



Article

# The Impact of Individual Mobility on Long-Term Exposure to Ambient PM<sub>2.5</sub>: Assessing Effect Modification by Travel Patterns and Spatial Variability of PM<sub>2.5</sub>

Eun-hye Yoo <sup>1,\*</sup>, Qiang Pu <sup>1</sup>, Youngseob Eum <sup>1</sup> and Xiangyu Jiang <sup>2</sup>

<sup>1</sup> Department of Geography, State University of New York at Buffalo, Buffalo, NY 14260, USA; qiangpu@buffalo.edu (Q.P.); yeum@buffalo.edu (Y.E.)

<sup>2</sup> Georgia Environmental Protection Division, Atlanta, GA 30354, USA; xiangyuj@buffalo.edu

\* Correspondence: eunhye@buffalo.edu

**Abstract:** The impact of individuals' mobility on the degree of error in estimates of exposure to ambient PM<sub>2.5</sub> concentrations is increasingly reported in the literature. However, the degree to which accounting for mobility reduces error likely varies as a function of two related factors—individuals' routine travel patterns and the local variations of air pollution fields. We investigated whether individuals' routine travel patterns moderate the impact of mobility on individual long-term exposure assessment. Here, we have used real-world time–activity data collected from 2013 participants in Erie/Niagara counties, New York, USA, matched with daily PM<sub>2.5</sub> predictions obtained from two spatial exposure models. We further examined the role of the spatiotemporal representation of ambient PM<sub>2.5</sub> as a second moderator in the relationship between an individual's mobility and the exposure measurement error using a random effect model. We found that the effect of mobility on the long-term exposure estimates was significant, but that this effect was modified by individuals' routine travel patterns. Further, this effect modification was pronounced when the local variations of ambient PM<sub>2.5</sub> concentrations were captured from multiple sources of air pollution data ('a multi-sourced exposure model'). In contrast, the mobility effect and its modification were not detected when ambient PM<sub>2.5</sub> concentration was estimated solely from sparse monitoring data ('a single-sourced exposure model'). This study showed that there was a significant association between individuals' mobility and the long-term exposure measurement error. However, the effect could be modified by individuals' routine travel patterns and the error-prone representation of spatiotemporal variability of PM<sub>2.5</sub>.

**Keywords:** long-term exposure to ambient PM<sub>2.5</sub>; uncertainty; mobility-based approach; spatial exposure models; routine travel patterns



**Citation:** Yoo, E.H.; Pu, Q.; Eum, Y.; Jiang, X. The Impact of Individual Mobility on Long-Term Exposure to Ambient PM<sub>2.5</sub>: Assessing Effect Modification by Travel Patterns and Spatial Variability of PM<sub>2.5</sub>. *Int. J. Environ. Res. Public Health* **2021**, *18*, 2194. <https://doi.org/10.3390/ijerph18042194>

Academic Editors: Daniel A. Griffith and Yongwan Chun

Received: 30 December 2020

Accepted: 12 February 2021

Published: 23 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Environmental health studies have routinely conducted exposure assessments using pollutant measurements from sparse networks of monitors matched with the locations of home addresses of individuals [1,2]. However, reliance on these networks leads to measurement error in exposure assessments because they are unable to capture high spatial heterogeneity in pollution fields [3,4]. Similarly, this approach, which focuses only on individuals' location of residences, ignores time-activities of individuals and fails to assess dynamic exposure to air pollution [5–9]. In sum, considerable uncertainties exist in residence-based exposure assessments relying on air monitoring networks.

As an effort to improve the resolution and availability of air pollution data, recent studies have incorporated additional data sources, such as satellite images and computer simulation outcomes, to supplement monitoring measurements. Particularly, satellite-derived aerosol optical depth (AOD) and the Community Multiscale Air Quality Modeling System (CMAQ) have received growing interests as they provide fine spatiotemporal data

with complete spatial coverage. It is worth noting, however, that the estimation of spatial and temporal variability of ambient air pollutant concentration alone cannot guarantee the accuracy and precision of personal exposure estimates [10–15].

Similarly, several epidemiological studies attempted to reduce exposure misclassification by accounting for the variations in the time people spend in different locations by using time-weighted pollutant concentrations in each microenvironment [16–19]. These studies typically collected participants' activity data and their locations using 24-h diaries in a survey [20], which involves high costs and raises concerns regarding reliability and validity [21,22]. Recent technological advancements, such as global positioning system (GPS) and mobile phones, and their adoption in exposure assessments have shifted the paradigm in exposure studies [23–26]. Unlike traditional travel surveys, mobile phones equipped with GPS provide detailed and accurate time–activity information with low burden to the phone carriers and low cost to researchers [27,28]. These mobile phones are also able to frequently measure and record their locations. Researchers can use the resulting geographic coordinates (latitude and longitude) to measure an individual's locations at multiple points throughout ones' daily life instead of relying solely on a person's residential location to estimate his or her air pollutant exposures [2,29,30].

Despite a continuous effort to understand the relationship between ambient air pollution and personal exposures, previous investigations of exposure measurement error and exposure misclassification have been confined to a single issue. Some studies have focused on the mischaracterization of the spatial heterogeneity of air pollution due to the limited availability and resolution of environmental pollution data [31–35]. Other studies investigated the effects of ignoring human mobility on air pollution exposures [36–41]. The uncertainty of personal exposure to ambient air pollution is likely affected by both the representation of spatiotemporal variability of air pollution and the use of time–activity data to capture dynamic nature of individuals' movements [28,42–44]. However, few studies systematically investigated the joint effect of the human mobility and spatial exposure models on the uncertainty of the personal exposure assessment.

This gap is important to fill because exposure measurement error is likely driven by a synergy of two related factors. Assessment based on location-specific exposure at the home address rather than mobility-based exposure contributes to exposure measurement error. However, the degree of accuracy versus error in these home-based estimates will likely vary based on individuals' routine travel patterns and/or how the spatiotemporal dynamics of air pollutant concentrations are modeled. For example, the exposure measurement error induced by failing to account for daily mobility will be much smaller for individuals who typically spend most time at home compared to people who regularly travel long distance and spend much time outside the home [45]. Similarly, personal exposure with and without taking one's daily mobility into account would make little to no difference if a spatial exposure model fails to accurately represent the spatiotemporal variability of air pollution.

An increasing number of studies have demonstrated the effect of mobility on individuals' exposure to air pollution, although most studies are based on either simulated or approximated mobility data, such as call detail records. Some studies have conducted using empirical data or accurate mobility data, but they often lack generalizability due to a small sample size and/or a short study term. For example, Dons et al. [46] assessed personal exposure to black carbon using real time location data collected from GPS loggers of 16 participants over 7 days, whereas Park and Kwan [38] compared individuals' exposure to O<sub>3</sub> with and without consideration of 80 participant's simulated activities on a single day. Other studies were conducted with a large number of participants, but individuals' location data are approximated by nearest cell tower locations [2,30,47]. This discrepancy between actual and approximated movement data may lead to exposure measurement error in longitudinal studies of air pollution and health, particularly if they were matched with spatially and temporally resolved air pollutant estimates. Lastly, most studies that accounted for individuals' mobility tend to focus on short-term exposure (from a few

minutes to 24 h at best) to air pollution unless simulated data are used [48–50]. It is clear that the effect of mobility on longer term exposure over months, referred to as long-term (or sub-chronic) exposure hereafter, using empirical and accurate time–activity data with sufficiently large participants has been understudied.

In the present study, we examined the effect of individuals' mobility on long-term air pollution exposure measurement error using accurate and real-world time–activity data collected from thousands of participants in the study area over a few months. We also investigated individuals' routine travel patterns and spatial exposure models as moderators of the effect of an individual's mobility. First, we examine whether individuals' routine travel patterns moderate the impact of mobility on individual long-term exposure assessment. Next, we examine the role of the spatiotemporal representation of ambient fine particulate matter (PM<sub>2.5</sub>) concentrations as a second moderator of the relationship between individuals' mobility and the exposure measurement error. We predict a three-way interaction effect of spatial exposure models, routine travel patterns, and individuals' mobility on long-term exposure measurement error to ambient PM<sub>2.5</sub> concentrations using the data collected in Erie and Niagara counties, New York, NY, USA.

## 2. Materials and Methods

### 2.1. Time–Activity Data

We used time–activity information of 2013 residents in the Erie and Niagara counties, New York State (NYS), from 1 December 2016 to 31 May 2017. Most participants began their survey in October or November of 2016, and participated in the study for a minimum of 72 days to 182 days. To ensure the continuous monitoring, we retained the mobile phone traces observed between December 2016 and May 2017. At the beginning of the study we collected the address of residence and workplace (if applicable) of participants. Individuals' whereabouts during the study period were collected using their own mobile phone and an application developed for Android and iOS systems by our research team. Specifically, mobile phone traces were periodically reported by the Android application every 35 min on average with a high accuracy (40 m of positional error). For iPhone users, their movement information was updated whenever the device moved more than 500 m from the last recorded location. The iOS system-based location data had a high positional uncertainty of 650 m, but they had systematic temporal gaps. We filled these gaps using a set of imputation strategies with a 30-min imputation interval, while taking into account self-reported place information provided by study participants and positional uncertainty of the data. Detailed information on the data and the imputation strategies are available in Yoo et al. [51].

Still, we found that substantial amounts of daily time–activity information were missing to calculate individuals' daily time-integrated exposure to ambient PM<sub>2.5</sub>. To address this issue, we developed a deep long short-term memory imputation model to predict the numerical representation of time–activity for the days. Here, we replaced the missing time–activity with the daily observations which are the most similar to the predicted numerical representation. We imputed missing time–activity data at every 30 min-interval and assigned the place at which the activity took place. We evaluated the performance of imputation in terms of the accuracy of reproducing time–activity patterns of individuals, such as daily average time spent at home, work, and shopping places each day.

### 2.2. Air Pollution Data

We considered two spatial exposure model scenarios in this study to reconstruct spatial variability of daily ambient PM<sub>2.5</sub> concentrations in 2016. First, we generated estimates of daily PM<sub>2.5</sub> concentrations using a single-sourced data, that is, measurements obtained from the monitoring networks [7]. There was a limited number of monitoring stations in the range of 1 to 5 stations operated in Erie/Niagara county during the study period of 2016. To improve the reliability of predictions, we used a total of 38 monitoring

stations present in NYS for model building. Monitoring data used in the present study were obtained from the U.S. Environmental Protection Agency (EPA, <http://www.epa.gov/>).

In the multi-sourced exposure model, we compensated the lack of spatial coverage of monitoring station data by incorporating both satellite-derived Aerosol Optical Depth (AOD) and the Community Multi-scale Air Quality Model (CMAQ) outputs. Satellite-derived AOD does not directly represent ground-level PM<sub>2.5</sub> concentrations, and thus they need to be calibrated by accounting for the influences from meteorological, topological and anthropogenic conditions [52,53]. In the present paper, we reconstructed a daily PM<sub>2.5</sub> concentration surface over the study area primarily based on the 1 km high resolution AOD derived by the algorithm of Multi-Angle Implementation of Atmospheric Correction (MAIAC) [54–56]. During the study period, however, there were many days without sufficient MAIAC AOD retrievals due to cloud/snow presence. To address this missing AOD problem, we imputed the missing values by incorporating multiple proxy data available at coarse spatial resolutions, including Modern-Era Retrospective Analysis for Research and Applications version 2 and the CMAQ model outputs [57].

### 2.3. Measures

#### 2.3.1. Individual’s Travel Patterns

We examined daily travel patterns of each participant using spatial and temporal metrics, namely, the characteristic distance travelled by individuals and the time allocated at places other than one’s home address [58–60]. Specifically, we calculated a radius of gyration (RoG), denoted by  $r_g^j$ , on a daily basis as

$$r_g^j = \sqrt{\frac{1}{N_j} \sum_{k \in L_j} n_k (\mathbf{u}_k - \mathbf{u}_{c_j})^2}, \tag{1}$$

to characterize the distance travelled by a participant on a day  $j$ . Here,  $L_j$  denotes the set of locations visited by a participant on the day  $j$ ;  $\mathbf{u}_k$  is a two-dimensional vector of geographic coordinates of location  $k$ ;  $n_k$  is the frequency of visits at the  $k$ -th location;  $N_j = \sum_{k \in L_j} n_k$  is the total number of trips made on the day  $j$ ; and  $\mathbf{u}_{c_j}$  is the center of mass of the individual on the day  $j$ . The daily characteristic distance of  $i$ -th participant ( $i = 1, \dots, 2013$ ) was further aggregated as an average RoG  $r_i = \frac{1}{n_i} \sum_{j=1}^{n_i} r_g^j$  over the number of days  $n_i$  that the participant  $i$  enrolled in the survey. We also calculated each participant’s daily averaged time spent away from home (referred to as non-home time), denoted as  $\tau_i$ , over the study period. Considering the positional error of mobile phone location data, we used a buffer of 2 km centered at each participant home address and excluded all GPS trajectories within the buffered zones in the calculation of  $\tau_i$ .

Based on these two metrics, we classified study participants into three groups: (1) individuals with static travel patterns who spend most time at home; (2) individuals with moderate travel patterns as they either travel long-distance but spend little time outside the home or travel short distances but spend many hours outside the home; (3) individuals with active travel patterns who regularly travel long-distance and spend large amount of time away from home. The classification scheme is summarized in Table 1, where we used the mean distance ( $R_c$ ) of daily RoGs in pooled study participants and the mean non-home time ( $\Gamma_c$ ) of the study participants during the study period, respectively.

**Table 1.** Stratification of individuals’ travel patterns.

Travel Patterns	Static	Moderate	Active
	$r_i < R_c, \tau_i < \Gamma_c$	$r_i \geq R_c, \tau_i < \Gamma_c$ $r_i < R_c, \tau_i \geq \Gamma_c$	$r_i \geq R_c, \tau_i \geq \Gamma_c$

### 2.3.2. Spatial Variability of Daily PM<sub>2.5</sub> Concentration

We used the two spatial exposure models to reconstruct the daily PM<sub>2.5</sub> variations across the study area. In the single-source exposure model, we estimated daily PM<sub>2.5</sub> concentrations throughout the study area using Ordinary Kriging (OK), known as a best linear unbiased estimator among spatial interpolation methods [61–63]. In the multi-sourced exposure model, we used an ensemble model that integrates multiple machine-learning algorithms based on its superior performance to a single machine learning algorithm [64,65].

The two exposure models differ substantially in terms of the data sources that they relied on rather than their statistical approaches. Both spatial interpolation techniques make predictions across space from air monitoring data and rely on relationships being identified within the data itself or with information regarding the surrounding environment [66,67]. Both models generate predictions of daily PM<sub>2.5</sub> concentration across the entire study area using some forms of spatial dependence, which are quantified as a mathematical function of the distance between monitoring sites. When sparsely sampled air monitoring data are used as the primary source in the single-sourced exposure model, an overly-smooth spatial structure of daily PM<sub>2.5</sub> concentrations is produced with high levels of spatial uncertainty. Meanwhile, the multi-sourced model captures local variations of PM<sub>2.5</sub> concentrations by incorporating additional information from AOD and auxiliary variables, such as land cover, population density, and meteorological conditions.

The performance of the two exposure models was evaluated using the 10-fold spatial cross validation (CV) method. Specifically, the monitoring sites within NYS in 2016 were randomly divided into 10 groups where sites from nine groups were used as training sites and the remaining sites were used for validation. We trained each model with the same training data aggregated from the training sites and made predictions at validation sites. The differences between predicted and known values at validation sites were summarized by metrics, such as root mean square error (RMSE) and mean prediction error (MPE). This process was repeated 10 times until all monitoring sites were selected as validation sites.

### 2.3.3. Personal Exposures to PM<sub>2.5</sub> Concentrations

We estimated personal exposure to ambient PM<sub>2.5</sub> using both a residence-based and mobility-based approach. Specifically, we matched home address and mobile-phone traces with a daily PM<sub>2.5</sub> concentration surface estimated from the two spatial exposure models (single-sourced and multi-sourced models), respectively. It should be noted that PM<sub>2.5</sub> monitoring data and satellite-derived AOD are obtained for the entire year of 2016, but the time–activity data are collected across the years of 2016 and 2017. Here, the primary focus lies on the measurement error by ignoring participants' whereabouts instead of the actual exposure estimates and thus we treated the time–activity data collected in 2017 (January to May) as observations in 2016 and matched them with corresponding PM<sub>2.5</sub> estimate surfaces.

For each participant we estimated four sets of daily exposure to PM<sub>2.5</sub> and aggregated them over time for long-term exposure. Let  $y_M^S(i, j)$  be the PM<sub>2.5</sub> exposure estimate of participant  $i = 1, \dots, 2013$  on day  $j = 1, \dots, n_i$  from a PM<sub>2.5</sub> surface estimated from the  $S$  exposure model. Here,  $n_i$  denotes the total number of days that the  $i$ -th subject participated the study. Each day we also estimated a pair of daily PM<sub>2.5</sub> exposures with ( $M = 1$ ) and without ( $M = 0$ ) one's time–activity patterns. The residence-based exposure was calculated by extracting the PM<sub>2.5</sub> concentration estimates at each participant's home address, where daily PM<sub>2.5</sub> surfaces were obtained from a single-sourced ( $S = 0$ ) and a multi-sourced ( $S = 1$ ) exposure model, respectively. For the mobility-based exposure, we used the general form of the equation where the time-weighted exposure was calculated as [68]

$$y_{M=1}^S(i, j) = \sum_{k \in L_j} C_k^S t_k, \quad (2)$$

where  $C_k^S$  and  $t_k$  are the PM<sub>2.5</sub> concentration estimated from the  $S$ -source exposure model ( $S = 0, 1$ ) and the total time spent at the  $k$ -th place located at  $\mathbf{u}_k$  on the day  $j$ , respectively.

The daily exposure to ambient PM<sub>2.5</sub> concentrations of each study participant was further summarized into a long-term exposure by taking the average of the *i*-th participant's daily exposure estimates as  $Y_M^S(i) = \frac{1}{n_i} \sum_{j=1}^{n_i} y_M^S(i, j)$  for  $S = 0, 1$  and  $M = 0, 1$ .

#### 2.4. Statistical Analyses

We developed a 3-way interaction model in a linear mixed-effect regression to study the effect of the mobility on exposures to ambient PM<sub>2.5</sub> concentrations. Here, we also explicitly accounted for the moderation by other two factors—individuals' routine travel patterns and spatial exposure models. Three sets of analyses were conducted to address each hypothesis below:

**Hypothesis 1.** *The consideration of individuals' mobility affects long-term exposure to ambient PM<sub>2.5</sub> concentrations.*

**Hypothesis 2.** *Routine travel patterns moderate the relationship between the mobility and estimates of the individuals' exposure to ambient PM<sub>2.5</sub> concentrations.*

**Hypothesis 3.** *The type of spatial exposure models (multi-sourced or single-sourced) moderates the two-way interaction between of mobility and the long-term travel patterns on estimates of individual exposure to ambient PM<sub>2.5</sub>.*

To test each hypothesis, we developed a multiple regression model with additional interaction terms to determine whether a moderating effect exists. We also accounted for the potential correlation among multiple exposure measures per participant by developing a random effect model.

In the statistical analyses, we used an indicator coding scheme for the mobility effect. It is worth noting that our primary focus was on the statistical significance of the relative difference of mobility-based exposures from residence-based exposures. We anticipated that the sign or the magnitude of the differences (i.e., exposure measurement errors) might vary for pollutants other than PM<sub>2.5</sub> or different study regions. However, the statistical significance of the differences implies that the consideration of individuals' mobility affects the exposure measurement error. For individuals' routine travel patterns, we used a numerical variable (referred to as 'activity-travel score'), which coded static, moderate and active travel patterns as a numerical value of 0, 1, and 2, respectively. This new coding scheme facilitates the interpretation of the statistical analysis. For example, we can relate the changes in the long-term exposure to ambient PM<sub>2.5</sub> as a function of individuals' travel pattern in terms of the travel distance and non-home time. This coding also enabled us to account for the larger differences between static and active travel patterns versus moderate and active travel patterns.

The statistical analyses, spatial prediction of PM<sub>2.5</sub> concentrations and time-activity pattern analyses were conducted in R version 4.0.2 (R Core Team, 2020). We used the nlme package (v.3.1-148) and the effects package (4.2-0) for statistical analyses and visualizations; and the gstat package (v.2.0.6) for a single-sourced exposure model. We used a fast and scalable machine learning platform, H2O (<https://www.h2o.ai/>) (accessed on 3 February 2021), to build the multi-sourced ensemble model.

### 3. Results

#### 3.1. Daily Mobility Patterns and Routine Travel Patterns

A total of 1013 participants' time-activity data collected between 1 December 2016 and 31 May 2017 were included in the study. For each participant we reconstructed time-activity patterns at 30-min intervals each day during the study period. During the six months of the study period, a total of 1612 to 2012 participants engaged in the survey each day. Based on the reconstructed time-activity trajectories, we characterized each participant's routine travel patterns by calculating both the averaged radius of gyration

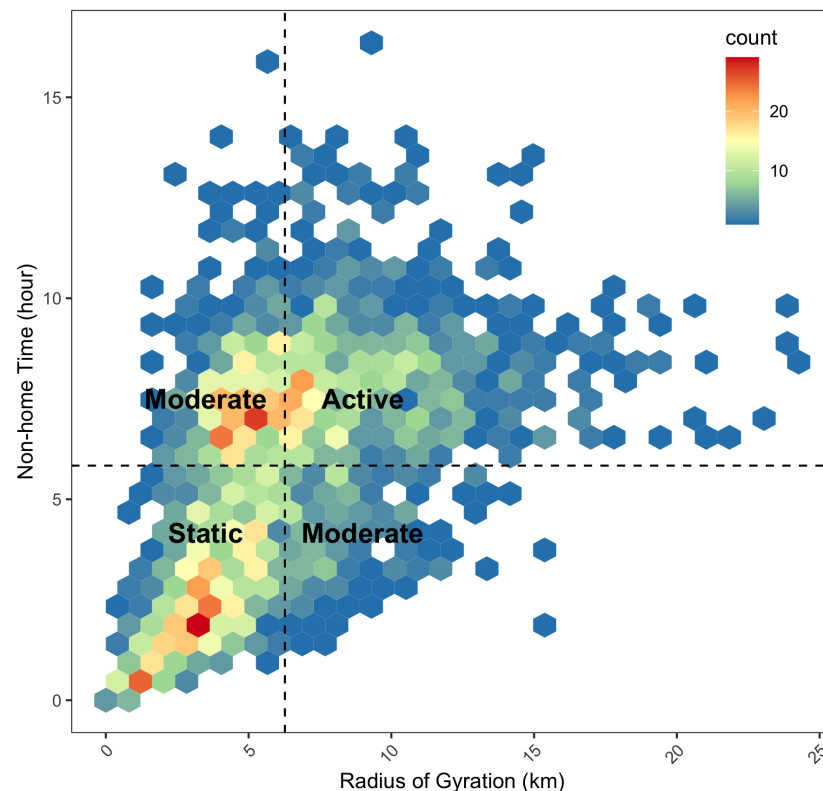
( $r_i$  in km) and the averaged time spent away from home ( $\tau_i$  in hour). The averaged daily travel distance measured in the averaged radius of gyration (RoG) and the time spent away from home (NhT) of each study participant are summarized in Table 2. The pooled and average travel distance, characterized by  $R_c = 6.27$  km with a standard deviation 3.60 km with a range of 0.08 to 24.34 km. The study participants also spent their time away from home on average  $\Gamma_c = 5.84$  h, with a standard deviation of 2.91 h.

**Table 2.** The summary statistics of a total of 2013 study participants’ travel patterns quantified by the Radius of Gyration (RoG) and Non-home Time (NhT).

	Mean	SD	Min	Q1	Median	Q3	Max
RoG	6.27	3.60	0.08	3.76	5.51	8.13	24.34
NhT	5.84	2.91	0.03	3.34	6.37	7.94	16.21

SD denotes standard deviation; Q1 and Q3 denote the first and third quartile, respectively.

Based on the averaged daily travel distance and time spent outside the home, we classified study participants into three groups. Study participants whose averaged RoG and NhT are greater than 6.27 km and 5.84 h were considered as active travelers (Active in Figure 1). In contrast, if their RoG and NhT are both less than 6.27 km and 5.84 h, we classified their mobility patterns into static (Static in Figure 1). The rest of the study participants were considered to have moderate travel patterns. The classification scheme is illustrated in Figure 1. The total number of participants for Static, Moderate, and Active group is 720 (35.8 %), 654 (32.0%), and 648 (32.2%), respectively.



**Figure 1.** Stratification of individuals’ travel patterns based on the averaged travel distance (RoG) of 6.27 km and time spent away from home (Non-home time) of 5.84 h.

3.2. Spatial and Temporal Variability of PM<sub>2.5</sub> Concentrations

Over the study period of six months (January to May, and December) in 2016, the spatial average of daily PM<sub>2.5</sub> concentration in the study region varied from 1.25 µg/m<sup>3</sup>

and  $16.58 \mu\text{g}/\text{m}^3$  with a mean of  $7.15 \mu\text{g}/\text{m}^3$  and standard deviation of  $3.11 \mu\text{g}/\text{m}^3$  from the single-sourced exposure model and a mean of  $5.60 \mu\text{g}/\text{m}^3$  and standard deviation of  $2.28 \mu\text{g}/\text{m}^3$  from the multi-sourced exposure model. The daily mean and standard deviation of  $\text{PM}_{2.5}$  estimates obtained from two models were calculated over the six months and we summarized their distributions in Table 3. Overall, the average of daily  $\text{PM}_{2.5}$  concentration estimated from the multi-sourced exposure model was lower than the averaged  $\text{PM}_{2.5}$  predictions obtained from the single-sourced model. However, the multi-sourced model captured higher levels of spatial variability (1.16 of the multi-sourced model versus 0.25 of the single-sourced model). The standard deviation (SD) of daily  $\text{PM}_{2.5}$  concentration estimates, which quantifies how daily  $\text{PM}_{2.5}$  predictions vary in space, was also substantially different between the two models. The multi-sourced exposure model predictions have higher variability with a mean of 1.16 than a single-sourced model estimates whose mean SD was as low as 0.25.

**Table 3.** The spatial mean and variability of  $\text{PM}_{2.5}$  concentration estimates from a single-sourced and multi-sourced exposure models.

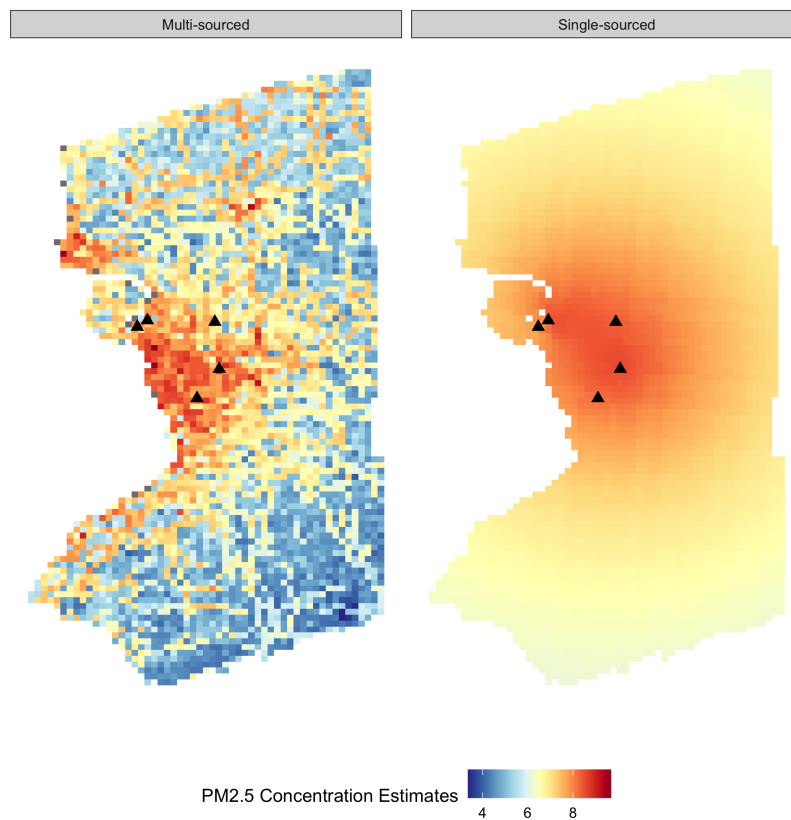
Model		Mean	SD	Min	Q1	Median	Q3	Max
Single-Sourced	DM	7.15	3.11	1.25	4.86	6.71	8.77	16.58
	DSD	0.25	0.35	0.00	0.06	0.15	0.29	2.39
Multi-Sourced	DM	5.60	2.28	1.14	3.96	5.24	6.83	14.66
	DSD	1.16	0.58	0.20	0.68	1.11	1.48	3.13

DM denotes the daily mean and DSD denotes the daily standard deviation.

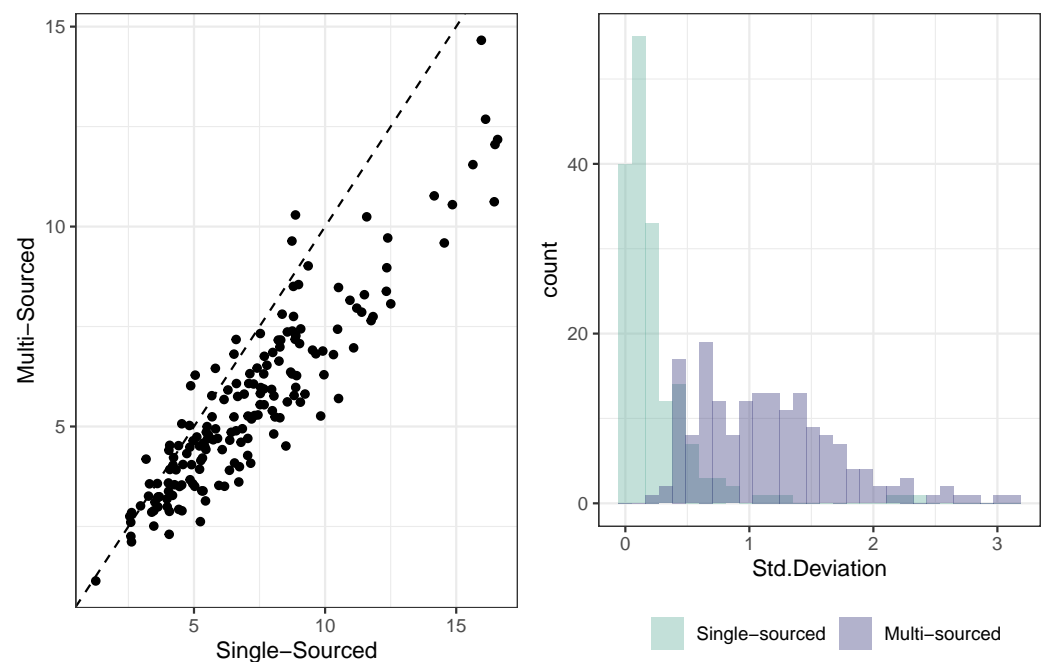
As illustrated in Figure 2, the  $\text{PM}_{2.5}$  predictions of a single day (12/01/2016) obtained from the single-sourced exposure model in the right panel resulted in smoother spatial structures than the multi-sourced exposure model in the left panel. This is highly anticipated because the smoothing effect of the Ordinary Kriging is well known by geostatisticians and is considered as a serious drawback [69]. Specifically, the smooth patterns of  $\text{PM}_{2.5}$  prediction in the right panel of Figure 2 is a result of making predictions from sparsely sampled air monitoring data (denoted as the symbol of triangles), where low levels of  $\text{PM}_{2.5}$  predictions were overestimated and high values were underestimated. Meanwhile, the  $\text{PM}_{2.5}$  prediction from the multi-sourced model in the left panel of Figure 2 depicted local variations of daily  $\text{PM}_{2.5}$  concentrations that were affected by the proximity to road networks, the spatial variations of population density and land use type. These visual inspections were also asserted by the spatial 10-fold cross validation (CV), where the multi-sourced model outperformed the single-sourced model with  $\text{CV-R}^2$  of 0.74 versus 0.66. A similar pattern was found in the comparison of RMSE/MPE of the two models:  $1.91/1.40 \mu\text{g}/\text{m}^3$  for the multi-sourced model and  $1.96/1.49 \mu\text{g}/\text{m}^3$  for the single-sourced model.

Lastly, we compared the spatial averaged  $\text{PM}_{2.5}$  concentration levels from both models at each day during the study period. This comparison is significant because the daily spatial variation of  $\text{PM}_{2.5}$  predictions could directly affect personal exposure estimates. As presented in Figure 3, the single-sourced model consistently yielded higher  $\text{PM}_{2.5}$  concentrations than the multi-sourced model except a few days (see the left panel). We also examined the distribution of spatial variation of the estimated daily  $\text{PM}_{2.5}$  concentrations from both models using the histogram in the right panel. The standard deviation of the estimates across the study area is much smaller for a single-sourced model compared to the multi-sourced model, suggesting that the single-sourced model might result in over-smoothed  $\text{PM}_{2.5}$  predictions.





**Figure 2.** PM<sub>2.5</sub> estimates obtained from both a multi-sourced and single-sourced exposure models on 1 December 2016 overlaid with the five air monitoring stations operated in 2016 (denoted as symbols of black triangle).



**Figure 3.** PM<sub>2.5</sub> concentration estimates obtained from both a multi-sourced and single-sourced exposure models and the comparison of model predictions using the daily spatial mean and standard deviations.

### 3.3. Mobility Effect on Long-Term Exposure to Ambient PM<sub>2.5</sub>

For each participant, we calculated four long-term exposure estimates obtained from two spatial exposure models of ambient PM<sub>2.5</sub> with and without the individual's mobility. Based on the four exposure estimates of 2013 participants, we investigated the effect of personal mobility on individual exposure to ambient PM<sub>2.5</sub>. Considering that the contribution of mobility in personal exposure assessment may vary per individuals' travel patterns and the spatiotemporal variability of air pollution fields, we formally tested their interactions in the random-effect regression model. We separately assessed their interactions per a spatial exposure model that captures different levels of spatial variability of PM<sub>2.5</sub> concentrations.

Specifically, we conducted multiple regression analyses to evaluate whether the consideration of mobility patterns affects the air pollution exposure assessment (i.e., measurement error). The results of analyses are summarized in Table 4. As shown in Step 1, Hypothesis 1 predicted that mobility-based exposure has a significant effect on individual exposure to ambient PM<sub>2.5</sub> ( $\beta = 0.03, p < 0.01$ ).

Hypothesis 2 predicted the effect of mobility on personal exposure measurement error would vary depending on individuals' routine travel patterns. For example, mobility-based exposure would be higher for individuals with active travel patterns than individuals who travel long distances but do not spend much time outside the home or who travel short distances although spend much time outside the home (moderate travel patterns). As shown in Table 4 (Step 2), the interaction effect of mobility and individuals' travel patterns were statistically significant ( $\beta = 0.03, p < 0.05$ ).

Hypothesis 3 predicted that the effect of mobility on personal exposure measurement error would be substantial when the PM<sub>2.5</sub> concentration is estimated from a multi-sourced exposure model and for individuals whose travel patterns are active—routinely travel long distances and spent a lot of time outside the home. As shown in the Step 3 of Table 4, the three-way interaction effect of mobility, active travel patterns of individuals, and a multi-sourced exposure model on personal exposure to PM<sub>2.5</sub> was significant ( $\beta = 0.05, p < 0.05$ ).

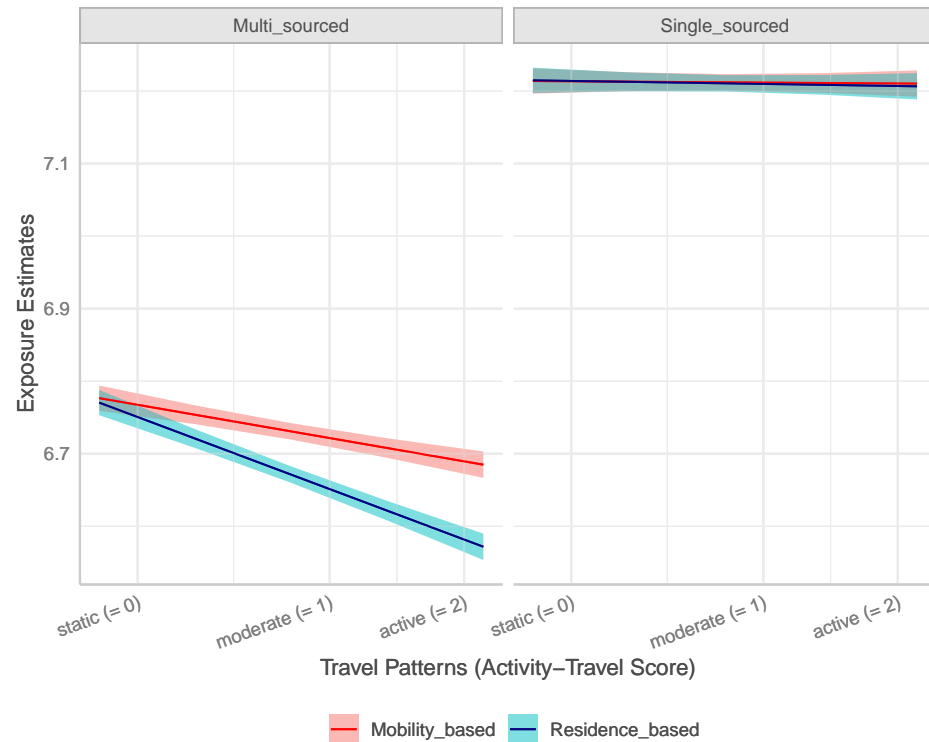
**Table 4.** Statistical models.

	Step 1	Step 2	Step 3
(Intercept)	7.23 ***	7.23 ***	7.21 ***
<i>Main Effects</i>			
Mobility-based Approach	0.03 **	−0.03	−0.00
Travel Patterns	−0.04 ***	−0.02	−0.00
Multi-sourced Exposure Model	−0.51 ***	−0.47 ***	−0.44 ***
<i>2-Way Interaction Terms</i>			
Mobility-based × Travel Patterns		0.03 *	0.00
Mobility-based × Multi-sourced		0.06 **	0.01
Travel Patterns × Multi-sourced		−0.07 ***	−0.10 ***
<i>3-Way Interaction Term</i>			
Mobility-based × Travel Patterns × Multi-sourced			0.05 *
AIC	10,702.71	10,672.50	10,674.81
BIC	10,744.66	10,735.42	10,744.73
Log Likelihood	−5345.35	−5327.25	−5327.41

\*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ .

To facilitate the understanding of this significant three-way interaction, we plotted the effect of mobility patterns on personal exposure in Figure 4. The long-term exposure estimates from the single-sourced model are substantially higher than the multi-sourced model. Further, there was no effect of travel pattern on exposure estimates when the single-sourced model was used regardless of whether they were based mobility or residence (in the right panel). In contrast, when the daily ambient PM<sub>2.5</sub> was predicted from the multi-sourced exposure model, the effect of mobility varied depending on whether an individual's travel pattern is static or active (in the left panel). Greater activity-travel scores

were associated with lower levels of exposure estimates, when individuals' long-term exposure was estimated at residence. That is, the time-weighted exposures taking into account both multiple places visited by individuals and the time spent at each place was higher than one's residence-based exposure.



**Figure 4.** The three-way interaction effects of individuals' daily mobility, routine travel patterns, and the exposure models on long-term exposure to ambient PM<sub>2.5</sub> concentrations.

#### 4. Discussion

To better understand the effect of mobility on individuals' long-term exposure to ambient PM<sub>2.5</sub>, we developed a statistical model for mobility, individuals' routine travel patterns and spatial exposure models. The results strongly suggested that mobility has a statistically significant effect on personal long-term exposure to ambient PM<sub>2.5</sub>. Our unique contribution to existing literature, however, lies in that we demonstrated that the mobility effect can be modified by individuals' routine travel patterns and the validity of spatial exposure models. For individuals with static travel patterns over an overly smooth PM<sub>2.5</sub> surface, we found that the mobility-based approach has no effect on the accuracy of long-term exposure estimates. On the other hand, the mobility-based approach will be the most effective to capture accurate exposure of individuals with highly active travel patterns over an air pollution field with high spatial and temporal variability.

In the first stage, we examined the effect of individuals' mobility on long-term exposure to ambient PM<sub>2.5</sub> by comparing a location-specific exposure at home with time-weighted exposure based on mobility data. The results suggested that the mobility effect on individuals' long-term exposure to ambient PM<sub>2.5</sub> was statistically significant. When mobility was taken into account, personal exposure to PM<sub>2.5</sub> was 0.03 µg/m<sup>3</sup> higher than overall mean exposure. Our finding on the significant effect of mobility on long-term exposure is in line with previous studies [40,41,50]. To our best knowledge, however, this is one of a few studies to examine and show the statistically significant effects of the mobility on long-term exposure to air pollution in a large sample over a period of multiple months.

In the second stage, we examined effect modification by individuals' routine travel patterns by introducing the two-way interaction term between the mobility and individuals' travel patterns in the multiple regression model. The results revealed that the

single-sourced exposure model contributed to higher long-term exposure estimates to ambient PM<sub>2.5</sub> than the multi-sourced exposure model. Based on the cross-validation results, we speculated that the higher values of exposure estimates were attributed to the overestimation of PM<sub>2.5</sub> predictions from the single-sourced exposure model. This overestimation from the single-sourced model is rather anticipated because the primary source of PM<sub>2.5</sub> measurements obtained from air quality monitoring sites operated in the study region in 2016 are clustered around highly polluted areas, including Tonawanda Coke Facility [70].

Another unique point of our study is that we have used complete and relatively accurate mobility data collected from thousands individuals during their daily life. As demonstrated in our comparison study, however, the effect of mobility on the measurement error of long-term air pollution exposure estimates is only pronounced when they are matched with spatially and temporally resolved PM<sub>2.5</sub> estimates [57]. The difference between residence-based and mobility-based exposure estimates have not been detected unless the spatial dynamics of ambient PM<sub>2.5</sub> were captured by the multi-sourced model. This indicates the measurement error of long-term exposure can be reduced only by improving the spatial and temporal resolution of both mobility data and the representation of air pollution fields simultaneously.

It is important to note that our conclusions only pertain to the uncertainties associated with personal long-term exposure to ambient PM<sub>2.5</sub>. In contrast, our finding cannot speak to the issue of the accuracy of exposure assessment. Given that the true exposure of individuals is unknown, the accuracy (or error) of exposure assessment can not be quantified [71]. However, we attempted to quantify and compare uncertainties in long-term PM<sub>2.5</sub> exposure estimates that arise from the oversimplified assumption ('residence-based approach') on individuals' spatial behaviors. We further examined how the uncertainties associated with individuals' mobility effect on one's long-term exposure are modified by individuals' routine spatial behaviors and the limited representation of spatiotemporal representation of air pollution fields [72].

The main strengths of this study are as follows—first, we examined the effect of mobility on individuals' long-term exposure to ambient PM<sub>2.5</sub> and highlighted the moderating effects of their routine travel patterns and the spatiotemporal dynamic of air pollution fields. Second, we conducted the study using both accurate and spatially and temporally resolved time–activity data from thousands of participants over several months, as well as high-resolution satellite-derived AODs and CMAQ model outputs to improve the availability and resolution of air pollution data [31]. As a result, the effects of small sample size and sample bias on the accuracy of personal exposure assessment were reduced. Third, we found that individuals' mobility has a significant effect on long-term exposure to ambient PM<sub>2.5</sub> in this study, although its effect was moderated by individuals' routine travel patterns and the spatial variability in the study area. Our findings may provide a scientific basis for long-term exposure modeling and the study designs of health effect estimation.

Several limitations should be acknowledged in this study. First, there is potential measurement error because we used measured PM<sub>2.5</sub> concentrations taken at a small number of fixed monitoring stations as a primary data source in both spatial exposure models. This concern was lessened in the multi-sourced exposure model by incorporating additional data, such as AOD, but the sparse monitoring stations operated in the study area likely contributed to the measurement errors. Furthermore, reliance on a sparse monitoring network might affect the observed mobility effects as well as the moderation of the mobility effect on personal exposure to PM<sub>2.5</sub>. Similarly, our findings might be subject to confounding factors, such as study participants' age, gender and occupational status. The effect of stratification of exposure to different subpopulation groups needs to be further investigated. Third, we have stratified individuals' long-term travel patterns into three groups based on both a spatial and temporal mobility metric, namely radius of gyration and time spent away from home. However, other mobility metrics, such as mean square displacement and entropy [73], could be used to quantify individuals' long-term

travel patterns, and its effect on the statistical analysis needs to be investigated. Lastly, our personal exposure assessment was limited to ambient PM<sub>2.5</sub> concentrations for the purpose of illustration without the health effect assessments. In future work, we plan to include other pollutants that are known to have severe adverse health impacts, such as NO<sub>2</sub> and O<sub>3</sub>, as well as the subsequent health outcomes, including flu symptoms.

## 5. Conclusions

This study suggests that there is a significant association between mobility and long-term exposure measurement error, although the mobility effect can be modified by individuals' routine travel patterns and the spatiotemporal variability of air pollutants of concern. Our findings on the modification effect of mobility suggest that even accurately measured mobility may have little to no effect on exposure assessment for individuals with static travel patterns over a relatively homogeneous air pollution surface. Meanwhile, our findings underscore the significance of a mobility-based approach when a study involves individuals who regularly travel long distances and spend longer hours outside the home over spatially and temporally dynamic air pollutant concentrations. These findings suggest that investigators should pay greater attention to both individuals' routine travel patterns and the spatial uncertainty of air pollution predictions to minimize long-term exposure measurement error.

**Author Contributions:** E.-h.Y.: Conceptualization, Data curation, Formal analysis, Methodology, Writing—original draft, Y.E.: Data curation, Writing—original draft, Q.P.: Data curation, Writing—original draft, X.J.: Methodology, Writing—original draft. All authors have read and agreed to the published version of the manuscript.

**Funding:** Eun-hye Yoo and Youngseob Eum are funded in part by NIH Grant #R01GM108731.

**Institutional Review Board Statement:** The study was approved by the Institutional Review Board (IRB) at the University at Buffalo.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets (study participants' mobility data) analyzed during the current study are not publicly available due to the privacy of individuals.

**Acknowledgments:** We are grateful to two guest editors, Daniel A. Griffith and Yongwan Chun, and two anonymous reviewers for helpful comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Chang, H.H.; Fuentes, M.; Frey, H.C. Time series analysis of personal exposure to ambient air pollution and mortality using an exposure simulator. *J. Exp. Sci. Environ. Epidemiol.* **2012**, *22*, 483–488. [[CrossRef](#)] [[PubMed](#)]
2. Dewulf, B.; Neutens, T.; Lefebvre, W.; Seynaeve, G.; Vanpoucke, C.; Beckx, C.; Van de Weghe, N. Dynamic assessment of exposure to air pollution using mobile phone data. *Int. J. Health Geogr.* **2016**, *15*, 1–14. [[CrossRef](#)] [[PubMed](#)]
3. Shaddick, G.; Zidek, J.V. A case study in preferential sampling: Long term monitoring of air pollution in the UK. *Spat. Stat.* **2014**, *9*, 51–65. [[CrossRef](#)]
4. Yoo, E.H.; Zammit-Mangion, A.; Chipeta, M.G. Adaptive spatial sampling design for environmental field prediction using low-cost sensing technologies. *Atmos. Environ.* **2020**, *221*, 117091. [[CrossRef](#)]
5. Brauer, M.; Hoek, G.; Van Vliet, P.; Meliefste, K.; Fischer, P.H.; Wijga, A.; Koopman, L.P.; Neijens, H.J.; Gerritsen, J.; Kerkhof, M.; et al. Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. *Am. J. Respir. Crit. Care Med.* **2002**, *166*, 1092–1098. [[CrossRef](#)]
6. Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P.; Gehring, U.; Heinrich, J.; Cyrys, J.; Bellander, T.; Lewne, M.; et al. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. *Epidemiology* **2003**, *14*, 228–239. [[CrossRef](#)]
7. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; Giovis, C. A review and evaluation of intraurban air pollution exposure models. *J. Exp. Sci. Environ. Epidemiol.* **2005**, *15*, 185–204. [[CrossRef](#)]
8. Liao, D.; Peuquet, D.J.; Duan, Y.; Whitsel, E.A.; Dou, J.; Smith, R.L.; Lin, H.M.; Chen, J.C.; Heiss, G. GIS approaches for the estimation of residential-level ambient PM concentrations. *Environ. Health Perspect.* **2006**, *114*, 1374–1380. [[CrossRef](#)]

9. Tang, R.; Tian, L.; Thach, T.Q.; Tsui, T.H.; Brauer, M.; Lee, M.; Allen, R.; Yuchi, W.; Lai, P.C.; Wong, P.; et al. Integrating travel behavior with land use regression to estimate dynamic air pollution exposure in Hong Kong. *Environ. Int.* **2018**, *113*, 100–108. [[CrossRef](#)]
10. Van Donkelaar, A.; Martin, R.V.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P.J. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environ. Health Perspect.* **2010**, *118*, 847–855. [[CrossRef](#)]
11. Szpiro, A.A.; Paciorek, C.J.; Sheppard, L. Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology* **2011**, *22*, 680. [[CrossRef](#)]
12. Liu, Y.; He, K.; Li, S.; Wang, Z.; Christiani, D.C.; Koutrakis, P. A statistical model to evaluate the effectiveness of PM<sub>2.5</sub> emissions control during the Beijing 2008 Olympic Games. *Environ. Int.* **2012**, *44*, 100–105. [[CrossRef](#)]
13. Brauer, M.; Freedman, G.; Frostad, J.; Van Donkelaar, A.; Martin, R.V.; Dentener, F.; Dingenen, R.v.; Estep, K.; Amini, H.; Apte, J.S.; et al. Ambient air pollution exposure estimation for the global burden of disease 2013. *Environ. Sci. Technol.* **2016**, *50*, 79–88. [[CrossRef](#)] [[PubMed](#)]
14. Jiang, X.; Yoo, E.H. Modeling wildland fire-specific PM<sub>2.5</sub> concentrations for uncertainty-aware health impact assessments. *Environ. Sci. Technol.* **2019**, *53*, 11828–11839. [[CrossRef](#)] [[PubMed](#)]
15. Pu, Q.; Yoo, E.H. Spatio-temporal modeling of PM<sub>2.5</sub> concentrations with missing data problem: A case study in Beijing, China. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 423–447. [[CrossRef](#)]
16. Duan, N. Models for human exposure to air pollution. *Environ. Int.* **1982**, *8*, 305–309. [[CrossRef](#)]
17. Ott, W.R. Concepts of human exposure to air pollution. *Environ. Int.* **1982**, *7*, 179–196. [[CrossRef](#)]
18. Burke, J.M.; Zufall, M.J.; Ozkaynak, H. A population exposure model for particulate matter: Case study results for PM<sub>2.5</sub> in Philadelphia, PA. *J. Expo. Anal. Environ. Epidemiol.* **2001**, *11*, 470–489. [[CrossRef](#)] [[PubMed](#)]
19. Flachsbart, P.G. *Exposure to Carbon Monoxide*; CRC Press: Boca Raton, FL, USA, 2007; pp. 113–146.
20. McCurdy, T.; Glen, G.; Smith, L.; Lakkadi, Y. The national exposure research laboratory’s consolidated human activity database. *J. Exp. Sci. Environ. Epidemiol.* **2000**, *10*, 566–578. [[CrossRef](#)]
21. Freeman, N.C.; de Tejada, S.S. Methods for collecting time/activity pattern information related to exposure to combustion products. *Chemosphere* **2002**, *49*, 979–992. [[CrossRef](#)]
22. Crosbie, T. Using activity diaries: Some methodological lessons. *J. Res. Pract.* **2006**, *2*, D1.
23. Breen, M.S.; Long, T.C.; Schultz, B.D.; Crooks, J.; Breen, M.; Langstaff, J.E.; Isaacs, K.K.; Tan, Y.M.; Williams, R.W.; Cao, Y.; et al. GPS-based microenvironment tracker (MicroTrac) model to estimate time–location of individuals for air pollution exposure assessments: Model evaluation in central North Carolina. *J. Exp. Sci. Environ. Epidemiol.* **2014**, *24*, 412–420. [[CrossRef](#)] [[PubMed](#)]
24. 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*, 2977–2982. [[CrossRef](#)] [[PubMed](#)]
25. Glasgow, M.; Rudra, C.; Yoo, E.-H.; Demirbas, M.; Rudra, C.; Mu, L. Using smartphones to collect time–activity data for long-term personal-level air pollution exposure assessment. *J. Exp. Sci. Environ. Epidemiol.* **2016**, *26*, 356. [[CrossRef](#)]
26. Picornell, M.; Ruiz, T.; Borge, R.; García-Albertos, P.; de la Paz, D.; Lumbreras, J. Population dynamics based on mobile phone data to improve air pollution exposure assessments. *J. Exp. Anal. Environ. Epidemiol.* **2019**, *29*, 278. [[CrossRef](#)]
27. 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. [[CrossRef](#)]
28. Steinle, S.; Reis, S.; Sabel, C.E. Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Sci. Total Environ.* **2013**, *443*, 184–193. [[CrossRef](#)]
29. Ouidir, M.; Giorgis-Allemand, L.; Lyon-Caen, S.; Morelli, X.; Cracowski, C.; Pontet, S.; Pin, I.; Lepeule, J.; Siroux, V.; Slama, R. Estimation of exposure to atmospheric pollutants during pregnancy integrating space–time activity and indoor air levels: Does it make a difference? *Environ. Int.* **2015**, *84*, 161–173. [[CrossRef](#)]
30. Nyhan, M.; Kloog, I.; Britter, R.; Ratti, C.; Koutrakis, P. Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data. *J. Exp. Anal. Environ. Epidemiol.* **2019**, *29*, 238–247. [[CrossRef](#)]
31. Alexeeff, S.E.; Schwartz, J.; Kloog, I.; Chudnovsky, A.; Koutrakis, P.; Coull, B.A. Consequences of kriging and land use regression for PM<sub>2.5</sub> predictions in epidemiologic analyses: Insights into spatial variability using high-resolution satellite data. *J. Exp. Sci. Environ. Epidemiol.* **2015**, *25*, 138–144. [[CrossRef](#)]
32. Keller, J.P.; Chang, H.H.; Strickland, M.J.; Szpiro, A.A. Measurement error correction for predicted spatiotemporal air pollution exposures. *Epidemiology* **2017**, *28*, 338. [[CrossRef](#)]
33. Schneider, P.; Castell, N.; Vogt, M.; Dauge, F.R.; Lahoz, W.A.; Bartonova, A. Mapping urban air quality in near real-time using observations from low-cost sensors and model information. *Environ. Int.* **2017**, *106*, 234–247. [[CrossRef](#)]
34. Wu, C.D.; Zeng, Y.T.; Lung, S.C.C. A hybrid kriging/land-use regression model to assess PM<sub>2.5</sub> spatial-temporal variability. *Sci. Total Environ.* **2018**, *645*, 1456–1464. [[CrossRef](#)]
35. Berrocal, V.J.; Guan, Y.; Muyskens, A.; Wang, H.; Reich, B.J.; Mulholland, J.A.; Chang, H.H. A comparison of statistical and machine learning methods for creating national daily maps of ambient PM<sub>2.5</sub> concentration. *Atmos. Environ.* **2020**, *222*, 117130. [[CrossRef](#)]

36. Setton, E.; Marshall, J.D.; Brauer, M.; Lundquist, K.R.; Hystad, P.; Keller, P.; Cloutier-Fisher, D. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. *J. Exp. Sci. Environ. Epidemiol.* **2011**, *21*, 42–48. [[CrossRef](#)]
37. Gariazzo, C.; Pelliccioni, A.; Bolignano, A. A dynamic urban air pollution population exposure assessment study using model and population density data derived by mobile phone traffic. *Atmos. Environ.* **2016**, *131*, 289–300. [[CrossRef](#)]
38. Park, Y.M.; Kwan, M.P. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health Place* **2017**, *43*, 85–94. [[CrossRef](#)]
39. Kwan, M.P. The neighborhood effect averaging problem (NEAP): An elusive confounder of the neighborhood effect. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1841. [[CrossRef](#)]
40. Yu, H.; Russell, A.; Mulholland, J.; Huang, Z. Using cell phone location to assess misclassification errors in air pollution exposure estimation. *Environ. Pollut.* **2018**, *233*, 261–266. [[CrossRef](#)]
41. Finazzi, F.; Paci, L. Quantifying personal exposure to air pollution from smartphone-based location data. *Biometrics* **2019**, *75*, 1356–1366. [[CrossRef](#)]
42. Yoo, E.H.; Rudra, C.; Glasgow, M.; Mu, L. Geospatial estimation of individual exposure to air pollutants: Moving from static monitoring to activity-based dynamic exposure assessment. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 915–926. [[CrossRef](#)]
43. Chen, B.; Song, Y.; Jiang, T.; Chen, Z.; Huang, B.; Xu, B. Real-time estimation of population exposure to PM<sub>2.5</sub> using mobile-and station-based big data. *Int. J. Environ. Res. Public Health* **2018**, *15*, 573. [[CrossRef](#)]
44. Dias, D.; Tchepel, O. Spatial and temporal dynamics in air pollution exposure assessment. *Int. J. Environ. Res. Public Health* **2018**, *15*, 558.
45. Gurrain, S.; Stuart, A.L.; Pinjari, A.R. Impacts of travel activity and urbanicity on exposures to ambient oxides of nitrogen and on exposure disparities. *Air Qual. Atmos. Health* **2015**, *8*, 97–114. [[CrossRef](#)]
46. Dons, E.; Panis, L.I.; 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. [[CrossRef](#)]
47. Guo, H.; Zhan, Q.; Ho, H.C.; Yao, F.; Zhou, X.; Wu, J.; Li, W. Coupling mobile phone data with machine learning: How misclassification errors in ambient PM<sub>2.5</sub> exposure estimates are produced? *Sci. Total Environ.* **2020**, *745*, 141034. [[CrossRef](#)] [[PubMed](#)]
48. Dhondt, S.; Beckx, C.; Degraeuwe, B.; Lefebvre, W.; Kochan, B.; Bellemans, T.; Panis, L.I.; Macharis, C.; Putman, K. Health impact assessment of air pollution using a dynamic exposure profile: Implications for exposure and health impact estimates. *Environ. Impact Assess. Rev.* **2012**, *36*, 42–51. [[CrossRef](#)]
49. Lu, M.; Schmitz, O.; Vaartjes, I.; Karssenbergh, D. Activity-based air pollution exposure assessment: Differences between homemakers and cycling commuters. *Health Place* **2019**, *60*, 102233. [[CrossRef](#)]
50. Wu, Y.; Song, G. The Impact of Activity-Based Mobility Pattern on Assessing Fine-Grained Traffic-Induced Air Pollution Exposure. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3291. [[CrossRef](#)]
51. Yoo, E.H.; Roberts, J.E.; Eum, Y.; Shi, Y. Quality of hybrid location data drawn from GPS-enabled mobile phones: Does it matter? *Trans. GIS* **2020**, *24*, 462–482. [[CrossRef](#)]
52. Hoff, R.M.; Christopher, S.A. Remote sensing of particulate pollution from space: Have we reached the promised land? *J. Air Waste Manag. Assoc.* **2009**, *59*, 645–675. [[CrossRef](#)]
53. Wang, J.; Christopher, S.A. Intercomparison between satellite-derived aerosol optical thickness and PM<sub>2.5</sub> mass: Implications for air quality studies. *Geophys. Res. Lett.* **2003**, *30*. [[CrossRef](#)]
54. Di, Q.; Kloog, I.; Koutrakis, P.; Lyapustin, A.; Wang, Y.; Schwartz, J. Assessing PM<sub>2.5</sub> exposures with high spatiotemporal resolution across the continental United States. *Environ. Sci. Technol.* **2016**, *50*, 4712–4721. [[CrossRef](#)]
55. Hu, X.; Belle, J.H.; Meng, X.; Wildani, A.; Waller, L.A.; Strickland, M.J.; Liu, Y. Estimating PM<sub>2.5</sub> concentrations in the conterminous United States using the random forest approach. *Environ. Sci. Technol.* **2017**, *51*, 6936–6944. [[CrossRef](#)] [[PubMed](#)]
56. Stafoggia, M.; Bellander, T.; Bucci, S.; Davoli, M.; De Hoogh, K.; De Donato, F.; Gariazzo, C.; Lyapustin, A.; Michelozzi, P.; Renzi, M.; et al. Estimation of daily PM<sub>10</sub> and PM<sub>2.5</sub> concentrations in Italy, 2013–2015. using a spatiotemporal land-use random-forest model. *Environ. Int.* **2019**, *124*, 170–179. [[CrossRef](#)] [[PubMed](#)]
57. Pu, Q.; Yoo, E.H. Ground PM<sub>2.5</sub> prediction using imputed MAIAC AOD with uncertainty quantification. *Environ. Pollut.* **2021**, *274*, 116574. [[CrossRef](#)] [[PubMed](#)]
58. Gonzalez, M.C.; Hidalgo, C.A.; Barabasi, A.L. Understanding individual human mobility patterns. *Nature* **2008**, *453*, 779–782. [[CrossRef](#)] [[PubMed](#)]
59. Song, C.; Qu, Z.; Blumm, N.; Barabási, A.L. Limits of Predictability in Human Mobility. *Science* **2010**, *327*, 1018–1021. [[CrossRef](#)]
60. Pappalardo, L.; Simini, F.; Rinzivillo, S.; Pedreschi, D.; Giannotti, F.; Barabási, A.L. Returners and explorers dichotomy in human mobility. *Nat. Commun.* **2015**, *6*, 8166. [[CrossRef](#)]
61. Cressie, N. Spatial prediction and Ordinary Kriging. *Math. Geol.* **1988**, *20*, 405–421. [[CrossRef](#)]
62. Armstrong, M. *Basic Linear Geostatistics*; Springer: Berlin, Germany, 1998; p. 153.
63. Schabenberger, O.; Gotway, G. *Statistical Methods for Spatial Data Analysis*; CRC Press: Boca Raton, FL, USA, 2005.
64. Murray, N.; Chang, H.H.; Holmes, H.; Liu, Y. Combining satellite imagery and numerical model simulation to estimate ambient air pollution: An ensemble averaging approach. *arXiv* **2018**, arXiv:1802.03077.
65. Shtein, A.; Kloog, I.; Schwartz, J.; Silibello, C.; Michelozzi, P.; Gariazzo, C.; Viegi, G.; Forastiere, F.; Karnieli, A.; Just, A.C.; et al. Estimating daily PM<sub>2.5</sub> and PM<sub>10</sub> over Italy using an ensemble model. *Environ. Sci. Technol.* **2019**, *54*, 120–128. [[PubMed](#)]

66. Wong, D.W.; Yuan, L.; Perlin, S.A. Comparison of spatial interpolation methods for the estimation of air quality data. *J. Exp. Sci. Environ. Epidemiol.* **2004**, *14*, 404–415. [[CrossRef](#)]
67. Pearce, J.L.; Rathbun, S.L.; Aguilar-Villalobos, M.; Naeher, L.P. Characterizing the spatiotemporal variability of PM<sub>2.5</sub> in Cusco, Peru using kriging with external drift. *Atmos. Environ.* **2009**, *43*, 2060–2069. [[CrossRef](#)]
68. Branco, P.; Alvim-Ferraz, M.; Martins, F.; Sousa, S. The microenvironmental modelling approach to assess children's exposure to air pollution—A review. *Environ. Res.* **2014**, *135*, 317–332. [[CrossRef](#)]
69. Yamamoto, J.K. Correcting the smoothing effect of ordinary kriging estimates. *Math. Geol.* **2005**, *37*, 69–94. [[CrossRef](#)]
70. Department of Environmental Conservation. *Tonawanda Community Air Quality Study*; Department of Environmental Conservation: Albany, NY, USA, 2020.
71. Zhang, J.; Goodchild, M.F. *Uncertainty in Geographical Information*; Taylor & Francis: London, UK, 2002.
72. Jerrett, M.; Burnett, R.T.; Beckerman, B.S.; Turner, M.C.; Krewski, D.; Thurston, G.; Martin, R.V.; van Donkelaar, A.; Hughes, E.; Shi, Y.; et al. Spatial analysis of air pollution and mortality in California. *Am. J. Respir. Crit. Care Med.* **2013**, *188*, 593–599. [[CrossRef](#)] [[PubMed](#)]
73. Barbosa, H.; Barthelemy, M.; Ghoshal, G.; James, C.R.; Lenormand, M.; Louail, T.; Menezes, R.; Ramasco, J.J.; Simini, F.; Tomasini, M. Human mobility: Models and applications. *Phys. Rep.* **2018**, *734*, 1–74. [[CrossRef](#)]