Environmental Science & Technology

Article

pubs.acs.org/est

Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, $PM_{2.5}$, Particle Size) Using Mobile Monitoring

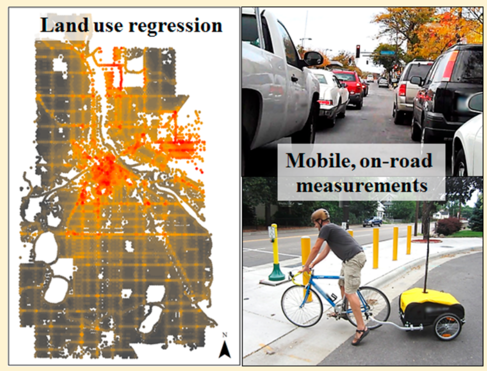
Steve Hankey*,† and Julian D. Marshall‡

†School of Public and International Affairs, Virginia Tech, 140 Otey Street, Blacksburg, Virginia 24061, United States

‡Department of Civil, Environmental, and Geo- Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, Minnesota 55455, United States

Supporting Information

ABSTRACT: Land Use Regression (LUR) models typically use fixed-site monitoring; here, we employ mobile monitoring as a cost-effective alternative for LUR development. We use bicycle-based, mobile measurements (~85 h) during rush-hour in Minneapolis, MN to build LUR models for particulate concentrations (particle number [PN], black carbon [BC], fine particulate matter [$PM_{2.5}$], particle size). We developed and examined 1224 separate LUR models by varying pollutant, time-of-day, and method of spatial and temporal smoothing of the time-series data. Our base-case LUR models had modest goodness-of-fit (adjusted R^2: ~0.5 [PN], ~0.4 [$PM_{2.5}$], 0.35 [BC], ~0.25 [particle size]), low bias (<4%) and absolute bias (2−18%), and included predictor variables that captured proximity to and density of emission sources. The spatial density of our measurements resulted in a large model-building data set (n = 1101 concentration estimates); ~25% of buffer variables were selected at spatial scales of <100m, suggesting that on-road particle concentrations change on small spatial scales. LUR model-R^2 improved as sampling runs were completed, with diminishing benefits after ~40 h of data collection. Spatial autocorrelation of model residuals indicated that models performed poorly where spatiotemporal resolution of emission sources (i.e., traffic congestion) was poor. Our findings suggest that LUR modeling from mobile measurements is possible, but that more work could usefully inform best practices.



1. INTRODUCTION

Land Use Regression (LUR) is an empirical approach to describing the spatial or spatiotemporal variability in air pollution concentrations.[1−4] Most LURs are built for one urban area, using field campaigns of long-term average concentrations at many (~20−200) locations, with the regression equation then used to estimate concentrations at locations without measurements. Urban scale models have been developed for cities in North America,[5,6] Europe,[7,8] South Asia,[9] East Asia,[11] and Australia,[12] and for national or international regions in North America,[13,14] Europe,[15,16] and Australia.[17] Several reviews summarize LUR and compare to other methods.[1−4,18,19]

A limitation of LUR is that input requirements are typically large;[20,21] for example, Hoek et al.[19] suggest a minimum of 40−80 sampling locations to properly specify an urban model. As such, most LUR models have been limited to pollutants for which inexpensive sampling is possible (most commonly NO or NO_2) or else suffer from potentially too few measurement sites for pollutants where sampling equipment is expensive (e.g., ultrafines, black carbon).[9,22,23] A variation of traditional LUR meant to reduce sampling input requirements is to rotate (short- and long-duration) fixed-site measurements and adjust for day-to-day differences in background concentrations with permanent fixed-sites.[9,10]

A potential alternative to traditional (i.e., fixed-site) LUR sampling is mobile monitoring. Mobile monitoring allows for measuring concentrations with good spatial coverage with a limited number of monitoring devices, but the researcher must control for temporal variability in concentrations, e.g., via separate models by time-of-day. A key advantage to mobile monitoring is the ability to investigate pollutants that are expensive to measure and for which LUR would be cost or labor prohibitive using traditional approaches (with devices currently available).

A limited number of studies have used mobile monitoring to explore spatial patterns of air pollution and build LUR models. Most studies use motor vehicles and are for specific sources (e.g., woodsmoke,[24−26] airports,[27] high-emitting vehicles)[28] or locations (e.g., freeways[29−31] or neighborhoods).[32] Some studies have measured air pollution on bicycle routes in

Received: March 9, 2015
Revised: June 30, 2015
Accepted: July 2, 2015
Published: July 2, 2015

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202

Table 1. Parameters Used To Generate Various Iterations of LUR Models

parameter	values for core models	values for sensitivity analyses
pollutant	particle number, black carbon, $PM_{2.5}$, particle size	(same)
time of day	morning, afternoon	(same)
statistical summary of concentrations at each aggregation location	mean, median	P10, P25, P75, P90
averaging time	lowest time average with sufficient model performance (60 s for BC, particle size; 1 s for PN, $PM_{2.5}$)	1, 10, 30, 60, 300 s
spatial resolution	100 m	50, 200 m
are concentrations (dependent variable) log-transformed?	yes	no

urban areas;[33−36] for the bicycle-based studies, regression models were developed to assess the impact of traffic levels and mix, but complete LUR models were usually not developed. Best practices for integrating mobile monitoring into LUR modeling do not exist.[37] We explore, for the first time, how mobile monitoring campaigns can be designed specifically for the purpose of building LUR models to make spatial estimates at locations without measurements for particulate air pollution.

In this paper, we develop LURs for on-road, rush-hour (7−9 am; 4−6 pm) particulate air pollution concentrations in Minneapolis, MN using bicycle-based, mobile monitoring. Our data set (3×10^5 individual measurements per instrument) allowed for development of many LUR models (n = 1224). We were able to assess important modeling choices (e.g., spatial and temporal smoothing) due to the spatial density of concentration estimates (n = 1101) tabulated from our mobile measurements. We explored how careful design of mobile measurement campaigns and data postprocessing to control for temporal variability impacts LUR model performance.

2. MATERIALS AND METHODS

2.1. Mobile Monitoring Data for LUR Development. We developed LUR models using mobile measurements of particulate air pollution.[38] Multiple aspects of particulates were measured using a bicycle-based sampling platform: particle number (PN) concentration (CPC 3007, TSI, Inc., Shoreview, MN), black carbon (BC) mass concentration (AE51, AethLabs, San Francisco, CA), fine particulate ($PM_{2.5}$) mass concentration (DustTrak 8320, TSI, Inc.), and particle size distributions (NanoScan, TSI Inc.). Mobile measurements were collected with the goal of LUR-development in mind; three sampling routes were selected to span various levels of traffic, types of land use, and cover a large number of neighborhoods. The routes were sampled repeatedly (thereby mitigating the effect of singular events during any one sampling run) during morning (7−9 am) and afternoon (4−6 pm) rush-hour (8/14−10/16/2012); reference measurements were collected at a central site before and after each sampling run to control for day-to-day variability in background concentrations among sampling days. In total, 42 sampling runs (~85 h of on-road measurements; 30 [12] runs in the afternoon [morning]) were completed among the sampling routes.

Measurements were corrected for sampling-artifacts and for temporal variability to assemble a consistent data set for spatial analysis. Specifically, we (1) collected reference-site measurements at a central site for 30 min before and after each sampling run and (2) developed pollutant-specific underwrite functions to estimate background concentrations during each sampling run. We adjusted our measurements for day-to-day differences in background concentrations using a 3-step process: (1) subtracting instantaneous background concentration estimates (via the underwrite function) from all instrument-reported concentrations, (2) calculating mean reference-site (i.e., background) measurements among all sampling days, and (3) adding the mean reference-site concentrations (from step 2) to the underwrite adjusted concentrations from step 1. This process was repeated separately for morning and afternoon sampling runs. See Hankey and Marshall[38] for details on the measurement campaign and data postprocessing. Summary statistics of the unadjusted and adjusted measurements are in Table S1 of the Supporting Information.

We performed sensitivity analyses to explore how various parameters (e.g., pollutant averaging time, spatial resolution of the aggregation of the mobile-measurements) impact LUR model performance. In total, we developed and examined 1224 separate LUR models; we discuss trends in model performance for these sensitivity analyses and describe how we chose base-case models for each pollutant (PN, BC, $PM_{2.5}$, particle size) and time-of-day (morning; afternoon).

2.2. LUR Dependent Variables: Spatial and Temporal Aspects of Data Aggregation. Few studies offer guidance on how best to aggregate and model mobile measurements using LUR.[37] Our approach is based on aggregating measurements at equal interval distances along the sampling routes. Then, we use estimates of concentrations from pooled data at each aggregation location to build LUR models for each pollutant. We develop separate models for morning and afternoon to account for temporal variability in the spatial patterns of concentrations.

2.2.1. Characteristics of the Measurement Distributions at Aggregation Locations. Our goal is to develop models for "typical" on-road exposure concentrations during morning and afternoon rush-hour; as such, our base-case models use estimates of central tendencies of the measurement distributions at each aggregation location. We also explored other distribution summaries to describe spatial patterns of "acute" (i.e., 75th and 90th percentiles) and "baseline" (i.e., 10th and 25th percentile) concentrations.

2.2.2. Temporal Smoothing of Mobile Measurement Data. We checked for differences in LUR model performance using various averaging times for the time-series concentration measurements (i.e., 1, 10, 30, 60, 300 s). We explore how temporal smoothing of mobile measurements may dampen spatial variability.

2.2.3. Distance between Aggregation Locations. We aggregated counts at varying distances (50, 100, 200 m) along the sampling routes to test how spatial resolution impacts model performance. As the distance between aggregation locations decreases, the spatial resolution increases (a benefit

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202

for LUR modeling) but the number of measurements per location decreases (a dis-benefit). Some aggregation locations using the 50 m version had small sample sizes (see Figures S2 and S3 of the Supporting Information). On the basis of the average block size in Minneapolis (~120 m) and requiring the median sample size at each aggregation location be sufficient to tabulate distributional parameters (>50 measurements per location), we chose 100 m as our base-case spatial resolution.

Parameters used to specify the dependent variable (i.e., pollutant concentration) in the LUR models are in Table 1. We primarily focused on three factors when selecting base-case models: (1) type of pollutant, (2) time of day (morning, afternoon), and (3) summary metric of the measurements at each aggregation location. Models employing variations among other parameters were used to test the sensitivity of the base-case models to other choices such as averaging time, spatial resolution, log transformation of concentrations, and modeling aspects other than central tendencies. We developed and examined 1224 separate LUR models to explore the effect of each parameter (i.e., 4 pollutants × 2 times of day × 6 distributional parameters × 5 averaging times × 3 spatial resolutions × 2 transformations = 1440 potential LUR models; because fewer averaging times [n = 2] were available for the NanoScan the total number of actual LUR models was 1224).

2.3. LUR Independent Variables. We assembled four categories of candidate independent variables: (1) traffic, (2) land use, (3) population dynamics, and (4) physical geography. Variables were either point estimates at a specific spatial location or (more commonly) based on buffers around a measurement location. We included 14 buffer variables (at 15 buffer lengths each) and 5 point variables, resulting in a total of 215 (i.e., 14 × 15 + 5) variables available for selection in each model (Table 2). Candidate independent variables were the same for all models; differences among models are in the dependent variables.

2.4. LUR Model Building Approach. We built LUR models using the stepwise regression technique given in Su et al.[39] Specifically, all 215 candidate independent variables are checked for strength of correlation with the dependent variable (pollutant concentrations); then, the independent variable most correlated with the dependent variable is added to the model. The regression is performed on the included variable and the remaining candidate variables are then tested for correlation with the model residuals; again, the most correlated variable is added to the model. This process repeats until the last added variable is either not significant (p > 0.05) or has Variance Inflation Factor (VIF; a check for multicollinearity with other independent variables) greater than 5. If a variable was selected at one buffer length we allowed the model to again select that variable at another buffer length. For each model, we calculated goodness-of-fit (adjusted R^2), mean absolute error, and mean absolute bias. We explored trends in these metrics when varying the dependent variable, as described above.

2.5. LUR Model Validation. We performed the following validation exercises:

Random Holdout Validation. We employed a Monte Carlo random hold out of 1/3 of the measurement data to use as a validation data set. We compared model estimates and observed concentrations (for the hold-out data set) by correlation to calculate validation R^2 values. We repeated this process 100 times for each model.

Systematic Validation by Sampling Route. We also applied a more rigorous, systematic validation approach, wherein we use two sampling routes for model-building, and the third route for model-testing. This process was repeated for each sampling route and each model.

Table 2. Independent Variables Included in Model Building

variable	units	buffer/point
traffic		
length of all roads	meters	buffer
length of freeways	meters	buffer
length of major roads	meters	buffer
length of local roads	meters	buffer
count of intersections	number	buffer
length of bus routes	meters	buffer
count of bus stops	number	buffer
traffic intensity	AADT m^{-2}	point
distance to freeway	meters	point
distance to major road	meters	point
land use		
industrial land use area	square meters	buffer
retail land use area	square meters	buffer
railway land use area	square meters	buffer
open space land use area[a]	square meters	buffer
population		
population density	people km^{-2}	buffer
housing unit density	house km^{-2}	buffer
median HH income	USD	buffer
physical geography		
elevation	meters	point
slope	grade (%)	point

buffers (meters): 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1500, 2000, 3000, 5000

[a]Includes parks, water, golf courses, and cemeteries.

Assessing Spatial Autocorrelation. We explored spatial autocorrelation among model residuals to assess overall model fit (using Moran's I) as well as any patterns in locations where model estimates were poor. To explore the local patterns of spatial autocorrelation, we used a Local Indicator of Spatial Analysis (LISA) developed by Anselin.[40]

3. RESULTS AND DISCUSSION

We developed a large number of LUR models (n = 1224) to assess how postprocessing of the mobile measurement data impacts model performance. Here we summarize key findings; detailed analyses are in the Supporting Information.

3.1. Model Results: Distributional Parameter at Aggregation Location. We used six distributional parameters at the aggregation locations (P10, P25, P75, P90, mean, median) as dependent variables in the LUR models. Models generally performed better for median than for mean concentration (see Figures S5−S8 of the Supporting Information). In some cases (for $PM_{2.5}$ and BC in the mornings), adjusted R^2 values were higher for the mean model than for the median model; however, in all cases absolute error and bias were either similar or lower for median than for mean concentrations. Thus, we chose to use median concentrations for the base-case models.

For the other distributional parameters, model performance (adj-R^2) was best for the lower ends of the distribution (i.e., 10th and 25th percentile) and worst at the upper ends of the distribution (i.e., 75th and 90th percentiles). This finding suggests that lower percentile (cleaner-air) conditions are more correlated with local land uses than higher percentile (dirtier-

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202

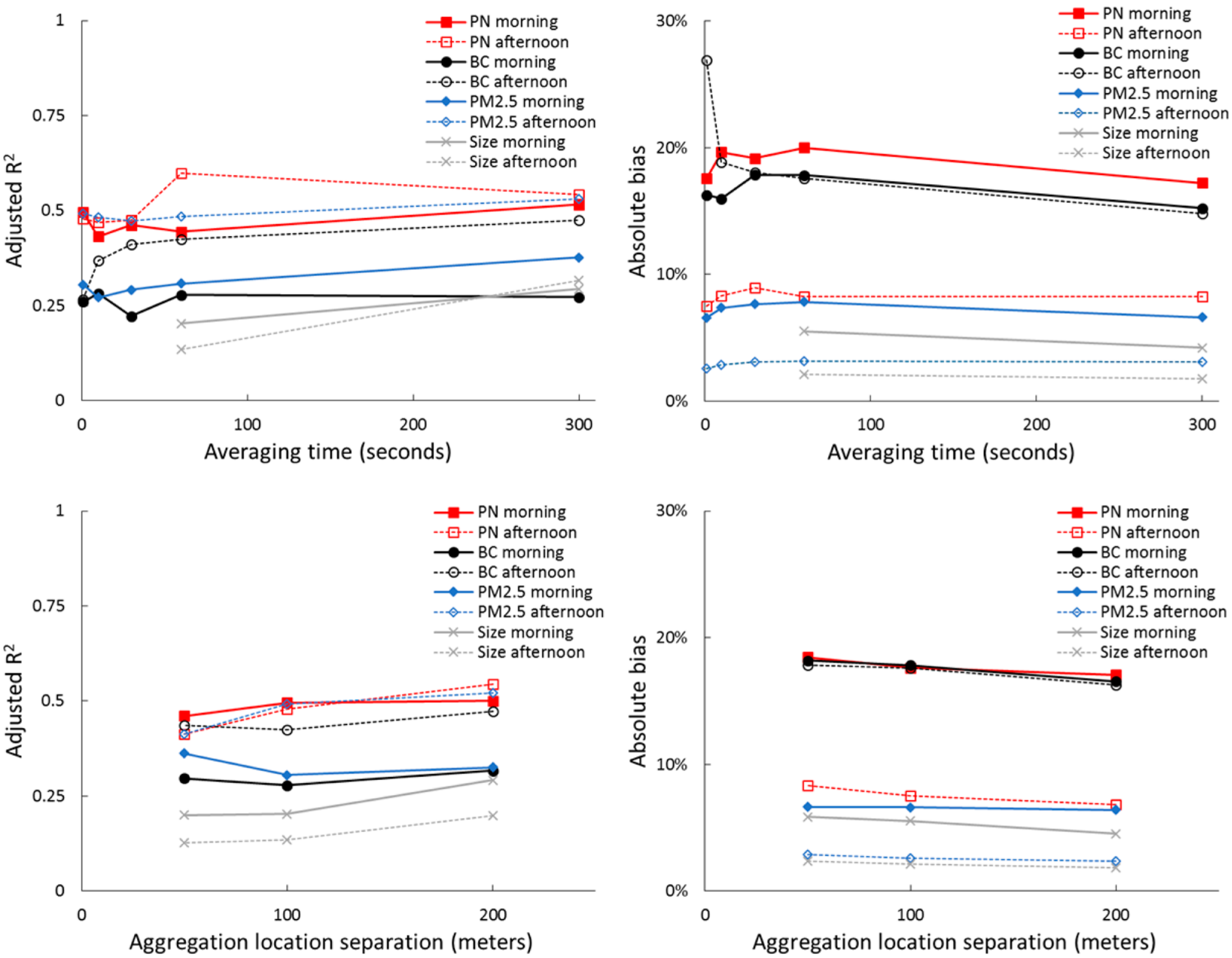


Figure 1. LUR model performance by averaging time and distance between aggregation locations.

air) conditions. Higher-percentile conditions may depend on idiosyncratic attributes (e.g., the specific types and distances of vehicles near the monitoring device).

3.2. Model Results: Temporal Smoothing and Spatial Aggregation. To explore the impacts to model performance of temporal smoothing, we varied the averaging time of the mobile measurements. A disadvantage to time-averaging is the corresponding spatial smoothing of the data; thus, our goal was to employ time-averaging only when doing so increased the LUR model performance. Model adjusted R^2 was mostly unchanged (slight increase) with greater time averages (absolute error and bias also were mostly unchanged). We chose small (or no) time-averaging to minimize spatial smoothing: for PN (1 s), $PM_{2.5}$ (1 s), and particle size (60 s). For BC, an averaging time of ~60 s was beneficial to model performance (see Figure 1); this improvement was likely attributable to smoothing inherent noise in the instrumentation (microaethalometers can have significant noise at small time scales, e.g., 1−30 s, especially for mobile monitoring).[41]

Model performance was generally better in the afternoons than mornings. This effect could be attributable to differences in the spatial patterns of pollutants during those times of day (e.g., lower background concentrations and spatial variability in the afternoons) or could be a result of fewer sampling runs in the morning (n = 12) than in the afternoon (n = 30). Model adjusted R^2 was highest for PN concentrations (adjusted R^2 ~ 0.5), moderate for both $PM_{2.5}$ and BC concentrations (adjusted R^2 ~ 0.3−0.5), and lowest for particle size (R^2 ~ 0.25).

As described above, we aggregated the mobile measurements at specific distances along the sampling routes. As a sensitivity analysis, we generated models for aggregation intervals of half (50 m) and double (200 m) the base-case, 100 m. We found (see Figure 1) little difference in model performance by the spatial resolution of aggregation locations. Figures S9−S12 of the Supporting Information summarize results by spatial resolution and time-averaging interval; in general, model performance was effected more by time-averaging than spatial resolution.

3.3. Base-Case LUR Models. We chose the spatial resolution of base-case models (i.e., 100 m aggregation) to ensure that aggregation locations provided sufficient sample sizes to estimate concentrations and by assessing changes in model performance under alternate spatial resolutions (50, 200 m). One exception is that, for particle size, we chose models based on the 200 m aggregation because the NanoScan's minimum time resolution is 1 min (a time that corresponds to ~260 m traveled at an average speed of 10 mph). As described above, we chose median (rather than mean) concentrations at aggregation locations, strived to minimize time-averaging when possible, and used log-transformed dependent variables. To

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202

Table 3. Base-Case Models Compared to "Best" Models As Measured by Adjusted R^2

model[a]	spatial resolution (m)	time average (s)	aggregation distribution parameter	R^2	Adj R^2	absolute error	bias (%)	absolute bias (%)	number of independent variables included
base-case models									
$PM_{2.5}$ morning	100	1	median	0.31	0.30	0.71 $\mu g\ m^{-3}$	<1	7	14
$PM_{2.5}$ afternoon	100	1	median	0.50	0.49	0.23 $\mu g\ m^{-3}$	<1	3	12
particle number morning	100	1	median	0.50	0.50	5290 pt cm^{-3}	3	18	18
particle number afternoon	100	1	median	0.48	0.48	1230 pt cm^{-3}	<1	8	7
black carbon morning	100	60	median	0.29	0.28	0.55 $\mu g\ m^{-3}$	3	18	16
black carbon afternoon	100	60	median	0.43	0.42	0.16 $\mu g\ m^{-3}$	3	18	15
particle size morning	200	60	median	0.30	0.29	1.97 nm	<1	5	8
particle size afternoon	200	60	median	0.22	0.20	0.72 nm	<1	2	11
"best" models									
$PM_{2.5}$ morning	200	60	mean	0.49	0.48	0.80 $\mu g\ m^{-3}$	<1	7	13
$PM_{2.5}$ afternoon	200	30	median	0.57	0.56	0.25 $\mu g\ m^{-3}$	<1	3	14
particle number morning	200	60	median	0.58	0.56	5570 pt cm^{-3}	2	17	19
particle number afternoon	100	60	median	0.61	0.60	1450 pt cm^{-3}	<1	8	26
black carbon morning	100	60	mean	0.35	0.35	0.58 $\mu g\ m^{-3}$	3	18	14
black carbon afternoon	200	60	median	0.49	0.47	0.15 $\mu g\ m^{-3}$	2	16	13
particle size morning	200	60	median	0.30	0.29	1.97 nm	<1	5	8
particle size afternoon	200	60	median	0.22	0.20	0.72 nm	<1	2	12

[a]All models shown here employ log-transformed concentrations as the dependent variable.

Table 4. Fully Normalized Coefficients and Buffer Sizes for the Base-Case LUR Models[a,b]

LUR model		industry	retail	open space	railway	bus routes	bus stops	traffic intensity	all roads	major roads	local roads	freeways	dist. to major roads	intersections	population density	house density	HH income	elevation
particle number	morning	0.24 (500)	0.15 (75)	-0.14 (25) -0.24 (1,500) 0.18 (5,000)	0.03 (100) 0.04 (400)	0.07 (100)	0.08 (100)	0.16		0.13 (25) -0.16 (750)	-0.08 (25) -0.38 (3,000)	-0.55 (3,000)		0.30 (2,000)		0.14 (500)		-0.12
	afternoon	0.35 (1,500)	0.15 (25)		0.05 (400)		0.13 (75)			0.14 (25) 0.21 (200)						0.24 (1,000)		
black carbon	morning	0.22 (400) 0.37 (3,000)	0.14 (75)	-0.16 (1,500) 0.29 (5,000)	0.09 (150)	0.12 (150)		0.13	-0.47 (3,000)	0.21 (25)	0.26 (5,000)	0.26 (1,000)	-0.11	-0.14 (300)		-0.10 (25) 0.54 (400)		
	afternoon	0.08 (75) 0.07 (1,000)	0.16 (25)	0.11 (3,000) 0.11 (5,000)	0.03 (100)	0.10 (75)				0.13 (25)	-0.37 (1,000)	0.15 (400) -0.32 (1,500) 0.13 (5,000)	-0.11			0.10 (50) -0.12 (1,000)		
$PM_{2.5}$	morning	0.08 (150) 0.19 (1,500)	0.07 (25)	-0.16 (25) 0.35 (3,000)	0.09 (150) 0.30 (3,000)		0.09 (100)	0.12		0.08 (75)				-0.17 (400) 0.21 (5,000)		0.31 (1,500)	-0.14 (25)	
	afternoon	0.16 (1,000)	0.22 (150)	0.12 (5,000)	0.09 (400)	0.10 (25)	0.21 (200)	0.06		0.14 (50)				-0.08 (150)	-0.20 (500)	0.09 (75)	-0.22 (500)	
particle size	morning	-0.08 (300)	-0.18 (100)	-0.21 (5,000)		0.30 (1,000)					0.25 (3,000)	-0.10 (100) -0.15 (400)					0.47 (3,000)	
	afternoon	-0.13 (500)	0.09 (25)		0.07 (500)	-0.52 (5,000)					0.30 (1,000) -0.31 (3,000)	-0.07 (150) 0.23 (2,000)		0.49 (5,000)		-0.32 (1,000)	-0.14 (100) 0.26 (1,500)	

[a]Here, all model coefficients were fully normalized by multiplying by the following factor: difference between 95th and 5th percentile independent variable values divided by the difference between 95th and 5th percentile pollutant concentration values. [b]Each entry in the table shows fully normalized coefficients with the corresponding buffer length in meters in parentheses.

evaluate our choice of base-case models, we also summarized models that performed best (based on adjusted R^2) for each of the eight model cases (i.e., each pollutant and time of day). Table 3 shows the 8 base-case models and the "best" models.

In general, base-case models performed similarly to the "best" models. Variables selected in the "best" models were similar to those in the base-case models. Adjusted R^2 was only slightly lower for the base-case models. However, absolute error, bias, and absolute bias were similar between the two groups of models; this result indicates that although model adjusted R^2 was slightly higher for the "best" models, error in concentration estimates was similar among the models. In most cases, the "best" models included more spatial (mostly 200 m) and temporal (mostly 60 s) smoothing of the data. Because error and bias were generally similar between the two groups of models, it is possible that the higher R^2 values in the "best" group is a result of a reduction in the variability of the model input concentrations from spatial and temporal smoothing. Absolute bias was low for all models (2−18%) and lower for pollutants with less spatial variability ($PM_{2.5}$ [∼5%], particle size [∼3%]) than for the traffic-related pollutants, which exhibited higher spatial variability (PN [∼13%], BC [∼18%]). Our base-case models had moderate R^2 values and small absolute bias; this result is perhaps explained by the limited spatial variability in our concentration estimates (especially for the afternoon). Repeating our methods in cities with more spatial variability in particulate concentrations may improve R^2 values, assuming those variations correlate well with land uses.

Table 4 shows independent variables (with corresponding buffer sizes) selected in each base-case model. To compare model results among pollutants we calculated fully normalized model coefficients by multiplying each coefficient by the following factor: difference between the 95th and 5th percentile independent variable value divided by the difference between the 95th and 5th percentile dependent variable value (i.e., pollutant concentration). The normalized model coefficients represent the number of 95th/5th percentile differences that pollutant concentrations will increase with each 95th/5th percentile difference increase in the corresponding independent variable. In most of the base-case models, a small number of the independent variables accounted for a large proportion of the variance explained. In six [two] of the eight models, the first

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202

two [three] variables accounted for over half of the final R^2 (Tables S2−S9 of the Supporting Information). We created plots of $\beta \times$ IQR vs the partial R^2 for each independent variable to visually depict how each independent variable contributed to the base-case models (Figures S13−S16 of the Supporting Information). Variables that were the most important (i.e., high $\beta \times$ IQR and high partial R^2) were those related to separation from emission sources (e.g., open space area, distance from major roads, length of low-traffic roads) and proximity to and density of emission sources (e.g., industrial area, house density, length of high-traffic roads). Figure 2 shows model estimates of afternoon rush-hour concentrations for the City of Minneapolis.

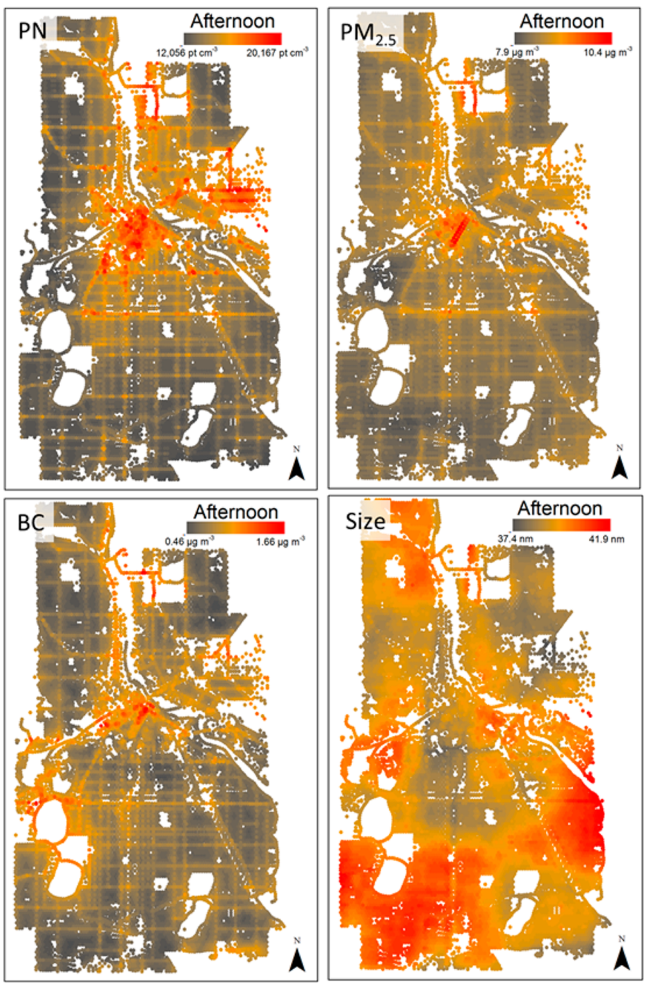


Figure 2. LUR model estimates for rush-hour in Minneapolis, MN.

A unique aspect of our approach to build LUR models is the density of concentration estimates in the study area (1101 locations here, versus 20−200 in a typical LUR). Among the base-case models, 23 variables (24% of the total) employ buffer lengths <100 m; the proportion was slightly higher for the traffic-related pollutants (BC [29%]; PN [30%]). These variables included length of major roads (n = 6), retail area (n = 6), house density (n = 3), and bus routes/stops (n = 3). Selection of these variables at a small spatial scale suggests that activity centers (e.g., origins [house density], destinations [retail areas]) and busy travel routes (e.g., major roads and bus routes/stops) are important to the small-scale variation in traffic-related pollutants.

Some variables switched direction at different buffer lengths. For example, when open space was selected at a buffer length of 1500 m or less, the coefficient was always negative and the magnitude was largest for the smaller buffer sizes (e.g., 25m). However, when open space was selected at a buffer length of >1500 m, the coefficients were positive. That result is likely attributable to confounding factors; for example, many of the off-street trails sampled (often colocated with open space) are old rail corridors, near freeways, or near activity centers. For Minneapolis, the open space variable (at large buffer sizes) may better capture these colocated emission sources than other variables. If we had forced variables to have a specific sign (e.g., open space must have a negative coefficient) we would not have uncovered this aspect of the data (yet the resulting model may be more applicable to other locations).

To test how our temporal adjustment for day-to-day differences in background concentrations (see Hankey and Marshall[38]) influenced model performance, we reran all base-case models using unadjusted concentrations. Mean adjusted R^2 decreased by 0.02 and mean error and absolute bias increased for all models (see Table S10 of the Supporting Information). We also reran all models using pooled measurements from morning and afternoon sampling runs. Model adjusted R^2 increased slightly (mean increase: 0.06); however, we chose to model each time period separately to preserve differences in spatial patterns of pollutant concentrations by time of day.

3.4. LUR Model Validation. In general, validation R^2 values (resulting from the Monte Carlo 1/3 hold-out analysis) were slightly lower than for the full models (average gap between model-building and model-testing R^2 is 0.07; for a 10% hold-out and the gap decreased slightly to 0.06), suggesting that goodness-of-fit statistics reported here may be slightly over-estimated (see Figure S17 of the Supporting Information).

Models were not robust to the systematic validation approach (i.e., using data from two routes to predict the third). R^2 values were low for predicted vs observed concentrations for nearly all validation models (see Figures S18−S25 of the Supporting Information). This finding is perhaps because of how we chose sampling routes. We strived to span the variable space for many predictor variables among all three routes. However, since sampling routes were completed in 2 h periods, we were not able to balance spanning all factors within each individual route. Thus, in this case the systematic validation approach may be an overly lofty criterion for an acceptable model based on our sampling method.

In general, model residuals showed moderate spatial autocorrelation using Moran's I (see Table S11 of the Supporting Information). As the search radius increased, Moran's I decreased suggesting that spatial autocorrelation is mostly an issue at small geographic scales. Most model residuals (~90%) were not flagged by the LISA procedure; among the flagged values, most were locations where the model under-estimated concentrations: at congested locations during rush-hour (e.g., high-traffic arterials near entrances to freeways; Figure S28−S29 of the Supporting Information). During sampling runs, we were sometimes unable to pass through the congestion at these locations and instead were forced to wait in idling traffic. Our models likely underestimate concentrations in the congested areas because temporally resolved spatial information on congestion was not available for

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202

use in the LUR models. Our traffic intensity metric, based on annual average daily traffic, does not capture hourly variations in traffic.

We tested whether bicycle speed could be used as a proxy for vehicle congestion by including mean bicycle speed as a candidate independent variable in the model building process. This approach did not change model performance (see Table S13 of the Supporting Information); adding this variable was likely not effective because bicycle speed is not only affected by motor vehicle congestion but also by other factors (e.g., hills; presence of a stop sign or traffic light).

3.5. Model Performance as a Function of Hours of Data Collected. We designed our study to capture "typical" concentrations by sampling the same routes repeatedly on different days. A question related to this approach is how many sampling runs are needed for concentration estimates to converge. To explore this issue we built LUR models separately depending on the number of sampling runs completed. We found that model adj-R^2 increased as the number of sampling runs completed increased. However, the magnitude of increase diminished as more sampling runs were completed (see Figure 3 for afternoon sampling runs; Figure S31 of the Supporting Information for morning sampling runs). This finding suggests that, for conditions considered here, additional sampling runs (beyond ~40 h in Figure 3) may only add incrementally to model performance. Further work is needed to replicate this finding in other urban areas.

3.6. Implications of Using Mobile Monitoring to Build LUR Models. There are a limited number of studies that use mobile monitoring to develop LUR models. Even fewer studies design measurement campaigns with the goal of using models to make estimates at all spatial locations in a study area. Table 5 summarizes prior studies by pollutant and model performance. Only three studies designed measurements with the goal of spatial extrapolation; all were for woodsmoke (two in Vancouver, Canada; one in upstate New York). Most studies collected measurements using motorized vehicles and many focused at least partly on freeways. Many of the studies included real-time traffic estimates as predictor variables; this approach precludes the possibility of extrapolating estimates to locations without automated traffic counts. Our models performed similarly to the other studies, which is reassuring given our limitations (nonmotorized vehicle; lack of real-time traffic data). Model R^2 values displayed in Table 5 reflect differences in reporting methods (R^2 versus adjusted R^2, cross-validated versus not) and suggest only modest performance: of the 20 R^2/adj-R^2 values listed in the table, only three are greater than 0.6.

A key advantage to mobile monitoring for LUR-development is the spatial density of measurements. Our data aggregation approach resulted in a large number of concentration estimates for modeling (i.e., n = 1101 for the base-case models; n = 550–2202 for the sensitivity analyses) relative to typical fixed-site LUR models (n ~ 20–200). Our approach allowed for two notable differences in how we assembled independent variables. First, our models included buffer sizes (25–100 m) that are smaller than is typical for LUR models making it possible to investigate small-scale changes in land use that may impact concentrations. Second, because our study area is for only one municipality (i.e., City of Minneapolis), we were able to assemble GIS data that has more detail than is typically available at larger scales (e.g., regional or national level models). For example, we were able to use better information on street

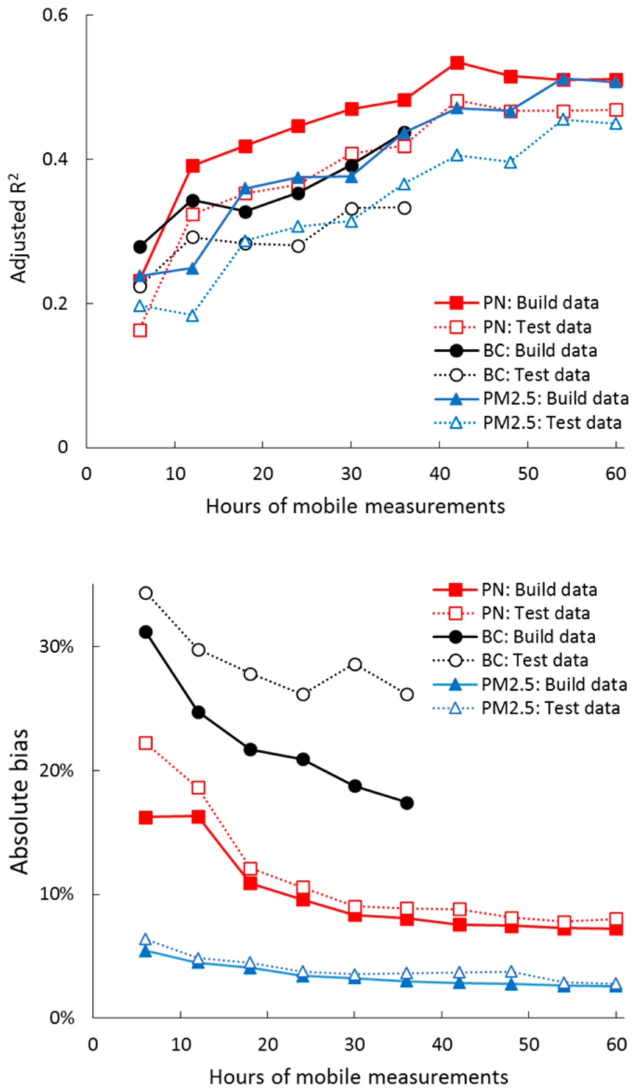


Figure 3. LUR model performance by hours of mobile measurements for afternoons. Included in both panels are values for the model building and validation (10% hold-out) data sets. Fewer sampling runs were completed for black carbon.

classification (freeway, arterial, collector, local) than is available from state or national databases (federal, state, municipal roads) and define specific land uses (e.g., retail area and industrial area) that may correlate with concentrations.

Merging temporal and spatial aspects of LUR modeling is an ongoing and worthwhile goal for explaining exposure patterns. We isolated specific times of day (i.e., morning and afternoon rush-hours) to control for hour-of-day differences in the spatial patterns of particulate concentrations. A limitation to modeling spatiotemporal patterns of pollutant concentrations is the necessary input data. For example, we used GIS based variables in our LUR models. We hypothesize that our models underestimate higher concentrations due to the poor spatiotemporal resolution of traffic-related input data (i.e., congestion patterns). This issue may be especially important for models that isolate specific times of day.[29] More work is needed to generate input data for LUR modeling that reflects the spatiotemporal nature of emission sources.

An important outstanding research question is whether short-duration (e.g., hourly) estimates of concentrations can be scaled to longer-term (e.g., annual-average) concentration

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194–9202

Table 5. Existing Studies Using Mobile Measurements To Develop LUR Models

study	pollutants	location	mode	geographic scale	for extrapolation[a]	regression	model R^2	notes
Larson et al. (2007)	particle light absorption coefficient	Vancouver, BC	motor vehicle	urban area	yes	linear	0.64	Measurements at night (winter). Model for emissions from woodsmoke.
Larson et al. (2009)	particle light absorption coefficient	Vancouver, BC	motor vehicle	urban area	yes	linear	0.56 (CV)	Same as Larson et al. (2007) but summertime measurements.
Su et al. (2011)	$PM_{2.5}$ (woodsmoke)	Upstate New York	motor vehicle	seven counties	yes	linear	0.58	Combined fixed-site and mobile monitoring. Sampling at night for woodsmoke.
Zwak et al. (2011)	PN $PM_{2.5}$	Brooklyn, NY	pedestrian	neighborhood	no	GAM	PN: 0.32 $PM_{2.5}$: 0.85	Walked neighborhood routes. Used real-time traffic counts.
Aggarwal et al. (2012)	PN	Minneapolis, MN	motor vehicle	urban area	no	linear	0.77 (adj-R^2)	Freeways only; used real-time traffic data.
Li et al. (2013)	PN $PM_{2.5}$ NO_x PB-PAH	Los Angeles, CA	hybrid vehicle	urban area	no	linear GAM	PN: 0.45 $PM_{2.5}$: 0.51 NO_x: 0.37 PB-PAH: 0.41	Good spatial coverage. Used real-time traffic data.
Patton et al. (2014)	PN	Somerville, MA	recreational vehicle	neighborhood (near freeway)	no	linear	0.38-0.47 (cross validated [CV])	Short sampling route near freeway. Included real-time traffic counts.
this study	PN BC $PM_{2.5}$ particle size	Minneapolis, MN	bicycle	urban area	yes	linear	Adj-R^2: morning (afternoon) PN: 0.50 (0.48) BC: 0.28 (0.42) $PM_{2.5}$: 0.30 (0.49) size: 0.29 (0.20)	Did not use real-time traffic inputs; goal of models is for extrapolation.

[a]Extrapolation refers to models that include predictor variables that allow for making spatial estimates at all locations in a study area for the time period studied.

estimates. For example, our LUR models were developed for specific time periods (7–9 am, 4–6 pm; late summer); however, most studies of air pollution and health use annual-average concentrations. A little-studied topic is what uncertainty exists with scaling concentration estimates and what number of mobile-monitoring measurements would be needed at a short-duration site to reliably scale to long-term averages (or from on-road concentrations to near-road concentrations).

Our measurement campaign was carried out in a relatively clean urban area. For some pollutants and times-of-day (i.e., afternoon $PM_{2.5}$ and BC) we observed limited spatial variability. This suggests that for certain combinations of pollutant, time-of-day, and urban area our approach may offer limited benefits over fixed-site monitoring. At the same time, our approach's ability to capture even small variability in concentrations is promising for using mobile measurements in LUR. Systematic evaluation of our approach in other cities, seasons, and times-of-day would help identify scenarios when our approach would offer the most benefit over fixed-site monitoring.

Our work suggests that LUR modeling from mobile measurements is possible; further investigation would help to determine best practices and make results more generalizable. Future research should aim to refine mobile measurement campaigns in systematic ways to tailor the data sets for use in LUR models.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional information as noted in text. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.5b01209.

■ AUTHOR INFORMATION

Corresponding Author

*S. Hankey. E-mail: hankey@vt.edu. Phone: 540-231-7508.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Jason Johnson (TSI, Inc.) for use of the NanoScan for this study, Jeff Blair (Aethlabs) for helpful guidance about the microaethalometer, and Dr. Josh Apte (UT Austin) for providing a program to postprocess the microaethalometer data.

■ REFERENCES

(1) Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology.* **2003**, *14* (2), 228–239.

(2) Briggs, D. J.; de Hoogh, C.; Gulliver, J.; Wills, J.; Elliott, P. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Sci. Total Environ.* **2000**, *253* (1–3), 151–167.

(3) Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D. A review and evaluation of intraurban air pollution exposure models. *J. Exposure Anal. Environ. Epidemiol.* **2005**, *15* (2), 185–204.

(4) Marshall, J. D.; Nethery, E.; Brauer, M. Within-urban variability in ambient air pollution: comparison of estimation methods. *Atmos. Environ.* **2008**, *42* (6), 1359–1369.

(5) Jerrett, M.; Arain, M. A.; Kanaroglou, B.; Beckerman, D.; Crouse, D.; Gilbert, N. L.; Brook, J. R.; Finkelstein, N.; Finkelstein, M. M. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. *J. Toxicol. Environ. Health, Part A* **2007**, *70* (3–4), 200–212.

(6) Ross, Z.; Jerrett, M.; Ito, K.; Tempalski, B.; Thurston, G. D. A land use regression for predicting fine particulate matter concentrations in the New York City region. *Atmos. Environ.* **2007**, *41* (11), 2255−2269.
(7) de Hoogh, K.; Wang, M.; Adam, M.; Badaloni, C.; Beelen, R.; Birk, M.; Cesaroni, G.; Cirach, M.; Declercq, C.; Dedele, A.; et al. Development of land use regression models for particle composition in twenty study areas in Europe. *Environ. Sci. Technol.* **2013**, *47* (11), 5778−5786.
(8) Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; Tsai, M.; Kunzli, N.; Schikowski, T.; Marcon, A.; et al. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe − the ESCAPE project. *Atmos. Environ.* **2013**, *72*, 10−23.
(9) Saraswat, A.; Apte, J. S.; Kandlikar, M.; Brauer, M.; Henderson, S. B.; Marshall, J. D. 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.
(10) Rivera, M.; Basagana, X.; Aguilera, I.; Agis, D.; Bouso, L.; Foraster, M.; Medina-Ramon, M.; Pey, J.; Kunzli, N.; Hoek, G. Spatial distribution of ultrafine particles in urban settings: a land use regression model. *Atmos. Environ.* **2012**, *54*, 657−666.
(11) Kashima, S.; Yorifuji, T.; Tsuda, T.; Doi, H. Application of land use regression to regulatory air quality data in Japan. *Sci. Total Environ.* **2009**, *407* (8), 3055−3062.
(12) Rose, N.; Cowie, C.; Gillett, R.; Marks, G. B. Validation of a spatiotemporal land use regression model incorporating fixed site monitors. *Environ. Sci. Technol.* **2011**, *45* (1), 294−299.
(13) Hystad, P.; Setton, E.; Cervantes-Larios, A.; Poplawski, K.; Deschenes, S.; Brauer, M.; van Donkelaar, A.; Lamsal, L.; Martin, R.; Jerrett, M.; et al. Creating national air pollution models for population exposure assessment in Canada. *Environ. Health Perspect.* **2011**, *119* (8), 1123−1129.
(14) Novotny, E. V.; Bechle, M. J.; Millet, D. B.; Marshall, J. D. National satellite-based land-use regression: NO2 in the United States. *Environ. Sci. Technol.* **2011**, *45* (10), 4407−4414.
(15) Gulliver, J.; de Hoogh, K.; Hansell, A.; Vienneau, D. Development and back-extrapolation of NO2 land use regression models for historic exposure assessment in Great Britain. *Environ. Sci. Technol.* **2013**, *47* (14), 7804−7811.
(16) Vienneau, D.; de Hoogh, K.; Bechle, M. J.; Beelen, R.; Donkelaar, A.; Martin, R. V.; Millet, D. B.; Hoek, G.; Marshall, J. D. Western European land use regression incorporating satellite- and ground-based measurements of NO_2 and PM10. *Environ. Sci. Technol.* **2013**, *47* (23), 13555−13564.
(17) Knibbs, L. D.; Hewson, M. G.; Bechle, M. J.; Marshall, J. D.; Barnett, A. G. A national satellite-based land-use regression model for air pollution exposure assessment in Australia. *Environ. Res.* **2014**, *135*, 204−211.
(18) Ryan, P. H.; LeMasters, G. K. A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicol.* **2007**, *19* (Suppl1), 127−133.
(19) 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.
(20) Karnaroglou, P. S.; Jerrett, M.; Morrison, J.; Beckerman, B.; Arain, M. A.; Gilbert, N. L.; Brook, J. R. Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. *Atmos. Environ.* **2005**, *39* (13), 2399−2409.
(21) Kumar, N. An optimal spatial sampling design for intra-urban population exposure assessment. *Atmos. Environ.* **2009**, *43* (5), 1153.
(22) Abernathy, R. C.; Allen, R. W.; McKendry, I. G.; Brauer, M. A land use regression model for ultrafine particles in Vancouver, Canada. *Environ. Sci. Technol.* **2013**, *47* (10), 5217−5225.
(23) Hoek, G.; Beelen, R.; Kos, G.; Dijkema, M.; Van Der Zee, S.; Fischer, P. H.; Brunekreef, B. Land use regression model for ultrafine particles in Amsterdam. *Environ. Sci. Technol.* **2011**, *45* (2), 622−628.
(24) 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.
(25) Larson, T.; Su, J.; Baribeau, A.; Buzzelli, M.; Setton, E.; Brauer, M. A spatial model of urban winter woodsmoke concentrations. *Environ. Sci. Technol.* **2007**, *41* (7), 2429−2436.
(26) Su, J. G.; Allen, G.; Miller, P. J.; Brauer, M. Spatial modeling of residential woodsmoke across a non-urban upstate New York region. *Air Qual., Atmos. Health* **2013**, *6* (1), 85−94.
(27) 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.
(28) Tan, Y.; Lipsky, E. M.; Saleh, R.; Robinson, A. L.; Presto, A. A. Characterizing the spatial variation of air pollutants and the contributions of high emitting vehicles in Pittsburgh, PA. *Environ. Sci. Technol.* **2014**, *48* (24), 14186−14194.
(29) Li, L.; Wu, J.; Hudda, N.; Sioutas, C.; Fruin, S. A.; Delfino, R. J. Modeling the concentrations of on-road air pollutants in Southern California. *Environ. Sci. Technol.* **2013**, *47* (16), 9291−9299.
(30) Aggarwal, S.; Jain, R.; Marshall, J. D. Real-time prediction of size-resolved ultrafine particulate matter on freeways. *Environ. Sci. Technol.* **2012**, *46* (4), 2234−2241.
(31) 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.
(32) Zwack, L. M.; Paciorek, C. J.; Spengler, J. D.; Levy, J. I. Modeling spatial patterns of traffic-related air pollutants in complex urban terrain. *Environ. Health Perspect.* **2011**, *119* (6), 852−859.
(33) Hatzopoulou, M.; Weichenthal, S.; Dugum, H.; Pickett, G.; Miranda-Moreno, L.; Kulka, R.; Andersen, R.; Goldberg, M. The impact of traffic volume, composition, and road geometry on personal air pollution exposures among cyclists in Montreal, Canada. *J. Exposure Sci. Environ. Epidemiol.* **2013**, *23* (1), 46−51.
(34) Hong, A. E.; Bae, C. Exposure of bicyclists to air pollution in Seattle, Washington: Hybrid analysis using personal monitoring and land use regression. *Transp. Res. Rec.* **2012**, *2270*, 59−66.
(35) Zuurbier, M.; Hoek, G.; Oldenwening, M.; Lenters, V.; Meliefste, K.; van den Hazel, P.; Brunekreef, B. Commuters' exposure to particulate matter air pollution is affected by mode of transport, fuel type, and route. *Environ. Health Perspect.* **2010**, *118* (6), 783−789.
(36) Van den Bossche, J.; Peters, J.; Verwaeren, J.; Botteldooren, D.; Theunis, J.; De Baets, B. Mobile monitoring for mapping spatial variation in urban air quality: development and validation of a methodology based on an extensive dataset. *Atmos. Environ.* **2015**, *105*, 148−161.
(37) Lightowlers, C.; Nelson, T.; Setton, E.; Keller, C. P. Determining the spatial scale for analyzing mobile measurements of air pollution. *Atmos. Environ.* **2008**, *42* (23), 5933−5937.
(38) Hankey, S.; Marshall, J. D. Exposure to on-road particulate air pollution (particle number, black carbon, $PM_{2.5}$, particle size) while cycling, in review.
(39) Su, J. G.; Jerrett, M.; Beckerman, B. A. A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. *Sci. Total Environ.* **2009**, *407* (12), 3890−3898.
(40) Anselin, L. Local Indicators of Spatial Association − LISA. *Geogr Anal* **1995**, *27* (2), 93−115.
(41) Apte, J. S.; Kirchstetter, T. W.; Reich, A. H.; Deshpande, S. J.; Kaushik, G.; Chel, A.; Marshall, J. D.; Nazaroff, W. W. Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Dehli, India. *Atmos. Environ.* **2011**, *45*, 4470−4480.

DOI: 10.1021/acs.est.5b01209
Environ. Sci. Technol. 2015, 49, 9194−9202