Supporting Information for

Intercomparison of six national empirical models for $PM_{2.5}$ air pollution in the contiguous US

Matthew J. Bechle[1], Michelle L. Bell[2], Daniel L. Goldberg[3], Steve Hankey[4], Tianjun Lu[5], Albert A. Presto[6], Allen L. Robinson[6], Joel Schwartz[7], Liuhua Shi[7,8], Yang Zhang[9], Julian D. Marshall[1,*]

[1] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA
[2] Yale School of Forestry & Environmental Studies, Yale University, New Haven, CT, USA
[3] Department of Environmental and Occupational Health, George Washington University, Washington, DC, USA
[4] School of Public and International Affairs, Virginia Tech, Blacksburg, VA, USA
[5] Department of Earth Science and Geography, California State University, Dominguez Hills, Carson, CA, USA
[6] Department of Mechanical Engineering and Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA
[7] Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
[8] Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA
[9] Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115, USA
[*] Corresponding author. Email: jdmarsh@uw.edu.

Correspondence to: jdmarsh@uw.edu

The supporting information include:

Introduction and background

Empirical models are used to understand and predict levels of outdoor air pollution, including at unmeasured locations and typically with good spatial precision. The name ("empirical") emphasizes that the models reflect empirical measurements. Such model results have been used, for example, in health risk assessment, environmental epidemiology, and environmental justice analysis.

Generating empirical-model results typically involves three steps: (1) *Model building*: generating an empirical model to predict measured concentrations (i.e., the dependent variable; the model is calibrated to and attempts to predict these), using several parameters that might correlate with concentrations (i.e., the potential independent variables). (2) *Model testing*, to quantify parameters such as uncertainty, robustness, error, and bias. If multiple models were built by a research group, the model-testing phase could involve a model-selection process. Hold-out cross-validation typically occurs in this step. (3) *Model application*, wherein the final selected model(s) is used to estimate concentrations throughout the domain of interest (e.g., at all Census Block centroids in the continuous US).

Early empirical models were developed at the urban-scale, using land-use variables (e.g., road locations, industrial locations) and linear regression, and hence were called "land-use regression" (LUR).[1–6] Subsequent developments include (1) adding many more independent variables, including microscale point-of-interest sources,[7,8] satellite-derived estimates for pollution (e.g., atmospheric column totals) or land-cover,[9–11] and predictions from chemical transport models,[10,12] (2) deriving independent variables from imagery (Google Street View images or satellite images) or using images directly via machine learning rather than first obtaining specific independent variables,[13–18] (3) employing more-advanced mathematics rather than linear regression,[19–23] (4) quantifying temporal variability,[24,25] and (5) using a national or international, rather than urban, spatial domain.[26–31] For the dependent variable, early models often employed purposefully-placed passive NO_2 samplers;[32–34] subsequent developments include using regulatory monitoring data,[27,28,30] mobile monitoring,[20,35–40] and freely-available data from ubiquitous low-cost sensors already deployed by the public.[41,42] Strengths of empirical models include excellent spatial

precision in model-predictions, and being based in observations and datasets (e.g., land-use datasets; satellite observations); weaknesses include the long list of input data required (which is unavailable for the future and for the distant past), inability to allow "what if?" questions (e.g., "what if emissions were reduced?"), and potential error/bias in the measurements and in the models. Empirical models reflect *typical* correlations with specific land uses, so they are weaker at detecting atypical conditions (e.g., an unexpected hotspot) or conditions not well correlated with a land-use in the model.

Studies to intercompare empirical models are scarce, especially for large geographies. Some studies have compared empirical models with mechanistic models (e.g., CMAQ),[4,43] satellite-based models (e.g., aerosol optical depth, AOD),[44,45] or hybrid models.[46,47] Other studies have compared results using different methods for model-building (e.g., LUR vs. machine learning vs. kriging vs. hybrid empirical models).[48–50] However, most prior comparisons were at the city or region level, and comparisons were generally within a single research team. We identified only one study that compared empirical models nationwide.[16]

Additional considerations regarding the methods

Our approach is to intercompare a sample of six national empirical models[51–59] for annual-average ambient PM2.5. We focused on annual-averages for fine particles (PM2.5) for several reasons: PM2.5 is an important criteria-pollutant, regulated by the US EPA through a health-based National Ambient Air Quality Standard (NAAQS); millions of people in the US live in areas that exceed the NAAQS (US EPA, 2022a); and the health effects associated with annual-average PM2.5 are large. Importantly, multiple national empirical models predict annual-average PM2.5 available for this pollutant.

In general, one way to intercompare models would be for all modelers to pre-agree to a set of model-building and model-testing observations. (Or, if there were a set of measurements that no model included in model-building — e.g., a dataset that was unknown or otherwise unused — then the outcome would be

similar: a dataset that could be used to test all of the models.) In this case, it would be possible to compare each method against the held-out cross-validation measurements. However, in the current intercomparison, each research group used their own held-out data, comparison metrics, and approach to investigate model uncertainty. Furthermore, the models incorporate the monitoring data in different ways (e.g., via a kriging component); for that reason, simply comparing the six models against observations (which were used during model-building) may not shed light on model reliability at locations without measurements.

Instead, we directly intercompare the models, without comparing against held-out measurements. We do not have "gold-standard" observations to compare against. Nevertheless, we believe that useful insights can be gained from the intercomparisons conducted.

Limitations of this research include the following. (1) We considered one set of spatiotemporal comparisons (annual-average; national/regional/urban-rural) and one set of metrics (RMSE, correlation), but did not compare all possible comparisons (e.g., did not investigate seasonal or daily models, nor sub-regional or local/community model results) or metrics. Other metrics or spatiotemporal representations of the models too may be useful for health, environmental justice, or risk analysis. (2) We have not specifically investigated the fitness of these models for specific purposes, including epidemiological studies, environmental justice studies, public outreach, regulatory analysis, or risk assessment. (3) As mentioned above, we did not compare against measurements; this paper presents only a model-model comparison. Model-model agreement is not the same as a model being "correct". (4) We have identified that the empirical models are relatively consistent with each other, but we have not investigated, within the models themselves, *why*. For example, it may be that the models use the same or similar independent variables; or, it may be that the similarities in model-prediction are despite large differences in independent variables employed.

Strengths of this research include the following. We inter-compared several models, and shed light on similarities and differences nationally, regionally, for urban/rural differences, by pollution level, and by

population density. This is, to our knowledge, the first intercomparison of national empirical models. As noted above, we did not compare against monitors; however, that aspect can partially be viewed as a strength, because the monitoring network is not evenly distributed spatially. Comparisons of models at monitor locations may or may not shed light on concentrations at unmonitored locations; the comparisons here are at Census geographics (Tracts) and so reflect locations where people live. Although $PM_{2.5}$ is generally considered to be a regional pollutant, when considering the whole US there is a wide range of concentrations (here: ~2 to ~15 $\mu g/m^3$, or approximately a 7-fold range).

The models employ different techniques for model building. Some are closer to a linear model, some use machine learning or highly complex mathematical relationships that would be difficult for a human to create or understand. They employ a wide variety of independent variables. However, all of the models use EPA monitoring station data as the model-building dataset. Whatever strengths or weaknesses exist in using EPA monitors (and their locations) for empirical models, those likely impact all of the models.

Sensitivity analyses

We conducted several sensitivity analyses. First, reflecting that SEARCH results are only available in the eastern half of the US, we generated pairwise scatterplots for only the eastern half of the US (Figure S1). Next, we generated separate scatterplots for urban-only (Figure S2) and urban-only in the eastern half of the US (Figure S3) and for rural-only (Figure S4) and for rural-only in the eastern half of the US (Figure S5). We find, for example, that the maximum RMSD is slightly larger for rural areas than for urban areas, a finding that may differ from expectations but is consistent with results described in the main paper (Figure 3E) and may reflect the lower density of monitors in rural areas or that the correlation between concentrations and land use may be lower in rural than in urban areas.

We repeated the analyses in Figure 3 but for the eastern half of the US (Figure S6 and S7). The findings are generally consistent with results in the main paper: the models generally agree with each other. The

range of predictions (a measure of model-model disagreement) is greater at lower-concentration locations than at high-concentration locations.

Figures S1 to S9

Figure S1: Scatterplot matrix for 2010 tract-level $PM_{2.5}$ (E-US only). Same as Fig 1, but for locations in E-US only. Scatterplots in the upper right show pairwise tract-level predictions from each model. Dashed line shows 1:1 line, red line shows linear trendline. Corresponding boxes in the bottom left show Pearson's correlation (*r*; unitless) and root mean squared difference (RMSD; $\mu g/m^3$) between model predictions.

Figure S2: Scatterplot matrix for 2010 urban tract-level $PM_{2.5}$. Same as Fig 1, but for urban locations only. Scatterplots in the upper right show pairwise tract-level predictions from each model. Dashed line shows 1:1 line, red line shows linear trendline. Corresponding boxes in the bottom left show Pearson's correlation (*r*; unitless) and root mean squared difference (RMSD; $\mu g/m^3$) between model predictions.

Figure S3: Scatterplot matrix for 2010 urban tract-level $PM_{2.5}$ (E-US only). Same as Fig 1, but for urban locations in E-US only. Scatterplots in the upper right show pairwise tract-level predictions from each model. Dashed line shows 1:1 line, red line shows linear trendline. Corresponding boxes in the bottom left show Pearson's correlation (*r*; unitless) and root mean squared difference (RMSD; $\mu g/m^3$) between model predictions.

Figure S4: Scatterplot matrix for 2010 rural tract-level $PM_{2.5}$. Same as Fig 1, but for rural locations only. Scatterplots in the upper right show pairwise tract-level predictions from each model. Dashed line shows 1:1 line, red line shows linear trendline. Corresponding boxes in the bottom left show Pearson's correlation (*r*; unitless) and root mean squared difference (RMSD; $\mu g/m^3$) between model predictions.

Figure S5: Scatterplot matrix for 2010 rural tract-level $PM_{2.5}$ (E-US only). Same as Fig 1, but for rural locations in E-US only. Scatterplots in the upper right show pairwise tract-level predictions from each model. Dashed line shows 1:1 line, red line shows linear trendline. Corresponding boxes in the bottom left show Pearson's correlation (*r*; unitless) and root mean squared difference (RMSD; $\mu g/m^3$) between model predictions.

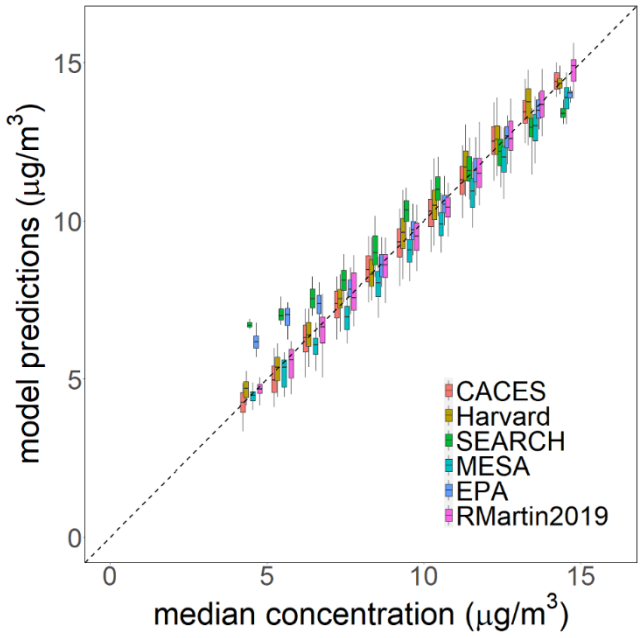

Figure S6: Predictions by model and concentration level (E-US only). Boxplots show range of tract-level model predictions (y-axis) as a function of the median concentration among model predictions within each tract, binned to 1 μg/m³ bins (x-axis). Horizontal bar shows the median, box shows the interquartile range, and vertical lines show the 5th and 95th percentiles of predicted concentrations for each model within each bin.

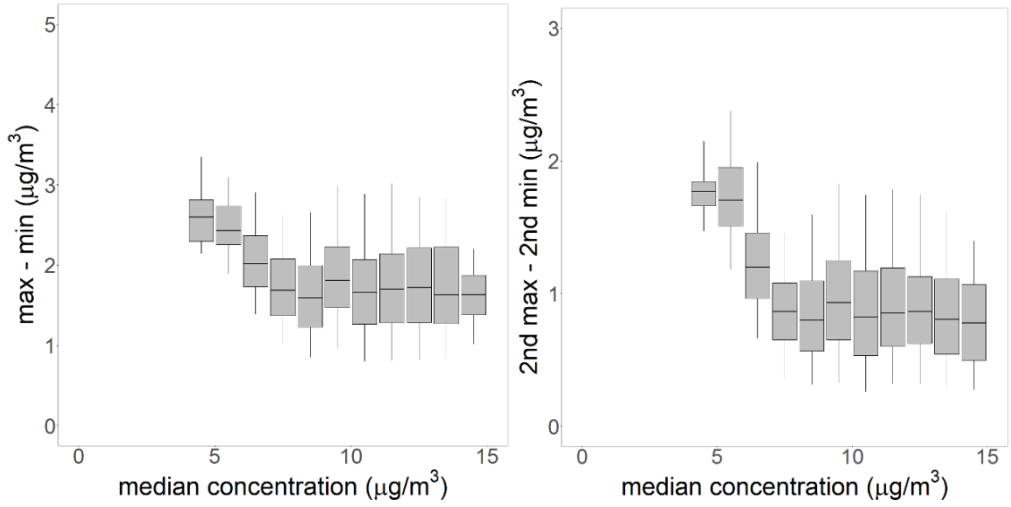

Figure S7: Variability by concentration level (E-US only). Boxplots show within-tract variation (y-axis) as a function of the median concentration among model predictions within each tract, binned to 1 μg/m³ bins (x-axis). Within-tract variation is calculated as either max minus the min (left) or 2nd max minus 2nd min (right) of model predictions within each tract. Horizontal bar shows the median, box shows the interquartile range, and vertical lines show the 5th and 95th percentiles of the variability for tracts within each bin.

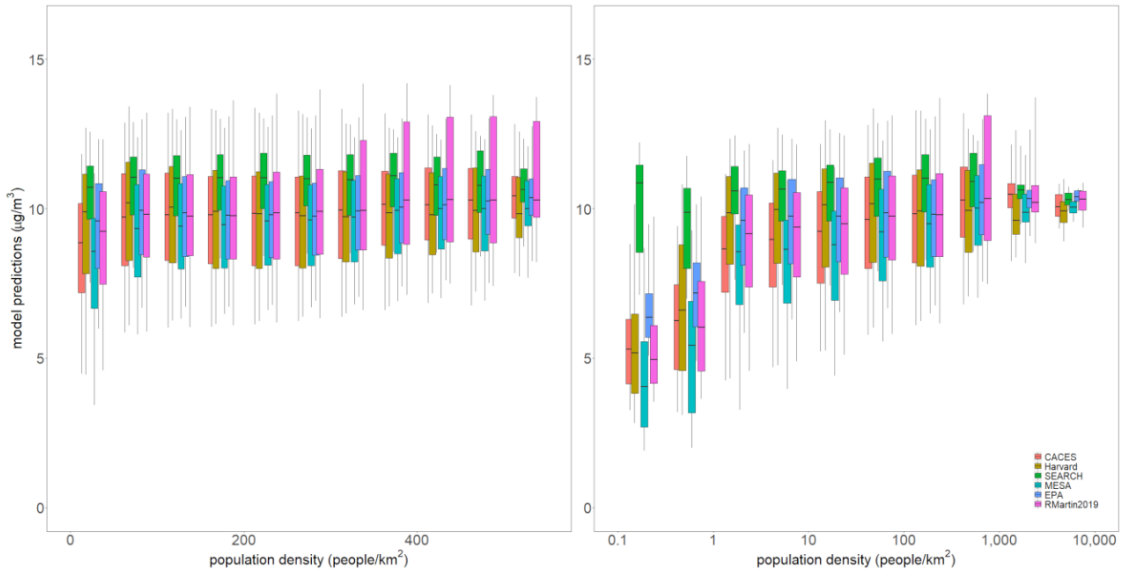

Figure S8: Predictions by model and population density. Boxplots show range of tract-level model predictions (y-axis) as a function of the tract-level population density (x-axis), binned to equally space linear- (left) and log-space (right) bins (note: last bin contains remaining data). Horizontal bar shows the median, box shows the interquartile range, and vertical lines show the 5th and 95th percentiles of predicted concentrations for each model within each bin.

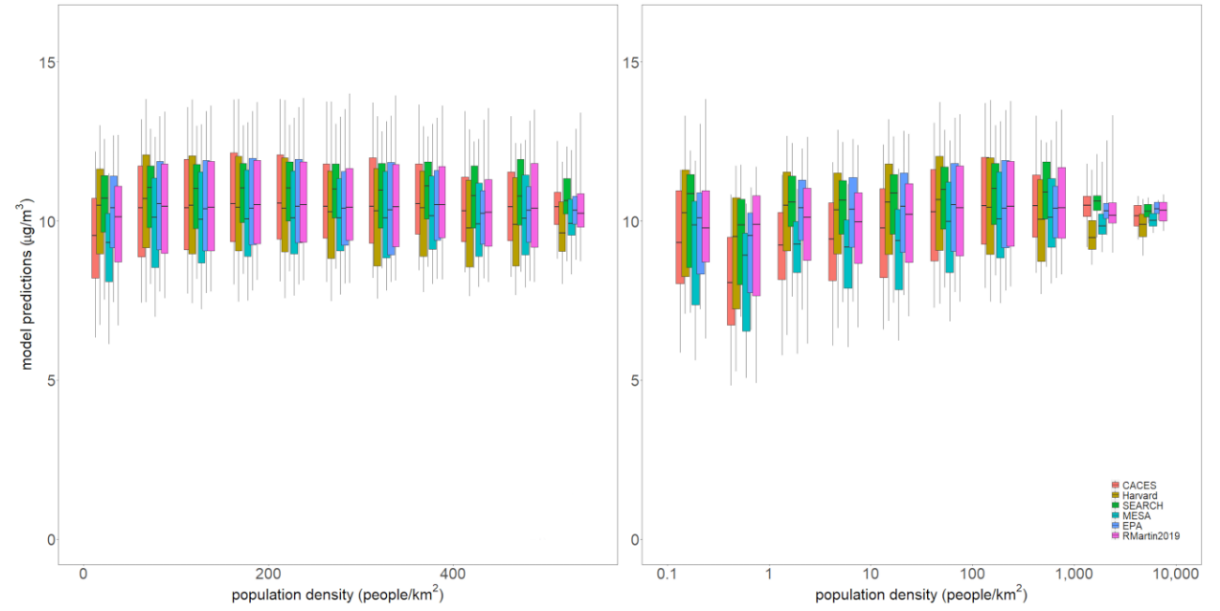

Figure S9: Predictions by model and population density (E-US only). Boxplots show range of tract-level model predictions (y-axis) as a function of the tract-level population density (x-axis), binned to equally space linear- (left) and log-space (right) bins (note: last bin contains remaining data). Horizontal bar shows the median, box shows the interquartile range, and vertical lines show the 5th and 95th percentiles of predicted concentrations for each model within each bin.

References

(1) Brauer, M.; Hoek, G.; van Vliet, P.; Meliefste, K.; Fischer, P.; Gehring, U.; Heinrich, J.; Cyrys, J.; Bellander, T.; Lewne, M. Estimating Long-Term Average Particulate Air Pollution Concentrations: Application of Traffic Indicators and Geographic Information Systems. *Epidemiology* **2003**, 228–239.

(2) 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. *Journal of Exposure Science & Environmental Epidemiology* **2005**, *15* (2), 185–204.

(3) 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. *Atmospheric environment* **2008**, *42* (33), 7561–7578.

(4) Marshall, J. D.; Nethery, E.; Brauer, M. Within-Urban Variability in Ambient Air Pollution: Comparison of Estimation Methods. *Atmospheric Environment* **2008**, *42* (6), 1359–1369.

(5) Su, J.; Buzzelli, M.; Brauer, M.; Gould, T.; Larson, T. Modeling Spatial Variability of Airborne Levoglucosan in Seattle, Washington. *Atmospheric Environment* **2008**, *42* (22), 5519–5525.

(6) Eeftens, M.; Tsai, M. Y.; Ampe, C.; Anwander, B.; Beelen, R.; Bellander, T.; Cesaroni, G.; Cirach, M.; Cyrys, J.; de Hoogh, K.; De Nazelle, A.; de Vocht, F.; Declercq, C.; Dedele, A.; Eriksen, K.; Galassi, C.; Gražulevičiene, R.; Grivas, G.; Heinrich, J.; Hoffmann, B.; Iakovides, M.; Ineichen, A.; Katsouyanni, K.; Korek, M.; Krämer, U.; Kuhlbusch, T.; Lanki, T.; Madsen, C.; Meliefste, K.; Mölter, A.; Mosler, G.; Nieuwenhuijsen, M.; Oldenwening, M.; Pennanen, A.; Probst-Hensch, N.; Quass, U.; Raaschou-Nielsen, O.; Ranzi, A.; Stephanou, E.; Sugiri, D.; Udvardy, O.; Vaskövi, É.; Weinmayr, G.; Brunekreef, B.; Hoek, G. Spatial Variation of PM2.5, PM10, PM2.5 Absorbance and PMcoarse Concentrations between and within 20 European Study Areas and the Relationship with NO2 - Results of the ESCAPE Project. *Atmospheric Environment* **2012**, *62*, 303–317. https://doi.org/10.1016/j.atmosenv.2012.08.038.

(7) Wu, C.-D.; Chen, Y.-C.; Pan, W.-C.; Zeng, Y.-T.; Chen, M.-J.; Guo, Y. L.; Lung, S.-C. C. Land-Use Regression with Long-Term Satellite-Based Greenness Index and Culture-Specific Sources to Model PM2. 5 Spatial-Temporal Variability. *Environmental Pollution* **2017**, *224*, 148–157.

(8) Lu, T.; Lansing, J.; Zhang, W.; Bechle, M. J.; Hankey, S. Land Use Regression Models for 60 Volatile Organic Compounds: Comparing Google Point of Interest (POI) and City Permit Data. *Science of the Total Environment* **2019**, *677* (2), 131–141. https://doi.org/10.1016/j.scitotenv.2019.04.285.

(9) 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. Satellite-Based Land-Use Regression for Continental-Scale Long-Term Ambient PM2. 5 Exposure Assessment in Australia. *Environmental science & technology* **2018**, *52* (21), 12445–12455.

(10) Bechle, M. J.; Millet, D. B.; Marshall, J. D. National Spatiotemporal Exposure Surface for NO2: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–2010. *Environmental science & technology* **2015**, *49* (20), 12297–12305.

(11) 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. Development of West-European PM2. 5 and NO2 Land Use Regression Models Incorporating Satellite-Derived and Chemical Transport Modelling Data. *Environmental research* **2016**, *151*, 1–10.

(12) Goldberg, D. L.; Gupta, P.; Wang, K.; Jena, C.; Zhang, Y.; Lu, Z.; Streets, D. G. Using Gap-Filled MAIAC AOD and WRF-Chem to Estimate Daily PM2. 5 Concentrations at 1 Km Resolution in the Eastern United States. *Atmospheric Environment* **2019**, *199*, 443–452.

(13) 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. *Environmental research* **2019**, *176*, 108513.

(14) 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. *Environment international* **2019**, *122*, 3–10.

(15) Ganji, A.; Minet, L.; Weichenthal, S.; Hatzopoulou, M. Predicting Traffic-Related Air Pollution Using Feature Extraction from Built Environment Images. *Environmental Science & Technology* **2020**, *54* (17), 10688–10699.

(16) Lu, T.; Marshall, J. D.; Zhang, W.; Hystad, P.; Kim, S.; Bechle, M. J.; Demuzere, M.; Hankey, S. National Empirical Models of Air Pollution Using Microscale Measures of the Urban Environment. *Environmental Science and Technology* **2021**, *55*, 15519–15530. https://doi.org/10.1021/acs.est.1c04047.

(17) Qi, M.; Hankey, S. Using Street View Imagery to Predict Street-Level Particulate Air Pollution. *Environmental Science & Technology* **2021**, *55* (4), 2695–2704.

(18) Qi, M.; Dixit, K.; Marshall, J. D.; Zhang, W.; Hankey, S. National Land Use Regression Model for NO2 Using Street View Imagery and Satellite Observations. *Environmental Science & Technology* **2022**, *56* (18), 13499–13509.

(19) 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. *Environmental science & technology* **2013**, *47* (13), 7233–7241.

(20) Weichenthal, S.; Van Ryswyk, K.; Goldstein, A.; Bagg, S.; Shekkarizfard, M.; Hatzopoulou, M. A Land Use Regression Model for Ambient Ultrafine Particles in Montreal, Canada: A Comparison of Linear Regression and a Machine Learning Approach. *Environmental research* **2016**, *146*, 65–72.

(21) Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M. B.; Choirat, C.; Koutrakis, P.; Lyapustin, A. An Ensemble-Based Model of PM2. 5 Concentration across the Contiguous United States with High Spatiotemporal Resolution. *Environment international* **2019**, *130*, 104909.

(22) Lautenschlager, F.; Becker, M.; Kobs, K.; Steininger, M.; Davidson, P.; Krause, A.; Hotho, A. OpenLUR: Off-the-Shelf Air Pollution Modeling with Open Features and Machine Learning. *Atmospheric environment* **2020**, *233*, 117535.

(23) Wong, P.-Y.; Lee, H.-Y.; Chen, Y.-C.; Zeng, Y.-T.; Chern, Y.-R.; Chen, N.-T.; Lung, S.-C. C.; Su, H.-J.; Wu, C.-D. Using a Land Use Regression Model with Machine Learning to Estimate Ground Level PM2. 5. *Environmental Pollution* **2021**, *277*, 116846.

(24) Wu, C.-D.; Zeng, Y.-T.; Lung, S.-C. C. A Hybrid Kriging/Land-Use Regression Model to Assess PM2. 5 Spatial-Temporal Variability. *Science of the Total Environment* **2018**, *645*, 1456–1464.

(25) Masiol, M.; Squizzato, S.; Chalupa, D.; Rich, D. Q.; Hopke, P. K. Spatial-Temporal Variations of Summertime Ozone Concentrations across a Metropolitan Area Using a Network of Low-Cost Monitors

to Develop 24 Hourly Land-Use Regression Models. *Science of The Total Environment* **2019**, *654*, 1167–1178.

(26) Hoek, G.; Eeftens, M.; Beelen, R.; Fischer, P.; Brunekreef, B.; Boersma, K. F.; Veefkind, P. Satellite NO2 Data Improve National Land Use Regression Models for Ambient NO2 in a Small Densely Populated Country. *Atmospheric Environment* **2015**, *105*, 173–180.

(27) Hystad, P.; Setton, E.; Cervantes, A.; Poplawski, K.; Deschenes, S.; Brauer, M.; van Donkelaar, A.; Lamsal, L.; Martin, R.; Jerrett, M. Creating National Air Pollution Models for Population Exposure Assessment in Canada. *Environmental health perspectives* **2011**, *119* (8), 1123–1129.

(28) Novotny, E. V.; Bechle, M. J.; Millet, D. B.; Marshall, J. D. National Satellite-Based Land-Use Regression: NO2 in the United States. *Environmental science & technology* **2011**, *45* (10), 4407–4414.

(29) 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. *Environmental research* **2014**, *135*, 204–211.

(30) Larkin, A.; Geddes, J. A.; Martin, R. V.; Xiao, Q.; Liu, Y.; Marshall, J. D.; Brauer, M.; Hystad, P. Global Land Use Regression Model for Nitrogen Dioxide Air Pollution. *Environmental science & technology* **2017**, *51* (12), 6957–6964.

(31) Saha, P. K.; Hankey, S.; Marshall, J. D.; Robinson, A. L.; Presto, A. A. High-Spatial-Resolution Estimates of Ultrafine Particle Concentrations across the Continental United States. *Environmental Science & Technology* **2021**, *55* (15), 10320–10331.

(32) Aguilera, I.; Sunyer, J.; Fernández-Patier, R.; Hoek, G.; Aguirre-Alfaro, A.; Meliefste, K.; Bomboi-Mingarro, M. T.; Nieuwenhuijsen, M. J.; Herce-Garraleta, D.; Brunekreef, B. Estimation of Outdoor NO x, NO2, and BTEX Exposure in a Cohort of Pregnant Women Using Land Use Regression Modeling. *Environmental science & technology* **2008**, *42* (3), 815–821.

(33) Wang, M.; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G.; De Hoogh, K.; Dedele, A.; Declercq, C.; Dimakopoulou, K.; Eeftens, M.; Forastiere, F.; Galassi, C.; Gražulevičiene, R.; Hoffmann, B.; Heinrich, J.; Iakovides, M.; Künzli, N.; Korek, M.; Lindley, S.; Mölter, A.; Mosler, G.; Madsen, C.; Nieuwenhuijsen, M.; Phuleria, H.; Pedeli, X.; Raaschou-Nielsen, O.; Ranzi, A.; Stephanou, E.; Sugiri, D.; Stempfelet, M.; Tsai, M. Y.; Lanki, T.; Udvardy, O.; Varró, M. J.; Wolf, K.; Weinmayr, G.; Yli-Tuomi, T.; Hoek, G.; Brunekreef, B. Evaluation of Land Use Regression Models for NO2 and Particulate Matter in 20 European Study Areas: The ESCAPE Project. *Environmental Science and Technology* **2013**, *47* (9), 4357–4364. https://doi.org/10.1021/es305129t.

(34) Lee, M.; Brauer, M.; Wong, P.; Tang, R.; Tsui, T. H.; Choi, C.; Cheng, W.; Lai, P.-C.; Tian, L.; Thach, T.-Q. Land Use Regression Modelling of Air Pollution in High Density High Rise Cities: A Case Study in Hong Kong. *Science of the Total Environment* **2017**, *592*, 306–315.

(35) Larson, T.; Su, J.; Baribeau, A.-M.; Buzzelli, M.; Setton, E.; Brauer, M. A Spatial Model of Urban Winter Woodsmoke Concentrations. *Environmental science & technology* **2007**, *41* (7), 2429–2436.

(36) Thai, A.; McKendry, I.; Brauer, M. Particulate Matter Exposure along Designated Bicycle Routes in Vancouver, British Columbia. *Science of the total environment* **2008**, *405* (1–3), 26–35.

(37) 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. *Environmental Science and Technology* **2015**, *49*, 9194–9202. https://doi.org/10.1021/acs.est.5b01209.

(38) Messier, K. P.; Chambliss, S. E.; Gani, S.; Alvarez, R.; Brauer, M.; Choi, J. J.; Hamburg, S. P.; Kerckho, 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 : E Ffi Cient Approaches with Mobile Monitoring and Land Use Regression. *Environmental Science and Technology* **2018**, *52*, 12563–12572. https://doi.org/10.1021/acs.est.8b03395.

(39) Minet, L.; Liu, R.; Valois, M.-F.; Xu, J.; Weichenthal, S.; Hatzopoulou, M. Development and Comparison of Air Pollution Exposure Surfaces Derived from On-Road Mobile Monitoring and Short-Term Stationary Sidewalk Measurements. *Environmental science & technology* **2018**, *52* (6), 3512–3519.

(40) Hankey, S.; Sforza, P.; Pierson, M. Using Mobile Monitoring to Develop Hourly Empirical Models of Particulate Air Pollution in a Rural Appalachian Community. *Environmental Science and Technology* **2019**, *53*, 4305–4315. https://doi.org/10.1021/acs.est.8b05249.

(41) Bi, J.; Wildani, A.; Chang, H. H.; Liu, Y. Incorporating Low-Cost Sensor Measurements into High-Resolution PM2.5 Modeling at a Large Spatial Scale. *Environmental Science and Technology* **2020**, *54* (4), 2152–2162. https://doi.org/10.1021/acs.est.9b06046.

(42) Lu, T.; Bechle, M. J.; Wan, Y.; Presto, A. A.; Hankey, S. Using Crowd-Sourced Low-Cost Sensors in a Land Use Regression of PM2.5 in 6 US Cities. *Air Quality, Atmosphere and Health* **2022**, *15* (4), 667–678. https://doi.org/10.1007/s11869-022-01162-7.

(43) Samoli, E.; Butland, B. K.; Rodopoulou, S.; Atkinson, R. W.; Barratt, B.; Beevers, S. D.; Beddows, A.; Dimakopoulou, K.; Schwartz, J. D.; Yazdi, M. D. The Impact of Measurement Error in Modeled Ambient Particles Exposures on Health Effect Estimates in Multilevel Analysis: A Simulation Study. *Environmental Epidemiology* **2020**, *4* (3).

(44) Yu, H.; Russell, A.; Mulholland, J.; Odman, T.; Hu, Y.; Chang, H. H.; Kumar, N. Cross-Comparison and Evaluation of Air Pollution Field Estimation Methods. *Atmospheric environment* **2018**, *179*, 49–60.

(45) Cowie, C. T.; Garden, F.; Jegasothy, E.; Knibbs, L. D.; Hanigan, I.; Morley, D.; Hansell, A.; Hoek, G.; Marks, G. B. Comparison of Model Estimates from an Intra-City Land Use Regression Model with a National Satellite-LUR and a Regional Bayesian Maximum Entropy Model, in Estimating NO2 for a Birth Cohort in Sydney, Australia. *Environmental research* **2019**, *174*, 24–34.

(46) Michanowicz, D. R.; Shmool, J. L.; Tunno, B. J.; Tripathy, S.; Gillooly, S.; Kinnee, E.; Clougherty, J. E. A Hybrid Land Use Regression/AERMOD Model for Predicting Intra-Urban Variation in PM2. 5. *Atmospheric environment* **2016**, *131*, 307–315.

(47) Zhang, X.; Just, A. C.; Hsu, H.-H. L.; Kloog, I.; Woody, M.; Mi, Z.; Rush, J.; Georgopoulos, P.; Wright, R. O.; Stroustrup, A. A Hybrid Approach to Predict Daily NO2 Concentrations at City Block Scale. *Science of the Total Environment* **2021**, *761*, 143279.

(48) Adam-Poupart, A.; Brand, A.; Fournier, M.; Jerrett, M.; Smargiassi, A. Spatiotemporal Modeling of Ozone Levels in Quebec (Canada): A Comparison of Kriging, Land-Use Regression (LUR), and Combined Bayesian Maximum Entropy–LUR Approaches. *Environmental health perspectives* **2014**, *122* (9), 970–976.

(49) Jain, S.; Presto, A. A.; Zimmerman, N. Spatial Modeling of Daily PM2. 5, NO2, and CO Concentrations Measured by a Low-Cost Sensor Network: Comparison of Linear, Machine Learning, and Hybrid Land Use Models. *Environmental Science & Technology* **2021**, *55* (13), 8631–8641.

(50) Dharmalingam, S.; Senthilkumar, N.; D'Souza, R. R.; Hu, Y.; Chang, H. H.; Ebelt, S.; Yu, H.; Kim, C. S.; Rohr, A. Developing Air Pollution Concentration Fields for Health Studies Using Multiple Methods: Cross-Comparison and Evaluation. *Environmental Research* **2022**, *207*, 112207.

(51) 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* (2), e0228535.

(52) Fused Air Quality Surface Using Downscaling (FAQSD), 2022. https://www.epa.gov/hesc/rsig-related-downloadable-data-files.

(53) Keller, J. P.; Olives, C.; Kim, S.-Y.; Sheppard, L.; Sampson, P. D.; Szpiro, A. A.; Oron, A. P.; Lindström, J.; Vedal, S.; Kaufman, J. D. A Unified Spatiotemporal Modeling Approach for Predicting Concentrations of Multiple Air Pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. *Environmental health perspectives* **2015**, *123* (4), 301–309.

(54) Van Donkelaar, A.; Martin, R. V.; Li, C.; Burnett, R. T. Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental science & technology* **2019**, *53* (5), 2595–2611.

(55) EPA Downscaler Model for Predicting Daily Air Pollution, 2017. https://19january2017snapshot.epa.gov/air-research/downscaler-model-predicting-daily-air-pollution_.html.

(56) Sampson, P. D.; Richards, M.; Szpiro, A. A.; Bergen, S.; Sheppard, L.; Larson, T. V.; Kaufman, J. D. A Regionalized National Universal Kriging Model Using Partial Least Squares Regression for Estimating Annual PM2. 5 Concentrations in Epidemiology. *Atmospheric environment* **2013**, *75*, 383–392.

(57) Young, M. T.; Bechle, M. J.; Sampson, P. D.; Szpiro, A. A.; Marshall, J. D.; Sheppard, L.; Kaufman, J. D. Satellite-Based NO2 and Model Validation in a National Prediction Model Based on Universal Kriging and Land-Use Regression. *Environmental science & technology* **2016**, *50* (7), 3686–3694.

(58) Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M. B.; Choirat, C.; Koutrakis, P.; Lyapustin, A. Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging. *Environmental science & technology* **2019**, *54* (3), 1372–1384.

(59) Requia, W. J.; Di, Q.; Silvern, R.; Kelly, J. T.; Koutrakis, P.; Mickley, L. J.; Sulprizio, M. P.; Amini, H.; Shi, L.; Schwartz, J. An Ensemble Learning Approach for Estimating High Spatiotemporal Resolution of Ground-Level Ozone in the Contiguous United States. *Environmental science & technology* **2020**, *54* (18), 11037–11047.