

Supplementary Information

Mobile monitoring of traffic-related air pollution for a prospective cohort study in the greater Seattle area

Magali N. Blanco,^a Amanda Gassett,^a Timothy Gould,^b Annie Doubleday,^a David L. Slager,^a
Elena Austin,^a Edmund Seto,^a Timothy Larson,^{a,b} Julian Marshall,^b Lianne Sheppard^{a,c}

^aDepartment of Environmental and Occupational Health Sciences, School of Public Health,
University of Washington, Hans Rosling Center for Population Health, 3980 15th Ave NE,
Seattle, WA 98195

^bDepartment of Civil & Environmental Engineering, College of Engineering, University of
Washington, 201 More Hall, Box 352700, Seattle, WA 98195

^cDepartment of Biostatistics, School of Public Health, University of Washington, Hans Rosling
Center for Population Health, 3980 15th Ave NE, Seattle, WA 98195

Table of Contents

1	METHODS	1
1.1	STUDY DESIGN	1
1.2	QUALITY ASSURANCE AND QUALITY CONTROL	8
1.3	COMPUTATION	15
2	RESULTS	16
2.1	DATA COLLECTED	16
2.2	SITE VISITS	19
2.3	ANNUAL AVERAGES	27
2.4	COLLOCATIONS AT REGULATORY MONITORING SITES.....	37
3	REFERENCES	39

List of Tables

TABLE S1.	ROUTE STATISTICS.	3
TABLE S2.	AIR POLLUTANTS AND OTHER PARAMETERS MEASURED WITH MOBILE MONITORING.	5
TABLE S3.	DISTRIBUTION OF CALIBRATION CURVE COEFFICIENT ESTIMATES.....	10
TABLE S4.	DISTRIBUTION OF TEMPERATURE AND RELATIVE HUMIDITY CONDITIONS INSIDE THE MANIFOLD DURING SITE VISITS (N=9,047 TOTAL). ^A	14
TABLE S5.	COLLOCATION REGULATORY SITES AND SIMILAR PARAMETERS MEASURED. ^A	14
TABLE S6.	DISTRIBUTION OF SITE VISIT POLLUTANT CONCENTRATIONS (N = 309 SITES X ~ 29 VISITS/SITE).....	17
TABLE S7.	ORIGINAL AND FINAL MOBILE MONITORING STOP MEASUREMENTS (~2 MIN EACH).	18
TABLE S8.	COMPARISON OF WITHIN- AND BETWEEN- SITE VARIABILITY.	21
TABLE S9.	PERCENT (%) OF VARIABILITY IN VISIT CONCENTRATIONS EXPLAINED BY SPATIAL AND TEMPORAL FACTORS. ^A	26
TABLE S10.	AVAILABLE GEOGRAPHIC COVARIATES (GEOCOVARIATES) AT MONITORING AND COHORT LOCATIONS. ...	30

List of Figures

FIGURE S1.	COVARIATE DISTRIBUTIONS OF MOBILE MONITORING ROAD SEGMENTS (N=6,786 100 M SEGMENTS) AND ACT COHORT LOCATIONS (N=10,330). NOTE THAT FIGURE 2 IN THE MAIN TEXT SHOWS THE DISTRIBUTION OF STOP MONITORING LOCATIONS RATHER THAN ROAD SEGMENTS.	2
FIGURE S2.	MEAN PERCENT ERROR ($(\text{ESTIMATED_CONC} - \text{TRUE_CONC})/\text{TRUE_CONC} * 100$) IN THE ESTIMATED NITROGEN MONOXIDE (NO) ANNUAL AVERAGE AT THE 10 TH & WELLER (AQS10W; NEAR-ROAD SITE) AND BEACON HILL (AQSBH; BACKGROUND SITE) REGULATORY SITES FROM REPEATED SHORT-TERM RANDOM SAMPLES (2- AND 60-MIN), WHEN COMPARED TO THE “TRUE” ANNUAL AVERAGE ESTIMATED FROM ALL THE AVAILABLE 2019 DATA. ESTIMATES ARE FOR 10,000 SIMULATIONS OF RANDOM SAMPLING WITHOUT REPLACEMENT. THE BLUE VERTICAL LINE IS FOR 25 REPEAT VISITS.	4
FIGURE S3.	IN-VEHICLE CONFIGURATION.....	7
FIGURE S4.	MANIFOLD SCHEMATICS.....	7
FIGURE S5.	PARTICLE INSTRUMENT RESPONSES TO FILTERED AIR (NEAR 0 PT/CM ³). DOTS SHOW MEDIAN, TWO-MINUTE INSTRUMENT READINGS. RED LINES ARE “LOW” AMBIENT CONCENTRATION REFERENCES, BASED ON THE 5 TH QUANTILE OF STOP CONCENTRATIONS FOR EACH POLLUTANT. RESPONSES FOR PRIMARY INSTRUMENTS ARE BASED ON MORE INSTRUMENT CHECKS (DATA) THROUGHOUT THE STUDY PERIOD.	12

FIGURE S6. COMPARISON OF TWO-MINUTE MEDIAN STOP CONCENTRATIONS FROM INSTRUMENT COLLOCATIONS. GAS VALUES ARE POST CALIBRATION. UFP INSTRUMENTS REPORT NUMBER CONCENTRATIONS: PARTICLES (PT) PER CM ³ .	13
FIGURE S7. NUMBER OF SITE VISITS PER TIME PERIOD. SHOWING UFP DATA, THOUGH ALL INSTRUMENTS WERE SIMILAR.	16
FIGURE S8. INTERQUARTILE RANGE (IQR) OF SITE VISIT CONCENTRATIONS FOR BC, CO ₂ , NO ₂ AND PM _{2.5} (N~29 VISITS PER SITE).	19
FIGURE S9. INTERQUARTILE RANGE (IQR) OF SITE OF VISIT CONCENTRATIONS (N~29 VISITS PER SITE) FOR UFP MEASURES WITH DIFFERENT INSTRUMENTS AND PARTICLE SIZE RANGES (NM).	20
FIGURE S10. MEDIAN SITE VISIT CONCENTRATIONS BY SEASON. LINES CONNECT MEDIAN CONCENTRATIONS. BOXES SHOW THE 25 TH , 50 TH AND 75 TH QUANTILE; WHISKERS SHOW THE 5 TH AND 95 TH QUANTILES.	22
FIGURE S11. MEDIAN SITE VISIT CONCENTRATIONS BY DAY OF THE WEEK. LINES CONNECT MEDIAN CONCENTRATIONS. BOXES SHOW THE 25 TH , 50 TH AND 75 TH QUANTILE; WHISKERS SHOW THE 5 TH AND 95 TH QUANTILES.	23
FIGURE S12. MEDIAN SITE VISIT CONCENTRATIONS BY HOUR OF THE DAY. LINES CONNECT MEDIAN CONCENTRATIONS. BOXES SHOW THE 25 TH , 50 TH AND 75 TH QUANTILE; WHISKERS SHOW THE 5 TH AND 95 TH QUANTILES.	24
FIGURE S13. SITE-SPECIFIC POLLUTANT CONCENTRATIONS OVER THE COURSE OF THE STUDY. THIN LINES SHOW SITE-SPECIFIC SMOOTH (LOESS) FITS FOR MEDIAN VISIT CONCENTRATIONS (N~29 VISITS/SITE). BLACK LINES SHOW THE OVERALL SMOOTH TRENDS FOR ALL OF THE SITES.	25
FIGURE S14. ANNUAL AVERAGE SITE CONCENTRATIONS (N=309). BOXPLOTS ARE TRADITIONAL AND SHOW ALL OF THE SITE AVERAGES SUCH THAT BOXES SHOW THE 25 TH , 50 TH AND 75 TH QUANTILE (Q); WHISKERS SHOW A 1.5*(Q75-Q25) DISTANCE AWAY FROM THE 25 TH OR 75 TH QUANTILE, AND DOTS ARE OUTLIERS.	27
FIGURE S15. ANNUAL AVERAGE POLLUTANT CORRELATIONS (N=309 SITES). LOWER PANELS SHOW SCATTERPLOTS WITH LOESS LINES AND 95% CONFIDENCE INTERVALS; UPPER PANELS SHOW PEARSON CORRELATIONS (R), WITH HIGHER VALUES IN DARKER REDS; DIAGONAL PANELS SHOW DENSITY PLOTS.	28
FIGURE S16. ANNUAL AVERAGE TOTAL UFP CONCENTRATIONS AT MONITORING SITES (N=309) FOR DIFFERENT UFP MEASURES.	29
FIGURE S17. ANNUAL AVERAGE TOTAL SCALED AIR POLLUTION, DEFINED AT EACH SITE AS: $Sum\ of\ scores = i = 1n(xi, j - medianXi) / IQR(Xi)$, WHERE X IS EACH INDIVIDUAL POLLUTANT (N=5) AND J INDEXES EACH INDIVIDUAL POLLUTANT MEASUREMENT. COLORS INDICATE WHERE AIR POLLUTION LEVELS WERE HIGHER OR LOWER THAN THE OVERALL AVERAGE (TOTAL SCALED AIR POLLUTION = 0). SHAPES INDICATE THE LEADING POLLUTANT (HIGHEST SCORE) AT EACH SITE. UFP LEVELS ARE FOR PARTICLES 20-1,000 NM (P-TRAK).	35
FIGURE S18. LASSO REGRESSION COEFFICIENT ESTIMATES FOR ANNUAL AVERAGE POLLUTANT CONCENTRATIONS. COVARIATES WITHOUT BUFFER VALUES (SHOWN IN THE LEGEND AS "0" (M)) ARE FOR PROXIMITY AND ELEVATION VARIABLES RATHER THAN VALUES FOR A BUFFERED RADIUS. SHOWING THE TOP 10 COVARIATES WITH THE LARGEST COEFFICIENT ESTIMATES FOR EACH POLLUTANT. SEE SI TABLE S8 FOR COVARIATE DEFINITIONS.	36
FIGURE S19. COMPARISON OF TWO-MINUTE MEDIAN CONCENTRATIONS FROM MOBILE MONITORING AND THE DEPARTMENT OF ECOLOGY (DOE) READINGS AT AIR QUALITY SYSTEM (AQS) COLLOCATION SITES. MOBILE MONITORING PM _{2.5} CONCENTRATIONS ARE FROM CALIBRATED NEPHELOMETER READINGS (SEE METHODS). DOE PM _{2.5} CONCENTRATIONS ARE FROM NEPHELOMETERS WHEN AVAILABLE (AQSD, AQSK, AQSTUK), OTHERWISE THEY ARE FROM GRAVIMETRIC AND BETA ATTENUATION (BAM) METHODS, WHICH ARE UPDATED LESS FREQUENTLY (AQSD, AQSK, AQSTUK – READINGS ARE BASED ON ROLLING 1-HOUR ESTIMATES UPDATED EVERY 6 MINUTES, AQSBH – READINGS ARE UPDATED HOURLY).	37
FIGURE S20. COMPARISON OF ANNUAL AVERAGE ESTIMATES FROM MOBILE MONITORING AND THE DEPARTMENT OF ECOLOGY (DOE) READINGS AT AIR QUALITY SYSTEM (AQS) COLLOCATION SITES. PLOTS COMPARE ESTIMATES USING MOBILE MONITORING STOP DATA, DOE DATA DURING THE SAME TWO-MINUTE TIME PERIODS, AND THE TRUE ANNUAL AVERAGES AT THOSE SITES USING ALL THE AVAILABLE REGULATORY MONITORING DATA FOR THE STUDY PERIOD. MOBILE MONITORING PM _{2.5} CONCENTRATIONS ARE FROM CALIBRATED NEPHELOMETER READINGS (SEE METHODS). DOE PM _{2.5} CONCENTRATIONS ARE FROM NEPHELOMETERS WHEN AVAILABLE (AQSD, AQSK, AQSTUK), OTHERWISE THEY ARE FROM GRAVIMETRIC AND BETA ATTENUATION (BAM) METHODS (AQSD, AQSK, AQSTUK – READINGS ARE BASED ON ROLLING 1-HOUR ESTIMATES UPDATED EVERY 6 MINUTES, AQSBH – READINGS ARE UPDATED HOURLY).	38

List of Equations

EQUATION S1. NEPHELOMETER LIGHT SCATTERING (BSCAT) CALIBRATION CURVE FOR PM _{2.5}	14
--	----

List of Notes

NOTE S1. SELECTING THE MONITORING LOCATIONS	1
NOTE S2. QUALITY ASSURANCE AND QUALITY CONTROL PROCEDURES	8
NOTE S3. SOFTWARE USED IN ANALYSES.	15

1 Methods

1.1 Study Design

Note S1. Selecting the monitoring locations

We used ArcMap to select 304 participant residences within our monitoring region that maximized spatial coverage. To do so, we first created a street network in ArcMap using Tiger/Line® shapefiles downloaded from the U.S. Census Bureau.^{1,2} These include all roads within the monitoring area. We divided the monitoring area into nine regions and selected approximately 34 participant residences within each of these regions (~34 locations/region x 9 regions = 304 total locations). These locations were meant to maximize spatial coverage by minimizing the distance between each selected location and all of the nearby participant locations. The initially selected locations were jittered (using the *jitter* function in R [v 3.5.1, using RStudio v 1.0.143]) to maintain participant confidentiality, and 304 new, nearby locations were identified as monitoring locations. The resulting locations were shifted anywhere from roughly a couple of houses to several blocks over. Locations were manually moved to the nearest home if the jittering caused locations to end up in a lake, park, etc. We used Google Maps Street View³ to ensure that a vehicle could safely park at each location, otherwise the stop was moved to the nearest location where it was safe to do so. We included 5 regulatory monitoring locations to obtain the final 309 monitoring locations.

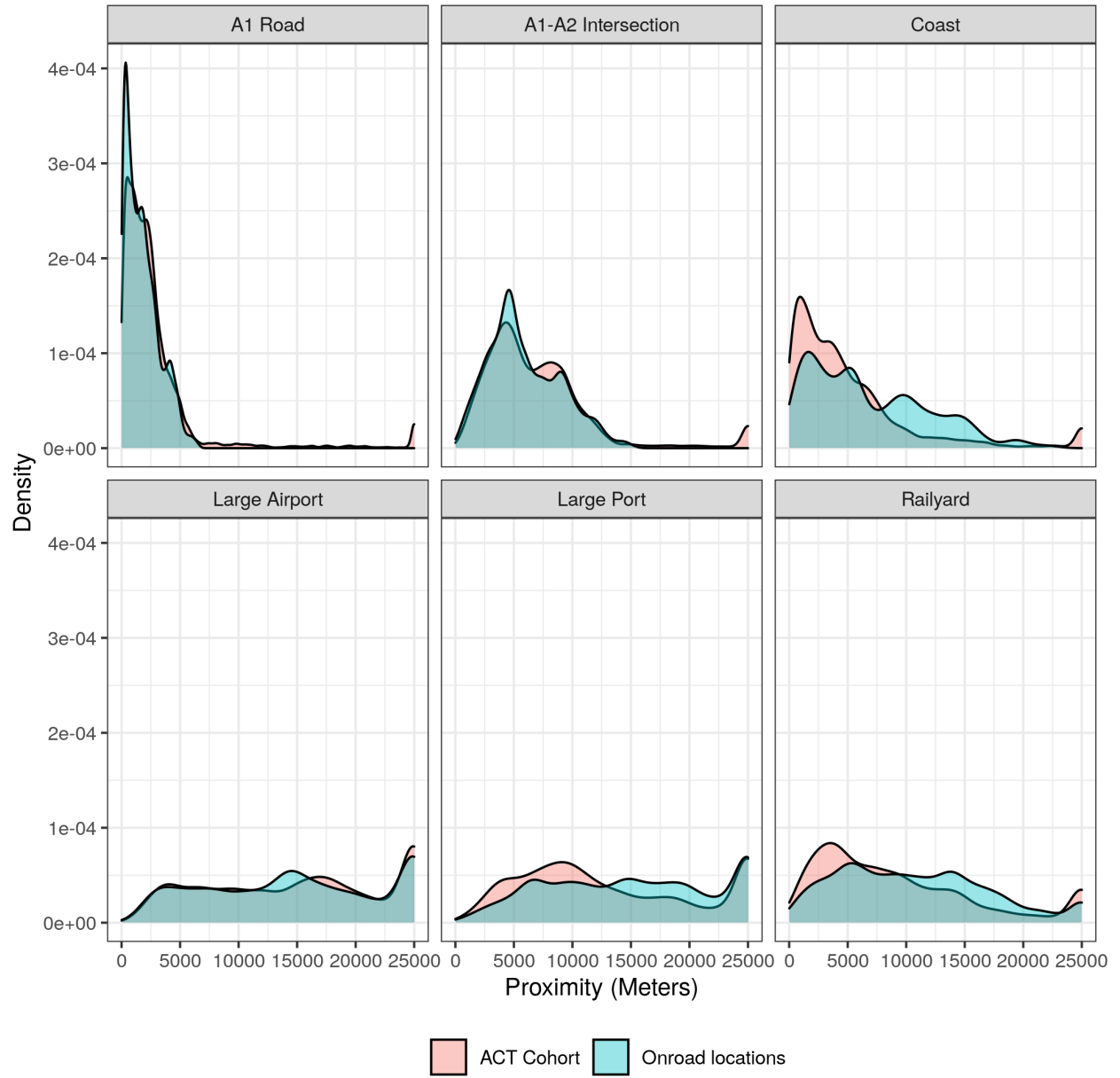


Figure S1. Covariate distributions of mobile monitoring road segments ($n=6,786$ 100 m segments) and ACT cohort locations ($n=10,330$). Note that Figure 2 in the main text shows the distribution of stop monitoring locations rather than road segments.

Table S1. Route Statistics.

Route ^a	No. Stops	No. Sampling Dates	Distance (mi)	Distance (km)	Total Distance (mi) ^b	Total Distance (km)	No. Road Segments ^c	Median (IQR) Drive Time (hr)	Total Drive Time (hr)
1	40	30	48	78	1,446	2,328	578	5.2 (4.8, 5.7)	153
2	35	30	47	75	1,397	2,247	517	4.8 (4.4, 5.1)	140
3	45	34	59	95	2,008	3,232	804	5.6 (4.9, 6.1)	170
4	33	33	66	107	2,190	3,525	619	5.0 (4.7, 5.3)	143
5	32	32	69	112	2,221	3,574	702	5.1 (4.7, 5.5)	143
6	35	32	88	142	2,829	4,554	856	5.4 (4.9, 5.8)	141
7	32	30	90	145	2,697	4,341	933	5.6 (5.2, 6.1)	163
8	28	32	104	168	3,332	5,362	942	5.7 (5.2, 6.1)	156
9	29	35	92	148	3,227	5,193	835	5.5 (5.0, 5.8)	158
Total	309	35	664	1,069	21,347	34,355	6,786	5.2 (4.8, 5.8)	1,367

^aThere were about 18 additional make-up routes (about 4-5 per quarter), each with stops from multiple routes.

^bTotal route driving distance is estimated from the route distance and the number of sampling dates. The exact distance varied based on makeup routes, route deviations, etc.

^cRoad segments are 100 m long. Segments were excluded if they were on A1 roads (e.g., I-5 and WA-520), had a median visit duration of less than 5s per visit, or had fewer than 23 visits.

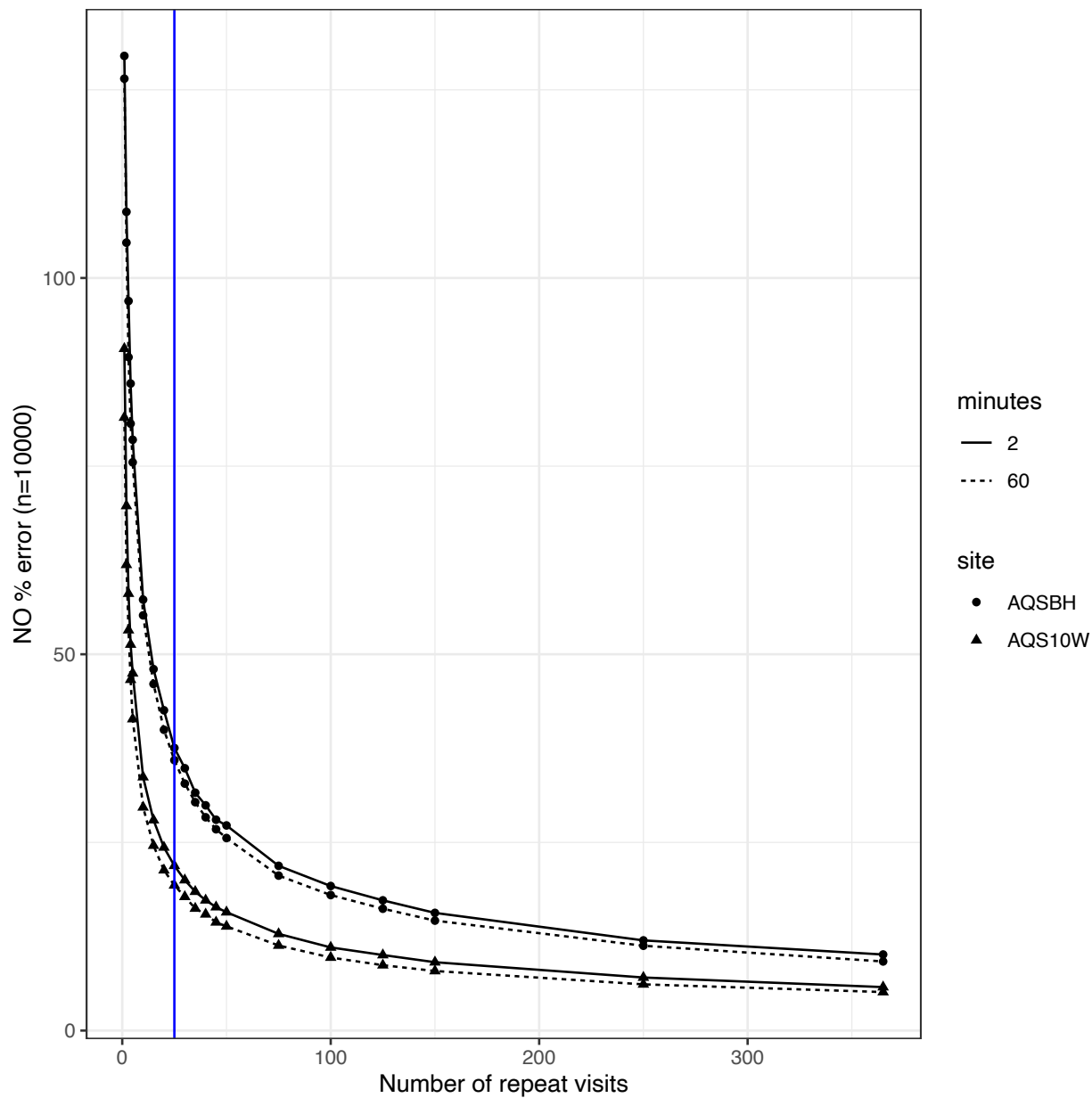


Figure S2. Mean percent error $(\text{estimated_conc} - \text{true_conc}) / \text{true_conc} * 100$ in the estimated nitrogen monoxide (NO) annual average at the 10th & Weller (AQS10W; near-road site) and Beacon Hill (AQSBH; background site) regulatory sites from repeated short-term random samples (2- and 60-min), when compared to the “true” annual average estimated from all the available 2019 data. Estimates are for 10,000 simulations of random sampling without replacement. The blue vertical line is for 25 repeat visits.

Table S2. Air pollutants and other parameters measured with mobile monitoring.

Parameter	Instrument	Manufacturer	Measurement Range	Limit of Quantification	Time Resolution
Particles (pt)					
UFP					
10-420 nm (13-bin PSD ^a)	NanoScan 3910	TSI	10 ² - 10 ⁶ pt/cm ³	10 pt/ cm ³	60 sec
10-700 nm	DiSCmini	Testo	10 ³ - 10 ⁶ pt/cm ³	500-2,000 pt/cm ³ ^b	1 sec
20-1,000 nm	PTRAK 8525	TSI	0 - 5x10 ⁵ pt/cm ³	1 pt/cm ³	1 sec
36-1,000 nm	PTRAK 8525, with diffusion screen	TSI	0 - 5x10 ⁵ pt/cm ³	1 pt/cm ³	1 sec
BC	microAeth MA200	AethLabs	0 - 10 ⁶ ng/m ³	30 ng BC/m ³ ^c	10 sec
Light scattering nephelometer (PM _{2.5})	M903	Radiance Research	0 - >1 km ⁻¹	10 ⁻⁶ m ⁻¹	10 sec
Gases					

NO ₂	CAPS NO ₂	Aerodyne Research, Inc.	0 – 2x10 ³ ppb	2 ppbv	1 sec
CO ₂	LI-850	Li-Cor	0-5x10 ³ ppm (vol)	100 ppmv	1 sec
CO ^d	CO Monitor T15N	Langan, Inc.	0-200 ppm	0.1 ppm	1 sec
Other					
Temperature	Onset UX100- 011	HOBO	-4-158°F		1 sec
Relative humidity	Onset UX100- 011	HOBO	0-95%		1 sec
Positioning & real-time tracking	DG-500	US GlobalSat	0-515 m/sec speed	2.5 m	1 sec

^a PSD: particle size distribution

^b estimate; detection limit is dependent on particle size

^c for a 5 min time base, 150 ml/min flow rate

^d CO measurements were collected but not utilized because they did not meet our quality assurance standards

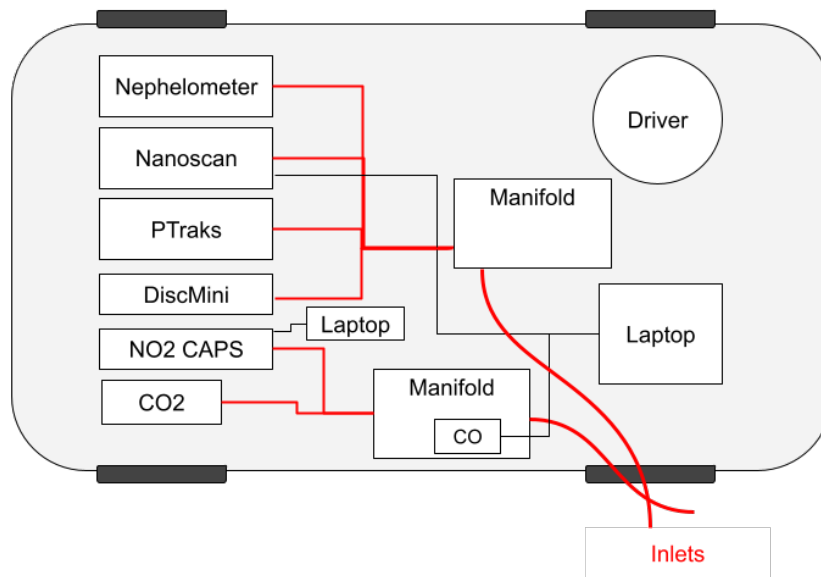


Figure S3. In-vehicle configuration

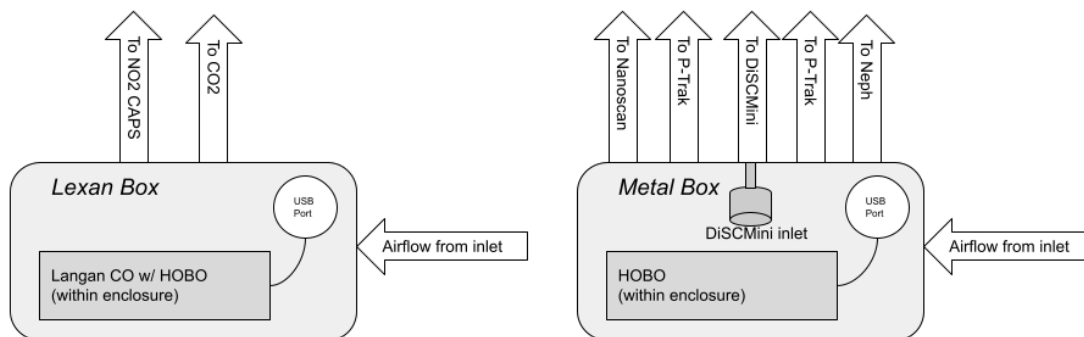


Figure S4. Manifold schematics

1.2 Quality Assurance and Quality Control

Note S2. Quality assurance and quality control procedures

Data Management System

After each drive, the field technician uploaded each instrument's file to the server using a standardized file naming convention. These files were automatically loaded into a MySQL database every morning at 4 a.m. A report was produced automatically that showed the time series plot of each instrument, counts of the times of day that each stop on the route had been visited to date, and a map that highlighted any missed stops. The driver reviewed these data before starting the next day's drive. Single missed stops could be visited on the way to another day's route or on a day dedicated to make-up stops. In addition, a project manager, information technology (IT) specialist, and several data analysts routinely reviewed and worked with the data, thus allowing for additional feedback.

We carried out an extensive independent code review of the database and made further improvements to the system before completely reloading all raw data files into the database and locking the final version.

Instrumentation

To ensure instrument accuracy, all gas instruments were calibrated before the campaign and regularly thereafter. Particle instruments were purchased new and arrived with calibration certifications, or they were compared to like instruments that had been serviced prior to the study.

Primary and backup instruments were collocated every few weeks on route to assess the precision (repeatability) of our measurements in different environments and over time.

Data Cleaning

We conducted various quality control procedures prior to conducting data analyses. We added a ten-second lag to all the instrument readings to account for the time required for a

volume of air to travel from the sampling inlet to each instrument. This was based on the manifold volumes and instrument flow rates.

Readings with instrument error codes were dropped. This included, for example, aethalometer (BC) pump flow errors and readings of NO₂ field baseline samples.

Aethalometers were checked to ensure that the filter attenuation was below 50% thus ensuring optimal instrument sensitivity at all times.^{4,5}

Gas instruments and nephelometers (which were checked for a response against CO₂ gas⁶) were calibrated various times during the study period. CO and nephelometer instruments were automatically reset during calibrations. CO₂ and NO₂ instruments were manually calibrated using least squares linear regression models with reference concentrations as the independent variable and instrument readings as the response variable.^{7,8} Particle instruments were checked for zero concentration responses by placing a high efficiency particle air (HEPA) filter on the instrument or manifold inlet.

We calculated stop visit medians (about 2 minutes worth of data) with the resulting data. Readings outside the instrument ranges, screened P-TRAK readings below 100 pt/cm³, and other UFP instrument readings below 300 pt/cm³ (NanoScans, unscreened P-TRAKS) were dropped.

We investigated collocated instrument readings to assess repeatability. Comparing instruments to one another is particularly common with particle instruments since there is no standard for the field calibration of these instruments. Backup NO₂ (“NO2_1”) and NanoScan (“PMSCAN_3”) instruments were adjusted based on readings from their respective primary instruments during the beginning of the study since these were used exclusively at the beginning of the study. We calibrated UFP readings from the two DiSCmini instruments used in this study to the mean of their responses.⁹ This was done to ensure consistency across instruments since these were equally used throughout the study period. In this approach, a calibration curve is established by fitting separate linear regression models to each instrument, with that instrument’s readings as the independent variable and the mean reading of duplicate instruments as the response variable.

The backup CO₂ instrument (CO2_19) was dropped since it produced unstable responses over time, it did not always correlate well with the primary instrument (CO2_14), and it was solely used as a collocation instrument (i.e., never on its own). All CO readings were dropped

since instruments produced unstable readings, and collocated instruments were poorly correlated with one another or observations from collocations at regulatory monitoring sites.

Quality Control Results Summary

SI Table S2 shows the calibration curve coefficient estimates used to manually adjust CO and NO₂.

In response to clean, filtered air, particle instruments generally reported near zero concentrations that were also lower than a “low” ambient concentration, as determined from the data (SI Figure S6). Some exceptions included the backup aethalometer (BC_0066), which reported negative readings, though this was based on very little data (two 2-min medians). The primary aethalometer (BC_0063) and nephelometer instruments (PM25_176), as well as the backup DiscMini instrument (PMDISC_8) additionally reported low ambient concentrations that were similar to some of their filtered air responses, suggesting that these instruments may be less sensitive to very low ambient concentrations.

Collocated instruments generally produced similar responses (SI Figure S7). As noted above, the backup CO₂ instrument (CO2_19) and all CO instruments were dropped because they did not meet quality assurance standards. Backup NO₂ and NanoScan instruments (NO2_1, PMSCAN_3) were adjusted to better align with primary instrument readings.

Temperature and relative humidity conditions inside the manifold during site visits are presented in SI Table S3.

Table S3. Distribution of calibration curve coefficient estimates.

Pollutant^a	Instrument ID	Term	N^b	Min	Median	Max
NO ₂ (ppb)	NO2_1	slope	6	1.09	1.14	1.20
NO ₂ (ppb)	NO2_1	intercept	6	-29.02	-3.15	-0.55
NO ₂ (ppb)	NO2_2	slope	18	0.55	1.05	1.23

NO ₂ (ppb)	NO2_2 ^c	intercept	18	0	0	0
CO (ppm)	CO_1	slope	16	1.05	1.10	1.16
CO (ppm)	CO_1	intercept	16	1.60	2.25	2.77
CO (ppm)	CO_190134	slope	5	0.74	0.84	0.94
CO (ppm)	CO_190134	intercept	5	1.46	2.19	2.27
CO (ppm)	CO_3	slope	13	0.20	0.41	0.82
CO (ppm)	CO_3	intercept	13	0.81	1.33	2.29

^a CO₂ was automatically reset after each calibration, and no additional adjustments were necessary.

^b N = number of calibration days.

^c A no intercept model was fit to instrument NO2_2, which reset after each baseline zero reading.

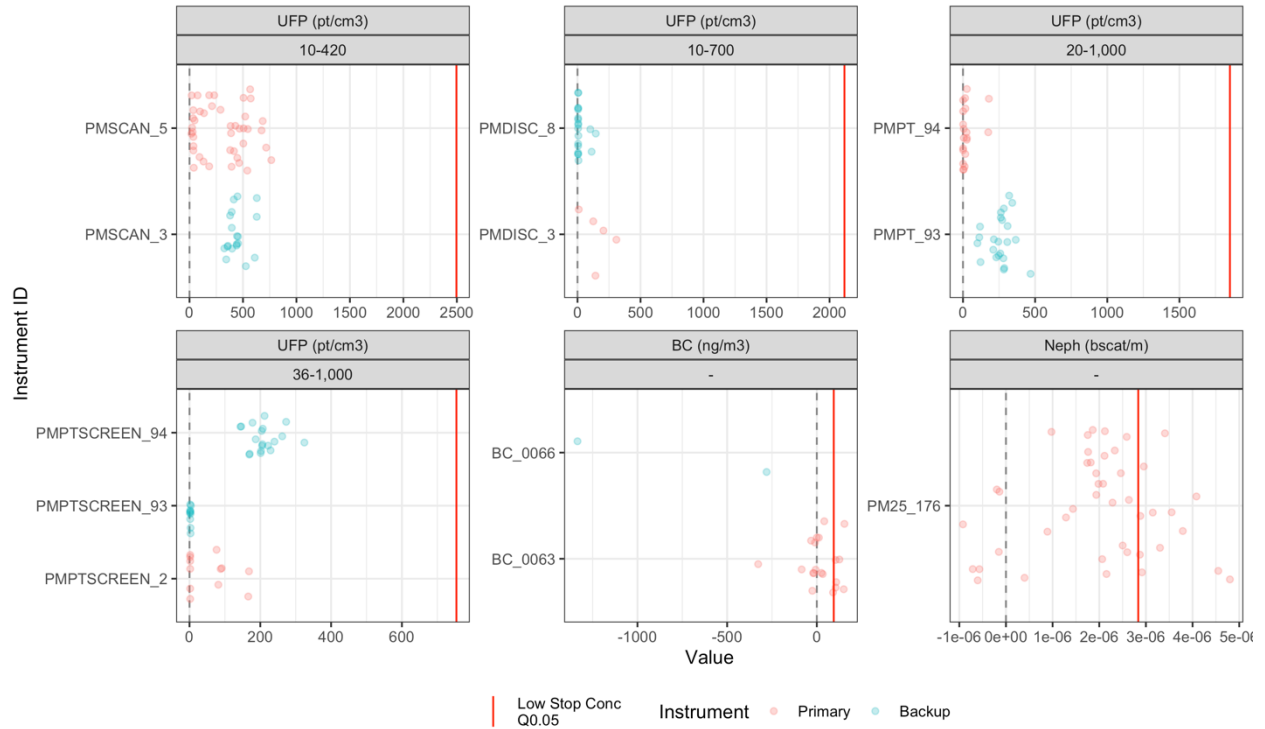


Figure S5. Particle instrument responses to filtered air (near 0 pt/cm³). Dots show median, two-minute instrument readings. Red lines are “low” ambient concentration references, based on the 5th quantile of stop concentrations for each pollutant. Responses for primary instruments are based on more instrument checks (data) throughout the study period.

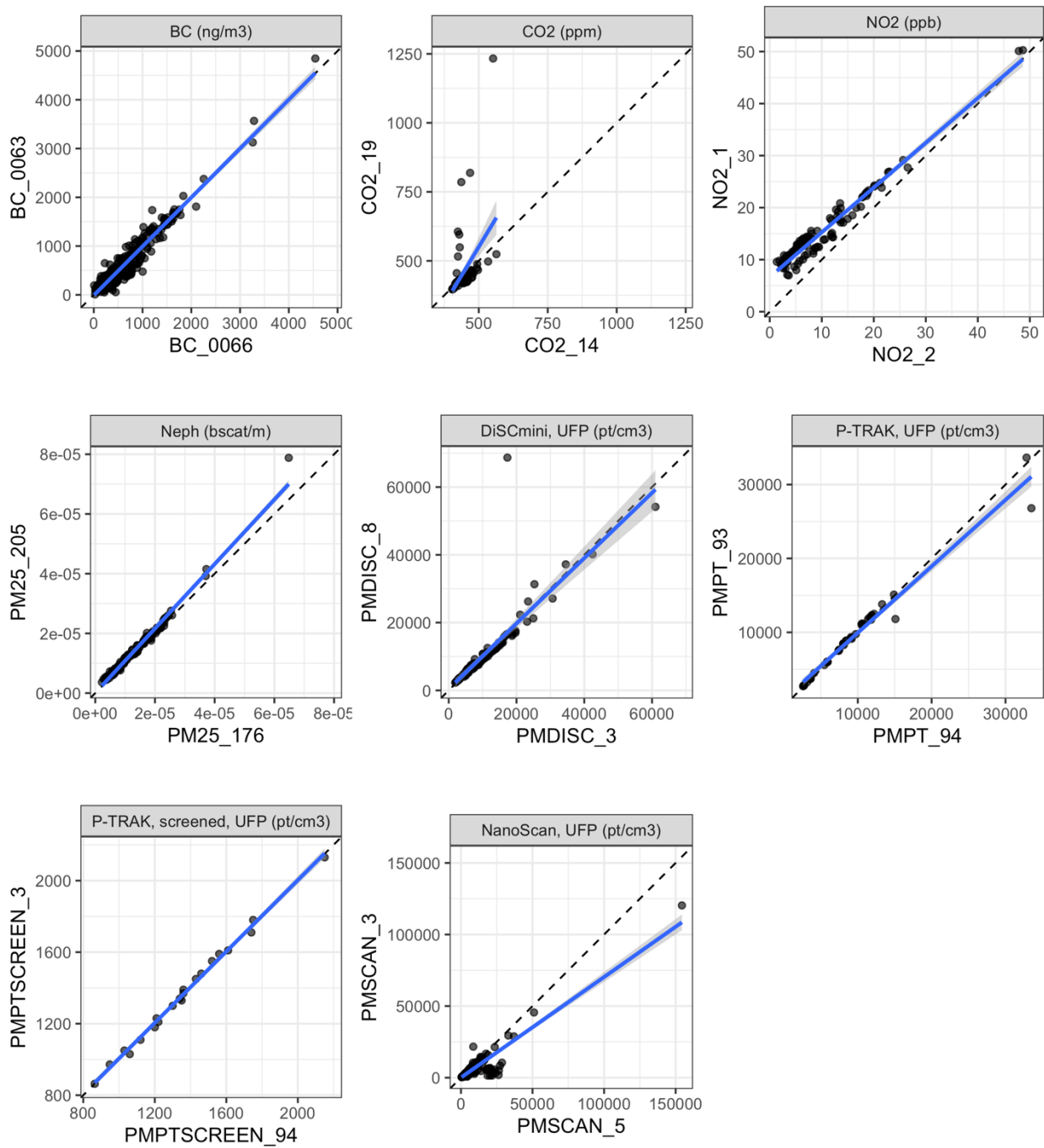


Figure S6. Comparison of two-minute median stop concentrations from instrument collocations. Gas values are post calibration. UFP instruments report number concentrations: particles (pt) per cm³.

Table S4. Distribution of temperature and relative humidity conditions inside the manifold during site visits (N=9,047 total).^a

Variable	Min	Q05	Q25	Median	Q75	Q95	Max
Relative Humidity (%)	12	26	37	44	51	61	78
Temperature (F)	50	60	64	67	72	80	96

^a Measurements are for 2-min medians from within the vehicle (manifold) for 9,047 site visits

Equation S1. Nephelometer light scattering (bscat) calibration curve for PM_{2.5}

$$PM_2 \left(\frac{ug}{m^3} \right) = 25.10(bscat \times 10^{-4}) + 1.06$$

Table S5. Collocation regulatory sites and similar parameters measured.^a

Station (ID)	Location	PM _{2.5} FRM	PM _{2.5} FEM	bscat & PM _{2.5} bscat	BC	NO ₂ ^b
10 th & Weller, Seattle (BK)	Urban Center; near-road	No	Yes	No	Yes	Yes
Tukwila Allentown (BL)	Suburban, industrial, residential	No	Yes	Yes	Yes	No
Beacon Hill (BW)	Suburban, commercial, residential	Yes	Yes	No	No	Yes
Duwamish (CE)	Urban center, industrial	No	Yes	Yes	Yes	No
James St & Central Ave, Kent (CW)	Suburban, commercial	No	Yes	Yes	Yes	No

^a FRM: federal reference method; FEM: federal equivalent method; bscat: beta light scattering; Temp: temperature (°F); RH: relative humidity

^b or NO_y-NO

1.3 Computation

Note S3. Software used in analyses.

We conducted all analyses using MySQL¹⁰ and R (v 3.6.2, using RStudio v 1.2.5033).¹¹ We used the R packages: Broom (v. 0.5.5),¹² colorspace (v. 1.4-1),¹³ cowplot (v. 1.0.0),¹⁴ dplyr (v. 1.0.6),¹⁵ fmsb (v. 0.7.1),¹⁶ forcats (v. 0.5.0),¹⁷ GGally (v. 2.1.1),¹⁸ ggmap (v. 3.0.0),¹⁹ ggplot2 (v. 3.3.3),^{20(p2)} ggpmisc (v. 0.4.0),²¹ ggpp (v. 0.4.0),²² ggpubr (v. 0.2.5),²³ ggrepel (v. 0.8.1),²⁴ ggspatial (v. 1.1.4),²⁵ gstat (v. 2.0-7),²⁶ kableExtra (v. 1.1.0),²⁷ knitr (v. 1.28),²⁸ lubridate (v. 1.7.10),²⁹ magrittr (v. 1.5),³⁰ purrr (v. 0.3.3),³¹ readr (v. 1.3.1),³² sf (v. 0.9-5),³³ spData (v. 0.3.10),³⁴ stringr (v. 1.4.0),³⁵ tibble (v. 3.1.2),³⁶ tidyr (v. 1.0.2),³⁷ tidyverse (v. 1.3.0),³⁸ units (v. 0.6-7)³⁹ and VCA (v. 1.4.2).⁴⁰ We created all maps with map tiles by Stamen Design⁴¹ under CC BY 3.0,⁴² using data by OpenStreetMap under ODbL.⁴³

2 Results

2.1 Data Collected

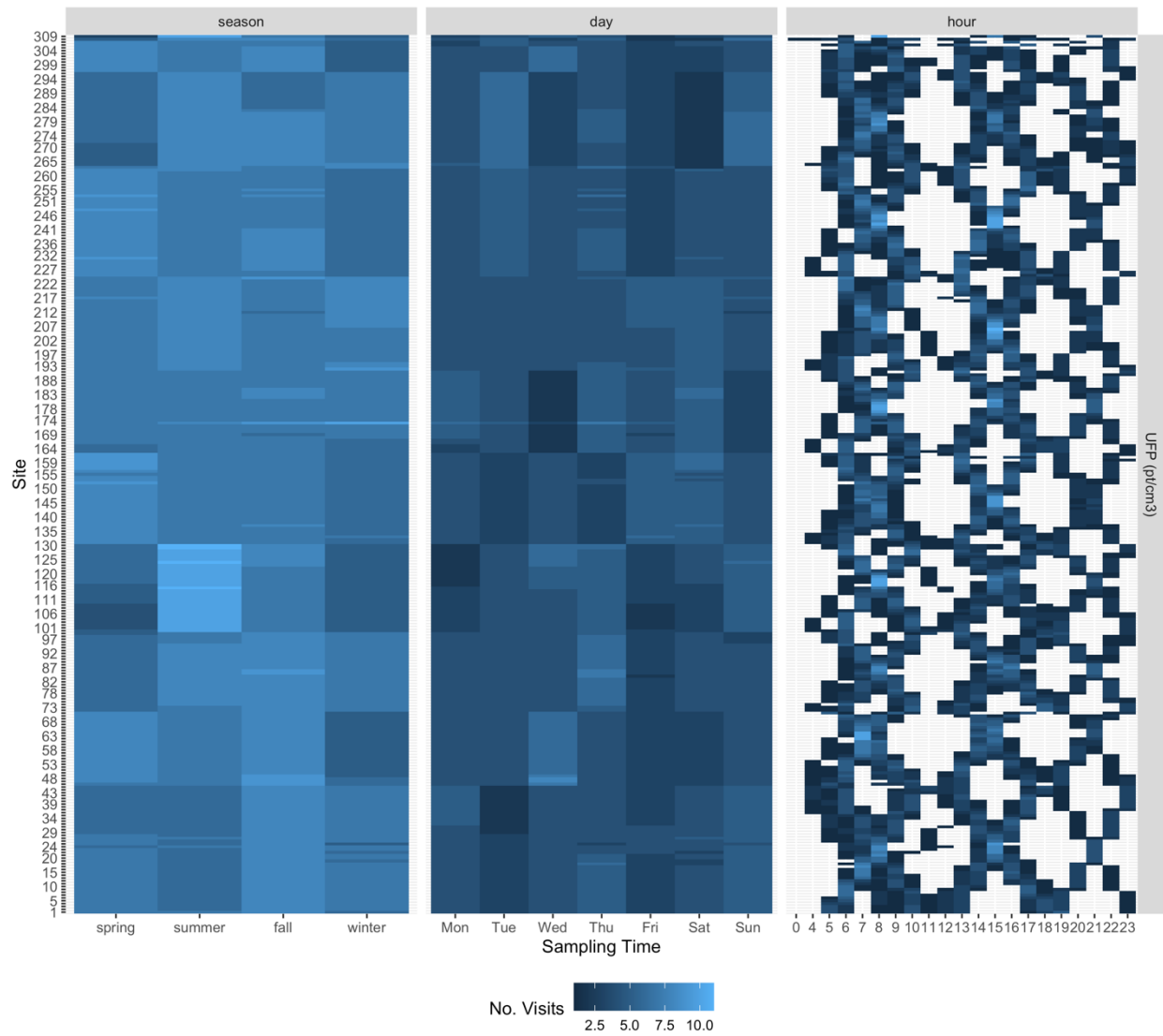


Figure S7. Number of site visits per time period. Showing UFP data, though all instruments were similar.

Table S6. Distribution of site visit pollutant concentrations (N = 309 sites x ~ 29 visits/site).

Pollutant^a	Instrument	N	Q05	Q25	Median	Q75	Q95	Q95/Q05
BC (ng/m ³)	MA200	8,861	94	242	402	694	1,736	18
CO ₂ (ppm)	SenseAir	8,950	405	415	425	441	478	1.2
NO ₂ (ppb)	CAPS	8,766	1.6	4	7.2	13	23	14
PM _{2.5} (ug/m ³)	M903	8,786	1.8	2.7	3.9	5.8	11	6.1
UFP (pt/cm ³) 10-420 nm	NanoScan	8,999	2,496	5,060	8,150	13,165	27,235	11
UFP (pt/cm ³) 10-700 nm	DiSCmini	8,697	2,118	4,336	7,028	11,413	24,575	12
UFP (pt/cm ³) 20-1,000 nm	P-TRAK	8,729	1,850	3,640	5,850	9,135	18,080	9.8
UFP (pt/cm ³) 36-1,000 nm	P-TRAK	8,908	754	1,580	2,520	4,050	8,136	11

^a UFP instruments report number concentrations for a specific particle size range: particles (pt) per cm³.

Table S7. Original and final mobile monitoring stop measurements (~2 min each).

Pollutant ^a	Original Stop Measurements ^b		Dropped Stop Measurements ^c		Final Stop Measurements ^b	
	N	%	N	%	N	%
CO ₂ (ppm)	8,982	99.28%	32	0.36%	8,950	98.93%
BC (ng/m ³)	9,005	99.54%	144	1.6%	8,861	97.94%
Neph (bscat/m)	8,802	97.29%	16	0.18%	8,786	97.12%
NO ₂ (ppb)	8,913	98.52%	147	1.65%	8,766	96.89%
UFP (pt/cm ³), 10-420 nm	9,000	99.48%	1	0.01%	8,999	99.47%
UFP (pt/cm ³), 10-700 nm	8,790	97.16%	93	1.06%	8,697	96.13%
UFP (pt/cm ³), 20-1,000 nm	8,731	96.51%	2	0.02%	8,729	96.49%
UFP (pt/cm ³), 36-1,000 nm	8,908	98.46%	0	0%	8,908	98.46%
TOTAL	71,131	98.28%	435	0.61%	70,696	97.68%

^a UFP instruments report number concentrations for a specific size range: particles (pt) per cm³.

^b Original and final stop measurement percents are based on the total number of stops that collected at least one 2-minute measurement in the campaign (9,047; Total = 72,376 = 9,047 stops x 8 instruments).

^c Measurements were dropped for various reasons: readings outside of each instrument's reporting range; NanoScans and non-screened P-TRAKS readings < 300 pt/cm³; backup CO₂ instrument (CO2_19) and all CO (both instruments) readings because these did not meet QC protocols (see Note S2). Dropped stops percents are based on the original stop measurements (the measurements actually collected).

2.2 Site Visits

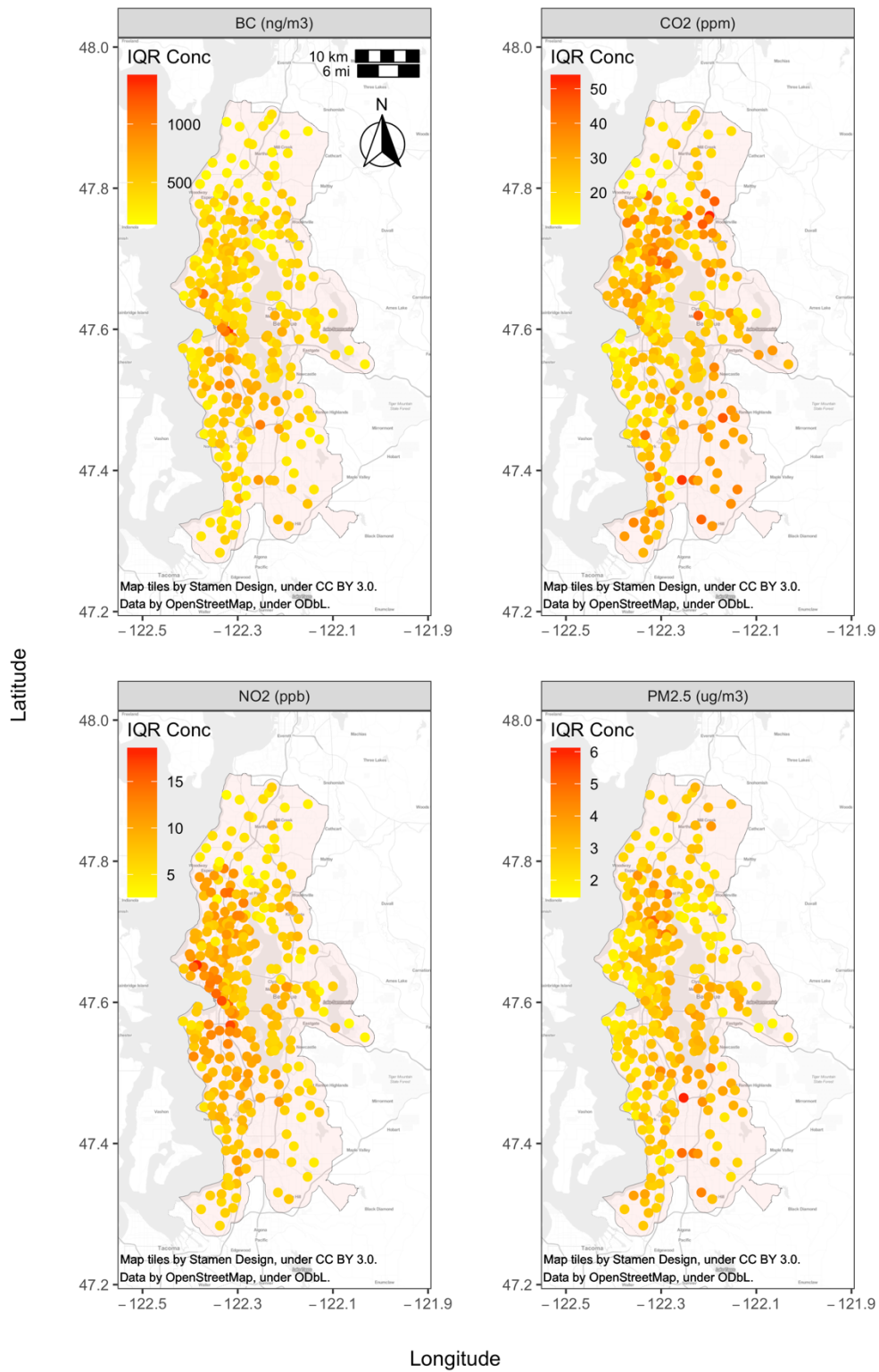


Figure S8. Interquartile range (IQR) of site visit concentrations for BC, CO₂, NO₂ and PM_{2.5} (N~29 visits per site).

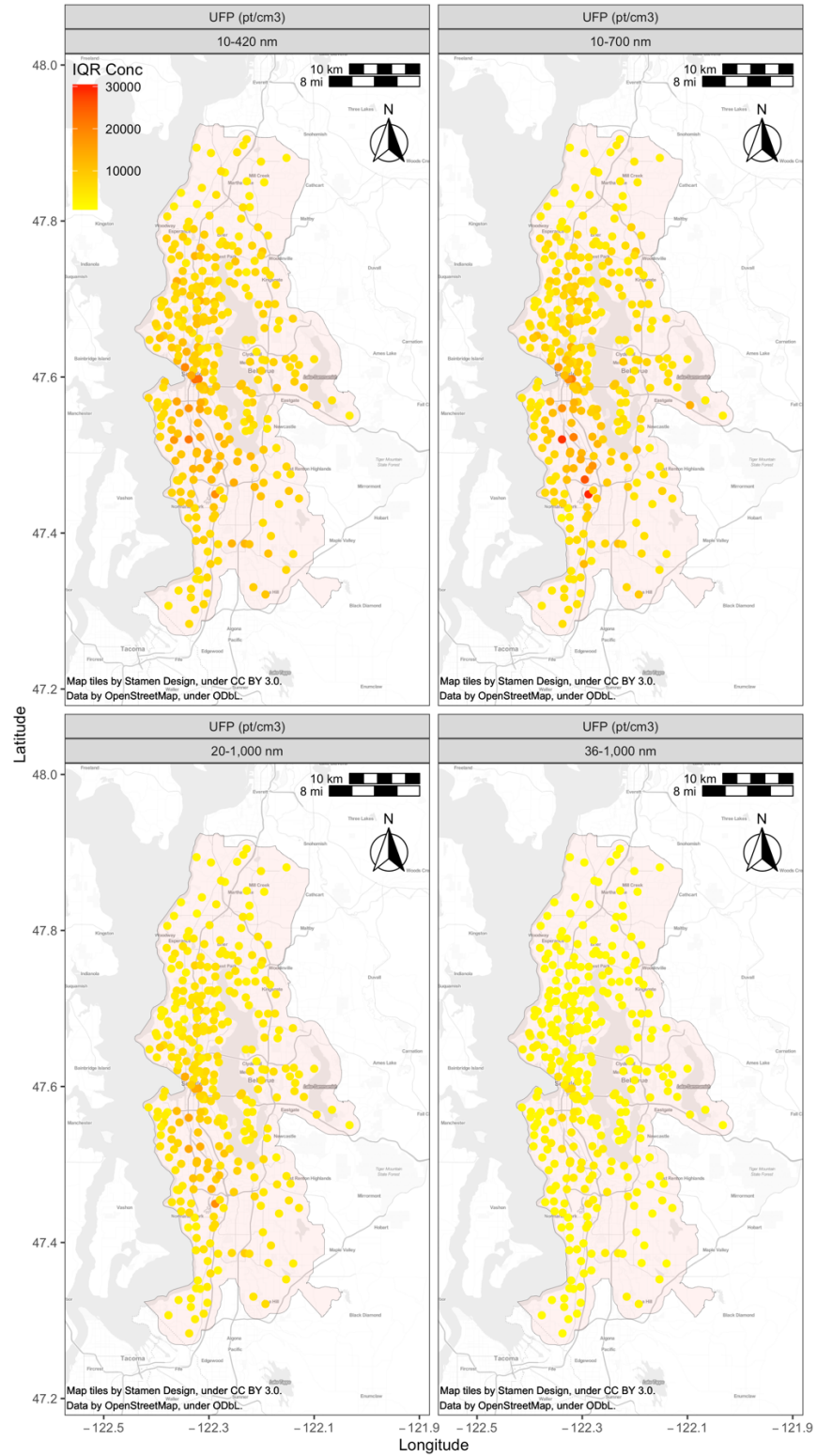


Figure S9. Interquartile range (IQR) of site of visit concentrations (N~29 visits per site) for UFP measures with different instruments and particle size ranges (nm).

Table S8. Comparison of within- and between- site variability.

Pollutant	Instrument	Within-Site Variability of Median Visit Conc's		Between-Site Variability of Annual Avg Conc's
		Median (Range) IQR	Max/Min IQR	IQR
BC (ng/m ³)	MA200	410 (142-1,421)	10.0	210
CO ₂ (ppm)	SenseAir	24 (11-54)	5.0	8
NO ₂ (ppb)	CAPS	7.7 (2.5-18.6)	7.3	3
PM _{2.5} (ug/m ³)	M903	2.7 (1.5-6.1)	4.2	0.8
UFP (pt/cm ³), 10-420	NanoScan	7,183 (2,834-22,625)	8.0	2,972
UFP (pt/cm ³), 10-700 nm	DiSCmini	6,103 (1,872-30,477)	16.3	3,192
UFP (pt/cm ³), 20-1,000 nm	P-TRAK	4,720 (1,711-18,755)	11.0	2,354
UFP (pt/cm ³), 36-1,000 nm	P-TRAK, Screen	2,185 (765-7,045)	9.2	985

^a Within-site variability is for site-specific IQR values of median visit concentrations (N = 309 site IQRs, each site has approximately 29 visits) and thus shows visit concentration data.

^b Between-site variability shows the IQR of annual average concentrations (N=309 sites), as shown in Table 1 of the main document.

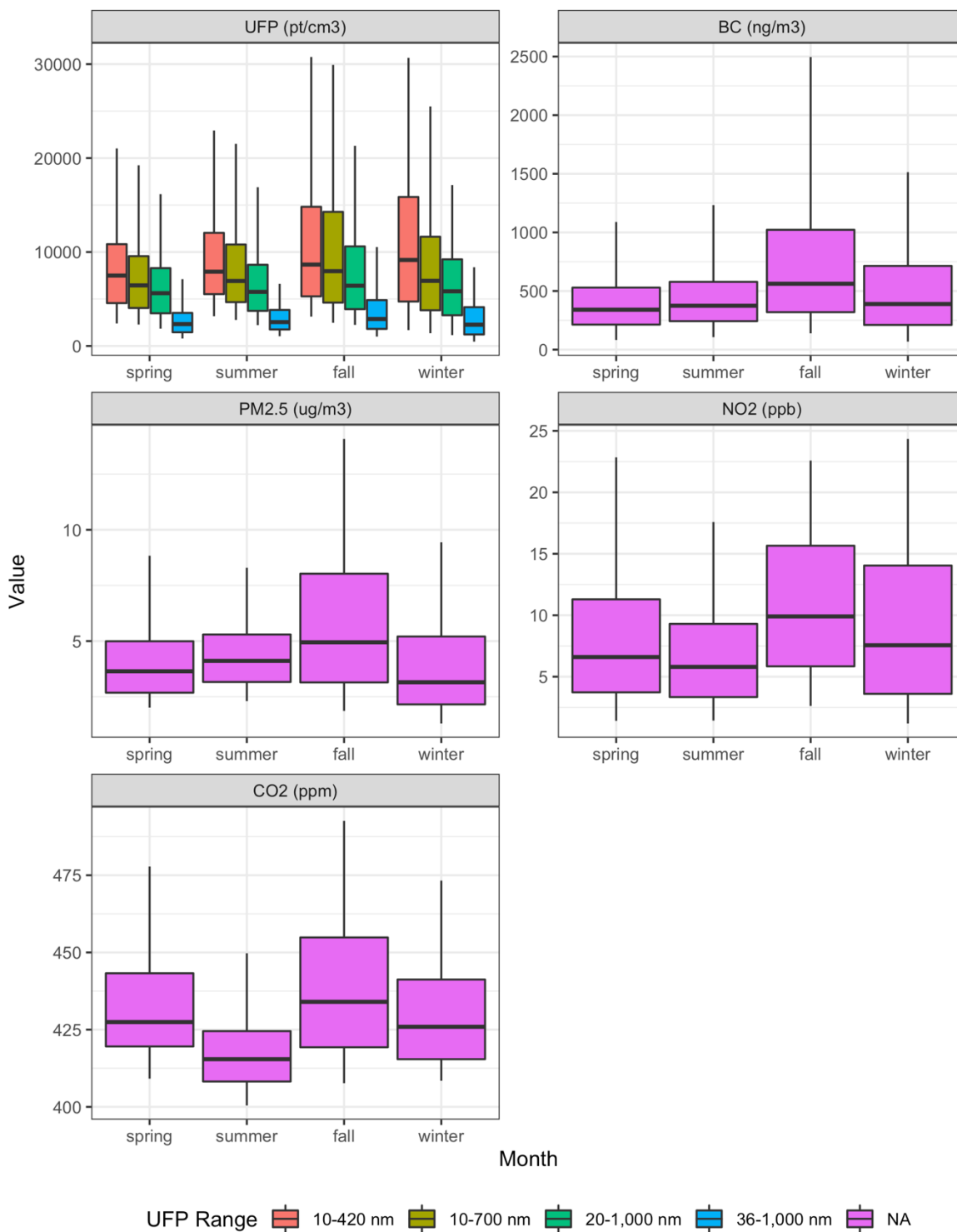


Figure S10. Median site visit concentrations by season. Lines connect median concentrations. Boxes show the 25th, 50th and 75th quantile; whiskers show the 5th and 95th quantiles.

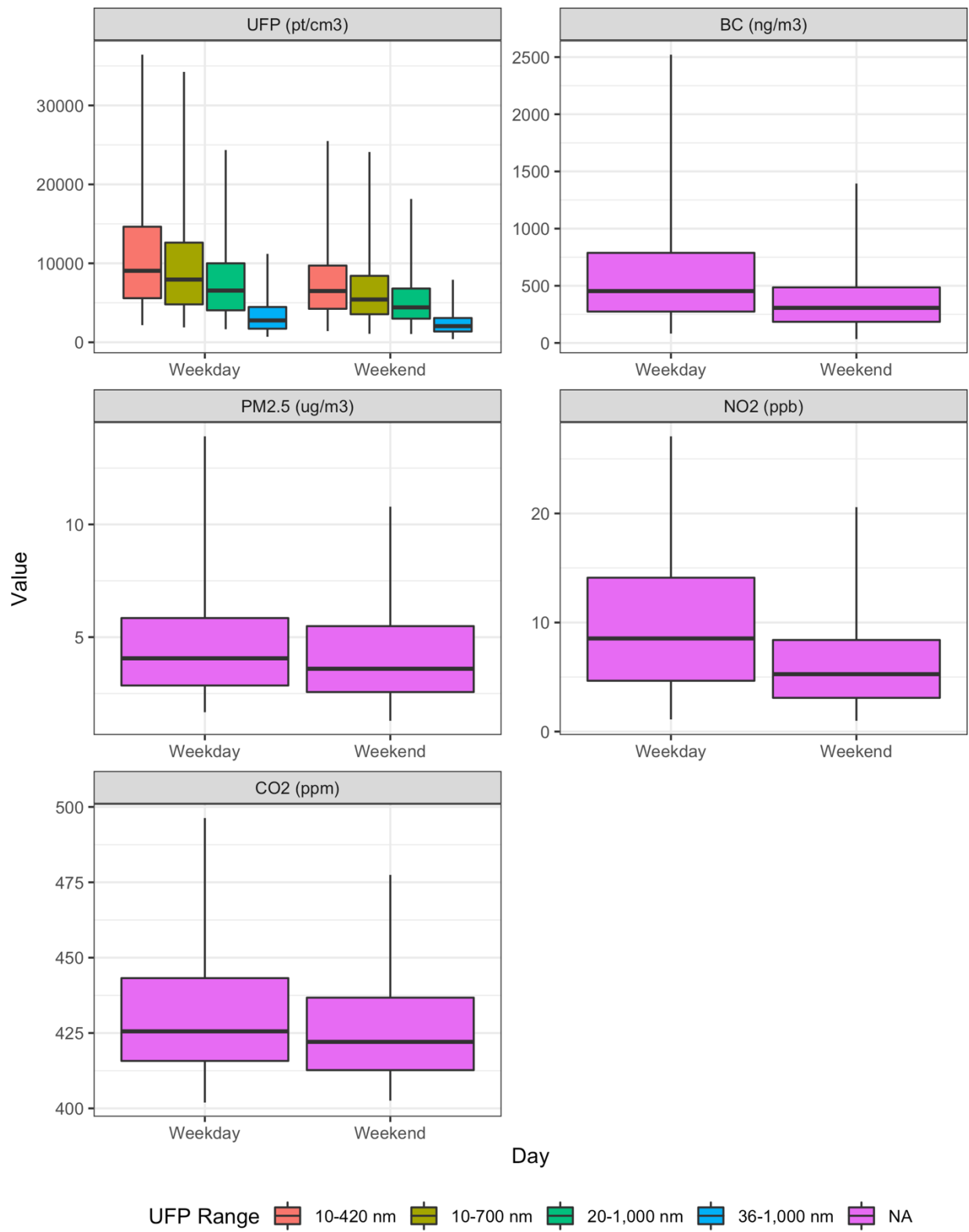


Figure S11. Median site visit concentrations by day of the week. Lines connect median concentrations. Boxes show the 25th, 50th and 75th quantile; whiskers show the 5th and 95th quantiles.

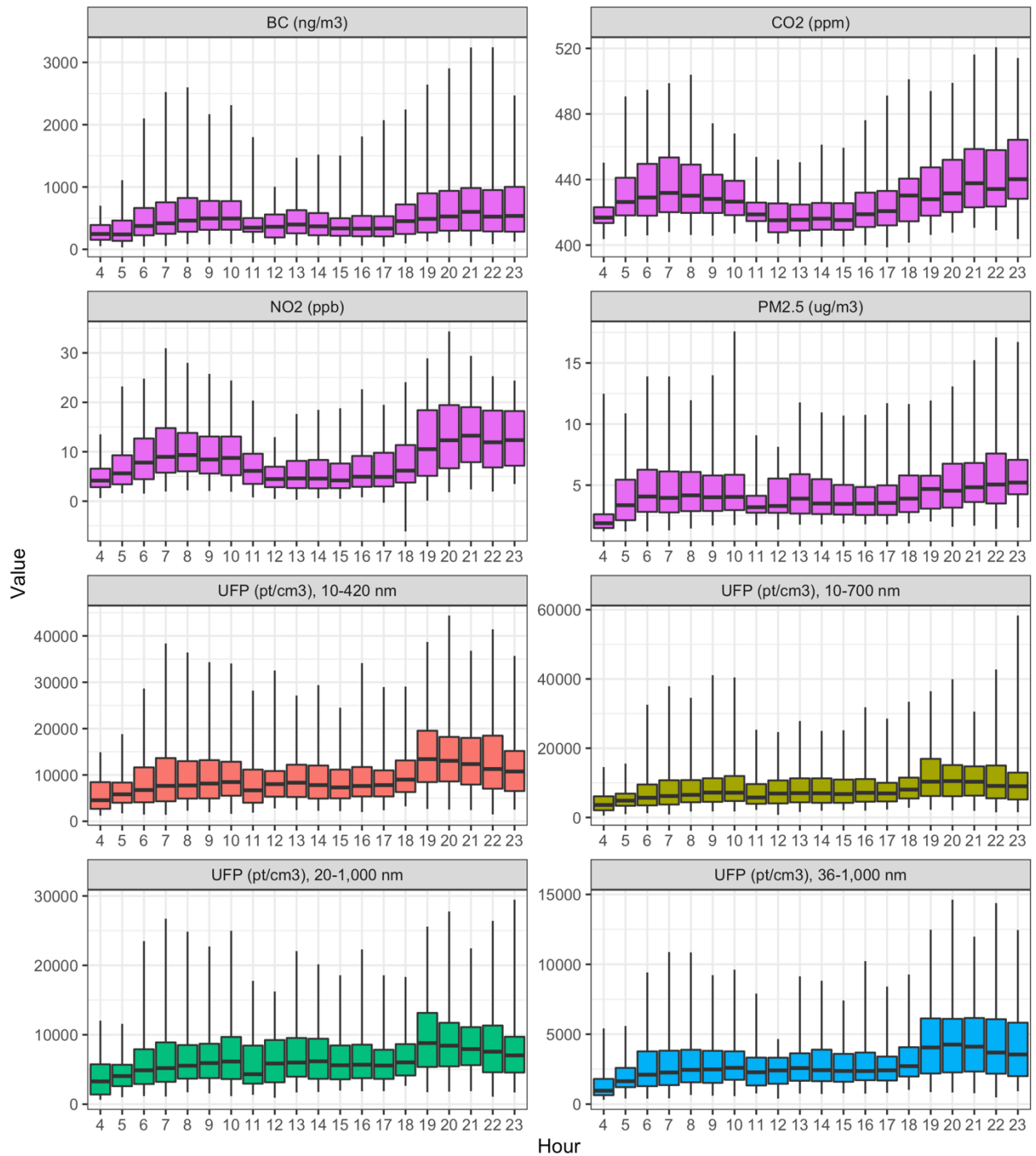


Figure S12. Median site visit concentrations by hour of the day. Lines connect median concentrations. Boxes show the 25th, 50th and 75th quantile; whiskers show the 5th and 95th quantiles.

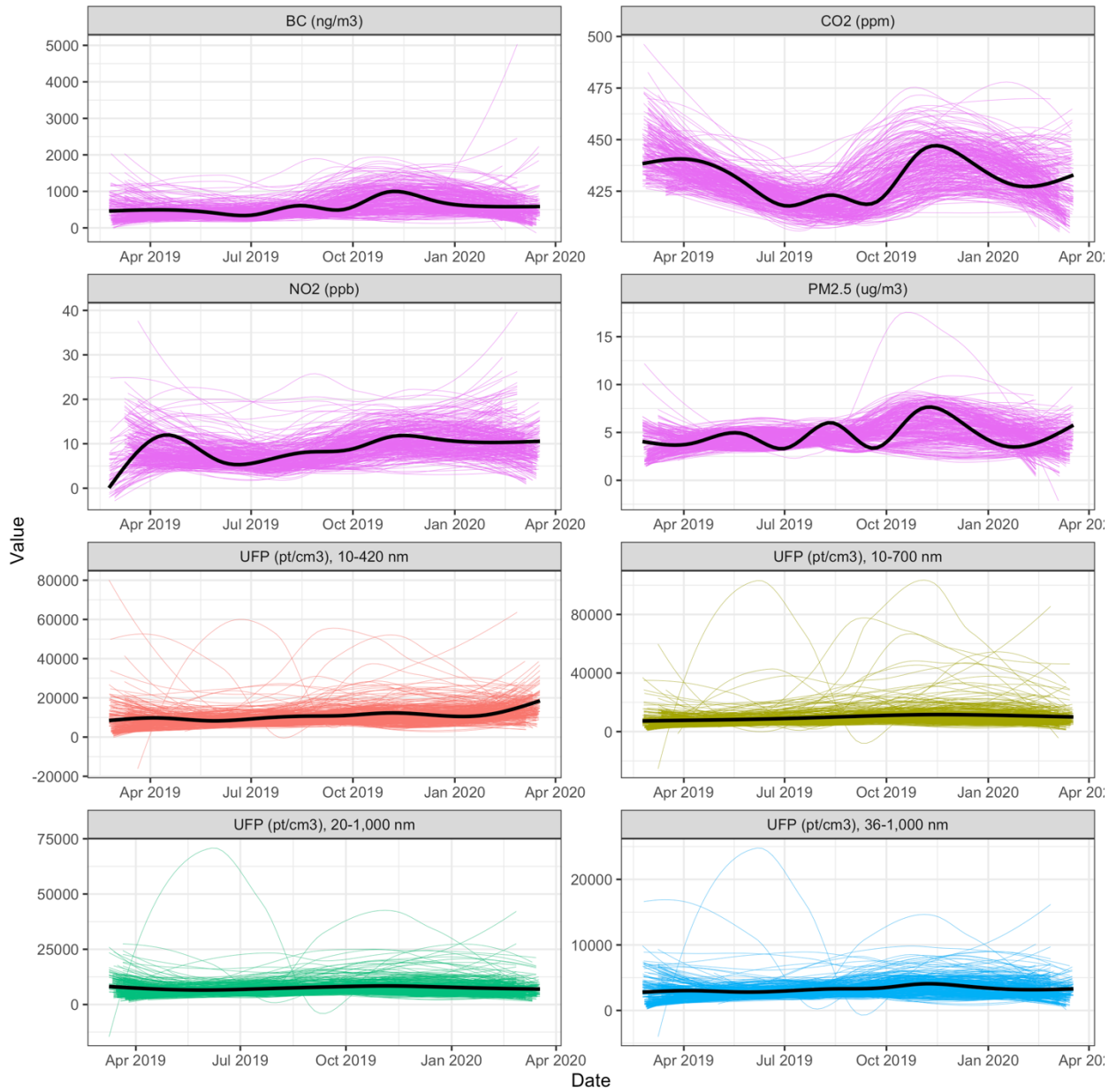


Figure S13. Site-specific pollutant concentrations over the course of the study. Thin lines show site-specific smooth (loess) fits for median visit concentrations ($N \sim 29$ visits/site). Black lines show the overall smooth trends for all of the sites.

Table S9. Percent (%) of variability in visit concentrations explained by spatial and temporal factors.^a

Pollutant	Instrument	Location (n=309)	Day of Week (n=7)	Hour of Day (n=21)	Season (N=4)	Residual Error
BC (ng/m ³)	MA200	4.5%	4.6%	3.1%	7.9%	79.9%
CO ₂ (ppm)	SenseAir	3.0%	1.7%	13.4%	11.6%	70.3%
NO ₂ (ppb)	CAPS	9.7%	6.1%	12.6%	4.5%	67.1%
PM _{2.5} (ug/m ³)	M903	0.2%	1.1%	2.8%	8.9%	87.0%
UFP (pt/cm ³), 10-420 nm	NanoScan	6.9%	3.1%	2.0%	1.8%	86.2%
UFP (pt/cm ³), 10-700 nm	DiSCmini	7.5%	1.6%	0.6%	0.9%	89.4%
UFP (pt/cm ³), 20-1,000 nm	P-TRAK	7.2%	2.4%	0.9%	0.6%	88.9%
UFP (pt/cm ³), 36-1,000 nm	P-TRAK, Screen	5.9%	2.9%	2.9%	1.6%	86.8%

^a Values are based on separate Analysis of Variance (ANOVA) models for each pollutant. The total percent is 100 for each pollutant. The location term represents between site variability across 309 sites. The day, hour and season terms represent temporal variability. The residual error term represents within-site variability across approximately 29 visits per site. The degrees of freedom for this error term varies by pollutant.

2.3 Annual Averages

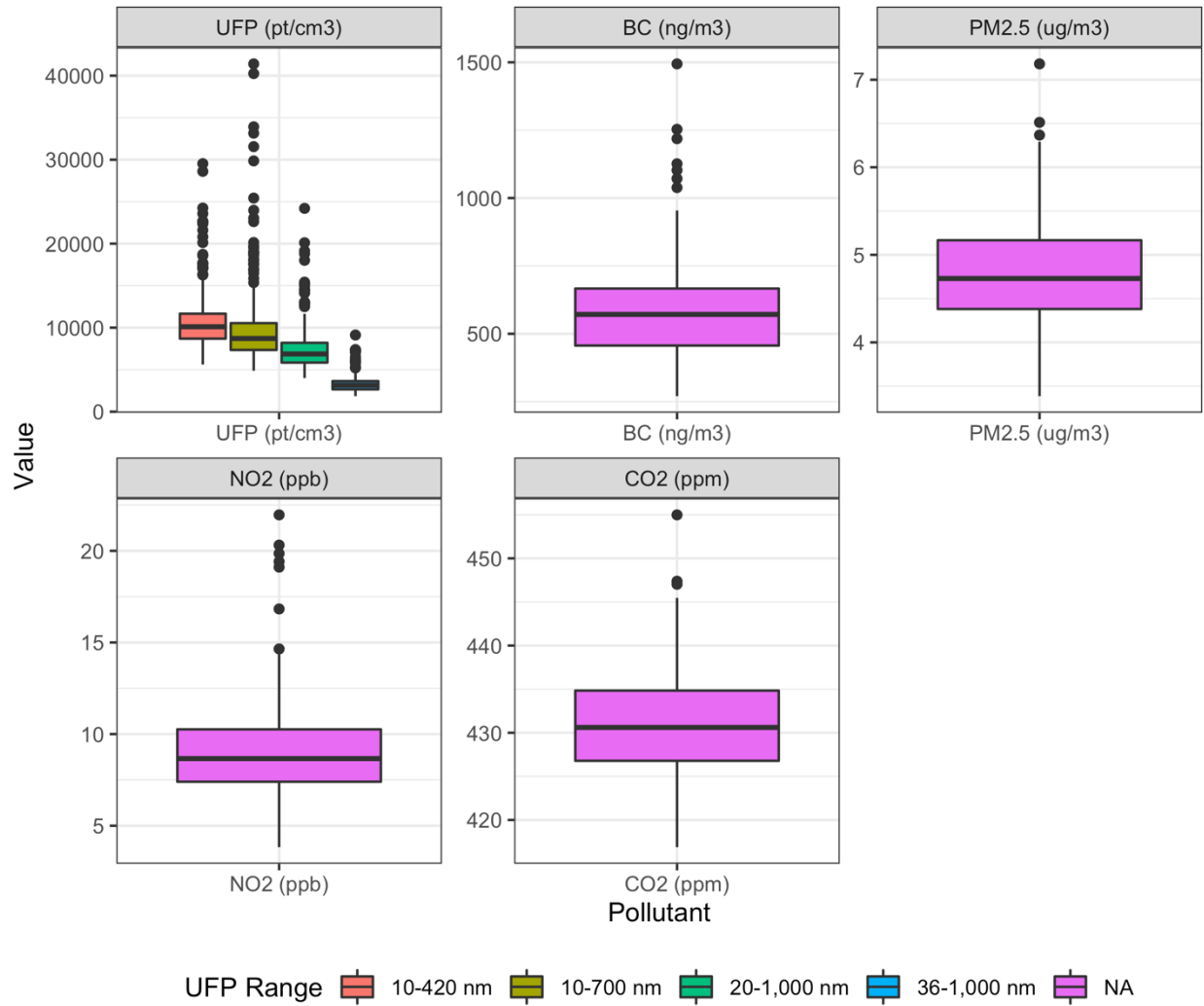


Figure S14. Annual average site concentrations (N=309). Boxplots are traditional and show all of the site averages such that boxes show the 25th, 50th and 75th quantile (Q); whiskers show a 1.5*(Q75-Q25) distance away from the 25th or 75th quantile, and dots are outliers.

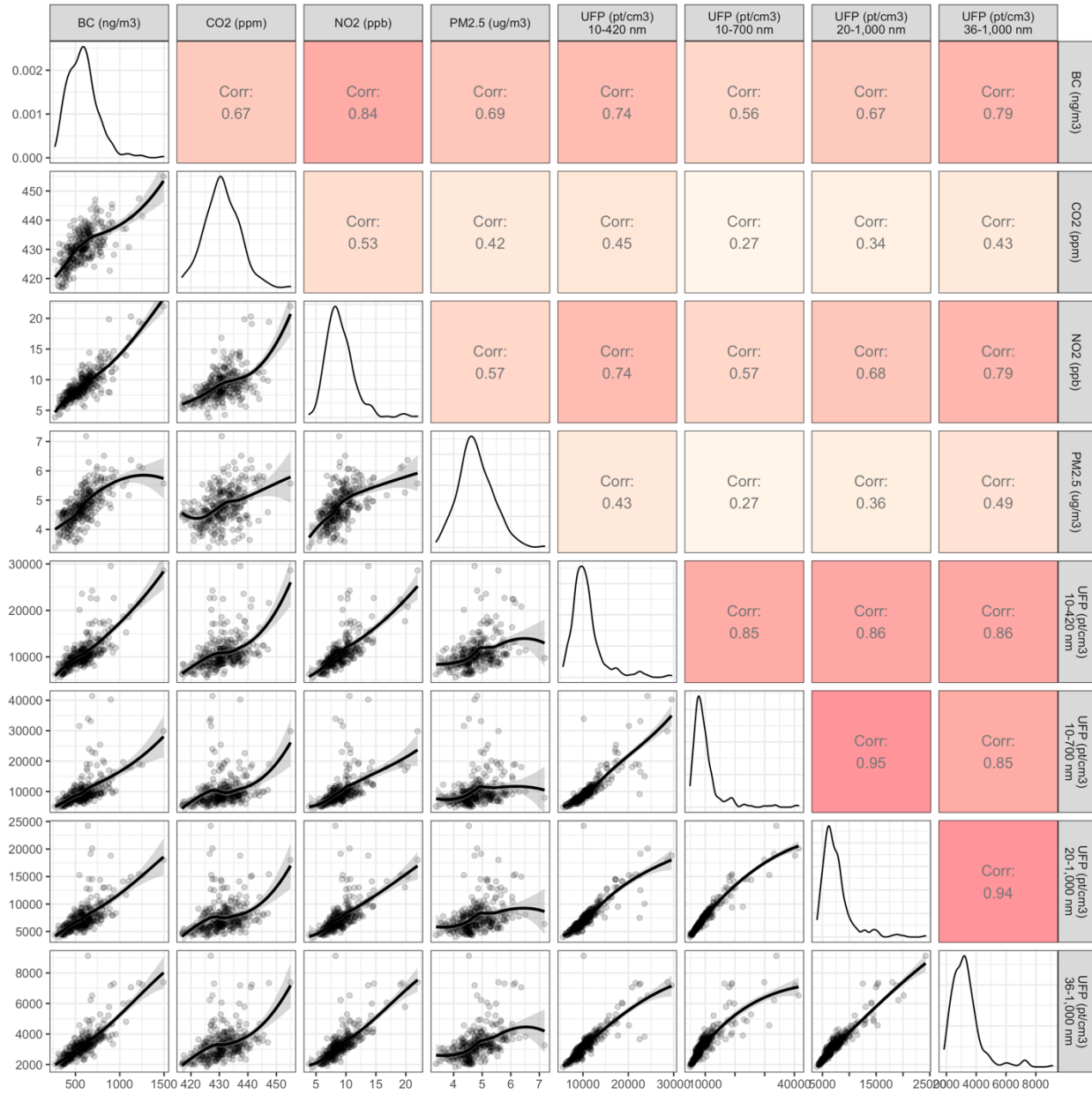


Figure S15. Annual average pollutant correlations ($N=309$ sites). Lower panels show scatterplots with loess lines and 95% confidence intervals; upper panels show Pearson correlations (R), with higher values in darker reds; diagonal panels show density plots.

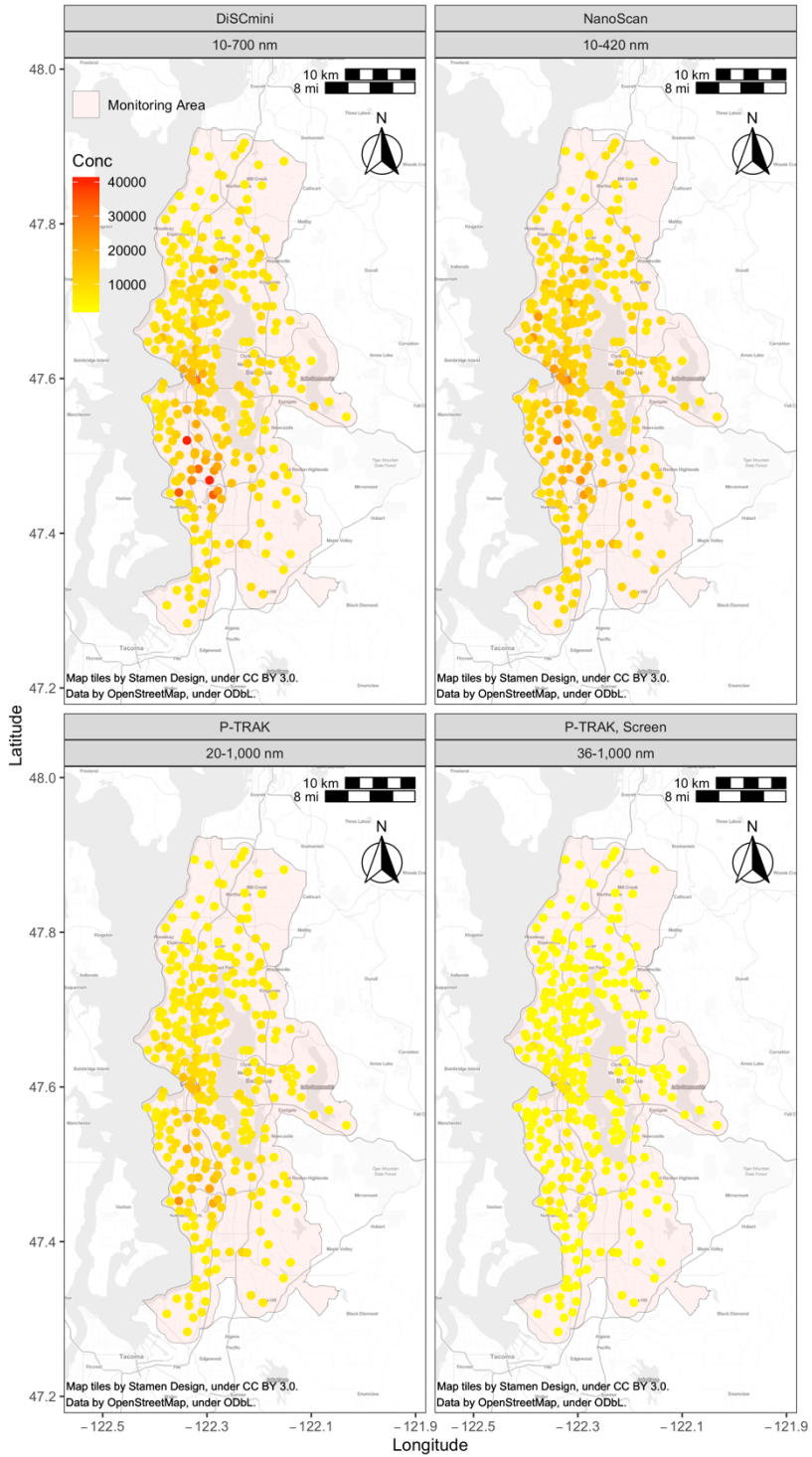


Figure S16. Annual average total UFP concentrations at monitoring sites ($N=309$) for different UFP measures.

Table S10. Available geographic covariates (geocovariates) at monitoring and cohort locations.

	Category	Covariate	Available Buffer Sizes	Description
1	airports	log_m_to_airp	NA	log meters to closest airport
2	airports	log_m_to_l_airp	NA	log meters to closest large airport
3	coast	log_m_to_coast	NA	log meters to closest coastline
4	commercial and services	log_m_to_comm	NA	log meters to closest commercial and services area
5	elevation	elev_above	1000, 5000	number of points (out of 24) more than 20 m and 50 m uphill of a location for a 1000 m and 5000 m buffer, respectively
6	elevation	elev_at_elev	1000, 5000	number of points (out of 24) within 20 m and 50 m of the location' elevation for a 1000 m and 5000 m buffer, respectively
7	elevation	elev_below	1000, 5000	number of points (out of 24) more than 20 m and 50 m downhill of a location for a 1000 m and 5000 m buffer, respectively
8	elevation	elev_elevation	NA	elevation above sea level in meters
9	elevation	elev_stdev	1000, 5000	standard deviation of elevation of 20 points surrounding the location
10	imperviousness	imp_a	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	average imperviousness

11	land use	rlu_barren_p	1000	proportion of barren land
12	land use	rlu_decid_forest_p	300, 400, 500, 750, 1000	proportion of deciduous forest
13	land use	rlu_dev_hi_p	100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of highly developed land (e.g., commercial and services; industrial; transportation, communication and utilities)
14	land use	rlu_dev_lo_p	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of low developed land (e.g., residential)
15	land use	rlu_dev_med_p	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of medium developed land (e.g., residential)
16	land use	rlu_dev_open_p	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of developed open land
17	land use	rlu_evergreen_p	150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of evergreen forest
18	land use	rlu_herb_wetland_p	750, 1000	proportion of herb (nonforested) wetland
19	land use	rlu_mix_forest_p	150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of mixed forest
20	land use	rlu_woody_wetland_p	750, 1000	proportion of woody wetland

21	NDVI	ndvi_q25_a	250, 500, 1000, 2500, 5000, 7500, 10000	NDVI (25th quantile)
22	NDVI	ndvi_q50_a	250, 500, 1000, 2500, 5000, 7500, 10000	NDVI (50th quantile)
23	NDVI	ndvi_q75_a	250, 500, 1000, 2500, 5000, 7500, 10000	NDVI (75th quantile)
24	NDVI	ndvi_summer_a	250, 500, 1000, 2500, 5000, 7500, 10000	average summer time NDVI
25	NDVI	ndvi_winter_a	250, 500, 1000, 2500, 5000, 7500, 10000	average winter time NDVI
26	population	pop10_s	500, 1000, 1500, 2000, 2500, 3000, 5000, 10000, 15000	2010 population density
27	port	log_m_to_l_port	NA	log meters to closest large port
28	port	log_m_to_m_port	NA	log meters to closest medium port
29	railroads, rail yards	log_m_to_rr	NA	log meters to closest railroad
30	railroads, rail yards	log_m_to_ry	NA	log meters to closest rail yard
31	roads	intersect_a1_a3_s	1000, 3000	number of a1-a3 road intersections
32	roads	intersect_a2_a2_s	3000	number of a2-a2 road intersections

33	roads	intersect_a2_a3_s	3000	number of a2-a3 road intersections
34	roads	intersect_a3_a3_s	500, 1000, 3000	number of a3-a3 road intersections
35	roads	ll_a1_s	500, 750, 1000, 1500, 3000, 5000	length of a1 roads
36	roads	ll_a23_s	50, 100, 150, 300, 400, 500, 750, 1000, 1500, 3000, 5000	length of a2 and a3 roads
37	roads	log_m_to_a1	NA	log meters to closest a1 road
38	roads	log_m_to_a1_a1_intersect	NA	log meters to closest a1-a1 road intersection
39	roads	log_m_to_a1_a2_intersect	NA	log_m_to_a1_a2_intersect
40	roads	log_m_to_a1_a3_intersect	NA	log meters to closest a1-a3 road intersection
41	roads	log_m_to_a123	NA	log meters to closest a1, a2 or a3 road
42	roads	log_m_to_a2_a2_intersect	NA	log meters to closest a2-a2 road intersection
43	roads	log_m_to_a2_a3_intersect	NA	log meters to closest a2-a3 road intersection
44	roads	log_m_to_a23	NA	log meters to closest a2 or a3 road
45	roads	log_m_to_a3_a3_intersect	NA	log meters to closest a3-a3 road intersection
46	truck routes	log_m_to_truck	NA	log meters to closest truck route

47	truck routes	tl_s	300, 400, 500, 750, 1000, 1500, 3000, 5000, 10000, 15000	length of truck routes
48	water	log_m_to_waterway	NA	log meters to closest waterway
49	water	rlu_water_p	500, 750, 1000, 3000, 5000	proportion of water

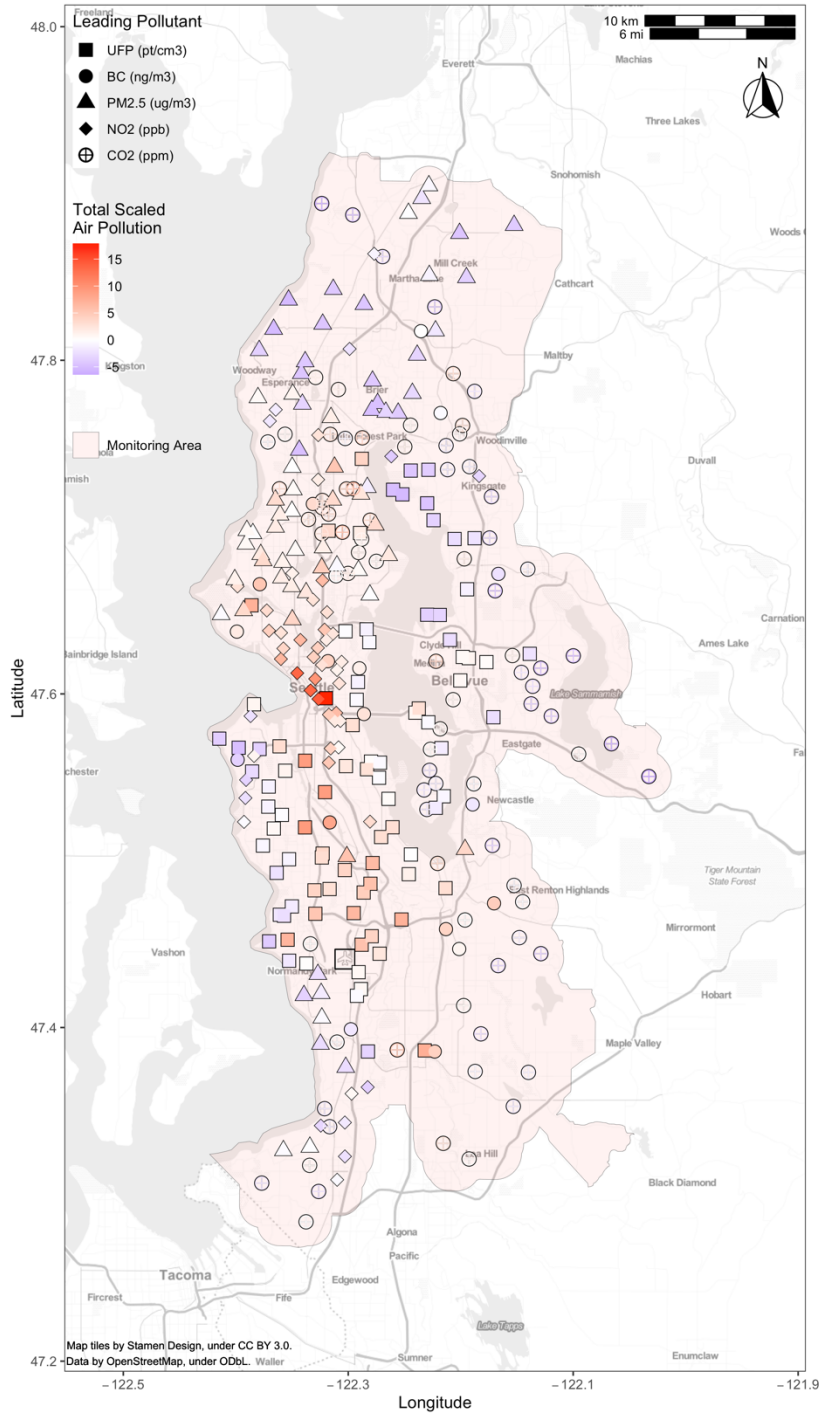


Figure S17. Annual average total scaled air pollution, defined at each site as: $\text{Sum of scores} = \sum_{i=1}^n (x_{i,j} - \text{median}(X_i)) / \text{IQR}(X_i)$, Where X is each individual pollutant ($n=5$) and j indexes each individual pollutant measurement. Colors indicate where air pollution levels were higher or lower than the overall average (total scaled air pollution = 0). Shapes indicate the leading pollutant (highest score) at each site. UFP levels are for particles 20-1,000 nm (P-TRAK).

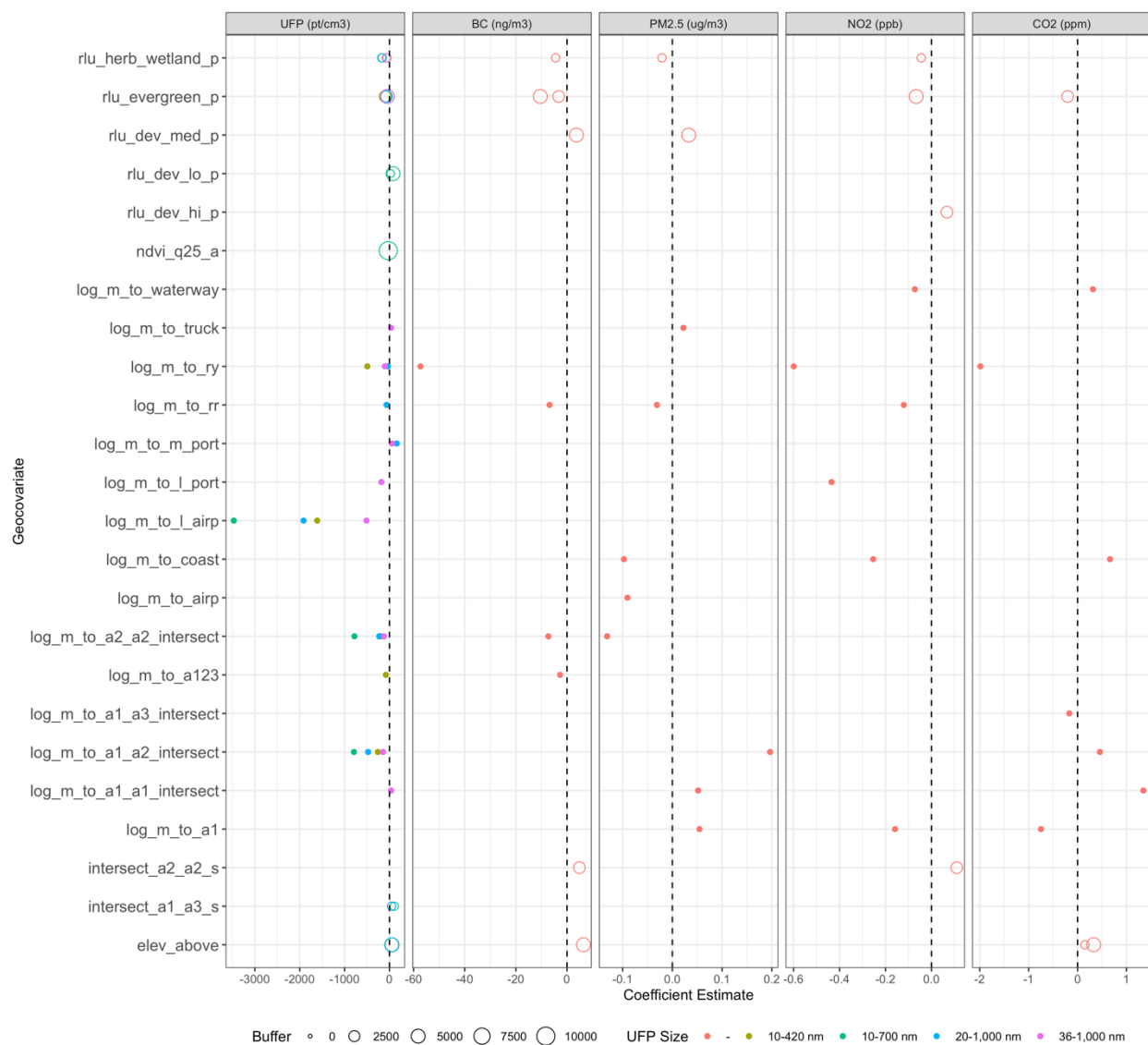


Figure S18. Lasso regression coefficient estimates for annual average pollutant concentrations. Covariates without buffer values (shown in the legend as “0” (m)) are for proximity and elevation variables rather than values for a buffered radius. Showing the top 10 covariates with the largest coefficient estimates for each pollutant. See SI Table S8 for covariate definitions.

2.4 Collocations at Regulatory Monitoring Sites

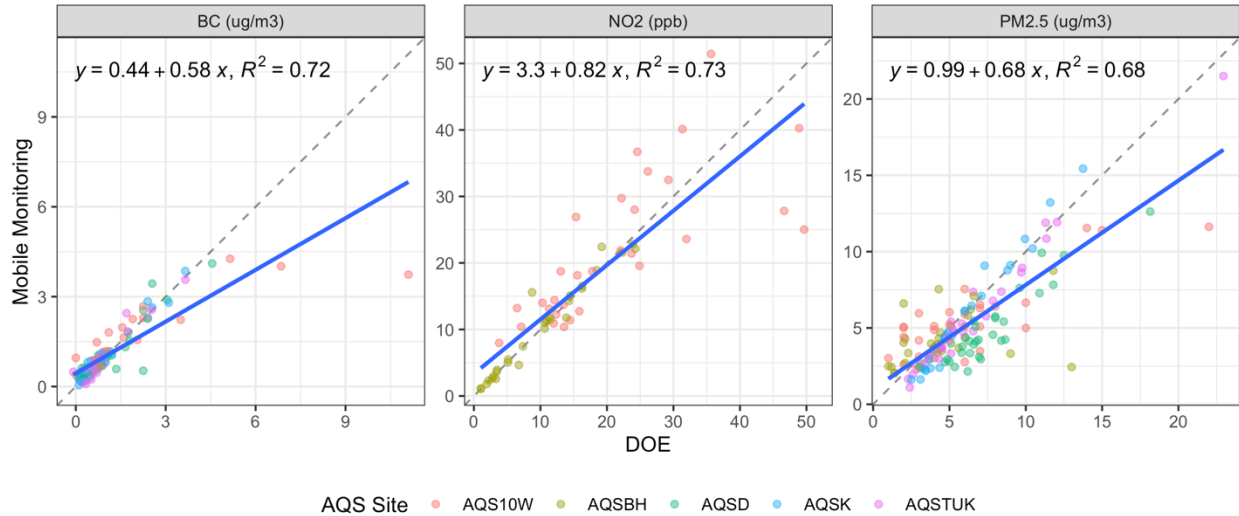


Figure S19. Comparison of two-minute median concentrations from mobile monitoring and the Department of Ecology (DOE) readings at air quality system (AQS) collocation sites. Mobile monitoring $PM_{2.5}$ concentrations are from calibrated nephelometer readings (see Methods). DOE $PM_{2.5}$ concentrations are from nephelometers when available (AQSD, AQSK, AQSTUK), otherwise they are from gravimetric and beta attenuation (BAM) methods, which are updated less frequently (AQS10W – readings are based on rolling 1-hour estimates updated every 6 minutes, AQSBH – readings are updated hourly).

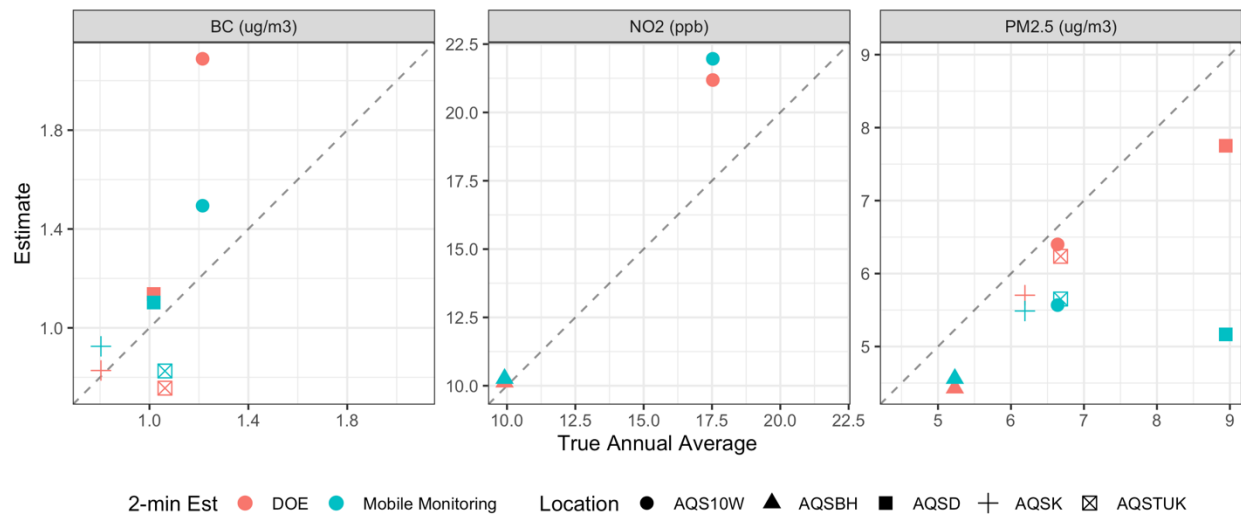


Figure S20. Comparison of annual average estimates from mobile monitoring and the Department of Ecology (DOE) readings at air quality system (AQS) collocation sites. Plots compare estimates using mobile monitoring stop data, DOE data during the same two-minute time periods, and the true annual averages at those sites using all the available regulatory monitoring data for the study period. Mobile monitoring PM_{2.5} concentrations are from calibrated nephelometer readings (see Methods). DOE PM_{2.5} concentrations are from nephelometers when available (AQSD, AQSK, AQSTUK), otherwise they are from gravimetric and beta attenuation (BAM) methods (AQS10W – readings are based on rolling 1-hour estimates updated every 6 minutes, AQSBH – readings are updated hourly).

3 References

1. US Census. TIGER/Line Shapefile, 2017, 2010 nation, U.S., 2010 Census Urban Area National. Published online 2021. Accessed April 29, 2021. <https://catalog.data.gov/dataset/tiger-line-shapefile-2017-2010-nation-u-s-2010-census-urban-area-national>
2. Esri. *ArcGIS Desktop*. Esri; 2019. <https://www.esri.com/about/newsroom/overview/?rmedium=NewsFallback&rsource=blogs.esri.com/Support/blogs/mappingcenter/archive/2010/12/03/using-and-citing-esri-data.aspx>
3. Google. *Google Maps*. Google Inc.; 2019. <https://www.maps.google.com>
4. Good N, Mölter A, Peel JL, Volckens J. An accurate filter loading correction is essential for assessing personal exposure to black carbon using an Aethalometer. *J Expo Sci Environ Epidemiol*. 2017;27(4):409-416. doi:10.1038/jes.2016.71
5. Stampfer O, Austin E, Ganuelas T, Fiander T, Seto E, Karr CJ. Use of low-cost PM monitors and a multi-wavelength aethalometer to characterize PM_{2.5} in the Yakama Nation reservation. *Atmos Environ*. 2020;224:117292. doi:10.1016/j.atmosenv.2020.117292
6. Anderson TL, Ogren JA. Determining aerosol radiative properties using the TSI 3563 integrating nephelometer. *Aerosol Sci Technol*. 1998;29(1):57-69.
7. Hibbert DB. The uncertainty of a result from a linear calibration. *Analyst*. 2006;131(12):1273-1278.
8. Tellinghuisen J. Inverse vs. classical calibration for small data sets. *Fresenius J Anal Chem*. 2000;368(6):585-588. doi:10.1007/s002160000556
9. Austin E, Xiang J, Gould TR, et al. Distinct Ultrafine Particle Profiles Associated with Aircraft and Roadway Traffic. *Environ Sci Technol*. 2021;55(5):2847-2858. doi:10.1021/acs.est.0c05933
10. Oracle Corporation. *MySQL*.; 2021. <https://dev.mysql.com>
11. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Published 2019. <https://www.r-project.org>
12. Robinson D, Hayes A. *Broom: Convert Statistical Analysis Objects into Tidy Tibbles*.; 2020. <https://CRAN.R-project.org/package=broom>
13. Zeileis A, Fisher JC, Hornik K, et al. *Colorspace: A Toolbox for Manipulating and Assessing Colors and Palettes*. arXiv.org E-Print Archive; 2019. <http://arxiv.org/abs/1903.06490>

14. Wilke CO. *Cowplot: Streamlined Plot Theme and Plot Annotations for “Ggplot2.”*; 2019. <https://CRAN.R-project.org/package=cowplot>
15. Wickham H, François R, Henry L, Müller K. *Dplyr: A Grammar of Data Manipulation.*; 2021. <https://CRAN.R-project.org/package=dplyr>
16. Nakazawa M. *Fmsb: Functions for Medical Statistics Book with Some Demographic Data.*; 2021. <https://CRAN.R-project.org/package=fmsb>
17. Wickham H. *Forcats: Tools for Working with Categorical Variables (Factors).*; 2020. <https://CRAN.R-project.org/package=forcats>
18. Schloerke B, Cook D, Larmarange J, et al. *GGally: Extension to “Ggplot2.”*; 2021. <https://CRAN.R-project.org/package=GGally>
19. Kahle D, Wickham H. ggmap: Spatial Visualization with ggplot2. *R J.* 2013;5(1):144-161.
20. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York; 2016. <https://ggplot2.tidyverse.org>
21. Aphalo PJ. *Ggpmisc: Miscellaneous Extensions to “Ggplot2.”*; 2021. <https://CRAN.R-project.org/package=ggpmisc>
22. Aphalo PJ. *Ggpp: Grammar Extensions to “Ggplot2.”*; 2021. <https://CRAN.R-project.org/package=ggpp>
23. Kassambara A. *Ggpubr: “ggplot2” Based Publication Ready Plots.*; 2020. <https://CRAN.R-project.org/package=ggpubr>
24. Slowikowski K. *Ggrepel: Automatically Position Non-Overlapping Text Labels with “Ggplot2.”*; 2019. <https://CRAN.R-project.org/package=ggrepel>
25. Dunnington D. *Ggspatial: Spatial Data Framework for Ggplot2.*; 2020. <https://CRAN.R-project.org/package=ggspatial>
26. Pebesma E, Graeler B. *Gstat: Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation.*; 2021. <https://CRAN.R-project.org/package=gstat>
27. Zhu H. *KableExtra: Construct Complex Table with “kable” and Pipe Syntax.*; 2019. <https://CRAN.R-project.org/package=kableExtra>
28. Xie Y. *Knitr: A General-Purpose Package for Dynamic Report Generation in R.*; 2020. <https://CRAN.R-project.org/package=knitr>
29. Grolemund G, Wickham H. Dates and Times Made Easy with lubridate. *J Stat Softw.* 2011;40(3):1-25.
30. Bache SM, Wickham H. *Magrittr: A Forward-Pipe Operator for R.*; 2014. <https://CRAN.R-project.org/package=magrittr>

31. Henry L, Wickham H. *Purrr: Functional Programming Tools.*; 2019. <https://CRAN.R-project.org/package=purrr>
32. Wickham H, Hester J, Francois R. *Readr: Read Rectangular Text Data.*; 2018. <https://CRAN.R-project.org/package=readr>
33. Pebesma E. *Sf: Simple Features for R.*; 2020. <https://CRAN.R-project.org/package=sf>
34. Bivand R, Nowosad J, Lovelace R. *SpData: Datasets for Spatial Analysis.*; 2021. <https://CRAN.R-project.org/package=spData>
35. Wickham H. *Stringr: Simple, Consistent Wrappers for Common String Operations.*; 2019. <https://CRAN.R-project.org/package=stringr>
36. Müller K, Wickham H. *Tibble: Simple Data Frames.*; 2021. <https://CRAN.R-project.org/package=tibble>
37. Wickham H, Henry L. *Tidyr: Tidy Messy Data.*; 2020. <https://CRAN.R-project.org/package=tidyr>
38. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *J Open Source Softw.* 2019;4(43):1686. doi:10.21105/joss.01686
39. Pebesma E, Mailund T, Kalinowski T. *Units: Measurement Units for R Vectors.*; 2020. <https://CRAN.R-project.org/package=units>
40. Schuetzenmeister A, Dufey F. *VCA: Variance Component Analysis.*; 2019. <https://CRAN.R-project.org/package=VCA>
41. Stamen Design. Published 2021. <http://maps.stamen.com/#terrain/12/37.7707/-122.3783>
42. Creative Commons. Attribution 3.0 Unprotected (CC BY 3.0). Published 2021. <https://creativecommons.org/licenses/by/3.0/>
43. OpenStreetMap contributors. Published 2021. <https://www.openstreetmap.org/copyright>