Environmental Science & Technology

pubs.acs.org/est

Article

Using Vehicles' Rendezvous for In Situ Calibration of Instruments in Fleet Vehicle-Based Air Pollution Mobile Monitoring

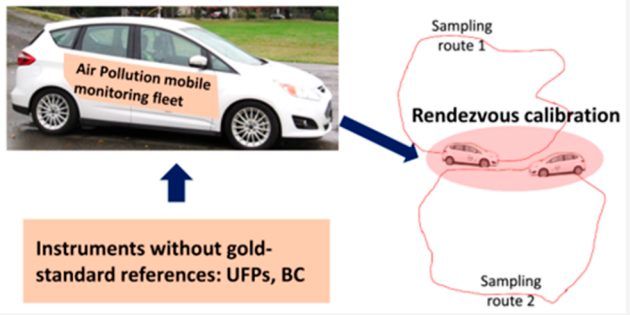
Jianbang Xiang,* Elena Austin, Timothy Gould, Timothy Larson, Michael Yost, Jeffry Shirai, Yisi Liu, Sukyong Yun, and Edmund Seto

Cite This: *Environ. Sci. Technol.* 2020, 54, 4286−4294

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: This study examines the feasibility of the in situ calibration of instruments for fleet vehicle-based mobile monitoring of ultrafine particles (UFPs) and black carbon (BC) by comparing rendezvous vehicle measurements. Two vehicles with identical makes and models of UFP and BC monitors as well as GPS receivers were sampled within 140 m of each other for 2 h in total during winter in Seattle, Washington. To identify an optimal intervehicle distance for rendezvous calibration, 6 different buffers within 0−140 m for UFP monitors and 5 different buffers within 0−90 m for BC monitors were chosen, and the results of calibration were compared against a reference scenario, which consisted of mobile colocation measurements with both sets of the UFP and BC monitors deployed in one of the vehicles. Results indicate that the optimal distances for rendezvous calibration are 10−80 m for UFP monitors and 0−30 m for BC monitors. In comparison with the mobile colocation calibration, the rendezvous calibration shows a normalized root mean squared deviation of 6−14% and a normalized mean absolute deviation of 4−8% for these monitors. Criteria for applying a rendezvous calibration approach are presented, and an extension of this approach to an instrumented fleet of mobile monitoring vehicles is discussed.



1. INTRODUCTION

The recent use of Google Street View[1,2] and other private/public mobile platforms (e.g., buses, cars, pedestrians)[3−11] for mobile air quality monitoring has prompted new opportunities and challenges with respect to large scale mobile measurements of environmental quality using all sorts of fleet vehicles. Notable among the challenges is the need to maintain data quality and measurement accuracy among large numbers of instruments in the field. Specifically, quantitative calibration of such instruments is of great importance in maintaining data quality.

Previous studies have evaluated the quantitative performance of air pollution monitors under controlled laboratory conditions[12−16] or ambient fixed-site deployment.[14,17−19] However, neither laboratory studies nor one-off static colocations adjacent to an off-road reference analyzer represent the likely typical in situ usage of these instruments for mobile measurements. While "calibration" typically refers to the adjustment of instrument settings or output to produce measurements that are consistent with a gold standard reference, which is possible for many gaseous pollutants, the notion of calibration is different for ambient particles, which vary in composition and concentrations across a sampling study domain. These variations are challenging to replicate in the laboratory. Consequently, there are no gold-standard reference instruments for some air pollutants, such as ultrafine particles (UFPs; particles with aerodynamic diameter <100 nm) and black carbon (BC; one of the major atmospheric aerosol pollutants from incomplete combustion of fossil fuel or biomass), which operate on optical measurement principles, other than comparisons against similar high-end instruments. Therefore, comparisons cannot be made with reference analyzers, and one is left with the averaged measured results of multiple colocated similar monitors. In situ mobile colocation with multiple sets of instruments deployed in the same vehicle seems to be a feasible way when there are only a few monitors. However, it quickly becomes logistically challenging and impractical as the numbers of vehicles, pollutants of interest, and instruments increase with mobile monitoring fleet deployments. For mobile measurements with many vehicles, vehicles may come in proximity to each other either by design of the sampling approach or by chance for large-scale unorganized sampling. When vehicles approach

Received: January 30, 2020
Revised: March 6, 2020
Accepted: March 9, 2020
Published: March 9, 2020

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

each other, these "rendezvous" events allow for comparisons between the measured concentrations of each air pollutant from each platform and thus, a calibration approach: "rendezvous calibration". However, it is unclear how well the measurements from these fleet vehicles agree with each other, especially for pollutants (e.g., UFPs and BC) which have shown high spatial variability within a small region.[1,20,21] There remain two essential questions: (a) how best can one optimize the "rendezvous", the optimal distance between measurement platforms and (b) how accurate is this calibration compared to a reference colocation method?

The present study leveraged the Mobile ObserVations of Ultrafine Particles (MOV-UP) Study, which aims to study the impacts of air traffic on air quality (e.g., UFPs and BC) for communities located near and below the flight paths of the Seattle-Tacoma (SeaTac) International Airport in the Seattle region of Washington, USA.[22] Using the MOV-UP Study data set, the present study aims to develop an in situ rendezvous calibration approach for instruments in fleet vehicle-based mobile monitoring of UFPs and BC and examines the feasibility and accuracy of this rendezvous calibration approach.

2. METHODS

Figure S1 presents an overview of the methodology of the present study. The optimal distances for rendezvous calibration for UFP and BC monitors were identified through a detailed multiple-buffer analysis. With the optimal distances determined, the accuracy of rendezvous calibration was then examined by comparing against a conventional colocation calibration approach. The sensitivity of the results to the response characteristics of the monitors was examined.

2.1. Mobile Monitoring Measurements. As described above, the present study leveraged part of the MOV-UP Study data set that was conducted from February 7, 2018 to January 11, 2019. Typically, the two vehicles sampled for 5 h from 12:00 to 17:00 on different routes: either on five transects that cross the flight paths north or on six transects south of the SeaTac airport. To examine the feasibility of the rendezvous calibration, the measurements were examined when these two vehicles were driven close to each other (within a buffer of 140 m) by design while sampling for 2 h in total during February 14−16, 2018. This three-day data set was named as the "rendezvous dataset".

A detailed description of the mobile platform is given elsewhere.[7,11] In brief, each mobile monitoring platform included a Toyota Prius gasoline-electric hybrid vehicle from University of Washington Fleet and several portable instruments for air pollution measurements (Figure S2). A GPS was mounted in the vehicle to record its position and speed. A sampling inlet was mounted on the roof on the driver's side of the vehicle in a forward position, with the sampling inlet positioned above the vehicle boundary layer, the zone of laminar flow directly associated with vehicle motion, and connecting tubes entering the vehicle through the otherwise sealed left rear window where they were connected to a manifold, with lines from the manifold then connected to the instruments. Particle loss was minimized by using stainless steel, copper, and conductive flexible tubing for the particle sampling inlet and tubing. The exhaust pipe from the vehicle's gasoline engine discharged low to the ground, away from the elevated, left-side air monitoring inlet. To further minimize the potential for self-pollution, the vehicle's gasoline engine was typically shut off when stopped at red traffic lights.

Throughout the campaign, UFP number concentrations were monitored with a Condensation Particle Counter (CPC, model 3007, TSI Inc., MN) and two P-Trak (Model 8525, TSI Inc., MN) condensation nuclei particle counters (one with an inlet diffusion screen) at 1-s intervals for each mobile monitoring platform. The CPC and P-Trak measured the total number concentrations of particles larger than 10 nm diameter and 20 nm diameter, respectively; while the P-Trak with the inlet diffusion screen used in this study measured the total number concentrations of particles larger than 36 nm diameter. All P-Trak, P-Trak screened, and CPC monitors are categorized as UFP monitors. All UFP monitors were factory calibrated by TSI before they were delivered. BC mass concentrations were monitored with a BC monitor (microAeth AE51, AethLabs, CA) at 10-s intervals throughout the three-day campaign. The position (longitude, latitude, and altitude) and speed of each vehicle were recorded at 1-s intervals with a GPS Receiver (BU-353-S4, GlobalSat, TW) that has an accuracy of ±5 m. Instrument quality assurance/quality control (QA/QC) objectives and evaluation methods have been described in detail for mobile monitoring generally elsewhere[7,11] and more specifically for our study as summarized in Table S1 (Supporting Information).

In addition to these mobile monitoring measurements with the two mobile platforms as described above, 4 days of mobile colocation measurements were conducted between March 6, 2018 and August 30, 2018 with pairwise UFP and BC monitors deployed in one vehicle to act as the reference calibration approach: "co-location calibration". Both pairwise monitors in the vehicle were sampled from the same sampling inlet as described above. Throughout the colocation mobile measurement campaign, all UFP and BC monitors as well as GPS receivers recorded data at 1-s intervals. The colocation measurement lasted about 1 h for the UFP monitors and 10 h in total for the BC monitors (Table S2), and occurred on the same routes as the other mobile measurements. This four-day data set was named as the "co-location dataset".

2.2. Rendezvous Calibration Analysis. The BC data were corrected using an Optimized Noise-reduction Averaging (ONA) algorithm developed by US EPA (with attenuation coefficient (ATN) threshold set to ΔATN = 0.05) to reduce potential instrumental optical and electronic noise.[23] After correcting the BC data, 10-s median BC concentrations were calculated for the colocation data set since the BC data for the rendezvous data set were 10-s averages and the sensitivity of the AE51 monitor in the 1-s operating mode is relatively low.[24] Thus, 10-s BC data and 1-s UFP data were used for the following analysis.

The distance between the two vehicles can be calculated based on their GPS information, as shown in eq 1

$$Distance(t) = \sqrt{Dis_{horizontal}(t)^2 + (Alti_{veh1}(t) - Alti_{veh2}(t))^2} \quad (1)$$

where $Dis_{horizontal}(t)$ is the horizontal distance between the two vehicles determined by their longitude and latitude at the time t using the Haversine formula with "geosphere" package in R;[25] $Alti_{veh1}(t)$ and $Alti_{veh2}(t)$ are the altitudes above sea level of the two vehicles at time t, respectively. To merge with the UFP and BC data for the rendezvous data set, we calculated 1-s distances for UFP monitors and 10-s distances based on 10-s median location data (longitudes, latitudes, and altitudes) for BC monitors. Note that no time rolling averaging smoothing

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

approach was applied to the pollutant and intervehicle distance data.

To quantify the effect of intervehicle distance on the performance of rendezvous calibration, the data set was segmented based on the intervehicle distance buffers, i.e., 6 different buffers (0−10, 10−20, 20−40, 40−60, 60−80, 80−140 m) for UFP monitors and 5 different buffers (0−10, 10−20, 20−30, 30−50, 50−90 m) for BC monitors. The buffers were chosen by balancing the sizes of each buffer-specific segmented data set as shown in Table S2. As there were some BC data missing for the rendezvous data set, different buffers were selected for the UFP and BC monitors. For the data within each buffer, a linear regression model (eq 2) was fitted in which the dependent variable was the mean of the two pairwise monitors, and the independent variable was one of the vehicle's instrument measures (separate models for each vehicle). A bootstrap analysis was applied for each buffer with 1000 replicates of 100 and 30 randomly sampled pairwise observations in each pseudodata set (about one tenth of total observations) to get normal distributions of the calibration coefficients for the UFP and BC monitors, respectively. The bootstrap approach is recommended when the coefficients are estimated from the same input data set.[26,27] Outliers were diagnosed by using Cook's distance value. The observations with a Cook's distance value greater than 0.5 were excluded from the regression analysis.[28,29] A similar analysis using the bootstrap (1000 replicates of 300 randomly sampled pairwise observations) and linear regression model was conducted for the mobile colocation data set to determine the reference calibration coefficients for each of those monitors.

$$C_{ref} = \beta_0 + \beta_1 C_i \tag{2}$$

where C_{ref} is the reference concentration, the mean of the two monitors; C_i (i = 1, 2) is the measured concentration of Monitor i; β_0 and β_1 are the calibration coefficients from the linear regression model.

The central estimates (means) of β_0 and β_1 determined from the bootstrap analyses together with linear regression analyses were taken as the best estimates of the calibration coefficients for each monitor. To test the accuracy of the rendezvous calibration approach, the calibration coefficients for each monitor from rendezvous calibration and colocation calibration were applied to calibrate the colocation data set as shown in eq 3. Then, the postcalibrated concentrations based on the rendezvous calibration were compared with those based on the colocation calibration as the colocation calibration was taken as a more accurate calibration approach. The mean absolute deviation (MAD), normalized mean absolute deviation (NMAD), root mean squared deviation (RMSD), and normalized root mean squared deviation (NRMSD) between the rendezvous-calibrated concentrations (C_{cal_ren}) and colocation-calibrated concentrations (C_{cal_colo}) were calculated based on eqs 4−7, respectively. Although MAD and RMSD are both commonly used to measure forecast accuracy, RMSD is a metric for measuring forecast accuracy based on means while MAD is a metric based on medians.[30] As MAD and RMSD are scale-dependent and the concentration units of UFPs and BC are different in the present study, NMAD and NRMSD were calculated which were scaled to the mean of the colocation-calibrated concentrations. In the present study, NRMSD was taken as the primary metric to determine the accuracy of rendezvous calibration approach compared with colocation calibration while the other three statistical metrics provide additional valuable information.

$$C_{cal} = \beta_0 + \beta_1 C_{raw} \tag{3}$$

where C_{raw} and C_{cal} are the pre- and postcalibrated concentrations for the colocation data set; β_0 and β_1 are the central-estimated calibration coefficients as described above, separate for the monitors and calibration scenarios.

$$\mathrm{MAD} = \frac{\sum_{t=1}^{N} |C_{cal\_ren,t} - C_{cal\_colo,t}|}{N} \tag{4}$$

$$\mathrm{NMAD} = \frac{\mathrm{MAD}}{\sum_{t=1}^{N} C_{cal\_colo,t}/N} \times 100\% \tag{5}$$

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{t=1}^{N} (C_{cal\_ren,t} - C_{cal\_colo,t})^2}{N}} \tag{6}$$

$$\mathrm{NRMSD} = \frac{\mathrm{RMSD}}{\sum_{t=1}^{N} C_{cal\_colo,t}/N} \times 100\% \tag{7}$$

where the subscript t represents the sequence number of the observation; N is the number of all observations for each monitor for the colocation data set.

Optimal distance was introduced to maximize the variability of the sampling data and defined as the pooled buffers with the smallest NRMSD in the buffer-based rendezvous calibration scenarios. With the optimal distance for each monitor determined, all the pairwise data within the optimal distance for each monitor were pooled. Similar bootstrap and linear regression analyses were then conducted to determine the optimal-distance-based calibration coefficients for all monitors. Similarly, the accuracy of the optimal-distance-based rendezvous calibration compared against the colocation calibration was examined based on MAD, NMAD, RMSD, and NRMSD.

2.3. Sensitivity Analysis. The UFP and BC monitors were manufacturer-calibrated against higher-class reference instruments prior to the mobile monitoring as shown in Table S1. However, the response characteristics from some particle and BC monitors can change over time or when used with accessories such as diffusion screens that may degrade over time.[31−34] The sensitivity of the results to the response characteristics of the monitors were examined. Simulations were conducted for five additional scenarios in which one of the pairwise monitors measured (i) 5%, (ii) 10%, (iii) 15%, (iv) 20%, and (v) 25% higher than those in the original scenario. A monitor with a large change of response characteristic (e.g., 30%) might be taken as a malfunctioning monitor. This will be further discussed in the Discussion section. In the original scenario, Monitor_1 for the four pairs of monitors generally measured higher concentrations than Monitor_2 (Monitor_1 and Monitor_2 were determined by their serial numbers which were constants in this study). Thus, Monitor_1 was assumed to measure higher to further increase the differences between the two monitors. Similar bootstrap and linear regression analyses were then conducted to determine the optimal distances and the accuracy of rendezvous calibration against colocation calibration for each monitor under each simulated scenario using the approach described above (see Figure S1 for more details).

All calculations and figures were made using "data.table", "geosphere", and "ggplot2" packages in R, Version 3.3.0

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

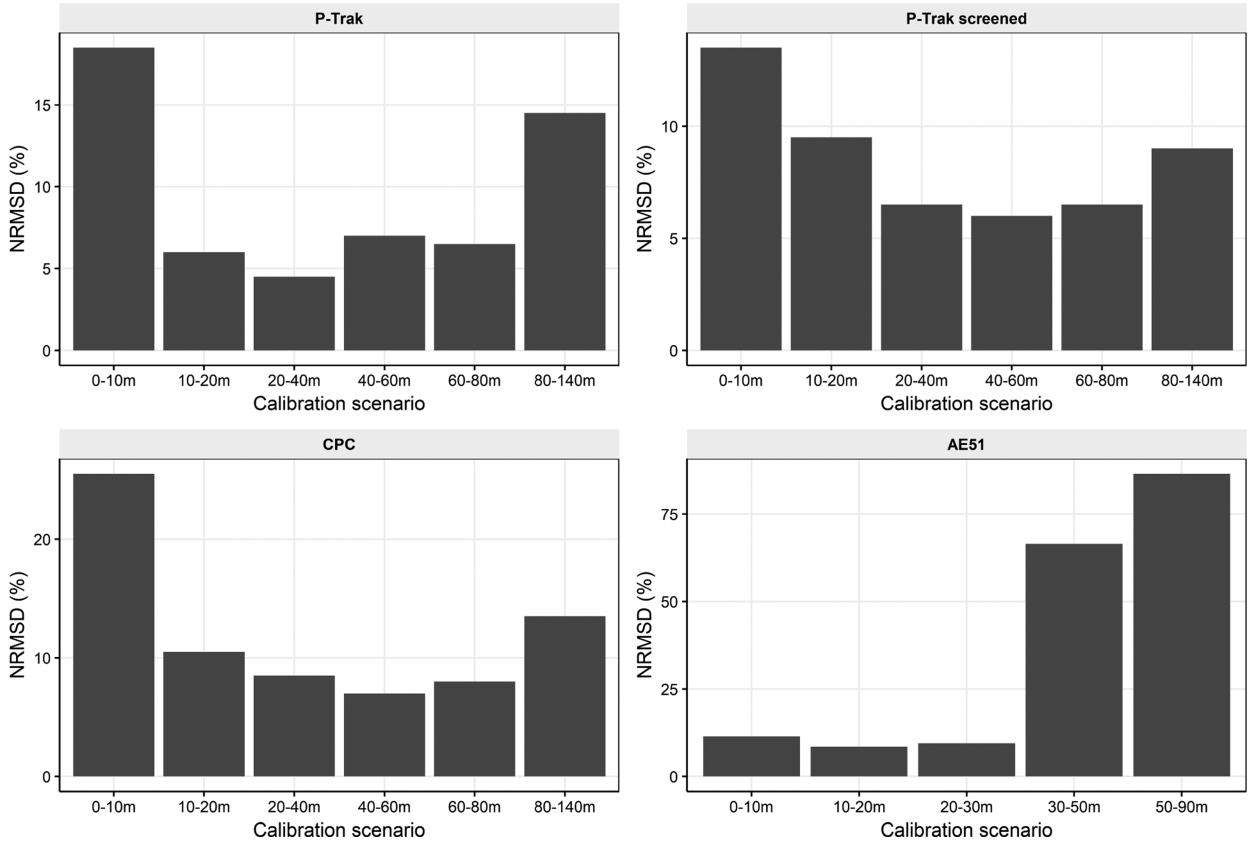


Figure 1. Averaged normalized root mean squared deviation (NRMSD) between the rendezvous-calibrated concentrations and colocation-calibrated concentrations for the colocation data set with different calibration scenarios. "0−10 m", "10−20 m", and other distance buffers represent rendezvous calibration with corresponding buffer-based segmented data sets. Results are averaged for the pairwise monitors.

embedded in RStudio Version 1.1.456. In the Results section, summed calculations are presented as the average of the two pairwise monitors. For each monitor (Monitor_1 and Monitor_2), specific calculations are presented in Tables S3−S7.

3. RESULTS

3.1. Optimal Distance Determination. Table S2 summarizes the statistic descriptions of pollutant levels and pairwise data observation numbers as well as sampling time for colocation and rendezvous calibration. The numbers of observations for each buffer-based segmented data set are over 700 for UFP monitors and over 70 for BC monitors, while that for colocation data set are over 3000. The numbers of observations were translated into a sampling time of 0.2−0.5 h for buffer-based segmented data sets and 0.9−10 h for colocation data set. In general, the UFP levels in the rendezvous data set are higher than those in the colocation data set. For instance, with intervehicle distances ≤140 m, the UFP number concentrations measured from CPC ranged from 2230 to 42450 particles/cm^3 with a median of 12270 particles/cm^3 for the colocation data set and from 2180 to 232640 particles/cm^3 with a median of 20130 particles/cm^3 for the entire rendezvous data set. In contrast, the BC mass concentrations in these two data sets are comparable, ranging from 40 to 13620 ng/m^3 with a median of 850 ng/m^3 for the colocation data set and from 220 to 12570 ng/m^3 with a median of 650 ng/m^3 for the entire rendezvous data set. The pollutant levels among different buffer-based segmented data sets are generally comparable, especially within a distance buffer of 10−80 m for UFPs and 10−50 m for BC.

Figure 1 shows the averaged NRMSD between the rendezvous-calibrated concentrations and colocation-calibrated concentrations for the colocation data set with different calibration scenarios for each type of instrument. Among all the rendezvous-calibration scenarios for UFP monitors, the NRMSDs for the scenario of a 0−10 m buffer are the largest with a range of 13.5−25.5%, while those for scenarios of 10−20 m, 20−40 m, 40−60 m, and 60−80 m buffers are the smallest with a range of 4.5−10.5%. The NRMSDs tend to increase notably when the intervehicle distance buffers are greater than 80 m, especially for P-Trak and CPC monitors, with a range of 13.5−14.5%. As for BC monitors, the NRMSDs for the scenarios of 0−10 m, 10−20 m and 20−30 m buffers are the smallest with a range of 8.5−11.5%. For those buffers greater than 30 m, the NRMSDs increase to 66.5−86.5%. On the basis of the definition of optimal distance, the optimal distances of rendezvous calibration for UFP and BC monitors are 10−80 m and 0−30 m, respectively.

In addition to the NRMSD, Table S3 shows the MAD, NMAD, and RMSD between the rendezvous-calibrated concentrations and colocation-calibrated concentrations for the colocation data set with different calibration scenarios for each monitor. The patterns of the MAD, NMAD, and RMSD among all the calibration scenarios are similar to those of the NRMSD. For instance, among all the rendezvous-calibration scenarios for P-Trak Monitor_1, the MAD for the scenario of a 0−10 m buffer is the largest (2304 particles/cm^3), followed by the scenario of a 80−140 m buffer (2007 particles/cm^3), while

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

those for scenarios of 10−20 m, 20−40 m, 40−60 m, and 60−80 m buffers are the smallest with a range of 361−937 particles/cm^3. As for AE51 Monitor_2, the MAD for the scenario of 30−50 m and 50−90 m buffers are the largest with a range of 458−539 ng/m^3, while those for scenarios of 0−10 m, 10−20 m, and 20−30 m buffers are the smallest with a range of 101−118 ng/m^3. Thus, the results for the MAD, NMAD, and RMSD also support the determination of the optimal distances: 10−80 m for UFP monitors and 0−30 m for BC monitors. Additionally, Table S3 shows that the results are slightly different between the pairwise monitors. However, the patterns of all these statistical metrics among all the calibration scenarios are similar for the two monitors.

3.2. Accuracy of Optimal-Distance-Based Rendezvous Calibration. The results of the linear regression models for colocation and optimal-distance-based rendezvous calibration are summarized in Table S4. The coefficient of determination (R^2) for the colocation and rendezvous calibration ranges from 0.96 to 0.99 and 0.95 to 0.98, respectively. The larger R^2 for the colocation calibration compared against rendezvous calibration indicates that the former is generally more accurate, which agrees with the fact that the colocation calibration was taken as a reference approach to determine the accuracy of the rendezvous calibration. With the measured concentration for the colocation data set calibrated with the colocation and optimal-distance-based calibration models, Figure S3 shows the scatter plot of the pre- and postcalibrated measured concentrations against the reference concentrations.

Table 1 presents the MAD, NMAD, RMSD, NRMSD between the optimal-distance-based rendezvous-calibrated and colocation-calibrated concentrations for the colocation data set. The NRMSDs for P-Trak, P-Trak screened, CPC, and AE51 monitors are 6%, 7%, 10%, and 14%, respectively, which are comparable to those based on buffer-specific rendezvous calibration shown in Figure 1. Similarly, the NMADs, with a range of 4−8%, are comparable to those based on buffer-specific rendezvous calibration shown in Table S3. The relatively small NRMSDs together with NMADs indicate a high accuracy of rendezvous calibration by taking both the mean and median concentrations as the optimization objective. Additionally, the MADs and RMSDs are <1500 particles/cm^3 for UFP monitors and <180 ng/m^3 for BC monitors. The results for each monitor (Monitor_1 and Monitor_2) are shown in Table S5.

Table 1. Statistical Description of the Accuracy of Optimal-Distance-Based Rendezvous Calibration Compared with Co-Location Calibration[a]

Instrument	MAD[b]	NMAD (%)	RMSD[b]	NRMSD (%)
P-Trak	497	4	645	6
P-Trak screened	211	5	340	7
CPC	1259	8	1495	10
AE51	95	8	173	14

[a]Abbreviations: MAD (mean absolute deviation), NMAD (normalized mean absolute deviation), RMSD (root mean squared deviation), and NRMSD (normalized root mean squared deviation). [b]Unit: particles/cm^3 for P-Trak, P-Trak screened, and CPC monitors; ng/m^3 for AE51 monitors.

3.3. Sensitivity Analysis. As described in the Methods, to evaluate the sensitivity of the results to the response characteristics of the monitors, the optimal distances and the accuracy of rendezvous calibration were evaluated for five additional scenarios in which Monitor_1 measured 5%, 10%, 15%, 20%, and 25% higher than those in the original scenario. As can be seen in Figure 2, although the differences of the NRMSD among all the calibration buffers become smaller as the response changes of Monitor_1 increase from 0% to 25%, the optimal distances generally remain 10−80 m for UFP monitors and 0−30 m for BC monitors for the five response-change scenarios. Table S6 presents the MAD, NMAD, RMSD, NRMSD between the optimal-distance-based rendezvous-calibrated and colocation-calibrated concentrations for five response-change scenarios. Generally, the changes of those statistical metrics are <10% for each monitor with the response of Monitor_1 increasing by 25%.

4. DISCUSSION

Previous studies have examined the calibration of air pollution monitors under controlled laboratory conditions,[12−16] ambient fixed-site atmosphere,[14,17−19] or in situ mobile colocation deployment.[35] However, the present study is the first to systematically examine the feasibility of in situ calibration of instruments for fleet vehicle-based mobile monitoring of UFPs and BC by comparing nearby vehicle measurements (rendezvous). This study reveals that the optimal intervehicle distances of rendezvous calibration for UFP and BC monitors are 10−80 m and 0−30 m, respectively. In comparison with a referenced mobile colocation calibration approach, the rendezvous calibration shows a NRMSD of 6−14% and a NMAD of 4−8% for these UFP and BC monitors. This rendezvous calibration approach is mainly developed for situations where the mobile colocation calibration approach becomes logistically infeasible, as the numbers of vehicles, pollutants of interest, and instruments increase with mobile monitoring fleet deployments.

As mentioned in the Results section, among all the rendezvous-calibration scenarios, the NRMSDs for the scenario of a 0−10 m buffer for the UFP monitors are notably larger than other scenarios within the optimal distances. This can be explained by the fact that the order of the vehicles was not randomized and one vehicle was consistently the lead during the mobile monitoring during the period of February 14−16, 2018. This makes the measured concentrations of UFPs from the following vehicle notably higher than those from the front vehicle when they were very close (<10m) due to the impacts of the exhaust of the front vehicle. According to previous studies, both heavy-duty diesel trucks and light-duty vehicles are the major source of on-road UFPs while the former is the major source of on-road BC.[36−38] This explains why the NRMSDs for the scenario of 0−10 m buffer compared with other scenarios within the optimal distances for the BC monitors are not as large as those for the UFP monitors. Although the corresponding NRMSDs for the scenario of a 0−10 m buffer for the UFP monitors might be smaller if the orders of the two vehicles were randomized when they were within 10 m, the situation where the following vehicle measured pollutants in the front vehicle's exhaust probably remained. Hence, excluding the data within the 0−10 m buffer is reasonable for optimal-distance rendezvous calibration. In addition to the difference for the 0−10 m buffer, the optimal distance for BC monitors is much closer than those for the UFP monitors in this study (0−30 m vs 10−80 m). This may be explained by the fact that diesel trucks generally account for

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

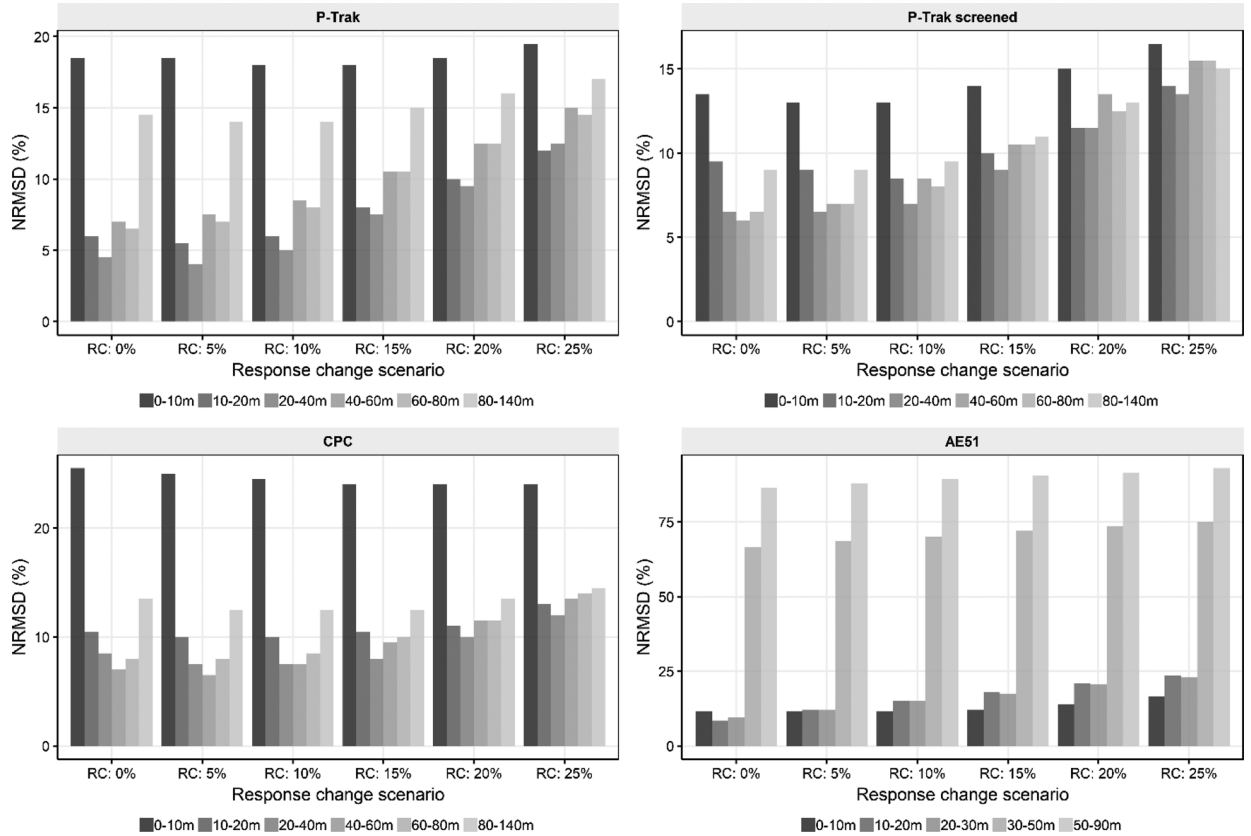


Figure 2. Averaged normalized root mean squared deviation (NRMSD) between the rendezvous-calibrated concentrations and colocation-calibrated concentrations with different calibration buffers under five additional scenarios in which Monitor_1 measured 5%, 10%, 15%, 20%, and 25% higher than those in the original scenario. "0−10 m", "10−20 m", and other distance buffers represent rendezvous calibration with corresponding buffer-based segmented data sets. "RC" represents "response change". Results are averaged for the pairwise monitors.

<10% of the on-road vehicles in Washington State,[39] which makes the distribution of on-road BC concentrations less smooth than that of UFP concentrations and thus leads to the difference of the optimal distance.

In real-world settings, the rendezvous of vehicles can be either natural or designed. It is likely that designed rendezvous would be a better method as it can ensure sufficient sampling time and pollutant concentration variabilities. In the present study, the rendezvous of the two vehicles was designed while sampling for 2 h in total within a buffer of 140 m, which was proven to be feasible. Actually, a minimum of 0.2-h sampling within the optimal distance was sufficient for the rendezvous calibration based on this study. The design of the rendezvous generally involves the meeting points of pairwise vehicles and their moving in the same direction within the optimal distance for a minimum of 0.2 h in total. As shown in Table S2, the sampling time for UFP monitors within a 60−80 m buffer and BC monitors within a 20−30 m buffer is 0.2 h. For this scenario, the NMADs and NRMSDs for both the UFP and BC monitors are <10%.

This study has illustrated the feasibility of calibration of portable monitors for multiple pollutants placed in two vehicles based on the proposed optimal-distance rendezvous calibration. This method can be extended to situations with a large fleet of vehicle-based air pollution mobile monitoring platforms, in which each vehicle includes one set of the monitors. For instance, assume that there are a fleet of six vehicles conducting air pollution mobile monitoring in a city with the sampling routes shown in Figure S5. In this case, each vehicle would approach another two or three vehicles during the monitoring. Taking Vehicle 2 as an example to illustrate the rendezvous calibration in such situations, Vehicle 2 can rendezvous with Vehicles 1, 3, and 5 during the sampling. Applying the proposed optimal-distance-based rendezvous calibration approach, three linear calibration models for monitors in Vehicle 2 can be obtained. The average of the postcalibrated concentrations with these three models are taken as the best-estimated concentrations (see more details in Figure S5). Similarly, the best-estimated concentrations for all the other monitors in other vehicles can be calculated. However, prior to the application of the rendezvous calibration approach to the fleet data set, there are two issues that need to be considered.

First, the measurements from malfunctioning monitors which measure considerably higher or lower than other pairwise monitors need to be excluded otherwise the measurements of the entire fleet would be driven higher or lower. There are several steps to identify a malfunctioning monitor. Whenever feasible, factory calibration should be conducted. Factory calibration is insufficient to identify a malfunctioning monitor since the factory reference concentrations (<6000 particles/cm^3) are generally much lower than in situ concentrations (>100000 particles/cm^3 sometimes) for UFP monitors. However, factory calibration makes sure that these instruments are consistent at lower concentrations. If factory calibration is inaccessible, there should be a lab colocation calibration with all instruments in prior to the fleet mobile monitoring. The lab colocation deployment mostly

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

provides an atmosphere of lower UFP concentrations similar to the factory calibration, which enables the possibility to identify those malfunctioning monitors which measure much lower or higher (e.g., 30%; depending on the needs) than the rest. Those malfunctioning monitors can be either sent back to factory for calibration or calibrated against the average of the "good" instruments. The factory or lab colocation calibration ensures the consistency of all instruments at lower concentrations. After that, we can make a further check using the rendezvous data set from the fleet mobile monitoring. Taking Vehicle 2 in Figure S5 as an example, there would be three rendezvous data sets relating to Vehicle 2. If the monitor in Vehicle 2 consistently measure much lower or higher (e.g., 30%; depending on the needs) than the other three monitors in those rendezvous data sets, the monitor in Vehicle 2 should be taken as a malfunctioning monitor. Similarly, other malfunctioning monitors can be identified using other rendezvous data sets. Although the real-world situations may be much more complicated than the aforementioned example, we illustrate that the malfunctioning monitors could be identified through a series of statistical analyses using the rendezvous data sets.

Second, there might exist a concern whether the optimal-distance-based rendezvous calibration improves the data quality in comparison with the precalibrated data. To address this concern, we examined the MAD, NMAD, RMSD, and NRMSD between the pre/post-calibrated concentrations and the averaged concentrations of the pairwise monitors (referenced concentrations) for the original and five additional simulated rendezvous data sets (see Table S7 for more details). As shown in Table S7, the NRMSDs and NMADs with the rendezvous calibration are no larger than those without the rendezvous calibration for all monitors under all scenarios except for the AE51 monitors under several simulated scenarios in which AE51 Monitor_1 measured 5−20% higher than in the original scenario. In addition, the differences of these statistical metrics between the scenarios with and without the rendezvous calibration increase with the response characteristic change of the UFP monitors. Hence, application of the rendezvous calibration approach would generally improve the quality of the data set, especially for the UFP monitors. As for future studies, such comparisons should be conducted to decide whether to apply the rendezvous calibration approach.

Our study has several limitations. *First*, this work focuses on portable monitors for UFPs and BC. Future analysis of rendezvous calibration for other monitors without gold-standard reference instruments, using an approach analogous to that applied in this study, is warranted. *Second*, this study has ignored the potential differences of on-road UFP and BC profiles between vertical and horizontal directions when calculating the intervehicle distances. However, the differences of altitudes of the two vehicles were relatively small (mostly <10 m), especially for those with the optimal distance:10−80 m for UFP monitors and 0−30 m for BC monitors (Table S8). *Third*, due to data constraints, the variation of portable monitor calibrations over time was not evaluated, which is of great importance to long-term mobile monitoring studies. *Fourth*, this field study was conducted in northwest US and mostly in winter. More studies are warranted to examine whether the results derived from this study would work for other locations and other times of year. *Finally*, this work used a two-vehicle-based air pollution monitoring platform to evaluate the rendezvous calibration method. Although this method should apply to mobile monitoring platforms with more fleet vehicles as discussed previously, future studies involving multiple-vehicle-based mobile monitoring platforms would be valuable.

Despite these limitations, this rendezvous calibration approach has promise for many mobile monitoring scenarios with more and more air pollution sensors being deployed and particularly lower-cost sensors that may not be well-calibrated from the manufacturer.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.0c00612.

Quality assurance, statistic descriptions of pollutant levels, statistic descriptions of accuracy of buffer-specific rendezvous calibration, summary of linear regression models, statistic descriptions of accuracy of optimal-distance rendezvous calibration, sensitivity analysis, rendezvous calibration vs no rendezvous calibration, statistic descriptions of intervehicle altitude differences, overview of methodology, experimental setup, scatter plot for the original scenario, scatter plot for the simulated scenario, illustration of rendezvous calibration of six fleet vehicle-based air pollution mobile monitoring (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jianbang Xiang − *Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0001-5196-2574; Phone: +1-206-518-8568; Email: jx56@uw.edu

Authors

Elena Austin − *Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0002-4724-1042

Timothy Gould − *Department of Civil & Environmental Engineering, University of Washington, Seattle, Washington 98195, United States*

Timothy Larson − *Department of Environmental & Occupational Health Sciences and Department of Civil & Environmental Engineering, University of Washington, Seattle, Washington 98195, United States*

Michael Yost − *Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, Washington 98195, United States*

Jeffry Shirai − *Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0002-9864-532X

Yisi Liu − *Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0001-6974-1462

Sukyong Yun − *Department of Civil & Environmental Engineering, University of Washington, Seattle, Washington 98195, United States*

Edmund Seto − *Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, Washington 98195, United States*

Complete contact information is available at:

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294

https://pubs.acs.org/10.1021/acs.est.0c00612

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to express special thanks to Mr. Dave Hardie for his efforts in the field work. The study was funded by the State of Washington for the MOV-UP study and Award Number 5P30 ES007033-23 from the National Institute of Environmental Health Sciences.

■ REFERENCES

(1) Apte, J. S.; Messier, K. P.; Gani, S.; Brauer, M.; Kirchstetter, T. W.; Lunden, M. M.; Marshall, J. D.; Portier, C. J.; Vermeulen, R. C.; Hamburg, S. P. High-resolution air pollution mapping with Google street view cars: exploiting big data. *Environ. Sci. Technol.* **2017**, *51* (12), 6999−7008.

(2) Messier, K. P.; Chambliss, S. E.; Gani, S.; Alvarez, R.; Brauer, M.; Choi, J. J.; Hamburg, S. P.; Kerckhoffs, J.; LaFranchi, B.; Lunden, M. M.; Marshall, J. D.; Portier, C. J.; Roy, A.; Szpiro, A. A.; Vermeulen, R. C. H.; Apte, J. S. Mapping Air Pollution with Google Street View Cars: Efficient Approaches with Mobile Monitoring and Land Use Regression. *Environ. Sci. Technol.* **2018**, *52* (21), 12563−12572.

(3) Hagemann, R.; Corsmeier, U.; Kottmeier, C.; Rinke, R.; Wieser, A.; Vogel, B. Spatial variability of particle number concentrations and NOx in the Karlsruhe (Germany) area obtained with the mobile laboratory 'AERO-TRAM'. *Atmos. Environ.* **2014**, *94*, 341−352.

(4) Biondi, S. M.; Catania, V.; Monteleone, S.; Polito, C. Bus as a sensor: A mobile sensor nodes network for the air quality monitoring. In *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*; 2017; IEEE: 2017; pp 272−277.

(5) Wallace, J.; Corr, D.; Deluca, P.; Kanaroglou, P.; McCarry, B. Mobile monitoring of air pollution in cities: the case of Hamilton, Ontario, Canada. *J. Environ. Monit.* **2009**, *11* (5), 998−1003.

(6) Zwack, L. M.; Paciorek, C. J.; Spengler, J. D.; Levy, J. I. Characterizing local traffic contributions to particulate air pollution in street canyons using mobile monitoring techniques. *Atmos. Environ.* **2011**, *45* (15), 2507−2514.

(7) Riley, E. A.; Banks, L.; Fintzi, J.; Gould, T. R.; Hartin, K.; Schaal, L.; Davey, M.; Sheppard, L.; Larson, T.; Yost, M. G.; Simpson, C. D. Multi-pollutant mobile platform measurements of air pollutants adjacent to a major roadway. *Atmos. Environ.* **2014**, *98*, 492−499.

(8) Yu, C. H.; Fan, Z. H.; Lioy, P. J.; Baptista, A.; Greenberg, M.; Laumbach, R. J. A novel mobile monitoring approach to characterize spatial and temporal variation in traffic-related air pollutants in an urban community. *Atmos. Environ.* **2016**, *141*, 161−173.

(9) Hudda, N.; Gould, T.; Hartin, K.; Larson, T. V.; Fruin, S. A. Emissions from an international airport increase particle number concentrations 4-fold at 10 km downwind. *Environ. Sci. Technol.* **2014**, *48* (12), 6628−6635.

(10) Hudda, N.; Fruin, S. International airport impacts to air quality: size and related properties of large increases in ultrafine particle number concentrations. *Environ. Sci. Technol.* **2016**, *50* (7), 3362−3370.

(11) Larson, T.; Gould, T.; Riley, E. A.; Austin, E.; Fintzi, J.; Sheppard, L.; Yost, M.; Simpson, C. Ambient air quality measurements from a continuously moving mobile platform: Estimation of area-wide, fuel-based, mobile source emission factors using absolute principal component scores. *Atmos. Environ.* **2017**, *152*, 201−211.

(12) Castell, N.; Dauge, F. R.; Schneider, P.; Vogt, M.; Lerner, U.; Fishbain, B.; Broday, D.; Bartonova, A. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environ. Int.* **2017**, *99*, 293−302.

(13) Spinelle, L.; Gerboles, M.; Aleixandre, M. Performance evaluation of amperometric sensors for the monitoring of O3 and NO2 in ambient air at ppb level. *Procedia Eng.* **2015**, *120*, 480−483.

(14) Lewis, A. C.; Lee, J. D.; Edwards, P. M.; Shaw, M. D.; Evans, M. J.; M?ller, S. J.; Smith, K. R.; Buckley, J. W.; Ellis, M.; Gillot, S. R. Evaluating the performance of low cost chemical sensors for air pollution research. *Faraday Discuss.* **2016**, *189*, 85−103.

(15) Manikonda, A.; Zíková, N.; Hopke, P. K.; Ferro, A. R. Laboratory assessment of low-cost PM monitors. *J. Aerosol Sci.* **2016**, *102*, 29−40.

(16) Gillooly, S. E.; Zhou, Y. L.; Vallarino, J.; Chu, M. T.; Michanowicz, D. R.; Levy, J. I.; Adamkiewicz, G. Development of an in-home, real-time air pollutant sensor platform and implications for community use. *Environ. Pollut.* **2019**, *244*, 440−450.

(17) Rai, A. C.; Kumar, P.; Pilla, F.; Skouloudis, A. N.; Di Sabatino, S.; Ratti, C.; Yasar, A.; Rickerby, D. End-user perspective of low-cost sensors for outdoor air pollution monitoring. *Sci. Total Environ.* **2017**, *607*, 691−705.

(18) Borrego, C.; Costa, A.; Ginja, J.; Amorim, M.; Coutinho, M.; Karatzas, K.; Sioumis, T.; Katsifarakis, N.; Konstantinidis, K.; De Vito, S. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmos. Environ.* **2016**, *147*, 246−263.

(19) Kizel, F.; Etzion, Y.; Shafran-Nathan, R.; Levy, I.; Fishbain, B.; Bartonova, A.; Broday, D. M. Node-to-node field calibration of wireless distributed air pollution sensor network. *Environ. Pollut.* **2018**, *233*, 900−909.

(20) Peters, J.; Van den Bossche, J.; Reggente, M.; Van Poppel, M.; De Baets, B.; Theunis, J. Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. *Atmos. Environ.* **2014**, *92*, 31−43.

(21) Karner, A. A.; Eisinger, D. S.; Niemeier, D. A. Near-roadway air quality: synthesizing the findings from real-world data. *Environ. Sci. Technol.* **2010**, *44* (14), 5334−5344.

(22) Austin, E.; Xiang, J.; Gould, T.; Shirai, J.; Yun, S.; Yost, M.; Larson, T.; Seto, E. *Mobile ObserVations of Ultrafine Particles: The MOV-UP study report*; University of Washington: Seattle, WA, 2019.

(23) Hagler, G. S.; Yelverton, T. L.; Vedantham, R.; Hansen, A. D.; Turner, J. R. Post-processing method to reduce noise while preserving high time resolution in aethalometer real-time black carbon data. *Aerosol Air Qual. Res.* **2011**, *11* (5), 539−546.

(24) Aethlabs, What are the detection limits given based on 1 second resolution time? https://aethlabs.com/faq. (accessed on March 2, 2020).

(25) Hijmans, R.; Williams, E.; Vennes, C., Spherical Trigonometry. https://cran.r-project.org/web/packages/geosphere/geosphere.pdf. (accessed on December 18, 2019).

(26) Härdle, W.; Horowitz, J.; Kreiss, J. P. Bootstrap methods for time series. *International Statistical Review* **2003**, *71* (2), 435−459.

(27) Politis, D. N. The impact of bootstrap methods on time series analysis. *Statistical Science* **2003**, *18*, 219−230.

(28) Cook, R. D. Detection of influential observation in linear regression. *Technometrics* **1977**, *19* (1), 15−18.

(29) Yerramshetty, J. S.; Akkus, O. The associations between mineral crystallinity and the mechanical properties of human cortical bone. *Bone* **2008**, *42* (3), 476−482.

(30) Vandeput, N., Forecast KPI: RMSE, MAE, MAPE and Bias. In *Data Science for Supply Chain Forecast*; Independently Published: 2018; p 15.

(31) Wallace, L. A.; Wheeler, A. J.; Kearney, J.; Van Ryswyk, K.; You, H.; Kulka, R. H.; Rasmussen, P. E.; Brook, J. R.; Xu, X. Validation of continuous particle monitors for personal, indoor, and outdoor exposures. *J. Exposure Sci. Environ. Epidemiol.* **2011**, *21* (1), 49.

(32) El Hadri, H.; Petersen, E. J.; Winchester, M. R. Impact of and correction for instrument sensitivity drift on nanoparticle size measurements by single-particle ICP-MS. *Anal. Bioanal. Chem.* **2016**, *408* (19), 5099−5108.

(33) Coffey, E. R.; Pfotenhauer, D.; Mukherjee, A.; Agao, D.; Moro, A.; Dalaba, M.; Begay, T.; Banacos, N.; Oduro, A.; Dickinson, K. L. Kitchen Area Air Quality Measurements in Northern Ghana: Evaluating the Performance of a Low-Cost Particulate Sensor within a Household Energy Study. *Atmosphere* **2019**, *10* (7), 400.

(34) Backman, J.; Schmeisser, L.; Virkkula, A.; Ogren, J. A.; Asmi, E.; Starkweather, S.; Sharma, S.; Eleftheriadis, K.; Uttal, T.; Jefferson, A. On Aethalometer measurement uncertainties and an instrument correction factor for the Arctic. *Atmos. Meas. Tech.* **2017**, *10*, 5039−5062.
(35) Lin, C.; Masey, N.; Wu, H.; Jackson, M.; Carruthers, D.; Reis, S.; Doherty, R.; Beverland, I.; Heal, M. Practical field calibration of portable monitors for mobile measurements of multiple air pollutants. *Atmosphere* **2017**, *8* (12), 231.
(36) Miguel, A. H.; Kirchstetter, T. W.; Harley, R. A.; Hering, S. V. On-road emissions of particulate polycyclic aromatic hydrocarbons and black carbon from gasoline and diesel vehicles. *Environ. Sci. Technol.* **1998**, *32* (4), 450−455.
(37) Zhai, W.; Wen, D.; Xiang, S.; Hu, Z.; Noll, K. E. Ultrafine-particle emission factors as a function of vehicle mode of operation for LDVs based on near-roadway monitoring. *Environ. Sci. Technol.* **2016**, *50* (2), 782−789.
(38) Zhu, Y.; Hinds, W. C.; Kim, S.; Shen, S.; Sioutas, C. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmos. Environ.* **2002**, *36* (27), 4323−4335.
(39) Washington State Department of Transportation. *The 2018 Corridor Capacity Report*; 2018.

https://dx.doi.org/10.1021/acs.est.0c00612
Environ. Sci. Technol. 2020, 54, 4286−4294