Environmental Science & Technology

pubs.acs.org/est

Article

National Land Use Regression Model for NO_2 Using Street View Imagery and Satellite Observations

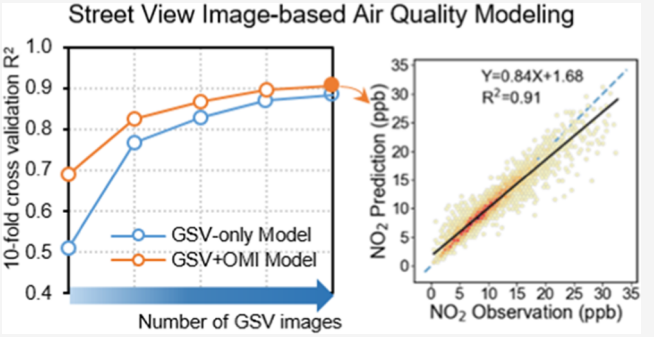
Meng Qi, Kuldeep Dixit, Julian D. Marshall, Wenwen Zhang, and Steve Hankey*

Cite This: *Environ. Sci. Technol.* 2022, 56, 13499−13509

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Land use regression (LUR) models are widely applied to estimate intra-urban air pollution concentrations. National-scale LURs typically employ predictors from multiple curated geodatabases at neighborhood scales. In this study, we instead developed national NO_2 models relying on innovative street-level predictors extracted from Google Street View [GSV] imagery. Using machine learning (random forest), we developed two types of models: (1) GSV-only models, which use only GSV features, and (2) GSV + OMI models, which also include satellite observations of NO_2. Our results suggest that street view imagery alone may provide sufficient information to explain NO_2 variation. Satellite observations can improve model performance, but the contribution decreases as more images are available. Random 10-fold cross-validation R^2 of our best models were 0.88 (GSV-only) and 0.91 (GSV + OMI)—a performance that is comparable to traditional LUR approaches. Importantly, our models show that street-level features might have the potential to better capture intra-urban variation of NO_2 pollution than traditional LUR. Collectively, our findings indicate that street view image-based modeling has great potential for building large-scale air quality models under a unified framework. Toward that goal, we describe a cost-effective image sampling strategy for future studies based on a systematic evaluation of image availability and model performance.



KEYWORDS: *empirical models, exposure assessment, air quality, image sampling and processing, computer vision, machine learning*

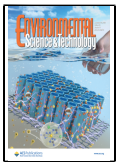
INTRODUCTION

Nitrogen dioxide (NO_2) is an important traffic-related air pollutant (TRAP) and a criteria pollutant. NO_2 has been associated with a wide range of adverse health effects, including premature mortality, asthma, lung cancer, and cardiovascular disease.[1−6] As a short-lived pollutant, NO_2 concentrations decay rapidly near emission sources, with published estimates of typical distance-decay gradients ranging from <500 to <2000 m, depending on the atmospheric dispersion conditions.[7−10] Capturing intra-urban variations of NO_2 concentrations is essential for accurately assessing human exposure to NO_2 and the corresponding health outcomes.[11,12] Among the many efforts to improve the spatial resolution of air quality models, land use regression (LUR) has been frequently applied as a cost-effective way to estimate fine-scale ambient air pollutant concentrations.[13−16] According to a review, LUR was the most popular method for NO_x studies in the last decade.[17]

National LUR models have been developed for, e.g., the United States,[18−21] Australia,[22,23] Europe,[24,25] and China.[26,27] Traditional LUR models rely on predictors collected from curated GIS-derived databases.[13,28] For example, Kim et al. collected geographic variables from 11 categories to build national air quality models, including traffic, population, land use, elevation, emissions, and satellite estimates.[18] Larkin et al. developed a global NO_2 LUR model using satellite estimates and land use variables, including normalized difference vegetation index, impervious surface, road length, and emission sources.[29] However, the availability of curated GIS-derived databases may vary across jurisdictions,[30] and preprocessing geographic predictors originating from various data sources can be a labor-intensive process. Additionally, there is evidence that traditional GIS variables may not capture street-level features that may be important determinants of air quality.[11,31]

In recent years, innovative image data sources (e.g., street view imagery, high-resolution satellite imagery) have emerged as possible tools to capture hyperlocal characteristics of the natural and built environment. Image-based data may be promising for replacing or augmenting traditional LUR predictors[32,33] when enabled by imagery processing techniques (i.e., computer vision) and advanced modeling (e.g., machine learning).[34] For example, information (e.g., traffic, land use,

Received: May 18, 2022
Revised: August 27, 2022
Accepted: August 29, 2022
Published: September 9, 2022



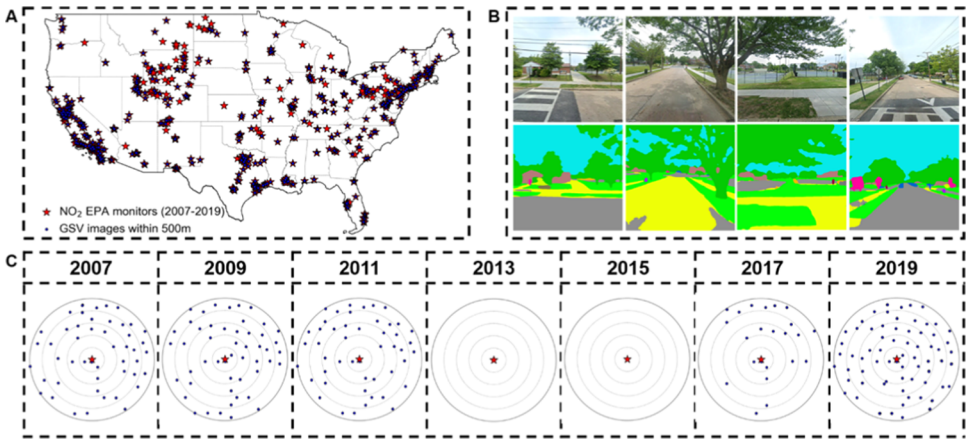


Figure 1. NO_2 measurement and GSV locations for the national NO_2 models. (A) Spatial distribution of NO_2 EPA monitors and GSV images within 500 m of monitors. (B) Example of GSV features extracted via image semantic segmentation. (C) Illustration of spatiotemporal variation in GSV image availability for one EPA monitor.

built environment features) provided by traditional LUR predictors are also encoded in high-resolution digital images[35] and can be extracted and quantified via advanced computer vision techniques. Compared to traditional curated geodatabases, image data sources allow consistent and uniform data collection and processing across jurisdictions.

As an emerging research topic, a small number of studies have explored using these new data sources to build empirical air quality models.[32,33,36−38] For example, Lu et al. compared national empirical models of six criteria pollutants using traditional LUR variables versus microscale variables (e.g., street view imagery, point of interest, local climate zones).[32] Similar model performance was achieved using the various combinations of variables, suggesting that microscale variables could be a suitable substitute for traditional predictors. Another study trained convolutional neural networks (CNN) on Google Maps satellite and street-level images, together with traditional LUR variables, to predict ultrafine particle (UFP) and black carbon (BC) concentrations.[33] Those results suggest that images may capture similar features as traditional GIS predictors while also allowing for the inclusion of higher-resolution (i.e., street-level) features that traditional GIS variables lack.[37] Another notable trend in recent air quality modeling studies is the adoption of advanced machine learning approaches to improve model performance, including random forest,[39−41] gradient boosting,[42,43] artificial neural network,[44−46] and hybrid algorithms.[47,48]

Our previous study successfully developed single-city LUR models solely using street view images.[36] In this study, we scale this approach to develop national NO_2 models for 2007−2019 using street view images and machine learning algorithms. Using the Environmental Protection Agency (EPA) fixed-site network as a basis (dependent variable), we employed the same image processing protocol developed in our previous study[36] to extract street-level features for modeling. We developed two types of random forest models: (1) using Google Street View (GSV) imagery, and (2) using GSV features plus satellite observations of NO_2 tropospheric vertical columns. To our knowledge, this is the first national NO_2 model which is based on street view imagery. We systematically assess the impact of GSV image availability on model performance, which may provide important evidence for optimal image sampling strategies for future image-based air quality modeling studies.

MATERIALS AND METHODS

Modeling Variables. Figure 1 shows the sampling and processing procedure for preparing variables for modeling. We collected annual average NO_2 concentrations at EPA monitors across the contiguous U.S. during 2007−2019 from the EPA AQS data set. For predictor variables, we used street-level features extracted from GSV images and satellite observations of NO_2 tropospheric column density. In this paper, we use the term "GSV image" to refer to GSV panoramic images with unique image IDs. Each unique "GSV image" has four flat images corresponding to the four cardinal directions.

GSV Image-Derived Variables. We downloaded 336,544 GSV flat images within 500 m of EPA NO_2 monitors. For image collection and processing, we followed a similar procedure as detailed in our previous study.[36] Specifically, we created 100 m × 100 m grids within 500 m of each EPA monitor and retrieved the metadata for all GSV images within the grid cells. To improve temporal consistency between the NO_2 observations and GSV features, we constrained the selection of GSV images such that only images taken during the same year or within the previous 4 years of the NO_2 observations were downloaded and processed. Further details about the choice of image inclusion by year are explained in the modeling approach section. After image retrieval, we used a deep learning model called Pyramid Scene Parsing Network (PSPNet),[49] to extract GSV features through image semantic segmentation (i.e., each pixel of the image is assigned to a feature category). A total of 150 GSV features were extracted through PSPNet processing. To improve the model interpretation, we removed features that are unrelated to air pollution (47 variables (Table S1) remained for model development). We categorized GSV features into seven subgroups: built environment (n = 8), transport network (n = 8), transport vehicles (n = 9), natural (n = 8), vegetation (n = 6), water (n = 6), and human (n = 2). To tabulate GSV features for modeling, we stratified the GSV images into different buffer radii. Within each buffer radius, we calculated the mean percentage of GSV features to represent the surrounding natural and built environment characteristics.

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

Unless specified otherwise, we used 250 and 500 m buffer radii as the default for modeling.

Satellite Variable. We also retrieved satellite observations of NO_2 from the Ozone Monitoring Instrument (OMI). OMI was launched in 2004 on the Aura platform and had a nadir footprint of approximately 13 km × 24 km.[50] Many studies have demonstrated the predictive power of satellite-based variables in NO_2 models.[19,21,43,51] Some studies found that directly adopting the satellite column density into the LUR model is sufficient for modeling, suggesting that the extra step for estimating the column-to-surface ratio may be unnecessary.[20,22] In our study, we directly used the satellite observations (column density) at the EPA sites as an additional model predictor. We obtained OMI NO_2 tropospheric column density from the OMI L3 product (OMI/Aura NO_2 Tropospheric, Stratospheric & Total Columns MINDS Daily L3 Global Gridded 0.25° × 0.25°)[52] provided by the NASA Goddard Earth Sciences Data and Information Services Center (GESDISC). For data quality control, we used the cloud-screened tropospheric vertical column density with effective cloud fraction (ECF) <0.3, solar zenith angle (SZA) <85°, and the primary summary quality flag indicating good data.

Modeling Approach. Since GSV images are not updated at a uniform rate across regions, the density of available images in the vicinity of NO_2 monitors varies depending on the site location and year. Figure 1C demonstrates the spatiotemporal variation of image availability using one EPA site for illustration; here (as with many locations), image availability varies by year, and in extreme cases, there may be no images for certain years. To systematically analyze the impact of GSV image availability on model performance, we designed five scenarios for temporal matching of images as model inputs (see Table S2). The strictest scenario (i.e., scenario 1) uses only GSV images from the same year of the NO_2 observations. Since the built environment changes slowly in developed countries such as the U.S., we relaxed the temporal match criteria in other scenarios. The relaxed scenarios (i.e., scenarios 2−5) further include GSV images from the previous 1−4 years before the NO_2 observations. As a result, the data set for a stricter scenario (i.e., fewer years) is always a subset of the more relaxed scenario. Data summaries for each scenario are shown in Table S2. Once all model variables were preprocessed, we developed two types of random forest models: (1) GSV-only models, which solely relied on GSV features, and (2) GSV + OMI models, which used both GSV features and the OMI variable. Both models included the year of NO_2 observation as a dummy variable in the model.

Model Evaluation. To compare model performance among different scenarios, we report cross-validation (CV) R^2, mean absolute error (MAE), and root-mean-squared error (RMSE). Our default cross-validation approach was a random 10-fold cross-validation for which all NO_2 records were randomly divided into 10 subsets. Each subset was sequentially held out as the test set, a random forest model was trained with the remaining nine subsets, and model performance was evaluated on the unseen test set. We iterated this process for each subset and calculated the overall R^2 after all subsets were predicted. In addition to the conventional random 10-fold cross-validation, we also report two additional cross-validation strategies: spatial hold out and temporal hold out. More details are in the Supporting Information (SI).

Sensitivity Analysis for Image Sampling Strategies. In addition to exploring the influence of temporal matching criteria for image−monitor pairs, we also examined the impact of image availability on model performance using different image sampling strategies. Our base models downloaded all available images; in this sensitivity analysis, we limited the number of GSV images retrieved for each 100 m × 100 m grid within the 500 m buffer of NO_2 monitors. We chose thresholds of 1, 2, or 3 GSV panoramic images (i.e., 4, 8, or 12 GSV flat images) for each grid cell. In this paper, we refer to these models as image-limited models. We refer to models with no image inclusion threshold, i.e., including all GSV images available within the grid cell for modeling, as image-unlimited models. For image-limited models, when multiple images from different years were available, we chose images with the best NO_2-GSV temporal match. We also examined image sampling strategies among different sampling grid sizes. We compared 150 m × 150 m, 200 m × 200 m, and 250 m × 250 m grids. Direct comparison between image-unlimited and image-limited models was feasible within the same temporal matching scenario because the model training size remained the same among different sampling strategies.

RESULTS AND DISCUSSION

Summary of Model Inputs. We collected 4867 unique NO_2 annual average records for the contiguous U.S. during the 13-year study period. NO_2 concentrations declined over time (Figure S1); the number of NO_2 EPA monitors slightly increased from 356 stations in 2007 to 407 in 2019 (average 374 stations/year during the study period).

GSV image availability varies across regions and years (see Figure S2). Hence, not every NO_2 site in every year has sufficient GSV images for model training. By including images from neighboring years, we were able to include a larger training data set while striving to capture similar built environment characteristics to the target year. In scenario 1, only 42% of NO_2 annual records had one or more images for the target year within 250 m of the monitor. Google started collecting street view images in 2007; image availability at EPA monitors varied by year. We found that 51% of NO_2 monitors had GSV images within 250 m in 2007; however, very few EPA monitors had images in 2008−2010 (e.g., only five locations had images in 2010). Relatively more images were collected in 2011, but the number of NO_2-GSV pairs dropped again from 2012 to 2014. In more recent years, the number of available GSV images generally increased. By relaxing the requirement for an exact temporal match (e.g., in scenario 5), far more images are available for modeling, leading to a more balanced training set across years. Overall, 72% of NO_2 annual records were successfully paired with GSV features in scenario 5. Among these GSV-NO_2 pairs, the median (P10/P25/P75/P90) of GSV images within 500 m of one EPA monitor was 53 (12/26/104/164).

Figure S3 shows the spatial distribution of NO_2-GSV pairs among years for scenario 5. GSV image availability near NO_2 monitors increased from early years to the present. Variation in GSV image availability among monitor sites also increased, i.e., for a given year, some monitors include hundreds of images while others only have a few images. We also examined the spatial distribution of NO_2 monitors, which were excluded from modeling due to insufficient images, and found that monitors located around the Rocky Mountains have the fewest GSV images. In general, the EPA locates NO_2 monitors for a variety of purposes (e.g., near-road monitoring, area-wide monitoring, or to protect susceptible and vulnerable

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

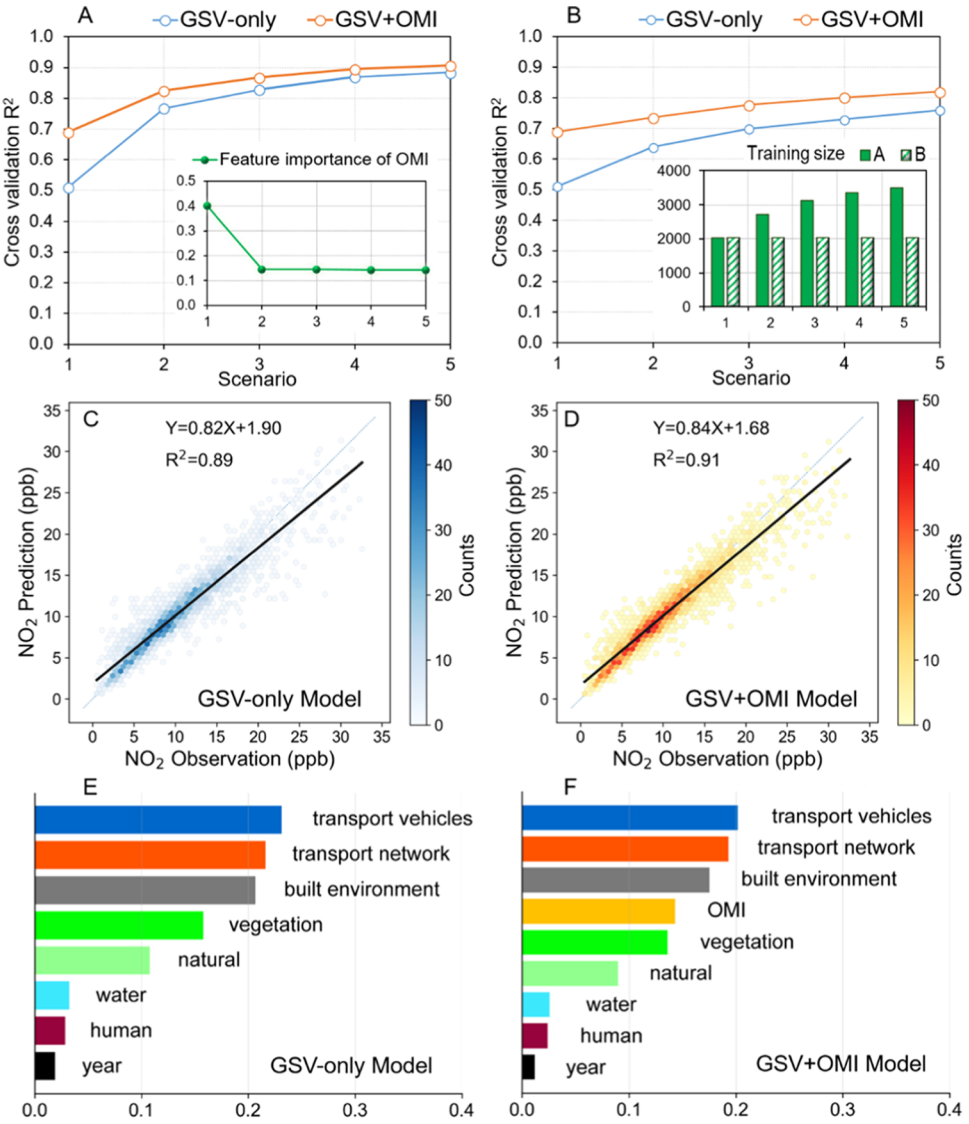


Figure 2. Comparison of random 10-fold cross-validation R^2 between GSV-only and GSV + OMI models. Panel (A) shows the results of image-unlimited models, while panel (B) shows the results of image-unlimited models with constrained training sizes. The training sizes of models in panels (A) and (B) are shown in the bar chart within panel (B). Panel (A) also shows the feature importance score of the OMI variable in the GSV + OMI models among the five scenarios. Both panels (A) and (B) show results for five scenarios, while subplots in panels (C) to (F) only show results by scenario 5 from models in panel (A). Panels (C) and (D) show hexagonal binned plots of model predictions versus observations. The color reflects the number of data points within each hexagon. Panels (E) and (F) show the rank of feature importance in the random forest models. Results of other scenarios are shown in Figures S5 and S7.

populations). Table S3 shows the distribution of NO_2 monitors by urbanicity (i.e., urban/suburban/rural) using 2019 for illustration: of the 407 monitors, 45% were located in urban areas versus 28% (suburban) and 27% (rural). For context, the U.S. population is 32% urban, 47% suburban, and 22% rural. Because of how GSV images are collected, more images were available in urban than in other areas: the proportion of monitors having GSV images within 500 m in scenario 5 was 96% for urban areas, 84% for suburban areas, and 38% for rural areas. Detailed descriptive statistics of GSV features for NO_2 monitors by urbanicity are listed in Table S4.

Model Results of Image-Unlimited Models. Using the image-unlimited models (i.e., including all images), we compared model performance among the different NO_2-GSV temporal match scenarios and investigated the ranking of feature importance among models.

Model Performance. Figure 2A shows the results of the random 10-fold cross-validation for both GSV-only and GSV + OMI models. Figure S4 shows the corresponding results for MAE and RMSE. Full model performance is shown in Table 1. Both GSV-only and GSV + OMI models showed increasing performance from scenario 1 to scenario 5. With the addition of satellite observations, the GSV + OMI models outperformed GSV-only models in all scenarios. In the strictest scenario, the GSV-only model showed moderate performance (random 10-fold CV R^2: 0.50; MAE: 2.95 ppb; RMSE: 3.90 ppb), while the corresponding GSV + OMI model showed improved results (random 10-fold CV R^2: 0.69; MAE: 2.26 ppb; RMSE: 3.05 ppb). However, the advantage of adding the satellite-based data gradually narrowed when more images were included in the model. From scenario 1 to scenario 2, the GSV-only model showed a dramatic jump in model fit, with CV R^2 increasing by 0.26. As more images were included in the model, the model

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

Table 1. Model Performance of Image-Unlimited Models

model type	scenario	random 10-fold CV R^2	mean absolute error (MAE)	root-mean-squared error (RMSE)
GSV-only	1	0.51	2.89	3.83
	2	0.77	1.86	2.64
	3	0.83	1.55	2.27
	4	0.87	1.33	1.99
	5	0.88	1.25	1.88
GSV + OMI	1	0.69	2.26	3.05
	2	0.83	1.61	2.29
	3	0.87	1.37	2.00
	4	0.90	1.19	1.78
	5	0.91	1.13	1.69

performance increased but plateaued by scenario 5. Correspondingly, the difference in the random 10-fold CV R^2 between GSV + OMI models and GSV-only models decreased from 0.18 in scenario 1 to 0.02 in scenario 5. A similar decreasing trend was observed for the feature importance of the OMI variable in the GSV + OMI models. Both the GSV-only model (10-fold CV R^2: 0.88; MAE: 1.25 ppb; RMSE: 1.88 ppb) and GSV + OMI model (10-fold CV R^2: 0.91; MAE: 1.13 ppb; RMSE: 1.69 ppb) achieved the best performance in scenario 5, where the advantage of the GSV + OMI model over the GSV-only model became marginal. Compared to other national NO_2 modeling studies (R^2: 0.72∼0.91),[18,21,32,53] the performance of our best models was consistent with or better than previous efforts. Our results indicate that with sufficient images, GSV imagery alone may explain the variation of long-term national NO_2 concentrations.

Relaxing the temporal matching criteria for GSV-NO_2 pairs increased both the number of GSV images included in the model and the available GSV-NO_2 pairs for model training. To separate the effects of the two factors, we further constrained the training size of each scenario by choosing the same NO_2 records for model development. Figure 2B illustrates that when limited to the same training data (i.e., the same NO_2 observations), adding more GSV images to the model still increases model performance significantly. For example, 10-fold CV R^2 for the GSV-only model in scenario 5 increased by 0.25 compared to scenario 1. Comparatively, the increase for models without constraints in training size was 0.37. As expected, the model performance was worse when we limited the training data size, and a larger gap between GSV-only and GSV + OMI models can be observed. The best model fits decreased to 0.76 for the GSV-only model and 0.82 for the GSV + OMI model. Full results are shown in Table S5.

We examined the model performance by comparing model predictions versus observations. Figure 2C,D illustrates the results for scenario 5, while the full results are shown in Figure S5. Overall, the trendlines for model predictions versus observations are closer to the 1:1 line when including more images (i.e., moving from scenario 1 to scenario 5). In scenario 5, both GSV-only and GSV + OMI models show good agreement between the predicted and observed NO_2 concentrations. No distinctive spatial pattern was observed in terms of monitors with the lowest or highest errors (Figure S6).

Rank of Feature Importance. Feature importance scores (Figure 2E,F) suggest that traffic-related features were the largest contributors to our NO_2 predictions. Generally, transport vehicle variables contributed the most, followed by transport network and built environment variables. This result is important as traditional GIS-derived variables are not able to resolve some of these differences (e.g., transport vehicles vs transport network) that imagery was able to identify. This may be important as some parts of the urban environment are fixed (e.g., buildings, roads), while others are transient (e.g., vehicles); policy solutions may differ depending on the nature of these features.

The granularity of GSV features allowed further comparison among different components of the urban environment. Figure S8 shows disaggregated model predictors using the top 10 features for illustration. The OMI variable was always the strongest single predictor in our GSV + OMI models. In scenario 1, the feature importance score of the OMI variable was 0.40, while the second highest score was only 0.06 for car. However, when we added more images into the model, the contribution of satellite data decreased dramatically, and GSV features contributed more to the models overall. This finding may be partially explained by the different spatial resolutions of GSV features and the OMI variable. The satellite observations have a coarser resolution (relative to GSV images) and thus contribute information on regional trends. Our preliminary analysis of an OMI-only model (i.e., using the OMI variable as the only predictor) showed that OMI alone could explain around half of the NO_2 variation (CV R^2: 0.49; MAE: 3.09 ppb; RMSE: 4.08 ppb). To further capture spatial trends of NO_2, predictors with much higher spatial resolution are needed.

Among GSV features, the most important features in both GSV-only and GSV + OMI models included transport vehicles (car, truck), transport network (sidewalk), built environment (wall, building), and vegetation (grass). These features have plausible explanations for inclusion in our models, e.g., traffic-related features reflect emission sources, built environment features may be associated with activity centers and pollution dispersion, and green spaces are typically low-emission areas. We examined how GSV features influence our models through accumulated local effects plots using several important features and the GSV-only model for illustration. As shown in Figure S9, car, truck, building, wall, and sidewalk positively contribute to NO_2 pollution, while grass has a negative contribution.

Comparison of Different Cross-Validation Strategies. Besides the conventional random 10-fold cross-validation, we used two additional cross-validation strategies, i.e., splitting the training and testing set spatially or temporally (full details are described in the SI). Briefly, we found that temporal cross-validation achieved a similar performance as the random cross-validation approach (Figure S10). However, spatial cross-validation showed degraded performance, especially for the GSV-only models (CV R^2: 0.46; MAE: 3.05 ppb; RMSE: 4.04 ppb). The GSV + OMI model (CV R^2: 0.59; MAE: 2.62 ppb; RMSE: 3.52 ppb) showed slightly better performance, indicating improved model generalizability with the inclusion of regional information (i.e., OMI variable) in the model. One explanation may be that we were only able to use a small search area (i.e., 500 m) to retrieve street view images near NO_2 monitors. As a result, only hyperlocal information was included in the national model—a potential issue for model extrapolation in regions where there are few EPA monitors. We speculate that if a denser air pollution monitoring network were used (e.g., mobile monitoring), this issue could be mitigated.

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

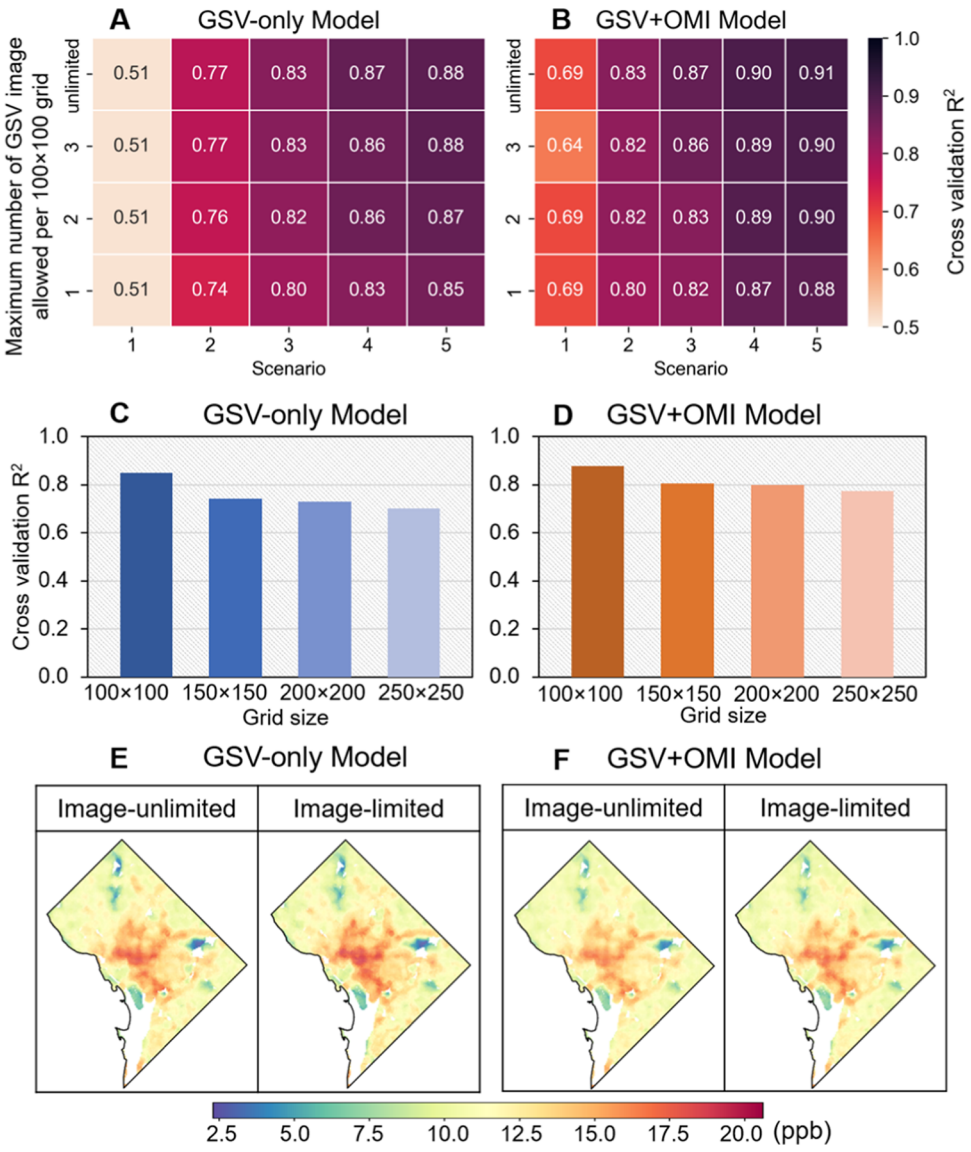


Figure 3. Results of sensitivity analyses among image sampling strategies. Panels (A) and (B) show model results among all scenarios, with varying maximum number of GSV images per 100 m × 100 m sampling grid. Panels (C) and (D) show model results in scenario 5 with varying grid sizes. Each grid cell is allowed one GSV image for sampling. Panels (E) and (F) show model predictions for Washington DC in 2019 using image-unlimited models and image-limited models, with maximum one GSV image per 100 m × 100 m grid.

Sensitivity Analysis Using Image-Limited Models. We systematically controlled the image sampling strategy by varying the number of images sampled for each grid cell (as well as the sampling grid size) to investigate the impact on model performance. Full model results are in Tables S6 and S7. As shown in Figure 3A,B, image-limited models generally had decreased performance compared to image-unlimited models. However, the reduction in model fit was relatively small. For example, in scenario 5, when collecting one GSV image per 100 m × 100 m grid, the random 10-fold CV R^2 was 0.85 (0.88) for the GSV-only (GSV + OMI) model; comparatively, it was 0.88 (0.91) for the GSV-only (GSV + OMI) for the image-unlimited models. This result was due to the fact that only a small portion of grid cells had repetitive GSV images, resulting in only 6.2% of GSV flat images being removed from the analysis.

Although only a small number of images were removed during model development, this finding is most important for the model application stage when the number of GSV images above a given threshold (e.g., 1, 2, or 3 per grid cell) could be huge. We used Washington DC as an example city to illustrate model prediction in 2019 at 100 m spatial resolution among different image sampling strategies (Figure 3E,F). Interestingly, image-unlimited and image-limited models generated similar prediction surfaces for NO_2, while the former models used twice the number of images (158,276 vs 79,612 GSV flat images). This finding is important when making predictions at national or global scales. Our results suggest that a parsimonious image sampling strategy may be sufficient and most cost-effective for model development and application. A rationale for this approach is that images distributed geographically close are more likely to provide redundant information, resulting in little influence on the tabulated model predictors, i.e., the mean of GSV features within a certain buffer area. This issue could be further explored in future research, especially for differences in transient (e.g.,

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

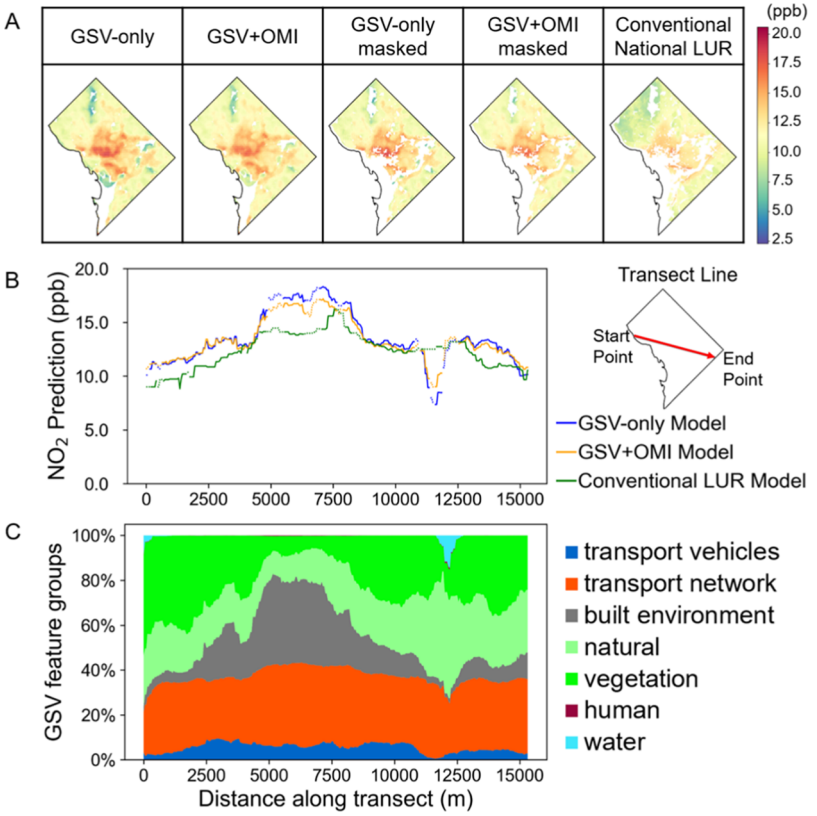


Figure 4. Comparison between GSV-based models and conventional national LUR models using prediction maps for Washington DC in 2015. (A) Model predictions for the GSV-only model, GSV + OMI model, and conventional national LUR model. For comparison, we masked GSV-based prediction maps based on where the conventional LUR predictions were available. (B) Predicted NO_2 concentrations among different models along the transect line. The solid lines represent actual model predictions; gaps in the conventional national LUR model (masked GSV-based models) were filled by interpolation (original GSV-based models) and represented by dotted lines. (C) Values of GSV feature groups along the transect line.

vehicles, people) vs fixed (e.g., roads, buildings) features of the urban environment.

We also explored the influence of various sampling grid sizes on model performance when sampling one GSV image per grid cell. Results for scenario 5 are in Figure 3C,D. Increasing the size of grid cells led to a significant deterioration in the model performance. When using a 150 m × 150 m (250 m × 250 m) grid and a maximum of one GSV image per grid cell, the random 10-fold CV R^2 of the GSV-only model dropped to 0.74 (0.70). This result reflects the importance of spatial resolution of GSV images used in the models. For example, the 100 m × 100 m grid generated ∼81 grid cells within a 500 m buffer, while the 150 m × 150 m (250 m × 250 m) grid generated ∼29 (∼7) grid cells. As expected, the GSV + OMI models were more robust to differences in grid size relative to the GSV-only models. Overall, our findings suggest that retrieving GSV images at 100 m resolution may be the most cost-effective way to establish national NO_2 models with high performance. Sparse resolution of image sampling may miss important hyperlocal features and lead to underfitting models. Overly dense sampling within a single grid cell may lead to an unnecessarily high workload—a cost-prohibitive (e.g., image downloading and processing) and time-consuming task for limited improvement in model fit.

Model Prediction and Comparison. Following the same image collection and processing protocol as in the model development stage, we collected another 338,084 GSV flat images to predict NO_2 air pollution for Washington DC from 2007 to 2019 (Figure S11). Both the GSV-only and GSV + OMI models captured fine-scale gradients of NO_2 concentrations in Washington DC. Consistent with the decreasing national trend in EPA-monitored NO_2 concentrations, our prediction maps also show reduced NO_2 concentrations over time. The most polluted area is located in the central city, and the concentrations gradually decrease elsewhere. Major gaps in the surface are over rivers or large green spaces since most GSV images are collected by GSV cars driving along roads. Generally, GSV imagery has good coverage in this region, and our GSV image-based models were able to predict 88.1 ± 1.6% of the city.

We compared our GSV models to another national NO_2 model developed using a conventional LUR approach (Figure 4). We found that the GSV-only model predicted the highest NO_2 concentrations in the central city in Washington DC (the GSV + OMI model predicted slightly lower concentrations). The image-based models predicted significantly higher concentrations than the conventional LUR for the same region (Figure 4B). Conversely, in the less polluted areas (e.g., city parks and rivers), the GSV-only model predicted the lowest NO_2 concentrations, followed by GSV + OMI model; the conventional model predicted the highest concentrations. Previous studies show the tendency of LURs to underestimate in polluted regions and overestimate in clean regions,[29,54,55] which could be partly the result of uncaptured predictors in the model (e.g., street-level features). Our results suggest that leveraging street view images for air quality modeling may have

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

the potential to mitigate this tendency and better capture intra-urban NO_2 variability than other LUR methods. We found similar differences in another example city—San Francisco (Figure S12). Future work could compare the results with mobile monitoring campaigns to investigate whether GSV image-based models capture more spatial variability of air pollution. Identifying the impact of different features on air pollution may provide important evidence for making local and regional air quality policies. We also plotted the distribution of GSV feature groups along a transect line (Figures 4C and S12C). NO_2 predictions appear to show intuitive associations with the built environment—a surrogate for human activities. Figure S13 shows the distribution of GSV feature groups over the DC area in 2019 for illustration. The spatial patterns of different feature groups appear to be reasonably consistent with the land use characteristics of DC. For example, the transport vehicles and built environment features are highlighted in the central downtown area, while transport networks are more dispersed.

LIMITATIONS AND IMPLICATIONS FOR FUTURE STUDIES

Our results demonstrate promising performance for GSV image-based air quality modeling, i.e., the potential to capture fine-scale NO_2 using a uniform modeling pipeline that can be easily scaled. The major limitation of our approach is the imbalance in image availability over time and space. This limitation is most prominent in early years and becomes less severe in more recent years, as more images are available. Most GSV images are collected by Google GSV cars (with a smaller number shared by Google users). Depending on priorities set by the entity collecting images (in this case, Google), the street-level imagery is updated at different frequencies over different regions and time periods. We are not able to fully utilize all monitoring sites from the EPA monitoring network for model development since some EPA sites (especially in the rural Midwest and Rocky Mountains areas) do not have GSV images. Similarly, the model application is affected by image availability, and gaps in prediction maps may exist. However, since a priority is to collect images in areas where people live, the imagery does include the vast majority of the U.S. population. Another limitation is our choice to relax the temporal matching criteria between air pollutant observations and GSV images. While we believe this approach is reasonable given the slow changes in the built environment in developed countries (and the results indicate that the added images increase model performance), it is difficult to quantify the model uncertainty resulting from this choice. The lack of temporal resolution in the GSV imagery makes it challenging to use our approach to develop monthly or daily average models.

The satellite-based NO_2 variable has limitations similar to other air quality models that include satellite-based observations. For example, OMI has a coarse spatial resolution, which is not capable of capturing local hot spots or intra-urban variation of NO_2 concentrations. In addition, satellite-derived products reflect but do not offer exact attributes of the ground-level concentrations; retrieval of satellite-derived data involves many assumptions which may introduce uncertainties in the model.[56] Data availability and quality can be affected by multiple factors, e.g., errors in NO_2 column retrievals are large during cloudy days or over snow/ice surfaces.[52] All of these uncertainties may affect model performance.

Our models could be enhanced in several ways. We were only able to apply our models for two cities for illustration. As retrieval of large volumes of images becomes more common, our future work may map NO_2 concentrations at the national scale. Our models show reductions in model performance when using spatial cross-validation. Although other studies also report decreases in performance when using spatial cross-validation, the reductions were smaller (0.09−0.20).[21] Future work may mitigate this issue by extracting GSV features from a larger spatial buffer around each monitoring station, including interpolation components (e.g., kriging) in the modeling framework or applying more advanced deep learning models to directly learn NO_2 pollution-related features from the street view images.[57]

Another promising direction is to utilize air pollution data from denser air quality monitoring networks that better match the granularity of the information in the street view images. For example, utilizing information from mobile monitoring studies that identify localized air quality patterns may unlock additional information in the street-level imagery that traditional, GIS-based predictors cannot. As mobile monitoring, low-cost sensing, and the frequency of image collection grows in future years, we expect this will be a significant area of research in the near future.

Our results suggest the promise of leveraging street view imagery and satellite data in building large-scale empirical air quality models (e.g., national or global) under a unified framework. Unlike traditional LUR predictors collected from multiple curated geodatabases, our approach requires few (e.g., single source for GSV-only models) data sources that achieve comparably high model performance to previous models. Adding satellite observations is a simple way to further boost model performance. Both data sources are publicly accessible and have a uniform data format, which makes them excellent candidates in establishing a consistent data collection and processing protocol for model development over large geographies. This may be especially useful for less developed areas where municipal geodatabases are sparse or absent. Our study also presents a cost-efficient image sampling strategy (i.e., one image every 100 m) developed through a systematic sensitivity analysis—this finding may be helpful for future image-based modeling work. Our modeling approach is not restricted to GSV imagery. Many other large open-access street view image databases have emerged globally, e.g., Bing Maps Streetside, Apple Look Around, Tencent Maps (China), Daum Maps (South Korea), CycloMedia (Europe and U.S.), etc.[58] Our modeling framework can be applied to any street-level image. We expect that street view image-based air quality modeling may play a more important role in the near future given the fast growth of image resources.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.2c03581.

Information on description of spatial and temporal cross-validation; results of spatial and temporal cross-validation; predictor variables for model development (Table S1); scenarios for GSV-NO_2 temporal match during model development (Table S2); distribution of NO_2 EPA monitors by urban, suburban, rural using 2019 for illustration (Table S3); descriptive statistics of GSV

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

features within 500 m of NO_2 monitors stratified by land use types (Table S4); model performance of image-unlimited models using the same training size in five scenarios (Table S5); model performance of image-limited models which used various maximum number of GSV images per 100 m × 100 m grid cell (Table S6); model performance of image-limited models using different grid sizes for sampling a maximum of one GSV image per grid cell (Table S7); decreasing trend of annual NO_2 concentrations during 2007 to 2019 from the EPA regulatory network (Figure S1); number of available NO_2-GSV pairs and GSV images in different years using scenario 1 and 5 for illustration (Figure S2); spatial distribution of NO_2-GSV pairs in different years using scenario 5 for illustration (Figure S3); comparison of root-mean-squared errors and mean absolute errors between GSV-only and GSV + OMI image-unlimited models (Figure S4); hexagonal binned plots of model predictions versus observations for both GSV and GSV + OMI models among different scenarios (Figure S5); spatial distribution of NO_2 monitors with lowest and highest 5% errors (Figure S6); rank of feature importance for GSV-only and GSV + OMI models among different scenarios (Figure S7); distribution of the top 10 important features of GSV-only and GSV + OMI models among different scenarios (Figure S8); accumulated local effects of several top important GSV features for the GSV-only model using scenario 5 for illustration (Figure S9); comparison of three cross validation strategies for GSV-only and GSV + OMI models using scenario 5 for illustration (Figure S10); estimated NO_2 concentrations for Washington DC by GSV-only and GSV-OMI models using scenario 5 (Figure S11); comparison between GSV-based models and conventional national LUR models using prediction maps for San Francisco in 2015 (Figure S12); spatial distribution of GSV feature groups for Washington DC in 2019 (Figure S13) (PDF)

AUTHOR INFORMATION

Corresponding Author

Steve Hankey − *School of Public and International Affairs, Virginia Tech, Blacksburg, Virginia 24061, United States;* orcid.org/0000-0002-7530-6077; Phone: 540.231.7508; Email: hankey@vt.edu

Authors

Meng Qi − *School of Public and International Affairs, Virginia Tech, Blacksburg, Virginia 24061, United States;* orcid.org/0000-0003-0700-2773

Kuldeep Dixit − *School of Public and International Affairs, Virginia Tech, Blacksburg, Virginia 24061, United States*

Julian D. Marshall − *Department of Civil & Environmental Engineering, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0003-4087-1209

Wenwen Zhang − *Edward J. Bloustein School of Planning and Public Policy, Rutgers University, New Brunswick, New Jersey 08901, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.2c03581

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant no. 1943705 and as part of CACES, which was supported under Assistance Agreement R835873 awarded by the U.S. Environmental Protection Agency (EPA). It has not been formally reviewed by EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication. The authors thank Drs. Matthew Bechle and Tianjun Lu for their valuable comments on collecting satellite variables.

REFERENCES

(1) Atkinson, R. W.; Butland, B. K.; Anderson, H. R.; Maynard, R. L. Long-term concentrations of nitrogen dioxide and mortality: a meta-analysis of cohort studies. *Epidemiology* **2018**, *29*, 460−472.

(2) Alotaibi, R.; Bechle, M.; Marshall, J. D.; Ramani, T.; Zietsman, J.; Nieuwenhuijsen, M. J.; Khreis, H. Traffic related air pollution and the burden of childhood asthma in the contiguous United States in 2000 and 2010. *Environ. Int.* **2019**, *127*, 858−867.

(3) Collart, P.; Dubourg, D.; Levêque, A.; Sierra, N. B.; Coppieters, Y. Short-term effects of nitrogen dioxide on hospital admissions for cardiovascular disease in Wallonia, Belgium. *Int. J. Cardiol.* **2018**, *255*, 231−236.

(4) He, M. Z.; Kinney, P. L.; Li, T. T.; Chen, C.; Sun, Q. H.; Ban, J.; Wang, J. N.; Liu, S. L.; Goldsmith, J.; Kioumourtzoglou, M. A. Short- and intermediate-term exposure to NO_2 and mortality: A multi-county analysis in China. *Environ. Pollut.* **2020**, *261*, No. 114165.

(5) Achakulwisut, P.; Brauer, M.; Hystad, P.; Anenberg, S. C. Global, national, and urban burdens of paediatric asthma incidence attributable to ambient NO2 pollution: estimates from global datasets. *Lancet Planet. Health* **2019**, *3*, E166−E178.

(6) Crouse, D. L.; Peters, P. A.; Hystad, P.; Brook, J. R.; van Donkelaar, A.; Martin, R. V.; Villeneuve, P. J.; Jerrett, M.; Goldberg, M. S.; Pope, C. A.; Brauer, M.; Brook, R. D.; Robichaud, A.; Menard, R.; Burnett, R. T. Ambient $PM_{2.5}$, O_3, and NO_2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ. Health Perspect.* **2015**, *123*, 1180−1186.

(7) Demetillo, M. A. G.; Navarro, A.; Knowles, K. K.; Fields, K. P.; Geddes, J. A.; Nowlan, C. R.; Janz, S. J.; Judd, L. M.; Al-Saadi, J.; Sun, K.; McDonald, B. C.; Diskin, G. S.; Pusede, S. E. Observing nitrogen dioxide air pollution inequality using high-spatial-resolution remote sensing measurements in Houston, Texas. *Environ. Sci. Technol.* **2020**, *54*, 9882−9895.

(8) Karner, A. A.; Eisinger, D. S.; Niemeier, D. A. Near-roadway air quality: synthesizing the findings from real-world data. *Environ. Sci. Technol.* **2010**, *44*, 5334−5344.

(9) Zhou, Y.; Levy, J. I. Factors influencing the spatial extent of mobile source air pollution impacts: a meta-analysis. *BMC Public Health* **2007**, *7*, No. 89.

(10) Apte, J. S.; Messier, K. P.; Gani, S.; Brauer, M.; Kirchstetter, T. W.; Lunden, M. M.; Marshall, J. D.; Portier, C. J.; Vermeulen, R. C. H.; Hamburg, S. P. High-resolution air pollution mapping with Google Street View cars: Exploiting big data. *Environ. Sci. Technol.* **2017**, *51*, 6999−7008.

(11) Chambliss, S. E.; Pinon, C. P.; Messier, K. P.; LaFranchi, B.; Upperman, C. R.; Lunden, M. M.; Robinson, A. L.; Marshall, J. D.; Apte, J. S. Local-and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2109249118.

(12) Southerland, V. A.; Anenberg, S. C.; Harris, M.; Apte, J.; Hystad, P.; van Donkelaar, A.; Martin, R. V.; Beyers, M.; Roy, A.

Assessing the distribution of air pollution health risks within cities: A neighborhood-scale analysis leveraging high-resolution data sets in the Bay Area, California. *Environ. Health Perspect.* **2021**, *129*, No. 037006.
(13) 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*, 7561−7578.
(14) Hankey, S.; Marshall, J. D. Land use regression models of on-road particulate air pollution (particle number, black carbon, $PM_{2.5}$, particle size) using mobile monitoring. *Environ. Sci. Technol.* **2015**, *49*, 9194−9202.
(15) De Hoogh, K.; Chen, J.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Bauwelinck, M.; van Donkelaar, A.; Hvidtfeldt, U. A.; Katsouyanni, K.; et al. Spatial $PM_{2.5}$, NO_2, O_3 and BC models for Western Europe−Evaluation of spatiotemporal stability. *Environ. Int.* **2018**, *120*, 81−92.
(16) Messier, K. P.; Chambliss, S. E.; Gani, S.; Alvarez, R.; Brauer, M.; Choi, J. J.; Hamburg, S. P.; Kerckhoffs, J.; LaFranchi, B.; Lunden, M. M.; Marshall, J. D.; Portier, C. J.; Roy, A.; Szpiro, A. A.; Vermeulen, R. C. H.; Apte, J. S. Mapping air pollution with Google Street View cars: Efficient approaches with mobile monitoring and land use regression. *Environ. Sci. Technol.* **2018**, *52*, 12563−12572.
(17) Liao, K.; Huang, X. H.; Dang, H. F.; Ren, Y.; Zuo, S. D.; Duan, C. S. Statistical approaches for forecasting primary air pollutants: A review. *Atmosphere* **2021**, *12*, No. 686.
(18) Kim, S. Y.; Bechle, M.; Hankey, S.; Sheppard, L.; Szpiro, A. A.; Marshall, J. D. Concentrations of criteria pollutants in the contiguous US, 1979-2015: Role of prediction model parsimony in integrated empirical geographic regression. *PLoS One* **2020**, *15*, No. e0228535.
(19) Novotny, E. V.; Bechle, M. J.; Millet, D. B.; Marshall, J. D. National satellite-based land-use regression: NO_2 in the United States. *Environ. Sci. Technol.* **2011**, *45*, 4407−4414.
(20) Bechle, M. J.; Millet, D. B.; Marshall, J. D. National spatiotemporal exposure surface for NO_2: monthly scaling of a satellite-derived land-use regression, 2000−2010. *Environ. Sci. Technol.* **2015**, *49*, 12297−12305.
(21) Young, M. T.; Bechle, M. J.; Sampson, P. D.; Szpiro, A. A.; Marshall, J. D.; Sheppard, L.; Kaufman, J. D. Satellite-based NO_2 and model validation in a national prediction model based on universal kriging and land-use regression. *Environ. Sci. Technol.* **2016**, *50*, 3686−3694.
(22) 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.
(23) Knibbs, L. D.; Van Donkelaar, A.; Martin, R. V.; Bechle, M. J.; Brauer, M.; Cohen, D. D.; Cowie, C. T.; Dirgawati, M.; Guo, Y.; Hanigan, I. C.; et al. Satellite-based land-use regression for continental-scale long-term ambient $PM_{2.5}$ exposure assessment in Australia. *Environ. Sci. Technol.* **2018**, *52*, 12445−12455.
(24) Vienneau, D.; De Hoogh, K.; Bechle, M. J.; Beelen, R.; Van 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 PM_{10}. *Environ. Sci. Technol.* **2013**, *47*, 13555−13564.
(25) de Hoogh, K.; Gulliver, J.; van Donkelaar, A.; Martin, R. V.; Marshall, J. D.; Bechle, M. J.; Cesaroni, G.; Pradas, M. C.; Dedele, A.; Eeftens, M.; et al. Development of West-European $PM_{2.5}$ and NO_2 land use regression models incorporating satellite-derived and chemical transport modelling data. *Environ. Res.* **2016**, *151*, 1−10.
(26) Xu, H.; Bechle, M. J.; Wang, M.; Szpiro, A. A.; Vedal, S.; Bai, Y.; Marshall, J. D. National $PM_{2.5}$ and NO_2 exposure models for China based on land use regression, satellite measurements, and universal kriging. *Sci. Total Environ.* **2019**, *655*, 423−433.
(27) Zhang, Z.; Wang, J.; Hart, J. E.; Laden, F.; Zhao, C.; Li, T.; Zheng, P.; Li, D.; Ye, Z.; Chen, K. National scale spatiotemporal land-use regression model for $PM_{2.5}$, PM_{10} and NO_2 concentration in China. *Atmos. Environ.* **2018**, *192*, 48−54.
(28) Ryan, P. H.; LeMasters, G. K. A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicol.* **2007**, *19*, 127−133.
(29) Larkin, A.; Geddes, J. A.; Martin, R. V.; Xiao, Q. Y.; Liu, Y.; Marshall, J. D.; Brauer, M.; Hystad, P. Global land use regression model for nitrogen dioxide air pollution. *Environ. Sci. Technol.* **2017**, *51*, 6957−6964.
(30) Hong, K. Y.; Pinheiro, P. O.; Minet, L.; Hatzopoulou, M.; Weichenthal, S. Extending the spatial scale of land use regression models for ambient ultrafine particles using satellite images and deep convolutional neural networks. *Environ. Res.* **2019**, *176*, No. 108513.
(31) Weissert, L.; Salmond, J.; Miskell, G.; Alavi-Shoshtari, M.; Williams, D. Development of a microscale land use regression model for predicting NO_2 concentrations at a heavy trafficked suburban area in Auckland, NZ. *Sci. Total Environ.* **2018**, *619−620*, 112−119.
(32) Lu, T.; Marshall, J. D.; Zhang, W.; Hystad, P.; Kim, S.-Y.; Bechle, M. J.; Demuzere, M.; Hankey, S. National empirical models of air pollution using microscale measures of the urban environment. *Environ. Sci. Technol.* **2021**, *55*, 15519−15530.
(33) Lloyd, M.; Carter, E.; Diaz, F. G.; Magara-Gomez, K. T.; Hong, K. Y.; Baumgartner, J.; Herrera G, V. cM.; Weichenthal, S. Predicting within-city spatial variations in outdoor ultrafine particle and black carbon concentrations in Bucaramanga, Colombia: A hybrid approach using open-source geographic data and digital images. *Environ. Sci. Technol.* **2021**, *55*, 12483−12492.
(34) Zhong, S.; Zhang, K.; Bagheri, M.; Burken, J. G.; Gu, A.; Li, B.; Ma, X.; Marrone, B. L.; Ren, Z. J.; Schrier, J.; et al. Machine learning: New ideas and tools in environmental science and engineering. *Environ. Sci. Technol.* **2021**, *55*, 12741−12754.
(35) Weichenthal, S.; Hatzopoulou, M.; Brauer, M. A picture tells a thousand···exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology. *Environ. Int.* **2019**, *122*, 3−10.
(36) Qi, M.; Hankey, S. Using street view imagery to predict street-level particulate air pollution. *Environ. Sci. Technol.* **2021**, *55*, 2695−2704.
(37) Hong, K. Y.; Pinheiro, P. O.; Weichenthal, S. Predicting outdoor ultrafine particle number concentrations, particle size, and noise using street-level images and audio data. *Environ. Int.* **2020**, *144*, No. 106044.
(38) Ganji, A.; Minet, L.; Weichenthal, S.; Hatzopoulou, M. Predicting traffic-related air pollution using feature extraction from built environment images. *Environ. Sci. Technol.* **2020**, *54*, 10688−10699.
(39) Jain, S.; Presto, A. A.; Zimmerman, N. Spatial Modeling of Daily $PM_{2.5}$, NO_2, and CO Concentrations Measured by a Low-Cost Sensor Network: Comparison of Linear, Machine Learning, and Hybrid Land Use Models. *Environ. Sci. Technol.* **2021**, *55*, 8631−8641.
(40) Lim, C. C.; Kim, H.; Vilcassim, M. J. R.; Thurston, G. D.; Gordon, T.; Chen, L. C.; Lee, K.; Heimbinder, M.; Kim, S. Y. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea. *Environ. Int.* **2019**, *131*, No. 105022.
(41) Vu, T. V.; Shi, Z. B.; Cheng, J.; Zhang, Q.; He, K. B.; Wang, S. X.; Harrison, R. M. Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique. *Atmos. Chem. Phys.* **2019**, *19*, 11303−11314.
(42) He, Q.; Qin, K.; Cohen, J. B.; Loyola, D.; Li, D.; Shi, J.; Xue, Y. Spatially and temporally coherent reconstruction of tropospheric NO_2 over China combining OMI and GOME-2B measurements. *Environ. Res. Lett.* **2020**, *15*, No. 125011.
(43) Kim, M.; Brunner, D.; Kuhlmann, G. Importance of satellite observations for high-resolution mapping of near-surface NO_2 by machine learning. *Remote Sens. Environ.* **2021**, *264*, No. 112573.
(44) Chan, K. L.; Khorsandi, E.; Liu, S.; Baier, F.; Valks, P. Estimation of surface NO_2 concentrations over Germany from TROPOMI satellite observations using a machine learning method. *Remote Sens.* **2021**, *13*, No. 969.

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509

(45) Wu, C.-l.; He, H. d.; Song, R.-f.; Peng, Z.-r. Prediction of air pollutants on roadside of the elevated roads with combination of pollutants periodicity and deep learning method. *Build. Environ.* **2022**, *207*, No. 108436.
(46) Kokkinos, K.; Karayannis, V.; Nathanail, E.; Moustakas, K. A comparative analysis of Statistical and Computational Intelligence methodologies for the prediction of traffic-induced fine particulate matter and NO_2. *J. Cleaner Prod.* **2021**, *328*, No. 129500.
(47) Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M. B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; et al. Assessing NO_2 concentration and model uncertainty with high spatiotemporal resolution across the contiguous United States using ensemble model averaging. *Environ. Sci. Technol.* **2020**, *54*, 1372−1384.
(48) Wong, P.-Y.; Su, H.-J.; Lee, H.-Y.; Chen, Y.-C.; Hsiao, Y.-P.; Huang, J.-W.; Teo, T.-A.; Wu, C.-D.; Spengler, J. D. Using land-use machine learning models to estimate daily NO_2 concentration variations in Taiwan. *J. Cleaner Prod.* **2021**, *317*, No. 128411.
(49) Zhao, H. S.; Shi, J. P.; Qi, X. J.; Wang, X. G.; Jia, J. Y. In *Pyramid Scene Parsing Network*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017; pp 2881−2890.
(50) Levelt, P. F.; Van Den Oord, G. H.; Dobber, M. R.; Malkki, A.; Visser, H.; De Vries, J.; Stammes, P.; Lundell, J. O.; Saari, H. The ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1093−1101.
(51) Hoek, G.; Eeftens, M.; Beelen, R.; Fischer, P.; Brunekreef, B.; Boersma, K. F.; Veefkind, P. Satellite NO_2 data improve national land use regression models for ambient NO2 in a small densely populated country. *Atmos. Environ.* **2015**, *105*, 173−180.
(52) Lamsal, L. N.; Krotkov, N. A.; Marchenko, S. V.; Joiner, J.; Oman, L.; Vasilkov, A.; Fisher, B.; Qin, W.; Yang, E.-S.; Fasnacht, Z.; Choi, S.; Leonard, P.; Haffner, D. *OMI/Aura NO_2 Tropospheric, Stratospheric & Total Columns MINDS Daily L3 Global Gridded 0.25 degree x 0.25 degree*; NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC), 2020.
(53) Hart, J. E.; Yanosky, J. D.; Puett, R. C.; Ryan, L.; Dockery, D. W.; Smith, T. J.; Garshick, E.; Laden, F. Spatial modeling of PM_{10} and NO_2 in the continental United States, 1985−2000. *Environ. Health Perspect.* **2009**, *117*, 1690−1696.
(54) Gulliver, J.; de Hoogh, K.; Hansell, A.; Vienneau, D. Development and Back-Extrapolation of NO_2 Land Use Regression Models for Historic Exposure Assessment in Great Britain. *Environ. Sci. Technol.* **2013**, *47*, 7804−7811.
(55) Cordioli, M.; Pironi, C.; De Munari, E.; Marmiroli, N.; Lauriola, P.; Ranzi, A. Combining land use regression models and fixed site monitoring to reconstruct spatiotemporal variability of NO_2 concentrations over a wide geographical area. *Sci. Total Environ.* **2017**, *574*, 1075−1084.
(56) Zheng, T.; Bergin, M. H.; Hu, S.; Miller, J.; Carlson, D. E. Estimating ground-level $PM_{2.5}$ using micro-satellite images by a convolutional neural network and random forest approach. *Atmos. Environ.* **2020**, *230*, No. 117451.
(57) Weichenthal, S.; Dons, E.; Hong, K. Y.; Pinheiro, P. O.; Meysman, F. J. Combining citizen science and deep learning for large-scale estimation of outdoor nitrogen dioxide concentrations. *Environ. Res.* **2021**, *196*, No. 110389.
(58) Lefevre, S.; Tuia, D.; Wegner, J. D.; Produit, T.; Nassar, A. S. Toward seamless multiview scene analysis from satellite to street level. *Proc. IEEE* **2017**, *105*, 1884−1899.

https://doi.org/10.1021/acs.est.2c03581
Environ. Sci. Technol. 2022, 56, 13499−13509