	2900	(1,814, 3900)
Tianjin	18,657	(15,551, 21,627)	12,500	(7,839, 16,757)	1971	(1,236, 2643)
Shijiazhuang	15,440	(12,915, 17,835)	8787	(5,569, 11,658)	2199	(1,394, 2917)
Tangshan	6966	(5,807, 8073)	3989	(2503, 5345)	998	(626, 1337)
Qinhuangdao	3002	(2,489, 3498)	1736	(1,072, 2364)	434	(268, 591)
Handan	10,388	(8,702, 11,982)	5896	(3,754, 7788)	1475	(939, 1949)
Xingtai	11,074	(9,271, 12,782)	6293	(3,998, 8329)	1575	(1,000, 2084)
Baoding	16,265	(13,642, 18,740)	9213	(5,885, 12,129)	2305	(1,473, 3035)
Zhangjiakou	3179	(2,626, 3719)	1852	(1,130, 2553)	463	(283, 639)
Chengde	4165	(3,439, 4872)	2426	(1,480, 3344)	607	(370, 837)
Cangzhou	10,106	(8,440, 11,693)	5769	(3,639, 7689)	1443	(911, 1924)
Langfang	4335	(3,620, 5015)	2474	(1,561, 3297)	619	(391, 825)
Hengshui	12,420	(10,409, 14,321)	7045	(4,490, 9297)	1763	(1,123, 2326)
Total	138,150	(115,358, 159,863)	80,945	(51,027, 107,982)	18,752	(11,828, 25,007)

number of premature deaths for each city ranged from 3002 to 22,153 for all-cause, 1736 to 12,965 for cardiovascular diseases, and 434 to 2900 for respiratory diseases in 2015. The total number of premature deaths in the BTH region was 138,150, 80,945, and 18,752 for all-cause, cardiovascular diseases, and respiratory diseases, respectively.

3.5. Comparison of five exposure models

Furthermore, the exposure levels and health outcomes (all-cause mortality) were estimated using other four exposure models based on different datasets (Table 3). The quantitative comparison shows that there are obvious disagreements in the estimated average exposure levels and the corresponding all-cause premature mortalities among the five exposure assessment models. For example, in Beijing, the minimum and maximum number of all-cause premature deaths were observed in Model-3 and Model-2 respectively. Similar situations were also found in Shijiazhuang, Tangshan, Xingtai, Baoding, Chengde, Cangzhou, and Hengshui. Compared with our model's results (Model-5), overestimates existed in the Model-1's and Model-2's results with a positive bias of 3309 and 4841 premature deaths, respectively; by contrast, underestimates existed in the Model-3's and Model-4's results with a negative bias of 6017 and 1881 premature deaths, respectively.

Table 3
Comparison of average exposure levels and premature deaths estimated by five exposure models.

City	Model-1		Model-2		Model-3		Model-4		Model-5	
	Average Exposure ^a	Premature Mortality	Average Exposure ^a	Premature Mortality	Average Exposure ^a	Premature Mortality	Average Exposure ^a	Premature Mortality	Average Exposure ^a	Premature Mortality
Beijing	68.561	21,459	72.862	22,855 ^c	63.301	19,721 ^b	70.328	22,035	70.691	22,153
Tianjin	75.414	19,081 ^c	74.006	18,718	74.367	18,812	73.319	18,541 ^b	73.768	18,657
Shijiazhuang	90.535	16,454	89.898	16,343 ^c	80.108	14,597 ^b	83.968	15,293	84.787	15,440
Tangshan	72.343	6,780 ^b	74.041	6,943 ^c	72.132	6,759 ^b	73.791	6,919	74.282	6,966 ^c
Qinghuangdao	57.188	2,966 ^b	57.266	2971	57.354	2976	57.579	2989	57.813	3,002 ^c
Handan	92.861	10,753	93.091	10,778 ^c	87.419	10,145 ^b	89.336	10,361	89.577	10,388
Xingtai	93.106	11,773	94.073	11,889 ^c	85.247	10,812 ^b	87.187	11,052	87.365	11,074
Baoding	95.914	16,719	98.639	17,165 ^c	83.634	14,650 ^b	92.743	16,194	93.169	16,265
Zhangjiakou	49.206	3,464 ^c	47.701	3340	44.639	3,087 ^b	45.629	3169	45.754	3179
Chengde	50.840	4724	49.440	4,574 ^c	44.673	4,057 ^b	45.196	4114	45.659	4165
Cangzhou	80.200	10,160	81.051	10,267 ^c	77.399	9,805 ^b	78.233	9911	79.773	10,106
Langfang	77.305	4191	77.135	4,182 ^b	80.677	4,374 ^c	79.941	4335	79.949	4335
Hengshui	94.578	12,935	94.813	12,966 ^c	90.016	12,338 ^b	90.146	12,356	90.637	12,420
Totals		141,459		142,991^c		132,133^b		137,269		138,150

^a $\mu\text{g}/\text{m}^3/\text{year}$.

^b The minimum number of premature mortalities among the five models in each city.

^c The maximum number of premature mortalities among the five models in each city.

4. Discussion

This study introduced a robust PM_{2.5} exposure model using the PM_{2.5} concentrations derived from the GWTR satellite-ground-hybrid model, and LBS-based dynamic population maps to assess the monthly exposure levels and the health outcomes at thirteen cities in the BTH region in 2015. Compared with previous exposure assessment methods, the newly proposed method adequately considered the estimation accuracy of PM_{2.5} concentrations and the spatiotemporal variability of population distribution. In addition, our method was compared with the other four PM_{2.5} exposure models based on commonly used datasets. The disagreements in exposure level and annual premature mortality among the five exposure models highlighted the effects in the environmental assessments when the complex and changeable interactions between the population and ambient environment were ignored.

The PM_{2.5} concentrations estimated by the GTWR satellite-station-hybrid model can more accurately capture the spatial variabilities and ensure the accuracy of subsequent assessments. Due to the limited number of ground stations, station-based observations alone cannot provide accurate spatial information on PM_{2.5} concentrations, thus may lead to potential biases in assessments regarding pollution exposure and health impacts. The differences in the estimated premature mortalities (or exposure levels) between

model-1 and model-3 (as well as model-2 and model-4) revealed this issue. Moreover, a similar comparative study conducted by [Jerrett et al. \(2016\)](#) also verified that simply using station-based exposure estimates may result in incorrect estimations of health risk. For monthly exposure estimations, regional differences also existed when comparing the five models ([Fig. S1](#) in Supplementary Material). For example, for these cities located in the south BTH-region (i.e., Baoding, Shijiazhuang, Xingtai, Handan, and Hengshui), the $PM_{2.5}$ exposure levels estimated by Model-1/2 were over-estimated compared with other three models' results in winter (around January). Opposite results were found in these cities located in the east BTH region (i.e., Qinhuangdao, Tangshan, Langfang, and Tianjin), where under-estimations were found in the exposure levels estimated by Model-1/2 compared with other models in summer (around July). These dissimilarities were likely caused by the accuracy of $PM_{2.5}$ estimations and the $PM_{2.5}$ variations in sources or composition. This further revealed that the satellite-station-hybrid model could more accurately capture the spatial variation of air pollution. Therefore, the usage of satellite-station-based $PM_{2.5}$ concentrations, with improved estimation accuracy and higher spatial-resolution information, should be a promising direction for future related studies.

More importantly, some potential biases and limitations in this study should be highlighted. First, LBS data (e.g., social media and mobile phone data) are regarded as non-representative data ([Kwan, 2016](#); [Zagheni and Weber, 2015](#)), and the density maps of LBS data actually present the distribution of active users rather than the real population density. This type of data tends to leave out some sections of society (e.g., children, the elderly, and the poor) since a lower proportion of these people actively use mobile internet services or intelligent terminals compared to the whole society. However, previous studies have demonstrated that potential sampling bias does not debilitate social-media-based LSBD's performance in characterizing dynamic population distribution. However, uncertainty still exists when using such datasets, and further validation against ground-truth datasets is needed. Therefore, we should remain cautious about the results and conclusions obtained using these data. Second, most human mortality datasets released by the Chinese government are recorded at the city scale. To maintain the sample size consistency, the health impact assessment of $PM_{2.5}$ exposure was conducted based on city boundaries in this study. However, the aggregation process may cause some potential biases (e.g., the zoning effect of MAUP). Therefore, mortality information with higher spatial resolutions (e.g., street-level) is needed to minimize the uncertainties in the environmental health assessments in China. Lastly, due to the limitations on satellite data sources and quality, the spatiotemporal resolutions (i.e., 3-km, daily) of the data used for assessments are not high enough. Even though the satellite-station-hybrid model used in this study shows satisfactory accuracy and high repeatability, $PM_{2.5}$ estimations with a high spatiotemporal resolution are still important, because they can effectively delineate the relevant environmental factors affecting people. Fortunately, several meteorological satellites with high spatiotemporal resolution (e.g., Himawari-8 at a 10-min temporal resolution and 500-m spatial resolution) have been launched and can help reduce uncertainties in future air pollution exposure assessment. Besides, some well-designed site-based (city-scale) models ([Shi et al., 2018](#)), as well as machine learning techniques ([Xu et al., 2018](#)), has been proved to significantly improve the estimations in ground-level $PM_{2.5}$ concentrations with more detailed spatial information. Therefore, how to better incorporate these site-based models and advanced techniques in region-scale $PM_{2.5}$ concentration mapping and exposure assessment remains to be an open question for future research.

5. Conclusions

This study tried to introduce an improved $PM_{2.5}$ exposure model by integrating the dynamic population distribution and satellite-station-based $PM_{2.5}$ concentrations. The results showed that the proposed method could effectively quantify the dynamic city-scale $PM_{2.5}$ exposure levels and the three related causes of annual premature mortalities in the BTH region, China. The comparative analysis of five exposure models further illustrated the importance of the estimation accuracy of $PM_{2.5}$ concentrations and dynamic changes in population distribution which should be carefully considered in environmental exposure or health assessments. Although some limitations exist, our study provided a robust method to quantify environmental exposure and related health burdens, which can be extended to exposure estimations to other environmental factors. These findings will help researchers and policymakers better understand the spatial patterns and effects of $PM_{2.5}$ pollution exposure and establish effective pollution-control measures.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

This study was supported by the Fundamental Research Funds for the Central Universities (WUT: 193108007). The authors thank two anonymous reviewers and the editor for providing helpful suggestions and thoughtful comments which helped to improve the manuscript. The authors also thank Sina Weibo for making the LBS data publicly available.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envpol.2019.06.057>.

References

- [Anenberg, S.C., Horowitz, L.W., Tong, D.Q., West, J.J., 2010.](#) An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling. *Environ. Health Perspect.* 118, 1189–1195.
- [Bai, Y., Wu, L., Qin, K., Zhang, Y., Shen, Y., Zhou, Y., 2016.](#) A geographically and temporally weighted regression model for ground-level $PM_{2.5}$ estimation from satellite-derived 500 m resolution AOD. *Rem. Sens.* 8, 262.
- [Cai, J., Huang, B., Song, Y., 2017.](#) Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sens. Environ.* 202, 210–221.
- [Chen, B., Song, Y., Jiang, T., Chen, Z., Huang, B., Xu, B., 2018a.](#) Real-time estimation of population exposure to $PM_{2.5}$ using mobile-and station-based big data. *Int. J. Environ. Res. Public Health* 15, 573.
- [Chen, B., Song, Y., Kwan, M.-P., Huang, B., Xu, B., 2018b.](#) How do people in different places experience different levels of air pollution? Using worldwide Chinese as a lens. *Environ. Pollut.* 238, 874–883.
- [Chen, H., Lin, Y., Su, Q., Cheng, L., 2017a.](#) Spatial variation of multiple air pollutants and their potential contributions to all-cause, respiratory, and cardiovascular mortality across China in 2015–2016. *Atmos. Environ.* 168, 23–35.
- [Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., Xu, X., Pei, F., 2017b.](#) Delineating urban functional areas with building-level social media data: a dynamic time warping (DTW) distance based k-medoids method. *Landsc. Urban Plan.* 160, 48–60.
- [China, N., 2016.](#) Statistical Communiqué of the People's Republic of China on the 2015 National Economic and Social Development. *National Bureau of Statistics of China*, Beijing.
- [Chu, H.-J., Huang, B., Lin, C.-Y., 2015.](#) Modeling the spatio-temporal heterogeneity in the PM_{10} - $PM_{2.5}$ relationship. *Atmos. Environ.* 102, 176–182.
- [Cohen, A.J., Anderson, H.R., Ostro, B., Pandey, K.D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Pope III, C.A., Romieu, I., Samet, J.M., 2004.](#) Urban Air Pollution. Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors, vol. 2, pp. 1353–1433.
- [Dobson, J.E., Bright, E.A., Coleman, P.R., Durfee, R.C., Worley, B.A., 2000.](#) LandScan: a global population database for estimating populations at risk. *Photogramm.*

- Eng. Rem. Sens. 66, 849–857.
- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris Jr., B.G., Speizer, F.E., 1993. An association between air pollution and mortality in six US cities. *N. Engl. J. Med.* 329, 1753–1759.
- Duan, F., Liu, X., Yu, T., Cachier, H., 2004. Identification and estimate of biomass burning contribution to the urban aerosol organic carbon concentrations in Beijing. *Atmos. Environ.* 38, 1275–1282.
- Fang, C., Wang, Z., Xu, G., 2016. Spatial-temporal characteristics of PM 2.5 in China: a city-level perspective analysis. *J. Geogr. Sci.* 26, 1519–1532.
- Fleischer, N.L., Meriardi, M., van Donkelaar, A., Vadillo-Ortega, F., Martin, R.V., Betran, A.P., Souza, J.P., Neill, M.S.O., 2014. Outdoor air pollution, preterm birth, and low birth weight: analysis of the world health organization global survey on maternal and perinatal health. *Environ. Health Perspect.* 122, 425–430.
- Gray, S.C., Edwards, S.E., Schultz, B.D., Miranda, M.L., 2014. Assessing the impact of race, social factors and air pollution on birth outcomes: a population-based study. *Environ. Health* 13, 4.
- Guo, Y., Tang, Q., Gong, D.-Y., Zhang, Z., 2017. Estimating ground-level PM2.5 concentrations in Beijing using a satellite-based geographically and temporally weighted regression model. *Remote Sens. Environ.* 198, 140–149.
- Hawelka, B., Sitko, I., Beinart, E., Sobolevsky, S., Kazakopoulos, P., Ratti, C., 2014. Geo-located Twitter as proxy for global mobility patterns. *Cartogr. Geogr. Inf. Sci.* 41, 260–271.
- He, Q., Huang, B., 2018a. Satellite-based high-resolution PM2.5 estimation over the Beijing-Tianjin-Hebei region of China using an improved geographically and temporally weighted regression model. *Environ. Pollut.* 236, 1027–1037.
- He, Q., Huang, B., 2018b. Satellite-based mapping of daily high-resolution ground PM2.5 in China via space-time regression modeling. *Remote Sens. Environ.* 206, 72–83.
- Ho, H., Knudby, A., Huang, W., 2015. A spatial framework to map heat health risks at multiple scales. *Int. J. Environ. Res. Public Health* 12, 16110–16123.
- Ho, H.C., Wong, M.S., Yang, L., Shi, W., Yang, J., Bilal, M., Chan, T.-C., 2018. Spatio-temporal influence of temperature, air quality, and urban environment on cause-specific mortality during hazy days. *Environ. Int.* 112, 10–22.
- Huang, W., Li, S., 2016. Understanding human activity patterns based on space-time-semantics. *ISPRS J. Photogrammetry Remote Sens.* 121, 1–10.
- Jerrett, M., Turner, M.C., Beckerman, B.S., Pope III, C.A., van Donkelaar, A., Martin, R.V., Serre, M., Crouse, D., Gapstur, S.M., Krewski, D., 2016. Comparing the health effects of ambient particulate matter estimated using ground-based versus remote sensing exposure estimates. *Environ. Health Perspect.* 125, 552–559.
- Jurdak, R., Zhao, K., Liu, J., Aboujaoude, M., Cameron, M., Newth, D., 2015. Understanding human mobility from Twitter. *PLoS One* 10, e0131469.
- Kan, H., Chen, R., Tong, S., 2012. Ambient air pollution, climate change, and population health in China. *Environ. Int.* 42, 10–19.
- Kwan, M.-P., 2016. Algorithmic geographies: big data, algorithmic uncertainty, and the production of geographic knowledge. *Ann. Assoc. Am. Geogr.* 106, 274–282.
- Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* 525, 367.
- Liao, Z., Gao, M., Sun, J., Fan, S., 2017. The impact of synoptic circulation on air quality and pollution-related human health in the Yangtze River Delta region. *Sci. Total Environ.* 607, 838–846.
- Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., Hong, Y., 2017. Classifying urban land use by integrating remote sensing and social media data. *Int. J. Geogr. Inf. Sci.* 31, 1675–1696.
- Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., Shi, L., 2015. Social sensing: a new approach to understanding our socioeconomic environments. *Ann. Assoc. Am. Geogr.* 105, 512–530.
- Lu, F., Xu, D., Cheng, Y., Dong, S., Guo, C., Jiang, X., Zheng, X., 2015. Systematic review and meta-analysis of the adverse health effects of ambient PM2.5 and PM10 pollution in the Chinese population. *Environ. Res.* 136, 196–204.
- Lu, X., Yao, T., Fung, J.C., Lin, C., 2016. Estimation of health and economic costs of air pollution over the Pearl River Delta region in China. *Sci. Total Environ.* 566, 134–143.
- Luo, F., Cao, G., Mulligan, K., Li, X., 2016. Explore spatiotemporal and demographic characteristics of human mobility via Twitter: a case study of Chicago. *Appl. Geogr.* 70, 11–25.
- Lv, B., Hu, Y., Chang, H.H., Russell, A.G., Cai, J., Xu, B., Bai, Y., 2017. Daily estimation of ground-level PM2.5 concentrations at 4 km resolution over Beijing-Tianjin-Hebei by fusing MODIS AOD and ground observations. *Sci. Total Environ.* 580, 235–244.
- Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating ground-level PM2.5 in China using satellite remote sensing. *Environ. Sci. Technol.* 48, 7436–7444.
- Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., Liu, Y., 2015. Satellite-based spatiotemporal trends in PM2.5 concentrations: China, 2004–2013. *Environ. Health Perspect.* 124, 184–192.
- Nel, A., 2005. Air pollution-related illness: effects of particles. *Science* 308, 804–806.
- Nyhan, M., Grauwil, S., Britter, R., Misstear, B., McNabola, A., Laden, F., Barrett, S.R., Ratti, C., 2016. “Exposure track”—the impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution. *Environ. Sci. Technol.* 50, 9671–9681.
- Park, Y.M., Kwan, M.-P., 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health Place* 43, 85–94.
- Peng, J., Chen, S., Lü, H., Liu, Y., Wu, J., 2016. Spatiotemporal patterns of remotely sensed PM2.5 concentration in China from 1999 to 2011. *Remote Sens. Environ.* 174, 109–121.
- Pope III, C.A., Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K., Thurston, G.D., 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama* 287, 1132–1141.
- Pope III, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. *J. Air Waste Manag. Assoc.* 56, 709–742.
- Rashidi, T.H., Abbasi, A., Maghrebi, M., Hasan, S., Waller, T.S., 2017. Exploring the capacity of social media data for modelling travel behaviour: opportunities and challenges. *Transp. Res. Part C* 75, 197–211.
- Shang, Y., Sun, Z., Cao, J., Wang, X., Zhong, L., Bi, X., Li, H., Liu, W., Zhu, T., Huang, W., 2013. Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. *Environ. Int.* 54, 100–111.
- Shen, F., Ge, X., Hu, J., Nie, D., Tian, L., Chen, M., 2017. Air pollution characteristics and health risks in Henan Province, China. *Environ. Res.* 156, 625–634.
- Shi, Y., Ho, H.C., Xu, Y., Ng, E., 2018. Improving satellite aerosol optical Depth-PM2.5 correlations using land use regression with microscale geographic predictors in a high-density urban context. *Atmos. Environ.* 190, 23–34.
- Song, W., Jia, H., Huang, J., Zhang, Y., 2014. A satellite-based geographically weighted regression model for regional PM2.5 estimation over the Pearl River Delta region in China. *Remote Sens. Environ.* 154, 1–7.
- Song, Y., Huang, B., Cai, J., Chen, B., 2018. Dynamic assessments of population exposure to urban greenspace using multi-source big data. *Sci. Total Environ.* 634, 1315–1325.
- Steiger, E., Westerholt, R., Resch, B., Zipf, A., 2015. Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. *Comput. Environ. Urban Syst.* 54, 255–265.
- Sun, L., Wei, J., Duan, D., Guo, Y., Yang, D., Jia, C., Mi, X., 2016. Impact of land-use and land-cover change on urban air quality in representative cities of China. *J. Atmos. Sol. Terr. Phys.* 142, 43–54.
- Van Donkelaar, A., Martin, R.V., Brauer, M., Boys, B.L., 2014. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environ. Health Perspect.* 123, 135–143.
- Wang, J., Christopher, S.A., 2003. Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: implications for air quality studies. *Geophys. Res. Lett.* 30.
- Wang, J., Kwan, M.-P., Chai, Y., 2018. An innovative context-based crystal-growth activity space method for environmental exposure assessment: a study using GIS and GPS trajectory data collected in Chicago. *Int. J. Environ. Res. Public Health* 15, 703.
- Wang, J., Wang, S., Voorhees, A.S., Zhao, B., Jang, C., Jiang, J., Fu, J.S., Ding, D., Zhu, Y., Hao, J., 2015. Assessment of short-term PM2.5-related mortality due to different emission sources in the Yangtze River Delta, China. *Atmos. Environ.* 123, 440–448.
- Wei, J., Li, Z., Peng, Y., Sun, L., 2019. MODIS Collection 6.1 aerosol optical depth products over land and ocean: validation and comparison. *Atmos. Environ.* 201, 428–440.
- Weibo-Corporation, 2015. Retrieved from. <http://www.nasdaq.com/press-release/weibo-reports-third-quarter-2015-results-20151118-01121>.
- Xie, Y., Wang, Y., Zhang, K., Dong, W., Lv, B., Bai, Y., 2015. Daily estimation of ground-level PM2.5 concentrations over Beijing using 3 km resolution MODIS AOD. *Environ. Sci. Technol.* 49, 12280–12288.
- Xu, Y., Ho, H.C., Wong, M.S., Deng, C., Shi, Y., Chan, T.-C., Knudby, A., 2018. Evaluation of machine learning techniques with multiple remote sensing datasets in estimating monthly concentrations of ground-level PM2.5. *Environ. Pollut.* 242, 1417–1426.
- Yang, H., Tao, W., Liu, Y., Qiu, M., Liu, J., Jiang, K., Yi, K., Xiao, Y., Tao, S., 2019. The contribution of the Beijing, Tianjin and Hebei region's iron and steel industry to local air pollution in winter. *Environ. Pollut.* 245, 1095–1106.
- Zagheni, E., Weber, I., 2015. Demographic research with non-representative internet data. *Int. J. Manpow.* 36, 13–25.
- Zhang, Q., Streets, D.G., Carmichael, G.R., He, K., Huo, H., Kannari, A., Klimont, Z., Park, I., Reddy, S., Fu, J., 2009. Asian emissions in 2006 for the NASA INTEX-B mission. *Atmos. Chem. Phys.* 9, 5131–5153.
- Zhang, S., Zhou, W., 2018. Recreational visits to urban parks and factors affecting park visits: evidence from geotagged social media data. *Landsc. Urban Plan.* 180, 27–35.
- Zhang, Z., He, Q., Zhu, S., 2017. Potentials of using social media to infer the longitudinal travel behavior: a sequential model-based clustering method. *Transp. Res. Part C* 85, 396–414.
- Zhao, B., Wang, S., Ding, D., Wu, W., Chang, X., Wang, J., Xing, J., Jang, C., Fu, J.S., Zhu, Y., 2019. Nonlinear relationships between air pollutant emissions and PM2.5-related health impacts in the Beijing-Tianjin-Hebei region. *Sci. Total Environ.* 661, 375–385.
- Zheng, S., Wang, J., Sun, C., Zhang, X., Kahn, M.E., 2019. Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nat. Hum. Behav.* 1.