Environmental Science & Technology

Article

pubs.acs.org/est

Comparison of Ultrafine Particle and Black Carbon Concentration Predictions from a Mobile and Short-Term Stationary Land-Use Regression Model

Jules Kerckhoffs,[*,†] Gerard Hoek,[†] Kyle P. Messier,[‡,§] Bert Brunekreef,[†,∥] Kees Meliefste,[†] Jochem O. Klompmaker,[†,⊥] and Roel Vermeulen[†]

[†]Institute for Risk Assessment Sciences (IRAS), Division of Environmental Epidemiology, Utrecht University, Utrecht 3584 CK, The Netherlands

[‡]Department of Civil, Architectural and Environmental Engineering, University of Texas, Austin, Texas 78712, United States

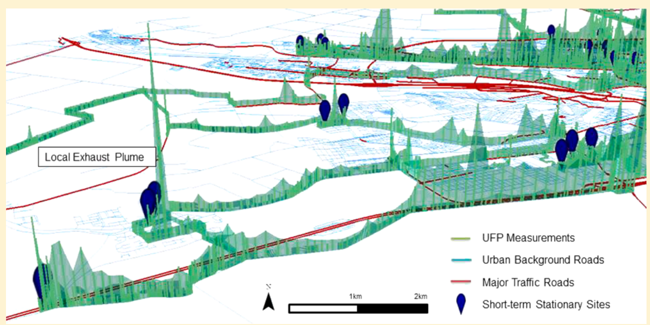
[§]Environmental Defense Fund, Austin, Texas 78701, United States

[∥]Julius Center for Health Sciences and Primary Care, University Medical Center, University of Utrecht, Utrecht, 3584 CK, The Netherlands

[⊥]National Institute for Public health and the Environment (RIVM), Bilthoven 3720 BA, The Netherlands

Supporting Information

ABSTRACT: Mobile and short-term monitoring campaigns are increasingly used to develop land-use regression (LUR) models for ultrafine particles (UFP) and black carbon (BC). It is not yet established whether LUR models based on mobile or short-term stationary measurements result in comparable models and concentration predictions. The goal of this paper is to compare LUR models based on stationary (30 min) and mobile UFP and BC measurements from a single campaign. An electric car collected both repeated stationary and mobile measurements in Amsterdam and Rotterdam, The Netherlands. A total of 2964 road segments and 161 stationary sites were sampled over two seasons. Our main comparison was based on predicted concentrations of the mobile and stationary monitoring LUR models at 12 682 residential addresses in Amsterdam. Predictor variables in the mobile and stationary LUR model were comparable, resulting in highly correlated predictions at external residential addresses (R^2 of 0.89 for UFP and 0.88 for BC). Mobile model predictions were, on average, 1.41 and 1.91 times higher than stationary model predictions for UFP and BC, respectively. LUR models based upon mobile and stationary monitoring predicted highly correlated UFP and BC concentration surfaces, but predicted concentrations based on mobile measurements were systematically higher.



1. INTRODUCTION

Multiple studies have shown negative relationships between outdoor particulate matter air pollution and health.[1] Both animal and human studies provide evidence for respiratory and cardiovascular effects associated with exposure to outdoor air pollution, with the ultrafine particle (UFP) fraction[2,3] and black carbon (BC)[4] as valuable indicators of pollution mixtures produced by local combustion sources.

Assessing spatial variation of UFP and BC can be challenging because these concentrations are highly variable in space and time, especially in urban environments.[5] Land-use regression (LUR) modeling has proven to be a useful tool to predict (long-term) intraurban spatial variation of outdoor air pollution.[6] LUR models for components such as nitrogen dioxide (NO_2), particles smaller than 2.5 μm ($PM_{2.5}$), and BC are usually based on 20−100 locations measured repeatedly over 1 or 2 week time periods.[6] For UFP, long-term sampling is complicated because instruments need frequent quality control and maintenance and are expensive compared to measurement instruments for NO_2 and $PM_{2.5}$.[7] Although long-term LUR models exist for UFP,[8] most monitoring campaigns are based on short-term measurements ranging from 15 min to 1 h per site[9−12] or true mobile monitoring with measurements obtained from a platform moving in traffic, often measuring 1 s to 1 min in interval.[13−17]

Short-term stationary and mobile monitoring reduces total measurement time significantly, allowing the measurement of a larger number of sites. However, they present challenges in separating spatial and temporal variability compared to study

Received: July 11, 2016
Revised: October 21, 2016
Accepted: November 3, 2016
Published: November 3, 2016

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902

designs with long-term fixed sites measured simultaneously. In mobile monitoring, this is complicated further by the even shorter time per location (typically road segments) compared to short-term monitoring. On-road monitoring may further not be representative for the modeling of residential exposures because dwellings can, depending on the urban topography, be several meters away from the roadside.

Previous studies have used either short-term stationary or mobile monitoring but not both in one study area. Short-term stationary monitoring campaigns for UFP and BC have been conducted in Vancouver, Canada;[10] Girona, Spain;[11] Amsterdam and Rotterdam, The Netherlands;[9] and New Delhi, India.[12] Typically, models explained a moderate amount of variation in measured concentrations (R^2: 30–50%). The moderate R^2 has been attributed to the large variability of short-term stationary measurements compared to campaigns using repeated 1–2 week measurement periods.[9] True mobile monitoring for UFP has been performed in Toronto and Montreal, Canada;[15−17] Somerville, MA;[14] and Minneapolis, MO.[7] Model R^2 was between 40 and 70%. Model R^2 can, however, not be easily compared between studies as some are spatial and some are spatiotemporal models, and sampling campaigns differ in duration and number of sites.

To date, there is no systematic comparison between LUR models of UFP and BC based on mobile and short-term stationary monitoring in the same study. We performed a measurement campaign in which we collected short-term stationary measurements (30 min) and measured in a mobile fashion in between the short-term stationary measurements using an electrical car. We previously showed that LUR models can be developed based on short-term stationary measurements in this region.[9,18] In this paper, we extend this work by developing LUR models based on mobile measurements and compare both methods in their ability to predict UFP and BC concentration surfaces of residences in the study area.

2. MATERIALS AND METHODS

2.1. Study Design. The monitoring campaign and site selection has been described previously.[18] Briefly, ultrafine particles were measured each second using a CPC 3007 (TSI Inc.). Black carbon was measured each minute using the micro Aethalometer (Aethlabs). Instruments were installed in the back of an electric vehicle (REVA, Mahindra Reva Electric Vehicles Pvt. Ltd., Bangalore, India), connected to inlets on the outside of the car. A Garmin eTrex Vista (Garmin Ltd. Kansas City, KS) global positioning system (GPS) measured the location of the electric vehicle in space and time. Sites were selected with a wide range of spatial contrast in traffic characteristics and land use.[9,18] Specifically, we selected traffic sites (>10 000 vehicles per day), urban background sites, sites near urban green, rural sites (outside the city), and sites near waterways.

Complete measurements of both air pollution and GPS were collected for 42 days in two seasons: winter and spring. A total of 2964 unique road segments were monitored between January 16th and May 22nd of 2013, taken by driving from one stationary measurement to the next. Measurements were performed between 9:15 AM and 16:00 PM to increase comparability between sites (avoiding rush hour). Of those 2964 road segments 745 (514) road segments had measurements of UFP and BC in both winter and spring. The abstract art shows a typical trajectory of the electric car.

2.2. Data Aggregation. Following previous mobile monitoring studies,[15,16] we averaged air pollution per road segment, defined as a part of a road between two consecutive intersections. Because all measurements were performed on the road, GPS points were snapped to the nearest road segment along the predetermined route to correct for small positional errors of the GPS. Road segments in tunnels or on bridges were deleted from the data set (n = 30) because they are not representative for residential streets. Road segments were on average 130 m long and comprised, on average, 12 s of UFP data. BC concentrations were based on 1 min values assigned to every road segment the car was on in that minute. On average, this means that three road segments were assigned the same BC measurement. A cutoff point of 10 000 vehicles per day was used to distinguish traffic and urban background road segments, following the definition for stationary sites.[9] Approximately 40% of road segments were considered as traffic and 60% as urban background road segments.

2.3. Air Pollution Data Processing. UFP data was removed if the concentration was below 500 particles per cm^3 and if the UFP concentration decreased or increased within a second by a factor 10, following the procedures of Klompmaker et al.[18] These criteria resulted in less than 1% of observations for short-term and mobile sites being removed. Concentrations lower than 500 particles per cm^3 were mostly 0, reflecting instrument malfunctioning.

For mobile monitoring, it is considered important to identify sampling events close to high-emission vehicles,[19] as these events can influence LUR model development. We defined samples influenced by local exhaust plumes if an UFP concentration was three standard deviations above the previous measurement second, determined based on the concentrations distribution for that day. Observations remained flagged until they dropped beneath the day average plus one standard deviation. This method is based on the method used by Drewnick et al.[20] For the main analyses, we used all measurements, including road segments with local exhaust plumes. For sensitivity analysis, we excluded them.

Our BC instrument generates minute averages, but this still can be a too short time period to produce reliable concentration levels because of a too-small change in attenuation (ATN). Data with a too-small ATN or negative ATN were corrected using the algorithm proposed by Hagler et al.[21] BC concentrations were only assigned when a threshold of 0.05 ATN change was fulfilled. Minute values with less than 0.05 ATN change were averaged over time until the criteria was met. The algorithm leads to significant noise reduction in our instrument while preserving high-resolution temporal variation. A total of 92% of the data has a 3 min time resolution or less. Only 20% of the data had a 1 min time resolution. Local exhaust plumes for BC were based on road segments for which UFP had a local exhaust plume, as 1 min BC measurements are too long to detect local concentration peaks.

2.4. Temporal Variation. A reference site with the same equipment as the electric vehicle was set up in Utrecht (about 30 km from Amsterdam and 50 km from Rotterdam), The Netherlands, to correct for temporal variation. We used the difference method for correcting the spatial data, following previous work in the stationary campaign.[18] First, the overall mean concentration of the entire campaign at the reference site was calculated. Next, for each minute at the reference site, an average of 30 min around time x was calculated that was subtracted from the overall mean concentration at the reference

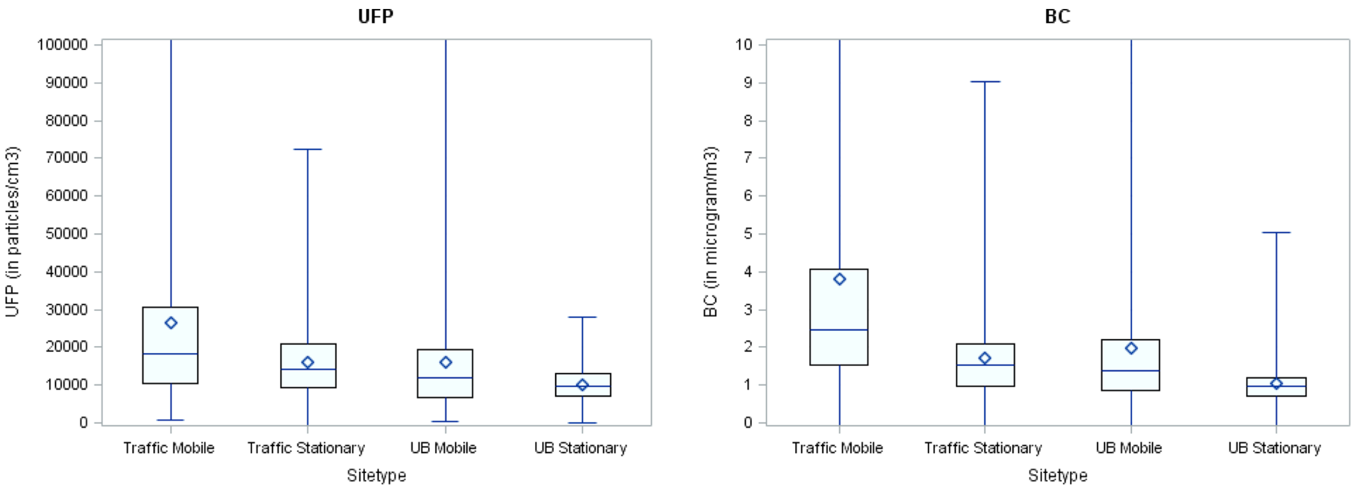


Figure 1. Distribution of average stationary (30 min) and road segment UFP and BC concentrations. Each box shows the median and the 25th and 75th percentiles. The diamond shape represents the group mean. Axes were truncated. There were 26 and 18 UFP observations higher than 100 000 p/cm^3 at traffic and background locations with a maximum of 221 243 p/cm^3. There were 57 and 32 BC observations higher than 10 μg/m^3 at traffic and background locations with a maximum of 43 μg/m^3.

site. The difference is then used to adjust the original concentration measured at the sampling locations.

2.5. Model Development. The midpoint of each road segment was identified and used as a coordinate for obtaining GIS predictors for LUR modeling. The average concentration adjusted for temporal variation of both pollutants per road segment was used as dependent variables in a linear regression analysis with multiple GIS variables as independent variables. We used the original concentration scale to predict arithmetic averages for epidemiological studies. We used data from all road segments, even when only one measurement was available. We observed similar standard errors of regression coefficients with this LUR modeling approach compared to using only the 745 (of 2964) road segments with at least two observations. GIS predictor variables were described previously[9] and summarized in Table S1. Briefly, a range of traffic variables was defined, including traffic intensity and road length variables (in 25 m to 1000 m buffers). Inverse distance to major roads was also used in the stationary model development but not in the mobile monitoring model, as this variable cannot be computed for major roads (distance zero for major road segments). Additionally, land use (e.g., port, industry, and urban green land) and population and household density in buffers from 100 to 5000 m were potential predictor variables. We additionally included airports as potential variable as recent studies found a 3- to 4-fold increase in UFP concentrations near airports.[22,23] These studies found elevated UFP concentrations up to 10 km from the airport, so for this variable we included a 5 and 10 km buffer as potential variable. The GIS variables were selected using a supervised stepwise selection procedure following our previous study.[9] The direction of the effect for the variables was determined a priori (Table S1), and the variable with the highest adjusted R^2 was entered in the model. Model building stopped when new variables were not able to improve the adjusted R^2. The variables in the resulting models were checked for p value (removed when p >0.10), collinearity (variance inflation factor >3 were removed), and influential observations (if Cook's D > 1, the model was further examined). We accounted for autocorrelation using a first order autoregressive (AR-1) term in the ARIMA procedure in SAS. Based upon the partial autocorrelation function, we determined that an AR-1 term was sufficient to characterize autocorrelation of the residuals. This correlation structure was also found to be most suitable in a mobile monitoring study by Farrell et al.[24] If after adding an AR-1 term to the identified model, variables were no longer significant (p > 0.10), they were removed from the model. We did not use universal Kriging to account for the autocorrelation of the data because in this method, the actual measured concentrations at the monitoring sites unduly influence model predictions. Measured concentrations are not precise due to the short duration of the measurements. The strength of mobile measurements is not the precision of individual observations but the amount of them.

For sensitivity analyses, we also developed models excluding observations flagged as influenced by local exhaust plumes. To test the impact of accounting for autocorrelation, we compared models with and without additional modeling of autocorrelation.

Because stationary models in the previous paper were based on three seasons and the true mobile monitoring models on two seasons, stationary LUR models were redeveloped with the same method as above, based on values in winter and spring and only used if both measurements were available. Montagne et al.[9] did not consider airports as potential variable, so we included the area of airports in a 5 and 10 km buffer in the new short-term stationary model.

2.6. Comparison between Short-Term Stationary and Mobile Monitoring. To compare stationary and mobile monitoring, four different analyses were performed: (1) stationary and mobile measurements were compared by identifying road segments adjacent to stationary sites. Measurement averages on these pseudo co-locations were compared using Pearson and Spearman correlation coefficients. For BC, this analysis was not performed because of the 1 min time resolution of the instrument. (2) Mobile LUR model predictions at the stationary measurement sites were compared with measured concentrations at these sites. (3) the stationary LUR model predictions were compared with concentrations measured at the mobile road segments. To perform this analysis, road segment coordinates with a value of zero for distance to nearest major road (mobile observations on a major road) were assigned a value of 10 m because inverse distance

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902

Table 1. Comparison between Average UFP and BC Concentrations of All Mobile and Short-Term Stationary Measurements

method		no. of observations	mean	5th percentile	median	95th percentile
stationary average[a]	UFP (in particles/cm^3)	128	12 630	5089	11 474	21 362
	BC (in μg/m^3)	141	1.35	0.56	1.13	2.67
mobile average[b]	UFP (in particles/cm^3)	2964	21 167	4391	15 057	59 628
	BC (in μg/m^3)	2336	2.83	0.48	1.79	8.41

[a]Stationary average consists of 2 × 30 min. [b]Mobile average consists, on average, of 18 s.

variables (predictor in the stationary models) could not be calculated otherwise. To check whether this number unduly affected the results, other distances (between 4 and 15m) were considered. (4) Both LUR models were used to predict concentration levels at residences using an external data set. GIS variables were used from a cohort study in Amsterdam, consisting of 12 682 residential addresses spread over the city, including urban background and traffic addresses with different land-use characteristics. We only used address information, and further details of this cohort can be found elsewhere.[25] The range of predictor variables was truncated to the range observed at the monitoring locations. Because application of the LUR models in epidemiological studies was the main goal of model development, this comparison was considered the central comparison between mobile and stationary monitoring models.

3. RESULTS

Figure 1 and Table 1 illustrate the variability of average UFP and BC concentrations from mobile and stationary monitoring, stratified by traffic and urban background (UB) sites. Mobile measurements were on average 1.7 times higher than stationary measurements for UFP (21 167 and 12 630 particles per cm^3 for mobile and stationary measurements, respectively). Black carbon concentrations collected on road were on average 2.1 times higher as stationary measurements, 2.83 and 1.35 μg/m^3, respectively. In particular, the higher percentiles of mobile measurements were about 3 times higher than the stationary measurements (2.8 for UFP and 3.1 for BC for the 95th percentile), likely related to the shorter averaging time of the mobile measurements. Mobile measurements were higher than the stationary measurements both for traffic and for urban background locations.

3.1. Mobile and Short-Term Stationary Monitoring LUR Models. The LUR models based upon mobile monitoring for UFP and BC are shown in Table 2. The model included predictor variables describing traffic (in small and large buffers) and population density for both pollutants. For UFP, airports and ports were also included in the model. Because models were developed including an AR-1 term, we cannot report standard R^2 values of our main models. Instead, the reported R^2 value is calculated by regressing the predicted concentration based on the parameter estimates of the mobile AR-1 model without the AR-1 term. R^2 values were low (0.13 and 0.12 for UFP and BC, respectively), reflecting the large temporal variability of the short duration measured concentrations. Models developed on road segments with at least two repeats (n = 745), increased the R^2 to 0.18 for UFP and 0.30 for BC (Table S2). Because the explained variance remains low, all further analyses are based on all road segments. Standard errors of the regression slopes were similar while losing about 75% of the data. We previously argued that a model with a low model R^2 can provide robust spatial models.[9,17]

When models were developed using only road segments without local exhaust plumes (n = 2907), models were very similar to the models including all road segments (Tables S3 and S5) with the exception that some small-scale traffic variables that were dropped from the UFP and BC models (traffic intensity on the nearest road). Not accounting for spatial autocorrelation changed the model only modestly (Tables S3 and S5). UFP and BC predictions based on the mobile models with and without AR-1 term and with and without local exhaust plumes were very highly correlated at the 12 682 addresses of the external data set (Tables S4 and S6), suggesting that these modeling choices do not affect the model predictions substantially.

Table 2. Land-Use Regression Models Based upon Mobile Measurements for UFP and BC

variables in LUR model	UFP (in particles/cm^3)[a]	BC (in μg/m^3)
intercept	5656 (2675)	0.48 (0.60)
population density in a 5000 m buffer	8064 (1947)	1.15 (0.48)
airport area in a 5000 m buffer	4669 (1185)	
port area in a 1000 m buffer	2499 (1248)	
nature area in 5000 m buffer	−2557 (1357)	
major road length in a 50 m buffer	6868 (1071)	0.61 (0.15)
traffic intensity on the nearest road		0.30 (0.14)
traffic load on major roads in a 100 m buffer	1928 (1095)	
traffic load in a 500 m buffer	2917 (1514)	
traffic load in a 1000 m buffer		0.88 (0.36)
R^2 of model	0.13[b]	0.12
no. of road segments used for model development	2964	2336

[a]Regression slopes and standard error (between parentheses), multiplied by the difference between 10th and 90th percentile for all predictors to allow comparison of the effect of predictors with different units and distribution on measured concentrations. [b]R^2 of AR-1 model without AR-1 term.

The stationary models developed based on two seasons (Table S7) were very similar to the models previously published based on three seasons[9] (Table S8). The stationary models contained fewer but similar predictor variables than the mobile monitoring models.

3.2. Comparison between Short-Term Stationary and Mobile Measurements. UFP measurements of both data sets were compared by creating a pseudo co-location of the stationary measurement. Of the possible 322 (161 × 2) stationary measurements, 184 also had a valid mobile measurement on the same road segment. Comparisons for UFP concentrations are given in Table 3 and Figure S1, showing a moderately high correlation between mobile and stationary measurements (R_p = 0.48). Correlations were higher for urban background than for traffic sites (R_p = 0.67 versus 0.39). Spearman correlations were higher than Pearson correlation values, indicating a nonlinear relation between stationary and mobile measurements. Mobile measurements at the pseudo co-location were on average 1.12 times higher than

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902

Table 3. Correlation between Mobile and Stationary UFP Measurements Sampled at the Same Road Segment

site type	number of observations	Pearson (Spearman)	median difference (particles/cm^3) and ratio
total	184	0.48 (0.70)	1674[a] (1.12)
traffic sites	100	0.39 (0.56)	2010 (1.12)
urban background sites	84	0.67 (0.81)	1450 (1.11)

[a]Difference is mobile minus stationary measurement.

stationary measurements, consistent with the comparison of all stationary sites and road segments (Table 1). However, the difference for the pseudo colocations is much lower than the overall difference. Mobile measurements were thus not only higher because of the distance to the road but also are related to the car using relatively busy roads to travel between stationary sites. As seen in Figure S1, the measurements were similar at the lower end of the concentration distributions and the difference increased at sites with high mobile measurements (hence, the higher Spearman correlation compared to the Pearson correlation).

3.3. Comparison between Mobile LUR Model Predictions and Short-Term Stationary Measurements. Predictions of the mobile LUR models were compared to the stationary measurements (Figure 2a). The mobile UFP model explained 26% of the spatial variability, two times more than the R^2 of the mobile model itself (Table 2). The mobile model explained the smaller variation of stationary measurements than the stationary model (Table S6) itself: 36%. For BC, the

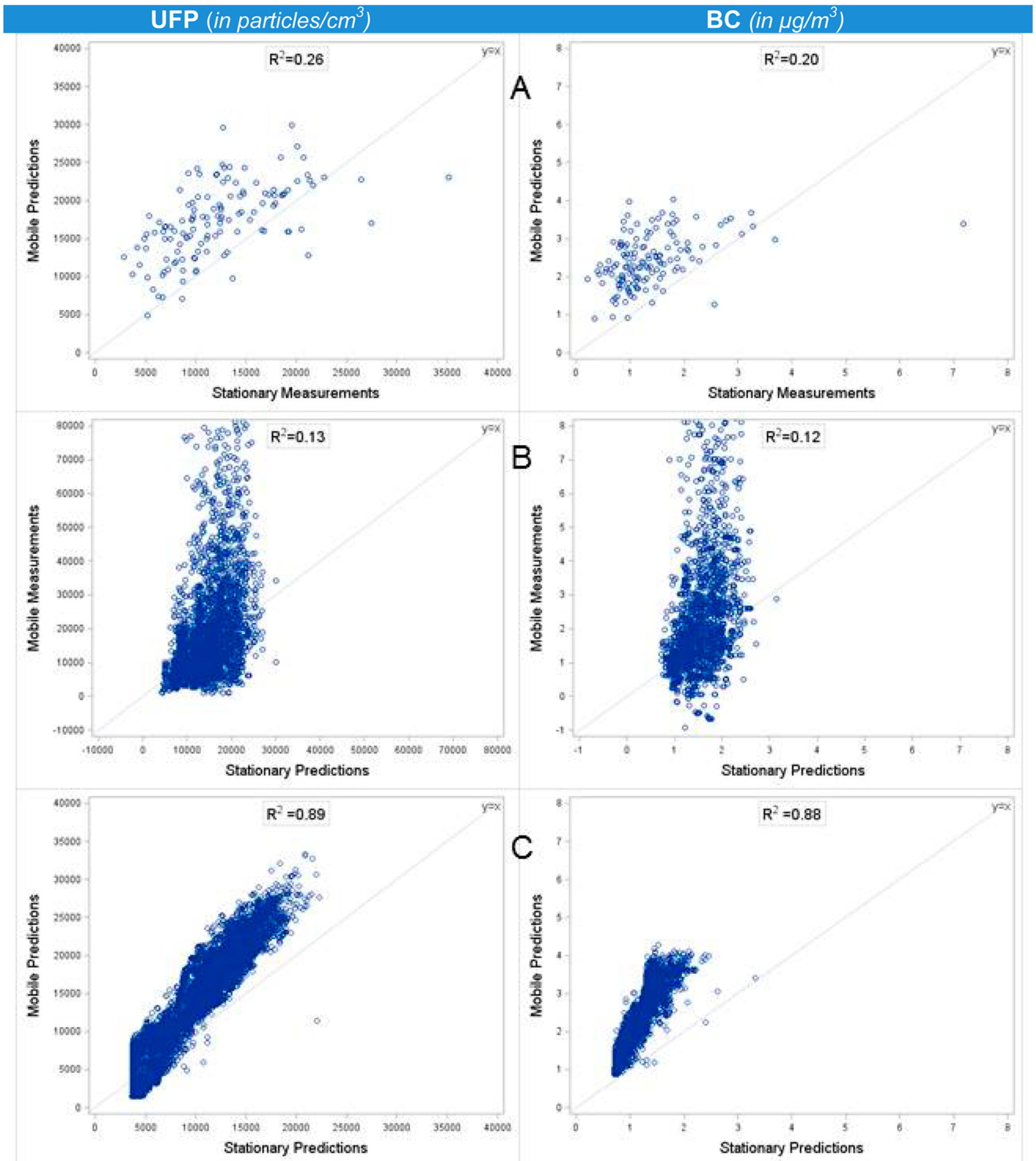


Figure 2. Comparison between stationary and mobile measurements and predictions for UFP and BC. (a) Predicted concentrations at stationary sites based on mobile model compared with stationary measurements. (b) Predicted concentrations at road segments based on the stationary model compared with mobile measurements. (c) Mobile vs stationary predicted UFP and BC concentrations at residences from an ongoing cohort study in Amsterdam consisting of 12 682 residential addresses.

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902

explained variance of stationary measurements was 20% for the mobile model (compared to 12% for the mobile BC model in Table 2) and 28% for stationary predictions, respectively. In this comparison, it should be noted that the stationary models were developed on the stationary measurement data. Both mobile models predicted higher concentrations than the measured concentrations at the short-term stationary sites. The mean difference was 4805 particles/cm^3 (95% CI: 3,336; 6,251) for UFP and 1.11 μg/m^3 (95% CI: 0.93; 1.28) for BC. Regressing the predicted UFP concentration based on the mobile measurements to the stationary measurements shows a slope of 0.70 (95% CI: 0.49; 0.91). For BC, the slope was 0.49 (95% CI: 0.32; 0.67).

3.4. Comparison between Mobile Measurements and Short-Term Stationary LUR Model Predictions. We also compared the stationary LUR model predictions to the mobile road segment average measurements, which resulted in a similar explained variance as for the mobile LUR models itself (R^2 0.13 versus 0.13 and 0.11 versus 0.12 for UFP and BC, respectively). Results are shown in Figure 2b. This comparison is somewhat hampered as an inverse distance is included in the stationary LUR model, and all mobile measurements were taken on road. We therefore set a minimum distance of 10 m to a major road and explored if varying this distance (4 to 15 m) resulted in marked differences. The results of these sensitivity analyses were essentially similar to the model assuming a minimum distance of 10 m.

3.5. Comparison between Mobile and Short-Term Stationary LUR Model Predictions on External Data Set. The developed stationary and mobile LUR models were used to predict concentration levels at residences from an ongoing cohort study in Amsterdam, consisting of 12 682 residential addresses. Predictions at these addresses were very highly correlated for both sets of LUR models (Figure 2c; R^2 of 0.89 for UFP and 0.88 for BC). Mobile models did, however, predict higher concentrations than stationary models, especially at higher concentrations. This difference is larger for BC (factor 1.91) than for UFP (factor 1.41) and reflective of the overall difference observed in measured concentrations in the stationary and mobile data sets (factor 2.1 and 1.7, respectively). Mean absolute difference was 4163 particles/cm^3 (95% CI: 4027; 4299) for UFP and 1.08 μg/m^3 (95% CI: 1.07; 1.10) for BC.

4. DISCUSSION

We developed LUR models based on mobile and short-term stationary measurements and compared both models in their ability to predict UFP and BC concentration levels at the stationary measurement locations and at the external data set with over 12 000 addresses in and around Amsterdam. Mobile monitoring LUR models included similar predictor variables compared to the short-term stationary models. Predictions of the mobile and stationary LUR models at residential addresses of the cohort study were very highly correlated (R^2 > 0.88), but predictions based on the mobile model were on average 1.41 (UFP) and 1.91 (BC) times higher than predictions of the stationary model.

4.1. Mobile versus Short-Term Stationary LUR Models. The variables in the mobile and stationary models were comparable. Both UFP models include the population density in a 5000 m buffer and port area, while the inverse distance to a major road variable in the stationary model is replaced by major road length in a 50 m buffer in the mobile model. Traffic intensity in a 100 m buffer is also very similar to the intensity of traffic on the nearest street, used in the mobile model. The mobile UFP model also included traffic load in a 500 m buffer and the area of airports in a 5000 m buffer. The presence of an airport in a 5 km buffer in the UFP model is consistent with recent studies evaluating the impact of airports on air pollutants.[22,23] These studies found elevated UFP (but not BC) concentrations up to 10 km downwind of airports. In agreement with these previous studies, our LUR model for BC did not include airports. Due to the similar structure of both models, the predicted UFP concentrations at over 12 000 addresses were highly correlated with an R^2 value of 0.89.

The mobile BC model predictions were also highly correlated with the stationary model predictions on the external data set (R^2 of 0.88). Variables in the BC mobile model were again similar to the stationary model variables. Here, population density and traffic intensity on the nearest road were present in both models, and the variable describing the inverse distance to major road in the stationary model was replaced by major road length in a 50 m buffer in the mobile model, as in the UFP model. These results indicate that both measurement approaches result in very similar model structures and highly correlated outdoor UFP and BC concentration estimates on a population level.

Explained variances of the mobile models were low and even lower than for the stationary models in this and our previous paper.[9] In a previous analysis, it was documented that the short stationary monitoring duration resulted in much higher within (temporal) to between (spatial) sites concentration variance ratios compared to monitoring campaigns with 1 or 2 week average samples.[9,18] Mobile monitoring involves even shorter monitoring per road segment and, hence, an even-less-favorable ratio between the ratio of concentration variance within to that between sites, resulting in low explained variances with the use of spatial predictor variables only. Consistently, the mobile UFP and BC models explained a larger percentage of the variation of the short-term stationary measurements than of the variability of the mobile measurements on which they were developed. We previously documented that, despite the low R^2, robust spatial models can be developed as the large number of sites in short-term stationary measurement campaigns compensates the low precision of the averages for each site.[9] The explanation offered was that measurement error in a continuous outcome variable in linear regression does not lead to biased regression slopes but does lead to lower R^2 values. In the present analyses, we extend this observation to mobile monitoring, showing that the mobile LUR models, on a population level, result in a similar rank-order of estimated outdoor concentrations, despite the fact that the mobile models systematically overestimated concentrations at residential addresses.

Consistent with the above reasoning and despite the low mobile model R^2 values, variables in the model were able to predict differences of up to 6997 particles/cm^3 for UFP and 1.15 μg/m^3 for BC between the 10th and 90th percentile of predictor variables. These are substantial contrasts, given the average urban background values of about 10 000 particles/cm^3 and 1.00 μg/m^3 for UFP and BC, respectively.

Some mobile and short-term monitoring studies reported higher model R^2 values than we report here. In a bicycle-based mobile monitoring study in Minneapolis, models explained about 50% of the particle number concentration (PNC) and 30−40% of the measured BC variability.[7] Averages consisted of

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902

12 to 30 runs on the same road segment, which is far more than our study in which most of the road segments were only sampled once. In studies by Weichenthal et al.[16,17] the average road segment consisted of 10 min of UFP data or consisted of at least 200 data points per segment, resulting in explained variances of 67% and 62%. Sabaliauskas et al.[15] only reported their R^2 of the mobile model compared with measurements at seven fixed sites to validate their model (R^2 of 0.68). Here, averaging time of UFP measurements was also between 5 and 10 min. Other differences between studies may be related to the overall variability of the measured concentrations, related to proximity to and magnitude of sources and the complexity of the study area.

4.2. Overestimation of Mobile Models. An average of 1.41 (UFP) and 1.91 (BC) times higher predicted values for mobile models compared to stationary models was found on residential addresses of the external data set. The overestimation is likely caused by the fact that mobile measurements are on-road, and the stationary measurements were taken at the sidewalks as close as possible to the facade of buildings. Although monitoring studies have documented large gradients of UFP and BC within meters of major roads,[3] our models contained buffer variables for traffic with radii of 50 to 1000 m, and these were unable to sufficiently explain differences between concentrations from on-road to residential addresses, typically located 5−20 m from the side of the road. We excluded inverse distance to major roads from modeling, as this variable cannot be calculated for on-road monitoring on major roads (distance zero). Assigning small distances to the road for on-road measurements was not an option, as the distance chosen is arbitrary and, if a small one is chosen, essentially results in a dummy variable for major roads versus nonmajor roads. In addition, it is unlikely that distance variables would be able to provide the needed scaling, given the limited precision of GIS calculated distances in compact urban areas.[8]

The higher overestimation for BC compared to UFP can possibly be ascribed to the coarser time resolution of the BC instrument (1 min versus 1 s for UFP) and the used data aggregation method. BC measurements were averaged over multiple minutes when the attenuation change was too low, while road segments with high concentration levels are more likely to remain on a 1 min resolution. Road segments with expected low concentrations are therefore assigned relative higher concentration levels. This method produces higher parameter estimates when modeling BC concentrations opposed to UFP. When UFP values were averaged with the same approach as the BC measurements (based on ATN values of the BC instrument), the absolute difference increased to 1.67 (Figure S2). Due to the precision of the instrument and data aggregation method, it seems that mobile monitoring while driving a car limits the application of our BC device. For mobile monitoring campaigns at lower driving speeds based, for example, upon cycling and upon walking in particular, the 1 min resolution may be appropriate. In study areas with higher concentrations than those observed in our study, the instrument will support shorter-duration measurements.

4.3. Mobile versus Stationary Monitoring. The advantage of mobile monitoring is that many locations can be measured in a relatively short amount of time with a limited number of monitoring devices. These locations include more-complex but realistic locations, such as near intersections.[9] An additional advantage is that little preparation is needed for site selection because the mobile platform does not need to be stationed anywhere for 30 min. However, selecting monitoring routes to cover relevant spatial settings and avoiding bias due to temporal variation remains important. We tried to minimize temporal bias by restricting to measurements outside the rush hours and by having a background reference site with identical equipment. Other approaches that could be taken is by smart-driving patterns in which several locations are revisited during the day or by having more than one platform driving at the same time.

Mobile monitoring may be affected by wind and vehicle speed effects, although the impact may be limited for submicrometer particles.[26] Figure S1 suggests this has not been a main issue in our study, as stationary and nearby mobile measurements did not differ substantially. Furthermore, we did not find a correlation between driving speed and measured concentrations in our data set.

One of the methodological challenges when using mobile data is to account for the inherent autocorrelation structure in the data. We used a first-order autoregressive model for the residuals, which assumes regular space and time intervals and that autocorrelation remains constant over time. However, as our measurements were not taken with the same lag-time in between road segments this method is unlikely to be optimal. We therefore performed sensitivity analyses in which we did not account for the autocorrelation. This indicated that the correlations on a population level between the models accounting or not accounting for autocorrelation are high: 0.99 for UFP and 0.97 for BC (Tables S4 and S6).

We aggregated the monitoring data to road segments and used the midpoint to obtain GIS predictor variables. This adds some uncertainty to the analysis, but in a study in Minneapolis, Hankey and co-workers found no difference in model performance for aggregation at 50, 100, or 200 m spatial resolution.[7]

The exclusion of road segments with local exhaust plumes for model development did not affect LUR models much in our study. An argument against removing local exhaust plumes of high-emitting vehicles is that busy road segments have a higher frequency of local exhaust plumes. In our modeling approach, we checked for the influence of potential outliers using Cook's D statistic. The large number of road segments in the data set probably prevented that the highest values influenced the developed model. Figure 1 illustrates that we measured few extreme concentrations. Consistently, only about 2% of our observations were flagged as local exhaust plume by our algorithm. In other study areas, this may be different.

An interesting option for model development is the combination of mobile and short-term stationary monitoring. This hybrid approach needs further work, likely involving definition of weights to take into account, the number of sites, and the time at each site and allowing for heterogeneity in variance structure. A further possible development is to build spatiotemporal models incorporating meteorology.

We focused our paper on the use of mobile and short-term stationary monitoring LUR models to characterize residential exposures for cohort studies. Mobile on-road monitoring models may also provide useful information with which to characterize commuter exposures. When coupled with time activity information, e.g., tracking by smartphone, individual exposure can be calculated by incorporating both residential and commuter exposures.[27]

4.4. Implications for Epidemiology. The spatial models that we developed may be useful for long-term exposure

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902

studies. On a population level, the predictions of the mobile and short-term model were highly correlated, implying that significant associations with health observed with one model would be detected with the other model as well. The overestimation of mobile models does not necessarily induce biased risk estimates in epidemiological studies if there is a systematic overestimation of the concentration predictions. Figure 2c, however, suggests that the relationship between the two model predictions is better described with a ratio than with a constant difference. Contrasts in exposure are higher for the mobile than for stationary-monitoring models, possibly leading to lower effect estimates per unit of exposure in epidemiological studies for the mobile models.

In addition, the overestimation of the mobile models was only modest: 30% for UFP and 50% for BC. Such differences can also be produced by measuring in one season only, excluding or including rush-hour traffic and measuring with different sampling devices. As such, the overestimation observed here is not likely to be an important factor. Our study thus suggests that mobile models can be used to predict exposures in epidemiological studies, taking into account that predictions on an absolute level may not reflect residential exposures fully.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.6b03476.

Definition of GIS predictors, sensitivity analyses for development of mobile and stationary LUR models, and stationary versus mobile UFP measurements on the same road segment. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: j.kerckhoffs@uu.nl. Phone: +31 30 253 9498.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was funded by the Environmental Defense Fund and The Netherlands Royal Academy of Sciences.

■ REFERENCES

(1) 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.; et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American heart association. *Circulation* **2010**, *121* (21), 2331−2378.

(2) Delfino, R. J.; Sioutas, C.; Malik, S. Potential role of ultrafine particles in associations between airborne particle mass and cardiovascular health. *Environ. Health Perspect.* **2005**, *113* (8), 934−946.

(3) HEI Review Panel. *Understanding the Health Effects of Ambient Ultrafine Particles*; Health Effects Institute: Boston, MA, 2013.

(4) Janssen, N. A. H.; Hoek, G.; Simic-Lawson, M.; Fischer, P.; van Bree, L.; ten Brink, H.; Keuken, M.; Atkinson, R. W.; Anderson, H. R.; Brunekreef, B.; et al. Black Carbon as an Additional Indicator of the Adverse Health Effects of Airborne Particles Compared with PM10 and PM2.5. *Environ. Health Perspect.* **2011**, *119* (12), 1691−1699.

(5) Zwack, L. M.; Hanna, S. R.; Spengler, J. D.; Levy, J. I. Using advanced dispersion models and mobile monitoring to characterize spatial patterns of ultrafine particles in an urban area. *Atmos. Environ.* **2011**, *45* (28), 4822−4829.

(6) Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* **2008**, *42* (33), 7561−7578.

(7) Hankey, S.; Marshall, J. D. Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, PM 2.5, Particle Size) Using Mobile Monitoring. *Environ. Sci. Technol.* **2015**, *49*, 9194−9202.

(8) Hoek, G.; Beelen, R.; Kos, G.; Dijkema, M.; van der Zee, S. C.; Fischer, P. H.; Brunekreef, B. Land use regression model for ultrafine particles in Amsterdam. *Environ. Sci. Technol.* **2011**, *45* (2), 622−628.

(9) Montagne, D. R.; Hoek, G.; Klompmaker, J. O.; Wang, M.; Meliefste, K.; Brunekreef, B. Land Use Regression Models for Ultrafine Particles and Black Carbon Based on Short-Term Monitoring Predict Past Spatial Variation. *Environ. Sci. Technol.* **2015**, *49* (14), 8712−8720.

(10) Abernethy, R. C.; Allen, R. W.; Mckendry, I. G.; Brauer, M. A Land Use Regression Model for Ultra fine Particles in Vancouver, Canada. *Environ. Sci. Technol.* **2013**, *47* (10), 5217−5225.

(11) Rivera, M.; Basagana, X.; Aguilera, I.; Agis, D.; Bouso, L.; Foraster, M.; Medina-Ramón, M.; Pey, J.; Künzli, N.; Hoek, G. Spatial distribution of ultrafine particles in urban settings: A land use regression model. *Atmos. Environ.* **2012**, *54*, 657−666.

(12) Saraswat, A.; Apte, J. S.; Kandlikar, M.; Brauer, M.; Henderson, S. B.; Marshall, J. D.; et al. Spatiotemporal land use regression models of fine, ultrafine, and black carbon particulate matter in New Delhi, India. *Environ. Sci. Technol.* **2013**, *47* (22), 12903−12911.

(13) Larson, T.; Henderson, S. B.; Brauer, M. Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environ. Sci. Technol.* **2009**, *43* (13), 4672−4678.

(14) Patton, A. P.; Collins, C.; Naumova, E. N.; Zamore, W.; Brugge, D.; Durant, J. L. An hourly regression model for ultrafine particles in a near-highway urban area. *Environ. Sci. Technol.* **2014**, *48* (6), 3272−3280.

(15) Sabaliauskas, K.; Jeong, C. H.; Yao, X.; Reali, C.; Sun, T.; Evans, G. J. Development of a land-use regression model for ultrafine particles in Toronto, Canada. *Atmos. Environ.* **2015**, *110*, 84−92.

(16) Weichenthal, S.; Van Ryswyk, K.; Goldstein, A.; Shekarrizfard, M.; Hatzopoulou, M. Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: A land use regression model. *Environ. Pollut.* **2016**, *208*, 241−248.

(17) Weichenthal, S.; Ryswyk, K.; Goldstein, A.; Bagg, S.; Shekkarizfard, M.; Hatzopoulou, M. A land use regression model for ambient ultrafine particles in Montreal, Canada: A comparison of linear regression and a machine learning approach. *Environ. Res.* **2016**, *146*, 65−72.

(18) Klompmaker, J.; Montagne, D.; Meliefste, K.; Hoek, G.; Brunekreef, B. Spatial variation of ultrafine particles and black carbon in two cities: results from a short-term measurement campaign. *Sci. Total Environ.* **2015**, *508*, 266−275.

(19) Brantley, H. L.; Hagler, G. S. W.; Kimbrough, E. S.; Williams, R. W.; Mukerjee, S.; Neas, L. M. Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. *Atmos. Meas. Tech.* **2014**, *7* (7), 2169−2183.

(20) Drewnick, F.; Böttger, T.; Von Der Weiden-Reinmüller, S. L.; Zorn, S. R.; Klimach, T.; Schneider, J.; Borrmann, S. Design of a mobile aerosol research laboratory and data processing tools for effective stationary and mobile field measurements. *Atmos. Meas. Tech.* **2012**, *5* (6), 1443−1457.

(21) Hagler, G. S. W.; Yelverton, T. L. B.; Vedantham, R.; Hansen, A. D. A.; Turner, J. R. Post-processing method to reduce noise while preserving high time resolution in aethalometer real-time black carbon data. *Aerosol Air Qual. Res.* **2011**, *11* (5), 539−546.

(22) Keuken, M. P.; Moerman, M.; Zandveld, P.; Henzing, J. S.; Hoek, G. Total and size-resolved particle number and black carbon concentrations in urban areas near Schiphol airport (the Netherlands). *Atmos. Environ.* **2015**, *104*, 132−142.

(23) Hudda, N.; Gould, T.; Hartin, K.; Larson, T. V.; Fruin, S. A. Emissions from an international airport increase particle number concentrations 4-fold at 10 km downwind. *Environ. Sci. Technol.* **2014**, *48* (12), 6628−6635.
(24) Farrell, W.; Weichenthal, S.; Goldberg, M.; Valois, M. F.; Shekarrizfard, M.; Hatzopoulou, M. Near roadway air pollution across a spatially extensive road and cycling network. *Environ. Pollut.* **2016**, *212*, 498−507.
(25) Van Eijsden, M.; Vrijkotte, T. G. M.; Gemke, R. J. B. J.; Van der Wal, M. F. Cohort profile: The Amsterdam born children and their development (ABCD) study. *Int. J. Epidemiol.* **2011**, *40* (5), 1176−1186.
(26) Morawska, L.; Salthammer, T. *Indoor Environment: Airborne Particles and Settled Dust*, Morawska, L., Salthammer, T., Eds.; Wiley: Hoboken, NJ, 2006.
(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.

DOI: 10.1021/acs.est.6b03476
Environ. Sci. Technol. 2016, 50, 12894−12902