

“Exposure Track”—The Impact of Mobile-Device-Based Mobility Patterns on Quantifying Population Exposure to Air Pollution

Marguerite Nyhan,^{*,†,‡,⊥} Sebastian Grauwin,[†] Rex Britter,[†] Bruce Misstear,[§] Aonghus McNabola,[§] Francine Laden,[‡] Steven R. H. Barrett,^{||} and Carlo Ratti[†]

[†]Massachusetts Institute of Technology, Senseable City Laboratory, Cambridge, Massachusetts 02139, United States

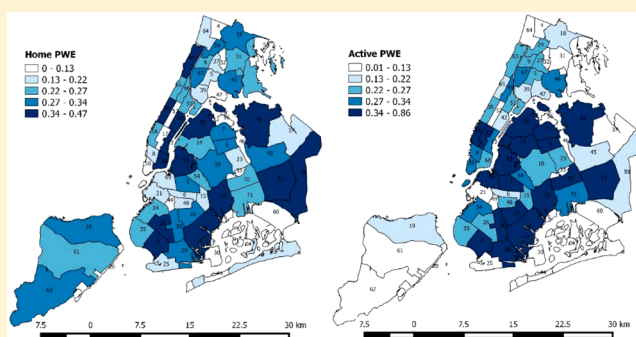
[‡]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts 02215, United States

[§]Department of Civil, Structural & Environmental Engineering, Trinity College Dublin, College Green, Dublin 2, Ireland

^{||}Massachusetts Institute of Technology, Department of Aeronautics & Astronautics, Cambridge, Massachusetts 02139, United States

S Supporting Information

ABSTRACT: Air pollution is now recognized as the world's single largest environmental and human health threat. Indeed, a large number of environmental epidemiological studies have quantified the health impacts of population exposure to pollution. In previous studies, exposure estimates at the population level have not considered spatially- and temporally varying populations present in study regions. Therefore, in the first study of its kind, we use measured population activity patterns representing several million people to evaluate population-weighted exposure to air pollution on a city-wide scale. Mobile and wireless devices yield information about where and when people are present, thus collective activity patterns were determined using counts of connections to the cellular network. Population-weighted exposure to PM_{2.5} in New York City (NYC), herein termed “Active Population Exposure” was evaluated using population activity patterns and spatiotemporal PM_{2.5} concentration levels, and compared to “Home Population Exposure”, which assumed a static population distribution as per Census data. Areas of relatively higher population-weighted exposures were concentrated in different districts within NYC in both scenarios. These were more centralized for the “Active Population Exposure” scenario. Population-weighted exposure computed in each district of NYC for the “Active” scenario were found to be statistically significantly ($p < 0.05$) different to the “Home” scenario for most districts. In investigating the temporal variability of the “Active” population-weighted exposures determined in districts, these were found to be significantly different ($p < 0.05$) during the daytime and the nighttime. Evaluating population exposure to air pollution using spatiotemporal population mobility patterns warrants consideration in future environmental epidemiological studies linking air quality and human health.



1. INTRODUCTION

A large number of epidemiological studies have been conducted, which quantify the health effects of population air pollution exposure. The negative human health effects associated with air pollution exposure are therefore widely documented in the literature.^{1–7} According to the World Health Organization, ambient air pollution contributes to approximately 3.7 million premature deaths annually, with particulate matter (PM) being of considerable concern.⁸ Among the different size classes of PM, PM_{2.5} (aerodynamic diameter < 2.5 μm) shows the strongest and most consistent association with adverse health effects.^{3,9} There is considerable epidemiological evidence to suggest a relationship between acute exposure to PM_{2.5} and increases in all-cause mortality,^{10–19} cardiovascular mortality,^{12,14,15,17,20–22} and respiratory-related mortality.^{12,15,17,20,21,23} Chronic exposure to PM_{2.5} has also been linked to increased incidences of all-cause mortal-

ity.^{4,24–28} Research suggests links between both short and long-term exposures to PM_{2.5} and morbidity,^{9,29} and studies have provided evidence linking exposure to air pollutants such as ultrafine particles, oxides of nitrogen, ozone, carbon monoxide, volatile organic compounds and sulfur dioxide to health effects^{1,30,31} including respiratory illnesses and lung cancer.³²

Previous research linking air pollution exposure to human health effects, at the population level, has been carried out using a variety of methods incorporating time-series studies¹³ and spatial analyses.^{14,27,36} Many studies however are limited by a lack of high resolution daily exposure data and have used measurements from single or a small number of fixed-site

Received: May 12, 2016

Revised: August 10, 2016

Accepted: August 11, 2016

Published: August 12, 2016

monitors to assess air pollution exposure.^{3,18,33–35} Some studies such as by Kloog et al.,²⁹ which evaluated the acute and chronic effects of particles on hospital admissions in New England, addressed this by incorporating a PM_{2.5} model, which accounted for the spatial variability of air pollution levels. Where air pollution modeling efforts have been applied, however, air pollution exposures have not been weighted according to spatially and temporally varying populations present in study regions. In spatial analyses linking population exposure to health impacts,^{27,36} population exposure estimates have not characterized or accounted for spatial and temporal population mobility patterns.

Several studies comparing personal exposure measurements to ambient monitoring at subject's residences have revealed significant variances between concentrations at home addresses and personal exposure concentrations, and large discrepancies between subjects and between studies.^{37,38} Other research has indicated that varying activities and microenvironments are major determinants of personal exposures.^{39–43} Dons et al.⁴¹ showed that commuting accounts for most variability in personal exposures between people exposed to equal concentrations at their residential location. A small number of models have been developed to estimate population exposure to air pollution using varying spatiotemporal population mobility estimates.^{44–48} The Stochastic Human Exposure Model and Dose Simulation (SHEDS-PM) predicts population exposure to PM⁴⁷ and the Air Pollutants Exposure model (APEX), which is part of the EPA's Total Risk Integrated Methodology (TRIM) model framework, predicts human health risks of air pollution exposure and inhalation.⁴⁸ Whereas these models are useful for estimating population exposure to air pollution, the mobility patterns of individuals are either simulated or randomly assigned, rather than directly measured. There is a lack of research whereby human exposure to air pollution has been assessed using location data collected from mobile phones, and where it has, these studies have been limited in terms of very low numbers of participants.^{49–51}

Spatial and temporal human mobility patterns can be characterized based on mobile phone trace data^{52,53} and techniques to extract mobility information from mobile phone traces have progressed in recent times.^{54–56} Data sets that can be used for human mobility analyses include geographically and time-referenced Call Detail Records (CDRs) or counts of connections to the cellular network. These types of data sets may include millions of anonymized records of mobile phone and wireless device usage and can yield detailed information about the activity patterns of large populations, especially where mobile phone and wireless device penetration rates are high relative to the population. Studies have shown that mobile phone trace data can also represent individual mobility patterns,^{55,57,58} and demonstrate advantages over traditional travel surveys used in human mobility studies, which are limited in terms of low response numbers, spatiotemporal scales, and limited update frequencies. To the author's knowledge, a study applying extensive spatiotemporal population mobility estimates from mobile phone data in the assessment of population exposure to environmental pollution over a substantial study domain has not been conducted previously.

This aim of this study was to quantify population-weighted exposure to air pollution by combining extensive population activity patterns and air pollution measurements. This would be considered for a substantial study domain and human population over a significant time period. The specific aim

was to use mobile device based mobility patterns representative of several million people and spatiotemporal PM_{2.5} concentration level estimates to evaluate population-weighted exposure to PM_{2.5} for New York City (NYC) and for 71 districts within the city. It was intended to improve the quantification of the spatial and temporal variations in population exposure to air pollution for potential application in future environmental epidemiological research studies.

2. METHODS

2.1. Study Protocol. Population-weighted PM_{2.5} exposure in the NYC region was examined from April to July 2013. Population-weighted exposure was calculated as a function of air pollution concentration in an area and the proportion of the total population of NYC exposed in that area. The spatial domain studied included the boroughs of Manhattan, Staten Island, Brooklyn, Queens and the Bronx. The total area of NYC is 1214 km². The total population of NYC is approximately 8.5 million people (2010 figure from the US Census Bureau, 2016). 71 separate districts of NYC were studied (see Figure S1) and residential population statistics were attained for each of the districts.⁵⁹ It was assumed that the hourly air quality level and the percentage of the total population present within each district were uniformly distributed. Each district had a geospatial centroid coordinate and hourly PM_{2.5} parameter levels to be inferred or already associated if having an air quality monitoring station located in it. Two scenarios of population-weighted exposure to air pollution were compared. The first was air pollution exposure weighted by population activity counts deciphered using extensive mobile device usage records, herein referred to as "Active Population Exposure". The second was air pollution exposure weighted by assuming people were always located at their home location, using a Census-defined spatial population distribution. This is referred to as "Home Population Exposure". For the "Active" scenario, the population mobility data varied hourly in contrast to the "Home" scenario in which the population was stationary over time.

2.2. Particulate Matter. PM_{2.5} concentration data was obtained from the New York City Community Air Survey (NYCCAS).⁶⁰ The design and implementation of the NYCCAS urban air monitoring program, which was designed to characterize intraurban spatial gradients and complement regulatory monitoring for informing local air quality management, is described in.⁶¹ The NYCCAS monitors PM_{2.5} and other criteria pollutant parameters at 155 locations throughout NYC during each season of the year (as shown in Figure S2). Data was collected over 2 week intervals at each of the 155 distributed locations once per season. Data was also collected at another five reference sites in two week periods year round for temporally adjusting distributed site data. These integrated sampling units provide high-quality air pollutant data, which exhibits significantly more geographic variation than is captured by regulatory monitoring.

Spatial distribution maps of air quality parameter levels were developed using the spatial interpolation technique of inverse-distance-weighting, as this had been used effectively in studies quantifying pollution exposures.⁶² PM_{2.5} values were therefore interpolated from the NYCCAS monitoring stations to each of the 71 district centroids, for each 2-week interval. To obtain daily values for use in our study, the 2 week interval data was adjusted to yield daily variation. This was completed using 24-h concentration measurements obtained from a United States Environmental Protection Agency (U.S. EPA) fixed-site

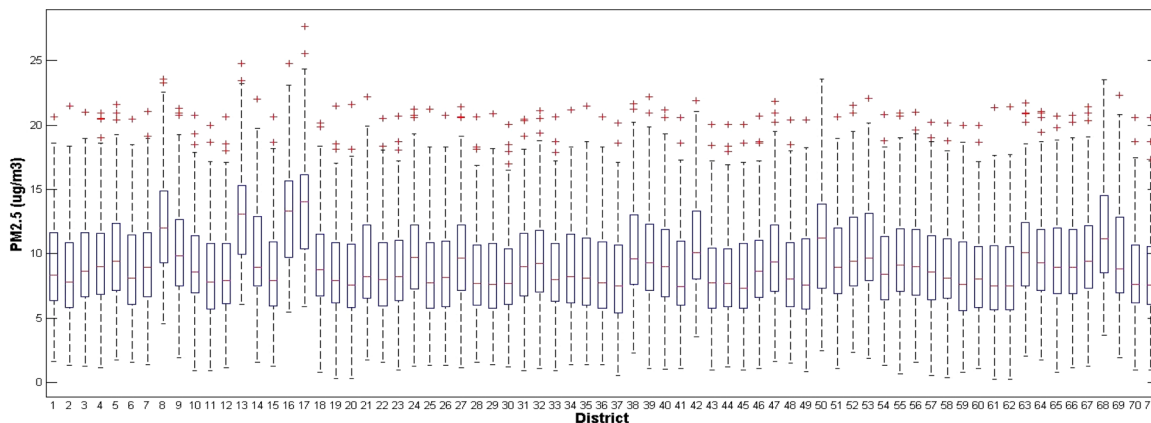


Figure 1. Boxplots of daily PM_{2.5} concentration levels for each district within NYC (*n* = 121 days).

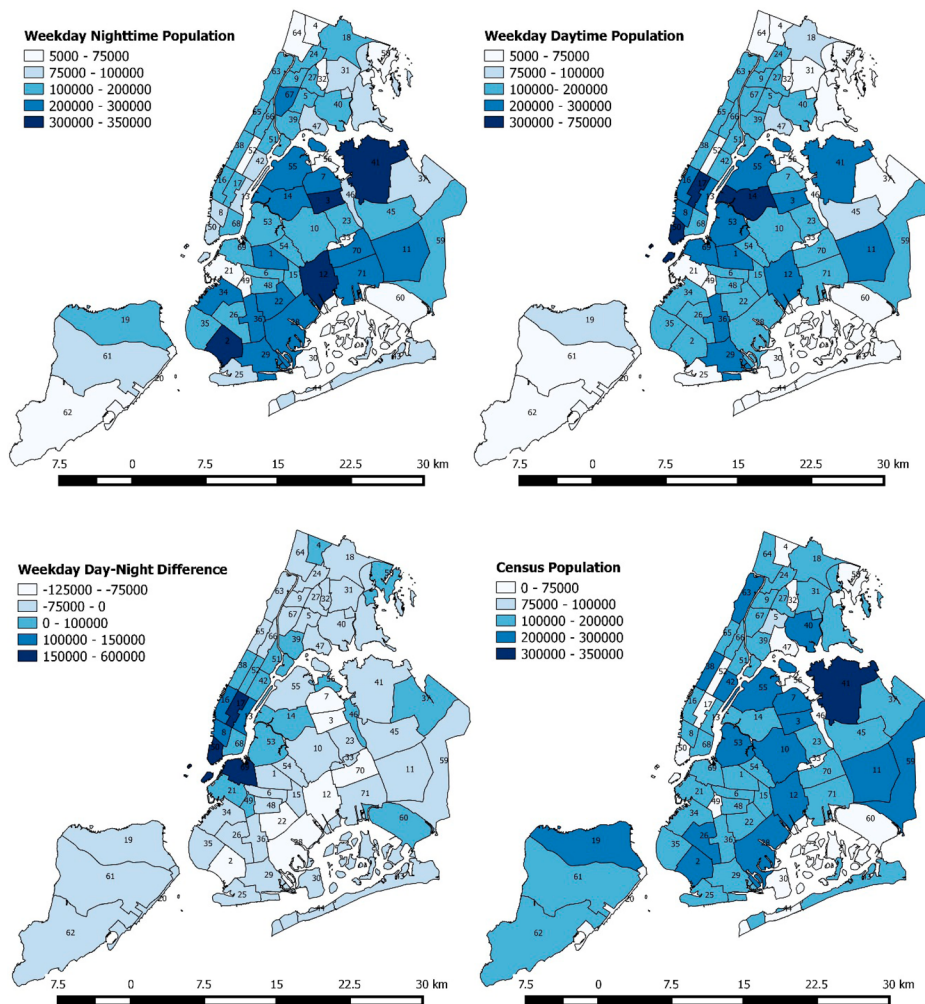


Figure 2. “Active Population” determined for each district of NYC: Mean nighttime population according to the counts of connections to the cellular network (top-left); mean daytime population according to counts of connections to the cellular network (top-right); mean difference between the daytime and nighttime population (lower-left), and the “Home Population” according to the Census (lower-right).

monitoring station centrally located in Manhattan, classified as appropriate for indicating population exposure to air pollution.⁶³ Figure 1 shows boxplots of the resulting daily PM_{2.5} concentration levels for each district within NYC. Details relating to the PM_{2.5} concentrations as measured by the NYCCAS, and further adjusted for daily variability using data from the USEPA, for districts are shown in Figure S3.

2.3. Mobile Device Based Population Activity Patterns. Geographically and time-referenced mobile traffic data were used to quantify hourly percentages of the total population of NYC present in each of the 71 districts of the city throughout the study period. The specific data set used was 3G mobile traffic data accrued from several operators and this included data from all types of mobile devices such as phones

and tablets. The data corresponded to several million subscribers which represented a statistically significant fraction of the total 3G mobile traffic in NYC (precise penetration rates are not given for confidentiality reasons). The data was assumed to be appropriate for extrapolating information about the spatial relative distribution of the entire population in terms of percentage of people that are present in certain areas. Included in the data set were counts related to data communication requests (phone-calls, SMS and passive data-requests). Normalized data which was aggregated at the cell tower level was provided, therefore users were anonymized. The service area of each cell tower which recorded information had a radius of approximately 100 m. To estimate population levels in different areas, the data corresponding to passive data-request activity were utilized. Passive data requests (such as applications running on the background of a phone automatically updating and syncing with the cellular network) track mobile devices without requiring active user interference. The data were spatially aggregated at the districts level (see Figure S1), aggregated hourly and then spatially normalized.

The data set enables the collective tracking of the spatial and temporal locations of the population of NYC. This would subsequently enable the assignment of corresponding air pollution exposures which also vary temporally and spatially. The population exposure computed would thus be compared to the traditional approach of assigning population exposures which assumes a stationary population distribution, derived using Census data.

Figure 2 shows the mean populations present in each district within NYC, detected using the mobile phone data and herein termed the “Active Population”. In this figure, a map of the mean nighttime population during the week (Monday–Friday), the mean daytime population determined during the week, and the mean difference between these daytime and nighttime populations are presented. A map of the Census population, herein termed the “Home Population” is also shown for comparison. Figure S4 shows boxplots of the “Active Population” along with the “Home Population” as a single point, for each district of NYC. Figure 3 shows the temporal variation of the number of people present in five separate districts selected from each of the five boroughs of NYC. It can be observed for District 17 in Manhattan, the number of people present peaks during the day, while in the other districts, the population present peaks during the night.

2.4. Population-Weighted Exposure. The population-weighted $PM_{2.5}$ exposures were computed for each district for both population scenarios of Active Population Exposure and Home Population Exposure. This enabled a relative comparison of populations exposure between the two scenarios. District level daily population-weighted exposure is defined as $E_i = \sum_{j=1}^{24} C_i P_{ij}$, where E_i is the daily total population-weighted exposure for each district i , C_i is the estimated concentration of $PM_{2.5}$ in district i on each day, and P_{ij} is the percent of the total population of NYC present in district i at time j (which is hour of the day) for both scenarios described already. Mann–Whitney U Tests were carried out to assess the statistical significance of differences between population weighted $PM_{2.5}$ exposures between scenarios. Weekday and nighttime population-weighted exposures were determined and compared also.

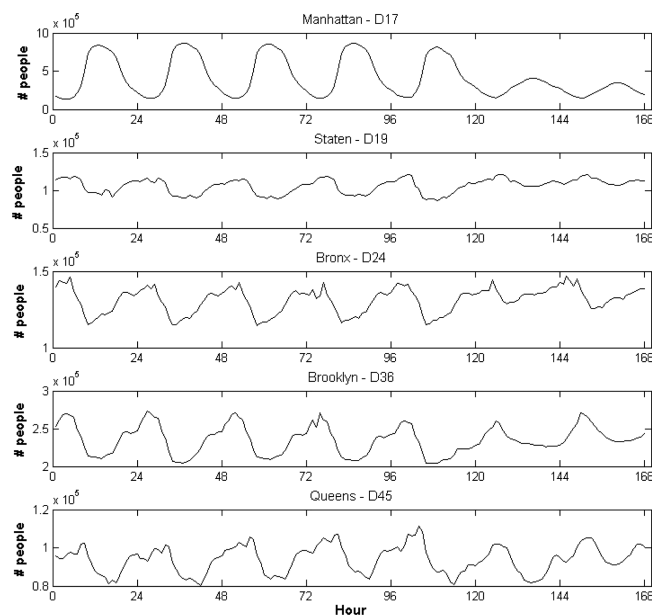


Figure 3. Temporal variability of the number of people detected in five districts located in each of the five boroughs of NYC during 1 week, according to spatially aggregated counts of connections to the cellular network in that district. From the figure, it is observed that in Manhattan the populations present in the Manhattan district (D17) reaches a peak during the day, while in the districts specified in the other four boroughs, the population present reaches a peak during the night.

3. RESULTS

3.1. Relative Difference in Population-Weighted Exposures by District. Population-weighted exposures to $PM_{2.5}$ were computed separately for each district, for the Active Population Exposure and Home Population Exposure scenarios. The relative influence of the population-weighted $PM_{2.5}$ exposures per district of NYC were thus examined. Figure 4 portrays the mean population-weighted $PM_{2.5}$ exposures computed in each district of NYC. In some districts, the population-weighted $PM_{2.5}$ exposure in the case of the Active Population Exposure or the Home Population Exposure are negligible. This was due to the estimated population statistics, which were used to calculate population-weighted exposures, being negligible in the relevant district. Figure 5 shows boxplots of the population-weighted $PM_{2.5}$ exposures for each district within NYC for both scenarios of exposure, and the relative difference between these.

In the case of the Home Population Exposure, Figure 4 shows the districts where the mean population-weighted $PM_{2.5}$ exposures computed are relatively higher than the rest of NYC. These are located within Manhattan, Queens and Brooklyn. For example, in Manhattan, District 63 and 38 (Hudson Heights and Upper East Side), District 42 (Upper East Side), District 13 (Midtown East), and in District 68 (East Village), the mean population-weighted $PM_{2.5}$ exposures are relatively higher. In Queens, District 41 (Flushing, Murray Hill, College Point, and Whitestone), District 11 (Jamaica, Rochdale Village, St. Albans, and Addisleigh Park Historic District in Queens), and District 59 (Cambria Heights, Laurelton, and Springfield Gardens) have relatively higher mean population-weighted exposures to $PM_{2.5}$. In Brooklyn, Districts 55 and 53 (Astoria and Greenpoint), Districts 2 and 26 (Bensonhurst and Borough Park), and

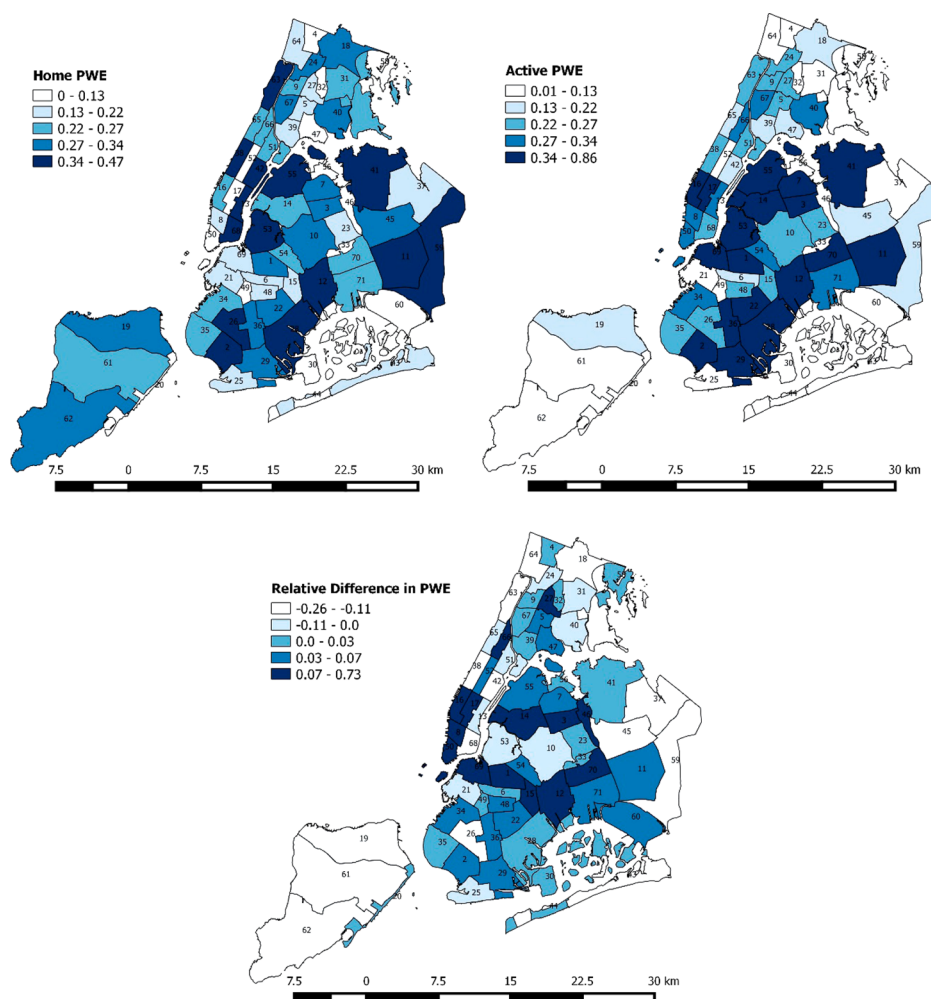


Figure 4. Map of mean population-weighted exposure (PWE) to $PM_{2.5}$ per district for the Home Population Exposure scenario (top-left), the Active Population Exposure scenario (top-right), and the relative difference between these two scenarios (lower). Units are $\mu g/m^3 \times$ percent of population present/district.

Districts 28 and 12 (Flatlands and East New York) have also been identified.

For the Active Population Exposure, which utilizes mobile-phone-based spatiotemporal population distributions to estimate population-weighted $PM_{2.5}$ exposures, Figure 4 shows Districts 17 and 16 (Midtown Manhattan) as having relatively higher $PM_{2.5}$ population-weighted exposures than the remaining districts in NYC. A cluster of districts in West Queens and North-West Brooklyn (Districts 1, 69, 53, 14, 55, 7, 3, and 41) have been identified as having relatively higher population-weighted exposures. In addition, a group of districts in South-East Brooklyn and South-West Queens (Districts 2, 29, 36, 28, 22, 12, 70, and 11) were found to exhibit relatively higher population-weighted exposures than the other districts of NYC.

Some areas in East Queens and South East Brooklyn were identified as areas where relatively greater population-weighted exposure occurs when $PM_{2.5}$ exposure was weighted by the Census population statistics (see Figure 4). However, when $PM_{2.5}$ exposure was weighted according to mobile-phone-based population activity patterns, it can be seen that districts with higher relative influence on the total $PM_{2.5}$ population exposure are located in the lower region of Manhattan and more centralized areas of Brooklyn and Queens (Figure 4). In

examining the relative changes in population-weighted $PM_{2.5}$ exposures calculated using a mobile phone based spatiotemporal population distribution (Active Population Exposure) compared with an exposure computed using a static Census-defined spatial population (Home Population Exposure), the largest increases in mean population-weighted exposures computed were observed in districts located in Lower Manhattan (Districts 16, 17, 8, and 50), and some districts of West Queens (Districts 14, 3, and 46) and North-East Brooklyn (Districts 69, 1, 15, 12, and 70) (Figure 4). District 17, which is located in Midtown Manhattan, was observed to have one of the lowest population-weighted $PM_{2.5}$ exposures of NYC when examining Home Population Exposure; however, it was ranked highest when examining Active Population Exposure. Differences in population-weighted $PM_{2.5}$ exposures computed in individual districts are shown in Figure S5 also.

Considering the population-weighted $PM_{2.5}$ exposures in the Active Population Exposure and the Home Population Exposure separately, the values computed for both were statistically significantly ($p < 0.05$) different in 68 of the 71 districts. The results of statistical tests assessing the differences in population-weighted $PM_{2.5}$ exposures between the two scenarios are shown in Table S1.

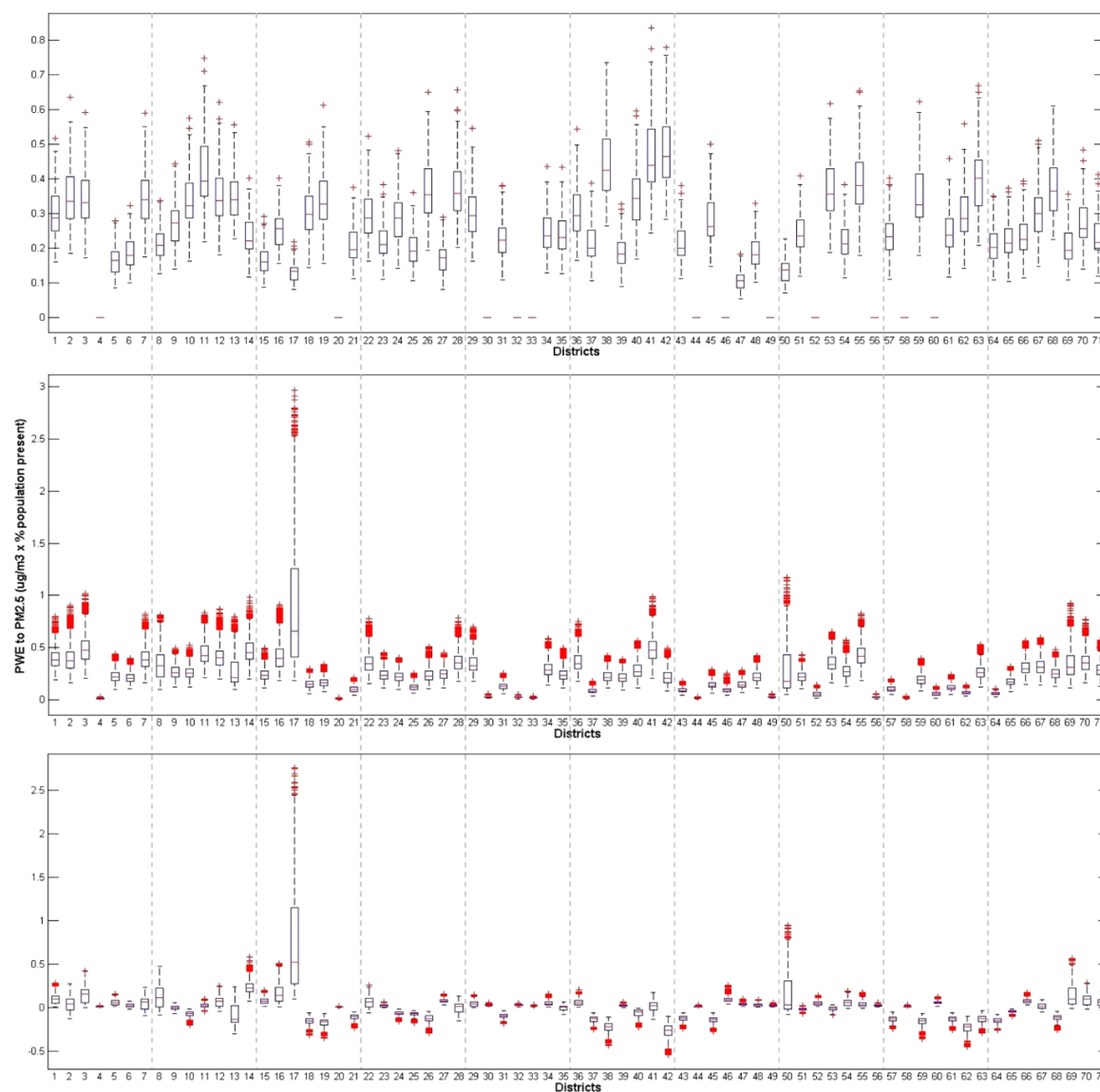


Figure 5. Boxplots of population-weighted exposure (PWE) to $\text{PM}_{2.5}$ for 71 districts of New York City for the (a) Home Population Exposure assuming the Census-defined spatial population distribution, (b) Active Population Exposure assuming the mobile-device-based spatiotemporal population distribution, and (c) relative difference between the Active Population Exposure and the Home Population Exposure. For the Home Population Exposure, in some cases, there are no reported residents in the relevant district according to the 2010 Census; therefore, the resulting PWE is negligible. Units of PWE are $\mu\text{g}/\text{m}^3 \times$ percent of population present/district.

3.2. Relative Temporal Differences in Population-Weighted Exposures.

In the case of the Active Population Exposure, Figure S6 shows the mean population-weighted $\text{PM}_{2.5}$ exposures computed during the daytime and the nighttime in districts within NYC, and the relative difference between the daytime and nighttime. The districts where the population-weighted $\text{PM}_{2.5}$ exposures are relatively higher are very clearly located within Midtown and Lower Manhattan, and centralized areas of Brooklyn and Queens. In analyzing population-weighted $\text{PM}_{2.5}$ exposure in the Active Population Exposure during the daytime and during the nighttime on weekdays, statistically significant ($p < 0.05$) differences occurred in 57 out of the 71 districts. During the weekend, statistically significant ($p < 0.05$) differences occurred in 31 out of 71 districts. These results are shown in Table S1.

3.3. Total Population Exposure for New York City. Figure 6 shows the daily cumulative population-weighted exposures determined for the Active Population Exposure

scenario for the entire study duration. A similar graph for the Home Population Exposure is seen in the Supporting Information (Figure S7). Figure S8 enables a graphical comparison of the daily cumulative population-weighted exposures for two population scenarios. It can be seen that the Home Population Exposure and Active Population Exposure distributions are similar. However, more incidences of $\text{PM}_{2.5}$ values lower than $10 \mu\text{g}/\text{m}^3$ in the Home scenario are observed in comparison to the Active scenario, in which more $\text{PM}_{2.5}$ exposure values greater than $10 \mu\text{g}/\text{m}^3$ are seen.

4. DISCUSSION

The impact of urban mobility patterns determined using cellular network data, in evaluating population exposure to air pollution for a large study domain and human population has not been previously investigated. This study applied population activity patterns representative of several million people to

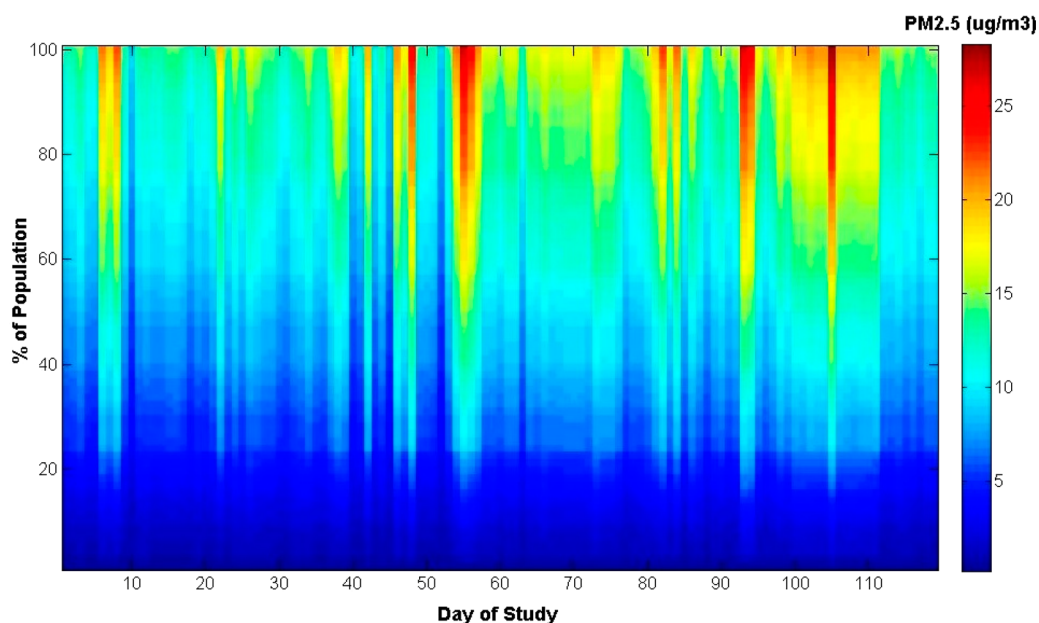


Figure 6. Cumulative percentage of population and $PM_{2.5}$ exposure determined for the Active Population Exposure scenario for each day of the study period of 121 days.

estimate population weighted exposure to air pollution in NYC. Spatiotemporal population statistics as defined by aggregated counts of connections to the cellular phone network, and $PM_{2.5}$ measurements were combined to estimate population-weighted $PM_{2.5}$ exposure, herein termed Active Population Exposure. This was compared to Home Population Exposure, which was calculated assuming a static Census-defined spatial population distribution.

The 71 districts of NYC were compared in terms of their relative contribution to the total population $PM_{2.5}$ exposure under both scenarios. When districts were examined on an individual basis, significantly different ($p < 0.05$) patterns of population-weighted $PM_{2.5}$ exposure were observed. In analyzing the Home Population Exposure, districts which contributed most to the overall population exposure of NYC tended to be located in areas of North-West and South-East Brooklyn, and within Queens. This was due to these districts having more residents than other areas according to the Census population estimates. However, when the Active Population Exposure was considered, districts with higher relative influence tended to be located in the lower regions of Manhattan and centralized areas of Brooklyn and centralized areas of Queens. This was a result of higher proportions of New Yorkers spending time in busy districts for employment, recreational, and social activities. For the Active Population Exposure scenario, the regions with relatively higher population-weighted exposures were concentrated in more distinct clusters of districts relative to the Home Population Exposure. This tentatively implies that efforts for reducing population exposure could be focused on these clusters of districts or regions within NYC. The relative differences between population-weighted exposures observed under varying population conditions, were larger where there were greater disparities between the number of people living in a region and the number of people most likely to spend time in an area, such as Lower Manhattan where a large increase was observed, and in West Queens and North-East Brooklyn.

By using varying spatiotemporal population metrics to investigate population exposure to air pollution, a new perspective on identifying areas of elevated population exposures is shown. Therefore, this research could aid in the prioritization of air pollution interventions (both infrastructural and policy orientated) for the protection of human health. Authorities could focus their air pollution monitoring and modeling efforts, for example by locating monitoring stations where populations are more likely to be exposed. Evidence suggests most air pollution interventions lead to health benefits, including reduced incidence of cardiovascular and respiratory mortality and morbidity.⁶⁴ In a review of the public health impacts of urban air pollution in 25 European cities, Pascal et al.⁶⁵ estimates that complying with the WHO guideline value of $10 \mu\text{g}/\text{m}^3$ in annual mean $PM_{2.5}$ exposure would add up to 22 months of life expectancy at age 30, depending on the city. This corresponded to a total of 19,000 deaths delayed for the regions studied. Further to this, the associated predicted monetary gain was approximately €31 billion annually, including savings on health expenditures, absenteeism, and intangible costs, such as well-being, life expectancy, and quality of life. While evaluating where people are exposed to air pollution in the future using mobile phone based population activity estimates, this could assist in identifying where people are being exposed to levels above the WHO recommended limits. Appropriate actions could then be taken to reduce this number, that is, assigning resources to prioritized areas, thereby maximizing public health and related societal and economic benefits.

The use of spatially interpolated $PM_{2.5}$ data in place of an air quality model calibrated using real data was a study limitation. Also, in future studies, a pollutant exhibiting more spatial heterogeneity such as nitrogen dioxide, may offer further evidence of the importance of considering mobility patterns in evaluating population-exposure to air pollution. Although municipal air quality monitoring is important for indicating general levels of population exposure in urban environments, it is often supplemented by physical air quality models such as the Operational Street Pollution Model (OSPM)⁶⁶ or the ADMS

Urban⁶⁷ to yield substantial spatial resolution. Techniques for estimating air pollution levels between monitoring sites include pollution dispersion,⁶⁸ Land Use Regression (LUR)⁶⁹ or hybrid models (satellite data and LUR).⁷⁰ Real-time air quality monitoring using networks of sensors⁷¹ have received some attention in recent years although the effectiveness of their use for epidemiological studies has yet to be evaluated. Future research should investigate whether air pollution exposure weighted by mobile phone and wireless device based spatiotemporal population distributions can offer improved prediction estimates of premature all-cause mortality, cardiovascular and respiratory mortality and morbidity, and other health outcomes. Previous environmental epidemiological studies have not incorporated dynamic human activity patterns into population exposure assessments. The methodology described in this study could be applied to studies predicting health end-points for different air pollution scenarios of exposure. In future research related to this study, population-weighted air pollution exposures for each district will be determined based on particular times of the day and days of the week. Probabilities of higher population exposure will therefore be compared temporally which may also be useful for city municipalities mitigating air pollution levels by, for example, using adaptive traffic management strategies to reduce vehicular emissions and air pollution concentration levels in localized areas at specific times.⁷² A study limitation is the potential bias in the population mobility patterns determined; in particular regarding omission of individuals who are less likely to carry mobile devices and to travel on a daily basis. Subpopulations such as young children and the elderly are less likely to carry mobile phones or mobile devices, and furthermore are less likely to partake in daily commuting or travel, relative to other demographics. As such, these subgroups may be under-represented in the mobile phone based population activity samples detected and the subsequent population-weighted exposures quantified in this study.

Previous modeling frameworks have aimed to estimate population exposure to air pollution using varying spatiotemporal population mobility estimates and activity patterns.^{44–48} For example, Burke et al.,⁴⁷ developed the Stochastic Human Exposure Model and Dose Simulation (SHEDS-PM) for modeling population exposure to PM. The population for the simulation were generated using demographic data at the census tract level and individual diaries of human activity were randomly assigned to each individual. These diaries were obtained from the EPA Consolidated Human Activity Database (CHAD) which includes 22 000 diary days collated from 10 separate surveys. The database contains various microenvironments which individuals visit and the various activities performed, while in each microenvironment. Further to this, the Air Pollutants Exposure model (APEX), which is part of the EPA's Total Risk Integrated Methodology (TRIM) modeling system predicts health risks due to air pollution exposure and inhalation.⁴⁸ APEX is a population-based, stochastic, micro-environmental model and also uses simulated individual profiles and randomly assigned activity patterns derived from the EPA's CHAD database. These models are functional for assessing human exposure to air pollution, however, include some limitations such as the mobility patterns and activities of populations being randomly assigned, rather than being based on measured mobility estimates for the populations considered. As demonstrated in this study, cellular network data can be used to measure the mobility patterns of millions of people.

Therefore, in future work, it may be possible to combine both approaches to enhance population exposure modeling capabilities.

The research methodology described in this study will enable a better quantification of population exposure on a city-wide scale, and can be applied across larger geographical regions where relevant data exists. Other geo-referenced data which would be indicative of human mobility patterns could be used (e.g., traces from various GPS-enabled devices). One of the novel future contributions from this research is that geo-referenced digital phone traces can also be used to decipher individual trajectories. Therefore, personal air pollution exposure studies could be conducted on cohorts incorporating locations of exposure through mobile phone and wireless device trace data. Accessible cellular network data are necessary for progressing research in light of the epidemiological and subsequent public health insights which can be gained. As ethical issues and concerns around individual data privacy emerge, however, appropriate stewarding and anonymization of data through spatial aggregation or otherwise needs to be ensured.

A novel perspective on population exposure is presented in this study, whereby population exposure to environmental air pollution is quantified using two dynamic variables, using spatiotemporal variations in air pollution levels and extensive spatiotemporal population activity patterns. This is the first study of its kind to use measured population activity patterns representative of several million people to quantify population-weighted exposure to air pollution on an urban scale, using cellular network data. This research is a novel contribution to environmental exposure science and the environmental epidemiology research sphere, and warrants consideration in future studies linking air quality and human health end-points.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.est.6b02385](https://doi.org/10.1021/acs.est.6b02385).

Information on study domain and NYCCAS monitoring sites, additional description of air pollution and population distributions, supplementary figures portraying comparisons of population exposure scenarios, and temporal variation, and results of statistical tests evaluating exposure differences between scenarios (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +16173848822. E-mail: nyhan@hsph.harvard.edu.

Present Address

[†]M.N.: Harvard University, Harvard School of Public Health, Boston, MA, USA 02215.

Notes

The authors declare no competing financial interest.

This study was conducted primarily at Massachusetts Institute of Technology.

■ ACKNOWLEDGMENTS

M.N. would like to thank Fulbright, the Irish Environment Protection Agency, and the Program for Research in Third Level Institutions in Ireland. We thank the Singapore National Research Foundation through the Singapore–MIT Alliance for

Research and Technology. We thank Ericsson for providing datasets for this research and for their support of the MIT Senseable City Laboratory. The authors thank the New York City Department of Health and Mental Hygiene, Queens College Center for the Biology of Natural Systems, and Zev Ross for granting access to the New York City Community Air Survey data-set. We also thank Ms. Siyi Zhang and Dr. Daniel Kondor for processing some of the data used in the study.

REFERENCES

- (1) Brunekreef, B.; Holgate, S. T. Air pollution and health. *Lancet* **2002**, *360*, 1233–1242.
- (2) Liao, D.; Duan, Y.; Whitsel, E. A.; Zheng, Z. J.; Heiss, G.; Chinchill, V. M.; Lin, H. M. Association of higher levels of ambient criteria pollutants with impaired cardiac autonomic control: A population-based study. *Am. J. Epidemiol.* **2004**, *159*, 768–777.
- (3) Pope, C. A., III; Dockery, D. W. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manage. Assoc.* **2006**, *56*, 709–742.
- (4) Laden, F.; Schwartz, J.; Speizer, F. E.; Dockery, D. W. Reduction in fine particulate air pollution and mortality extended follow-up of the Harvard six cities study. *Am. J. Respir. Crit. Care Med.* **2006**, *173*, 667–672.
- (5) Pope, C. A., III; Ezzati, M.; Dockery, D. W. Fine-particulate air pollution and life expectancy in the United States. *N. Engl. J. Med.* **2009**, *360*, 376–86.
- (6) Gurjar, B. R.; Jain, A.; Sharma, A.; Agarwal, A.; Gupta, P.; Nagpure, A. S.; Lelieveld, J. Human health risks in megacities due to air pollution. *Atmos. Environ.* **2010**, *44*, 4606–4613.
- (7) Nyhan, M.; Misstear, B. D.; McNabola, A. Comparison of particulate matter dose and acute heart rate variability response in cyclists, pedestrians, bus and train passengers. *Sci. Total Environ.* **2014**, *468–469*, 821–831.
- (8) WHO (World Health Organization). 7 million premature deaths annually linked to air pollution. Press release, World Health Organization Website, 25th March 2014. <http://www.who.int>.
- (9) Brook, R. D.; Rajagopalan, S.; Pope, C. A.; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; Holguin, F.; Hong, Y.; Luepker, R. V.; Mittleman, M. A.; Peters, A.; Siscovick, D.; Smith, S. C.; Whitsel, L.; Kaufman, J. D. Particulate matter air pollution and cardiovascular disease, an update to the scientific statement from the American Heart Association. *Circulation* **2010**, *121*, 2331–2378.
- (10) Lee, J.-T.; Kim, H.; Hong, Y. C.; Kwon, H. J.; Schwartz, J.; Christiani, D. C. Air pollution and daily mortality in seven major cities of Korea, 1991–1997. *Environ. Res.* **2000**, *84*, 247–254.
- (11) Levy, J. I.; Hammit, J. K.; Spengler, J. D. Estimating the mortality impacts of particulate matter: What can be learned from between-study variability? *Environ. Health Persp.* **2000**, *108*, 109–117.
- (12) Le Tertre, A.; Quenel, P.; Eilstein, D.; Medina, S.; Prouvost, H.; Pascal, L.; Boumghar, A.; Saviuc, P.; Zeghnoun, A.; Filleul, L.; Declercq, C.; Cassadou, S.; Le Goaster, C. Short-term effects of air pollution on mortality in nine French cities: a quantitative summary. *Arch. Environ. Health* **2002**, *57*, 311–319.
- (13) Burnett, R. T.; Goldberg, M. S. Size-Fractionated Particulate Mass and Daily Mortality in Eight Canadian Cities. *Revised Analyses of Time-Series of Air Pollution and Health: Special Report*; Health Effects Institute: Boston, MA, 2003; pp 85–90.
- (14) Dominici, F.; McDermott, A.; Zeger, S. L.; Samet, J. M. National maps of the effects of particulate matter on mortality: exploring geographical variation. *Environ. Health Persp.* **2003**, *111*, 39–43.
- (15) Omori, T.; Fujimoto, G.; Yoshimura, I.; Nitta, H.; Ono, M. Effects of particulate matter on daily mortality in 13 Japanese cities. *J. Epidemiol.* **2003**, *13*, 314–322.
- (16) Anderson, H. R.; Atkinson, R. W.; Peacock, J. L.; Sweeting, M. J.; Marston, L. Ambient particulate matter and health effects: publication bias in studies of short-term associations. *Epidemiology* **2005**, *16*, 155–163.
- (17) Analitis, A.; Katsouyanni, K.; Dimakopoulou, K.; Samoli, E.; Nikoloulopoulos, A. K.; Petasakis, Y.; Touloumi, G.; Schwartz, J.; Anderson, H. R.; Cambra, K.; Forastiere, F.; Zmirou, D.; Vonk, J. M.; Clancy, L.; Kriz, B.; Bobvos, J.; Pekkanen, J. Short-term effects of ambient particles on cardiovascular and respiratory mortality. *Epidemiology* **2006**, *17*, 230–233.
- (18) Zanobetti, A.; Schwartz, J. The effects of fine and coarse particulate air pollution on mortality: a national analysis. *Environ. Health Persp.* **2009**, *117*, 898–903.
- (19) Janssen, N. A. H.; Fischer, P.; Marra, M.; Ameling, C.; Cassee, F. R. Short-term effects of PM_{2.5}, PM₁₀ and PM_{2.5–10} on daily mortality in the Netherlands. *Sci. Total Environ.* **2013**, *463–464*, 20–26.
- (20) Schwartz, J. Airborne particles and daily deaths in 10 US cities. In *Revised Analyses of Time-Series of Air Pollution and Health: Special Report*; Health Effects Institute: Boston, MA, 2003; pp 211–218.
- (21) Schwartz, J. The effects of particulate air pollution on daily deaths: A multi-city case-crossover analysis. *Occup. Environ. Med.* **2004**, *61*, 956–961.
- (22) COMEAP. *Cardiovascular Disease and Air Pollution*, A report by the Committee on the Medical Effects of Air Pollutant's Cardiovascular Sub-Group; United Kingdom Department of Health: London, United Kingdom, 2006.
- (23) Katsouyanni, K.; Touloumi, G.; Samoli, E.; Petasakis, Y.; Analitis, A.; Le Tertre, A.; Rossi, G.; Zmirou, D.; Ballester, F.; Boumghar, A.; Anderson, H. R.; Wojtyniak, B.; Paldy, A.; Braunstein, R.; Pekkanen, J.; Schindler, C.; Schwartz, J. Sensitivity analysis of various models of short-term effects of ambient particles on total mortality in 29 cities in APHEA2. *Revised Analyses of Time-Series of Air Pollution and Health, Technical Report*; Health Effects Institute: Boston, MA, 2003; pp 157–164.
- (24) Dockery, D. W.; Pope, C. A.; Xu, X.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. G.; Speizer, F. E. An Association between air pollution and mortality in six US cities. *N. Engl. J. Med.* **1993**, *329* (24), 1753–1759.
- (25) Hoek, G.; Brunekreef, B.; Goldbohm, S.; Fischer, P.; van den Brandt, P. A. Association between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. *Lancet* **2002**, *360*, 1203–1209.
- (26) Pope, C. A., III; Burnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; Thurston, G. D. Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution. *J. Am. Med. Assoc.* **2002**, *287*, 1132–1141.
- (27) Jerrett, M.; Burnett, R. T.; Ma, R.; Pope, C. A., III; Krewski, D.; Newbold, K. B.; Thurston, G.; Shi, Y.; Finkelstein, N.; Calle, E. E.; Thun, M. J. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005**, *16*, 727–736.
- (28) Krewski, D.; Jerrett, M.; Burnett, R. T.; Ma, R.; Hughes, E.; Shi, Y.; Turner, M. C.; Pope, C. A., III; Thurston, G.; Calle, E. E.; Thun, M. J.; Beckerman, B.; DeLuca, P.; Finkelstein, N.; Ito, K.; Moore, D. K.; Newbold, K. B.; Ramsay, T.; Ross, Z.; Shin, H.; Tempalski, B. *Extended Follow-up and Spatial Analysis of the American Cancer Society Study Linking Particulate Matter and Mortality*, Research Report from the Health Effects Institute; Health Effects Institute: Boston, MA, 2009; pp 5–114, 115–136.
- (29) Kloog, L.; Coull, B. A.; Zanobetti, A.; Koutrakis, P.; Schwartz, J. D. Acute and chronic effects of particles on hospital admissions in New England. *PLoS One* **2012**, *7* (4), e34664.
- (30) Knibbs, L. D.; Cole-Hunter, T.; Morawska, L. A review of commuter exposure to ultrafine particles and its health effects. *Atmos. Environ.* **2011**, *45*, 2611–2622.
- (31) Hoek, G.; Krishnan, R. M.; Beelen, R.; Peters, A.; Ostro, B.; Brunekreef, B.; Kaufman, J. D. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ. Health* **2013**, *12*, 43.
- (32) Nyberg, F.; Gustavsson, P.; Jarup, L.; Bellander, T.; Berglund, N.; Jakobsson, R.; Pershagen, G. Urban air pollution and lung cancer in Stockholm. *Epidemiology* **2000**, *11*, 487–495.
- (33) Ma, Y.; Chen, R.; Pan, G.; Xu, X.; Song, W.; et al. Fine particulate air pollution and daily mortality in Shenyang, China. *Sci. Total Environ.* **2011**, *409*, 2473–2477.

- (34) Steinle, S.; Reis, S.; Sabel, E. C. Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Sci. Total Environ.* **2013**, *443*, 184–193.
- (35) Lepeule, J.; Laden, F.; Dockery, D. W.; Schwartz, J. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities Study from 1974 to 2009. *Environ. Health Persp.* **2012**, *120* (7), 965–70.
- (36) Puett, R. C.; Hart, J. E.; Yanosky, J. D.; Paciorek, C.; Schwartz, J.; Suh, H.; Speizer, F. E.; Laden, F. Chronic fine and coarse particulate exposure, mortality and coronary heart disease in the Nurses' Health Study. *Environ. Health Persp.* **2009**, *117* (11), 1697–701.
- (37) Avery, C. L.; Mills, K. T.; Williams, R.; McGraw, K. A.; Poole, C.; Smith, R. L.; Whitsel, E. A. Estimating error in using ambient PM_{2.5} concentrations as proxies for personal exposures: a review. *Epidemiology* **2010**, *21*, 215–223.
- (38) Nyhan, M.; Misstear, B.; McNabola, A. Evaluating artificial neural networks for predicting minute ventilation and lung deposited dose in commuting cyclists. *J. Transport Health* **2014**, *1* (4), 305–315.
- (39) Valero, N.; Aguilera, I.; Llop, S.; Esplugues, A.; De Nazelle, A.; Ballester, F.; Sunyer, J. Concentrations and determinants of outdoor, indoor and personal nitrogen dioxide in pregnant women from two Spanish birth cohorts. *Environ. Int.* **2009**, *35*, 1196–1201.
- (40) Schembari, A.; Triguero-Mas, M.; De Nazelle, A.; Dadvand, P.; Vrijheid, M.; Cirach, M.; Martinez, D.; Figueras, F.; Querol, X.; Basagana, X.I.; Eeftens, M.; Meliefste, K.; Nieuwenhuijsen, M. J. Personal, indoor and outdoor air pollution levels among pregnant women. *Atmos. Environ.* **2013**, *64*, 287–295.
- (41) Dons, E.; Int Panis, L.; Van Poppel, M.; Theunis, J.; Willems, H.; Torfs, R.; Wets, G. Impact of time-activity patterns on personal exposure to black carbon. *Atmos. Environ.* **2011**, *45*, 3594–3602.
- (42) Nieuwenhuijsen, M. J.; Donaire-Gonzalez, D.; Rivas, I.; de Castro, M.; Cirach, M.; Hoek, G.; Seto, E.; Jerrett, M.; Sunyer, J. Variability in and Agreement between Modeled and Personal Continuously Measured Black Carbon Levels Using Novel Smartphone and Sensor Technologies. *Environ. Sci. Technol.* **2015**, *49* (5), 2977–2982.
- (43) Kaur, S.; Nieuwenhuijsen, M. J. Determinants of Personal Exposure to PM_{2.5}, Ultrafine Particle Counts and CO in a Transport Microenvironment. *Environ. Sci. Technol.* **2009**, *43*, 4737–4743.
- (44) Beckx, C.; Int Panis, L.; Arentze, T.; Janssens, D.; Torfs, R.; Broekx, S.; Wets, G. A dynamic activity-based population modelling approach to evaluate exposure to air pollution: methods and application to a Dutch urban area. *Environ. Impact Asses.* **2009**, *29*, 179–185.
- (45) Dhondt, S.; Beckx, C.; Degraeuwe, B.; Lefebvre, W.; Kochan, B.; Bellemans, T.; Int Panis, L.; Macharis, C.; Putman, K. Health impact assessment of air pollution using a dynamic exposure profile: implications for exposure and health impact estimates. *Environ. Impact Asses.* **2012**, *36*, 42–51.
- (46) Hatzopoulou, M.; Miller, E. J. Linking an activity-based travel demand model with traffic emission and dispersion models: Transport's contribution to air pollution in Toronto. *Transport. Res. D: Tr. E* **2010**, *15*, 315–325.
- (47) Burke, J. M.; Zufall, M.; Özkaynak, H. A population exposure model for particulate matter: case study results for PM_{2.5} in Philadelphia, PA. *J. Exposure Anal. Environ. Epidemiol.* **2001**, *11*, 470–489.
- (48) U.S. EPA (United States Environmental Protection Agency). *Total Risk Integrated Methodology (TRIM) Air Pollutants Exposure Model (APEX) Documentation: TRIM.Expo/APEX, Version 4: User Guide*; U.S. EPA, 2012.
- (49) De Nazelle, A.; Seto, E.; Donaire-Gonzalez, D.; Mendez, M.; Matamala, J.; Nieuwenhuijsen, M. J.; Jerrett, M. Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environ. Pollut.* **2013**, *176*, 92–99.
- (50) Glasgow, M. L.; Rudra, C. B.; Yoo, E.; Demirbas, M.; Merriman, J.; Nayak, P.; Crabtree-Ide, C.; Szpiro, A. A.; Rudra, A.; Wactawski-Wende, J.; Mu, L. Using smartphones to collect time-activity data for long-term personal-level air pollution exposure assessment. *J. Exposure Sci. Environ. Epidemiol.* **2016**, *26* (4), 356–64.
- (51) Su, J. G.; Jerrett, M.; Meng, Y.; Pickett, M.; Ritz, B. Integrating smart-phone based momentary location tracking with fixed site air quality monitoring for personal exposure assessment. *Sci. Total Environ.* **2015**, *S06–S07*, S18–S26.
- (52) Candia, J.; Gonzalez, M. G.; Wang, P.; Schoenharl, T.; Madey, G.; Barabasi, A.-L. Uncovering individual and collective human dynamics from mobile phone records. *J. Phys. A: Math. Theor.* **2008**, *41*, 224015.
- (53) Gonzalez, M.; Hidalgo, C.; Barabasi, A. L. Understanding individual human mobility patterns. *Nature* **2008**, *453* (7196), 779–782.
- (54) Jiang, S.; Ferreira, J.; Gonzalez, M. Clustering daily patterns of human activities in the city. *Data Min. Knowl. Disc.: Special Issue: Data Mining Technologies for Computational Social Science* **2012**, *25* (3), 478–510.
- (55) Calabrese, F.; Diao, M.; Di Lorenzo, G.; Ferreira, J., Jr.; Ratti, C. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transport Res. C-Emer.* **2013**, *26*, 301–313.
- (56) Jiang, S.; Fiore, G. A.; Yang, Y.; Ferreira, J., Jr.; Frazzoli, E.; Gonzalez, M. C. A review of urban computing for mobile phone traces: Current methods, challenges and opportunities. *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, Chicago, IL, USA, 2013.
- (57) Calabrese, F.; Colonna, M.; Lovisolo, P.; Parata, D.; Ratti, C. Real-time urban monitoring using cell phones: a case study in Rome. *IEEE Transactions on Intelligent Transportation Systems* **2011**, *12* (1), 141–151.
- (58) Calabrese, F.; Di Lorenzo, G.; Liu, L.; Ratti, C. Estimating origin-destination flows using opportunistically collected mobile phone location data from one million users in Boston Metropolitan Area. *IEEE Pervasive Computing* **2011**, *10*, 36–44.
- (59) U.S. CB (United States Census Bureau). Population Statistics, U.S. Census Bureau, 2016.
- (60) New York City Department of Health & Mental Hygiene; Queens College Center for the Biology of Natural Systems; Zev Ross Spatial Analysis. *New York City Community Air Survey (NYCCAS) dataset*, 2016.
- (61) Matte, T. D.; Ross, Z.; Kheirbek, I.; Eisl, H.; Johnson, S.; Gorczynski, J. E.; Kass, D.; Markowitz, S.; Pezeshki, G.; Clougherty, J. E. Monitoring intraurban spatial patterns of multiple combustion air pollutants in New York City: Design and implementation. *J. Exposure Sci. Environ. Epidemiol.* **2013**, *23*, 223–231.
- (62) Li, L.; Losser, T.; Yorke, C.; Piltner, R. Fast inverse distance weighting-based spatiotemporal interpolation: a web-based application of interpolating daily fine particulate matter PM_{2.5} in the contiguous US using parallel programming and k-d tree. *Int. J. Environ. Res. Public Health* **2014**, *11* (9), 9101–9141.
- (63) U.S. EPA (United States Environmental Protection Agency). Air Quality System Data. United States Environmental Protection Agency Website. www.epa.gov.
- (64) Henschel, S.; Atkinson, R.; Zeka, A.; Le Tertre, A.; Analitis, A.; Katsouyanni, K.; Chanel, O.; Pascal, M.; Forsberg, B.; Medina, S.; Goodman, P. G. Air pollution interventions and their impact of public health. *Int. J. Public Health* **2012**, *57*, 757–768.
- (65) Pascal, M.; Corso, M.; Chanel, O.; Declercq, C.; Badaloni, C.; Cesaroni, G.; Henschel, S.; Meister, K.; Haluza, D.; Martin-Olmedo, P.; Medina, S. Assessing the public health impacts of urban air pollution in 25 European cities: Results of the Apekom project. *Sci. Total Environ.* **2013**, *449*, 390–400.
- (66) *Operational Street Performance Model (OSPM)*; NERI, Aarhus University: The Netherlands.
- (67) *CERC: ADMS Urban*; Cambridge Environmental Research Consultants: Cambridge, United Kingdom.
- (68) Kumar, P.; Ketzler, M.; Vardoulakis, S.; Pirjola, L.; Britter, R. Dynamics and dispersion modelling of nanoparticles from road traffic

in the urban atmospheric environment - A review. *J. Aerosol Sci.* **2011**, *42*, 580–603.

(69) Dons, E.; Van Poppel, M.; Kochan, B.; Wets, G.; Int Panis, L. Modeling temporal and spatial variability of traffic-related air pollution: Hourly land use regression models for black carbon. *Atmos. Environ.* **2013**, *74*, 237–246.

(70) Chudnovsky, A.; Lyapustin, Y.; Wang, Y.; Tang, C.; Schwartz, J.; Koutrakis, P. High resolution aerosol data from MODIS satellite for urban air quality studies. *Open Geosciences* **2014**, *6*, 17–26.

(71) Raju, H. P.; Partheeban, P.; Hemamalini, R. R. Urban mobile air quality monitoring using GIS, GPS, sensors and internet. *Int. J. Environ. Sci. Dev.* **2012**, *3* (4), 323–327.

(72) Nyhan, M.; Sobolevsky, S.; Kang, C.; Robinson, P.; Corti, A.; Szell, M.; Streets, D.; Lu, Z.; Britter, R.; Barrett, S. R. H.; Ratti, C. Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model. *Atmos. Environ.* **2016**, *140*, 352–363.