



Who are more exposed to PM2.5 pollution: A mobile phone data approach

Huagui Guo^{a,b}, Weifeng Li^{a,b,*}, Fei Yao^c, Jiansheng Wu^{d,e}, Xingang Zhou^f, Yang Yue^g, Anthony G.O. Yeh^{a,b}



^a Department of Urban Planning and Design, The University of Hong Kong, Hong Kong, China

^b Shenzhen Institute of Research and Innovation, The University of Hong Kong, Shenzhen 518057, PR China

^c School of GeoSciences, The University of Edinburgh, Edinburgh EH9 3FF, United Kingdom

^d Key Laboratory for Urban Habitat Environmental Science and Technology, Shenzhen Graduate School, Peking University, Shenzhen 518055, PR China

^e Key Laboratory for Earth Surface Processes, Ministry of Education, College of Urban and Environmental Sciences, Peking University, Beijing 100871, PR China

^f College of Architecture and Urban Planning, Tongji University, Shanghai 200092, PR China

^g Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518052, PR China

ARTICLE INFO

Handling Editor: Xavier Querol

Keywords:

Mobile phone location data
PM2.5 exposure
Economic status
Multi-temporal scales

ABSTRACT

Background: Few studies have examined exposure disparity to ambient air pollution outside North America and Europe. Moreover, very few studies have investigated exposure disparity in terms of individual-level data or at multi-temporal scales.

Objectives: This work aims to examine the associations between individual- and neighbourhood-level economic statuses and individual exposure to PM2.5 across multi-temporal scales.

Methods: The study population included 742,220 mobile phone users on a weekday in Shenzhen, China. A geo-informed backward propagation neural network model was developed to estimate hourly PM2.5 concentrations by the use of remote sensing and geospatial big data, which were then combined with individual trajectories to estimate individual total exposure during weekdays at multi-temporal scales. Coupling the estimated PM2.5 exposure with housing price, we examined the associations between individual- and neighbourhood-level economic statuses and individual exposure using linear regression and two-level hierarchical linear models. Furthermore, we performed five sensitivity analyses to test the robustness of the two-level effects.

Results: We found positive associations between individual- and neighbourhood-level economic statuses and individual PM2.5 exposure at a daytime, daily, weekly, monthly, seasonal or annual scale. Findings on the effects of the two-level economic statuses were generally robust in the five sensitivity analyses. In particular, despite the insignificant effects observed in three of newly selected time periods in the sensitivity analysis, individual- and neighbourhood-level economic statuses were still positively associated with individual total exposure during each of other newly selected periods (including three other seasons).

Conclusions: There are statistically positive associations of individual PM2.5 exposure with individual- and neighbourhood-level economic statuses. That is, people living in areas with higher residential property prices are more exposed to PM2.5 pollution. Findings emphasize the need for public health intervention and urban planning initiatives targeting socio-economic disparity in ambient air pollution exposure, thus alleviating health disparities across socioeconomic groups.

1. Introduction

Severe air pollution has become a global environmental problem and major risk to human health. Air pollution accelerates health disparities across socioeconomic groups because it may cause a greater suffering among the poor who are more exposed and may be socioeconomically vulnerable to air pollution (Evans and Kantrowitz, 2002;

O'Neill et al., 2003; Sacks et al., 2011). Accordingly, an in-depth understanding of exposure disparity is crucial to assist policy making aiming at attenuating health disparity. However, the associations between socioeconomic status and the exposure to air pollution have yet to be well understood in China.

Literature on the disparity in exposure to ambient air pollution is concentrated mainly in developed countries. Empirically, there remains

* Corresponding author at: Department of Urban Planning and Design, The University of Hong Kong, Pokfulam Road, Hong Kong, China.

E-mail addresses: huaguiguo@hku.hk (H. Guo), wfli@hku.hk (W. Li), Fei.Yao@ed.ac.uk (F. Yao), wujis@pkusz.edu.cn (J. Wu), zdg@tongji.edu.cn (X. Zhou), yueyang@szu.edu.cn (Y. Yue), hdxugoy@hkucc.hku.hk (A.G.O. Yeh).

the debated argument that areas inhabited by groups with low socioeconomic statuses experience high concentrations of air pollution. Most findings from North American studies tend to support this argument (Hajat et al., 2013; Gray, Edwards and Miranda, 2013; Bravo et al., 2016; Collins and Grineski, 2019). With respect to the findings of European studies, they are extremely mixed. A few European studies found the disparity pattern similar to that in North America (Chai et al., 2006; Samoli et al., 2019). Several European studies reported a nonlinear or insignificant association between socioeconomic status and air pollution concentration (Havard et al., 2009; Richardson et al., 2013; Fernández-Somoano and Tardon, 2014). Very few European studies found an opposite association (Buzzelli and Jerrett, 2007; Cesaroni et al., 2010). By contrast, research from developing countries is quite limited, although findings are generally similar to those of North American studies (Fan et al., 2012; Zhao et al., 2018a,b; Huang et al., 2019).

Moreover, more efforts and improvements are required in previous studies to deeply understand exposure disparity. Firstly, very few studies examine exposure disparity in terms of individual-level data. Most studies use the aggregated-level data (Buzzelli et al., 2003; Gray et al., 2013; Bravo et al., 2016; Ouyang et al., 2018). One of the dominant problems in aggregated-level examinations is the modifiable areal unit problem (i.e. MAUP). The MAUP means that the findings may be different if examinations are performed using different geographic units or at different spatial scales (Openshaw, 1984; Fotheringham and Wong, 1991). Consequently, the findings from aggregated-level examinations may be sensitive to the different choices of geographic unit. Such problems can be overcome to a certain extent if exposure disparity is examined using individual-level data. However, despite some efforts (Marshall, 2008; Hajat et al., 2013), studies using individual-level data are still rather limited. Moreover, very few have examined the effects of individual- and neighbourhood-level socioeconomic statuses (Cesaroni et al., 2010; Hajat et al., 2013).

Secondly, aggregated-level examinations encounter a severe problem of the misclassification of exposure to air pollution. Such examinations usually employ the average concentration of air pollution within a fixed geographic area where the residents reside as a proxy of individual exposure. On the one hand, it is problematic to assume that people living in the same spatial unit are exposed to the same level of air pollution, because air pollution concentrations usually vary over space and time within the same area. The missing consideration of air pollution variations can bias the estimate of individual exposure (Yoo et al., 2015; Park and Kwan, 2017). On the other hand, ignoring individual mobility can generate the neighbourhood effect averaging problem (i.e. NEAP) which is induced by the lack consideration of individual mobility and exposure in non-residential contexts (Kwan, 2018b). People move over space and time, especially those who spend a considerable amount of time out of home. Ignoring individual mobility can cause misclassification errors in exposure estimates (Nyhan et al., 2016; Tang et al., 2018; Kwan, 2018b). These two mechanisms can collectively lead to the misclassification of individual exposure to air pollution.

Thirdly, most studies examine exposure disparity at a single long-term scale, while the multi-temporal examinations are rather limited. One of the problems of single-scale examinations is temporal uncertainty. The findings may differ when exposure disparity is examined at different temporal scales. However, most studies focus on the examination at an annual scale (Bravo et al., 2016; Zhao et al., 2018a,b; Collins and Grineski, 2019; Schoolman and Ma, 2012). Thus, it is far from clear whether there exists the disparity in exposure to air pollution at the short-term (e.g. daily) or mid-term scale (e.g. monthly). Moreover, it is still unknown whether the pattern of exposure disparity on an annual basis is the same as that observed at other temporal scales. Multi-temporal examinations may enable researchers to obtain new findings that have not been observed at a single annual scale, thereby creating a more nuanced picture of associations between socioeconomic

status and air pollution exposure. Nonetheless, such multi-temporal examinations are rather limited.

Increasingly available mobile phone data demonstrate considerable potential to alleviate some of the aforementioned limitations. Firstly, the nature of this dataset's large sample size, which to a large extent can be a proxy of an entire population, and broad area coverage can provide the rationale for city-wide examinations on exposure disparity. The significance of the use of large-scale datasets to address long-standing questions, such as socioeconomic disparities in air pollution exposure, has been stressed in many studies (Graham and Shelton, 2013; Kwan, 2018a). Secondly, given that every individual (i.e. phone user) acts as an analysis unit, findings do not encounter the problems of MAUP and ecological fallacy that are derived from the selection of a geographic unit as the analysis unit in most previous studies. This enables the examinations on exposure disparity moving from a home-based to people-based research paradigm, thereby making investigations more comprehensive and scientific. Thirdly, the estimation of air pollution exposure at the individual level contributes to a better consideration of spatiotemporal variations of air pollution and individual trajectory, which can largely alleviate exposure misclassifications. Fourthly, the continuous information on hourly geospatial locations of phone users can benefit the examinations of exposure disparity at multi-temporal scales.

To date, limited effort has been exerted to use large-scale mobile phone data to examine the disparity in exposure to air pollution. In particular, using call detail records (i.e. CDRs) and data detail records (i.e. DDRs) in a mobile phone dataset to examine exposure disparity in Beijing, Xu et al. (2019) suggested that compared with commuters with high economic status, those with low economic status are more exposed to PM2.5 pollution per hour when staying at home during winter, but less exposed to PM2.5 when commuting the same distance. However, studies using mobile phone data to examine the disparity in air pollution exposure are rather limited (Xu et al., 2019), despite related attempts having used such data to assess air pollution exposure (Nazelle et al., 2013; Picornell et al., 2019) and examine the disparity in individual mobility across different socioeconomic groups (Silm and Ahas, 2014; Järv et al., 2015; Xu et al., 2015; Xu et al., 2018). Moreover, CDRs and DDRs data used in previous studies, which record geospatial locations in a non-continuous manner, may not capture the complete individual trajectories and thus may be inadequate to support the comprehensive examinations. Furthermore, to our knowledge, the disparity in exposure to air pollution has seldom been examined at multi-temporal scales.

To fill the gaps above, with mobile phone location data on a weekday in Shenzhen as an example, the present study aims to examine whether there is an economic disparity in exposure to PM2.5 pollution in China where exposure disparity has not been well understood. We developed a geo-informed backward propagation neural network (i.e. Geo-BPNN) model to estimate hourly PM2.5 concentrations at 1 km² spatial resolution in Shenzhen in terms of aerosol optical depth (i.e. AOD), in combination with traffic flow, meteorological variables, land use, elevation and spatiotemporally informative terms. Coupling the estimated PM2.5 concentrations with individual trajectories, individual total exposures at multi-temporal scales (daytime to annual) were estimated by summing up hourly PM2.5 exposures during the corresponding time period. As our mobile phone data source is available only on a weekday, the daytime, day, week, month, season and year that contain (or are within) such a weekday were selected for the multi-temporal examinations. Weekend days and holidays were thereby excluded because of the differences between individuals' weekday and weekend mobility patterns (Dewulf et al., 2016; Siła-Nowicka et al., 2016). Then, in terms of a linear regression model and three two-level hierarchical linear models controlling for location, socioeconomic, built environment and other covariates, we examined the associations between individual- and neighbourhood-level economic statuses and individual total exposure during weekdays at multi-temporal scales. We

further conducted five sensitivity analyses to test the robustness of the findings.

2. Materials and methods

2.1. Research area

The present study covers the main built-up areas in Shenzhen. Shenzhen is located in the southeast of Guangdong Province, China. It is one of the four first-tier metropolises in China, which covers a total area of approximately 1953 km² and a population of around 15 million. We excluded areas in Pingshan and Dapeng because mountains mostly cover these two districts. Hence, areas in the total of 8 districts were selected. They are Futian, Luohu, and Nanshan, which are downtown areas. Baoan, Longhua and Longgang are suburbs, while Guangming and Pingshan are rural areas. The annual mean PM2.5 pollution of seven monitoring stations in Shenzhen were 36.39 µg/m³ and 40.58 µg/m³ in 2012 and 2013, respectively, which were more than three times higher than 12 µg/m³ stated in the air quality guidelines of World Health Organization (World Health Organization, 2006). The maximum hourly PM2.5 concentration in 2012 and 2013 were 229 µg/m³ and 377 µg/m³, respectively. The leading industries in Shenzhen are electronics and new materials, electronic communication equipment, and electronic equipment and chemical industry, which are located in the west, middle and east of Shenzhen, respectively. Shenzhen has a highly developed road network (i.e. highways, primary road and secondary road) with a total length of around 1659 km (Chen et al., 2019). Private cars and taxi occupy 47% and 7% of the daily travels in Shenzhen, respectively (Chen et al., 2019).

2.2. Data

2.2.1. Mobile phone location data

Time-location information of individuals on a weekday (March 23, 2012) was derived from the mobile phone location dataset in Shenzhen, China. This dataset was provided by the largest mobile phone operator in Shenzhen for academic purposes. The use of mobile phone data in the present study has obtained the ethical approval from the Human Research Ethics Committee of The University of Hong Kong (Reference number: EA2003008). Initially, the number of mobile phone users is more than 12.4 million, representing a substantial portion of Shenzhen's population of around 15 million. Each time-location record

in the dataset includes information on: (1) user ID, which has been anonymously processed; (2) date and time; (3) longitude and latitude coordinates of the corresponding mobile phone towers providing the mobile phone services. These records were actively and continuously recorded at around one-hour interval when mobile phones were active. The service area of a mobile phone tower nearest to the user (usually represented by the Thiessen polygon) was used to estimate the user's location. The average service area of mobile phone towers in Shenzhen was 0.28 km² (standard deviation = 0.58 km²).

The dataset we obtained in the present study has been processed. More details of data processing, including the identification of home and workplace of phone users, can refer to Zhou et al. (2018). Briefly, activity locations of home and workplace were identified according to the place-starting time-duration model (Long and Thill, 2015). Home location was identified on the basis of mobile phone location records from 0 am to 6 am with duration of stay at this location of at least 4 hours. Similarly, workplace was identified according to the location records of two periods, namely, (8–12) am and (2–6) pm, with duration of stay at this location of at least 5 hours. Location records outside of home and workplace locations were identified as other activity locations. After processing, the number of mobile phone users with continuous and non-continuous 24-hour location records was nearly 1.3 million. As validated, the spatial distributions of residents and workers derived from mobile phone data can be a good proxy of those from statistical data with R² = 0.81 and 0.69, respectively (Zhou et al., 2018). To fit the research design in the present study properly, we derived the location records of mobile phone users whose 24-hour location records were continuously reported, and subsequently obtained 757143 mobile phone users. Also, because of the availability of socio-economic data, the final number of mobile phone users used in the present study is 742220.

2.2.2. Air pollution exposure

2.2.2.1. Estimating hourly PM2.5 concentrations at 1 km² spatial resolution. Hourly PM2.5 concentrations at 1 km² spatial resolution were estimated from March 9, 2012 to December 31, 2013. Hourly ground-level PM2.5 concentrations observed at seven national-level environmental monitoring stations were obtained from the Human Settlements and Environment Commission of Shenzhen Municipality. Calibrations and quality controls of PM2.5 measurements were performed according to the Chinese environmental protection standards (i.e. GB 3095–2012 and HJ 663–2013). The initial PM2.5

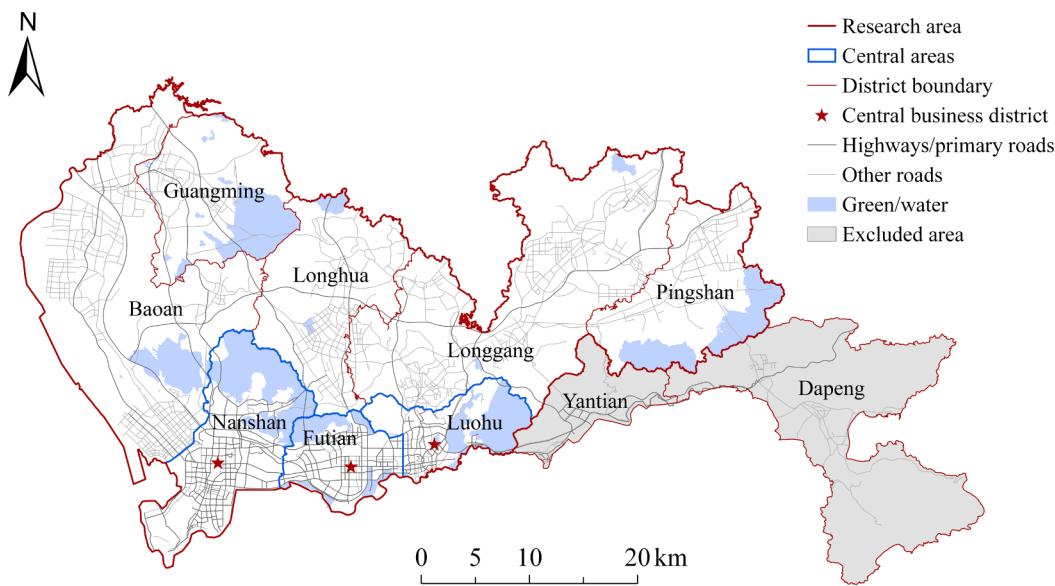


Fig. 1. Research area.

monitoring data in Shenzhen were publicly released at 8 am on March 8, 2012. Hence, we estimated the hourly PM2.5 concentrations from 0 am of March 9, 2012.

A geo-informed backward propagation neural network model (i.e. Geo-BPNN will be discussed in the following parts) was developed to estimate the hourly PM2.5 concentrations. Backward propagation (i.e. BP) is currently a widely used neural network architecture in the field of air pollution modelling (Di, Koutrakis and Schwartz, 2016; Li et al., 2017a,b), because of its capacities in modelling the complex and non-linear associations between predictors (e.g. traffic flow) and outcome variables (e.g. PM2.5 concentrations) (Hecht-Nielsen, 1992). More details of the BP neural network model can refer to Hecht-Nielsen (1992). Satellite-retrieved aerosol optical depth (i.e. AOD) has been widely used in PM2.5 estimates because of its large spatial coverage and high temporal resolution. Hence, we included a new AOD product, namely, MAIAC (i.e. Multi-Angle Implementation of Atmospheric Correction) AOD retrieved from the measurements of the Aqua (overpass at ~1:30 pm local time) and Terra (overpass at ~10:30 am local time) MODIS instruments. MAIAC AOD was used as an important input in our BP model. This type of AOD has fine-scale spatiotemporal resolutions (daily at 1 km²) (Lyapustin et al., 2011a; Lyapustin et al., 2011b).

However, the original MAIAC AOD dataset is incomplete. Like other AOD products, data missing problems are mainly caused by cloud contamination or the existence of bright surfaces (Liu et al., 2009b). To obtain full-coverage AOD, we followed and combined several practices of previous studies. Firstly, for those raster grids with both Terra and Aqua AOD retrievals, we directly averaged them as the daily AOD (Hu et al., 2014; Huang et al., 2018; Wei et al., 2019). Secondly, for those raster grids with either Terra or Aqua AOD retrievals, we fitted linear regression models to obtain the relationship between the Terra (Aqua) and Aqua (Terra) AOD using their match-ups in those raster grids where they were both available. This practice was conducted on a seasonal basis considering the seasonal variations of PM2.5 pollution and AOD (Lv et al., 2016; Ma et al., 2016). The obtained relationships, together with available Terra (Aqua) AOD, were then used to predict the missing Aqua (Terra) AOD, which was further averaged with Terra (Aqua) AOD to obtain the daily AOD.

Thirdly, for those raster grids without either Terra nor Aqua AOD retrievals, a backward propagation neural network (i.e. BPNN) model was developed to estimate AOD. The dependent variable was the daily AOD collected from those raster grids with both Aqua and Terra AOD retrievals. Following prior studies (Xiao et al., 2017; Zhao et al., 2019), the input variables included PM2.5 emissions, daily meteorological variables, elevation, land use factors, geolocation, transportation and population. During model fitting, the statistical indicators of coefficient of determination (i.e. R²) reached 0.89, while the root-mean-square error (i.e. RMSE) was 0.043 (original AOD value: 0–4). The corresponding values during ten-fold cross-validation were 0.85 and 0.049, respectively. This indicates that the model's over-fitting problem is slight. The trained BPNN model, along with the input variables, was used to estimate AOD for those raster grids without any MAIAC AOD retrieval. Finally, the daily AOD derived in the three categories of raster grids described above were merged into a full-coverage AOD product, preparing for the subsequent use in PM2.5 estimates.

Apart from full-coverage AOD, input variables in the neural network for estimating PM2.5 concentrations also included the: (1) hourly traffic flow released by the Shenzhen Transportation Committee (<http://szmap.sutpc.com/roadcong.aspx>); (2) Meteorological variables: u-component of wind 10 m above ground, v-component of wind 10 m above ground, temperature 2 m above ground, planetary boundary layer height, evaporation, high cloud cover, low cloud cover, mean surface downward short-wave radiation flux, surface air pressure, and total precipitation. They were extracted from the ERA5 hourly data on single levels (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>) and ERA Interim, Daily (<https://apps.ecmwf.int/datasets/data/interim-full-daily/>

levType=sfc/) at 6-hour temporal resolution; (3) land use indicators: the proportion of forest land use and impervious surface fraction derived from the dataset of FROM-GLC10-2017v01 (http://data.ess.tsinghua.edu.cn/fromglc10_2017v01.html) and normalized difference vegetation index (i.e. NDVI) derived from the dataset of MOD13Q1 - MODIS/Terra Vegetation Indices 16-Day L3 Global 250 m SIN Grid (<https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD13Q1/>); (4) urban form: building density and road density. They were calculated from the building and road network data provided by the Shenzhen Urban Planning and Land Resource Committee; (5) Elevation derived from the Geospatial Data Cloud (<https://www.gscloud.cn/>); (6) two spatiotemporally informative terms. It has been suggested that incorporating the auto-correlation terms as input variables can significantly improve the accuracy of PM2.5 estimates (Di, Koutrakis and Schwartz, 2016; Li et al., 2017a,b), because there exists the spatiotemporal autocorrelations of PM2.5 pollution itself. Hence, like these studies, we incorporated a spatially informative term to consider the spatial autocorrelation of PM2.5 concentration, which refers to the weighted average of PM2.5 measurements of nearby monitoring stations (reversely weighted by spatial distance). Similarly, we added a temporally informative term, which refers to the weighted average of PM2.5 measurements of prior time at the same grid (reversely weighted by temporal distance), to consider the temporal autocorrelation.

Regarding the performance of the Geo-BPNN model in estimating hourly PM2.5 concentrations, we found high consistency between the predicted PM 2.5 concentrations and ground-level measurements during model fitting ($R^2 = 0.70$ and root-mean-square error (RMSE) = 12.62 $\mu\text{g}/\text{m}^3$). With respect to model validations, we used both the sample-based ten-fold cross-validation and site-based cross-validation methods that have been widely used in AOD-based PM2.5 estimation studies (Xie et al., 2015; Li et al., 2017a,b; Yao et al., 2019). Sample-based validation approach can assess the overall predictive power of our Geo-BPNN model, while site-based method (spatial hold-out cross-validation strategy) can well evaluate the spatial predictive power of our model. The statistical indicators of R^2 and RMSE were equal to 0.69 and 13.25 $\mu\text{g}/\text{m}^3$, respectively, during sample-based ten-fold cross-validation. The corresponding values for site-based cross-validation were slightly worse, with $R^2 = 0.67$ and RMSE = 15.09 $\mu\text{g}/\text{m}^3$, respectively. Then, the trained neural network was used to estimate the hourly PM2.5 concentrations at 1 km² grid cells with full coverage in Shenzhen. Fig. 2 presents the spatial distributions of hourly PM2.5 concentrations on March 23, 2012 (the date of available mobile phone data).

2.2.2.2. Assessing individual total exposure at multi-temporal scales. Estimated hourly PM2.5 concentrations were combined with hourly mobile phone location data to calculate individual exposures at multi-temporal scales. Hourly PM2.5 exposure in a certain activity location of a mobile phone user was extracted from the PM2.5 concentration map at the corresponding hour. The total PM2.5 exposure of a mobile phone user in the course of a day was calculated by summing up the hourly PM2.5 exposures during the day. Similarly, we calculated the total PM2.5 exposures of a mobile phone user during a daytime, week, month, season and year, respectively. We assume that the daily mobility pattern (especially for the weekday) of an individual basically remains constant during a certain time period because of capacity constraints, coupling constraints and authority constraints placed on an individual's spatiotemporal behavior (Ilägerstrand, 1970). As our mobile phone location data source is available only on a weekday (i.e. March 23, 2012), weekend days and holidays were thereby excluded because of the differences between individuals' weekday and weekend mobility patterns as revealed by literature (Liu et al., 2009a; Dewulf et al., 2016; Siła-Nowicka et al., 2016). The daytime, day, week, month, season and year that contain (or are within) the day of the available mobile phone

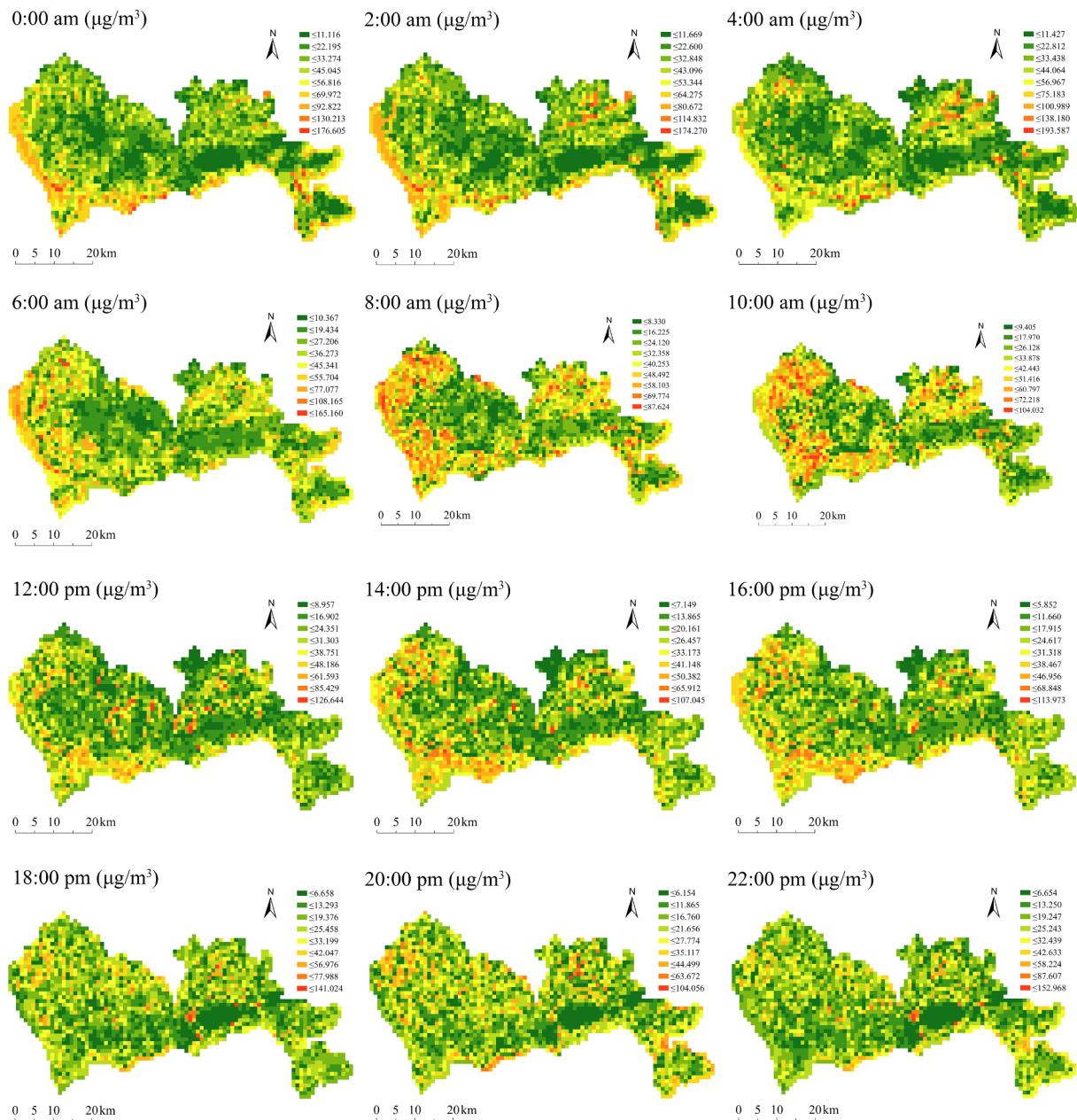


Fig. 2. Spatial distributions of hourly PM2.5 concentrations on March 23, 2012.

data were selected to calculate individual total exposure in the course of the corresponding time period, respectively. Given that our examination might be sensitive to the different selections of time periods, a day, week and month were also random selected according to each season to calculate individual total exposures in the sensitivity analysis owing to the seasonal variations of PM2.5 pollution (Wang et al., 2014; Li et al., 2017a,b). Individual total exposure during each of three other seasons was also calculated to test the robustness of the findings. Fig. 3 (a-f) show the spatial distributions of individual total exposure during weekdays aggregated in the residential committee (i.e. geographic unit in Shenzhen with the finest spatial resolution). Generally, when the distance to central business districts (i.e. CBDs, close to the southernmost part of the research area) increased, there was a trend of decreasing exposure levels. It should be noted that high exposure levels were also concentrated in the northeastern areas in Shenzhen where more industries are located in.

2.2.3. Individual- and neighbourhood-level socioeconomic statuses

At the individual level, we used housing price as a proxy of economic status of mobile phone users. The housing price dataset of 6811 residence communities was collected from the Real Estate Network platform of Fangtianxia (<https://sz.esf.fang.com/>). Fangtianxia is one of the largest companies in China that provides a comprehensive map-based search of housing properties, such as housing price and location. Each housing price record of residence communities contains information, such as the average housing price, total price and geographic coordinates of latitude and longitude. The average housing price of the residence community nearest to the phone user's home location was attributed to the user as a proxy of his (her) individual-level economic status. We assume that people living in residence communities with higher average housing prices are more likely to be richer. Then, similar to the methods of prior studies (Xu et al., 2018; Xu et al., 2019), neighbourhood-level economic status was calculated by averaging the

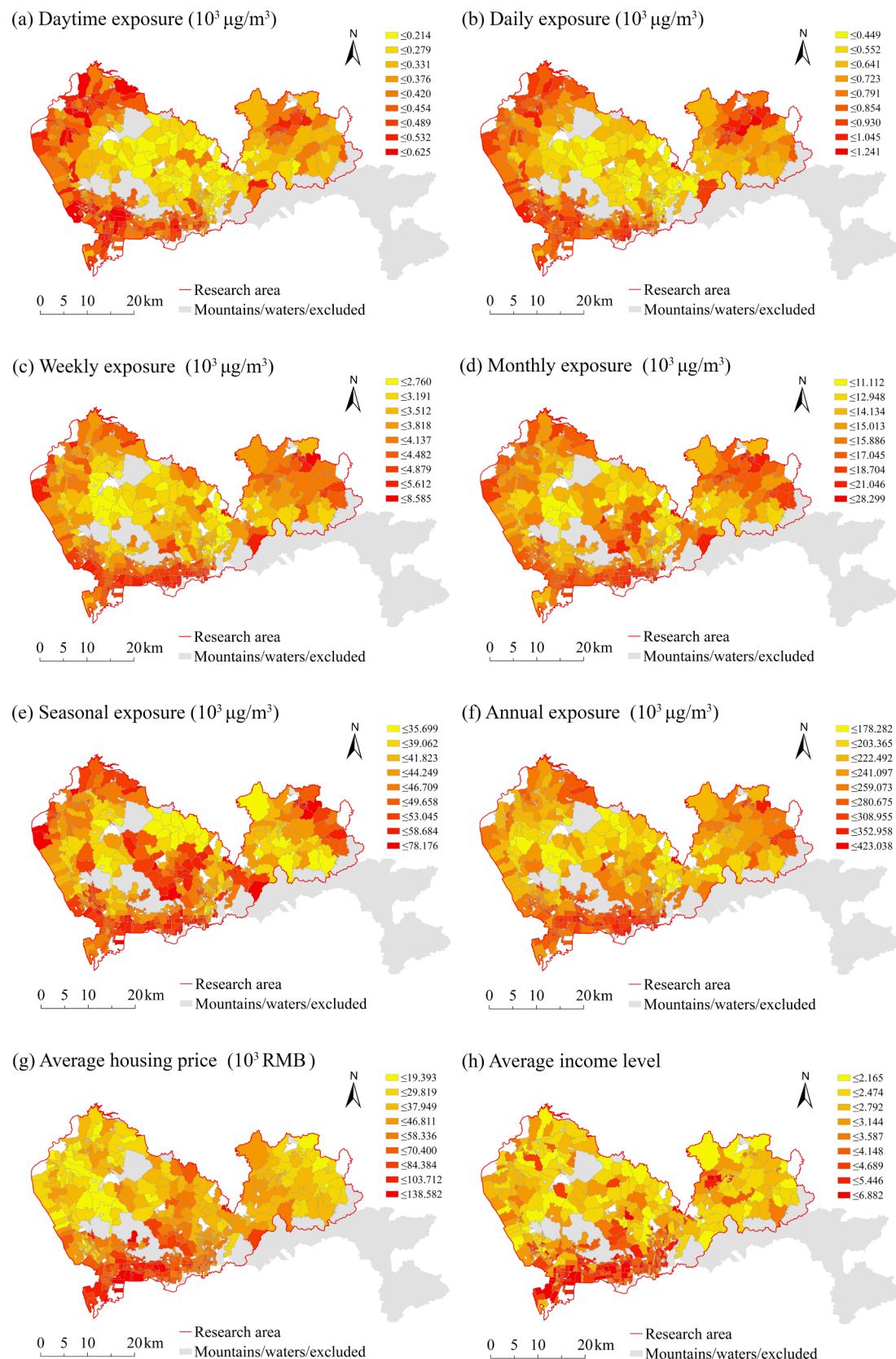


Fig. 3. Spatial distributions of individual total PM_{2.5} exposure at a daytime, daily, weekly, monthly, seasonal or annual scale (a-f) and economic status (g-h) aggregated in residential committee in Shenzhen.

mean housing prices of residence communities located in the same residential committee (i.e. neighbourhood). Fig. 3 (g) shows the spatial distribution of residential-committee-aggregated mean housing price. Generally, with the increase in distance to the city centre (i.e. CBDs), there is a decrease in economic status.

With respect to the socioeconomic data of residential committee (i.e. neighbourhood), they were built in terms of the population-representative dataset of the 2010 travel survey in Shenzhen (Urban Planning Land and Resource Commission of Shenzhen, 2011). This travel survey collected the travel diary and socioeconomic characteristics of more than 220,000 respondents in Shenzhen (around 2% sampling rate). In this dataset, residential committee is the finest geographic unit that can be linked to each respondent. Therefore, we calculated the socioeconomic characteristics of residential committee, that is, the average household income, the average age, the average transport evaluation, sex (i.e. proportion of males), job (i.e. proportion of construction workers) and housing types (i.e. proportions of residents with rental and self-purchasing housing). Notably, the average household income was calculated by averaging the household income levels amongst individuals (respondents) living in the same residential committee. The income level ranged from 1 to 9, and higher values indicated higher income levels. We calculated the average age and transport evaluation (values ranged from 1 to 9) for each residential committee in a similar manner. With respect to the proportion of males, it was calculated using the number of male respondents to divide the number of total respondents in the same residential committee. We calculated the variables of the proportion of construction workers and the proportions of residents with rental and self-purchasing housing for each residential committee in a similar manner. Fig. 3 (h) shows the spatial distribution of income as a proxy of neighbourhood-level economic status. Generally, the pattern of income distribution is similar to that of housing price distribution. That is, the high economic-status group tend to concentrate in the city centre (i.e. close to the

southernmost part of the research area), while the low economic-status group live far from the city centre. Fig. 4 presents the spatial distribution of certain socioeconomic covariates.

2.2.4. Location, built environment and other covariates

The variables of distance to CBD, road density, building density and commuting distance were selected to control the effects of location, built environment and other covariates. Among these, road and building data were acquired from the Shenzhen Urban Planning and Land Resource Committee. Distance to CBD was operationalized as the distance between a user's home location and the location of CBD nearest to the user's home (unit: km). Similarly, commuting distance was operationalized as the distance between a user's home and workplace locations (unit: km). Road density was operationalized as the total length of roads in each residential committee area (unit: km/km²). Building density was operationalized as the percentage of building area in each residential committee area (unit: %).

2.3. Statistical analysis

In the descriptive statistics, boxplots were used to demonstrate differences in individual total PM2.5 exposure across nine economic classes. Notably, the nine division here was used mainly to effectively show the differences among different economic-status groups, which can also be presented in terms of the widely used quantile division. The nine division was conducted according to the three-stratum model which identifies three main socioeconomic groups and three sub-groups in each of three main groups (Saunders, 2006). As in prior studies (Leo et al., 2016; Xu et al., 2018), the economic statuses of mobile phone users were firstly sorted in ascending order. Then, the cumulative sum of economic status in each of nine classes was calculated so that the cumulative sum in each class is the same. Compared with the widely used quantile division, social-stratum-model-based division classifies

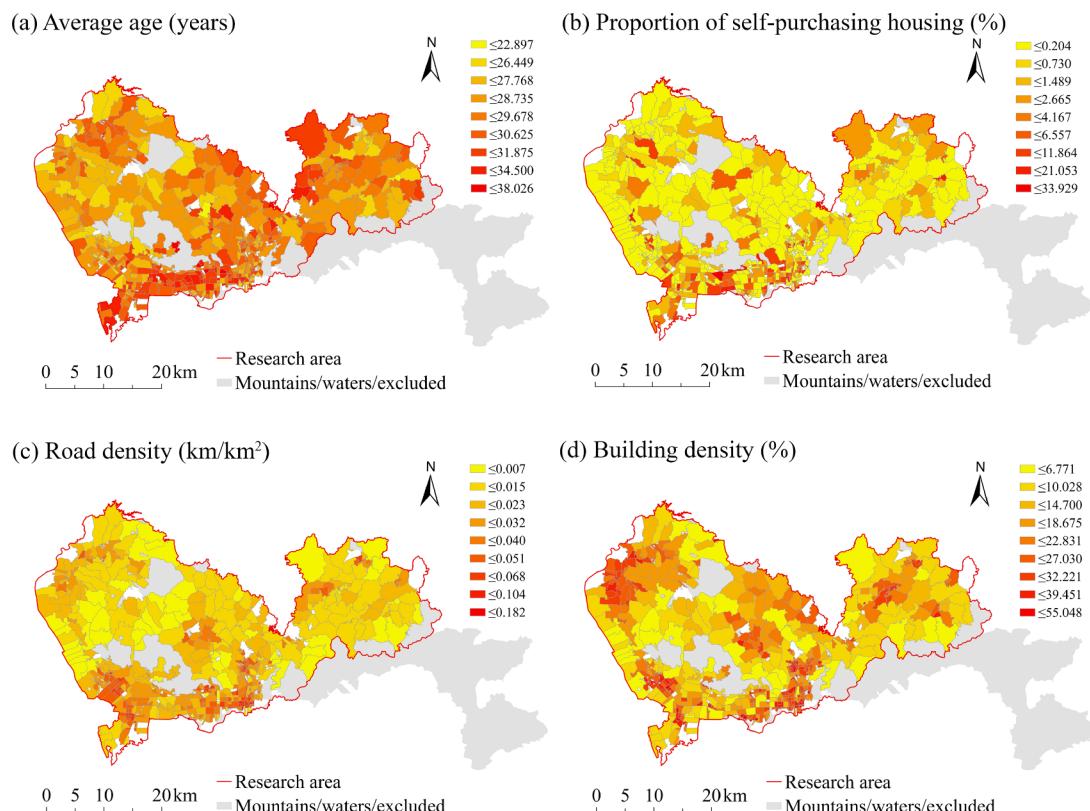


Fig. 4. Spatial distributions of socioeconomic indicators of residential committee in Shenzhen.

mobile phone users into classes with decreasing sizes, which thereby can more effectively exhibit the differences in PM2.5 exposures among economic classes.

Subsequently, we examined the short-term, mid-term and long-term associations of individual total PM2.5 exposure with individual- and neighbourhood-level economic statuses using four statistical models. Model 1 (i.e. a linear regression model) only included individual-level economic status to examine whether there is a significant effect of this variable. In model 2, because of the nested data structure in the present study, we built the first two-level hierarchical linear model (i.e. HLM1) to examine whether the effect of individual-level economic status remains significant. Model 2 adjusted for the socioeconomic indicators including the identity of urban-village resident, location factor including distance to CBD (i.e., central business district), built environment characteristics including road and building densities, and other covariates including commuting distance and transport evaluation.

Model 3 (i.e. HLM2) further added the variable of neighbourhood-level economic status to examine whether there is a significant association between neighbourhood-level economic status and individual exposure. Model 4 (i.e. HLM3) further controlled for neighbourhood-level socioeconomic indicators, that is, age, proportion of males, proportion of construction workers, proportions of residents with rental and self-purchasing housing. The selection of covariates adjusted was based on the comprehensive considerations of univariate correlation strength, covariate collinearity and literatures on exposure disparity (Gee and Payne-Sturges, 2004; Mohai, Pellow and Roberts, 2009; Hajat et al., 2013; Hajat, Hsia and O'Neill, 2015).

Then, we performed five sensitivity analyses. Firstly, we investigated whether the findings on the association between neighbourhood-level economic status and individual exposure were robust when individual covariates were not controlled in the regression models. Secondly, we examined whether the effects of the two-level economic statuses were still significant in the situation of income as a further proxy of neighbourhood-level economic status. Thirdly, we examined whether there exists economic disparities in PM2.5 exposures when the classic (small-sample) statistics approach was used. The small-sample method, in combination with a big data approach, contributes to a more robust examination of economic disparities in PM2.5 exposures. One solution is to aggregate individual total exposures to the neighborhood to examine the association between economic status and PM2.5 exposure at the neighborhood level.

A second solution involves the random selection of a sample of mobile phone users. Consistent with the sampling rate of the 2010 travel survey in Shenzhen (Urban Planning Land and Resource Commission of Shenzhen, 2011), 2% of mobile phone users were random selected. The effects of individual- and neighbourhood-level economic statuses were examined in terms of the aforementioned Models 1–4. Finally, since the daytime, week, month, season and year that contain (or are within) the day when the mobile phone data is available were selected for the multi-temporal examinations, we further examined whether the effects of the two-level economic statuses are sensitive to the different selections of time periods. More details of the selection of different time periods has been specified in Section 2.2.2.2.

3. Results

3.1. Descriptive analysis

Fig. 5 provides the summary statistics of individual total exposure for each of nine economic-status groups at multi-temporal scales. As stated in the statistical section, the nine division was obtained according to the three-stratum model, which is used mainly to more effectively exhibit differences in PM2.5 exposures across economic-status groups. Generally, with the increase of economic level, there was a slight increase in the median of total PM 2.5 exposure at a daytime, daily, weekly, monthly, seasonal or annual scale. As shown in Fig. 5 (c),

for example, the median of total weekly exposure to PM2.5 pollution for the first and second economic strata was approximately $3.7 \times 10^3 \mu\text{g}/\text{m}^3$, which was lower than around $4.3 \times 10^3 \mu\text{g}/\text{m}^3$ for the eighth and ninth economic strata. Similarly, the median of total annual exposure for the two lowest economic-status groups was approximately $2.2 \times 10^5 \mu\text{g}/\text{m}^3$, which was lower than around $2.7 \times 10^5 \mu\text{g}/\text{m}^3$ for the two highest economic-status groups (Fig. 5 (f)).

3.2. Short-term examination of exposure disparity: Daytime and daily scales

Tables 1 and 2 present the results of the short-term association between economic status and individual total exposure. Generally, there were positive associations between individual- and neighbourhood-level economic statuses and individual total exposure at a daytime or daily scale. A 10^3 RMB increase in housing price as a proxy of individual-level economic status was positively associated with a $0.932 \mu\text{g}/\text{m}^3$ ($P < 0.01$) increase in total PM2.5 exposure during daytime in Model 1. When adjusting for socioeconomic, location, built environment and other covariates in Model 2, individual-level economic status was still positively associated with total daytime exposure to PM2.5 pollution ($\beta = 0.265$, $P < 0.01$). When adding the variable of neighbourhood-level economic status in Model 3, we found that the associations between individual- and neighbourhood-level economic statuses and total daytime exposure were still significant, with the coefficients of 0.264 ($P < 0.01$) and 0.567 ($P < 0.01$), respectively. When further controlling for neighbourhood-level socioeconomic characteristics in Model 4, both individual- and neighbourhood-level economic statuses demonstrated significant effects on total daytime exposure, with the coefficients of 0.264 ($P < 0.01$) and 0.481 ($P < 0.1$), respectively. A similar pattern of results was observed in the examinations at a daily scale (Table 2). That is, both individual- and neighbourhood-level economic statuses were positively associated with total daily exposure in terms of Models 1–4 with different covariate controls (Table 2).

3.3. Mid-term examination of exposure disparity: Weekly and monthly scales

The results of the mid-term examination on exposure disparity are shown in Tables 3 and 4. Overall, the two-level economic statuses were positively associated with individual total weekly or monthly exposure. Regarding the weekly examinations (Table 3), when housing price, as a proxy of individual-level economic status, changed by 10^3 RMB, total weekly exposure significantly changed by $0.937 \mu\text{g}/\text{m}^3$ ($P < 0.01$) in Model 1 and $0.920 \mu\text{g}/\text{m}^3$ ($P < 0.01$) in Model 2 (with covariate adjustment). When adding the variable of neighbourhood-level economic status in Model 3, a significant effect of this variable on total weekly exposure was observed ($\beta = 27.500$, $P < 0.01$) (Table 3). When further adjusting for neighbourhood-level socioeconomic indicators, the significant effects of the two-level economic statuses on individual total weekly exposure remained. With respect to the monthly examinations, there were positive associations between the two-level economic statuses and individual total monthly exposure in models with different specifications (Table 4). In particular, as shown in Table 4, we observed the significant effects of individual- and neighbourhood-level economic statuses on total monthly exposure in Model 4, with the coefficients of 1.777 ($P < 0.01$) and 60.762 ($P < 0.01$), respectively.

3.4. Long-term examination of exposure disparity: Seasonal and annual scales

There were positive correlations between the two-level economic statuses and individual total exposure at a long-term scale. Specifically, a positive effect of individual-level economic status on total seasonal (Table 5) or annual exposure (Table 6) was observed in Models 1 and 2

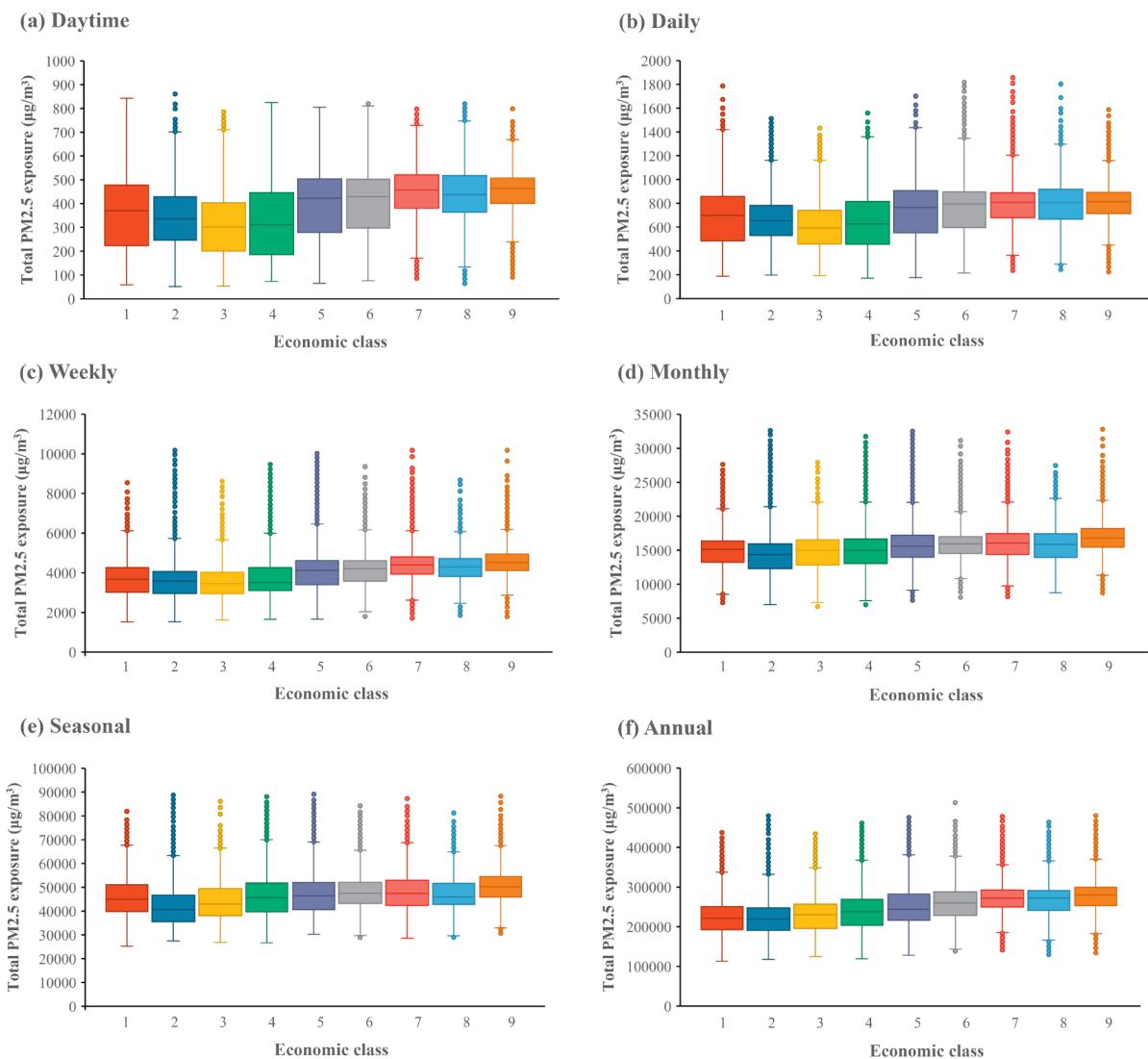


Fig. 5. Boxplots of individual total exposure to PM2.5 pollution for each economic-status group across multi-temporal scales.

Table 1
Daytime examination of economic disparity in PM2.5 exposure.

	Model 1 (OLS)	Model 2 (HLM1)	Model 3 (HLM2)	Model 4 (HLM3)
Economic status (I)	0.932***	0.265***	0.264***	0.264***
Urban-village resident		11.677***	11.677***	11.675***
Distance to CBD		1.220***	1.323***	1.329***
Road density ^a		0.015***	0.013***	0.013***
Building density ^a		0.005	0.025	0.030
Commuting distance		0.525***	0.525***	0.525***
Transport evaluation ^b		-0.076***	-0.065***	-0.067***
Economics status (N)			0.567***	0.481**
Age ^b				0.001
Male% ^b				0.092
Job_construct% ^b				0.038
Housing_rent% ^b				0.066
Housing_selfbuy% ^b				-0.093

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

Table 2
Daily examination of economic disparity in PM2.5 exposure.

	Model 1 (OLS)	Model 2 (HLM1)	Model 3 (HLM2)	Model 4 (HLM3)
Economic status (I)	1.368***	0.109***	0.109***	0.109***
Urban-village resident		6.035***	6.034***	6.030***
Distance to CBD		5.827***	6.106***	6.123***
Road density ^a		0.026***	0.021***	0.021***
Building density ^a		0.127	0.196**	0.209***
Commuting distance		-0.284***	-0.284***	-0.284***
Transport evaluation ^b		-0.166***	-0.130***	-0.132***
Economics status (N)				1.924***
Age ^b				0.005
Male% ^b				0.258**
Job_construct% ^b				0.049
Housing_rent% ^b				0.117
Housing_selfbuy% ^b				0.045

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

Table 3

Weekly examination of economic disparity in PM2.5 exposure.

	Model 1 (OLS)	Model 2 (HLM1)	Model 3 (HLM2)	Model 4 (HLM3)
Economic status (I)	9.347***	0.920***	0.914***	0.913***
Urban-village resident		41.266***	41.272***	41.246***
Distance to CBD		81.732***	82.587***	82.683***
Road density ^a	0.246***	0.175***	0.161***	
Building density ^a	0.824*	1.783***	1.980***	
Commuting distance	4.244***	4.240***	4.241***	
Transport evaluation ^b	-0.945***	-0.427***	-0.374***	
Economics status (N)		27.500***	25.247***	
Age ^b		0.022		
Male% ^b		-0.233		
Job_construct% ^b		-0.069		
Housing_rent% ^b		1.467***		
Housing_selfbuy% ^b		2.162**		

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

(with covariate adjustment). For instance, a unit increase in individual-level economic status (i.e. a 10^3 RMB increase in housing price) was significantly associated with a $43.532 \mu\text{g}/\text{m}^3$ and $0.668 \mu\text{g}/\text{m}^3$ increase in total seasonal exposure in Models 1 and 2, respectively (Table 5). When examining the effect of neighbourhood-level economic status, we

observed its statistically positive association with total seasonal or annual exposure in Models 3 and 4 (Tables 5 and 6). A significant effect of individual-level economic status in these two models was also observed. For example, individual- and neighbourhood-level economic statuses were positively associated with individual total annual exposure in Model 4, with the coefficients of 5.369 ($P < 0.01$) and 670.640 ($P < 0.01$), respectively (Table 6).

3.5. Sensitivity analysis

3.5.1. Independent effect of neighbourhood-level economic status

Fig. 6 (a-b) show the first sensitivity analysis of the independent effect of neighbourhood-level economic status. Overall, the effect of neighbourhood-level economic status was still significant when individual-level covariates were not considered in the regression models. In the univariate analysis, as shown in Fig. 6 (a), there was a positive association between neighbourhood-level economic status and individual total daytime exposure; when further controlling for neighbourhood-level socioeconomic indicators, a significant effect of neighbourhood-level economic status on individual total daytime exposure was still observed. A similar pattern of results was seen in the examinations at a daily, weekly, monthly, seasonal or annual scale (Fig. 6 (a-b)).

3.5.2. Income as a proxy of neighbourhood-level economic status

Tables 7, 8 and Fig. 6 (c-d) show the sensitivity analysis using income as a further proxy of neighbourhood-level economic status.

Table 4

Monthly examination of economic disparity in PM2.5 exposure.

	Model 1 (OLS)	Model 2 (HLM1)	Model 3 (HLM2)	Model 4 (HLM3)
Economic status (I)	20.101***	1.797***	1.777***	1.777***
Urban-village resident		4.806	4.777	4.694
Distance to CBD		200.472***	205.419***	205.873***
Road density ^a	0.596***	0.431***	0.400***	
Building density ^a	2.366**	4.663***	5.085***	
Commuting distance	16.997***	16.983***	16.985***	
Transport evaluation ^b	-2.281***	-1.056***	-0.945***	
Economics status (N)			65.474***	60.762***
Age ^b				0.034
Male% ^b				1.016
Job_construct% ^b				0.509
Housing_rent% ^b				3.395***
Housing_selfbuy% ^b				4.795

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

Table 5

Seasonal examination of economic disparity in PM2.5 exposure.

	Model 1 (OLS)	Model 2 (HLM1)	Model 3 (HLM2)	Model 4 (HLM3)
Economic status (I)	43.532***	0.668*	0.613*	0.612*
Urban-village resident		670.416***	670.377***	670.201***
Distance to CBD		477.082***	491.414***	492.314***
Road density ^a	1.597***	1.181***	1.119***	
Building density ^a	0.924	6.811**	7.855**	
Commuting distance	45.991	45.955***	45.960***	
Transport evaluation ^b	-5.019***	-1.910**	-1.736*	
Economics status (N)		0.668***	166.385***	154.657***
Age ^b				0.144
Male% ^b				5.403
Job_construct% ^b				4.194
Housing_rent% ^b				5.053
Housing_selfbuy% ^b				13.892

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

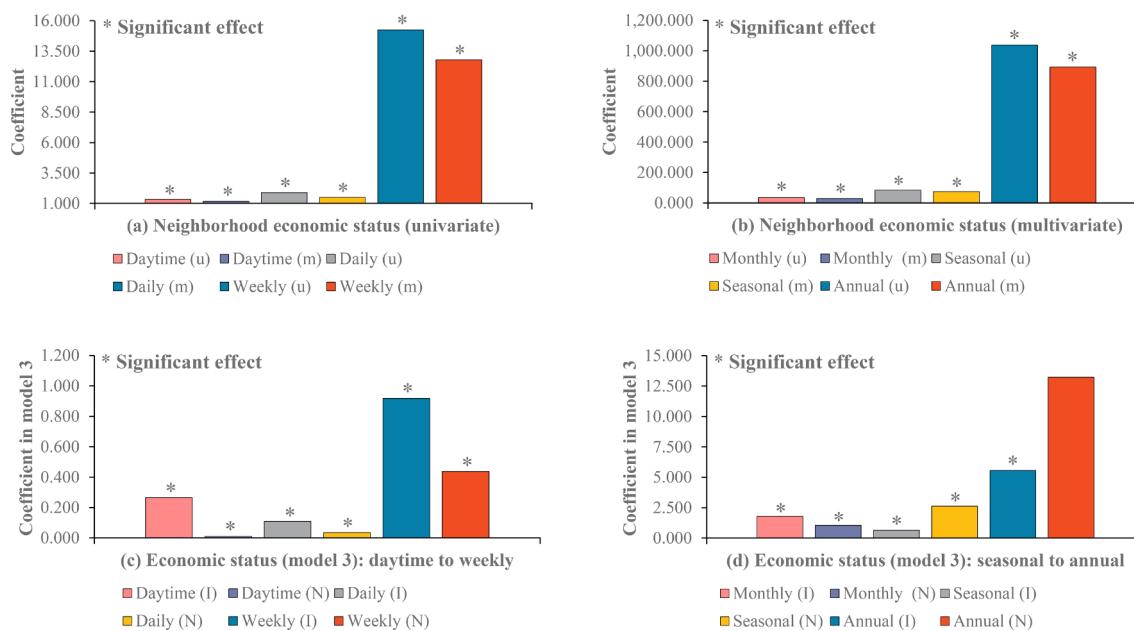
Table 6

Annual examination of economic disparity in PM2.5 exposure.

	Model 1 (OLS)	Model 2 (HLM1)	Model 3 (HLM2)	Model 4 (HLM3)
Economic status (I)	658.889***	5.632***	5.373***	5.369***
Urban-village resident		254.612**	253.941**	252.592**
Distance to CBD		18.761	115.918*	124.701**
Road density ^a		6.445***	4.641***	4.049***
Building density ^a		32.916*	59.197***	67.687***
Commuting distance		213.297***	213.137***	213.189***
Transport evaluation ^b		-37.369***	-23.609***	-22.164***
Economics status (N)			742.202***	670.640***
Age ^b				0.424
Male% ^b				23.283
Job_construct% ^b				46.139
Housing_rent% ^b				52.149***
Housing_selfbuy% ^b				65.774

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value *10⁶, ^b for value = original value *10³.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

**Fig. 6.** Independent effect of neighbourhood-level economic status (a-b) and income as a further proxy of neighbourhood-level economic status (c-d). u and m: univariate and multivariate analysis, respectively. I and N: individual and neighbourhood level, respectively.

Generally, the effects of individual- and neighborhood-level economic statuses were still significant. In Model 3, positive associations existed between the two-level economic statuses and individual total daytime exposure (Fig. 6 (c)). When further controlling for neighbourhood-level socioeconomic characteristics in Model 4, as shown in Table 7, the two-level economic statuses still exerted their significant effects on individual total exposure at a daytime scale. A similar pattern of results was observed in the examinations at each of five other temporal scales (Fig. 6 (c-d), Tables 7 and 8). In particular, as shown in Table 8, both individual- and neighbourhood-level economic statuses were positively associated with individual total annual exposure, with the coefficients of 5.546 ($P < 0.01$) and 11.100 ($P < 0.01$), respectively.

3.5.3. Aggregating individual exposure to the neighbourhood

A significant effect of economic status was observed when individual total exposure was aggregated to the neighbourhood. As shown in the univariate analysis (Fig. 7 (a), Table 9), a positive association existed between economic status and daytime total exposure at the neighbourhood level ($\beta = 0.990$, $P < 0.01$); when further adjusting for socioeconomic, built environment, location and other variables,

despite the slight fluctuation in effect size, we still observed the significant effect of economic status on total PM2.5 exposure at a daytime scale ($\beta = 0.04$, $P < 0.01$). A similar pattern of results was seen in the examinations at each of five other temporal scales (Fig. 7 (b-f), Tables 9 and 10). As shown in Table 10, for example, the association between economic status and total monthly PM2.5 exposure was still significant in the model with covariate control ($\beta = 31.694$, $P < 0.01$).

3.5.4. Random sampling with a small sample size

Fig. 8 presents the sensitivity analysis using a random selected small sample. Overall, the effects of individual- and neighbourhood-level economic statuses were still significant. As shown in Fig. 8 (a), there were positive associations between the two-level economic statuses and individual total daytime exposure. A similar pattern of results was observed in the examinations at a daily, weekly, monthly, seasonal or annual scale (Fig. 8 (b-f)). As shown in Fig. 8 (f), for example, both individual- and neighbourhood-level economic statuses were statistically and positively associated with total annual exposure when a random selected small sample was used.

Table 7

Income as a further proxy of neighbourhood-level economic status: Daytime to weekly.

	Daytime	Daily	Weekly
Economic status (I)	0.265***	0.109***	0.918***
Urban-village resident	11.673***	6.021***	41.172***
Distance to CBD	1.299***	6.047***	82.262***
Road density ^a	0.013***	0.020***	0.160***
Building density ^a	0.025	0.192**	1.663***
Commuting distance	0.525***	-0.284***	4.244***
Transport evaluation ^b	-0.070***	-0.141***	-0.561***
Economics status (N)	0.008*	0.030***	0.379***
Age ^b	0.002	0.006	0.056***
Male% ^b	0.101	0.291**	0.258
Job_construct% ^b	0.010	-0.055	-1.482**
Housing_rent% ^b	0.064	0.105	1.408***
Housing_selfbuy% ^b	-0.078	0.097	2.946**

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

Table 8

Income as a further proxy of neighbourhood-level economic status: Monthly to annual.

	Monthly	Seasonal	Annual
Economic status (I)	1.791***	0.650*	5.546***
Urban-village resident	4.458	669.575***	249.225**
Distance to CBD	203.691***	486.421***	89.593
Road density ^a	0.395***	1.109***	3.846***
Building density ^a	4.315***	5.858	59.779***
Commuting distance	16.995***	45.985***	213.303***
Transport evaluation ^b	-1.383***	-2.863***	-26.353***
Economics status (N)	0.922***	2.328***	11.100***
Age ^b	0.113**	0.345**	1.153
Male% ^b	2.200	8.422	36.460
Job_construct% ^b	-2.897	-4.423	6.662
Housing_rent% ^b	3.224**	4.641	48.743**
Housing_selfbuy% ^b	6.667**	18.645**	86.377*

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

Economic status (I) and Economic status (N): individual- and neighbourhood-level economic status, respectively.

3.5.5. Different selections of time periods

Sensitivity analysis of selecting different time periods is presented in Fig. 9. Generally, individual- and neighbourhood-level economic statuses were positively associated with individual total exposure in most random selected time periods. There were positive associations between the two-level economic statuses and individual total daily exposure in the day selected in spring, summer or autumn (Fig. 9 (a-c)); however, such significant associations were not observed in the day selected in winter (Fig. 9 (d)). Regarding the weekly examinations, the two-level economic statuses were positively associated with individual total weekly exposure in the random selected weeks except the insignificant effect of neighbourhood-level economic status observed in the week selected in autumn (Fig. 9 (h)). Similarly, individual-level economic status had a significant effect in all the selected months (Fig. 9 (i-l)), while a significant effect of neighborhood-level economic status was observed in the month selected in spring, summer or winter (Fig. 9 (e-g)) but not in autumn (Fig. 9 (l)). With respect to the seasonal examinations, there were seasonal variations of individual total exposure, with the largest total exposure occurring in winter, followed by autumn, spring and summer (Fig. S1), which is consistent with the seasonal pattern of PM2.5 pollution in previous studies (Ma et al., 2016; Li et al., 2017a,b). Concerning the patterns of exposure disparities across

seasons, as shown in Fig. 9 (m-p), positive associations existed between the two-level economic statuses and individual total exposure during spring, summer, autumn and winter.

4. Discussions

An in-depth understanding of the disparity in exposure to air pollution is essential to assist policy making to alleviate health disparity. However, whether low socioeconomic groups are disproportionately exposed to air pollution remains uncertain. Moreover, most studies on exposure disparity to air pollution are conducted in developed countries, while research from developing countries is quite limited. Furthermore, very few studies have examined exposure disparity in terms of individual-level data or at multi-temporal scales.

To the best of our knowledge, this is the first study using mobile phone big data (not CDRs) to examine the associations between individual- and neighbourhood-level economic statuses and individual exposure to PM2.5 pollutant at multi-temporal scales in China. Our findings contribute to the literature on exposure disparity in a developing setting where exposure disparity has not been well understood and the pattern of exposure disparity may be so different from that in Western countries (Especially for North America).

Theoretically, socioeconomic disparities in exposure to air pollution are mainly interpreted by facility-siting-based explanations which are popular in the United States, while very few European studies perform the explanations from an air-pollutant-distribution-based aspect. The former explains exposure disparity from two perspectives, that is, market dynamics and racial discrimination (Brulle and Pellow, 2006; Mohai, Pellow and Roberts, 2009). The air-pollutant-distribution-based explanations interpret exposure disparity from multiple perspectives including urban expansion (Cesaroni et al., 2010), local housing policy (Havard et al., 2009) and discrete choice (Padilla et al., 2014). However, such theories, which are developed mostly in America, may not be applicable for other countries to understand and explain exposure disparity. As stressed by the European studies, the determinants of exposure disparity in European cities are concerned with the differences in socioeconomic resources rather than racism which is popular in the United States (Havard et al., 2009; Padilla et al., 2014). Thus, examining exposure disparity in other contexts is necessary to promote theory development. Moreover, current European studies indicate the role of urbanisation, socioeconomic differences and contextual factors in understanding exposure disparity. However, none of these factors share a universal pattern in a country or region (Havard et al., 2009; Cesaroni et al., 2010; Hajat et al., 2013), especially in Chinese cities. Therefore, no single pattern of exposure disparity exists within a region or country. This key message further highlights not only the need for research but also the significance of examining the representative patterns of exposure disparity in other countries.

We found that people living in areas with higher residential property prices are more exposed to PM2.5 pollution in Shenzhen. This finding is not in line with that from most Western studies. With respect to the findings from Western countries, most of them indicated that it is the low socioeconomic-status group who are disproportionately exposed to air pollution (Yanosky, Schwartz and Suh, 2008; Bell and Ebisu, 2012; Hajat et al., 2013). Notably, the underlying mechanisms of exposure disparity are very complex and beyond the scope of the present study. Naturally, it is an outcome-based examination in the present study instead of the process-based studies. The former is performed mainly to examine whether there exists exposure disparity to air pollution, while the latter is conducted mainly to explain how exposure disparity is produced. One potential explanation for the different findings is the difference in the pattern of urbanization-induced residential segregation across different socioeconomic groups. In Western countries, a typical socio-spatial pattern involves the rich (e.g. people with high incomes) who tend to live far from city centre, while the poor having to live in central areas. In contrast, a typical pattern of

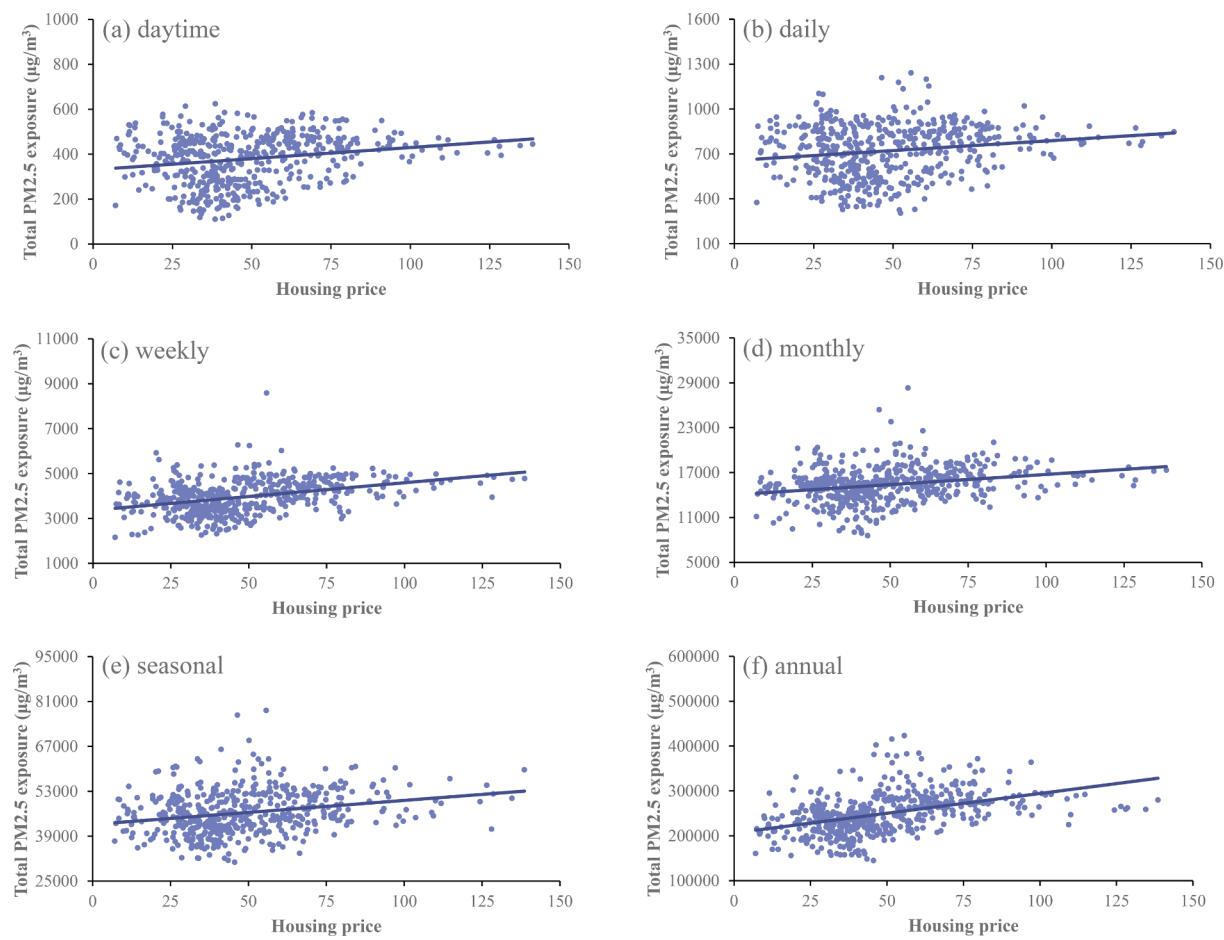


Fig. 7. Scatter plots of neighbourhood-aggregated economic status (i.e. housing price) and total PM2.5 exposure at different temporal scales.

Table 9

Aggregating individual total exposure to the neighbourhood: Daytime to weekly.

	Daytime	Daily	weekly
Model (univariate)			
Economics status	0.990***	1.325***	12.261***
Model (multivariate)			
Economics status	1.106***	1.614***	10.856***
Age ^b	0.001	0.005	0.018
Male% ^b	0.013	0.050**	0.719
Job_construct% ^b	0.097	0.117	0.620
Housing_rent% ^b	-0.047*	0.043	0.939**
Housing_selfbuy% ^b	2.968	5.607	13.482
Distance to CBD	0.015***	0.020***	0.083***
Road density ^a	0.037***	0.198***	1.177***
Building density ^a	-0.064	-0.133**	-0.525***
Transport evaluation ^b	36.460***	36.460***	36.460***

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

residential segregation in relation to urbanization in urban China involves the concentration of the rich in city centres and the resultant displacement of the poor to the urban fringe (Zheng, Fu and Liu, 2006) owing to the concentration of high-quality public facilities and services in central areas. At the same time, rapid urbanisation in urban China usually results in higher population density and higher intensity of human activities in central areas than in outskirts (Zheng and Kahn, 2008), which usually leads to a more severe air pollution situation in the city centre. The uneven distributions of socioeconomic groups and

Table 10

Aggregating individual total exposure to the neighbourhood: Monthly to annual.

	Monthly	Seasonal	Annual
Model (no control)			
Economics status	27.125***	74.919***	881.028***
Model (full control)			
Economics status	31.694***	45.540***	604.773***
Age ^b	0.027	0.121	0.461
Male% ^b	2.118	9.908*	49.218
Job_construct% ^b	1.607	-1.274**	46.844*
Housing_rent% ^b	2.109	4.644	57.240**
Housing_selfbuy% ^b	64.635	-20.011	-248.954
Distance to CBD	0.235***	0.545	3.544
Road density ^a	3.643***	1.452***	66.100***
Building density ^a	-1.252***	-2.879	-23.000***
Transport evaluation ^b	36.460***	36.460***	36.460***

* for $P < 0.1$, ** for $P < 0.05$ and *** for $P < 0.01$. ^a for value = original value $\times 10^6$, ^b for value = original value $\times 10^3$.

air pollution concentrations collectively result in the disparity in exposure to air pollution. That is, the high economic-status group are disproportionately exposed to air pollution (i.e. a typical pattern of exposure disparity in urban China).

A similar pattern of results can be seen in our case study in Shenzhen. People living in areas with higher residential property prices tend to concentrate in city centres (i.e. Futian, Luohu and Nanshan districts), as shown in Fig. 1 and Fig. 3 (g-h), while the poor live in suburban and rural areas (the five other districts). Similar to the spatial

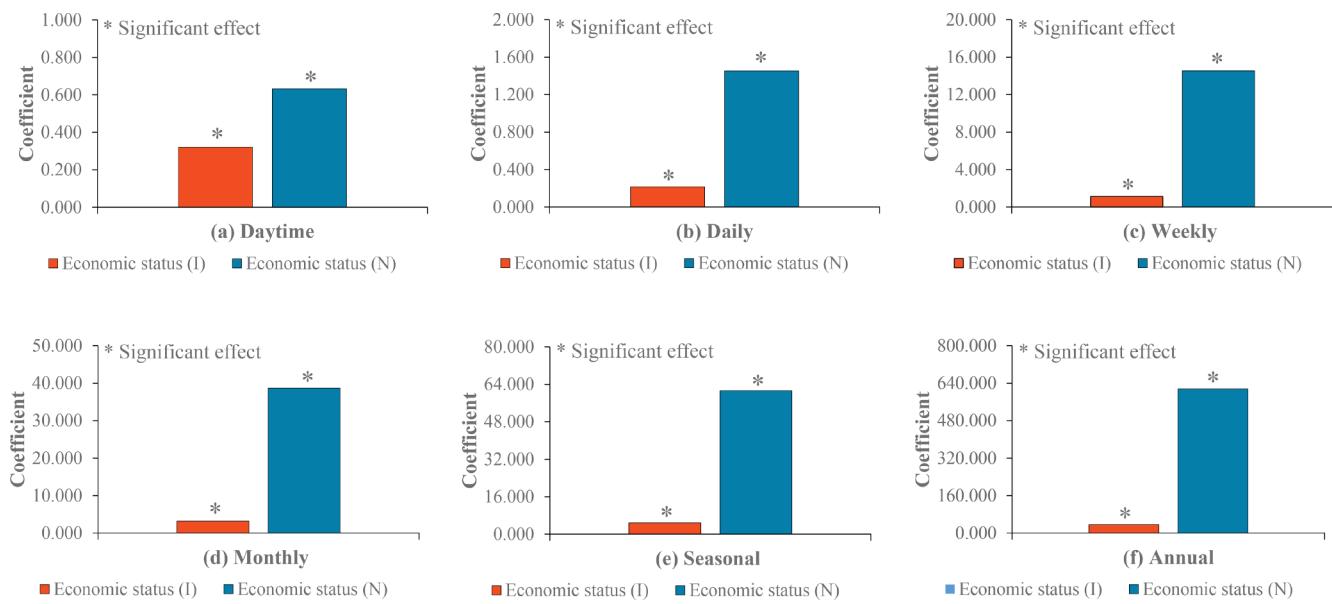


Fig. 8. The associations between individual- and neighbourhood-level economic statuses and individual total exposure at multi-temporal scales when using a random selected small sample. I and N refer to the individual and neighbourhood level, respectively.

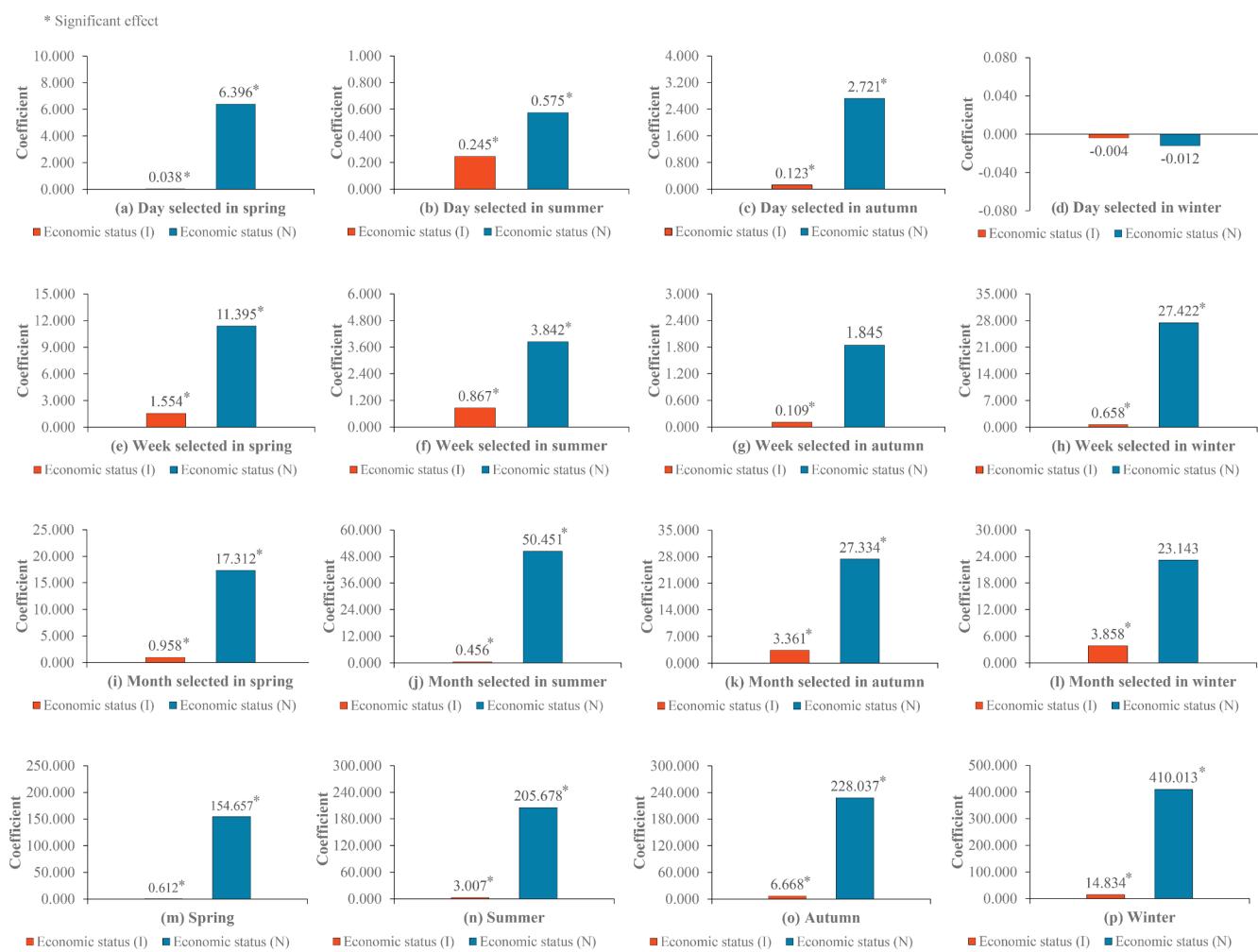


Fig. 9. Random selected time periods (including different seasons) to examine the associations between individual- and neighbourhood-level economic statuses and individual total exposure at multi-temporal scales. I and N: individual and neighbourhood level, respectively.

distribution pattern of economic groups, the air pollution situation in most cases is much more severe in central areas than in rural areas in Shenzhen (Fig. 2). This situation leads to the more exposures for individuals living in or close to city centres than those living in the urban fringe, which can be seen in the distribution of neighbourhood-aggregated PM2.5 total exposures. That is, with the increase in distance to city centres, generally, there is a decrease in total PM2.5 exposure at a daytime, daily, weekly, monthly, seasonal or annual scale (Fig. 3 (a-f)). These two similar patterns of disparities in socioeconomic groups and air pollution concentrations result in exposure disparity to air pollution in Shenzhen. This finding contributes to the literature on exposure disparity that the argument that low socioeconomic-status groups are disproportionately exposed to air pollution may not be applicable in many Chinese cities. In China, one of the dominant patterns of exposure disparity may be the rich who are more exposed to air pollution, which extends previous findings in the literature.

Several strengths in the present study should be noted. Firstly, we demonstrated how the large-scale mobile phone data bear the potential to uncover the patterns of exposure disparity to advance research on exposure disparity shifting from a place (i.e. home)-based to people-based research paradigm. Secondly, we performed the examination at multi-temporal scales (i.e. daytime to annual) instead of the solely long-term (e.g. annual) scale in most previous studies. This allows us to understand exposure differences across different economic-status groups from a multi-layered time perspective to better comprehend exposure disparity. Thirdly, we examined exposure disparity in terms of individual-level data instead of the aggregated-level data that are mostly used in previous studies, which benefits the examination conducted in a more scientific way.

Fourthly, a more accurate estimation of individual exposure benefits our scientific examination of exposure disparity. Individual exposure is determined by not only air pollution variations but also individual mobility, thereby making the accurate assessment of individual exposure much more challenging. Consequently, individual exposure may be biasedly estimated if the spatiotemporal variations of air pollution and individual mobility are not approximately considered. In the present study, the state-of-the-art air pollution modelling approach (i.e. Geo-BPNN) was used to model the hourly PM2.5 concentrations at 1 km² spatial resolution to well capture the spatiotemporal variations of air pollution. Regarding the consideration of individual trajectories, mobile phone big data provides the continuous information on geospatial locations on an hourly basis, which enables us to capture individual mobility effectively. The two high-resolution space-time datasets contribute to a more accurate estimation of individual exposure, thus benefiting a more scientific examination of exposure disparity.

Several limitations and future directions should also be considered. Firstly, as in many studies (Cesaroni et al., 2010; Chaix et al., 2006; Hajat et al., 2013), our examination of the effect of neighbourhood-level economic status on individual exposure may encounter the uncertain geographic context problem (Kwan, 2012). Secondly, as our mobile phone data source is available only on a weekday, our findings may not be suitable on weekend days or holidays because of the differences between individuals' weekday and weekend mobility patterns as revealed in the literature (Liu et al., 2009a; Dewulf et al., 2016; Siłanowicka et al., 2016). If mobile phone data over a long time period (e.g. an entire week or month) is available, it would be much more preferable. Thirdly, there is a privacy concern for studies using individual-level mobile phone data. Technically, a combination of mobile phone location records with other datasets may be able to de-anonymize a mobile phone user (De Montjoye et al., 2013). Our mobile phone data has been anonymously processed and is available only on a weekday, which makes the identification of an individual extremely difficult. Moreover, ethical approval on the use of mobile phone data in the present study has been granted. However, research design and data processing should still be handled carefully to well protect personal privacy in future work.

Fourthly, our examination on exposure disparity is performed in a polycentric and ribbon urban form; thus, future studies should extend the examination to more urban forms, such as the mono-centric spatial structure, thus uncovering more representative patterns of exposure disparity. Fifthly, as an extension, future research can target the exposure disparity at certain activity locations (e.g. workplace) which has been less understood, thus advancing exposure disparity research moving from a home-based to activity-location-based research paradigm. Sixthly, the average neighbourhood income was used in the present study as a further proxy of neighbourhood-level economic status. However, given the impact of investment properties that are used mainly for lease, future studies can combine housing price with income as a measure of economic status to reduce such impact. Lastly, given that the present study is an outcome-based research in nature, more process-based research is highly required to determine why the rich are more exposed to air pollution.

5. Conclusions

There are positive associations between individual- and neighbourhood-level economic statuses and individual exposure to PM2.5 pollution. That is, people living in areas with higher residential property prices are more exposed to PM2.5 pollution. To our knowledge, this is the first study using mobile phone big data (not call detail records of mobile phone data) to examine the disparity in exposure to PM 2.5 pollution across multi-temporal scales in China. The findings point to the need for public health intervention and urban planning initiatives that are targeted to alleviate economic disparity in PM2.5 exposure. Future research, which can use health outcome data at the individual level to examine air pollution-induced health effects, should account for the difference in not only individual-level economic status but also neighbourhood-level economic status.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2019YFB2102000), China; Chan To-Haan Endowed Professorship Fund of the University of Hong Kong, Hong Kong, China; and National Natural Science Foundation of China (41471378, 41971408, 41801147, 71961137003), China.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2020.105821>.

References

- Bell, M.L., Ebisu, K., 2012. Environmental inequality in exposures to airborne particulate matter components in the United States. *Environ. Health Perspect.* 120 (12), 1699–1704.
- Bravo, M.A., Anthopoulos, R., Bell, M.L., Miranda, M.L., 2016. Racial isolation and exposure to airborne particulate matter and ozone in understudied US populations: environmental justice applications of downscaled numerical model output. *Environ. Int.* 92, 247–255.
- Brulle, R.J., Pellow, D.N., 2006. Environmental justice: human health and environmental inequalities. *Annu. Rev. Public Health* 27, 103–124.
- Buzzelli, M., Jerrett, M., 2007. Geographies of susceptibility and exposure in the city: environmental inequity of traffic-related air pollution in Toronto. *Can. J. Regional Sci.* 30 (2).
- Buzzelli, M., Jerrett, M., Burnett, R., Finklestein, N., 2003. Spatiotemporal perspectives on air pollution and environmental justice in Hamilton, Canada, 1985–1996. *Ann. Assoc. Am. Geogr.* 93 (3), 557–573.

- Cesaroni, G., Badaloni, C., Romano, V., Donato, E., Perucci, C.A., Forastiere, F., 2010. Socioeconomic position and health status of people who live near busy roads: the Rome Longitudinal Study (RoLS). *Environ. Health* 9 (1), 41.
- Chaix, B., Gustafsson, S., Jerrett, M., Kristersson, H., Lithman, T., Boalt, Å., Merlo, J., 2006. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *J. Epidemiol. Community Health* 60 (3), 234–241.
- Chen, B.Y., Wang, Y., Wang, D., Lam, W.H., 2019. Understanding travel time uncertainty impacts on the equity of individual accessibility. *Transportation Res. Part D: Transport Environ.* 75, 156–169.
- Collins, T.W., Grineski, S.E., 2019. Environmental Injustice and religion: Outdoor air pollution disparities in metropolitan Salt Lake City, Utah. *Ann. Am. Assoc. Geographers* 109 (5), 1597–1617.
- De Montjoye, Y.A., Hidalgo, C.A., Verleyen, M., Blondel, V.D., 2013. Unique in the crowd: The privacy bounds of human mobility. *Sci. Rep.* 3, 1376.
- De Nazelle, A., Seto, E., Donaire-Gonzalez, D., Mendez, M., Matamala, J., Nieuwenhuijsen, M.J., Jerrett, M., 2013. Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environ. Pollut.* 176, 92–99.
- Dewulf, B., Neutens, T., Lefebvre, W., Seynaeve, G., Vanpoucke, C., Beckx, C., Van de Weghe, N., 2016. Dynamic assessment of exposure to air pollution using mobile phone data. *Int. J. Health Geographics* 15 (1), 14.
- Di, Q., Kourakis, P., Schwartz, J., 2016. A hybrid prediction model for PM_{2.5} mass and components using a chemical transport model and land use regression. *Atmos. Environ.* 131, 390–399.
- Evans, G.W., Kantrowitz, E., 2002. Socioeconomic status and health: the potential role of environmental risk exposure. *Annu. Rev. Public Health* 23 (1), 303–331.
- Fan, X., Lam, K.C., Yu, Q., 2012. Differential exposure of the urban population to vehicular air pollution in Hong Kong. *Sci. Total Environ.* 426, 211–219.
- Fernández-Somoano, A., Tardon, A., 2014. Socioeconomic status and exposure to outdoor NO₂ and benzene in the Asturias INMA birth cohort, Spain. *J. Epidemiol. Community Health* 68 (1), 29–36.
- Fotheringham, A.S., Wong, D.W., 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environ. planning A* 23 (7), 1025–1044.
- Gee, G.C., Payne-Sturges, D.C., 2004. Environmental health disparities: a framework integrating psychosocial and environmental concepts. *Environ. Health Perspect.* 112 (17), 1645–1653.
- Graham, M., Shelton, T., 2013. Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography* 3 (3), 255–261.
- Gray, S.C., Edwards, S.E., Miranda, M.L., 2013. Race, socioeconomic status, and air pollution exposure in North Carolina. *Environ. Res.* 126, 152–158.
- Hajat, A., Diez-Roux, A.V., Adar, S.D., Auchincloss, A.H., Lovasi, G.S., O'Neill, M.S., Kaufman, J.D., 2013. Air pollution and individual and neighbourhood socioeconomic status: evidence from the Multi-Ethnic Study of Atherosclerosis (MESA). *Environ. Health Perspect.* 121 (11–12), 1325–1333.
- Hajat, A., Hsia, C., O'Neill, M.S., 2015. Socioeconomic disparities and air pollution exposure: a global review. *Curr. Environ. Health Rep.* 2 (4), 440–450.
- Havard, S., Deguen, S., Zmirou-Navier, D., Schillinger, C., Bard, D., 2009. Traffic-related air pollution and socioeconomic status: a spatial autocorrelation study to assess environmental equity on a small-area scale. *Epidemiology* 223–230.
- Hecht-Nielsen, R., 1992. Theory of the backpropagation neural network. In: *Neural networks for perception*. Academic Press, pp. 65–93.
- Hu, X., Waller, L.A., Lyapustin, A., Wang, Y., Al-Hamdan, M.Z., Crosson, W.L., Liu, Y., 2014. Estimating ground-level PM_{2.5} concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sens. Environ.* 140, 220–232.
- Huang, G., Zhou, W., Qian, Y., Fisher, B., 2019. Breathing the same air? Socioeconomic disparities in PM_{2.5} exposure and the potential benefits from air filtration. *Sci. Total Environ.* 657, 619–626.
- Huang, K., Xiao, Q., Meng, X., Geng, G., Wang, Y., Lyapustin, A., Liu, Y., 2018. Predicting monthly high-resolution PM_{2.5} concentrations with random forest model in the North China Plain. *Environ. Pollut.* 242, 675–683.
- Ilägrstrand, T., 1970. What about people in regional science?. *Papers of the Regional Science Association* 24.
- Järv, O., Müürisepp, K., Ahas, R., Derudder, B., Witlox, F., 2015. Ethnic differences in activity spaces as a characteristic of segregation: A study based on mobile phone usage in Tallinn, Estonia. *Urban Stud.* 52 (14), 2680–2698.
- Kwan, M.P., 2012. The uncertain geographic context problem. *Ann. Assoc. Am. Geogr.* 102 (5), 958–968.
- Kwan, M.P., 2018a. The limits of the neighbourhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Ann. Am. Assoc. Geographers* 108 (6), 1482–1490.
- Kwan, M.P., 2018b. The neighbourhood effect averaging problem (NEAP): An elusive confounder of the neighbourhood effect. *Int. J. Environ. Res. Public Health* 15 (9), 1841.
- Leo, Y., Fleury, E., Alvarez-Hamelin, J.I., Sarraute, C., Karsai, M., 2016. Socioeconomic correlations and stratification in social-communication networks. *J. R. Soc. Interface* 13 (125), 20160598.
- Li, R., Cui, L., Li, J., Zhao, A., Fu, H., Wu, Y., Chen, J., 2017a. Spatial and temporal variation of particulate matter and gaseous pollutants in China during 2014–2016. *Atmos. Environ.* 161, 235–246.
- Li, T., Shen, H., Yuan, Q., Zhang, X., Zhang, L., 2017b. Estimating ground-level PM_{2.5} by fusing satellite and station observations: a geo-intelligent deep learning approach. *Geophys. Res. Lett.* 44 (23), 11–985.
- Liu, L., Hou, A., Biderman, A., Ratti, C., Chen, J., 2009a. Understanding individual and collective mobility patterns from smart card records: A case study in Shenzhen. In: *2009 12th International IEEE Conference on Intelligent Transportation Systems*. IEEE, pp. 1–6.
- Liu, Y., Paciorek, C.J., Koutrakis, P., 2009b. Estimating regional spatial and temporal variability of PM_{2.5} concentrations using satellite data, meteorology, and land use information. *Environ. Health Perspect.* 117 (6), 886–892.
- Long, Y., Thill, J.C., 2015. Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing. *Comput. Environ. Urban Syst.* 53, 19–35.
- Lv, B., Hu, Y., Chang, H.H., Russell, A.G., Bai, Y., 2016. Improving the accuracy of daily PM_{2.5} distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China. *Environ. Sci. Technol.* 50 (9), 4752–4759.
- Lyapustin, A., Martonchik, J., Wang, Y., Laszlo, I., Korkin, S., 2011a. Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. *J. Geophys. Res. Atmos.* 116(D3).
- Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Reid, J.S., 2011b. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *J. Geophys. Res. Atmos.* 116(D3).
- Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., ... & Liu, Y., 2016. Satellite-based spatiotemporal trends in PM_{2.5} concentrations: China, 2004–2013. *Environ. Health Perspectives*, 124(2), 184–192.
- Marshall, J.D., 2008. Environmental inequality: air pollution exposures in California's South Coast Air Basin. *Atmos. Environ.* 42 (21), 5499–5503.
- Mohai, P., Pellow, D., Roberts, J.T., 2009. Environmental justice. *Annu. Rev. Environ. Resour.* 34, 405–430.
- Nyhan, M., Grauwin, S., Britter, R., Misstear, B., McNabola, A., Laden, F., Ratti, C., 2016. "Exposure track": the impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution. *Environ. Sci. Technol.* 50 (17), 9671–9681.
- O'Neill, M.S., Jerrett, M., Kawachi, I., Levy, J.I., Cohen, A.J., Gouveia, N., ... Workshop on Air Pollution and Socioeconomic Conditions, 2003. Health, wealth, and air pollution: advancing theory and methods. *Environ. Health Perspect.* 111 (16), 1861–1870.
- Openshaw, S., 1984. The Modifiable Areal Unit Problem Concepts and Techniques in Modern Geography 38. Geobooks, Norwich.
- Ouyang, W., Gao, B., Cheng, H., Hao, Z., Wu, N., 2018. Exposure inequality assessment for PM_{2.5} and the potential association with environmental health in Beijing. *Sci. Total Environ.* 635, 769–778.
- Padilla, C.M., Kihal-Talantikite, W., Vieira, V.M., Rossello, P., Le Nir, G., Zmiro-Navier, D., Deguen, S., 2014. Air quality and social deprivation in four French metropolitan areas—A localized spatio-temporal environmental inequality analysis. *Environ. Res.* 134, 315–324.
- 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.
- Picornell, M., Ruiz, T., Borge, R., García-Albertos, P., de la Paz, D., Lumbreiras, J., 2019. Population dynamics based on mobile phone data to improve air pollution exposure assessments. *J. Exposure Sci. Environ. Epidemiol.* 29 (2), 278.
- Richardson, E.A., Pearce, J., Tunstall, H., Mitchell, R., Shortt, N.K., 2013. Particulate air pollution and health inequalities: a Europe-wide ecological analysis. *Int. J. Health Geographics* 12 (1), 34.
- Sacks, J.D., Stanek, L.W., Luben, T.J., Johns, D.O., Buckley, B.J., Brown, J.S., Ross, M., 2011. Particulate matter-induced health effects: who is susceptible? *Environ. Health Perspect.* 119 (4), 446–454.
- Samoli, E., Stergiopoulou, A., Santana, P., Rodopoulou, S., Mitsakou, C., Dimitropoulou, C., Corma, D., 2019. Spatial variability in air pollution exposure in relation to socioeconomic indicators in nine European metropolitan areas: A study on environmental inequality. *Environ. Pollut.* 249, 345–353.
- Saunders, P., 2006. Social Class and Stratification. Routledge.
- Schoolman, E.D., Ma, C., 2012. Migration, class and environmental inequality: Exposure to pollution in China's Jiangsu Province. *Ecol. Econ.* 75, 140–151.
- Sila-Nowicka, K., Vandrol, J., Oshan, T., Long, J.A., Demšar, U., Fotheringham, A.S., 2016. Analysis of human mobility patterns from GPS trajectories and contextual information. *Int. J. Geographical Inf. Sci.* 30 (5), 881–906.
- Silm, S., Ahas, R., 2014. Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data. *Ann. Assoc. Am. Geogr.* 104 (3), 542–559.
- Tang, R., Tian, L., Thach, T.Q., Tsui, T.H., Brauer, M., Lee, M., Barratt, B., 2018. Integrating travel behavior with land use regression to estimate dynamic air pollution exposure in Hong Kong. *Environ. Int.* 113, 100–108.
- Urban Planning Land and Resource Commission of Shenzhen (2011). 2010 travel survey in Shenzhen.
- Wang, Y., Ying, Q., Hu, J., Zhang, H., 2014. Spatial and temporal variations of six criteria air pollutants in 31 provincial capital cities in China during 2013–2014. *Environ. Int.* 73, 413–422.
- Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019. Estimating 1-km-resolution PM_{2.5} concentrations across China using the space-time random forest approach. *Remote Sens. Environ.* 231, 111221.
- World Health Organization, 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment (No. WHO/SDE/PHE/OEH/06.02). World Health Organization, Geneva.
- Xiao, Q., Wang, Y., Chang, H.H., Meng, X., Geng, G., Lyapustin, A., Liu, Y., 2017. Full-coverage high-resolution daily PM_{2.5} estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sens. Environ.* 199, 437–446.
- Xie, Y., Wang, Y., Zhang, K., Dong, W., Lv, B., Bai, Y., 2015. Daily estimation of ground-level PM_{2.5} concentrations over Beijing using 3 km resolution MODIS AOD. *Environ. Sci. Technol.* 49 (20), 12280–12288.
- Xu, Y., Belyi, A., Bojic, I., Ratti, C., 2018. Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Comput. Environ. Urban Syst.* 72, 51–67.

- Xu, Y., Jiang, S., Li, R., Zhang, J., Zhao, J., Abbar, S., González, M.C., 2019. Unraveling environmental justice in ambient PM2.5 exposure in Beijing: A big data approach. *Comput. Environ. Urban Syst.* 75, 12–21.
- Xu, Y., Shaw, S.L., Zhao, Z., Yin, L., Fang, Z., Li, Q., 2015. Understanding aggregate human mobility patterns using passive mobile phone location data: A home-based approach. *Transportation* 42 (4), 625–646.
- Yanosky, J.D., Schwartz, J., Suh, H.H., 2008. Associations between measures of socio-economic position and chronic nitrogen dioxide exposure in Worcester, Massachusetts. *J. Toxicol. Environ. Health, Part A* 71 (24), 1593–1602.
- Yao, F., Wu, J., Li, W., Peng, J., 2019. A spatially structured adaptive two-stage model for retrieving ground-level PM2.5 concentrations from VIIRS AOD in China. *ISPRS J. Photogramm. Remote Sens.* 151, 263–276.
- Yoo, E., Rudra, C., Glasgow, M., Mu, L., 2015. Geospatial estimation of individual exposure to air pollutants: Moving from static monitoring to activity-based dynamic exposure assessment. *Ann. Assoc. Am. Geogr.* 105 (5), 915–926.
- Zhao, C., Liu, Z., Wang, Q., Ban, J., Chen, N.X., Li, T., 2019. High-resolution daily AOD estimated to full coverage using the random forest model approach in the Beijing-Tianjin-Hebei region. *Atmos. Environ.* 203, 70–78.
- Zhao, N., Liu, Y., Vanos, J.K., Cao, G., 2018a. Day-of-week and seasonal patterns of PM2.5 concentrations over the United States: Time-series analyses using the Prophet procedure. *Atmos. Environ.* 192, 116–127.
- Zhao, X., Cheng, H., He, S., Cui, X., Pu, X., Lu, L., 2018b. Spatial associations between social groups and ozone air pollution exposure in the Beijing urban area. *Environ. Res.* 164, 173–183.
- Zheng, S., Kahn, M.E., 2008. Land and residential property markets in a booming economy: New evidence from Beijing. *J. Urban Econ.* 63 (2), 743–757.
- Zheng, S., Fu, Y., Liu, H., 2006. Housing-choice hindrances and urban spatial structure: Evidence from matched location and location-preference data in Chinese cities. *J. Urban Econ.* 60 (3), 535–557.
- Zhou, X., Yeh, A.G., Li, W., Yue, Y., 2018. A commuting spectrum analysis of the job-housing balance and self-containment of employment with mobile phone location big data. *Environ. Planning B: Urban Anal. City Sci.* 45 (3), 434–451.