Environmental Pollution 238 (2018) 874–883

Contents lists available at ScienceDirect

Environmental Pollution

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

How do people in different places experience different levels of air pollution? Using worldwide Chinese as a lens☆

Bin Chen [a, b], Yimeng Song [c], Mei-Po Kwan [d, e], Bo Huang [c], Bing Xu [a, f, g, *]

[a] Ministry of Education Key Laboratory for Earth System Modelling, Department of Earth System Science, Tsinghua University, Beijing 100084, China
[b] Department of Land, Air, and Water Resources, University of California, Davis, CA 95616, USA
[c] Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, Hong Kong
[d] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[e] Department of Human Geography and Spatial Planning, Utrecht University, 3508 TC Utrecht, The Netherlands
[f] State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China
[g] Department of Geography, University of Utah, 260 S. Central Campus Dr., Salt Lake City, UT 841112, USA

ARTICLE INFO

Article history:
Received 25 January 2018
Received in revised form
17 March 2018
Accepted 26 March 2018

Keywords:
$PM_{2.5}$
Exposure risk
Spatiotemporal difference
Worldwide chinese
Mobile phone location data

ABSTRACT

Air pollution, being especially severe in the fast-growing developing world, continues to post a threat to public health. Yet, few studies are capable of quantifying well how different groups of people in different places experience different levels of air pollution at the global scale. In this paper, we use worldwide Chinese as a lens to quantify the spatiotemporal variations and geographic differences in $PM_{2.5}$ exposures using unprecedented mobile phone big data and air pollution records. The results show that Chinese in South and East Asia suffer relatively serious $PM_{2.5}$ exposures, where the Chinese in China have the highest $PM_{2.5}$ exposures (52.8 $\mu g/m^3$/year), which is fourfold higher than the exposures in the United States (10.7 $\mu g/m^3$/year). Overall, the Chinese in Asian cities (35.5 $\mu g/m^3$/year) experienced the most serious $PM_{2.5}$ exposures when compared with the Chinese in the cities of other continents. These results, partly presented as a spatiotemporally explicit map of $PM_{2.5}$ exposures for worldwide Chinese, help researchers and governments to consider how to address the effects of air pollution on public health with respect to different population groups and geographic locations.

1. Introduction

Air pollutants, especially fine particulate matter such as $PM_{2.5}$ (particles with an aerodynamic diameter of less than 2.5 μm), have been the focus of increasing public concern because of their potential adverse impacts on human health (Apte et al., 2015; Franklin et al., 2007; Kioumourtzoglou et al., 2016; Kloog et al., 2013; Pope III et al., 2009). Numerous epidemiologic studies have established robust associations between long-term exposure to $PM_{2.5}$ and premature mortality associated with various health conditions—such as heart disease, cardiovascular and respiratory diseases, and lung cancer—that substantially reduce life expectancy (Apte et al., 2015; Franklin et al., 2007; Kioumourtzoglou et al., 2016; Kloog et al., 2013; Pope III et al., 2009). Previous air pollution exposure studies have worked to obtain refined exposure estimates with fine spatiotemporal resolutions (Apte et al., 2015; Han et al., 2016; Ma Z 2016; Park and Kwan, 2017; Van Donkelaar et al., 2010) in order to better address public health issues associated with $PM_{2.5}$ exposure (Di et al., 2017; Kioumourtzoglou et al., 2016; Kloog et al., 2013; Pope III et al., 2009). However, assessing how people in different places experience different levels of air pollution is still a major challenge, especially for specific groups of population at the regional or global scale.

Currently, demographic data based on administrative boundaries is the most widely used data for estimating people's exposures to air pollution (Fleischer et al., 2014; Gray et al., 2014). It provides accurate population census information over a certain period based on the smallest administrative unit (e.g., census block). However, such kind of data has limitations for comparing the exposures of the people in different countries since the data collection procedures used to collect demographic information may not be consistent among different nations. In addition, census data

☆ This paper has been recommended for acceptance by Haidong Kan.

* Corresponding author. Ministry of Education Key Laboratory for Earth System Modelling, Department of Earth System Science, Tsinghua University, Beijing 100084, China.
E-mail address: bingxu@tsinghua.edu.cn (B. Xu).

https://doi.org/10.1016/j.envpol.2018.03.093
0269-7491/

was collected based on census units. It does not consider the spatial variability in population distribution and has very low update frequency (5–10 years), which inevitably introduce considerable uncertainty to the accurate exposure assessment. In contrast, global gridded population data such as LandScan Global Population (Dobson et al., 2000), and Gridded Population of the World (GPW) series (Center for International Earth Science Information Network, 2016) are able to depict the geographic distribution of humans at large scales (e.g., covering many countries) with finer spatial details. However, the main data source of these gridded population data is census population data. Therefore, these datasets will still be affected by the potential concerns mentioned above (Center for International Earth Science Information Network, 2016; Dobson et al., 2000). More importantly, these datasets just regard population as a homogeneous entity and thus cannot be used to identify and differentiate population groups with respect to ethnicity or other socio-economic attributes, which prevents previous studies from investigating the spatial differences and the long-term cumulative differences in air pollution exposures for specific groups of the population. Fortunately, the popularization of mobile computing platforms (e.g., smartphones) and rapid growth of mobile apps (applications designed to run on smartphones and other mobile devices for social networking, navigation, shopping, and dining) produce large amounts of geotagged information and provide researchers with unprecedented opportunities to discover the spatiotemporal characteristics of human activities (Cheng et al., 2011; Lee and Sumiya, 2010; Preoţiuc-Pietro and Cohn, 2013). Moreover, the high geographic correlation between the distribution of geo-spatial big data (e.g., mobile phone records, social media check-in records, taxi trajectories, and smart card records) and human distribution has been widely revealed in previous studies (Fang et al., 2013; Lelieveld et al., 2015; Van Donkelaar et al., 2015; Zhou et al., 2010). The advent of geospatial big data has led to the promising direction of incorporating the socio-economic attributes and mobility of human beings in environmental exposure assessments to discover more specific facets of population exposure to air pollution. From GPS trajectory data of cars to social media check-in records and mobile phone data, a growing number of data sources have been used in relevant studies (Dewulf et al., 2016; Gariazzo et al., 2016; Nyhan et al., 2016). By identifying individuals' behavioral patterns in continuous space-time, geospatial big data may also be used to address the uncertain geographic context problem (UGCoP), which is a common problem in environmental health research because data with coarse spatial and temporal resolution cannot accurately assess individuals' actual environmental exposures (Kwan, 2012; Park and Kwan, 2017).

Meanwhile, some existing caveats concerning the use of geospatial big data should be pointed out here. First, geotagged information derived from multi-source platforms tends to contain considerable noises and discrepancies caused by the different number and composition of active users in terms of ethnicity, culture, education, occupation, income, and age groups. It is thus difficult to obtain geospatial datasets adequately characterize the socio-economic attributes of different groups of population at the regional or global scales. Second, most open-source geospatial datasets have limited spatial and temporal coverage, which hinders the characterization of long-term human activity patterns at the global scale. Third, despite the above issues, limited studies have attempted to combine geospatial big data to investigate how different groups of people who live in different places (at a global scale) experience different levels of air pollution.

Globalization of the 21st century has ushered in an era of fluidity and openness, in which changes in transportation, technology and culture are encouraging people to move across national borders with multiple purposes (e.g., work, settlement, study, professional advancement, marriage, retirement, or lifestyle change) (Castles, 2010), thus leading to the movement of different groups of people among different geographic areas at various spatial scales. For example, with roughly a 12.6% increase (~5 million) in overseas Chinese population during the period 2001–2011, this trend is significant for China (Poston and Wong, 2016). However, knowledge of the fine-resolution distribution of Chinese worldwide still remains limited despite national or international consensus on the distribution and size of oversea Chinese population. Similarly, knowledge of the geographic distribution of other population groups at the global scale is also highly limited to date.

This study seeks to address the challenges in assessing air pollution risks from the perspective of a specific population group and the difficulties in characterizing the geographic distribution of different population groups at the global scale. It uses the global Chinese population as a lens to quantify how people living in different places in the world experience different levels of air pollution. Specifically, this paper presents a spatiotemporal analysis of $PM_{2.5}$ exposures for the global Chinese population using unprecedented mobile phone big data and air pollution records.

2. Data and methods

2.1. Mobile phone location-based big data

In this study, we use the location information in a big mobile phone dataset from Tencent (China) to portray the geographic distribution of the global Chinese population. All of the location data is produced by Tencent through retrieving real-time locations of active mobile phone users when they are using Tencent applications and Tencent's location-based service (LBS) invoked by other mobile apps. As one of the world's largest internet service providers for ethnic Chinese, and given the widespread use of Tencent's service and apps (e.g., Wechat, QQ, etc.), the daily location records have reached 38 billion from more than 450 million users globally in 2016 (Tencent, 2016). It can be argued that the geography of Tencent location data presents a unique geographic distribution of worldwide Chinese. The dataset in this study was collected from March 14, 2016, to August 13, 2016. It has a spatial resolution of 36 arc-second (~1.2 km) and a temporal resolution of 5-min, and is retrieved and updated using the application program interface (API) from the Tencent location big data website (http://heat.qq.com). All the information regarding users' identities and privacies were removed from the public released dataset.

2.2. Global $PM_{2.5}$ concentration dataset

The time-series $PM_{2.5}$ observation records used in this study came from the global $PM_{2.5}$ concentration dataset (Van Donkelaar et al., 2010). Using a simulation of GEOS-Chem chemical transport model, this $PM_{2.5}$ concentration dataset was estimated from an integration of Moderate Resolution Imaging Spectroradiometer (MODIS) and Multi-angle Imaging SpectroRadiometer (MISR) aerosol optical depth (AOD) data with aerosol vertical profiles and scattering properties (Van Donkelaar et al. 2010, 2013). It has a spatial resolution of 10 km and a temporal coverage from 1999 to 2011 as a 3-year moving average, which was applied to reduce the retrieval biases of the annual $PM_{2.5}$ concentrations. Additionally, this dataset has been validated with ground observations at the global scale (Van Donkelaar et al. 2010, 2013, 2015). Experimental tests show that there is a significant agreement between satellite-based estimates and ground-based measurements across different continents (Van Donkelaar et al. 2010, 2013, 2015). The dataset thus provides us a spatially explicit and temporally consistent $PM_{2.5}$ concentration dataset for this study.