Mapping Total Exceedance PM$_{2.5}$ Exposure Risk by Coupling Social Media Data and Population Modeling Data

Zheng Cao$^{1,2}$, Guanhua Guo$^{1,2}$, Zhiheng Wu$^{1,2}$, Shaoying Li$^{1,2}$, Hui Sun$^{1,2}$, and Wenchuan Guan$^1$

$^1$School of Geographical Sciences, Guangzhou University, Guangzhou, China, $^2$Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China

Abstract  The PM$_{2.5}$ exposure risk assessment is the foundation to reduce its adverse effects. Population survey-related data have been deficient in high spatiotemporal detailed descriptions. Social media data can quantify the PM$_{2.5}$ exposure risk at high spatiotemporal resolutions. However, due to the no-sample characteristics of social media data, PM$_{2.5}$ exposure risk for older adults is absent. We proposed combining social media data and population survey-derived data to map the total PM$_{2.5}$ exposure risk. Hourly exceedance PM$_{2.5}$ exposure risk indicators based on population modeling (HEPE$_{psd}$) and social media data (HEPE$_{sm}$) were developed. Daily accumulative HEPE$_{sm}$ and HEPE$_{psd}$ ranged from 0 to 0.009 and 0 to 0.026, respectively. Three peaks of HEPE$_{sm}$ and HEPE$_{psd}$ were observed at 13:00, 18:00, and 22:00. The peak value of HEPE$_{sm}$ increased with time, which exhibited a reverse trend to HEPE$_{psd}$. The spatial center of HEPE$_{sm}$ moved from the northwest of the study area to the center. The spatial center of HEPE$_{psd}$ moved from the northwest of the study area to the southeast of the study area. The expansion area of HEPE$_{sm}$ was nearly 1.5 times larger than that of HEPE$_{psd}$. The expansion areas of HEPE$_{sm}$ and HEPE$_{psd}$ aggregated in the old downtown, in which the contribution of HEPE$_{psd}$ was greater than 90%. Thus, this study introduced various source data to build an easier and reliable method to map total exceedance PM$_{2.5}$ exposure risk. Consequently, exposure risk results provided foundations to develop PM$_{2.5}$ pollution mitigation strategies as well as scientific supports for sustainability and eco-health achievement.

Plain Language Summary  PM$_{2.5}$ is reported as the first concerned air pollution materials, which caused more than 1 million mortalities directly or indirectly. How to reduce the adverse effect of PM$_{2.5}$ on public health is crucially urgent. Here, we developed a model considering social media data and population modeling data to map total population PM$_{2.5}$ exceedance PM$_{2.5}$ exposure risk. The results indicated that the aged population groups were more vulnerable to air pollutions. PM$_{2.5}$ exposure risk for all population groups had specific spatiotemporal variation patterns, which were useful for mitigation strategy development.

1. Introduction

The adverse effects of PM$_{2.5}$ on public health have become a worldwide concern since the last century (Cohen et al., 2017; Dockery et al., 1993; Kiourmourtzoglou et al., 2015; Lelieveld et al., 2020). As one of the largest developing countries, a nearly twofold increase in the population-weighted PM$_{2.5}$ exposure risk has been observed in China since 1990 (Brauer et al., 2015; Y. Chen et al., 2013; Huang et al., 2014). The projection of PM$_{2.5}$ suggests that under all emission scenarios, the PM$_{2.5}$ concentration continues to increase, which implies that air pollution is a crucial threat to public health (Apte et al., 2015; Jie et al., 2015; Qin et al., 2015). These findings provide vital information on the estimation of PM$_{2.5}$ exposure risk to public health.

In previous studies, daily mean PM$_{2.5}$ concentration and population survey data were applied to construct indicators to calculate the population-weighted PM$_{2.5}$ exposure risk (Franklin et al., 2015; Pascal et al., 2014; Vodonos et al., 2015; Y. Wang et al., 2015). There are two disadvantages to this method. First, the daily mean PM$_{2.5}$ concentration value cannot fully illustrate the hourly variations in PM$_{2.5}$. Moreover, the hourly PM$_{2.5}$ concentration value might be higher than the daily mean PM$_{2.5}$ concentration as well as the PM$_{2.5}$ health guideline threshold set by the World Health Organization, which leads to an underestimation of its adverse effects on public health (Lin et al., 2017, 2018). Second, this method is confined to the low spatiotemporal...
resolution of the population survey data. The precise details of the PM$_{2.5}$ exposure risk are not well described. Therefore, more scientific indicators are necessary to improve the assessment results of the PM$_{2.5}$ exposure risk.

Therefore, social media data have been introduced into related investigations, which refer to the online population footprints collected by smartphones and facilities. With the prevalence of social media, these data are widely applied in population mobility-related investigations, such as urban function zone extractions, urban expansion, and population commuting (Li, Lyu, Huang, et al., 2020; Li, Lyu, Liu et al., 2020; Shelton et al., 2015; Shen & Karimi, 2016; Q. Wang et al., 2018; Ye et al., 2020). Combined with the definition of the exceedance PM$_{2.5}$ exposure risk and social media data, a novel indicator called hourly exceedance PM$_{2.5}$ exposure risk (HEPE) was constructed. A high spatiotemporal resolution of population-weighted PM$_{2.5}$ exposure risk variations can be obtained (Cao et al., 2020, 2021). However, social media data are regarded as nonrepresentative and non-sample data. Social media data are collected from smartphones, which are used less often by older adults. This situation results in uncertainties in the population-weighted PM$_{2.5}$ exposure risk assessment (Song et al., 2019; Yuan et al., 2020). Therefore, the population-weighted PM$_{2.5}$ exposure risk, relying only on social media data or population survey data, cannot fully reflect the total risk. A quantitative assessment of the total population-weighted PM$_{2.5}$ exposure risk at a high spatiotemporal resolution remains unsolved.

Therefore, we plan to assess the total PM$_{2.5}$ exposure risk by combining population survey-related data and social media data using the indicator designed in our latest investigation named the HEPE (Cao et al., 2020). First, we constructed the HEPE using population survey-related data and social media data. Then, the spatiotemporal variations in the HEPE considering different data sources were quantified. Finally, the contribution of the HEPE considering the population survey-related data to the total HEPE was evaluated. The findings from this study provide new insights that can be combined with different data sources to conduct public health-related investigations. By considering different data sources, maps of air pollution for different population groups were obtained. This is a vital foundation for air pollution mitigation strategies.

2. Study Area

The Tianhe District is the economic center of Guangzhou, located between 23.24°N–23.04°N and 113.18°E–113.45°E (Figure 1). Economic development has been clearly observed since 1979. In Guangzhou, the proportion of the gross domestic product in Tianhe increased from less than 10.00% in 1976 to 21.36% in 2020. The significant economic development has generated great demands for energy consumption and individual wealth, which has resulted in an increase in air pollution, such as PM$_{2.5}$. As a central downtown area, mostly indigenous people and immigrants inhabit this area. Most of them were older adults with low education levels. Thus, an awareness of air pollution protection is absent. Therefore, conducting air pollution assessments in this area is essential and urgent.

3. Data and Methodology

3.1. Data

3.1.1. Population Survey-Related Data

The global age-group composition data for 2018 were obtained from WorldPop (https://www.worldpop.org/). The study population was divided into 36 groups, separated by gender, in the age groups of less than 1 year old, 1–20 years old, 20–25 years old, 25–30 years old, 30–35 years old, 35–40 years old, 40–45 years old, 45–50 years old, 50–55 years old, 55–60 years old, 60–65 years old, 65–70 years old, 70–75 years old, 75–80 years old, and older than 80 years old. The spatial resolution was 100 m. To map the accurate and precise population age group data sets, the random forest method was used considering the population census database, land cover, human settlement, and other 300 input data (Stevens et al., 2015; Tatem et al., 2013). For the sake of accuracy and high spatial resolution, these data are widely used in population-based investigations, such as infectious transmission risk assessment, population distribution mapping validations, and specific population group location mapping (Giles et al., 2020; Lai et al., 2020; Leasure et al., 2020; Lloyd
et al., 2020; Ruktanonchai et al., 2020). Furthermore, individuals older than 60 years old, female and male, were included in this investigation as the representatives of the older adult group.

3.1.2. Social Media Data

Tencent user density data, which were collected from the Tencent platform, were used as the social media data in this study. These data were generated when smartphone users activated Tencent applications, such as WeChat (a social chatting application), Tencent QQ (an immediate message application), and Tencent map (a navigation application). Tencent applications cover more than 93% of smartphone users in Guangzhou. The spatiotemporal resolution of this data was 25 m and 1 hr. The prevalence of this data has been widely applied in population clustering pattern recognition analysis (Y. Chen et al., 2019; Niu et al., 2020), urban transportation analysis (Li, Lyu, Huang, et al., 2020; Li, Lyu, Liu et al., 2020), and public health risk assessments (Cao et al., 2020; Zheng et al., 2020).

3.1.3. Air Pollution Monitoring Data

Hourly PM$_{2.5}$ monitoring data were obtained from the Guangdong Environmental Monitoring Platform (http://gdee.gd.gov.cn). The hourly mean PM$_{2.5}$ concentration was collected from 11 stations for May 17, 2019. Data quality was controlled, and the invalid and missing data were deleted.
3.2. Methodology

3.2.1. HEPE Indicator Development

The HEPE developed in our previous investigation was applied in this study (Zheng et al., 2020). This indicator was constructed considering the exceedance PM$_{2.5}$. Two HEPE indicators were calculated based on the Tencent user density data (TUD) and population survey-related data. The calculations are as follows:

$$\text{HEPE}_{\text{sm}} = \frac{\text{NTUD}_i}{\sum_i \text{NTUD}_i} \times 100\% \times \text{EP}_i$$

(1)

$$\text{HEPE}_{\text{psd}} = \frac{\text{NPSD}_i}{\sum_i \text{NPSD}_i} \times 100\% \times \text{EP}_i$$

(2)

where HEPE$_{\text{sm}}$ refers to the HEPE calculation results based on the TUD data, HEPE$_{\text{psd}}$ refers to the HEPE calculation results based on the population modeling data, and EP$_i$ refers to the exceedance PM$_{2.5}$ concentration beyond 25 µg/m$^3$ at grid $i$. Also, 25 µg/m$^3$ has been set as the health safety level guided by the World Health Organization. NTUD$_i$ and NPSD$_i$ were the normalized results of the Tencent user density data and population modeling data, respectively, for the sake of different dimensions.

3.2.2. Spatiotemporal Pattern Detecting of HEPE

Spatiotemporal variations were detected using the standard deviation ellipse method (SDEM). Three dimensions are considered in the SDEM: major extension direction, secondary extension direction, and spatial center. Long and short diameters represent the major extension and secondary extension directions, respectively. The location information of the spatial center denotes the distribution center of the HEPE. The calculation is as follows:

$$M\left(\bar{X}, \bar{Y}\right) = \left(\frac{\sum_i w_i x_i}{\sum_i w_i}, \frac{\sum_i w_i y_i}{\sum_i w_i}\right)$$

(3)

$$\tan \theta = \frac{\left(\sum_i w_i^2 x_i^2 - \sum_i w_i^2 y_i^2\right) + \sqrt{\left(\sum_i w_i^2 x_i^2 - \sum_i w_i^2 y_i^2\right)^2 + 4\left(\sum_i w_i^2 x_i y_i\right)^2}}{2\sum_i w_i^2 x_i y_i}$$

(4)

$$\sigma_x = \sqrt{\frac{\sum_i w_i^2 \left(x_i \cos \theta - y_i \sin \theta\right)^2}{\sum_i w_i^2}}$$

(5)

$$\sigma_y = \sqrt{\frac{\sum_i w_i^2 \left(x_i \sin \theta - y_i \cos \theta\right)^2}{\sum_i w_i^2}}$$

(6)

where $M\left(\bar{X}, \bar{Y}\right)$ represents the coordinate information of spatial center, $\theta$ represents the angle between the long diameter and north direction, $x_i$ and $y_i$ represent the coordinate information of grid $i$, $x'_i$ and $y'_i$ represent the standard deviation of the spatial distance standard between grid $i$ and the spatial center, and $\sigma_x$ and $\sigma_y$ represent the standard deviation along the X axis and Y axis, respectively.

4. Results

4.1. Dynamics of Exceedance PM$_{2.5}$ Exposure Risk Considering Social Media Data and Population Modeling Data

As Figure 2 illustrated, the linear regression model showed increasing trends of temporal variation characteristics of PM$_{2.5}$ concentration at hourly level. The temporal trend trend slope was 0.87, which indicated a
0.87 μg/m³ increasing per hour. The first exceedance PM$_{2.5}$ concentration value was observed at 10:00. The lowest exceedance PM$_{2.5}$ concentration was found at 17:00 with a value of 0.25 μg/m³. The highest exceedance PM$_{2.5}$ concentration was found at 22:00 with a value of 14.09 μg/m³.

Figure 3 illustrates the dynamics of HEPE$_{sm}$ and HEPE$_{psd}$ during the study period. The mean values of HEPE$_{sm}$ and HEPE$_{psd}$ ranged between $1 \times 10^{-8}$ and $2 \times 10^{-3}$ and $1 \times 10^{-8}$ and $1 \times 10^{-9}$, respectively. The peak values of HEPE$_{sm}$ and HEPE$_{psd}$ ranged between $1 \times 10^{-8}$ and $2 \times 10^{-3}$ and $1 \times 10^{-8}$ and $1 \times 10^{-9}$, respectively. Temporal variations in HEPE$_{sm}$ exhibited three peaks at 13:00, 18:00, and 22:00. The peaks of HEPE$_{psd}$ were observed at the same time as that of HEPE$_{sm}$. Although the peaks occurred simultaneously, differences were observed. The peaks of HEPE$_{sm}$ lagged behind the peaks of HEPE$_{psd}$.

4.2. Spatial Patterns of Exceedance PM$_{2.5}$ Exposure Using Different Population Data

Spatial patterns of hourly HEPE$_{sm}$ and HEPE$_{psd}$ were detected using the SDEM (Figure 4). The spatial centers of HEPE$_{sm}$ were observed at 113.33°E and 23.17°N at 10:00. The spatial center of HEPE$_{sm}$ moved to the northeast of the study area at approximately 113.35°E and 23.18°N. Eventually, the spatial center of HEPE$_{sm}$ was at 113.36°E and 23.16°N. Compared with the spatial center of HEPE$_{sm}$, the spatial center of HEPE$_{psd}$ was located south of the spatial center of HEPE$_{sm}$. It was first observed at 113.31°E and 23.13°N at 10:00. Then, it moved to 113.33°E and 23.16°N. The spatial center of HEPE$_{psd}$ was finally observed at 113.34°E and 23.14°N. The movement of the spatial center of HEPE$_{sm}$ and HEPE$_{psd}$ implied that the public PM$_{2.5}$ exposure risk considering the population modeling data was aggregated to the southwest of the public PM$_{2.5}$ exposure risk considering the social media data.

The major extension direction of HEPE$_{sm}$ was first detected northeast-southwest at 10:00 and 19:00, then it turned to approximately north-south. The distance of the long diameter ranged between 1.55 and 10.26 km. The secondary extension direction of HEPE$_{sm}$ was first detected northwest-southeast at 10:00 and 19:00, then it turned to approximately west-east. The distance of the short diameter ranged between 0.91 and 7.29 km. The major extension direction and secondary extension direction of HEPE$_{psd}$ were consistent with those of HEPE$_{sm}$. The spatial distribution of HEPE$_{psd}$ was more aggregated with the long diameter ranging between 1.50 and 10.20 km and the short diameter ranging between 0.39 and 7.09 km.

4.3. Contribution of Exceedance PM$_{2.5}$ Exposure Risk to the Older Population to the Total

The contribution of HEPE$_{psd}$ to the total public exceedance PM$_{2.5}$ exposure risk varied spatiotemporally (Figure 5). The average contribution of HEPE$_{psd}$ to the total HEPE ranged from 70.1% to 95.3%. The temporal variations in the contribution of HEPE$_{psd}$ demonstrated two peaks and two troughs. The peaks were observed at 14:00 and 17:00, and the troughs were observed at 12:00 and 16:00. Although the average contribution of HEPE$_{psd}$ was high, the standard deviation was significant. The minimum contribution of HEPE$_{psd}$ ranged from 10.0% to 57.2%. The maximum contribution of HEPS$_{psd}$ was 100%. Four hot spots of contribution were detected, including Xinghua Township, Linhe Township, Xiancun Township, and Liede Township. Cold spots were detected in Shahe Township, Tangxia Township, and Yuancun Township.
5. Discussion and Conclusion

Previous studies have documented that the continuing increase in PM$_{2.5}$ poses various health threats, such as premature mortality and excess morbidity, which provide significant information for measuring the harmful effects of ambient air pollution (S. Chen et al., 2020; S. Liu et al., 2020; Lubczyńska et al., 2017; Mortamais et al., 2021; Xue et al., 2019; Yang et al., 2020). The scientific assessment of PM$_{2.5}$ exposure risk is the foundation for these investigations. In our latest investigations, we used social media data to propose a novel indicator named HEPE to provide the significant spatiotemporal characteristics of individual PM$_{2.5}$ exposure risk information. However, due to the nonrepresentative and non-sample properties, the HEPE for the older adults groups was absent. The total PM$_{2.5}$ exceedance exposure cannot be fully reflected only relying on social media data. Therefore, we proposed to map the total exceedance PM$_{2.5}$ exposure risk by combining the social media data and population modeling data. The theoretical and management implications are as follows.

5.1. Theoretical Implications

In previous studies, we first developed the indicator HEPE. Compared with previous indicators, such as daily mean PM$_{2.5}$ or daily peak PM$_{2.5}$, the advantage of HEPE was that it could represent the different exposure intensities and durations within one day, even with the same daily mean concentration. One study conducted in the Pearl River Delta demonstrated significant variations in HEPE, ranging between 50 and 110 units, with a similar daily mean PM$_{2.5}$ that was monitored in four tropical cities. This variation was associated with a maximum mortality rate of 4.43% and a minimum mortality increase of 2.86% in different...
Figure 4. Spatial variations of exceedance PM$_{2.5}$ exposure risk assessed considering social media data and population modeling data. Panels (a–n) show a spatial variation of exceedance PM$_{2.5}$ exposure risk assessed considering the social media data, and panels (o–ab) show a spatial variation of exceedance PM$_{2.5}$ exposure risk assessed considering the population modeling data.
cities. Therefore, the implementation of the definition of exceedance PM$_{2.5}$ is helpful in quantifying the high spatiotemporal associations between environmental exposure and public health outcomes.

In this study, another theoretical implication was to couple the multisource data on public health-related topics. Social media data and population modeling data are the newly developed data and the earliest used data in public-related topics, respectively. Owing to their advantages such as high spatiotemporal resolution and high data accuracy, they have been widely used in urban planning and public health-related topics (Grasso et al., 2017; Gu et al., 2016; Jung et al., 2019; X. Liu et al., 2017; Martí et al., 2019; Sun, 2020; Tu et al., 2017). However, the social media and population modeling data were used individually. The combination of these two kinds of data was rarely seen due to the differences in the data source, spatial resolution, information representation method, and information expression contents. By developing the HEPE indicator, we quantified the relative individual exposure risks. Because HEPE results describe the relative exposure risk, the normalized results of social media data and population modeling data avoid the mismatches caused by the differences in data source, spatial resolution, information representation method, and information expression contents. Therefore, this study can help researchers gain insights into theory development for the combination of social media data and traditional population data.

Moreover, the limitation and uncertainty caused by different source data should be addressed. Due to the restriction of Tencent user density data, only May 17, 2019 participated in analysis. The methodology in this study was adaptable for different study areas or periods theoretically. However, considering great changes of population mobility patterns on weekdays or weekends, various temporal scales should be considered in the future, such as daily, weekly, monthly, and seasonal, to map the total population exceedance PM$_{2.5}$ exposure risk comprehensively. Therefore, this could avoid two aspects of uncertainty. The first was caused by the heterogeneity of population mobility at different temporal scales. The second was caused by the variation of PM$_{2.5}$, in case of the influence of unpredictable meteorological events. The other uncertainty was caused by

Figure 5. Spatiotemporal variations of contribution of hourly exceedance PM$_{2.5}$ exposure risk (HEPE)$_{pd}$ to the total exceedance PM$_{2.5}$ exposure risk. Panel (a) shows the temporal variations of contribution of HEPS$_{pd}$ to the total exceedance PM$_{2.5}$ exposure risk, and panel (b) indicates the spatial characteristics of contribution of HEPE$_{pd}$ to the total exceedance PM$_{2.5}$ exposure risk.
the accuracy of aged population group data. The aged population group data were developed based on aged population survey data and residential area data. Aged population survey data were census data, which were relative accurate and precise. However, due to the statistical method and surveyors’ professional skills differences, incidental errors were unavoidable. Moreover, the residential area data were obtained from government or remote sensing data; this data was updated with delays resulting in the system errors of spatial distribution of aged population groups. However, the spatial resolution of this data was 100 m, which was relatively a large spatial scale that smoothed the incidental and system errors. This widely used data proved that local and global accuracy of this data can satisfy the population-related topic investigation.

5.2. Management Implications

The life expectancy with improved air quality has been addressed in previous studies (Qi et al., 2020). When the ambient air pollution guideline of PM$_{2.5}$ from the World Health Organization (25 µg/m$^3$) was applied, compared with the Chinese National Ambient Air Quality Standard (75 µg/m$^3$), 0.14 years of life expectancy is gained. For the older adult groups, increasing PM$_{2.5}$ concentrations are correlated with atherosclerotic plaque systemic oxidative stress and inflammation, which results in a high risk of mortality and morbidity (Brook et al., 2010). Therefore, an exposure risk assessment of the older adult groups and targeting hot spots is crucial for the development of air pollution mitigation strategies.

In this study, the HEPE$_{psd}$ was used to monitor the HEPE for the older adults. We observed a stable spatial distribution area of HEPE$_{psd}$ during the study period. High-value areas of HEPE$_{psd}$ were constricted within a circle with 10.2 km, which were located in the primitive downtown of Tianhe District. Two conclusions can be drawn: First, the PM$_{2.5}$ exposure risk is related to the mobility characteristics of the older adult groups. Compared with HEPE$_{sm}$, the spatial pattern of HEPE$_{psd}$ shrank. HEPE$_{sm}$ represents the group of young people. This group of people had periodic commuting characteristics. In the morning, the trajectory of HEPE$_{sm}$ begins from the home and ends at the workplace. In the afternoon, the trajectory of HEPE$_{sm}$ begins at the workplace and ends at home. This trajectory forms the cross-region results of the HEPE$_{sm}$. In contrast, the trip distance of the older population was constricted around their homes. Second, people in the old city center are exposed to higher air pollution risks, especially for the older adults. As urbanization progresses, high-quality settlement environments, industries, and medical resources aggregate to new urban centers. Due to cheap rent and a low threshold of employment opportunities, the old urban center is experiencing a significant population growth. A large number of older adults reside in this area, generating a large vulnerable population to PM$_{2.5}$. Therefore, the development of a PM$_{2.5}$ exposure risk mitigation strategy for the older adult should consider two key points. First, the development of a PM$_{2.5}$ exposure risk mitigation strategy should focus on small spatial scales. Targeting the older adult groups, green spaces, such as water bodies or plants, should be built within the common mobility distances of the older resident. Second, more effort to reduce the adverse effects of PM$_{2.5}$ on public health should be made targeting the old urban center.

In China, the daily PM$_{2.5}$ concentration threshold was 75 µg/m$^3$, which is three times larger than that guided by the WHO. Relative investigations have demonstrated that stricter ambient quality standards have led to more health benefits. In 2016, the Healthy China 2030 blueprint was released. In this blueprint, life expectancy of 79 years by 2030 is one of the most significant goals. To achieve this goal, strict PM$_{2.5}$ guideline standards and the exceedance effects of PM$_{2.5}$ should be conducted. Our study provides new evidence that the exceedance effects of PM$_{2.5}$ are a significant indicator for assessing the PM$_{2.5}$ exposure risk. We suggest that the findings of this study are helpful for policy-making.

A few limitations of this study must be addressed. Although we mapped the total exceedance PM$_{2.5}$ exposure risk by combining multisource data, HEPE$_{sm}$ and HEPE$_{psd}$ reflected the relative exposure risk rather than the actual exposure risk. TUD represents the relative population density. To map the real total exceedance PM$_{2.5}$ exposure risk, smartphone or social media data that provide counts in real time are urgently needed. Moreover, seasonal variations in PM$_{2.5}$ and the climatic background influence the exposure risk assessment results. In the future, HEPE$_{sm}$ and HEPE$_{psd}$ in different seasons under different climatic conditions should be further investigated.
Conflict of Interest
All authors declare no financial or personal relationships with other people or organizations that can inappropriately influence their work. There is no professional or personal interest of any nature or type in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled, “Mapping total exceedance PM$_{2.5}$ exposure risk by coupling social media data and population modeling data.”

Data Availability Statement
The hourly PM$_{2.5}$ monitoring data can be downloaded from China National Environmental Monitoring Center (http://www.cnemc.cn/). TUD data are obtained from Tencent Internet Corporation. Population modeling data are downloaded from Worldpop (https://www.worldpop.org/). Population modeling data are free of charge.

References
Apte, J. S., Marshall, J. D., Cohen, A. J., & Brauer, M. (2015). Addressing global mortality from ambient PM$_{2.5}$. Environmental Science & Technology, 49, 8057–8066. https://doi.org/10.1021/acs.est.5b02136
Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R. V., Dentener, F., et al. (2015). Ambient air pollution exposure estimation for the global burden of disease 2013. Environmental Science & Technology, 50, 79–81.
Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., et al. (2010). Particulate matter air pollution and cardiovascular disease an update to the scientific statement from the American Heart Association. Circulation, 121, 2331–2378. https://doi.org/10.1161/cir.0b013e318131bbece1
Cao, Z., Gao, F., Li, S., Wu, Z., Guan, W., & Ho, H. C. (2021). Ridership exceedance exposure risk: Novel indicators to assess PM$_{2.5}$ health exposure of bike sharing riders. Environmental Research, 197, 110100. https://doi.org/10.1016/j.envres.2021.110100
Cao, Z., Wu, Z., Li, S., Ma, W., Deng, Y., Sun, H., & Guan, W. (2020). Exploring spatiotemporal variation characteristics of exceedance air pollution risk using social media big data. Environmental Research Letters. https://doi.org/10.1088/1748-9326/abbd62
Chen, S., Li, D., Wu, X., Chen, L., Zhang, B., Tan, Y., et al. (2020). Application of cell-based biological bioassays for health risk assessment of PM$_{2.5}$ exposure in three megacities, China. Environment International, 139. https://doi.org/10.1016/j.envint.2020.105703
Chen, Y., Ebenstein, A., Greenstone, M., & Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from China’s Huai River policy. Proceedings of the National Academy of Sciences of the United States of America, 110, 12936–12941. https://doi.org/10.1073/pnas.1300018110
Chen, Y., Li, X., Liu, X., Zhang, Y., & Huang, M. (2019). Quantifying the teleconnections between local consumption and domestic land uses in China. Landscape and Urban Planning, 187, 60–69. https://doi.org/10.1016/j.landurbplan.2019.03.011
Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., et al. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. The Lancet, 389, 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6
Dockery, D., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., et al. (1993). An association between airborne particles and mortality in six U.S. cities. New England Journal of Medicine, 329, 1753–1759. https://doi.org/10.1056/nejm199312093292401
Franklin, M., Koutrakis, P., & Schwartz, J. (2008). The role of particle composition on the association between PM$_{2.5}$ and mortality. Epidemiology, 19, 680–689. https://doi.org/10.1097/ede.0b013e3181812bb7
Giles, J. R., zu Erbach-Schoenberg, E., Tatem, A. J., Gardner, L., Bjørnstad, O. N., Metcalf, C. J. E., & Wesolowski, A. (2020). The duration of travel impacts the spatial dynamics of infectious diseases. Proceedings of the National Academy of Sciences of the United States of America, 117, 22572–22579. https://doi.org/10.1073/pnas.1922663117
Grasso, V., Crisci, A., Morabito, M., Nesi, P., & Pantaleo, G. (2017). Public crowding of heat waves by social media data. Advances in Science and Research, 14, 217–226. https://doi.org/10.5194/asr-14-217-2017
Gu, Y., Qian, Z., & Chen, F. (2016). From Twitter to detector: Real-time traffic incident detection using social media data. Transportation Research Part C: Emerging Technologies, 67, 321–342. https://doi.org/10.1016/j.trc.2016.02.011
Huang, R. J., Zhang, Y., Bozetti, C., Ho, K.-F., Cao, J.-J., Han, Y., et al. (2014). High secondary aerosol contribution to particulate pollution during haze events in China. Nature, 514, 218–222. https://doi.org/10.1038/nature13774
Jie, W., Ying, X., & Bing, Z. (2015). Projection of PM$_{2.5}$ and ozone concentration changes over the Jing-Jin-Ji region in China. Atmospheric and Oceanic Science Letters, 8, 143–146. https://doi.org/10.1086/674283.15.11447251
Jung, J., Uejo, C., Duclo, C., & Jordan, M. (2019). Using web data to improve surveillance for heat sensitive health outcomes. Environmental Health, 18. https://doi.org/10.1186/s12940-019-0499-x
Kluoumouzitoglou, M.-A., Schwartz, J. D., Weisskopf, M. C., Melly, S. J., Wang, Y., Dominici, F., & Zanobetti, A. (2015). Long-term PM$_{2.5}$ exposure and neurological hospital admissions in the Northeastern United States. Environmental Health Perspectives, 123, 23–29.
Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. E., et al. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature, 585, 410–413. https://doi.org/10.1038/s41586-020-2293-x
Leasure, D. R., Jochem, W. C., Weber, E. M., Seaman, V., & Tatem, A. J. (2020). National population mapping from sparse survey data: A hierarchical Bayesian modeling framework to account for uncertainty. Proceedings of the National Academy of Sciences of the United States of America, 117, 24173–24179. https://doi.org/10.1073/pnas.191050117
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, 367–371. https://doi.org/10.1038/nature15371
Li, S., Liu, D., Huang, G., Zhang, X., Gao, F., Chen, Y., & Liu, X. (2020). Spatially varying impacts of built environment factors on rail transit ridership at station level: A case study in Guangzhou, China. Journal of Transport Geography, 82, 102631. https://doi.org/10.1016/j.jtrangeo.2019.102631
Li, S., Lyu, D., Liu, X., Tan, Z., Gao, F., Huang, G., & Wu, Z. (2020). The varying patterns of rail transit ridership and their relationships with fine-scale built environment factors: Big data analytics from Guangzhou. Cities, 99. https://doi.org/10.1016/j.cities.2019.102580

Lin, H., Ma, W., Qiu, H., Wang, X., Trevathan, E., Yao, Z., et al. (2017). Using daily excessive concentration hours to explore the short-term mortality effects of ambient PM2.5 in Hong Kong. Environmental Pollution, 229, 896–901. https://doi.org/10.1016/j.envpol.2017.07.060

Lin, H., Wang, X., Qian, Z., Guo, S., Yao, Z., Vaughn, M. G., et al. (2018). Daily exceedance concentration hours: A novel indicator to measure acute cardiovascular effects of PM2.5 in six Chinese suburban cities. Environment International, 111, 117–123. https://doi.org/10.1016/j.envint.2017.11.022

Liu, S., Xing, J., Wang, S., Ding, D., Chen, L., & Hao, J. (2020). Revealing the impacts of transboundary pollution on PM2.5-related deaths in China. Environment International, 134. https://doi.org/10.1016/j.envint.2019.105323

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. International Journal of Geographical Information Science, 31, 1675–1696. https://doi.org/10.1080/13658816.2017.1324976

Lloyd, C. T., Sturrock, H. J. W., Leasure, D. R., Jochem, W. C., Lázár, A. N., & Tatem, A. (2020). Using GIS and machine learning to classify residential status of urban buildings in low and middle income settings. Remote Sensing, 12. https://doi.org/10.3390/rs12233847

Lucbeynioka, M. J., Sunyer, J., Tiemeier, H., Porta, D., Kasper-Sonnenberg, M., Jaddoe, V. W. V., et al. (2017). Exposure to elemental composition of outdoor PM2.5 at birth and cognitive and psychomotor function in childhood in four European birth cohorts. Environment International, 109, 170–180.

Martí, P., Serrano-Estrada, L., & Nolasco-Cirugeda, A. (2019). Social Media data: Challenges, opportunities and limitations in urban studies. Computers, Environment and Urban Systems, 74, 161–174. https://doi.org/10.1016/j.compenvurbs.2018.11.001

Mortmamis, M., Gutierrez, L. A., de Hoogh, K., Chen, J., Vienneau, D., Carrière, I., et al. (2021). Long-term exposure to ambient air pollution and risk of dementia: Results of the prospective Three-City Study. Environment International, 148, 106376–106376. https://doi.org/10.1016/j.envint.2020.106376

Niu, T., Chen, Y., & Yuan, Y. (2020). Measuring urban poverty using multi-source data and a random forest algorithm: A case study in Guangzhou. Sustainable Cities and Society, 54. https://doi.org/10.1016/j.scs.2020.102014

Pascal, M., Falq, G., Wagner, V., Chatignoux, E., Corso, M., Blanchard, M., et al. (2014). Short-term impacts of particulate matter (PM2.5, PM10) on mortality in nine French cities. Atmospheric Environment, 85, 175–184. https://doi.org/10.1016/j.atmosenv.2014.06.030

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., et al. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Journal of the American Medical Association, 287, 1132–1141. https://doi.org/10.1001/jama.287.9.1132

Qi, J., Ruan, Z., Qian, Z. M., Yin, P., Yang, Y., Acharya, B. K., et al. (2020). Potential gains in life expectancy by attaining daily ambient fine particulate matter pollution standards in mainland China: A modeling study based on nationwide data. PLoS Medicine, 17. https://doi.org/10.1371/journal.pmed.1003027

Qin, S., Liu, F., Wang, C., Song, Y., & Qu, J. (2015). Spatial-temporal analysis and projection of extreme particulate matter (PM10, PM2.5) levels using association rules: A case study of the Jing-Jin-Ji region, China. Atmospheric Environment, 120, 339–350. https://doi.org/10.1016/j.atmosenv.2015.09.006

Ruktanonchai, N. W., Floyd, J. R., Lai, S., Ruktanonchai, C. W., Sadilek, A., Rente-Lourenco, P., et al. (2020). Assessing the impact of coor

Tu, W., Cao, J., Yue, Y., Shaw, S.-L., Zhou, M., Wang, Z., et al. (2017). Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. International Journal of Geographical Information Science, 31, 2331–2358. https://doi.org/10.1080/13658816.2017.1356464

Vodonom, A., Awad, Y. A., & Schwartz, J. (2018). The concentration-response between long-term PM2.5 exposure and mortality: A meta-regression analysis. Environmental Research, 166, 677–689. https://doi.org/10.1016/j.envres.2018.08.021

Wang, H., Wang, J., Wang, H., & Xu, W. (2020). Measuring urban air quality and air pollution risk using social media big data. Environmental Research Letters, 15, 10–2.5

Yang, M., Guo, Y.-M., Bloom, M. S., Dharmagee, S. C., Morawska, L., Joachim, H., et al. (2020). Is PM2.5 similar to PM2.5? A new insight into the association of PM2.5 and children’s lung function. Environment International, 145. https://doi.org/10.1016/j.envint.2020.106092

Ye, C., Zhang, F., Mu, L., Gao, Y., & Liu, Y. (2020). Urban function recognition by integrating social media and street-level imagery. Environment and Planning B: Urban Analytics and City Science.

Yuan, Y., Lu, Y., Chow, T. E., Ye, C., Alyaqout, A., & Liu, Y. (2020). The missing parts from social media–enabled smart cities: Who, where, when, and what? Annals of the Association of American Geographers, 110, 462–475. https://doi.org/10.1080/24694452.2019.1631144

Zhang, C., Wu, Z., Li, S., Ma, W., Deng, Y., Sun, H., et al. (2020). Exploring spatiotemporal variation characteristics of exceedance air pollution risk using social media big data. Environmental Research Letters.