



## On-highway vehicle emission factors, and spatial patterns, based on mobile monitoring and absolute principal component score

Yifan Wen<sup>a</sup>, Hui Wang<sup>a</sup>, Timothy Larson<sup>b,c</sup>, Makoto Kelp<sup>d</sup>, Shaojun Zhang<sup>a,e,\*</sup>,  
Ye Wu<sup>a,f,\*</sup>, Julian D. Marshall<sup>b,\*\*</sup>

<sup>a</sup> School of Environment, State Key Joint Laboratory of Environment Simulation and Pollution Control, Tsinghua University, Beijing 100084, PR China

<sup>b</sup> Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, United States

<sup>c</sup> Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA 98195, United States

<sup>d</sup> Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, United States

<sup>e</sup> Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY 14853, United States

<sup>f</sup> State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Beijing 100084, PR China

### HIGHLIGHTS

- Absolute principal component score model is used to analyze mobile monitoring data.
- Fleet-level, fuel-based emission factors are estimated for NO<sub>x</sub>, CO, BC, and PN.
- Higher CO emissions for gasoline vehicles than previous results are possibly due to high emitters.
- High-emitters among heavy-duty diesel trucks are 3–13 times dirtier than normal.

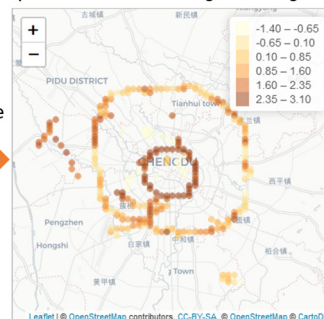
### GRAPHICAL ABSTRACT

#### Mobile monitoring platform



Absolute  
PCA

#### Spatial distribution of high-emitting HDDT



### ARTICLE INFO

#### Article history:

Received 3 March 2019

Received in revised form 7 April 2019

Accepted 11 April 2019

Available online 15 April 2019

Editor: Lidia Morawska

#### Keywords:

Vehicle emission factors

Mobile monitoring

Principal component analysis

Traffic-related pollutants

High emitters

### ABSTRACT

An important component of air quality engineering is quantifying in-use, fleet-average emission factors, and the spatial patterns of vehicle emissions. We report here that an absolute principal component score (APCS) analysis of on-road mobile measurements is a straightforward, efficient method for identifying the major contributors of traffic-related pollutants, deriving fuel-based emission factors, and mapping spatial patterns. Specifically, we applied the APCS model to on-highway measurements of nitrogen oxides (NO<sub>x</sub>), carbon monoxide (CO), carbon dioxide (CO<sub>2</sub>), black carbon (BC), and particle number (PN) obtained from a mobile platform deployed over a 5-day sampling period in Chengdu, China. Data were collected for (1) heavy-duty diesel truck (HDDT) plumes (“chase data”) and (2) the general on-road environment (“non-chase data”). The bootstrapped APCS model was used to estimate area-wide, fuel-based average emission factors and their respective 95% confidence intervals. Two components representing diesel trucks and gasoline vehicles were extracted from non-chase data, accounting for 67% of the variance of the on-highway concentrations. Two additional principal components extracted from HDDT chase data, representing normal and high emission features, further separating the emissions characteristics of HDDTs. The fleet-average emission factors for NO<sub>x</sub>, CO, BC, and PN were 2.2, 50.3, 0.023 g/kg, and  $0.32 \times 10^{15}$  particles/kg for gasoline-powered vehicles, respectively; 33, 3.7, 0.19 g/kg, and  $3.3 \times 10^{15}$  particles/kg fuel for HDDTs’

\* Corresponding authors at: School of Environment, State Key Joint Laboratory of Environment Simulation and Pollution Control, Tsinghua University, Beijing 100084, PR China.

\*\* Corresponding author.

E-mail addresses: [sz262@cornell.edu](mailto:sz262@cornell.edu) (S. Zhang), [ywu@tsinghua.edu.cn](mailto:ywu@tsinghua.edu.cn) (Y. Wu), [jdmars@uw.edu](mailto:jdmars@uw.edu) (J.D. Marshall).

normal emission feature, respectively; and 105, 29, 2.5 g/kg fuel, and  $16 \times 10^{15}$  particles/kg fuel for HDDTs' high emission feature, respectively. APCS results for chase data revealed the existence of high emitters among Chengdu's HDDT fleet, with emission factors 3 to 13 times higher than the normal HDDT vehicles. Although the high emitters are a minority of the fleet, they disproportionately contribute to the overall emissions; emission control policies may wish to target such vehicles.

© 2019 Elsevier B.V. All rights reserved.

## 1. Introduction

Air pollution is a leading risk factor for death and disease globally (Cohen et al., 2017). On-road vehicle emissions are main contributors to air pollutants such as fine particulate matter (PM<sub>2.5</sub>), carbon monoxide (CO), nitrogen oxide (NO<sub>x</sub>) and volatile organic compounds (VOCs), posing severe health risks on people.

Emission factors may differ for gasoline- vs. diesel-powered vehicles. For example, light-duty gasoline vehicles (LDGVs) are the largest anthropogenic source of CO in the US (Office of Air Quality Planning Standards), but on-road diesel engines are estimated to be the single largest anthropogenic source of nitrogen oxide (NO<sub>x</sub>) emissions and an important on-road source of primary particulate matter emissions in the US (Dallmann et al., 2012; Dallmann and Harley, 2010) and large Chinese cities (Wu et al., 2017; Wu et al., 2012; xinhuanet, 2018; Zhang et al., 2014). Also, spatiotemporal patterns of driving – and therefore of emissions – may differ by vehicle type in Chinese cities: light-duty vehicles (LDVs) are more typically on arterials and local roads in urban areas, during daytime; heavy-duty trucks (HDTs) more typically operate at night on highways outside urban areas (Zhang et al., 2018).

Advancements in pollutant measurement instruments, such as finer time-resolution and improved portability (Brantley et al., 2014), have enabled mobile platforms that characterize vehicle emission patterns and pollutant concentrations in traffic environments (Kozawa et al., 2014; Larson et al., 2017; Park et al., 2016; Vogt et al., 2003; Wang et al., 2009; Wang et al., 2011). Compared with in-situ monitoring of tailpipe exhaust, mobile monitoring offers efficient, spatially-resolved area-wide measurements of pollution patterns. For example, the mobile platform employed here integrates devices for measurement of several pollutants. Its mobility enables researchers to take measurements representing a large number of individual vehicles (“fleet-average characteristic”) during each drive, saving time and effort on data sampling. Also, the mobile platform can be used to chase vehicles while driving on roads, reflecting emission patterns in real world driving conditions. The pollution monitoring applications of mobile platforms can be largely categorized as either “vehicle-targeted” or “ambient-targeted”. One vehicle-targeted application of a mobile platform is a “chase study” for analyzing exhaust plumes from individual vehicles (Park et al., 2016; Vogt et al., 2003; Wang et al., 2011). When the mobile platform is not being used for chase a vehicle, it can be used to resolve the ambient pollutant concentrations at broad spatial coverage and characterize their spatial variation.

Traditional vehicle emission measurements generally test or analyze pollutants separately to generate the emission factors for each, requiring significant resources for data acquisition and analysis. For example, many researchers have used chase studies to generate emission factors for particular vehicle categories (Park et al., 2016; Vogt et al., 2003; Wang et al., 2011). This experimental design requires accurate measurements of pollutant concentrations from plumes and ambient air, as well as separate analyses for various pollutants and vehicle types. Several investigators have further examined multivariate correlations between simultaneously measured pollutants using principal component analysis (PCA) (Larson et al., 2017; Riley et al., 2014), an efficient tool for identification of the major sources of pollutant emissions and selection of statistically independent source tracers. The absolute principal component scores (APCS) model is traditionally applied to source apportionment analyses for ambient air pollutants (Thurston and

Spengler, 1987) and has also been applied to mobile measurements of traffic-related pollutants (Larson et al., 2017; Riley et al., 2014). One advantage of APCS is that species weights are based on their overall variance rather than their measurement uncertainty, potentially reducing the influence of meteorologically driven day to day variation in the species concentration in these on-road plumes (Larson et al., 2017). By using the APCS model, the most significant contributors can be identified from the simultaneously measured pollutants data collected from a mobile platform without complicated tests and analyses.

Using Chengdu as an example city, we explored estimating fleet-wide fuel-based average vehicle emission factors by applying the APCS receptor model to measurements obtained from a mobile platform. Based on the application of the APCS method for analyzing “ambient-targeted” data (Larson et al., 2017), we dispatched a mobile platform to intendly measure on-road plumes of HDDTs. A combination of “chase data” for HDDT plumes and “non-chase data” for the generalized traffic environment was expected to provide more emission characteristics of HDDTs. A total of 121,990 s (~34 h) of 1-Hz data were recorded and analyzed. Comparing with traditional laboratory-based measurements, which require abundant tests and analyses, APCS analysis of on-road mobile measurements is solely based on the observed concentrations without reliance on independent traffic information, greatly improving the efficiency of vehicle emission identification and monitoring. Also, the visualization product, an APCS distribution map, can help identify the emission patterns of various vehicle categories, identify emission hotspots, estimate the contribution from high emitters, and thus potentially provide support for the design and implementation of policies for vehicle emission management.

## 2. Methods and data

### 2.1. Research domain and data sampling

We investigated highways in Chengdu, Sichuan, a city with 16 million people and 4.9 million registered vehicles (second only to Beijing) by 2018 (Chengdu Statistical Yearbook, 2018). A previous 2015 source apportionment study (ChengduDaily, 2016) identified on-road vehicles as one of the largest contributors to ambient PM<sub>2.5</sub> concentrations in Chengdu with a contribution of 27% among all the local sources.

Chengdu has six ring-roads forming concentric circles and connected by radial arterials. Our study focused on the area within the fourth ring-road (G4201), which is the main urban area of Chengdu. The sampling route (Fig. 1A) covered all road types within the G4201. The second and fourth ring-roads are major highways, with substantial HDDT traffic; the radial arterials have comparatively greater LDGV traffic.

We measured concentrations of five pollutants: NO<sub>x</sub>, CO, CO<sub>2</sub>, black carbon (BC), and particle number (PN). Those pollutants were selected based on (1) the important contribution of traffic sources to their emissions and concentrations, and (2) the relative importance of these species in distinguishing emissions from gasoline and diesel-powered vehicles. Diesel trucks are important contributors to NO<sub>x</sub>, PN, and BC, whereas CO emissions typically are greater for gasoline than for diesel vehicles. CO<sub>2</sub> is a measurement of fuel consumption. Pollutant concentrations were used to estimate average fuel-based emission factors by source-related features in the study area.



**Fig. 1.** Research domain and sampling equipment. (A) Map of the main sampling routes. Freeways are shown in red, expressways in orange, and arterial roads in blue; (B) the mobile platform used in this study.

Monitors (Table 1) and a global positioning system (GPS) to record speed and position were transported on a gasoline-powered Buick GL8 (Fig. 1B). The resulting measurements were integrated into a database at a frequency of 1 record per 10 s (1-Hz raw data). The information recorded included test time, longitude, latitude and instantaneous speed from the GPS, and 10-s averaged concentrations of NO<sub>x</sub>, BC, CO, PN, and CO<sub>2</sub> collected simultaneously from five different instruments.

Monitoring was conducted during January 12–18, 2018. Routes (Fig. 1) were covered two to three times each. The platform sampled at the speed of surrounding traffic, with a target speed of ~30 km/h for busy expressways, ~70 km/h for highways. Approximately one-third of the data were “chase data” (chasing HDDTs); the rest were non-chase (i.e., not following a certain vehicle or vehicle-type). Driving consisted of chasing (following) an on-road HDDT for ~2 min, followed by non-chasing driving (i.e., sampling the ambient on-road air) while seeking the next HDDT to chase. Chased HDDTs were selected based on convenience: being nearby to the mobile platform. A total of 300 HDDTs were chased, with driving routes relatively uniformly distributed within the research domain.

## 2.2. Data processing and assumptions

We used three steps to reduce white noise and eliminate the influence of background concentrations in our data (Larson et al., 2017). First, we smoothed the concentrations by taking a moving block average of consecutive observations in a 70-s interval centered on each 10-s observation. Second, we estimated the background concentration ( $B_{i,t}$ , described below) associated with each 10-s observation period and

subtracted this background value from the 70-s moving block average concentration.

The smoothed, background-adjusted concentration ( $C_{i,t}^*$ ) for the  $i$ th species in period  $t$  is defined as

$$C_{i,t}^* = C_{i,t} - B_{i,t} \quad (1)$$

where  $C_{i,t}$  is the 70-s moving average concentration between  $t-30$ s and  $t+30$ s centered on the 10-s period  $t$ .  $B_{i,t}$  is a rolling minimum of  $C_{i,t}$  centered on the 10-s period  $t$ , defined as

$$B_{i,t} = \min\{C_{i,t-\tau}, \dots, C_{i,t+\tau}\}, \tau = 300s \quad (2)$$

We chose  $2\tau = 600$ s as a rolling minimum because the mobile monitoring platform travels approximately 4 km during this period, close to the upper bound of the neighborhood monitoring scale (Larson et al., 2017).

Third, we formed a set of adjusted concentrations ( $C_{i,t}^{adj}$ ) by removing samples from the  $C_{i,t}^*$  dataset using the following criteria: 1)  $CO_{2,t}^*$  is  $<5$  ppm, the instrument precision limit; 2) value of concentration is above the 95th daily percentile value for each pollutant. We numbered the adjusted concentrations in chronological order to create the final dataset  $C_{i,k}^{adj}$  used for further calculation, where  $k = 1, \dots, N$ , where  $N$  is the total number of the residual 10-s observations.

## 2.3. Absolute principal component score (APCS) model

We use the APCS model initially described by Thurston and Spengler (1987) and subsequently applied by Larson et al. (2017) to examine the

**Table 1**  
Mobile monitoring platform instruments.

Parameter	Instrument	Manufacturer	Measurement range
Black carbon (BC)	AE33 880 nm channel	Magee	0.01–100 $\mu\text{g}/\text{m}^3$
Particle number concentration (PN)	CPC 3007	TSI, Inc.	0.01–1 $\mu\text{m}$ ; 0–100,000 $\#/\text{cm}^3$
NO <sub>x</sub>	CLD 66	Eco Physics AG, Switzerland.	50–25,000 ppb
CO	Series 7100FM	SIGNAL USA	0–1000 ppm
CO <sub>2</sub>	LI-820	LI-COR, Inc.	0–20,000 ppm
Position and real-time tracking	Hemisphere GPS		Accuracy $\pm 5$ m

Multivariate correlations among the simultaneously measured pollutants from mobile platform and to identify the most significant contributors of on-road emissions. APCS was applied to adjusted, rather than to unadjusted, concentrations, to more equally weight each pollutant. The standardization of adjusted concentrations for pollutant  $i$  is shown in Eq. (3).

$$Z_{i,k} = \frac{C_{i,k}^{adj} - \bar{C}_i^{adj}}{\sigma_i} \quad (i = 1, \dots, m; k = 1, \dots, N) \quad (3)$$

$\bar{C}_i^{adj}$  and  $\sigma_i$  are the mean and standard deviation of all records in the final observation dataset for pollutant  $i$ . The number of concerned species is  $m$ . We retain  $p$  ( $p \leq m$ ) principal components based on eigenvalues  $>0.9$  and apply Varimax rotation to these components to maximize differences between them. APCS for Varimax rotated components are calculated from the scores,  $S_{j,k}$ , for the  $k^{\text{th}}$  observation of the  $j^{\text{th}}$  component as follows:

$$ACPS_{j,k} = S_{j,k} - (S_0)_j \quad (j = 1, \dots, p) \quad (4)$$

where  $S_{j,k}$  is the score derived from the  $Z_{i,k}$ , and  $(S_0)_j$  is the predicted value of the zero vector using the rotated PCA model. ‘Absolute’ is achieved by subtracting the zero vector’s score from the initial scores. Then  $ACPS_{j,k}$  are regressed against the  $C_{i,k}^{adj}$ .

$$C_{i,k}^{adj} = (b_0)_i + \sum_{j=1}^p b_{i,j} (ACPS_{j,k}) + \varepsilon_{i,k} \quad (5)$$

The intercept in Eq. (5) is the contribution to the adjusted concentrations from sources, which cannot be explained by principal components derived from PCA. The predicted concentration of pollutant  $i$ ,  $\hat{Y}_i$ , contributed by component  $j$  to the  $k^{\text{th}}$  sample is then defined by Eq. (6).

$$\hat{Y}_i = b_{i,j} (ACPS_{j,k}) \quad (6)$$

#### 2.4. Fuel-based emission factors

Fuel-based average emission factors (EF) were computed as

$$EF_{i,j} = \frac{\alpha (W_c)_j}{N} \sum_{k=1}^N \left( \frac{\hat{Y}_{i,j,k}}{(\hat{Y}_{CO_2})_{j,k} + (\hat{Y}_{CO})_{j,k}} \right) \quad (7)$$

where  $EF_{i,j}$  is the average fuel-based emission factor in grams (for PN: number) of pollutant  $i$  per kilogram of fuel burned for component  $j$ ,  $N$  is the total number of samples,  $(W_c)_j$  is the carbon weight fraction of the fuel corresponding to the  $j^{\text{th}}$  component and  $\alpha$  is a unit conversion factor ( $1 \mu\text{g}/\text{m}^3$  for CO, BC, and  $\text{NO}_x$  and  $10^{12}$  number/ $\text{cm}^3$  for PN). Because the concentration of BC is lower than  $\text{CO}_2$  and CO by 5 and 3 orders of magnitudes, respectively (according to Table 2), BC is a negligible species comparing with  $\text{CO}_2$  and CO to be considered in the denominator of Eq. (7).

To estimate the uncertainties in EFs, a blocked bootstrap was applied to the above model to minimize potential autocorrelations due to correlated background values not accounted for in our APCS model. We

**Table 2**  
Summary of trimmed, adjusted concentrations of non-chase and chase data.

	Non-chase data	Chase data
# of observations	4749	2290
PN median (mean) [ $\#/\text{cm}^3$ ]	16,965 (22,442)	34,484 (44,398)
$\text{NO}_x$ median (mean) [ppbv]	118 (207)	451 (643)
BC median (mean) [ $\text{ng}/\text{m}^3$ ]	3342 (4028)	4645 (6520)
CO median (mean) [ppmv]	0.71 (0.88)	0.61 (0.87)
$\text{CO}_2$ median (mean) [ppmv]	440 (508)	549 (721)

randomly sampled with replacement from non-overlapping blocks with optimal univariate block sizes determined using the ‘b.star’ function within the ‘np’ package in R. The maximum of the set of five univariate block sizes, corresponding to each of the five species, was chosen for bootstrap sampling. The bootstrap routine was repeated 10,000 times; 95% confidence limits were determined from the distribution of average EFs estimated from each of the 10,000 bootstrap iterations.

### 3. Result and discussion

#### 3.1. Varimax-rotated absolute principal components

An example of measurement data series for ‘chasing’ and ‘non-chasing’ periods is shown in Fig. S1. Mean and median trimmed, background-adjusted pollutant concentrations are summarized in Table 2. Mean/median concentrations for PN,  $\text{NO}_x$ , BC, and  $\text{CO}_2$  are higher for chase than for non-chase data because the chase data targeted HDDTs, while CO concentrations are comparable for two datasets in terms of the mean values. Non-chase data represents the ambient traffic environment, effected by gasoline- and diesel-powered vehicle emissions, as well as background air.

For non-chase data, two principal components were extracted from the PCA according to the criterion that the eigenvalues  $>0.9$  (see Table 3). The principal components are linear combinations of the pollutants’ concentrations. Higher loadings indicate the higher correlation between them; the significance of the principal components is represented by the eigenvalues and the proportion variance, with a larger value indicates that a component is making a larger contribution to the observed concentration.  $\text{CO}_2$ , as a measurement of fuel consumption, mainly reflects total traffic volume. Therefore, we only evaluate the correlation of the components with the other four pollutants. PN,  $\text{NO}_x$  and BC were heavily loaded on the first feature (the ‘PN- $\text{NO}_x$ -BC-rich’ feature); CO was the main species associated with the second feature (‘CO-rich’). Because  $\text{NO}_x$ , PN and BC tend to reflect diesel vehicles and CO is a typical pollutant for gasoline vehicles, the first and second features can be considered to represent diesel and gasoline-featured vehicles, respectively. Combined, the two features account for 67% of the variance of the non-chase data. Importantly, we did not manually select a diesel and a gasoline component in the data, nor did we predesign that those two types would come from this analysis; instead, results for those two vehicle types arise on their own from PCA of on-road measurements.

Further separating the emission characteristics of HDDTs, two principal components were extracted from the PCA for HDDT chase data (loadings, eigenvalues and proportion variances in Table 4). Based on the correlation between components and pollutants, the two components were referred to as ‘ $\text{NO}_x$ -rich’ and ‘PN-BC-rich.’ They represent two different emission features of HDDTs and describe 65% of the overall variance of the chase data.

For the chase data for HDDTs, the best-fit line for scatterplots of concentration variances between adjacent observations versus  $\text{CO}_2$  (Fig. 2) sheds light on the approximate emission factor for that component. Specifically, the slope of the best-fit line (APCS  $> 2$  for each feature)

**Table 3**  
Loadings, eigenvalues and proportion variances of two principal components from PCA for the non-chase data.

Pollutants	PN- $\text{NO}_x$ -BC-rich feature	CO-rich feature
PN	0.829	0.031
$\text{NO}_x$	0.834	0.008
BC	0.789	0.112
CO	0.020	0.849
$\text{CO}_2$	0.244	0.750
Eigenvalues	2.18	1.19
Proportion variances	0.413	0.261

**Table 4**  
Loadings, eigenvalues and proportion variances of two principal components for the chase data for HDDT plumes.

Pollutants	NO <sub>x</sub> -rich feature	PN-BC-rich feature
PN	0.335	0.775
NO <sub>x</sub>	0.910	0.160
BC	-0.116	0.877
CO	0.132	0.002
CO <sub>2</sub>	0.941	0.008
Eigenvalues	1.97	1.29
Proportion variances	0.371	0.280

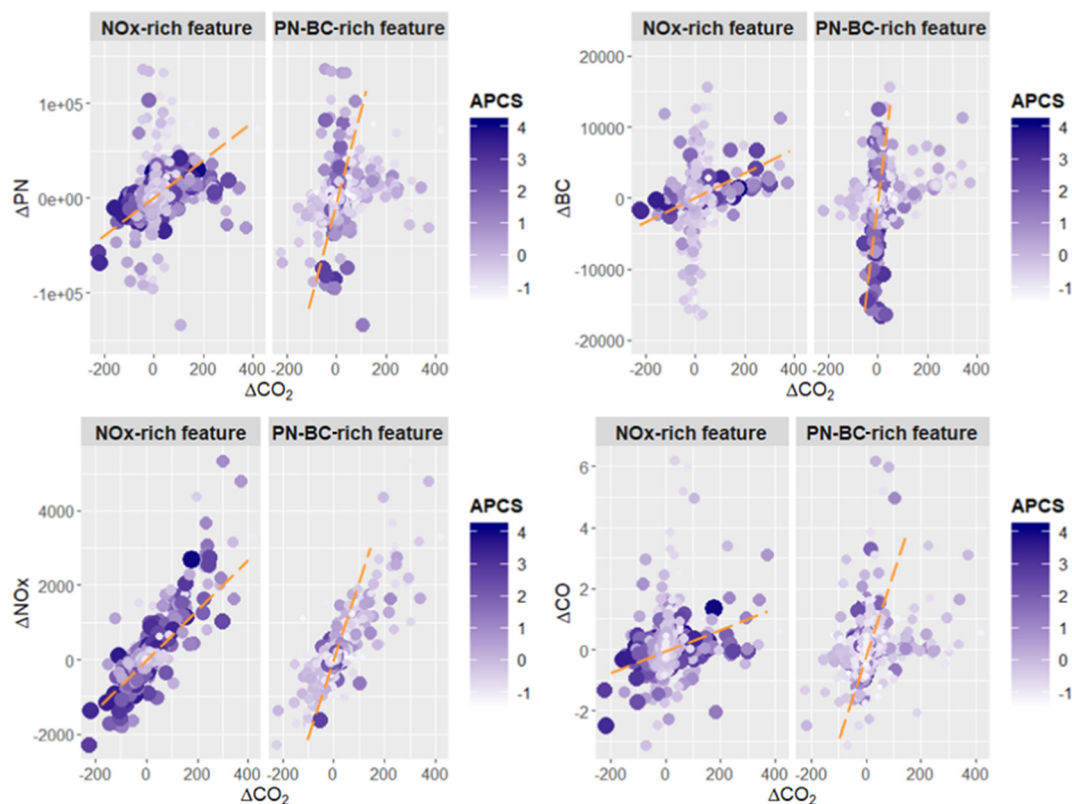
reflects typical fuel-based emission factor for that feature, as CO<sub>2</sub> is a measurement of fuel consumption. (The slope is only used as indicative of the emission factor, not as a calculation of the emission factor.). The fitted gradient for each pollutant in the PN-BC-rich feature is higher than that of the NO<sub>x</sub>-rich feature, by 5, 13, 3, and 8 times for PN, BC, NO<sub>x</sub>, and CO respectively. This result is consistent with the fact that average emission factors for the top 5% of high-emitting HDDTs can be ~7–18 times higher than the normal-emitting HDDTs for all pollutants (Ban-Weiss et al., 2009; Park et al., 2011; Wang et al., 2011). Therefore, the NO<sub>x</sub>-rich and PN-BC-rich features reflect normal and high emission patterns for HDDTs, respectively.

The loadings indicate the magnitude or importance of the pollutants for each component but not the absolute scale of their emission factors. For example, the label “NO<sub>x</sub>-rich feature” refers to NO<sub>x</sub> being a strong component of that feature; however, that does not necessarily mean that the emission factor is higher for that feature (“NO<sub>x</sub>-rich”) than for the other feature (“PN-BC-rich”). According to the results of the PCA, the high emission feature for HDDTs is significantly correlated with concentrations of PN and BC, which is consistent with the conclusion from a previous chase study of heavy-duty trucks that found that the distribution of PN and BC emission are more skewed than NO<sub>x</sub> (Wang et al.,

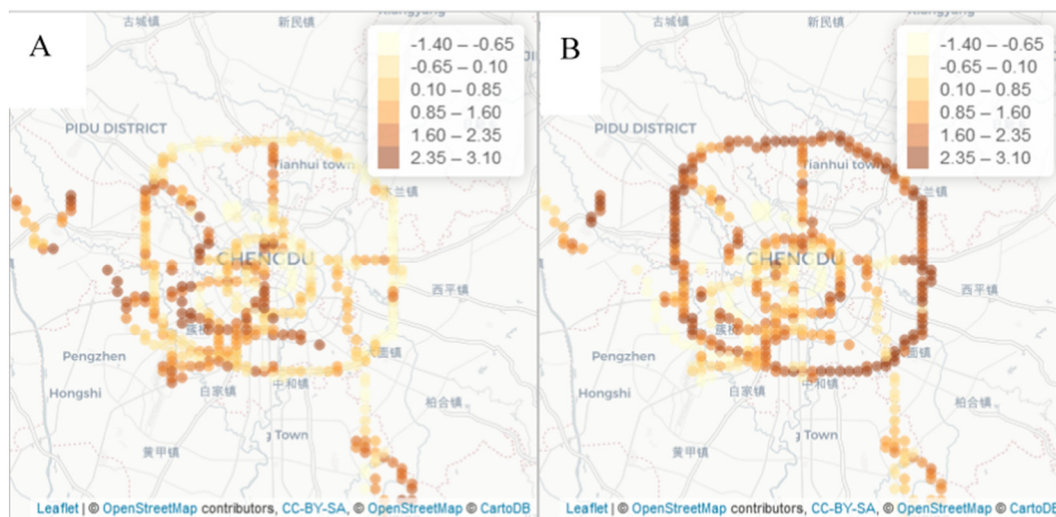
2012). Due to lax inspection processes in many regions of China other than a few cities (e.g., Beijing), some manufacturers have illegally adopted cheaper mechanical pump engines in vehicles presumed to comply with the China III or China IV standards, greatly increasing the emission disparity between high and normal emitters for PM and BC (Zheng et al., 2015). According to Wang et al. (2012), the top 10% highest-emitting trucks are responsible for ~18% of NO<sub>x</sub> emissions and 43%–64% of BC emissions in two Chinese cities. Similarly, our data indicate that the gap in emission factors between high-emitters and normal-emitters is greater for PN and BC than for NO<sub>x</sub>. For normal emitters, the change in emissions between the China II and China IV standards was relatively large for diesel particulate matters (Wu et al., 2012), but relatively minor for NO<sub>x</sub> (Wu et al., 2012). Even for selective catalyst reduction (SCR)-equipped HDDTs, prior research in China (Wu et al., 2012; Yang, 2018) indicates only modest improvements in NO<sub>x</sub> emissions. Because HDDTs that can normally satisfy the modern emission standards have more significantly reduced their PN/BC emissions than NO<sub>x</sub>, the PN-BC-rich feature is less prominent than the NO<sub>x</sub>-rich feature.

### 3.2. Spatial distribution of typical emission features

Varimax-rotated APCS features (Figs. 3 and 4) exhibit different spatial patterns. In non-chase data (Fig. 3), the CO-rich feature (likely representing gasoline vehicles) is more significant on radial arterial roads; the PN-NO<sub>x</sub>-BC-rich feature (likely representing diesel trucks) is more significant on belt expressways, especially the Fourth Ring Road. Those findings are consistent with typical patterns for gasoline and diesel vehicles based on field observations. On the one hand, huge commuting demand happens on radial arterial roads connecting CBD and outer urban areas, causing busy activities for LDGVs (of which the majority are gasoline vehicles). On the other hand, HDDTs are restricted



**Fig. 2.** Scatterplots of the concentration variances between adjacent observations for each pollutant versus CO<sub>2</sub>, and the gradients fitted using points whose APCS exceed 2 for each feature. The data point for each observation is colored and sized based on APCS value.



**Fig. 3.** Spatial distribution of the Varimax-rotated features for non-chase data. (A) CO-rich feature; (B) PN-NO<sub>x</sub>-BC-rich feature. Color is proportional to relative significance of the feature at that location.

from entering the local roads in main urban areas, so the proportion of diesel trucks on belt expressways is much higher than on arterial roads.

In the chase data for HDDT plumes (Fig. 4), the NO<sub>x</sub>-rich (normal-emitting) feature is more prominent on the second-ring expressway, while the PN-BC-rich (high-emitting) feature is more prominent on the fourth-ring freeway. Those spatial distributions are consistent with a policy enacted in 2018 that HDDTs not meeting China IV emission standards cannot enter the area within the Third Ring Road. Implementation of more stringent emission standards for HDDTs is expected to significantly reduce BC emissions.

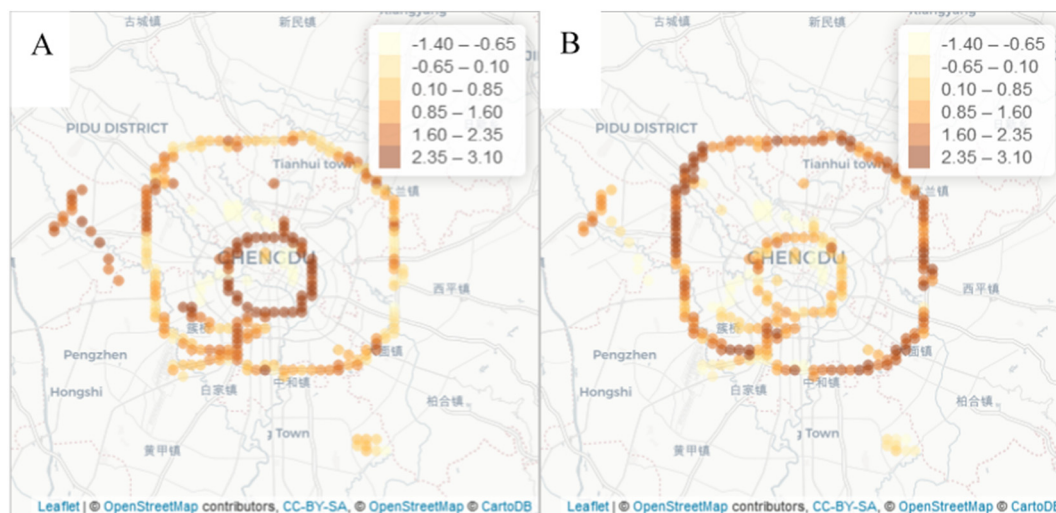
### 3.3. Estimated fuel-based emission factors

Fuel-based emission factors calculated here, based on Varimax-rotated APCS (Figs. 5 and 6; Tables S2 and S3) and employing Eq. 7 (above, Section 2.4), indicate emission factors that are consistent with recent field studies and the COPERT model ((EEA), 2018). 95% confidence limits are estimated from blocked bootstrap.

The fuel-based emission factors of the CO-rich feature for NO<sub>x</sub>, CO, BC and PN are 2.2, 50.3, 0.023 g/kg, and  $0.32 \times 10^{15}$  particles/kg, respectively, representing the fleet-average level of gasoline vehicles in Chengdu. According to the GPS data of mobile monitoring platform during non-chasing periods, the average speed is about 30 km/h, reflecting

the average speed of the fleet. So we consider the fuel-based emission factors of the CO-rich feature obtained here is under an average speed of 30 km/h. The average NO<sub>x</sub> emission factor of Chengdu's gasoline vehicle fleet was comparable to the China 3 level estimated by the EMBEV model (i.e., the archetype model for China's National Emission Inventory Guidebook) (Wu et al., 2017), and was higher than the China 4 level from the EMBEV model and other measurement studies (Huang et al., 2017; Huo et al., 2012; Yang, 2018). We also see that the fleet-level NO<sub>x</sub> emissions for gasoline vehicles in Chengdu were comparable to the average results derived for the vehicle fleet (a dominant mix by gasoline vehicles) in the US, and were between the emission level of Tier 1 and Tier 2 light-duty vehicles (May et al., 2014). For CO, our estimation is moderately higher than the average level of China 3 standard for LDGVs in Chinese studies, Tier 1 standard in the US, as well as the fleet-average LDGV emission level in recent US studies (Dallmann et al., 2013; Kozawa et al., 2014; Larson et al., 2017; Park et al., 2016). In comparison with Europe, the emission factors for both NO<sub>x</sub> and CO are higher than the Euro 3 standard in the COPERT model ((EEA), 2018) and the remote sensing results in London, UK (Carslaw et al., 2011).

According to Chengdu's vehicle statistics in 2017, the most common emission standard of the LDGV fleet is China 4 (53.3%), followed by China 5 (28.1%) and China 3 (12.4%), and some remaining older vehicles



**Fig. 4.** Spatial distribution of the Varimax-rotated features for HDDTs' chase data. (A) NO<sub>x</sub>-rich feature; (B) PN-BC-rich feature.

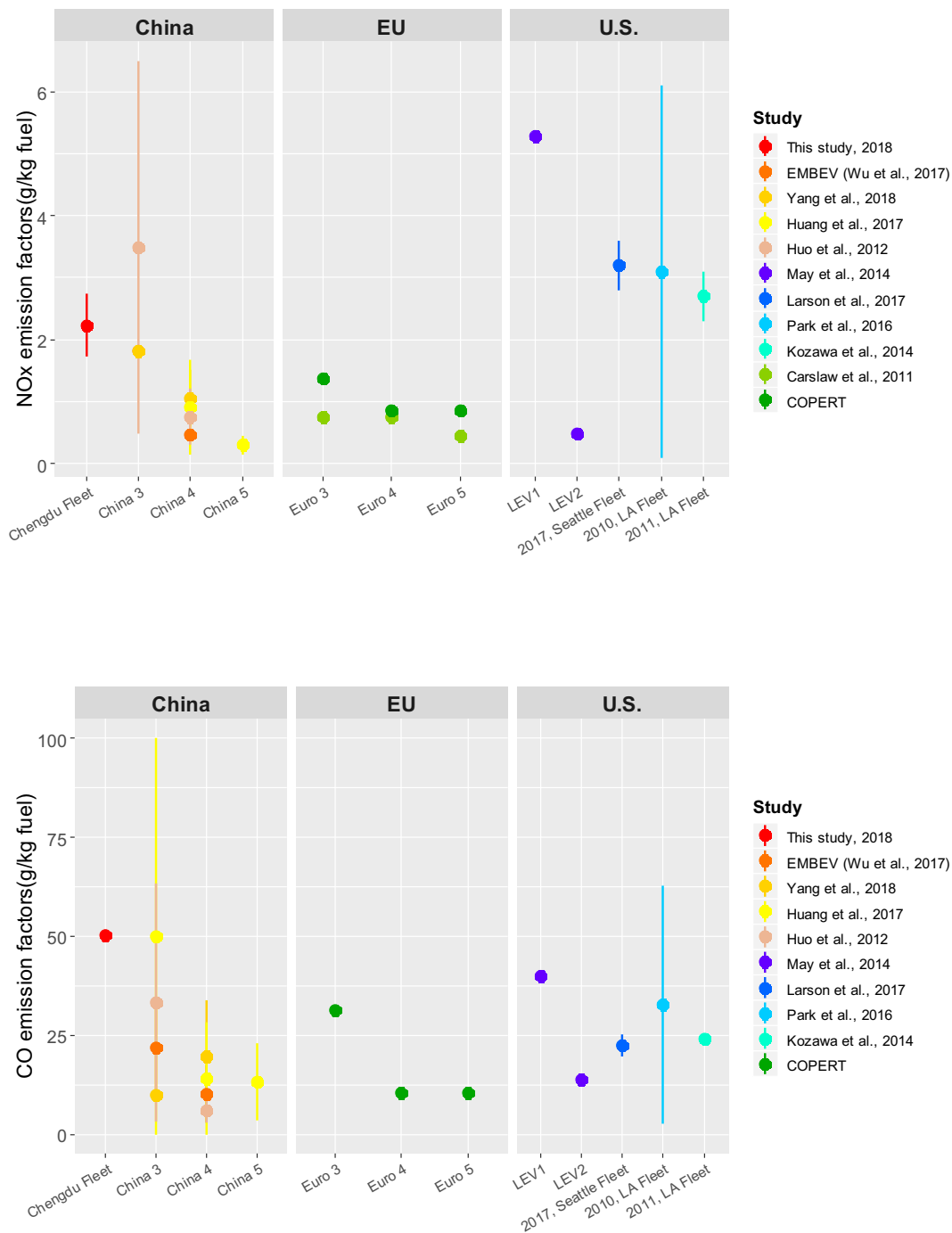


Fig. 5. NO<sub>x</sub> and CO emissions factors of gasoline vehicles predicted by this study, other studies in China and the US, and the COPERT model.

below China 2 (6.2%). Thus, the emission level of Chengdu's LDGV fleet is expected to be closed to the China 4 standard. In contrast, it is notable that our estimation of CO emission is higher than the China 3 level reported in other studies. According to Park et al. (2011), among all LDGV pollutants, CO levels are most sensitive to high emitters that contribute disproportionately to the overall fleet emissions, resulting in a skewed emission distribution. The same study also concluded that if the highest 5% of emitting LDGVs are removed from use, the average emission factors for NO<sub>x</sub> would be reduced by 34%, whereas the CO emission factor would be reduced by 50%. Therefore, the higher estimated CO emission factor observed in this study may reflect (and, may be especially sensitive to the number and characteristics of)

LDGV high emitters among Chengdu's fleet. High emitters can result from malfunction, poor or incorrect maintenance, modification of the vehicle, and/or removal of emission controls (Bishop et al., 2016), and the malfunction of three-way catalyst (TWC) can be a significant cause for high-mileage LDGVs (He et al., 2019).

Emission factors calculated here for diesel vehicles (Fig. 6) are as follows: for NO<sub>x</sub>, CO, BC and PN, respective emission factors were 33.3, 3.7, 0.19 g/kg fuel, and  $3.3 \times 10^{15}$  particles/kg fuel for normal emission conditions, and 104.5, 28.7, 2.52 g/kg fuel, and  $15.7 \times 10^{15}$  particles/kg fuel for high emission conditions. The average speed for chasing HDDTs in this study was about 55 km/h. For both NO<sub>x</sub> and BC, the emission factors of normal emission feature were comparable to emission levels

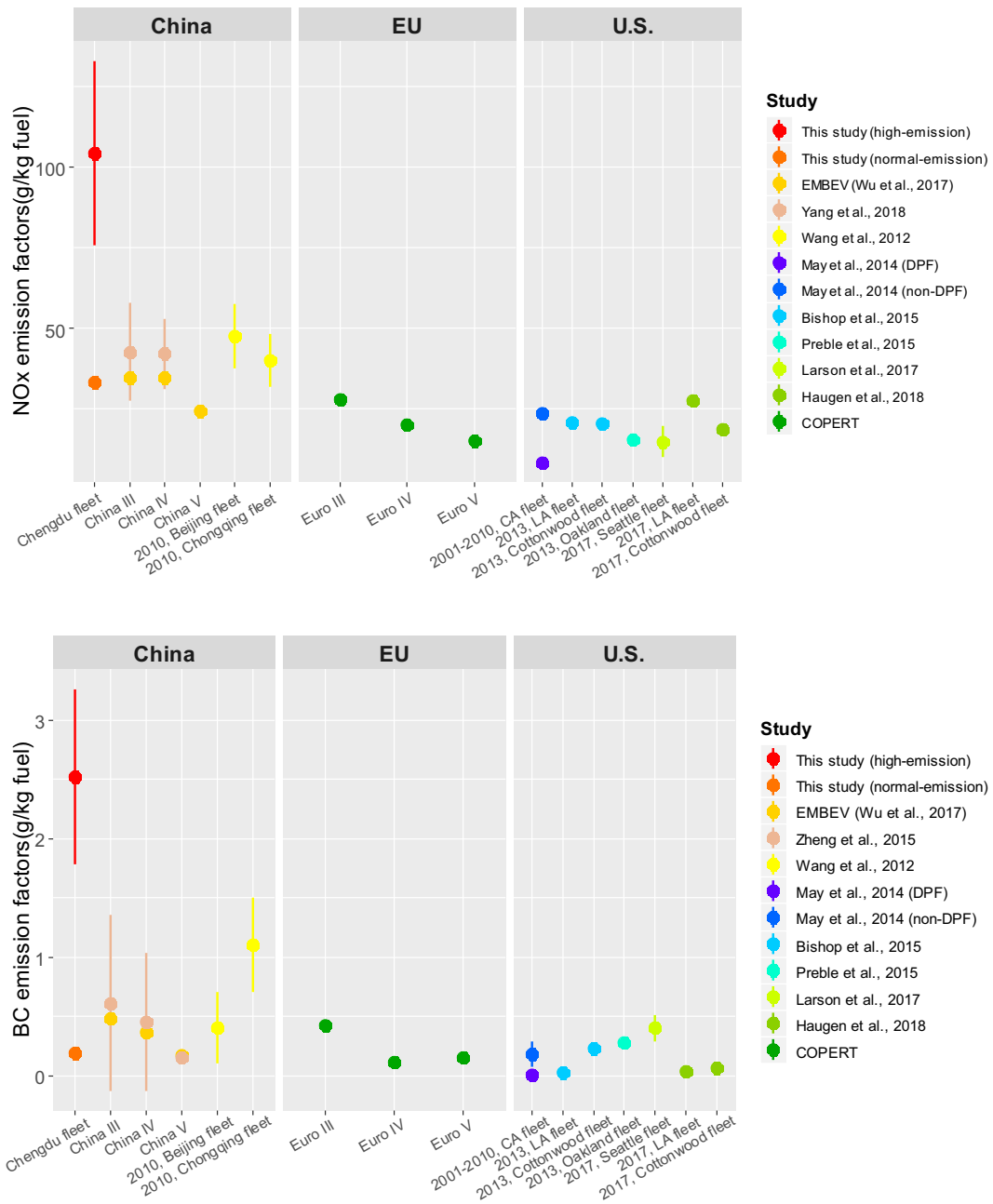


Fig. 6. Comparison of estimated NO<sub>x</sub> and BC emissions factors of diesel vehicles predicted by this study, other studies in China and the US, and the COPERT model.

between China III and China IV in EMBEV model (Wu et al., 2017) and other Chinese studies by using portable emissions measurement system (PEMS) (Yang, 2018; Zheng et al., 2015), consistent with the fact that most of Chengdu's HDDT fleet is compliant with China III and China IV standards. The normal emission level for BC is comparable to an emission level between Euro III and Euro IV in the COPERT model ((EEA), 2018) and to the average emission level observed in recent US studies (Bishop et al., 2015; Haugen and Bishop, 2018; Larson et al., 2017; May et al., 2014; Preble et al., 2015), while the normal emission level of NO<sub>x</sub> is higher than the Euro III standard in the COPERT model and the results of US studies. Zhang et al. (2014) found no significant improvement in NO<sub>x</sub> emissions associated with updating of emission standards, even for China IV diesel vehicles equipped with SCR systems due to the low-speed conditions. The poor performance of SCR systems has been observed in many studies (He et al., 2017; Wu et al., 2012), which

has raised concern among policy-makers and researchers regarding the value of SCR-equipped HDDTs in real-world applications.

The high-emission factors for NO<sub>x</sub>, CO, BC and PN were 3 to 13 times higher than the normal ones, which is consistent with previous observations that the average emission factors for the top 5% of HDDT high emitters can be 7 to 18 times higher than the normal-emitting HDDTs for all pollutants (Ban-Weiss et al., 2009; Park et al., 2011). These high emitters can have a great effect on the average emission level of the fleet, accounting for 40% to 50% of the average emission factors of HDDTs (Ban-Weiss et al., 2009; Park et al., 2011; Wang et al., 2012; Wang et al., 2011). The average emission factors of HDDT fleets are expected to decrease sharply if the high emitters are removed. Thus a small fraction of the most polluting vehicles will become critical for vehicle emission control in the near future, and precise identification and monitoring of high emitters will be essential.



As a sensitivity analysis on the HDDT data, we repeated the PCA analyses for chase data but with high-emitters removed. Specifically, using a definition of high emitters from a related study (Ban-Weiss et al., 2009), we removed the 230 observations with the highest 10% APCS for high emission features. The NO<sub>x</sub>-rich and PN-BC-rich features, and the resulting fuel-based emission factors that appear in the loadings, eigenvalues and proportion variances from this new, high-emitter-removed dataset (Tables S4) are similar to the normal-emission results from the main dataset. Correlations between the components and pollutants were nearly unchanged after removing the 230 high-emitter observations. These results suggest that the normal emission factor reported here is relatively robust to the presence of high-emitters in the dataset.

#### 4. Conclusion

We used PCA to derive fuel-based emission factors for gasoline and diesel vehicles from mobile monitoring data gathered in Chengdu, China. Compared with recent studies conducted in China that used other methods, our study highlights the existence of high emitters among Chengdu's fleet. Of note, the high-emitting HDDTs have emission factors 3 to 13 times higher than normal emission factors. The proportion of total emissions attributable to high emitters is expected to rise over time as the overall fleet becomes cleaner with more stringent regulations (Haugen and Bishop, 2018; Park et al., 2011); our work demonstrates the utility of mobile monitoring to identify and monitor these high emitters.

The application of APCS to data collected simultaneously on multiple pollutants using a mobile platform can be an efficient method to identify the major contributors of traffic-related pollutants and derive the fuel-based emission factors for each source without complicated tests and analyses. The APCS distribution map can also help identify emission patterns of different vehicle categories, identify emission hotspots, estimate the contribution from high emitters, thus potentially provide support for the design and implementation of policies for vehicle emission management.

Future research could (1) chase additional categories of vehicles, and thereby estimate emissions for additional categories of vehicles, (2) record details about the vehicles being chased, to better understand predictors of emissions and potentially to identify characteristics of high-emitters, and (3) record road conditions (e.g., traffic speed and volume, traffic density, road grade, number of lanes, vehicle composition of the surrounding fleet) while sampling, to potentially investigate how emissions vary with the conditions at the time measurements are made.

#### Acknowledgments

Ye Wu acknowledges support from and the National Key Research and Development Program of China (2017YFC0212100), the National Natural Science Foundation of China (91544222) and the Ministry of Science and Technology of China's International Science and Technology Cooperation Program (2016YFE0106300). Yifan Wen was sponsored by Tsinghua University Tutor Research Fund for summer research training. We sincerely thank Susan Cottrell for editorial assistance in the manuscript preparation.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2019.04.185>.

#### References

European Environment Agency (EEA), 2018. *Air Pollutant Emission Inventory Guidebook 2016 Part B: Sectoral Guidance Chapters -Road Transport 2018*.

- Ban-Weiss, G.A., Lunden, M.M., Kirchstetter, T.W., Harley, R.A., 2009. Measurement of black carbon and particle number emission factors from individual heavy-duty trucks. *Environmental Science & Technology* 43, 1419–1424.
- Bishop, G.A., Hottoraguindin, R., Stedman, D.H., McClintock, P., Theobald, E., Johnson, J.D., et al., 2015. On-road heavy-duty vehicle emissions monitoring system. *Environmental Science & Technology* 49, 1639–1645.
- Bishop, G.A., Stedman, D.H., Burgard, D.A., Atkinson, O., 2016. High-mileage light-duty fleet vehicle emissions: their potentially overlooked importance. *Environ Sci Technol* 50, 5405–5411.
- Brantley, H.L., Hagler, G.S.W., Kimbrough, E.S., Williams, R.W., Mukerjee, S., Neas, L.M., 2014. Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. *Atmospheric Measurement Techniques* 7, 2169–2183.
- Carlsaw, D.C., Beever, S.D., Tate, J.E., Westmoreland, E.J., Williams, M.L., 2011. Recent evidence concerning higher NO<sub>x</sub> emissions from passenger cars and light duty vehicles. *Atmos. Environ.* 45, 7053–7063.
- ChengduDaily, 2016. *Mobile Sources Contribute Most to PM<sub>2.5</sub> in Chengdu among the Five Major Sources of Air Pollution*. 2017. The Official Website of Chengdu Municipal People's Government.
- Chengdu Bureau of Statistics, 2018. *Chengdu Statistical Yearbook*.
- Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., et al., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* 389, 1907–1918.
- Dallmann, T.R., Harley, R.A., 2010. Evaluation of mobile source emission trends in the United States. *J. Geophys. Res.-Atmos.* 115.
- Dallmann, T.R., DeMartini, S.J., Kirchstetter, T.W., Herndon, S.C., Onasch, T.B., Wood, E.C., et al., 2012. On-road measurement of gas and particle phase pollutant emission factors for individual heavy-duty diesel trucks. *Environ Sci Technol* 46, 8511–8518.
- Dallmann, T.R., Kirchstetter, T.W., DeMartini, S.J., Harley, R.A., 2013. Quantifying on-road emissions from gasoline-powered motor vehicles: accounting for the presence of medium- and heavy-duty diesel trucks. *Environ Sci Technol* 47, 13873–13881.
- Haugen, M.J., Bishop, G.A., 2018. Long-term fuel-specific NO<sub>x</sub> and particle emission trends for in-use heavy-duty vehicles in California. *Environ Sci Technol* 52, 6070–6076.
- He, L., Hu, J., Zhang, S., Wu, Y., Guo, X., Guo, X., et al., 2017. Investigating real-world emissions of China's heavy-duty diesel trucks: can SCR effectively mitigate NO<sub>x</sub> emissions for highway trucks? *Aerosol Air Qual. Res.* 17, 2585–2594.
- He, L.Q., Hu, J.N., Yang, L.H.Z., Li, Z.H., Zheng, X., Xie, S.X., et al., 2019. Real-world gaseous emissions of high-mileage taxi fleets in China. *Sci. Total Environ.* 659, 267–274.
- Huang, C., Tao, S., Lou, S., Hu, Q., Wang, H., Wang, Q., et al., 2017. Evaluation of emission factors for light-duty gasoline vehicles based on chassis dynamometer and tunnel studies in Shanghai, China. *Atmos. Environ.* 169, 193–203.
- Huo, H., Yao, Z., Zhang, Y., Shen, X., Zhang, Q., Ding, Y., et al., 2012. On-board measurements of emissions from light-duty gasoline vehicles in three mega-cities of China. *Atmos. Environ.* 49, 371–377.
- Kozawa, K.H., Park, S.S., Mara, S.L., Herner, J.D., 2014. Verifying emission reductions from heavy-duty diesel trucks operating on Southern California freeways. *Environ Sci Technol* 48, 1475–1483.
- Larson, T., Gould, T., Riley, E.A., Austin, E., Fintzi, J., Sheppard, L., et al., 2017. 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.* 152, 201–211.
- May, A.A., Nguyen, N.T., Presto, A.A., Gordon, T.D., Lipsky, E.M., Karve, M., et al., 2014. Gas- and particle-phase primary emissions from in-use, on-road gasoline and diesel vehicles. *Atmos. Environ.* 88, 247–260.
- United States Environmental Protection Agency (EPA), 2018. *National Emissions Inventory (NEI) Air Pollutant Emissions Trends Data*.
- Park, S.S., Kozawa, K., Fruin, S., Mara, S., Hsu, Y.-K., Jakober, C., et al., 2011. Emission factors for high-emitting vehicles based on on-road measurements of individual vehicle exhaust with a mobile measurement platform. *J. Air Waste Manage. Assoc.* 61, 1046–1056.
- Park, S.S., Vijayan, A., Mara, S.L., Herner, J.D., 2016. Investigating the real-world emission characteristics of light-duty gasoline vehicles and their relationship to local socioeconomic conditions in three communities in Los Angeles, California. *J Air Waste Manag Assoc* 66, 1031–1044.
- Preble, C.V., Dallmann, T.R., Kreisberg, N.M., Hering, S.V., Harley, R.A., Kirchstetter, T.W., 2015. Effects of particle filters and selective catalytic reduction on heavy-duty diesel drayage truck emissions at the port of Oakland. *Environ Sci Technol* 49, 8864–8871.
- Riley, E.A., Banks, L., Fintzi, J., Gould, T.R., Hartin, K., Schaal, L., et al., 2014. Multi-pollutant mobile platform measurements of air pollutants adjacent to a major roadway. *Atmos. Environ.* 98 (1994), 492–499.
- Thurston, G.D., Spengler, J.D., 1987. A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmos. Environ.* 19, 9–25.
- Vogt, R., Scheer, V., Casati, Roberto, And, f., Benter, Thorsten, 2003. On-road measurement of particle emission in the exhaust plume of a diesel passenger car. *Environmental Science & Technology* 37, 4070–4076.
- Wang, X., Westerdahl, D., Chen, L.C., Wu, Y., Hao, J., Pan, X., et al., 2009. Evaluating the air quality impacts of the 2008 Beijing Olympic Games: on-road emission factors and black carbon profiles. *Atmos. Environ.* 43, 4535–4543.
- Wang, X., Westerdahl, D., Wu, Y., Pan, X., Zhang, K.M., 2011. On-road emission factor distributions of individual diesel vehicles in and around Beijing, China. *Atmos. Environ.* 45, 503–513.
- Wang, X., Westerdahl, D., Hu, J., Wu, Y., Yin, H., Pan, X., et al., 2012. On-road diesel vehicle emission factors for nitrogen oxides and black carbon in two Chinese cities. *Atmos. Environ.* 46, 45–55.

- Wu, Y., Zhang, S.J., Li, M.L., Ge, Y.S., Shu, J.W., Zhou, Y., et al., 2012. The challenge to NO<sub>x</sub> emission control for heavy-duty diesel vehicles in China. *Atmos. Chem. Phys.* 12, 9365–9379.
- Wu, Y., Zhang, S., Hao, J., Liu, H., Wu, X., Hu, J., et al., 2017. On-road vehicle emissions and their control in China: a review and outlook. *Sci. Total Environ.* 574, 332–349.
- xinhuanet, 2018. Beijing Released the Latest Round of PM<sub>2.5</sub> Source Analysis, Mainly from Motor Vehicles. 2018.
- Yang, L., 2018. Real-world Emissions in China: A Meta-study of PEMS Emissions Data From China 0 to China 5/V Light- and Heavy-duty Vehicles.
- Zhang, S., Wu, Y., Hu, J., Huang, R., Zhou, Y., Bao, X., et al., 2014. Can Euro V heavy-duty diesel engines, diesel hybrid and alternative fuel technologies mitigate NO<sub>x</sub> emissions? New evidence from on-road tests of buses in China. *Appl. Energy* 132, 118–126.
- Zhang, S., Niu, T., Wu, Y., Zhang, K.M., Wallington, T.J., Xie, Q., et al., 2018. Fine-grained vehicle emission management using intelligent transportation system data. *Environ. Pollut.* 241, 1027–1037.
- Zheng, X., Wu, Y., Jiang, J., Zhang, S., Liu, H., Song, S., et al., 2015. Characteristics of on-road diesel vehicles: black carbon emissions in Chinese cities based on portable emissions measurement. *Environ Sci Technol* 49, 13492–13500.