Environmental Science & Technology

Article

pubs.acs.org/est

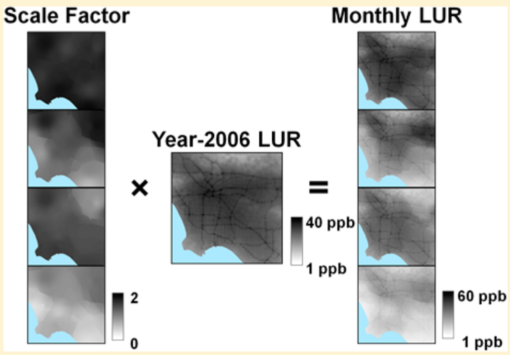
National Spatiotemporal Exposure Surface for NO_2: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–2010

Matthew J. Bechle,[†] Dylan B. Millet,[†,‡] and Julian D. Marshall*[,†]

[†]Department of Civil, Environmental, and Geo- Engineering and [‡]Department of Soil, Water, and Climate, University of Minnesota, Minneapolis, Minnesota 55455, United States

Supporting Information

ABSTRACT: Land-use regression (LUR) is widely used for estimating within-urban variability in air pollution. While LUR has recently been extended to national and continental scales, these models are typically for long-term averages. Here we present NO_2 surfaces for the continental United States with excellent spatial resolution (∼100 m) and monthly average concentrations for one decade. We investigate multiple potential data sources (e.g., satellite column and surface estimates, high- and standard-resolution satellite data, and a mechanistic model [WRF-Chem]), approaches to model building (e.g., one model for the whole country versus having separate models for urban and rural areas, monthly LURs versus temporal scaling of a spatial LUR), and spatial interpolation methods for temporal scaling factors (e.g., kriging versus inverse distance weighted). Our core approach uses NO_2 measurements from U.S. EPA monitors (2000–2010) to build a spatial LUR and to calculate spatially varying temporal scaling factors. The model captures 82% of the spatial and 76% of the temporal variability (population-weighted average) of monthly mean NO_2 concentrations from U.S. EPA monitors with low average bias (21%) and error (2.4 ppb). Model performance in absolute terms is similar near versus far from monitors, and in urban, suburban, and rural locations (mean absolute error 2–3 ppb); since low-density locations generally experience lower concentrations, model performance in relative terms is better near monitors than far from monitors (mean bias 3% versus 40%) and is better for urban and suburban locations (1–6%) than for rural locations (78%, reflecting the relatively clean conditions in many rural areas). During 2000–2010, population-weighted mean NO_2 exposure decreased 42% (1.0 ppb [∼5.2%] per year), from 23.2 ppb (year 2000) to 13.5 ppb (year 2010). We apply our approach to all U.S. Census blocks in the contiguous United States to provide 132 months of publicly available, high-resolution NO_2 concentration estimates.



1. INTRODUCTION

Nitrogen dioxide (NO_2) is a key component of urban air pollution generally associated with traffic-related emissions. Epidemiological studies have linked NO_2 to several adverse health outcomes, including premature mortality,[1−3] lung cancer,[4] and asthma exacerbations.[5,6] Predicting spatial and temporal variability in outdoor air pollution over large geographical areas has become an important goal for population exposure assessment, health studies, environmental justice, and public policy research. Regulatory monitors, which provide the basis for many investigations, can provide good temporal resolution, but generally are unable to capture within-urban variability in pollutant concentration owing to the limited number of monitors and their proximity.[7,8]

Land-use regression (LUR) has emerged as a useful tool for exploring within-urban variability in outdoor air pollution.[9] LUR is an empirical-statistical modeling approach that employs in situ measurements and geographic information system (GIS) variables to predict concentrations at nonmeasurement locations. The technique has been used extensively to assess within-city variability in outdoor air pollution, typically at the urban level.[7,10,11] Recent work has focused on using LUR and similar GIS techniques to model fine-scale air pollution concentrations over large geographic regions; we identified 21 such LUR studies (see Table S1). With recent improvements in data quality, including pollutant measurements from satellite instruments, LUR has been successfully applied at the continental scale for the United States,[12−14] Canada,[15] Europe,[16,17] and Australia.[18] Such models are useful for broad-scale cross-sectional exploration of air pollution concentrations and exposures.

Nevertheless, while small-scale LUR techniques have been studied extensively,[9] continental-scale LUR modeling requires further examination. Several studies have suggested 40–100 monitoring locations are necessary to obtain a robust LUR model; however, these results are based on LUR for urban areas and small countries rather than continental LUR.[9,19−21] Here we seek to test several aspects of continental LUR, with the

Received: June 13, 2015
Revised: September 22, 2015
Accepted: September 23, 2015
Published: September 23, 2015

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

goal of improving a prior national-scale spatial LUR model of annual-mean NO_2 for the United States.[13]

While the primary objective of LUR has generally been to improve the spatial resolution of air pollution estimates, improving the temporal resolution of LUR models has been a topic of recent interest.[22−32] Most large-scale studies we found (Table S1) focus on long-term (1−5 year mean) concentrations. A few studies estimate monthly[14,33−35] and daily[36−38] PM_{10} and $PM_{2.5}$ concentrations in the United States, one study estimates monthly NO_2 for Australia,[18] and one study estimates daily NO_2 for the northeastern United States.[39] Here we seek to develop an ex post facto approach, employing regulatory monitoring data to temporally scale estimates from an improved version of the prior national-scale spatial LUR model. This general approach has been used to extend the temporal coverage of urban-scale LUR[40−42] and national-scale satellite-based estimates,[43] but to our knowledge has not been employed for national-scale LUR models. An advantage to this approach is its low computational requirements; the alternative approach of LUR model building for each month is prohibitively computationally intensive, owing to the fact that geographic predictors need to be calculated (1) at additional monitor locations for each month and (2) for all Census blocks whenever a new variable is selected during model building. Moreover, our approach (using temporal trends derived from regulatory monitors) may be easily used to extend temporal coverage to similar long-term LUR models.

We seek to create spatial NO_2 surfaces in the continental United States that provide the excellent spatial resolution typical for urban-scale LURs (∼100 m scale), cover 100% of U.S. Census blocks, and provide monthly average concentrations for one decade. We focus on NO_2 because it has been shown to be a useful indicator of fresh combustion emissions, likely representing a mix of toxic species,[1,2] and because NO_2 concentrations have been shown to correlate with land use, making it a good candidate pollutant for new LUR models. Contributions of this paper to the literature include advancing LUR as a method by (1) testing model robustness to the number of training sites used for building a continental LUR, (2) comparing model sensitivity to inclusion of various regional NO_2 predictor variables (e.g., satellite column, satellite surface, high-resolution satellite, model surfaces from a mechanistic model), and (3) by presenting a large-scale ex post facto approach to obtain temporal estimates from an existing national-scale LUR.

2. METHODS

2.1. Spatial LUR Model. *2.1.1. Existing LUR Model.* We build on a previously developed year 2006 LUR for ground-level NO_2 for the contiguous United States.[13] Distinct features of the model relative to typical LURs included (1) broad geographic coverage, (2) inclusion of satellite-derived estimates of ground-level NO_2 and of regulatory monitoring data rather than de novo measurements, and (3) greater temporal coverage and precision. Model building incorporated six land-use characteristics (impervious surface, tree canopy, population, and major/minor/total road length) evaluated for 22 buffers (100 m to 10 km) and three point-based values (elevation, distance to coast, and satellite NO_2) (Table S2). Model formulation employed conventional stepwise forward regression.[44] This year 2006 annual mean model explained 78% of the spatial variability, with low mean bias (22% [overall], 5% [urban and suburban]). Model R^2 and internal leave-one-out validation may overestimate LUR performance; holdout cross-validation may provide better evaluation.[19,20] The base model is robust to holdout cross-validation (90% of monitoring data for model building, 10% for model testing; Monte Carlo random sampling (n = 500, median R^2 = 0.76)). Additional details for the prior LUR are available elsewhere.[13]

2.1.2. Regional NO_2 Covariates. Novotny et al. report a ∼0.1 increase in model R^2 when including satellite-based NO_2 as a predictor; other continental LURs report similar improvements (R^2 increase ∼0.02−0.15) .[15,17,18] Recent work suggests that satellite column abundance (total concentration within a vertical column) may be sufficient to track spatial patterns in ground-level NO_2,[8] in which case converting column abundance to surface concentrations may be unnecessary for LUR. Knibbs et al. reported LUR R^2 values of 0.81 when using column abundance, versus 0.79 with satellite-based surface estimates.[18]

Here, we compare model performance for LUR employing various regional NO_2 covariates. We consider column abundance and surface concentrations from two NO_2 algorithms for the Ozone Monitoring Instrument (OMI): (1) a global product (DOMINO, version 1.0.2, collection 3, http://www.temis.nl) and (2) a higher resolution U.S.-only product (BEHR, version 2.0A, http://behr.cchem.berkeley.edu/TheBEHRProduct.aspx).[45] For California and Nevada, NO_2 estimates were 30% lower for BEHR than for DOMINO over remote regions, but 8% higher (BEHR vs DOMINO) over polluted regions.[45] Surface concentrations are estimated from GEOS-Chem surface-to-column ratios, as in Novotny et al.[13]

We also include modeled NO_2 from a 12 km WRF-Chem chemical transport model (CTM) simulation for North America.[46] Finally, we include linear combinations of satellite + WRF-Chem NO_2 surfaces, calculated as the sum of standardized surfaces. We do this because when satellite and WRF-Chem NO_2 covariates are offered together during model building, both are selected into the final models. Standardized surfaces are obtained by dividing each grid estimated by the U.S. spatial mean concentration.

2.1.3. Alternative Models. Using relative measures (e.g., bias), the Novotny et al. model exhibits better model performance in urban and suburban (mean bias: 5%) locations than rural (71%) locations.[13] Moreover, Figure 1 (top panels) shows the Novotny et al. LUR NO_2 surface for the continental United States compared to the WRF-Chem and DOMINO NO_2 surfaces. This figure indicates broad agreement among methods, but marked by regional differences between the methods, including overestimation by the LUR of the low concentrations in the rural and remote Mountain West. Potential causes of the discrepancies include differences in the association between land use and NO_2 for urban/suburban versus for rural locations or monitor locations that do not adequately span the (independent) variable space. Regional discrepancies in Figure 1 generally correspond to elevation (e.g., the Rocky Mountains, California's Central Valley) and tree canopy (see Figure S1).

Monitor locations are not representative of elevations in the continental United States; the interquartile range (IQR) of the monitors' elevation is 0−300 m, yet 65% of the continental U.S. land area is >300 m. Additionally, the capacity of tree canopy as a sink for NO_2 is likely small;[47,48] however, tree canopy has a negative coefficient, and the magnitude of its effect ($\beta \times$ IQR = −0.91) is similar to that of major roadways ($\beta \times$ IQR = 0.97).[13] In the model, tree canopy likely represents a land use with little

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

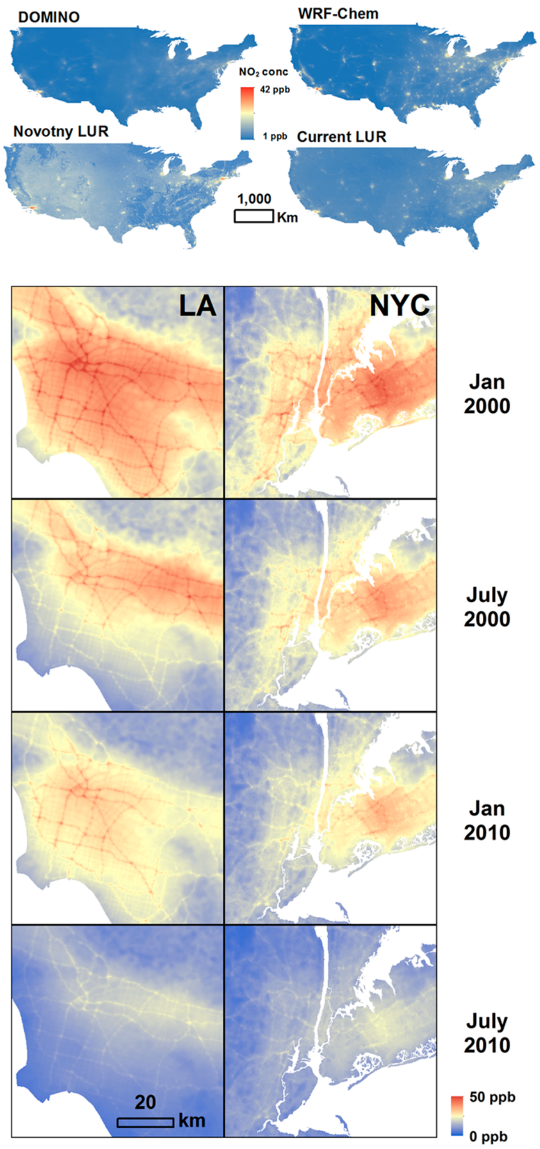


Figure 1. Top panels: ground-level NO_2 concentrations for the continental United States from 0.1° DOMINO satellite-derived estimates (top left), the 12 km WRF-Chem chemical transport model (top right), the Novotny et al. LUR model (bottom left), and the final spatial LUR model used here (bottom right). Bottom panels: modeled winter (January) and summer (July) ground-level monthly mean NO_2 concentrations for Los Angeles (LA) and New York (NYC) employing a 100 m grid for display purposes.

or no combustion sources in urban/suburban areas, but in rural areas this variable is effectively acting (incorrectly in the model) as a large sink for NO_2. Here, we address these discrepancies by (1) truncating elevation to the IQR (all values over 300 m are set to 300 m) and (2) removing tree canopy as a potential independent variable.

We attempt to further address poor model performance in rural areas by testing two alternative modeling approaches: (1) We employ the natural logarithm of NO_2 concentration as the dependent variable. Here, model building is as above, but with the dependent variable changed to the natural logarithm of year 2006 annual average NO_2 concentration. (2) We develop separate models for urban versus rural areas, by first subdividing EPA monitoring data into urban (plus suburban) versus rural.

2.1.4. Spatial Model Evaluation. We evaluate all spatial models on the basis of R^2, adjusted R^2, mean error, absolute error, mean bias, and absolute bias (eqs S1−S4). After selecting a final base spatial LUR model, as an additional sensitivity analysis, we use a Monte Carlo random sampling approach to explore model stability as a function of the number of training locations used for model building. Briefly, we conduct model building using random subsets of monitoring data (5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%), and evaluate the model's ability to predict concentrations at the remaining locations. For each subset size, performance metrics are calculated for 500 iterations.

2.2. Monthly NO_2 Surfaces. *2.2.1. Overview of Temporal Scaling.* We also expand on our prior work by incorporating an order of magnitude more data for the dependent variable (11 years instead of 1 year) and by developing an approach that reflects spatiotemporal variability. We start with 11 years (2000−2010) of hourly in situ NO_2 measurements for all U.S. EPA regulatory monitoring stations in the contiguous United States.[49] We calculate monthly mean concentrations for each monitor for months meeting the EPA reliability criterion of at least 75% of hourly values.

We calculate monthly scaling factors, which account for monthly deviations relative to the reference year, at each monitoring location using the following equation:

$$\mathrm{SF}_{im} = \frac{C_{im}}{\mathrm{LUR}_{2006}(x_i)} \tag{1}$$

Here, SF_{im} is the scaling factor for monitor i and month m, C_{im} is the monthly mean monitor concentration, and $\mathrm{LUR}_{2006}(x_i)$ is the year 2006 LUR model estimate at the location (x_i) of monitor i. We then create a scaling surface for each month (for a total of 132 months) by spatially interpolating the monthly scaling factors; we tested three interpolation techniques: (1) kriging (Bayesian kriging, ArcGIS 10.2), (2) inverse-distance weighting (IDW), and (3) nearest-neighbor (NN). (Further derivation of these interpolated scaling surfaces is given in the Supporting Information.) These monthly scaling surfaces are applied to the year 2006 LUR to create monthly NO_2 surfaces:

$$\mathrm{LUR}(x)_m = \mathrm{LUR}(x)_{2006} \times \mathrm{SS}(x)_m \tag{2}$$

where $\mathrm{LUR}(x)_m$ is the LUR model estimate at point x for month m, $\mathrm{LUR}(x)_{2006}$ is the year 2006 LUR model estimate at point x, and $\mathrm{SS}(x)_m$ is the (spatially varying) scaling surface value for month m at location x interpolated from the scaling factor (SF_{im}) values at monitor locations.

2.2.2. Monthly NO_2 Surface Evaluation. To determine the goodness of fit of our monthly NO_2 surfaces, we estimate the concentrations at each monitor location (i.e., excluding the monitor where the estimate is being made and using the remaining data to predict that at this location) and compare these estimates with the observed monitor concentrations. We evaluate model performance via spatial-only, temporal-only, and combined spatiotemporal comparisons. To determine spatial-only performance, we calculate R^2 statistics for each month (132 months) using monthly mean concentrations, and for each year (11 years) using annual mean concentrations. To determine annual and long-term temporal-only performance, we calculate R^2 statistics of monthly mean concentrations for each monitor location with at least 75% of months (at least 9 months for annual and 98 months for 11 years) of measurements. We determine long-term (132 months) and

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

Table 1. Final Model Using the Traditional LUR Approach

parameter	units	β	std error	$p > \|t\|$	partial R^2	IQR	$\beta \times$ IQR
intercept	ppb	2.44	0.41	<0.01			
DOMINO + WRF-Chem NO_2	unitless	0.72	0.036	<0.01	0.62	6.3	6.0
impervious (800 m)	%	0.085	0.0094	<0.01	0.75	43	4.9
elevation (truncated)	km	11.1	0.17	<0.01	0.76	0.27	1.9
major roads (800 m)	km	0.30	0.056	<0.01	0.78	3.2	0.91
residential roads (100 m)	km	2.82	1.04	0.01	0.78	0.27	0.77
distance to coast	km	-1.2×10^{-3}	3.8×10^{-4}	<0.01	0.79	630	−0.72

annual spatiotemporal performance using all monitor months. We report the mean error, absolute error, mean bias, and absolute bias (eqs S1−S4). We also investigate summary statistics by distance to the nearest monitor (<10, <25, 25−50, and >50 km), and by U.S. EPA-designated location type (urban, suburban, rural). To assess the performance of the model where there are people, we present population-weighted evaluations of the model performance, based on the U.S. Census population within a 1 km buffer of the monitor location. We assess within-city performance by calculating the above performance metrics for all monitors located within a single urban area, for the 10 urban areas with the most monitors. Monthly residuals were tested for Moran's *I* spatial autocorrelation using ArcGIS 10.2. Finally, as a sensitivity analysis, we compare the monthly NO_2 surface approach described above to (1) monthly LUR models for year 2006, (2) estimates from satellite-based measurements, and (3) a standard non-LUR approach: IDW interpolation of the monthly mean monitor values.

3. RESULTS

3.1. Spatial LUR Model. *3.1.1. Regional NO_2 Covariates.* We tested seven approaches for estimating regional NO_2 (DOMINO satellite column and surface estimates, BEHR satellite column and surface estimates, the 12 km WRF-Chem model, linear combinations of DOMINO and WRF-Chem, linear combinations of BEHR and WRF-Chem). We find that the continental LUR employed here is relatively insensitive to the choice of regional NO_2 covariate, but including no regional NO_2 covariate degrades performance (R^2 = 0.77−0.81 among models with regional NO_2, versus R^2 = 0.66 if no regional NO_2 variable is included). Models using satellite column NO_2 perform slightly better than those employing satellite-based surface estimates (see Table S3); this finding is consistent with a previous LUR of NO_2 for Australia[18] and suggests that the extra steps needed to estimate surface concentrations from satellite column measurements may be unnecessary for LUR models. We find that a model using a linear combination of satellite + WRF-Chem provides the best predictive power (R^2 = 0.81); we employ that regional NO_2 covariate (DOMINO + WRF-Chem) for analyses below.

3.1.2. Alternative Models and Final Model Selection. As mentioned above, we attempted to addresses the prior overestimation of low concentrations in the rural and remote mountain West by truncating the elevation variable at the IQR of monitor locations and by removing tree canopy as a land-use variable. Table 1 shows the resulting model. With these changes, the overall model performance diminishes slightly (R^2 = 0.79, mean absolute error 2.3 ppb, mean bias 18%, mean absolute bias 34%) and the performance at rural monitors is largely unchanged (see Table 2); however, Figure 1 (top panels) shows that the resulting model better captures the regional NO_2 patterns exhibited by the WRF-Chem and DOMINO estimates.

Table 2. Model Performance for Final Spatial LUR Model

	R^2	adj R^2	mean error (ppb)	mean abs error (ppb)	mean bias (%)	mean abs bias (%)
all	0.79	0.79	−0.30	2.3	18	34
urban	0.76	0.76	−0.78	2.4	−1	18
rural	0.50	0.49	0.39	2.4	57	75
population-weighted	0.81	0.81	−0.71	2.4	−1	17

Tables S4 and S5 show the two alternative LUR models (natural logarithm, urban + rural). Table S6 shows the model performance for the same two models and for our core LUR. Model performance is worse when using the natural logarithm of NO_2 (overall R^2 = 0.64 vs 0.79 for the traditional model), and similar (0.80) for the urban + rural model. Further investigation of the alternative models is given in the Supporting Information; overall, the alternative models were not strongly superior to the core model. We employ the traditional LUR as the base spatial model for the remainder of the analysis.

Figure 2 shows the model performance (median and IQR R^2 among n = 500 Monte Carlo iterations per comparison) as a function of the number of monitor locations used in model building. Figure S2 shows similar plots for absolute error and bias. Our final spatial LUR (derived from n = 369 monitoring

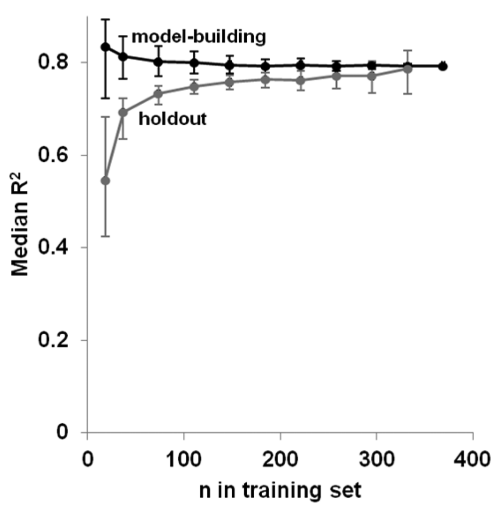


Figure 2. Median and interquartile range R^2 for Monte Carlo random sampling for n training monitors employed in model building out of 369 possible monitor locations. Above ∼150−200 monitors, the model building R^2 is stable and is consistent with the holdout R^2.

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

Table 3. Summary of Monthly Mean NO_2 Estimates Using Kriging Temporal Scaling[a] (2000−2010)

	mean spatial R^2	mean temporal R^2	spatiotemporal R^2	mean concn (ppb)	mean error (ppb)	mean abs error (ppb)	mean bias (%)	mean abs bias (%)	mean (IQR[b]) distance to nearest monitor (km)
all									
population-weighted	0.82	0.76	0.85	19.6	−0.3	2.7	2	18	32
unweighted	0.81	0.73	0.84	12.4	−0.05	2.4	21	38	48 (13−50)
distance to nearest monitor									
<10 k	0.80	0.72	0.82	16.3	−0.5	2.6	3	22	6 (4−8)
<25 km	0.82	0.73	0.84	15.4	−0.1	2.5	12	28	13 (8−18)
25−50 km	0.70	0.79	0.77	10.6	−0.06	2.1	23	40	36 (29−40)
>50 km	0.71	0.69	0.75	9.5	0.1	2.2	40	57	133 (65−158)
urban classification									
urban	0.76	0.76	0.80	16.2	−0.5	2.7	1	20	12 (7−16)
suburban	0.76	0.76	0.81	13.6	−0.02	2.4	6	22	30 (21−38)
rural	0.63	0.63	0.69	6.0	0.4	2.0	78	95	104 (36−136)

[a]All values except those in the first row are unweighted metrics. [b]IQR is the interquartile range.

locations) is robust to large holdout evaluations; for example, median R^2 values for model building vs holdout data sets are 0.79 vs 0.77 with 20% holdout (n = 295 monitors in model building) and 0.79 vs 0.76 with 50% holdout (n = 184 monitors in model building). For cases with more than ∼150−200 monitors in model building, the model building R^2 is stable and is similar to the holdout R^2. This suggests that large-scale LUR may require more locations than the ∼40−100 locations suggested for smaller scale LUR[19−21] to obtain a robust model. The IQR of the holdout R^2 is ∼0.1 for holdout subsets of 10−20%, suggesting a need to use large holdout subsets (here, ∼50%) or systematic holdouts (e.g., 10-fold 10% holdout) or to conduct Monte Carlo approaches so model evaluation is not effected by unusual holdout subsets.

3.2. Monthly NO_2 Surfaces. On the basis of EPA data for 11 years, on average 370 (IQR = 362−381) monitors meet the reliability criteria each month. The average (IQR) monthly mean NO_2 concentration (ppb) during the 11 years is 12.4 (9.9−14.6), and decreased from 15.2 (13.1−17.6) in year 2000 to 9.3 (7.2−11.4) in year 2010; population-weighted values have a mean of 19.5 (IQR: 16.1−22.3) for the 11 years, reflecting a 42% decrease overall, from 23.2 (21.1−25.0) in year 2000 to 13.5 (10.6−16.3) in year 2010. Mean concentrations over the 11 years are higher for urban (16.2) and suburban (13.6) locations than for rural (6.0) monitors. NO_2 concentration decreases observed here (for unweighted concentrations: 0.59 ppb y^{-1}, ∼4.7% y^{-1}; for population-weighted concentrations: 1.0 ppb y^{-1}, 5.2% y^{-1}) correspond to significant reductions in NO_x emissions throughout the United States.[50] The long-term temporal trend varies spatially (IQR: 29−42% [3.3−7.3 ppb] decrease in annual average NO_2 concentration), and includes increasing concentrations at a small number of rural locations (5% of monitors). These trends are consistent with findings from satellite-based measurements, and support the use of a spatially varying temporal scaling surface.[51,53] Table 3 and Tables S7 and S8 summarize monthly NO_2 surface estimates for the three temporal scaling techniques. All three methods perform similarly well, with the kriging method performing slightly better than IDW and NN. The resulting NO_2 surfaces (kriging, IDW, NN) predict, on a population-weighted average, 82% (kriging), 79% (IDW), and 80% (NN) of the spatial-only variability, 76%, 73%, and 74% of the annual temporal-only variability, and 85%, 84%, and 84% of the combined spatiotemporal variability in U.S. EPA-monitored NO_2 concentrations. (Unweighted R^2 values are 0.81, 0.78, and 0.79 [spatial only], 0.73, 0.69, and 0.70 [temporal only], and 0.84, 0.81, and 0.82 [spatiotemporal].) From here forward, we will focus primarily on the monthly NO_2 estimates employing kriging temporal scaling. The model mean bias and absolute error are low (21% and 2.4 ppb, respectively), consistent with the previously published year 2006 spatial LUR model. For illustrative purposes, we apply the final national NO_2 surface estimates to a 100 m grid for two cities (Los Angeles, CA, and New York, NY; see Figure 1). This figure illustrates month-to-month variability in the modeled spatial patterns.

Table 3 and Figure S3 show predicted versus observed monthly mean monitor values by distance to the nearest monitor. Model performance is generally better for locations within 10 km (mean spatial R^2 = 0.80, mean temporal R^2 = 0.72, spatiotemporal R^2 = 0.82, mean absolute error 2.6 ppb, mean bias 3%) and within 25 km (0.82, 0.73, 0.84, 2.5 ppb, 12%) than for locations between 25 and 50 km (0.70, 0.79, 0.77, 2.1 ppb, 23%) and further from (>50 km) nearby monitors (0.71, 0.69, 0.75, 2.2 ppb, 40%); differences among those four groups are greater for the mean bias than for the R^2 value. The mean absolute error is similar among the four location types. The spatial-only year 2006 LUR explains, on a population-weighted average, 71% (unweighted 72%) of the spatial variability in monthly mean concentrations for the years 2000−2010. Figure S4 exhibits spatial-only, temporal-only, and spatiotemporal R^2 values for each year using monthly mean concentrations for each monitor. Model performance is good, with R^2 values generally >0.70. The annual temporal R^2 shown is typically greater than for the long-term (2000−2011) temporal performance (R^2 = 0.72 [unweighted], 0.77 [weighted]). Monthly spatial-only correlation coefficients are given in Figures S5 and S8.

Figures S6 and S7 show plots analogous to those in Figures S3 and S4 but for urban, suburban, and rural monitor locations. Monitor performance is better for urban (mean spatial R^2 = 0.76, mean temporal R^2 = 0.76, mean spatiotemporal R^2 = 0.80, mean absolute error 2.6 ppb, mean bias 1%) and suburban (0.76, 0.76, 0.81, 2.4 ppb, 6%) locations than for rural locations (0.63, 0.63, 0.69, 2.0 ppb, 78%). As above, differences in the mean absolute error are small, and differences in the mean bias are larger. Urban and suburban locations typically outperform

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

Table 4. Summary of Within-City Monthly Mean LUR Estimates (2000−2010) for the 10 U.S. Cities with the Largest Number of NO_2 Monitors

city	mean (range) number of monitors	mean spatial R^2	mean temporal R^2	spatiotemporal R^2	mean error (ppb)	mean abs error (ppb)	mean bias (%)	mean abs bias (%)
Los Angeles	16 (11−18)	0.59	0.80	0.77	0.4	3.6	10	20
New York City	12 (9−15)	0.75	0.73	0.79	0.6	3.2	8	19
Houston	10 (6−11)	0.61	0.86	0.78	0.2	1.8	7	18
Washington, DC	9 (6−10)	0.61	0.81	0.77	−0.6	2.4	0.2	17
Chicago	7 (6−10)	0.57	0.52	0.62	−0.7	3.6	−2	17
Philadelphia	7 (4−8)	0.66	0.76	0.73	0.4	2.2	5	14
St. Louis	6 (1−9)	0.66	0.65	0.76	−0.1	1.9	3	15
San Diego	6 (5−8)	0.32	0.79	0.61	0.2	3.4	7	27
Pittsburgh	6 (4−6)	0.38	0.74	0.62	−0.7	2.5	−2	18
Phoenix	5 (1−7)	0.25	0.90	0.67	−0.01	3.4	4	16
median	7	0.60	0.78	0.75	0.1	2.9	5	17

rural locations for each year (Figure S7), except for the mean absolute error. Population-weighted performance metrics (see above and Table 1) are similar to urban- and suburban-only metrics, in part because a minority (∼19%) of the population lives in rural areas.[53] This fact emphasizes the potential utility of these models for estimating exposures. We found limited evidence of spatial autocorrelation in model residuals. For the months with significant values (102 of 132 months with $p < 0.05$), Moran's I coefficients were negative and small (values range from −0.05 to −0.22).

We assess the within-urban performance of our monthly NO_2 estimates by considering model performance separately in the 10 urban areas (UAs) with the most available monitors. Model performance varies among UAs (Table 4), but generally suggests reasonable within-urban predictive power (spatial R^2 = 0.25−0.75, temporal R^2 = 0.52−0.90, spatiotemporal R^2 = 0.61−0.79), even when considering only one UA at a time, as in Table 4. For the three cities with an average of at least 10 monitors (Los Angeles, New York, Houston), within-urban model performance is similar to the nationwide model performance. Scatterplots in Figure S9 of predicted vs observed concentrations for LA and NYC (see Figure 1) show strong model−measurement agreement.

As a sensitivity analysis, and to compare to the temporal scaling approach developed here, for year 2006, we developed 12 monthly LURs (i.e., using the same LUR model building procedure outlined above, but with monthly mean NO_2 monitor concentrations as the dependent variable). Model performance for the monthly LUR models (and the corresponding performance for spatial LUR + temporal scaling) are given in Table S9. The base-case spatial LUR + temporal scaling approach exhibits better model performance than the individual monthly LUR models (spatial R^2 = 0.77−0.84 [LUR + temporal scaling] vs 0.65−0.76 [monthly LUR], mean bias 14−26% vs 19−32%, mean absolute error 2.0−2.7 ppb vs 2.3−3.4 ppb). This finding highlights the potential difficulty in fitting monthly LUR models, and the utility of our temporal scaling approach.

We also tested the value of our approach by comparing our monthly estimates against (1) straightforward satellite-only estimates of surface concentrations (DOMINO) and (2) standard IDW interpolation of monitoring-station-only data. Summary statistics are in Tables S10 (DOMINO, 2005−2007) and S11 (IDW, 2000−2010). As expected, our core LUR + temporal scaling approach outperforms the two alternatives for predicting temporal variability (mean temporal R^2 = 0.73 [LUR + temporal scaling], 0.42, [DOMINO], 0.64 [IDW]), spatial variability (mean spatial R^2 = 0.81 vs 0.45 and 0.51), and spatiotemporal variability (R^2 = 0.84 vs 0.35 and 0.59). The mean bias is approximately 3 times lower for the LUR + temporal scaling than for IDW (21% vs 64%), and the mean absolute error is 1.7 times lower for the LUR + temporal scaling than for IDW (2.4 ppb vs 4.1 pbb). These findings illustrate the utility of the land-use information and LUR for predicting spatial variability, and the utility of the temporal scaling approach described here for predicting spatiotemporal variability.

4. DISCUSSION

We created national-scale monthly NO_2 surfaces for the contiguous United States using fixed-site regulatory monitors, satellite-derived NO_2 estimates, and GIS-derived land-use data. The resulting surfaces exhibit good spatial, temporal, and spatiotemporal predictive performance (overall R^2 = 0.84), with spatial resolution typical of urban LUR models (∼100 m). The large spatial and temporal coverage of our model is useful for national-scale longitudinal research on outdoor air pollution (e.g., population exposure assessment, epidemiology, environmental justice, surveillance, public policy) with excellent spatial resolution. For example, a cohort study of exposures during pregnancy could use NO_2 estimates provided here to explore impacts of exposure by pregnancy month or trimester. To facilitate future research, we estimated monthly NO_2 concentrations at the centroid locations for all ∼8 million Census blocks in the contiguous United States (blocks are the smallest area enumerated by the U.S. Census); this data set is publicly and freely available online [see the Supporting Information].

With relative measures (% rather than ppb), the model exhibits the greatest predictive power in urban and suburban locations. This finding supports the use of our LUR + temporal scaling model for population-based research. (Most [∼81%] of the U.S. population lives in urban or suburban areas.[53]) Rural monitors are, on average, further from nearby monitors (Table 3) and experience cleaner concentrations, and may suffer from greater relative error in estimating monthly scaling factors. In absolute terms (ppb rather than %), errors appear similar for urban, suburban, and rural areas. The correlation between NO_2 emissions (and concentrations) and GIS-derived land uses may be greater in urban areas than in suburban areas. Here, model

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

performance generally is similar or better for population-weighted metrics than unweighted metrics. The three methods of creating temporal scaling surfaces (kriging, IDW, NN) performed similarly well, suggesting that temporal scaling is relatively robust to the selection of interpolation technique. Moreover, this approach for temporal scaling (interpolating temporal trends from regulatory monitors) is straightforward and simple to implement. Alternative methods of temporal scaling (e.g., use of a CTM) may be more complex and require greater expert knowledge and computational resources to run. Satellite measurements offer another potential source of temporal scale factors; however, temporal scaling with that approach may be complicated by the need to combine data from several instruments with differing overpass times and retrieval algorithms.[52] Improved spatial and temporal resolution from future satellite measurements (e.g., Sentinel-5 Precursor, TEMPO) may improve the spatial performance of satellite-based LUR models and could provide an alternative temporal scaling approach.[54,55]

The model spatial-only performance (mean R^2 = 0.81) is consistent with that of other continental-scale NO_2 models (R^2 = 0.61−0.78; see Table S1)[13,15−17] and most urban- and national-scale LUR models for long-term average NO_2.[9] Spatiotemporal performance based on monthly mean estimates (R^2 = 0.84) is consistent with published monthly PM_{10} (R^2 = 0.76) and $PM_{2.5}$ (R^2 = 0.79−0.85) estimates in the United States.[14,34,35] Hart et al.[12] estimated annual average NO_2 concentrations for the years 1985−2000 in the United States with similar modeled spatiotemporal performance (R^2 = 0.88 vs 0.84 here, using annual mean concentrations). Kloog et al.[36,37] predicted daily mean $PM_{2.5}$ for the years 2000−2008 in New England and the Mid-Atlantic states with similar spatial R^2 (0.69−0.86) and spatiotemporal R^2 (0.73−0.90) and slightly better temporal R^2 (0.73−0.91). Lee et al.[39] predicted daily mean NO_2 for the years 2005−2010 in New England with similar spatiotemporal R^2 (0.79). These daily models, using a mixed effects modeling approach, offer excellent temporal coverage; to our knowledge, they have not been applied on a national scale.

Our results provide useful information for future continental LUR building. We found that model performance was relatively insensitive to the type of regional NO_2 covariate used (here, satellite-based ground level, satellite-based column total, DOMINO versus BEHR satellite data, 12 km CTM); the satellite column NO_2 slightly outperformed satellite-based surface estimates. These findings suggest that, for the purposes of LUR, it might be unnecessary to employ a CTM to convert satellite column measurements to surface estimates. Moreover, our findings indicate that the number of monitors needed to build a robust continental LUR (>150−200) is greater than that for a smaller scale LUR, and that large holdout subsets, systematic holdout, or Monte Carlo approaches may be needed to adequately evaluate the model performance.

Our model combines the spatial predictive power of LUR with the temporal coverage of the EPA monitoring network. The resulting model exhibits slightly better temporal predictive performance than IDW interpolation of monthly mean monitor values (on average, temporal-only R^2 = 0.73 [LUR + temporal scaling] vs 0.64 [IDW]), but with the greatly improved spatial performance (on average, spatial-only R^2 = 0.81 vs 0.51, spatiotemporal R^2 = 0.84 vs 0.59) typical of LUR.[7]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.5b02882.

Tables S1−S11, Figures S1−S9, eqs S1−4, derivation of interpolated spatiotemporal scaling surfaces, results of alternative spatial LUR models, and details on how to download monthly NO_2 concentration estimates for all U.S. Census block centroids (PDF)

Summary of existing large scale LUR models (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 612-625-2397; fax: 612-626-7750; e-mail: julian@umn.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant No. 1236800 and by the Minnesota Supercomputing Institute.

■ REFERENCES

(1) Burnett, R. T.; Stieb, D.; Brook, J. R.; Cakmak, S.; Dales, R.; Raizenne, M.; Vincent, R.; Dann, T. Associations between short-term changes in nitrogen dioxide and mortality in Canadian cities. *Arch. Environ. Health* **2004**, *59* (5), 228−236.

(2) Brook, J. R.; Burnett, R. T.; Dann, T. F.; Cakmak, S.; Goldberg, M. S.; Fan, X.; Wheeler, A. J. Further interpretation of the acute effect of nitrogen dioxide observed in Canadian time-series studies. *J. Exposure Sci. Environ. Epidemiol.* **2007**, *17* (Suppl. 2), S36−S44.

(3) Crouse, D. L.; Peters, P. A.; Villeneuve, P. J.; Proux, M.-O.; Shin, H. H.; Goldberg, M. S.; Johnson, M.; Wheeler, A. J.; Allen, R. W.; Atari, D. O.; et al. Within- and between-city contrasts in nitrogen dioxide and mortality in 10 Canadian cities; a subset of the Canadian Census Health and Environment Cohort (CanCHEC). *J. Exposure Sci. Environ. Epidemiol.* **2015**, *25* (5), 482−489.

(4) Filleul, L.; Rondeau, V.; Vandentorren, S.; Le Moual, N.; Cantagrel, A.; Annesi-Maesano, I.; Charpin, D.; Declercq, C.; Neukirch, F.; Paris, C.; et al. Twenty five year mortality and air pollution: results from the French PAARC survey. *Occup. Environ. Med.* **2005**, *62* (7), 453−460.

(5) Perez, L.; Lurmann, F.; Wilson, J.; Pastor, M.; Brandt, S. J.; Künzli, N.; McConnell, R. Near-roadway pollution and childhood asthma: implications for developing "win-win" compact urban development and clean vehicle strategies. *Environ. Health Perspect.* **2012**, *120* (11), 1619−1626.

(6) Gauderman, W. J.; Avol, E.; Lurmann, F.; Kuenzli, N.; Gilliland, F.; Peters, J.; McConnell, R. Childhood asthma and exposure to traffic and nitrogen dioxide. *Epidemiology* **2005**, *16* (6), 737−743.

(7) 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.

(8) Bechle, M. J.; Millet, D. B.; Marshall, J. D. Remote sensing of exposure to NO2: satellite versus ground-based measurement in a large urban area. *Atmos. Environ.* **2013**, *69* (2), 345−353.

(9) Hoek, G.; Beelen, R.; Dehoogh, 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.

(10) Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; Giovis, C. A review and evaluation of intraurban air pollution exposure models. *J. Exposure Anal. Environ. Epidemiol.* **2005**, *15* (2), 185−204.

(11) Ross, Z.; Jerrett, M.; Ito, K.; Tempalski, B.; Thurston, G. A land use regression for predicting fine particulate matter concentrations in the New York City region. *Atmos. Environ.* **2007**, *41* (11), 2255−2269.
(12) Hart, J. E.; Yanosky, J. D.; Puett, R. C.; Ryan, L.; Dockery, D. W.; Smith, T. J.; Garshick, E.; Laden, F. Spatial modeling of PM10 and NO2 in the continental United States, 1985−2000. *Environ. Health Perspect.* **2009**, *117* (11), 1690−1696.
(13) 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.
(14) Beckerman, B. S.; Jerrett, M.; Serre, M.; Martin, R. V.; Lee, S.-J.; van Donkelaar, A.; Ross, Z.; Su, J.; Burnett, R. T. A hybrid approach to estimating national scale spatiotemporal variability of PM2.5 in the contiguous United States. *Environ. Sci. Technol.* **2013**, *47* (13), 7233−7241.
(15) Hystad, P.; Setton, E.; Cervantes, 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.
(16) Beelen, R.; Hoek, G.; Pebesma, E.; Vienneau, D.; de Hoogh, K.; Briggs, D. J. Mapping of background air pollution at a fine spatial scale across the European Union. *Sci. Total Environ.* **2009**, *407* (6), 1852−1867.
(17) Vienneau, D.; de Hoogh, K.; Bechle, M. J.; Beelen, R.; van Donkelaar, A.; Martin, D. B.; Millet, D. B.; Hoek, G.; Marshall, J. D. Western European land use regression incorporating satellite- and ground-based measurements of NO_2 and PM_{10}. *Environ. Sci. Technol.* **2013**, *47* (23), 13555−13564.
(18) 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.
(19) Basagaña, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; Foraster, M.; de Nazelle, A.; Nieuwenhuijsen, M.; Vila, J.; et al. Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmos. Environ.* **2012**, *54*, 634−642.
(20) Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B. Systematic evaluation of land use regression models for NO_2. *Environ. Sci. Technol.* **2012**, *46* (8), 4481−4489.
(21) Johnson, M.; Isakov, V.; Touma, J. S.; Mukerjee, S.; Özkaynak, H. Evaluation of land-use regression models used to predict air quality concentrations in an urban area. *Atmos. Environ.* **2010**, *44* (30), 3660−3668.
(22) Chen, H.; Goldberg, M. S.; Crouse, D. L.; Burnett, R. T.; Jerrett, M.; Villeneuve, P. J.; Wheeler, A. J.; Labrèche, F.; Ross, N. A. Back-extrapolation of estimates of exposure from current land-use regression models. *Atmos. Environ.* **2010**, *44* (35), 4346−4354.
(23) Eeftens, M.; Beelen, R.; Fischer, P.; Brunekreef, B.; Meliefste, K.; Hoek, G. Stability of measured and modelled spatial contrasts in NO_2 over time. *Occup. Environ. Med.* **2011**, *68* (10), 765−770.
(24) Fang, T. B.; Lu, Y. Constructing a near real-time space-time cube to depict urban ambient air pollution scenario. *Trans. GIS* **2011**, *15* (5), 635−649.
(25) Mölter, A.; Lindley, S.; de Vocht, F.; Simpson, A.; Agius, R. Modelling air pollution for epidemiologic research–part II: predicting temporal variation through land use regression. *Sci. Total Environ.* **2010**, *409* (1), 211−217.
(26) 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.
(27) Sampson, P. D.; Szpiro, A. A.; Sheppard, L.; Lindström, J.; Kaufman, J. D. Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmos. Environ.* **2011**, *45* (36), 6593−6606.
(28) Su, J. G.; Brauer, M.; Ainslie, B.; Steyn, D.; Larson, T.; Buzzelli, M. An innovative land use regression model incorporating meteorology for exposure analysis. *Sci. Total Environ.* **2008**, *390* (2−3), 520−529.
(29) Maynard, D.; Coull, B. A.; Gryparis, A.; Schwartz, J. Mortality risk associated with short-term exposure to traffic particles and sulfates. *Environ. Health Perspect.* **2007**, *115* (5), 751−755.
(30) Dons, E.; Van Poppel, M.; Kochan, B.; Wets, G.; Int Panis, L. Modeling temporal and spatial variability of traffic-related air pollution: Hourly land use regression models for black carbon. *Atmos. Environ.* **2013**, *74*, 237−246.
(31) Gulliver, J.; Morris, C.; Lee, K.; Vienneau, D.; Briggs, D.; Hansell, A. Land use regression modeling to estimate historic (1962−1991) concentrations of black smoke and sulfur dioxide for Great Britain. *Environ. Sci. Technol.* **2011**, *45* (8), 3526−3532.
(32) 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.
(33) Mao, L.; Qiu, Y.; Kusano, C.; Xu, X. Predicting regional space-time variation of PM2.5 with land-use regression model and MODIS data. *Environ. Sci. Pollut. Res.* **2012**, *19* (1), 128−138.
(34) Yanosky, J. D.; Paciorek, C. J.; Schwartz, J.; Laden, F.; Puett, R.; Suh, H. H. Spatio-temporal modeling of chronic PM10 exposure for the Nurses' Health Study. *Atmos. Environ.* **2008**, *42* (18), 4047−4062.
(35) Yanosky, J. D.; Paciorek, C. J.; Suh, H. H. Predicting chronic fine and coarse particulate exposures using spatiotemporal models for the Northeastern and Midwestern United States. *Environ. Health Perspect.* **2009**, *117* (4), 522−529.
(36) Kloog, I.; Nordio, F.; Coull, B. A.; Schwartz, J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{2.5}$ exposures in the Mid-Atlantic states. *Environ. Sci. Technol.* **2012**, *46* (21), 11913−11921.
(37) Kloog, I.; Koutrakis, P.; Coull, B. A.; Lee, H. J.; Schwartz, J. Assessing temporally and spatially resolved PM2.5 exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos. Environ.* **2011**, *45* (35), 6267−6275.
(38) Liu, Y.; Paciorek, C. J.; Koutrakis, P. Estimating Regional Spatial and Temporal Variability of PM2.5 Concentrations Using Satellite Data, Meteorology, and Land Use Information. *Environ. Health Perspect.* **2009**, *117* (6), 886−892.
(39) Lee, H. J.; Koutrakis, P. Daily Ambient NO2 Concentration Predictions Using Satellite Ozone Monitoring Instrument NO2 Data and Land Use Regression. *Environ. Sci. Technol.* **2014**, *48* (4), 2305−2311.
(40) Gan, W. Q.; Koehoorn, M.; Davies, H. W.; Demers, P. A.; Tamburic, L.; Brauer, M. Long-term exposure to traffic-related air pollution and the risk of coronary heart disease hospitalization and mortality. *Environ. Health Perspect.* **2011**, *119* (4), 501−507.
(41) Slama, R.; Morgenstern, V.; Cyrys, J.; Zutavern, A.; Herbarth, O.; Wichmann, H.-E.; Heinrich, J. Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model. *Environ. Health Perspect.* **2007**, *115* (9), 1283−1292.
(42) Nethery, E.; Teschke, K.; Brauer, M. Predicting personal exposure of pregnant women to traffic-related air pollutants. *Sci. Total Environ.* **2008**, *395* (1), 11−22.
(43) Hystad, P.; Demers, P. A.; Johnson, K. C.; Brook, J.; van Donkelaar, A.; Lamsal, L.; Martin, R.; Brauer, M. Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study. *Environ. Health* **2012**, *11* (1), 22.
(44) Su, J. G.; Jerrett, M.; Beckerman, B. A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. *Sci. Total Environ.* **2009**, *407* (12), 3890−3898.
(45) Russell, A. R.; Perring, A. E.; Valin, L. C.; Bucsela, E. J.; Browne, E. C.; Wooldridge, P. J.; Cohen, R. C. A high spatial resolution retrieval of NO_2 column densities from OMI: method and evaluation. *Atmos. Chem. Phys.* **2011**, *11* (16), 8543−8554.
(46) Tessum, C. W.; Hill, J. D.; Marshall, J. D. Twelve-month, 12 km resolution North American WRF-Chem v3.4 air quality simulation: performance evaluation. *Geosci. Model Dev.* **2015**, *8* (4), 957−973.

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305

(47) Grundström, M.; Pleijel, H. Limited effect of urban tree vegetation on NO2 and O3 concentrations near a traffic route. *Environ. Pollut.* **2014**, *189*, 73−76.
(48) Pataki, D. E.; Carreiro, M. M.; Cherrier, J.; Grulke, N. E.; Jennings, V.; Pincetl, S.; Pouyat, R. V.; Whitlow, T. H.; Zipperer, W. C. Coupling biogeochemical cycles in urban environments: ecosystem services, green solutions, and misconceptions. *Front. Ecol. Environ.* **2011**, *9* (1), 27−36.
(49) U.S. Enivironmental Protection Agency. Air Quality System Data Mart. http://www.epa.gov/ttn/airs/aqsdatamart/. (Accessed 03/21/2012).
(50) U.S. Enivironmental Protection Agency. *Our Nation's Air: Status and Trends Through 2010*; Research Triangle Park, NC, 2012.
(51) Russell, A. R.; Valin, L. C.; Cohen, R. C. Trends in OMI NO2 observations over the United States: effects of emission control technology and the economic recession. *Atmos. Chem. Phys.* **2012**, *12* (24), 12197−12209.
(52) Kharol, S. K.; Martin, R. V.; Philip, S.; Boys, B.; Lamsal, L. N.; Jerrett, M.; Brauer, M.; Crouse, D. L.; McLinden, C.; Burnett, R. T. Assessment of the Magnitude and Recent Trends in Satellite-Derived Ground-Level Nitrogen Dioxide over North America. *Atmos. Environ.* **2015**, *118*, 236−245.
(53) U.S. Census Bureau. 2010 Urban Area Facts. http://www.census.gov/geo/www/ua/uafacts.html (Accessed 09/29/2015).
(54) NASA. Missions: TEMPO. http://science.nasa.gov/missions/tempo/ (Accessed 11/12/2013).
(55) ESA. sentinel-5 precursor. http://esamultimedia.esa.int/docs/S5-prec_Data_Sheet.pdf (Accessed 11/12/2013).

DOI: 10.1021/acs.est.5b02882
Environ. Sci. Technol. 2015, 49, 12297−12305