



Full length article

Cokriging with a low-cost sensor network to estimate spatial variation of brake and tire-wear metals and oxidative stress potential in Southern California

Jonathan Liu^a, Sudipto Banerjee^b, Farzan Oroumihyeh^a, Jiaqi Shen^c, Irish del Rosario^d, Jonah Lipsitt^a, Suzanne Paulson^c, Beate Ritz^d, Jason Su^e, Scott Weichenthal^f, Pascale Lakey^g, Manabu Shiraiwa^g, Yifang Zhu^a, Michael Jerrett^{a,*}

^a Department of Environmental Health Sciences, Jonathan and Karin Fielding School of Public Health, University of California, Los Angeles, 650 Charles E Young Dr S, Los Angeles, CA 90095, United States

^b Department of Biostatistics, Jonathan and Karin Fielding School of Public Health, 650 Charles E Young Dr S, Los Angeles, CA 90095, United States

^c Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, 520 Portola Plaza, Los Angeles, CA 90095, United States

^d Department of Epidemiology, Jonathan and Karin Fielding School of Public Health, 650 Charles E Young Dr S, Los Angeles, CA 90095, United States

^e Division of Environmental Health Sciences, School of Public Health, University of California at Berkeley, 2121 Berkeley Way, Berkeley, CA, United States

^f Department of Epidemiology, Biostatistics, and Occupational Health, Faculty of Medicine and Health Sciences, McGill University, 2001 McGill College, Suite 1200, Montreal, QC H3A 1G1, Canada

^g Department of Chemistry, University of California, Irvine, Natural Sciences II, 1102, Irvine, CA 92617, United States

ARTICLE INFO

Handling Editor: Adrian Covaci

Keywords:

Non-exhaust emissions
Low-cost sensors
Land-use regression
Cokriging
Exposure modeling
Heavy metals

ABSTRACT

Due to regulations and technological advancements reducing tailpipe emissions, an increasing proportion of emissions arise from brake and tire wear particulate matter (PM). PM from these non-tailpipe sources contains heavy metals capable of generating oxidative stress in the lung. Although important, these particles remain understudied because the high cost of actively collecting filter samples. Improvements in electrical engineering, internet connectivity, and an increased public concern over air pollution have led to a proliferation of dense low-cost air sensor networks such as the PurpleAir monitors, which primarily measure unspicified fine particulate matter (PM_{2.5}). In this study, we model the concentrations of Ba, Zn, black carbon, reactive oxygen species concentration in the epithelial lining fluid, dithiothreitol (DTT) loss, and OH formation. We use a cokriging approach, incorporating data from the PurpleAir network as a secondary predictor variable and a land-use regression (LUR) as an external drift. For most pollutant species, cokriging models produced more accurate predictions than an LUR model, which did not incorporate data from the PurpleAir monitors. This finding suggests that low-cost sensors can enhance predictions of pollutants that are costly to measure extensively in the field.

1. Introduction

Exposure to particulate matter in fine (PM_{2.5}) and coarse (PM_{2.5–10}) size ranges have been associated with a wide array of adverse human health effects (Sharma et al., 2020). The composition of PM, as a complex, multisource pollutant, varies depending on factors such as season, geography, and time of day (de Jesus et al., 2019; Wang et al., 2020; Wang et al., 2019). Traffic is an important source of ambient PM_{2.5} and

is a dominant contributor to metals found in PM (Grieshop et al., 2006; Habre et al., 2020). Recent vehicle emissions regulations have decreased average tailpipe emissions, but not those from brake and tire wear (CARB, 2020). As a result, the relative importance of non-exhaust vehicle emissions, namely those from brake and tire wear as well as road wear, has increased in recent years (Jeong et al., 2020; Pant and Harrison, 2013). This trend is expected to continue particularly as the sale of electric vehicles continues to grow (Fussell et al., 2022).

* Corresponding author.

E-mail addresses: jliu95@ucla.edu (J. Liu), sudipto@ucla.edu (S. Banerjee), farzor@ucla.edu (F. Oroumihyeh), jqshen@ucla.edu (J. Shen), irishcdr@ucla.edu (I. del Rosario), jonahlipsitt@gmail.com (J. Lipsitt), paulson@atmos.ucla.edu (S. Paulson), britz@ucla.edu (B. Ritz), jasonsu@berkeley.edu (J. Su), scottandrew.weichenthal@mcgill.ca (S. Weichenthal), lakey@uci.edu (P. Lakey), m.shiraiwa@uci.edu (M. Shiraiwa), yifang@ucla.edu (Y. Zhu), mjerrett@ucla.edu (M. Jerrett).

<https://doi.org/10.1016/j.envint.2022.107481>

Received 1 April 2022; Received in revised form 22 July 2022; Accepted 16 August 2022

Available online 23 August 2022

0160-4120/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Brake and tire wear particles are generated due to friction and abrasion caused by braking and normal driving conditions. Studies have found that brake and tire wear particles exist in both the fine and coarse ranges, with diameters ranging from 1–10 μm . Sharp spikes in temperature associated with braking action also generate ultrafine particles, provided that temperatures achieve a critical levels ($140^\circ\text{C} < T < 240^\circ$) (Fussell et al., 2022). Prior studies on roadways and in the laboratory have identified specific metals that can serve as tracers with varying specificity for identifying non-exhaust emissions, including Ba, Cu, and Sb for brake wear particles and Zn for tire wear particles (Amato et al., 2011; Grigoratos and Martini, 2015; Milani et al., 2004; Oroumiyeh et al., 2022; Wang et al., 2021). These metals, in addition to serving as tracers for brake and tire wear, are themselves highly reactive. In surrogate lung fluid, heavy metals such as Cu cause oxidative stress when reacting with H_2O_2 to form OH radicals in a process known as the Fenton reaction (Valko et al., 2005). The presence of ultrafine particulates emitted by brake wear that contain highly-reactive heavy metals may have human health implications, as ultrafine particles when inhaled travel deeper in the lung and can enter the bloodstream. Knowledge concerning the spatial distribution of such elements, however, remains limited.

Land-use regression (LUR) is a well-established technique used for exposure assessment of intraurban air pollution, where regression models include land use, traffic, physical geography, and business density as predictors of measured air pollution concentrations (Briggs, 2005; Hoek et al., 2008; Jerrett et al., 2005). Model derived coefficients can then be applied to generate an interpolated exposure surface for an entire study area.

LUR has been used extensively to study of traffic-related pollutants including unspicated particulate matter and NO_2 (Hoek et al., 2008), but relatively few studies have applied such methods to particulate matter constituents. In an early study, de Hoogh et al. (2013) analyzed spatial data from 400 sampling locations in 20 study areas across Europe and constructed LUR models for each site and each metal. They found that traffic-related variables, including traffic intensity and road length were often associated with metals identified as brake and tire tracers, such as Cu and Zn (de Hoogh et al., 2013). Zhang et al. (2015) collected PM_{10} data at 54 sampling locations (25 in the summer, 29 in winter) in Calgary, Alberta and reported that auto- and traffic-related variables contributed strongly to the predictions of monitored Ba, Zn, Cu, and Sb (Zhang et al., 2015). Ito et al. (2016) sampled at 99 locations in New York City and used speciated particle data to produce LUR models, identifying Cu, Fe, and Ti as markers of near-road pollution (Ito et al., 2016). Weichenthal et al. (2018) used LUR to model Cu, Fe, and generated a reactive oxygen species (ROS) concentration in lung fluid based on these two metals using 67 sites at which they monitored in two seasons in Toronto and found strong associations between road and traffic-related variables such as distance to and length of highways, length of major roads, and traffic counts and all three of their dependent variables, while also finding strong correlations between Cu, Ba, and Fe. (Weichenthal et al., 2018).

Other studies expanded on the traditional LUR approach by altering either the statistical methodology or including additional predictors in their LUR model. Brokamp et al. (2017) collected $\text{PM}_{2.5}$ samples at 24 sites in Cincinnati, Ohio, and modeled metal concentrations using both linear land use regression models and machine learning-based random forest regression models. Their approach resulted in better fit compared to previous models. For metals associated with brake and tire wear, including Cu and Zn, they found that traffic or road variables such as length of road and truck traffic were significant predictors (Brokamp et al., 2017). Tripathy et al. (2019) continued with a linear regression approach but integrated a Gaussian atmospheric dispersion model into their predictions, also finding improved model fit compared to only using land-use variables (Tripathy et al., 2019).

A general barrier of measuring speciated $\text{PM}_{2.5}$ elements is the time and labor involved in the sample collection and analysis. Collection of air samples on filters followed by inductively coupled mass spectrometry

analysis is the current gold standard, offering precise information on multiple metal constituents and isotopes. Collecting many samples simultaneously can be very challenging due to equipment and labor demands. As a result, assimilating other forms of data to assist with interpolation has clear advantages.

A potential solution to this problem involves using low-cost air sensor networks, such as PurpleAir (PurpleAir, 2021). Low-cost air sensor networks have grown following improvements in wireless connectivity technology and electrical engineering and have received increased attention from government regulatory agencies, such as the U. S. Environmental Protection Agency and the California Air Quality Management District (Johnson et al., 2019; Snyder et al., 2013). The PurpleAir device consists of two Plantower PMS5003 laser particle (channels A and B), which use light scattering of a laser to count suspended particulates in the air at various sizes and provides the mass concentration of $\text{PM}_{1.0}$, $\text{PM}_{2.5}$, and PM_{10} (PurpleAir, 2021). Devices also include wireless internet connectivity, and with the consent of the user, uploads time series data to the PurpleAir website, which provides data free of charge. PurpleAir sensors are not reference grade, although they correlate well with federal reference monitors (Kosmopoulos et al., 2020). In addition, these low-cost sensors do not measure speciated $\text{PM}_{2.5}$ constituents. The network, however, is far denser in major metropolitan areas compared to regulatory monitors and monitors deployed in academic studies, and information is collected and stored in real-time.

In this study we leverage the PurpleAir network to enhance the prediction of metal species and indicators of oxidative stress. We implemented geostatistical modeling employing the cokriging method to predict a target, gold standard variable (i.e., filter-based particle samples) based on auxiliary measures from more widely distributed PurpleAir sensors. While widely used in mining and petroleum engineering, cokriging has previously seen limited, but effective use in air pollution studies (Li et al., 2012). With a focus on brake and tire wear derived particles, this study is the first to model speciated $\text{PM}_{2.5}$ integrating information from a low-cost air sensor network.

2. Methods

2.1. Site selection, $\text{PM}_{2.5}$ speciation data

Our related publication details the site selection and sampling methods of this study and thus are described only briefly here (Oroumiyeh et al., 2022). Our study targeted the greater Los Angeles area, a geographically diverse and heavily populated region of 11.4 million people. The region is characterized as a Mediterranean climate, and is mild to hot year-round, with mean temperatures ranging from 15–17 $^\circ\text{C}$ in the winter to 21–23 $^\circ\text{C}$ in the summer. Average rainfall varies seasonally, with average rainfall at between 29–92 mm in the winter and 0–2.3 mm in the summer. We conducted two sampling campaigns, one in summer 2019 and the other in winter 2020. In each campaign, monitors simultaneously collected two-week samples at 46 different locations in the Greater Los Angeles region, with four sites serving as repeat locations. Our site selection methods, detailed in a prior report, involved a Multi-Criteria Decision Analysis, combining variables hypothesized a priori to be associated with brake and tire wear such as traffic intensity, road slope variance, and intersection density, as well as expert understanding of local geographies and PurpleAir densities to improve cokriging (Jerrett, 2021). The study analyzes data on ambient fine and coarse particulate matter (PM) samples using Harvard cascade impactors (CI) and personal environmental monitors (PEM).

In addition to chemical analyses described by Oroumiyeh et al. (2022), PEM samples were analyzed for black carbon and oxidative stress activity at UCLA with absorption at 880 and 370 nm, using an optical transmissometer (Magee Scientific). Oxidative stress potential was measured with two assays: OH radical formation and dithiothreitol (DTT) loss. We measured OH formation by incubating samples in

Table 1
List of independent variables in LUR model.

Variable Category	Covariates	Data Sources
Land Use	Tree canopy (%)	2016 National Land Use Consortium Database (MRLC 2016)
	Impervious (%)	
	Green space (%)	
	Non-green space (%)	
Traffic and/or Road	Intersection density (#/m ²)	US Census Bureau (Bureau 2019)
	Slope variance (%)	
	Road length (m)	
	Annual average daily traffic (Veh/day)	Federal Highway Administration (Roff 2020)
	Annual average daily heavy-duty traffic (Veh/day)	
Distance	Distance to rail (m)	US Census Bureau
	Distance to coast (m)	
	Distance to major road (m)	
Commercial	Brake and tire-related businesses (count)	ESRI Business Analyst (ESRI, 2018)

surrogate lung fluid containing an OH probe disodium terephthalate, and monitoring formation the reaction product of OH and terephthalate, which is highly fluorescent (Gonzalez et al., 2017). DTT loss was measured by incubating filters in phosphate buffer containing DTT, as detailed in elsewhere (Cho et al., 2005).

Prior to deployment, a preliminary list of metals hypothesized to serve as specific tracers for different sources (brake wear, tire wear, tailpipe emissions, and dust and soil) based on a prior literature search. To finalize a list of potential dependent outcomes, we calculated Spearman correlations between hypothesized tracers and selected Ba (brake wear), Zn (tire wear), DTT loss, OH formation, and black carbon, and a calculated ROS formation based on the kinetic multilayer model for surface and bulk chemistry in the epithelial lining fluid (KM-SUB-ELF) (Lakey et al., 2016), as well as unspiciated PM_{2.5}. KM-SUB-ELF uses concentrations of Cu and Fe, metals both associated with brake and tire wear, to calculate the rate of production of ROS in the epithelial lining fluid. Except for gravimetric PM_{2.5} and KM-SUB-ELF, we generated predictive surfaces for all metals and oxidative stress markers based on both volume-normalized and mass-normalized concentrations.

2.2. Land-use data

We downloaded land-use data from the following sources: (i) raster data on impervious surfaces, tree canopy, and green space data from the 2016 National Land Cover Database (NLCD) (MRLC, 2019); (ii) 2018 annual average daily traffic data from the federal highway administration (Roff, 2020); (iii) road network and rail network information from the U.S. Census Bureau (Bureau, 2019); and (iv) businesses likely to emit metals, such as auto repair shops and metal processing units, from the ESRI business analyst (ESRI, 2018).

NLCD data came in three raster files, each at a resolution of 30 by 30 meters: impervious percent, tree canopy percent, and a dataset which classified each pixel into type of land use (MRLC, 2019). Following a previous study, we created green space by grouping the following variables together: open water, ice/snow, developed/open space, developed/low intensity, deciduous forest, evergreen forest, mixed forest, dwarf scrub, shrub/scrub, grassland/herbaceous, sedge/herbaceous, pasture/hay, and cultivated crops (Akpınar et al., 2016). We classified the following as non-green space: medium intensity land cover, high intensity land cover, and barren land (rock/sand/clay) (Table S1).

Information on all businesses in the study area were downloaded from the ESRI Business Analyst. We selected businesses that potentially emit heavy metals and those related to brake and tire wear such as auto

shops and brake and tire manufacturers based on North American Industry Classification System code. Subsections for codes selected are available in Table S2.

2.3. Buffering and zonal statistics

We assigned land-use data to our sampling sites using nested circular buffers (Henderson et al., 2007; Klot, 2011). Around each monitoring site, we constructed circle-shaped areal units of 100-, 200-, 300-, 400-, 500-, 750-, and 1000-meter radii. For each circle, we calculate either the average value for land use for NLCD and traffic data or the density for intersection and businesses data. The complete list of potential variables and their sources are found in Table 1.

2.4. Land-use regression model building and validation

In line with other studies, we approach our data assuming a linear relationship between independent and dependent variables (Zhang et al., 2015). We began by plotting histograms and univariate plots to investigate the distribution of dependent and independent variables. If a histogram revealed the distribution of a variable to be skewed, or if univariate plots showed a non-linear relationship between independent and dependent variable, we log transformed our variables. To reduce collinearity between variables of the same class but different buffer radii, for each class of variable, we selected the buffer distance with the highest univariate correlation with the outcome of interest. After generating variables for each point, we used a method previously used to predict PM_{2.5} concentrations in the United States, selecting variables using a deletion-substitution-addition (DSA) algorithm (Beckerman et al., 2013). The DSA algorithm selected variables based on prediction accuracy with a L2 loss cross-validation function (Sinisi and van der Laan, 2004). After the DSA algorithm chose the optimal set of predictors of sizes 1–10 variables, we conducted another round of K-fold cross-validation with five folds to select an optimal number of variables in the final model based on the mean square predictor error.

To adjust for seasonality, we included a dummy variable indicating the sampling season (0 = summer, 1 = winter). In our final prediction surfaces, we calculated the average of the summer and winter surfaces by dividing the season coefficient by two and adding the resultant term to the remainder of the model.

2.5. PurpleAir data

Within the study area, we downloaded hourly-averaged PurpleAir data temporally aligned with both study periods from the PurpleAir website. PurpleAir sensors contain two sensors, which separately report PM 2.5 measurements through two channels, A and B. We restricted our sample to outdoor sensors. Downloaded data contained date and time of measurement, estimated particle matter mass concentration at 1, 2.5, and 10 microns reported by the two sensors in each device labeled channels A and B, temperature in °F, and relative humidity. Initial data quality assurance and quality control (QA/QC) steps included excluding sensors if they:

- Had a missing data rate of 10% or higher
- Did not report temperature

Next, we verified that our sensors were outdoor by creating time series temperature plots. If sensors reported a very low range of temperature, we assumed that they were mislabeled indoor sensors and subsequently deleted the sensor.

After obtaining a final list of PurpleAir sensors, we conducted further QA/QC per Plantower's factory standards (AQMD 2015). Data points were removed if they did not meet the following criteria:

- PM_{2.5} mass concentrations above 500 µg/m³

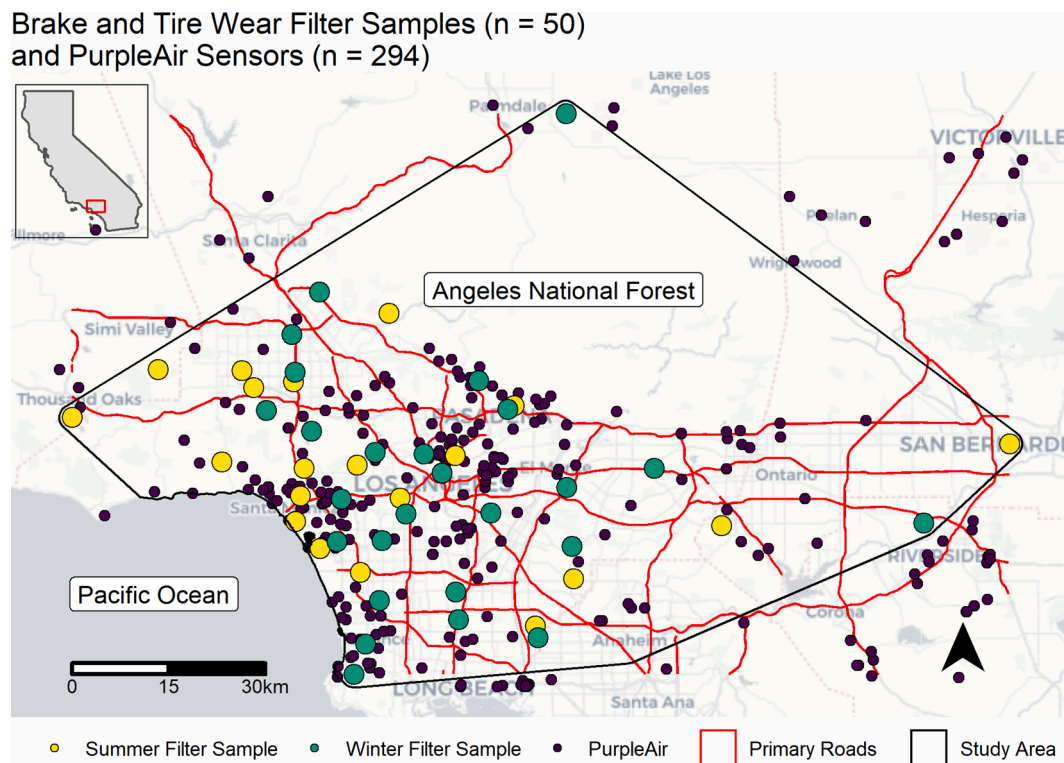


Fig. 1. Locations of sampling locations (green and yellow) and PurpleAir sensors (purple). The study location within the state of California is illustrated on the top-left, with the study location outlined in red. The area to be interpolated, defined by the convex hull of the pump boxes is outlined in black. Primary roads in the study area's bounding box are shown in red.

- For observations below $100 \mu\text{g}/\text{m}^3$, remove rows if the difference between $\text{PM}_{2.5}$ measurements reported in channels A and B exceeds $10 \mu\text{g}/\text{m}^3$
- For observations greater than $100 \mu\text{g}/\text{m}^3$, remove rows if the difference between $\text{PM}_{2.5}$ measurements reported in channels A and B exceeds 10%

Finally, we also removed a limited number of rows with extreme temperature and relative humidity values ($\text{RH}\% \leq 0$ or $\text{RH}\% \geq 100$; temperature $\leq -200^\circ\text{F}$ (-129°C) or temperature $\geq 1000^\circ\text{F}$ (537°C)). Once cleaned, for each sensor, we obtained a final measurement by averaging measurements from the A and B sensors.

We selected PurpleAir estimated $\text{PM}_{2.5}$ concentrations as our auxiliary variable in our cokriging model, as it exhibited the highest univariate correlations with our speciated $\text{PM}_{2.5}$ data. Locations of both filter samplers and PurpleAir sensors are plotted on Fig. 1. Our final dataset used data from 50 filter samplers across two seasons and 294 PurpleAir sensors.

2.6. Co-kriging with external drift

Co-kriging (Journal and Huijbregts, 1978; Matheron, 1979) refers to predicting the value of an outcome of interest at a spatial location exploiting information from (i) direct measurements of the primary variable; and (ii) measurements of other variables, sometimes referred to as auxiliary variables or covariables, with which the outcome is assumed to be statistically dependent. The key advantage of cokriging over ordinary kriging is that we incorporate the dependence between the primary variable and the auxiliary variable as well as the spatial dependence exhibited by all the variables. The external drift refers to a set of linear predictors, in this case our LUR models. By incorporating spatial dependencies between our measurements and PurpleAir estimates, we can model using spatial relationships unaccounted for with standard LUR models.

Co-kriging with external drift (CED) uses a multivariate spatial regression model

$$y_i(s) = x_i^T(s)\beta + w_i(s) + \epsilon_i(s), \quad i = 1, 2, \dots, q, \quad (1)$$

where $y_i(s)$ is the i -th outcome at location s , $x_i(s)$ is the external drift, a $p \times 1$ vector of land-use covariates or predictors corresponding to the i -th outcome at location s , $w_i(s)$ is a zero-centered spatial process capturing spatial dependence and $\epsilon_i(s)$ is a white-noise process independent of $w_i(s)$ capturing micro-scale variation arising from measurement errors or other fine resolution dependencies. Further details regarding cokriging is available in the supplemental information.

To avoid singularity issues, for co-located points in both speciated metals and PurpleAir data, we shifted longitude and latitude coordinates by a random distance between 0 and 5 meters. For each exposure (outcome variable), we estimated $w_i(s)$, the spatial process, as a covariogram modeling the cross-covariance between the exposure and auxiliary variable as a function of distance. After fitting initial theoretical covariograms, we evaluated the performance by simultaneously conducting 1000 iterations of 10-fold cross-validation for both the original LUR model as well as the CED model. We then used two-sample t-tests to compare the mean and standard deviation of the 1000 subsequent mean square errors generated and made manual adjustments to the cross co-variogram parameters. We repeated this process with the goal of minimizing both the mean and standard deviation of the MSE as much as possible.

We implemented co-kriging using the *gstat* library (Pebesma and Wesseling, 1998) in the R statistical computing environment to implement co-kriging. Locations of the filter samplers and PurpleAir monitors used in the study are mapped in Fig. 1.

2.7. Generating exposure surfaces

After completing LUR model building, we created raster surfaces for

Table 2
Summary statistics for dependent variables.

Outcome	Volume-Normalized Mean (SD)	Mass-Normalized Mean (SD)
PM _{2.5}	9.4 (2.4) $\mu\text{g}/\text{m}^3$	
Ba	17 (9.9) ng/m^3	1700 (750) ng/g
Zn	9.3 (4.6) ng/m^3	930 (320) ng/g
KM-SUB-ELF ROS	100 (26) nmol/L	
OH Formation	4.95 (2.1) $\text{pmol}/\text{min}/\text{m}^3$	0.56 (0.14) $\text{pmol}/\text{min}/\mu\text{g}$
DTT Loss	550 (208) $\text{pmol}/\text{min}/\text{m}^3$	61.9 (11.8) $\text{pmol}/\text{min}/\mu\text{g}$
Black Carbon	0.41 (0.17) $\mu\text{g}/\text{m}^3$	4.6 (1.2) %

predictor variables using the focal statistics function in ArcMap 10.8 (ESRI, Redlands, CA). In our convex hull, we used R 3.6.3 (R Foundation, Vienna, Austria) to generate a 30-meter by 30-meter grid of our study area, extract independent variable values at each point in the grid, and predict with both the LUR and CED models. We avoided predicting exposures in the Angeles National Forest, a national forest with a very few permanent inhabitants not represented in our study sample.

3. Results and Discussion

3.1. Summary statistics

We built land use regression and cokriging models for six different components/measures in the PM_{2.5} size fraction: Ba, Zn, OH formation, DTT loss, black carbon, and KM-SUB-ELF ROS. Table 2 summarizes both the volume-normalized concentration and mass-normalized (or equivalent) measurement for each model outcome, when available.

Our summary statistics mostly agree with findings from previous studies. Our mean gravimetric PM_{2.5} concentration of $9.4 \mu\text{g}/\text{m}^3$ was slightly lower, but overall comparable to the 2019 Los Angeles County average of $12.7 \mu\text{g}/\text{m}^3$, likely because our 2-week sampling periods in two seasons only covered four of the 52 weeks. Zn measurements are similar to the lower end of measurements reported by de Hoogh et al. (2013); Brokamp et al. (2017), and Kuang et al. (2020) but are much

lower than those reported by Tripathy et al. (2019), despite all of these studies having taken place in either European or American metropolitan areas, including some in Los Angeles. Zn is a product of some industrial processes such as steel production, as well as being a component of tires, which may cause the differences in measurements, both spatially and temporally. A reason might be that Zn emissions, while a product with car tire wear, is a multisource pollutant also associated with the lubrication oil and steel production (Brokamp et al., 2017; Kuang et al., 2020; Tripathy et al., 2019).

For black carbon, our average concentration of $0.41 \mu\text{g}/\text{m}^3$ was lower compared to a different study that sampled in the same region, who reported an average concentration of $0.59 \mu\text{g}/\text{m}^3$ (Jones et al., 2020). The discrepancy in black carbon levels is likely due differences in sampling strategy; Jones et al. (2020) focused on clustering monitoring sites around major freeways in the Los Angeles area, whereas we sampled in both high- and low-traffic areas. Additionally, we collected a single 14-day sample which included sampling during the day and night, with the latter having reduced levels of traffic. Jones et al. (2020) instead measured for shorter periods during the daytime, albeit outside of rush hour traffic. Regarding DTT loss, our average loss rate of $61.9 \text{ pmol}/\text{min}/\mu\text{g}$ is comparable, but on the higher end of findings published by Charrier et al. (2012), who analyzed PM_{2.5} samples from the San Joaquin Valley in California, US and found DTT loss rates between 17 and $70 \text{ pmol}/\text{min}/\mu\text{g}$ (Charrier and Anastasio, 2012). DTT was also studied by Yang et al. (2015). Compared to our mean volume-normalized mean DTT loss rate of $550 \text{ pmol}/\text{min}/\text{m}^3$, they measured generally higher rates of DTT loss, ranging from 700 to $2000 \text{ pmol}/\text{min}/\mu\text{g}$, (Yang et al., 2015). Ba is less commonly reported, however, while PM_{2.5}, Zn and many other elements measured in our study are much lower than measured in a multi-site air quality campaign in Los Angeles in 1987, Ba concentrations (in ng/m^3) fall in the same range ($17 \text{ ng}/\text{m}^3$ in our study vs. $13\text{--}32 \text{ ng}/\text{m}^3$) as the measurements made in 1987. Two earlier studies have reported OH formation in simulated lung fluid from ambient particles, for particles collected at one site during summer in Los Angeles and the other in Beijing and Wangdu in Northern China

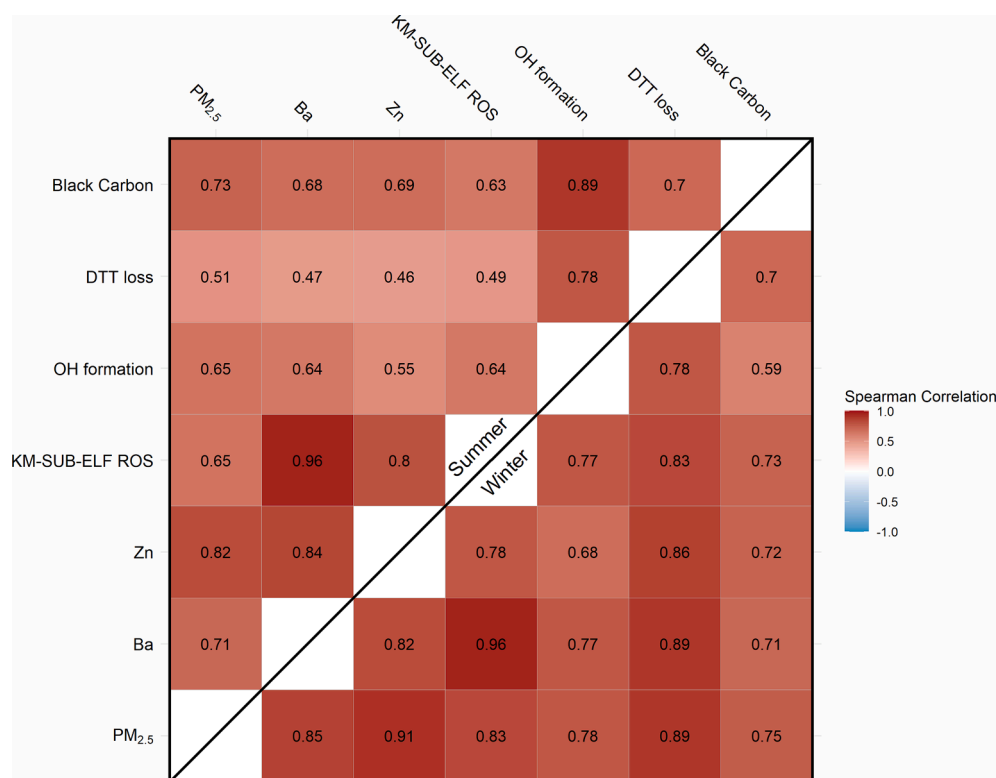


Fig. 2. Spearman correlation matrix between volume-normalized study outcomes during the summer (top/left) and winter (bottom/right) sampling campaigns.

Table 3

Model summaries for volume-normalized (left) and mass-normalized measurements (right). Numbers contained in parentheses next to model parameters indicates the buffer radius in the LUR model, if applicable. * indicates that the dependent variable was log-transformed.

Pollutant	Volume-normalized measurement			Mass-normalized measurement		
	Adj. R ²	Model Parameters	Coef.	Adj. R ²	Model Parameters	Coef.
PM _{2.5}	0.41	Log annual average daily traffic (500)	0.32	-	-	-
		Log distance to rail	−0.75			
		Log distance to major road	−0.38			
		Impervious % (1000)	0.05			
		Tree canopy % (400)	0.08			
Ba	0.60	Period	9.146	0.56	Period	725.79
		Tree canopy % (400)	0.763		Tree canopy % (750)	47.42
		Log distance to rail	−3.40		Log distance to rail	−162.9
		Impervious % (200)	0.259		Impervious % (200)	10.55
		Business count (750)	0.313		Business count (750)	25.9
Zn	0.47			0.42*	Log annual average daily heavy-duty traffic (400)	59.8
		Period	3.69		Period	6.95
		Log distance to rail	1.40		Tree canopy % (750)	0.026
		Log annual average daily traffic (300)	0.64		Log distance to rail	−0.12
					Impervious % (100)	0.0056
KM-SUB-ELF ROS	0.58	Period	81.77	-	-	-
		Tree canopy % (750)	1.57			
		Log distance to rail	−5.49			
		Impervious % (200)	0.55			
		Business count (750)	0.87			
OH Formation	0.56	Log annual average daily heavy-duty traffic (400)	2.42	0.37		
		Period	1.88		Period	0.38
		Business count (750)	0.07		Business count (750)	0.01
		Log annual average daily heavy-duty traffic (400)	0.28		Log annual average daily heavy-duty traffic (400)	0.02
		Tree canopy % (1000)	0.07			
DTT Loss	0.60	Impervious % (200)	0.04	0.68		
		Period	215.1		Period	17.49
		Business count (300)	34.1		Business count (300)	2.02
		Log annual average daily heavy-duty traffic (400)	22.1		Log annual average daily heavy-duty traffic (400)	1.04
		Tree canopy % (1000)	10.16		Tree canopy % (1000)	0.34
Black Carbon	0.64	Impervious % (1000)	4.26	0.56		
		Period	0.15		Period	1.25
		Impervious % (1000)	0.01		Impervious % (1000)	0.02
		Tree canopy % (400)	0.01		Tree canopy % (400)	0.06
		Log distance to rail	−0.06		Log distance to rail	−0.46

respectively (Kuang et al., 2017; Li et al., 2019). Our mass-normalized average value of 0.55 pM/min/ μ g is reasonably close to the earlier Los Angeles value (0.32 pM/min/ μ g, Kuang et al. 2017) and falls in the range of values observed in China (0.22 – 1.13 pM/min/ μ g). Finally, we found higher levels of ROS generation estimated by KM-SUB-ELF (100 nmol/L), compared to Weichenthal et al. (2018), who sampled during the summer and winter seasons in Toronto, Canada, and reported an annual average ROS concentration of 52 nmol/L, possibly due to generally lower particulate concentrations in Toronto (Weichenthal et al., 2018).

We calculated univariate Spearman correlation coefficients between our six outcome variables, summarized in Fig. 2. Observing changes in correlation coefficients between summer and winter we have divided our correlation matrix into summer (top/left) and winter (bottom/right) and added a dummy variable indicating sampling period (0 for summer, 1 for winter) to our land use regression. We find Spearman correlations are very high (ρ : 0.8–0.98) for metal concentrations or measurements based on metal concentrations (Ba, Zn, and ROS). Such high correlations between brake and tire-wear associated metal-based variables are consistent with a study of speciated PM₁ elsewhere (Zhang et al., 2015). For other measurements, correlations are positive, but differ between summer and winter. DTT loss is most highly correlated with other measures in the winter sample (ρ : 0.51–0.73) and only moderately in the summer (ρ : 0.24–0.47), which are consistent with a prior study finding Spearman correlations ranging from 0.24 to 0.52 (Yang et al., 2015). In contrast, for OH formation correlations with other measures were slightly higher during the summer (ρ : 0.53–0.71) compared to the winter (0.46–0.60).

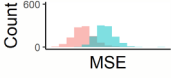
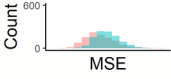
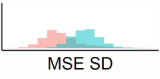
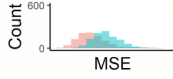
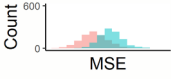
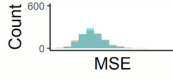
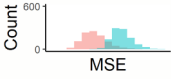
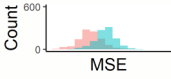
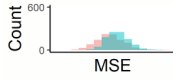
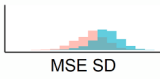
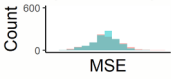
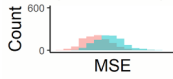
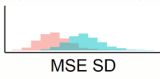
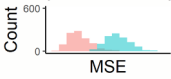
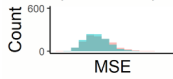
3.2. Variable selection and LUR model results

Figure S1 shows the cross-validation errors for models of each model size chosen by the DSA search algorithm resulting from a 5-fold out-of-sample cross validation process. For each exposure, we chose the model with the minimum error, regardless of model size. Table 3 summarizes the results of variable selection while building of 12 different land use regression models for seven exposures, including predictors selected, regression coefficients, and adjusted R². The most common predictor, which appears in all LUR models except for PM_{2.5}, is the dummy variable for monitoring period, indicating seasonal difference among most exposures. All models featured at least one traffic-related variable, either distance to railway, heavy duty traffic, or both. Other common predictors included the land use variables tree canopy cover percent, impervious surface percent, and number of businesses within the buffer distance. The best performing models were for volumized concentrations of Ba, black carbon, and DTT loss, with adjusted R² statistics of 0.60, 0.64, and 0.68, respectively. Despite mainly relying on cross-validation error, our model R² values are similar to those obtained in prior literature (Yang et al., 2015; Zhang et al., 2015). Our CV method has previously been shown to be optimized for predictive accuracy, which allows us to minimize the risk of overfitting in our modeling common among studies that choose models based on an R² statistic (Davies and van der Laan, 2016).

Across our 12 LUR models, the direction of effect associated with each variable remained consistent, which we interpret as a sign of internal consistency in our results. For example, a positive period variable indicates that exposure concentrations were greater in the winter compared to the summer. Negative distance to rail coefficients indicates that higher levels of exposure closer to railways. Business count co-

Table 4

Percent changes in mean and standard deviation of MSE after implementing CED compared to LUR for volume-normalized (left) and mass-normalized (right) measurements. Under each estimate are histograms illustrating the 10-fold cross-validation performance differences between LUR (blue) and CED (pink) modelling approaches after 1000 iterations. A darker area indicates overlap between LUR and CED results.

Pollutant	Volume-normalized measurement		Mass-normalized measurement	
	Mean MSE Change (95% CI)	Mean MSE SD Change (95% CI)	Mean MSE Change (95% CI)	Mean MSE SD Change (95% CI)
PM _{2.5}	-11.7% (-12.2%, -11.2%) 	-21.8% (-22.4%, -21.3%) 		
Ba	-3.1% (-3.6%, -2.6%) 	-7.9% (-8.4%, -7.4%) 	-5.7% (-6.2%, -5.2%) 	-8.8% (-9.4%, -8.2%) 
Zn	-6.9% (-7.3%, -6.4%) 	-4.9% (-5.5%, -4.3%) 	-0.1% (-0.6%, 0.4%) 	-3.5% (-4.0%, -3.0%) 
KM-SUB-ELF ROS	-14.1% (-14.6%, -13.6%) 	-14.3% (-14.8%, -13.7%) 		
OH Formation	-6.3% (-6.7%, -5.9%) 	-15.8% (-16.2%, -15.3%) 	-2.1% (-2.4%, -1.8%) 	-3.1% (-3.4%, -2.9%) 
DTT Loss	0.3% (-0.2%, 0.8%) 	0.2% (-0.4%, 0.8%) 	-6.1% (-6.6%, -5.5%) 	-13.8% (-14.5%, -13.2%) 
Black Carbon	-12.6% (-12.9%, -12.2%) 	-36.0% (-36.3%, -35.7%) 	0.7% (0.3%, 1.0%) 	-8.1% (-8.4%, -7.9%) 

efficients, regardless of buffer distance, were uniformly positive, matching a prior study finding that auto-repair and related businesses are positively correlated with brake and tire wear metals or that they may also be destinations correlated with heavier traffic in the surrounding area (Hasheminassab et al., 2020). Similarly, the coefficient for heavy duty traffic, a variable very highly correlated with passenger traffic, was uniformly positive across our 12 models. Unexpectedly, tree canopy percentage coefficients were not only frequently selected as predictors into models, but were also uniformly positive. Although tree canopy cover may serve as a marker for green space, our study almost exclusively monitored in urban or residential areas, where it is common to plant trees alongside sidewalks near roadways. In our data, we found that log distance to road and tree canopy cover were positively correlated (Pearson $\rho = 0.34$ – 0.51), which may suggest that in a large, highly developed urban setting, tree canopy serves as a proxy not for density of green space, but instead proximity to the roads.

After completing LUR modeling, we assessed spatial autocorrelation by examining spatial residual maps, constructing and visually inspecting Thiessen polygons, calculating the Moran's I statistic, and modeling the semivariogram of the residuals. All four methods did not show evidence of significant spatial autocorrelation among LUR residuals.

3.3. Changes in cross-validation accuracy between LUR and CED

Cross covariograms between LUR residuals and PurpleAir PM_{2.5}

showed a degree of cross-covariance (Figures S5 and S6). We find this reflected in the results of iterated 10-fold cross validation, where after optimizing the cross covariogram parameters, a statistically significant improvement in prediction accuracy among the majority of our exposures (Table 4).

After conducting 1000 iterations of 10-fold cross validation comparing CED and LUR, we found statistically significant reductions in the mean and standard deviation of the MSE for 8 out of the 12 the study's exposures, as shown on Table 4. There were statistically significant changes in neither mean or standard deviation of the MSE for mass-normalized Zn concentrations and volumized DTT loss. While black carbon percentage had a small but statistically significant increase in mean MSE, standard deviation of MSE decreased, suggesting increased precision but slightly reduced accuracy. For other outcomes, we found mean reductions in MSE ranging between 2.1% and 14.1%, and reductions in MSE standard deviation between 3.1% and 36%. These results suggest that for most brake and tire wear-related outcomes, incorporating low-cost PM_{2.5} measurements via co-kriging leads to improved predictive accuracy and precision. While co-kriging methods have been used in the past to improve PM₁₀ based on incorporating wind speed curves, to our knowledge our study is the first to show that compared to conventional LUR methods, co-kriging with a low-cost air sensor network improves predictions of PM_{2.5} species and markers of oxidative stress.

Modeled surfaces for selected dependent variables comparing CED

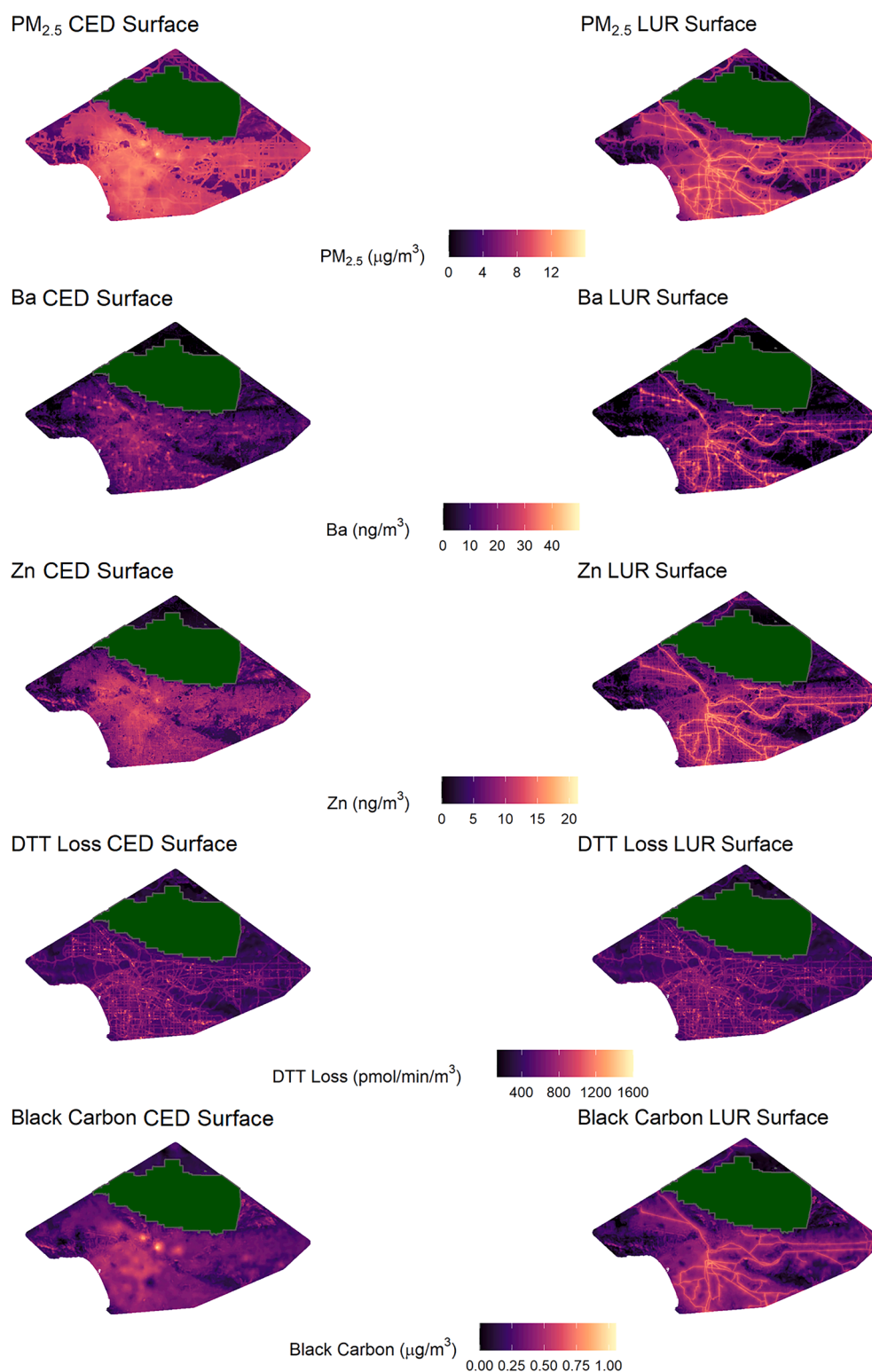


Fig. 3. Exposure surfaces generated by CED (left) and LUR (models) for (from top to bottom) volumized estimates of $PM_{2.5}$, Ba concentration, Zn concentration, and loss rates, and black carbon, which represent brake wear, tire wear, oxidative stress potential, and combustion, respectively. The Angeles National Forest, where we do not interpolate, is shown in green.

and LUR are illustrated in Fig. 3. In addition to generally lower MSE and MSE standard deviation, visual inspection reveals that CED appears to act as a smoother for the LUR model. This is particularly noticeable when mapping out the volumized exposure surfaces for of Zn, $PM_{2.5}$ and black carbon. LUR models for those pollutants shows hotspots along rail lines and major roadways, which result from the distance to rail and

traffic-related predictors. Potential explanations include the choice of modeling method and auxiliary variable. The auxiliary variable in the co-kriging model is $PM_{2.5}$, which as a multisource mixture, generally displays smoother spatial patterns than primary particles from brake and tire wear. Its inclusion as a variable in our prediction model may consequently over-smooth our exposure surfaces. Thus, co-kriging may

lead to a sacrifice in small-area signal, in exchange for more accurate and precise predictions of overall concentrations. Notably, greater degrees of smoothing occurred among exposures that exhibited higher increases of predictive accuracy and precision, such as $PM_{2.5}$, Zn, and black carbon, compared to DTT loss, which did not experience a significant change in predictive accuracy.

Environmental features may also be responsible for differences in spatial distribution that may be captured by PurpleAir data but not in standalone LUR models. Although we did not observe systematic patterns in LUR spatial residuals, factors such as prevailing wind direction may have affected spatial patterns, leading to smoothing. Other hyper-local environmental features such as the presence of street canyons may have affected local wind direction, leading to smoothing effects beyond wind direction alone. Traffic activity may also lead to continued resuspension of settled particles, which when entering the atmosphere, leading to more smoothing. Further study of these effects aimed at better understanding the factors behind smoothing include potential simulation studies, repeated campaigns with increased sample size, or alternative choices of auxiliary variables.

Our study has a few limitations. Although the number of sensors deployed matches that of some prior studies, such as Zhang et al. (2015), our study area is much larger than that of prior studies, and the network of monitoring sites is not very dense for the area covered, which may have led to over-smoothing of the prediction surfaces due to sparse data support. We decided to prioritize covering a larger study area over having many repeated monitoring sites, which resulted in very low temporal resolution, with repeated measurements at only four sites. Adding in the PurpleAir network improves cross-validation for most outcomes and improves the spatial resolution of our data, but there may be concerns about this network and its data. While low-cost, the PurpleAir sensor remains a relative luxury and network density is consequently biased towards high-income households (Liang et al., 2021). As a low-cost sensor maintained by consumers, despite attempts at laboratory and field calibration, PurpleAir sensors when inadequately maintained are at risk of unaccountable sensor drift and distortion of results due to environmental factors (AQMD, 2015; Carvlin, 2019). Our results show significant improvements in predictive accuracy and precision, though future similar studies using well-maintained sensor networks may see even greater improvements in predictive accuracy. With regards to speciated data, most of our filter samplers were placed at residences of participants of an existing birth cohort study, who were largely high SES homeowners. As a result, our model predictions in and around disadvantaged communities, which historically have higher levels of air pollution and lower levels of homeownership, likely have higher levels of uncertainty. Additionally, based on the observed improvements in our predictive models when adding in PurpleAir data, our study suggests that adding in temporally aligned data may help address a fundamental issue among LUR studies.

4. Conclusion

Here we presented a novel method for integrating low-cost sensor data with reference grade data to estimate surfaces of brake and tire wear across Southern California. We constructed several exposure surfaces over a large area for use in population health studies with information from filter samples, land-use data, and the PurpleAir low-cost sensor network. Our study demonstrates that as sensor networks continue to expand, the ease of data collection and high spatial resolution associated with such networks may have the potential to improve exposure modeling and subsequent air pollution epidemiology. Compared a conventional LUR, CED addresses spatial relationships between filter samples and PurpleAir sensors across the study area and may account for spatial distribution of brake and tire wear-related particles more accurately than an LUR. Our results reflect normal conditions because the monitoring did not take place during major recent events affecting air quality, such as the California wildfires and the

COVID-19 pandemic.

Despite limitations noted above, our model has identified various land use variables, such as traffic, business density, and impervious surfaces as important predictors of brake and tire wear metals and oxidative potential measures in a large study area. We report the first study to incorporate low-cost sensor data to model speciated $PM_{2.5}$, as well as the largest-area study covering the Los Angeles metropolitan area. Our study results suggest that such sensor networks, particularly when further expanding, have potential for use in exposure modeling research. When dense enough, such networks have the potential to complement existing research-grade monitoring, particularly when studying pollutants or their markers that are difficult or expensive to measure, such as the constituents of $PM_{2.5}$ or oxidative stress markers. Future research directions in the exposure field may include taking advantage of the real-time nature of the PurpleAir network to scale speciated $PM_{2.5}$ exposures at different time periods. Beyond exposure modeling, we plan in the future to apply exposure surfaces to health studies, including an existing pregnancy cohort and a study of adverse birth outcomes based on state records.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the California Air Resources Board [contract number 17RD012]. M.S. acknowledges funding from the Health Effects Institute (HEI) (Walter A. Rosenblith New Investigator Award, No. 4964-RFA17-3/18–6).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.envint.2022.107481>.

References

- Akpınar, A., Barbosa-Leiker, C., Brooks, K.R., 2016. Does green space matter? Exploring relationships between green space type and health indicators. *Urban Forest. Urban Green.* 20, 407–418. <https://doi.org/10.1016/j.ufug.2016.10.013>.
- Amato, F., Pandolfi, M., Moreno, T., Furger, M., Pey, J., Alastuey, A., Bukowiecki, N., Prevot, A.S.H., Baltensperger, U., Querol, X., 2011. Sources and variability of inhalable road dust particles in three European cities. *Atmos. Environ.* 45, 6777–6787. <https://doi.org/10.1016/j.atmosenv.2011.06.003>.
- AQMD, S., 2015. PurpleAir PA-II. South Coast Air Quality Management District.
- Beckerman, B.S., Jerrett, M., Serre, M., Martin, R.V., Lee, S.-J., van Donkelaar, A., Ross, Z., Su, J., Burnett, R.T., 2013. A Hybrid Approach to Estimating National Scale Spatiotemporal Variability of $PM_{2.5}$ in the Contiguous United States. *Environ. Sci. Technol.* 47, 7233–7241. <https://doi.org/10.1021/es400039u>.
- Briggs, D., 2005. The Role of GIS: Coping With Space (And Time) in Air Pollution Exposure Assessment. *J. Toxicol. Environ. Health, Part A* 68, 1243–1261. <https://doi.org/10.1080/15287390590936094>.
- Brokamp, C., Jandarov, R., Rao, M.B., LeMasters, G., Ryan, P., 2017. Exposure assessment models for elemental components of particulate matter in an urban environment: A comparison of regression and random forest approaches. *Atmospheric Environment (Oxford, England)* 151, 1–11. doi:10.1016/j.atmosenv.2016.11.066.
- Bureau, U.C., 2019. TIGER/Line Shapefiles. The United States Census Bureau.
- CARB, 2020. 2020 Mobile Source Strategy: A Vision for Clean Air.
- Carvlin, G., 2019. Purple Air Calibration: Effect of space, time and environment.
- Charrier, J.G., Anastasio, C., 2012. On dithiothreitol (DTT) as a measure of oxidative potential for ambient particles: Evidence for the importance of soluble transition metals. *Atmos. Chem. Phys. (Print)* 12, 11317–11350. <https://doi.org/10.5194/acpd-12-11317-2012>.
- Cho, A.K., Sioutas, C., Miguel, A.H., Kumagai, Y., Schmitz, D.A., Singh, M., Eiguen-Fernandez, A., Froines, J.R., 2005. Redox activity of airborne particulate matter at different sites in the Los Angeles Basin. *Environ. Res.* 99, 40–47. <https://doi.org/10.1016/j.envres.2005.01.003>.
- Davies, M.M., van der Laan, M.J., 2016. Optimal Spatial Prediction Using Ensemble Machine Learning. *Int. J. Biostat.* 12, 179–201. <https://doi.org/10.1515/ijb-2014-0060>.

- de Hoogh, K., Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Declercq, C., Dedelè, A., Dons, E., de Nazelle, A., Eeftens, M., Eriksen, K., Eriksson, C., Fischer, P., Gražulevičienė, R., Gryparis, A., Hoffmann, B., Jerrett, M., Katsoyanni, K., Iakovides, M., Lanki, T., Lindley, S., Madsen, C., Mölter, A., Mosler, G., Nádor, G., Nieuwenhuijsen, M., Pershagen, G., Peters, A., Phuleria, H., Probst-Hensch, N., Raaschou-Nielsen, O., Quass, U., Ranzi, A., Stephanou, E., Sugiri, D., Schwarze, P., Tsai, M.-Y., Yli-Tuomi, T., Varró, M.J., Vienneau, D., Weinmayr, G., Brunekreef, B., Hoek, G., 2013. Development of Land Use Regression Models for Particle Composition in Twenty Study Areas in Europe. *Environ. Sci. Technol.* 47, 5778–5786. <https://doi.org/10.1021/es400156t>.
- de Jesus, A.L., Rahman, M.M., Mazaheri, M., Thompson, H., Knibbs, L.D., Jeong, C., Evans, G., Nei, W., Ding, A., Qiao, L., Li, L., Portin, H., Niemi, J.V., Timonen, H., Luoma, K., Petäjä, T., Kulmala, M., Kowalski, M., Peters, A., Cyrys, J., Ferrero, L., Manigrasso, M., Avino, P., Buonano, G., Reche, C., Querol, X., Beddows, D., Harrison, R.M., Sowlat, M.H., Sioutas, C., Morawska, L., 2019. Ultrafine particles and PM_{2.5} in the air of cities around the world: Are they representative of each other? *Environ. Int.* 129, 118–135. <https://doi.org/10.1016/j.envint.2019.05.021>.
- ESRI, 2018. Business Analyst Data.
- Fussell, J.C., Franklin, M., Green, D.C., Gustafsson, M., Harrison, R.M., Hicks, W., Kelly, F.J., Kishta, F., Miller, M.R., Mudway, I.S., Oroumijeh, F., Selley, L., Wang, M., Zhu, Y., 2022. A Review of Road Traffic-Derived Non-Exhaust Particles: Emissions, Physicochemical Characteristics, Health Risks, and Mitigation Measures. *Environ. Sci. Technol.* <https://doi.org/10.1021/acs.est.2c01072>.
- Gonzalez, D.H., Cala, C.K., Peng, Q., Paulson, S.E., 2017. HULIS Enhancement of Hydroxyl Radical Formation from Fe(II): Kinetics of Fulvic Acid-Fe(II) Complexes in the Presence of Lung Antioxidants. *Environmental Science & Technology* 51, 7676–7685. <https://doi.org/10.1021/acs.est.7b01299>.
- Grieshop, A.P., Lipsky, E.M., Pekney, N.J., Takahama, S., Robinson, A.L., 2006. Fine particle emission factors from vehicles in a highway tunnel: Effects of fleet composition and season. *Atmospheric Environment, Particulate Matter Supersites Program and Related Studies* 40, 287–298. <https://doi.org/10.1016/j.atmosenv.2006.03.064>.
- Grigoratos, T., Martini, G., 2015. Brake wear particle emissions: A review. *Environ. Sci. Pollut. Res.* 22, 2491–2504. <https://doi.org/10.1007/s11356-014-3696-8>.
- Habre, R., Girguis, M., Urman, R., Fruin, S., Lurmann, F., Shafer, M., Gorski, P., Franklin, M., McConnell, R., Avol, E., Gilliland, F., 2020. Contribution of Tailpipe and Non-tailpipe Traffic Sources to Quasi-Ultrafine, Fine and Coarse Particulate Matter in Southern California. *Journal of the Air & Waste Management Association* null. <https://doi.org/10.1080/10962247.2020.1826366>.
- Hasheminassab, S., Sowlat, M.H., Pakbin, P., Katzenstein, A., Low, J., Polidori, A., 2020. High time-resolution and time-integrated measurements of particulate metals and elements in an environmental justice community within the Los Angeles Basin: Spatio-temporal trends and source apportionment. *Atmospheric Environment: X* 7, 100089. <https://doi.org/10.1016/j.aeaoa.2020.100089>.
- Henderson, S.B., Beckerman, B., Jerrett, M., Brauer, M., 2007. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology* 41, 2422–2428. <https://doi.org/10.1021/es0606780>.
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* 42, 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>.
- Ito, K., Johnson, S., Kheirbek, L., Clougherty, J., Pesheski, G., Ross, Z., Eisl, H., Matte, T. D., 2016. Intraurban Variation of Fine Particle Elemental Concentrations in New York City. *Environmental Science & Technology* 50, 7517–7526. <https://doi.org/10.1021/acs.est.6b00599>.
- Jeong, C.-H., Traub, A., Huang, A., Hilker, N., Wang, J.M., Herod, D., Dabek-Zlotorzynska, E., Celò, V., Evans, G.J., 2020. Long-term analysis of PM_{2.5} from 2004 to 2017 in Toronto: Composition, sources, and oxidative potential. *Environ. Pollut.* 263, 114652. <https://doi.org/10.1016/j.envpol.2020.114652>.
- Jerrett, M., 2021. Air Monitoring Site Selection Methodology for Particulate Matter from Brake and Tire Wear. *California Air Resources Board*.
- Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Science & Environmental Epidemiology* 15, 185–204. <https://doi.org/10.1038/sj.jea.7500388>.
- Johnson, K., Gantt, B., VonWald, I., Clemments, A.L., 2019. PurpleAir PM_{2.5} performance across the U.S.
- Jones, R.R., Hoek, G., Fisher, J.A., Hasheminassab, S., Wang, D., Ward, M.H., Sioutas, C., Vermeulen, R., Silverman, D.T., 2020. Land use regression models for ultrafine particles, fine particles, and black carbon in Southern California. *Science of The Total Environment* 699, 134234. <https://doi.org/10.1016/j.scitotenv.2019.134234>.
- Journal, A.G., Huijbregts, C.J., 1978. Mining geostatistics. Academic Press, London, New York.
- Klot, S., 2011. Equivalence of using nested buffers and concentric adjacent rings as predictors in land use regression models. *Atmospheric Environment - ATMOS ENVIRON* 45, 4108–4110. <https://doi.org/10.1016/j.atmosenv.2011.04.069>.
- Kosmopoulos, G., Salamalikis, V., Pandis, S.N., Yannopoulos, P., Bloutsos, A.A., Kazantzidis, A., 2020. Low-cost sensors for measuring airborne particulate matter: Field evaluation and calibration at a South-Eastern European site. *The Science of the Total Environment* 748, 141396. <https://doi.org/10.1016/j.scitotenv.2020.141396>.
- Kuang, X.M., Gonzalez, D.H., Scott, J.A., Vu, K., Hasson, A., Charbouillot, T., Hawkins, L., Paulson, S.E., 2020. Cloud Water Chemistry Associated with Urban Aerosols: Rapid Hydroxyl Radical Formation, Soluble Metals, Fe(II), Fe(III), and Quinones. *ACS Earth and Space Chemistry* 4, 67–76. <https://doi.org/10.1021/acsearthspacechem.9b00243>.
- Kuang, X.M., Scott, J.A., da Rocha, G.O., Betha, R., Price, D.J., Russell, L.M., Cocker, D. R., Paulson, S.E., 2017. Hydroxyl radical formation and soluble trace metal content in particulate matter from renewable diesel and ultra low sulfur diesel in at-sea operations of a research vessel. *Aerosol Sci. Technol.* 51, 147–158. <https://doi.org/10.1080/02786826.2016.1271938>.
- Lakey, P.S.J., Berkemeier, T., Tong, H., Arangio, A.M., Lucas, K., Pöschl, U., Shiraiwa, M., 2016. Chemical exposure-response relationship between air pollutants and reactive oxygen species in the human respiratory tract. *Scientific Reports* 6, 32916. <https://doi.org/10.1038/srep32916>.
- Li, L., Wu, J., Wilhelm, M., Ritz, B., 2012. Use of generalized additive models and cokriging of spatial residuals to improve land-use regression estimates of nitrogen oxides in Southern California. *Atmospheric Environment (Oxford, England: 1994)* 55, 220–228. <https://doi.org/10.1016/j.atmosenv.2012.03.035>.
- Li, X., Kuang, X.M., Yan, C., Ma, S., Paulson, S.E., Zhu, T., Zhang, Y., Zheng, M., 2019. Oxidative Potential by PM_{2.5} in the North China Plain: Generation of Hydroxyl Radical. *Environmental Science & Technology* 53, 512–520. <https://doi.org/10.1021/acs.est.8b05253>.
- Liang, Y., Sengupta, D., Campmier, M.J., Apte, J., Lunderberg, D.M., Goldstein, A., 2021. Wildfire Smoke Impacts on Indoor Air Quality Assessed Using Crowdsourced Data in California. <https://doi.org/10.26434/chemrxiv.14739219.v1>.
- Matheron, G., 1979. Recherche de simplification dans un problème de cokrigage. Publication N-628, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- Milani, M., Pucillo, F.P., Ballerini, M., Camatini, M., Gualtieri, M., Martino, S., 2004. First evidence of tyre debris characterization at the nanoscale by focused ion beam. *Mater. Charact.* 52, 283–288. <https://doi.org/10.1016/j.matchar.2004.06.001>.
- MRLC, 2019. National Land Cover Database 2019 (NLCD2019) Legend — Multi-Resolution Land Characteristics (MRLC) Consortium. Multi-Resolution Land Characteristics Consortium.
- Oroumijeh, F., Jerrett, M., Del Rosario, I., Lipsitt, J., Liu, J., Paulson, S.E., Ritz, B., Schauer, J.J., Shafer, M.M., Shen, J., Weichenthal, S., Banerjee, S., Zhu, Y., 2022. Elemental composition of fine and coarse particles across the greater Los Angeles area: Spatial variation and contributing sources. *Environ. Pollut.* 292, 118356. <https://doi.org/10.1016/j.envpol.2021.118356>.
- Pant, P., Harrison, R.M., 2013. Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: A review. *Atmos. Environ.* 77, 78–97. <https://doi.org/10.1016/j.atmosenv.2013.04.028>.
- Pebesma, E.J., Wesseling, C.G., 1998. Gstat: A program for geostatistical modelling, prediction and simulation. *Computers & Geosciences* 24, 17–31. [https://doi.org/10.1016/S0098-3004\(97\)00082-4](https://doi.org/10.1016/S0098-3004(97)00082-4).
- PurpleAir, 2021. PurpleAir — Real Time Air Quality Monitoring. PurpleAir, Inc.
- Roff, T., 2020. HPMS Public Release of Geospatial Data in Shapefile Format - Policy — Federal Highway Administration. US Department of Transportation Federal Highway Administration.
- Sharma, S., Zhang, M., Anshika, Gao, J., Zhang, H., Kota, S.H., 2020. Effect of restricted emissions during COVID-19 on air quality in India. *Science of the Total Environment* 728, 138878. <https://doi.org/10.1016/j.scitotenv.2020.138878>.
- Sinisi, S.E., van der Laan, M.J., 2004. Deletion/substitution/addition algorithm in learning with applications in genomics. *Statistical Applications in Genetics and Molecular Biology* 3. <https://doi.org/10.2202/1544-6115.1069>. Article18.
- Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., 2013. The Changing Paradigm of Air Pollution Monitoring. *Environmental Science & Technology* 47, 11369–11377. <https://doi.org/10.1021/es4022602>.
- Tripathy, S., Tunno, B.J., Michanowicz, D.R., Kinnee, E., Shmool, J.L.C., Gillooly, S., Clougherty, J.E., 2019. Hybrid land use regression modeling for estimating spatio-temporal exposures to PM_{2.5}, BC, and metal components across a metropolitan area of complex terrain and industrial sources. *The Science of the Total Environment* 673, 54–63. <https://doi.org/10.1016/j.scitotenv.2019.03.453>.
- Valko, M., Morris, H., Cronin, M.T.D., 2005. Metals, toxicity and oxidative stress. *Curr. Med. Chem.* 12, 1161–1208. <https://doi.org/10.2174/0929867053764635>.
- Wang, J.M., Jeong, C.-H., Hilker, N., Healy, R.M., Sofowote, U., Debosz, J., Su, Y., Munoz, A., Evans, G.J., 2021. Quantifying metal emissions from vehicular traffic using real world emission factors. *Environ. Pollut.* 268, 115805. <https://doi.org/10.1016/j.envpol.2020.115805>.
- Wang, Q., Li, J., Yang, J., Chen, Y., Li, Y., Li, S., Xie, C., Chen, C., Wang, L., Wang, L., Wang, W., Tong, S., Sun, Y., 2020. Seasonal characterization of aerosol composition and sources in a polluted city in Central China. *Chemosphere* 258, 127310. <https://doi.org/10.1016/j.chemosphere.2020.127310>.
- Wang, Y., Wang, Q., Ye, J., Yan, M., Qin, Q., Prevot, A.S.H., Cao, J., 2019. A Review of Aerosol Chemical Composition and Sources in Representative Regions of China during Wintertime. *Atmosphere* 10, 277. <https://doi.org/10.3390/atmos10050277>.
- Weichenthal, S., Shekarzifard, M., Kulka, R., Lakey, P.S.J., Al-Rijleh, K., Anowar, S., Shiraiwa, M., Hatzopoulou, M., 2018. Spatial variations in the estimated production of reactive oxygen species in the epithelial lung lining fluid by iron and copper in fine particulate air pollution. *Environmental Epidemiology* 2, e020. <https://doi.org/10.1097/EE9.0000000000000020>.
- Yang, A., Wang, M., Eeftens, M., Beelen, R., Dons, E., Leseman, D.L.A.C., Brunekreef, B., Cassee, F.R., Janssen, N.A.H., Hoek, G., 2015. Spatial Variation and Land Use Regression Modeling of the Oxidative Potential of Fine Particles. *Environ. Health Perspect.* 123, 1187–1192. <https://doi.org/10.1289/ehp.1408916>.
- Zhang, J.J.Y., Sun, L., Barrett, O., Bertazzon, S., Underwood, F.E., Johnson, M., 2015. Development of land-use regression models for metals associated with airborne particulate matter in a North American city. *Atmos. Environ.* 106, 165–177. <https://doi.org/10.1016/j.atmosenv.2015.01.008>.