

# Mobile Air Monitoring Data Processing Strategies and Effects on Spatial Air Pollution Trends

H. L. Brantley<sup>1,2</sup>, G. S. W. Hagler<sup>1,\*</sup>, S. Kimbrough<sup>1</sup>, R. W. Williams<sup>3</sup>, S. Mukerjee<sup>3</sup>, and L. M. Neas<sup>4</sup>

<sup>1</sup>U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Research Triangle Park, North Carolina, USA

<sup>2</sup>Student Services Contractor, Research Triangle Park, North Carolina, USA

<sup>3</sup>U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory, Research Triangle Park, North Carolina, USA

<sup>4</sup>U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Chapel Hill, North Carolina, USA

\*Corresponding Author: 109 T.W. Alexander Dr., Research Triangle Park, NC, USA 27711; phone: 1-919-541-2827; fax: 1-919-541-0359; hagler.gayle@epa.gov

*Correspondence to:* G. S. W. Hagler  
(hagler.gayle@epa.gov)

**Abstract.** The collection of real-time air quality measurements while in motion (i.e., mobile monitoring) is currently conducted worldwide to evaluate in situ emissions, local air quality trends, and air pollutant exposure. This measurement strategy pushes the limits of traditional data analysis with complex second-by-second multipollutant data varying as a function of time and location. Data reduction and filtering techniques are often applied to deduce trends, such as pollutant spatial gradients downwind of a highway. However, rarely do mobile monitoring studies report the sensitivity of their results to the chosen data processing approaches. The study being reported here utilized 40 hours (> 140,000 observations) of mobile monitoring data collected on a roadway network in central North Carolina to explore common data processing strategies including local emission plume detection, background estimation, and averaging techniques for spatial trend analyses. One-second time resolution measurements of ultrafine particles (UFPs), black carbon (BC), particulate matter (PM), carbon monoxide (CO), and nitrogen dioxide (NO<sub>2</sub>) were collected on twelve unique driving routes that were each sampled repeatedly. The route with the highest number of repetitions was used to compare local exhaust plume detection and averaging methods. Analyses demonstrate that the multiple local exhaust plume detection strategies reported produce generally similar results and that utilizing a median of measurements taken within a specified route segment (as opposed to a mean) may be suffi-

cient to avoid bias in near-source spatial trends. A time-series based method of estimating background concentrations was shown to produce similar, but slightly lower estimates than a location-based method. For the complete dataset the estimated contributions of the background to the mean pollutant concentrations were: BC (15%), UFPs (26%), CO (41%), PM<sub>2.5-10</sub> (45%), NO<sub>2</sub> (57%), PM<sub>10</sub> (60%), PM<sub>2.5</sub> (68%). Lastly, while temporal smoothing (e.g., 5 second averages) results in weak pair-wise correlation and the blurring of spatial trends, spatial averaging (e.g., 10 m) is demonstrated to increase correlation and refine spatial trends.

## 1 Introduction

Air quality research has been revolutionized in recent years by the development and application of mobile platforms capable of resolving air pollutant concentrations in real time. These platforms – including instrumented cars, vans, bicycles, and handheld devices – have been enabled by advancements in air monitoring instrumentation, such as higher time resolution and greater portability, as well as improvements in location resolution using commercially available global positioning systems (GPSs). The mobile measurement strategy has been utilized for a diversity of applications, which can be loosely categorized as: 1) emissions characterization, 2)

near-source assessment, and 3) general air quality surveying (Table 1).

Mobile monitoring is often chosen over other methods for its ability to efficiently obtain data at a high spatial resolution under a variety of different conditions. Vehicle emission factor estimation can be conducted using a number of methods, including chassis dynamometer experiments, tunnel studies, and remote sensing, but mobile monitoring methods are often selected because they enable researchers to characterize in-use emissions of individual vehicles under a variety of operating conditions (Park et al., 2011; Wang et al., 2011, 2012; Westerdahl et al., 2009; Wang et al., 2009).

In near-source environments and general air quality surveys, pollutant concentrations attributable to local sources can vary on the scale of tens of meters or smaller. To characterize this spatial variation, dense networks of stationary monitors can be deployed, but mobile monitoring is often preferred because of the increased spatial flexibility. (Baldauf et al., 2008; Choi et al., 2012; Durant et al., 2010; Hagler et al., 2012; Kozawa et al., 2009; Zwack et al., 2011a; Rooney et al., 2012; Westerdahl et al., 2005; Drewnick et al., 2012; Massoli et al., 2012). Broader surveys of ambient air quality are also frequently conducted using mobile monitoring on a scale ranging from neighborhood to country to characterize regional concentrations or locate previously unknown hotspots (Hagler et al., 2012, 2010; Arku et al., 2008; Adams et al., 2012; Farrell et al., 2013; Drewnick et al., 2012; Van Poppel et al., 2013; Hu et al., 2012).

In this study we considered three components of mobile monitoring data: 1) local exhaust plumes (i.e., tail pipe exhaust near the sampling inlet); 2) local air pollution (e.g., traffic-related air pollution); 3) urban-suburban background (i.e., ambient air quality in the area sampled). Gas and aerosol concentrations change in a continuum of spatial and temporal scales, from the point of emissions to ultimate fate in the environment. Our definitions of local exhaust plumes, local air pollution, and background were derived from the various investigations that have utilized mobile monitoring (Table 1). Local exhaust plumes are defined as short-term events characterized by extremely high pollutant concentrations that can be attributed to directly sampling exhaust from a nearby vehicle. Local air pollution is defined here as well-mixed air that is affected by one or more known sources and modulated by local wind, such as air flow from a major highway to local residential areas. Finally, urban-suburban background, henceforth called “background” for simplicity, is defined on the scale of the route (5–20 km) as representative of ambient air quality conditions without detectable impact of a nearby source.

With this study’s primary focus on spatial air pollution trend analysis to characterize general air quality trends and near-source air pollution, analyses to follow demonstrate the effects of various data processing strategies on resulting trends. In order to extract meaningful information from mobile monitoring data, the full design of the experiment from

the point of monitoring route selection to data processing strategies needs to be taken into account. For example, isolation of local air pollution trends may be simplified by site selection in an environment where roadways surrounding the source of interest have no traffic and the route incorporates a section representative of the background. However, such ideal conditions are rare and often studies need to compensate for local exhaust plumes or imperfect sampling of the background. To isolate key features of interest, past studies have employed a myriad of post-processing strategies (Table 1) to isolate local exhaust plumes, account for the background, and reduce data for visual representation of trends. The selection of post-processing strategies depends on the experimental design and research questions driving the analysis (Fig. 1).

Isolation of local exhaust plumes is of interest for studies that seek to estimate emission factors but also to minimize the impact of sporadic proximate exhaust when determining spatial trends of near-source air pollution. For spatial trend analysis, a variety of strategies have been utilized to minimize bias from incidental local exhaust (Table 2), including using summary statistics less affected by outliers (e.g., percentiles) or algorithms developed to detect brief excursions in the time series. Estimating background is a second key feature of interest to isolate in mobile air monitoring time series. Background air quality often varies diurnally and daily due to meteorology and long-distance transport of pollution. Accounting for the variable background may be conducted through optimal sampling design where an area representative of background is frequently sampled (e.g., Van Poppel et al., 2013). However, when a route completion exceeds the time frame within which the regional background changes or comparisons are being made between multiple routes measured on different days, additional strategies are needed. An alternative approach is to assume the baseline of the time series – represented simply as a low percentile of the data range or a more sophisticated time-varying baseline – is representative of background.

As a final data processing step, temporal or spatial smoothing is often applied either to reduce variation due to atmospheric variability or more effectively display trends (Westerdahl et al., 2005; Weijers et al., 2004; Pirjola et al., 2012). Applying a rolling median or mean can be used to maintain the temporal resolution while reducing the amount of instrument noise and influence of extreme outliers. Aggregating the data to a longer time window can be used to reduce the degree of autocorrelation among the measurements. Types of spatial smoothing include calculating median or mean values along fixed length intervals of the route or within a fixed radius of locations of interest.

Recently, efforts have been made to study the mobile monitoring approach. For example, Van Poppel et al. (2013) evaluated how many sampling route repeats were required to develop a representative data set. However, a rigorous examination of mobile monitoring data processing steps and the

implications for the derived results is needed. This study utilizes a robust multipollutant mobile monitoring dataset collected on a roadway network in North Carolina, USA to evaluate common data-processing methods, including local exhaust plume detection, background estimation, and spatial and temporal smoothing. The dataset consists of 40 hours of mobile monitoring data collected during weekday morning rush hour on 24 days and spanning 12 routes that covered areas of traffic delay, high traffic volume, transit routes, and urban background.

## 2 Methods

### 2.1 Experimental Data

An intensive mobile monitoring campaign was conducted in the Research Triangle Area of North Carolina in the summer of 2012 as part of the Research Triangle Area Mobile Source Emission Study (RAMSES). Measurements were collected using a converted all-electric PT Cruiser. Six instruments were securely installed on board the vehicle: an engine exhaust particle sizer (EEPS) (Model 3090, TSI, Shoreview, MN, USA) which provided size-resolved ultrafine and accumulation mode particle counts, an aerodynamic particle sizer (APS) (Model 3321, TSI, Shoreview, MN, USA) for size-resolved particle counts in fine to coarse mode, a portable aethalometer (AE42, Magee Scientific, Berkeley, CA, USA) that measured black carbon (BC), a dual quantum cascade laser (QCL) (Aerodyne Research, Inc., Billerica, MA, USA) that measured carbon monoxide (CO), a cavity attenuated phase shift (CAPS) monitor that measured nitrogen dioxide (NO<sub>2</sub>) (Aerodyne Research, Inc., Billerica, MA, USA), and a non-dispersive infrared (NDIR) gas analyzer that measured carbon dioxide (CO<sub>2</sub>) (Li-COR 820, LiCOR Biosciences, Lincoln, NE, USA). Due to an observation that the CO<sub>2</sub> data exhibited inexplicable periodic substantial drops in concentration during some of the runs, it was not incorporated into the analyses.

Calibration checks were routinely performed before and after each run. All instruments utilized minimal tubing length (<2 m) and pulled from manifolds connected to two co-located inlets mounted through a side passenger window location. Particle instruments utilized antistatic tubing with minimal bends to avoid particle loss. Further information on the general sampling vehicle set-up is available in Hagler et al. (2010). Wind speed and direction were measured with a highly sensitive 3-dimensional ultrasonic anemometer (Model 81000, RM Young Company, Traverse City, Michigan) placed at a stationary sampling site on each route.

For the current instrument setup, the time between a concentration change (high efficiency particulate air filter for particle instruments, gas standard for gas instruments) at the inlet and visual inspection of instrument response ranged from 0 s to 5 s for both real-time gas and particle instruments.

The response time of the QCL (CO) and APS (particle count in fine to coarse range) was less than 1 s, the CAPS (NO<sub>2</sub>) and aethalometer (BC) was 4 s, and the EEPS (UFP) was 5 s.

After applying the lags determined using the concentration change at the inlet, the correlation between the measurements at various time lags was used to fine-tune the alignment. Because the pollutants are co-emitted, the best estimate of the difference in response times between the instruments can be assumed to correspond with the lag time that produces the maximum correlation coefficient (Choi et al., 2012). CO was chosen as the reference measurement because the quantum cascade laser was the most sensitive instrument with the fastest response time. Because the primary source of CO and BC in the study area was vehicle exhaust, it was assumed that the maximum correlation would occur when the measurements were perfectly aligned. The measured BC concentration was found to lag the CO concentration by 3 s. The other particle instruments were also found to lag the CO measurement by 3 s. The only pollutant measured that was not strongly correlated with CO at a specific lag time was NO<sub>2</sub>; however, NO<sub>2</sub> was strongly correlated with UFPs at a lag of 0 s, so the lag used for UFPs (3 s) was also applied to NO<sub>2</sub>.

The campaign included 12 routes within Wake, Durham, and Orange counties, North Carolina (Fig. 2). The routes covered areas that had previously been classified using modeled traffic data as areas of traffic delay, high traffic volume, transit routes, high signal light density, and urban area. Mobile monitoring was conducted during morning rush hour (7–10:30 am) on 24 weekdays between August 23rd and October 11th, 2012. Each run consisted of approximately an hour and a half of mobile sampling and 45 minutes of stationary sampling. Each route was covered on two sampling days with at least three laps per day. The routes ranged from 5.2–18.1 km in length.

### 2.2 Data processing methods

Mobile monitoring data were processed and displayed using MATLAB (2012), ArcGIS (ESRI, 2011) and R version 2.15.1 (R Core Team, 2012) along with the R packages ggplot2 (Wickham, 2009), openair (Carslaw and Ropkins, 2012), and mcgv (Wood, 2003). A noise-reduction algorithm was applied to black carbon concentrations to reduce the frequency of negative values (Hagler et al., 2011). Examples of near-source air quality gradients and general air quality surveying were selected to illustrate the implications of the following data processing steps: background standardization, local exhaust plume detection, spatial smoothing and temporal smoothing.

Four methods of removing the influence of local exhaust plumes were compared: the running coefficient of variation (COV) method used by Hagler et al. (2012), the standard deviation of the background (SD) method used by Drewnick et al. (2012), the rolling 25th percentile method used by Choi

et al. (2012), and aggregating the data by route segment using outlier resistant statistics such as the median. The first two methods: the COV method (Hagler et al., 2012) and the SD method (Drewnick et al., 2012) are both methods of detecting and flagging local exhaust plumes. For studies characterizing near-source air pollution spatial gradients, one approach may be to remove these flagged periods to avoid confounding influence from side road traffic. Studies focused on personal or localized exposure, however, may not want to remove the influence of the local exhaust plumes. For studies emphasizing emissions characterization, the time periods where local exhaust is detected may be of most interest to isolate and further analyze. These methods are most effective for conditions where an individual vehicle's emissions causes a significant deviation in an otherwise low emissions environment, such as a truck passing the mobile monitoring vehicle on a low traffic residential road. In recent history, these approaches have been developed specifically for understanding local-scale air pollution effects from a nearby source, such as a major roadway, with the mobile sampling vehicle being driven along low-traffic side roads. Applying these approaches in environments with higher traffic, such as while driving on highways, likely only detects major outliers as the within-source pollutant levels are likely consistently high and dynamic. Measurements of local exhaust tend to be both higher and more variable than measurements of well-mixed air. Both the COV method (Hagler et al., 2012) and the SD method (Drewnick et al., 2012) rely on the high variability as well as the magnitude of measurements of local exhaust. The running COV method (Hagler et al., 2012) was developed using UFP concentrations and consists of calculating the rolling 5 s standard deviation (2 s before and after the center data point) and dividing it by the mean concentration of the sampling run. The 99th percentile of the calculated COV is used as a threshold (in Hagler et al. (2012) the threshold COV for UFPs was 2) and any data points with a COV above this threshold are flagged along with the data points 2 s before and after. In the SD method (Drewnick et al., 2012), the standard deviation of measurements (UFP or CO<sub>2</sub>) below the median is calculated ( $\sigma_b$ ). Any measurement more than  $3\sigma_b$  greater than the previous measurement is flagged. Subsequently, all measurements with concentration  $C_i$  that meets the following criteria are flagged as local exhaust plumes:

$$C_i > C_{uf} + 3\sigma_b + \sqrt{n} \times \sigma_b \quad (1)$$

Where  $C_{uf}$  is the concentration of the last unflagged measurement and  $n$  is the number of measurements between  $C_{uf}$  and  $C_i$ .

The rolling 25th percentile method (Choi et al., 2012) does not detect the local exhaust plumes but is used to reduced their effect on spatial gradients and involves calculating the 25th percentile of various time windows. Choi et al. (2012) used a 53 s time window (26 s before and after the center data point) when the sampling vehicle was more than 1 km

from away from a freeway, 31 s (15 s before and after) for distances between 300 m and 1 km, and 3 s (1 s before and after) within 300 m of a freeway. Because the majority of the data used in this comparison was between 300 m and 1 km, a 31 s window was used for the entire dataset to simplify the calculation. One final method of reducing the effect of local exhaust plumes on spatial gradient estimations is to use outlier resistant statistics when aggregating data by route segment such as the median instead of the mean.

A single run conducted on October 11, 2012 on Route B, was chosen to compare these methods because of the large number of laps conducted (12) and favorable wind conditions (from the highway towards the transect). The mean wind speed during mobile sampling was 0.56 m/s and the mean wind direction was 285 deg (from the NW). The route included a section of highway with an Annual Average Daily Traffic (AADT) of 109,000, an approximately 900 m transect running at an angle to the highway with moderate traffic (AADT 32,000), a low traffic road considered urban background, and a shorter transect (Fig. 3a). As an illustration, gradients of CO, UFP, BC, and NO<sub>2</sub> along the highway transect were used to compare the effect of the local exhaust plume removal methods on the 50 m mean concentrations and with the 50 m median concentration of the unfiltered data.

Estimating background concentrations presents a challenge for mobile monitoring studies. For most research groups, replicating the identical instruments (e.g., an engine exhaust particle sizer for ultrafine particles or quantum cascade laser for carbon monoxide) and positioning them in a background location is not feasible. Using alternative instruments for comparison can introduce error into the analysis by using a slower and less-sensitive instrument as the benchmark. An alternate approach used in previous mobile monitoring studies (Hagler et al., 2012; Van Poppel et al., 2013) is the location-based method. This method involves defining areas along the route that have low traffic and are far from any known source as background. The mean or median concentrations measured in the designated background sections are considered representative of background concentrations. Another approach is a time-series based method which relies on elements of the time series itself. One time-series based method is to calculate a single value for each sampling run to use to normalize the concentrations. This value can be a fixed concentration such as the 1st or 10th percentile of the measurements (Bukowiecki et al., 2002). A rolling minimum is a time series-based approach that produces a time-varying background (Kolb et al., 2004). Zwack et al. (2011a,b) also used a time-varying background estimation based on the time-series alone, but instead of estimating background concentrations separately, a smooth function of time over each sampling run was included as a term in the linear regression used to determine concentration differences.

The spline of minimums, which is a time-series based method explored in this study, consists of 3 steps: 1) apply-

ing a rolling 30 s mean to smooth the measurements; 2) dividing the time series into discrete 10 min windows and locating the minimum concentration in each window; 3) fitting a smooth thin plate regression spline through the minimum concentrations. Using a single run conducted on September 21, 2012 on Route B with 14 laps, the spline of 10 min minimums was compared with the location-based method and other time-series based methods: the use of a low percentile (Bukowiecki et al., 2002), a rolling minimum, and the spline of 5 min minimums (Fig. 5). The spline of 10 min minimums was determined to approximate the background better than the other time-series based methods and was further compared to the location-based method using the eight routes with designated background areas. To investigate how well the spline of minimums could estimate the background if the route did not include designated background areas, the concentrations measured in the defined background areas were artificially set to missing before the spline of minimums method was applied. The results were then compared to the location-based method.

Ultimately, the results of spatial and temporal smoothing were compared using all of the measurements collected on Route B. The average speed of the monitoring vehicle on the route was approximately 10 m/s. The smoothing intervals chosen for comparison were 10 m, 50 m, and 100 m segments and the time intervals necessary to traverse each distance at the average speed which equate to 1 s (raw data), 5 s, and 10 s, respectively. Spearman correlation coefficients were calculated for CO, BC, UFPs, NO<sub>2</sub>, and PM<sub>2.5</sub> before and after temporal and spatial smoothing.

### 3 Results and Discussion

The results described in this paper focus on data from a few of the routes and the implications of various data processing steps. Route B, which had the highest number of repetitions, was used to compare local exhaust plume detection and spatial and temporal smoothing methods. Eight of the twelve routes – those which had designated background sections – were utilized to compare how background may be estimated using a purely time-series based approach versus a location based approach. The entire data set (12 routes) was utilized to estimate overall background contribution to the measured concentration of each pollutant. The complexity of the pre-processing and analysis of mobile monitoring data precludes a detailed assessment of all study results in this paper.

#### 3.1 Comparison of methods of local exhaust plume detection

Four methods of removing the influence of local exhaust plumes were compared: the COV method used by Hagler et al. (2012), the SD method used by Drewnick et al. (2012), the rolling 25th percentile method used by Choi et al. (2012),

and aggregating the data by route segment using outlier resistant statistics such as the median. Fig. 3 illustrates the potential of local exhaust plumes to affect the characterization of near-source spatial trends. Using the COV method (Hagler et al., 2012) for both CO and UFPs, several local exhaust plumes were identified (Fig. 3b and c). Spatially aggregating the measurements without removing the influence of the plumes at 7:46 and 7:53 may erroneously lead to the conclusion that concentrations are generally greater along the transect than on the highway (Fig. 3).

For near-source air monitoring studies, a common analysis is to consider concentrations as a function of distance from the source of interest (e.g., edge of road) (Karner et al., 2010). Similar to previous studies, elevated concentrations of mobile source pollutants were observed on the highway (boxplots in Fig. 4), and measured concentrations decreased with increased distance from the highway (Fig. 4). However, the mean 50 m concentrations along the transect are clearly affected by local exhaust plumes, as is evidenced by the mean concentrations of UFP, BC, NO<sub>2</sub> and CO at 250 m (Fig. 4). Using any of the methods of separating measurements of well-mixed air from local exhaust plumes substantially reduces the influence of these events. The 25th percentile method (Choi et al., 2012) results in the lowest estimates of concentrations along the transect because it affects all of the measurements, not just those influenced by local exhaust. The 25th percentile filter (Choi et al., 2012) also results in the smoothest estimate of the gradient along the transect (Fig. 4).

Another important consideration is that different exhaust plumes contain different pollutant mixtures. For example, the plume that was encountered at 250 m caused spikes in all four exhaust indicators, while the plume encountered at 800 m caused increases in CO and UFPs but not in BC or NO<sub>2</sub> (Fig. 4). The measurements used as indicators of local exhaust must be chosen carefully to adequately remove the spikes while retaining the majority of the data. For this run, by using both CO and UFP as indicators, the spikes in NO<sub>2</sub> and BC were successfully removed.

#### 3.2 Comparison of background estimation methods

Several time-series based methods of estimating background were compared with the location-based method. One time-series based method is to calculate a single value for each sampling run using a low percentile of the measurements. However, in the present study over the course of a two hour sampling period, the baseline of the CO time series decreased from 400 ppb to 200 ppb (Fig. 5). During this run the wind speed increased from a mean of 0.3 m/s during the first half hour to a mean of 0.7 m/s for the last half hour and the wind direction was fairly consistently from the southwest: mean wind direction was 217 deg and 249 deg during the first and last half hour, respectively. The decrease in background concentrations over the two hour time span is likely related to an

increase in the atmospheric mixing height during the morning period, however further analysis would be required to fully explore the causes of background variation. Depending on the research question and the pollutant of interest, using a single value to normalize the data may introduce unnecessary error. A rolling minimum did not appear to be a good alternative to a fixed concentration. A 60 s rolling minimum is a better descriptor of variation in well-mixed air concentrations than variation in the background (Fig. 5b). A 300 s rolling minimum results in a more drastic stair-step pattern which is not descriptive of the change in background over time which generally changes very gradually. The spline of minimums best represented the smooth change in the time series baseline over time (Fig. 5c), and the choice of time window (5 min versus 10 min) did not cause a noticeable difference.

The spline of 10 min minimums was used to estimate the background of six pollutants: BC, CO, NO<sub>2</sub>, PM<sub>2.5</sub>, and UFPs on a total of 16 runs covering 8 routes. The background concentrations were estimated using the spline of minimums and compared with the median concentration measured during each pass of the background (Fig. 6). The spline of minimums estimates were positively correlated but consistently underestimated the median of measured background concentrations. The strongest agreement was between the PM<sub>2.5</sub> estimates, likely due to the fact that PM<sub>2.5</sub> concentrations are least likely to be influenced by incidental traffic in the background areas. Despite the use of the an outlier resistant statistic, the occurrences of median measured background values that are significantly higher than the background estimated by the spline of minimums are likely a result of traffic in the designated background area or other local sources (e.g., lawnmower emissions).

To investigate whether the spline of minimums method could be applied to routes that did not include identifiable background areas, the concentrations measured in the background areas were artificially removed (set to missing) and the background was re-estimated. The spline of minimums method was still able to estimate the background concentrations with the results more evenly distributed around the median concentrations (Fig. 6).

The spline of minimums proved to be an effective method for routes spanning a range of distances and under a variety of meteorological conditions. The average wind speed measured during the runs with designated background areas ranged from 0.4 (m/s) to 1.3 (m/s). The wind direction ranged from fairly consistent to highly variable with an average standard deviation of wind direction (Yamartino, 1984) ranging from 30 deg to 86 deg. The effectiveness of the spline of minimums method at estimating the background concentrations for multiple pollutants across various routes and meteorological conditions will enable researchers to compare routes measured on different days. One of the difficulties in using the location-based method is determining whether the inclusion of a background section in the route is feasible given the

study priorities. By using the spline of minimums method the analysis is simplified.

To illustrate the possibility of comparing different routes sampled on different days, we standardized the background of the concentrations measured on 4 routes on 8 different days by subtracting the estimates produced by the spline of minimums method from the measured concentrations. We then compared the PM<sub>2.5</sub> concentrations with and without background standardization. Before background standardization, the regional background variation obscured the variation in PM<sub>2.5</sub> due to the highway (I-40) (Fig. 7a). The PM<sub>2.5</sub> concentrations measured on Route B on a highway with an AADT of 109,000 were below the 50th percentile when compared with all of the measurements made over the course of the field campaign, while measurements collected on Route A on a road with an AADT of 18,000 and Route C on a road with an AADT of 17,000 were all above the 50th percentile (Fig. 7a). After the background was standardized, the influence of the highway, which is an established source of PM<sub>2.5</sub>, became much more evident: the majority of the measurements collected on highways (AADT ≥ 100,000) fall in the higher percentiles of PM<sub>2.5</sub> concentrations, and measurements made on roads with less traffic fall in the lower percentiles of the dataset (Fig. 7b).

Background standardization will have the greatest effect on measurements of pollutants that have a high regional background concentration relative to the concentrations emitted by the source of interest. The mean of the background concentration of all 24 runs and the contribution of the background to the total concentration was calculated for BC, UFP, CO, NO<sub>2</sub>, PM<sub>10</sub>, and PM<sub>2.5</sub> using the background estimated using the spline of minimums method (Table 3). Of the mobile source pollutants measured in this study, PM<sub>2.5</sub>, PM<sub>10</sub>, and NO<sub>2</sub> all fall into the category of co-emitted pollutants with high regional background concentrations (≥ 50% of the mean measured concentrations, Table 3). In contrast, CO, BC, and UFPs can all be classified as co-emitted pollutants with low regional background concentrations (Table 3).

Kimbrough et al. (2013), in a near road study conducted in Las Vegas, NV, also found that the background contribution of NO<sub>2</sub> to the total concentration was higher than the background contribution of CO and BC, with measured upwind concentrations approximately 69%, 63%, and 44% of downwind concentrations for NO<sub>2</sub>, CO, and BC, respectively. The background contributions measured by Kimbrough et al. (2013) are higher than those calculated for the current study, likely because the downwind measurements made by Kimbrough et al. (2013) were collected 20 m from the road, while many of the measurements in the current study were collected on the highway or on roads with high traffic volume causing the total concentrations to be higher and the fraction attributable to regional background to be lower. Upwind concentrations of UFPs measured by Hagler et al. (2009) were roughly 30% of the nearest downwind site and about 50% of the levels observed at 100 m from the road.

To compare the variation in background concentrations estimated using the spline of minimums method, the mean background value for each run was calculated and the between-run standard deviation (SD) was determined from the resulting 24 mean background values. Additionally, the within-run SDs of the estimated background concentrations were calculated by first calculating the SD of the background concentrations for each run and then taking the range of those values (Table 3). The large differences in within-run SD are likely due to variations in the stability of meteorological conditions. For this reason, the range is given instead of the mean. For CO and NO<sub>2</sub>, the between-run SD was greater than all of the within-run SDs (Table 3), indicating that the between run variation of these pollutants was greater than the hourly variation. For the rest of the pollutants measured, the between-run SD fell within the range of the within-run SD.

### 3.3 Temporal and spatial smoothing methods

The influence of temporal and spatial smoothing on the estimates of the NO<sub>2</sub> concentration gradient along the 900 m highway transect in Route B (analyzed in section 3.2) was also compared. The data shown were collected on September 21, 2012 and October 11, 2012 comprising a total of 26 laps and were filtered using the COV method (Hagler et al., 2012). On September 21, 2012 winds were generally calm with a mean wind speed of 0.4 m/s. On October 11, 2012 winds were slightly stronger and generally from the highway with a mean wind speed of 0.56 m/s and mean wind direction of 285 deg (from the NW). The model  $NO_2 = m \times \log(\text{distance}) + b$  was fit for each smoothing case because previous studies have found that pollutant concentrations tend to decrease exponentially with distance from a major source Karner et al. (2010).

When compared to the raw data (Fig. 8a), spatial smoothing alone clarified the spatial trend (Fig. 8b). In contrast, although temporal smoothing reduces the number of data points, the spatial trend was still obscured (Fig. 8c). Furthermore, while spatial smoothing alone resulted in a fairly smooth gradient and the degree of spatial smoothing did not have a significant effect on the fitted curve (Fig. 8b), aggregating the data to a larger time scale before applying spatial smoothing introduces additional variation (Fig. 8d). This variation is due to the error introduced into the estimation of location by using a longer time scale. The slight increase in concentrations at 500 m is due to a busy intersection at this location.

The same data set comprising 26 laps was used to compare the effect of smoothing on pollutant correlations. The results indicated that both spatial and temporal smoothing causes pollutant concentrations to become more correlated as measured by the Spearman correlation coefficients (Table 4). The average speed of the car on this route was approximately 10 m/s. The Spearman correlation coefficients were calculated for BC, CO, NO<sub>2</sub>, PM<sub>2.5</sub> and UFPs after

applying the COV filter (Hagler et al., 2012) and after calculating 5 s and 10 s averages (discrete windows) and dividing the route into 10 m, 50 m, and 100 m segments and calculating the average of the measurements in each segment. Spatial smoothing resulted in much stronger correlations compared to temporal smoothing (Table 4). After 10 m averaging, all of the pollutants were correlated with coefficients greater than 0.7. After 50 m averaging all of the correlation coefficients were greater than 0.8, but increasing the averaging interval to 100 m did not change any coefficients by more than 0.02. Spatial smoothing results in a smaller sample size used to determine the correlations compared with temporal smoothing due to the repeated laps. While a smaller sample size does not necessarily lead to higher correlation, this study demonstrates that spatial versus temporal averaging of mobile air monitoring observations does appear to provide higher correlation values. The results indicate that spatial averaging may more clearly isolate trends of higher versus lower air pollution environments (highways versus background).

## 4 Conclusions

The recent increase in the number of studies that employ mobile monitoring and the variety of applications demonstrate both the utility and versatility of mobile monitoring. As air monitoring instrumentation continues to advance toward greater portability, higher time resolution, greater capacity for operating autonomously, and lower costs, it is likely that these types of studies will become even more ubiquitous (Snyder et al., 2013). The greater temporal and geographic coverage of air pollution measurements can in turn lead to better protection of health and the environment. However, as was shown in this study, this new wealth of data requires the implementation of innovative data processing techniques to extract meaningful information and develop intuitive visuals. This study investigated the sensitivity of final analysis results to the data processing steps chosen.

A variety of research questions and the corresponding data processing strategies were discussed, and a framework for deciding which strategies to apply was presented. Removing the influence of local exhaust plumes can substantially change a near-source gradient, but the various methods compared resulted in similar results. A time-series based method for estimating background concentrations was compared with the location-based estimation of background. The time-series based method was found to slightly underestimate the background concentrations when compared with the location-based method, possibly due to traffic in the designated background areas. Background standardization was particularly important for pollutants with a high background concentration relative to the total concentration, and estimated background concentrations were shown to vary with time. Spatial averaging (50 m) resulted in smoother concen-

tration gradients and stronger correlations than temporal averaging (5 s).

The results demonstrate the vast amount of information contained in datasets collected using mobile monitoring and the myriad of research questions that can be answered using these data, as well as the sensitivity of the conclusions to the data processing approach utilized.

#### 4.1 Disclaimer

This document has been reviewed in accordance with the U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. The views expressed in this journal article are those of the authors and do not necessarily reflect the views or policies or the U.S. Environmental Protection Agency.

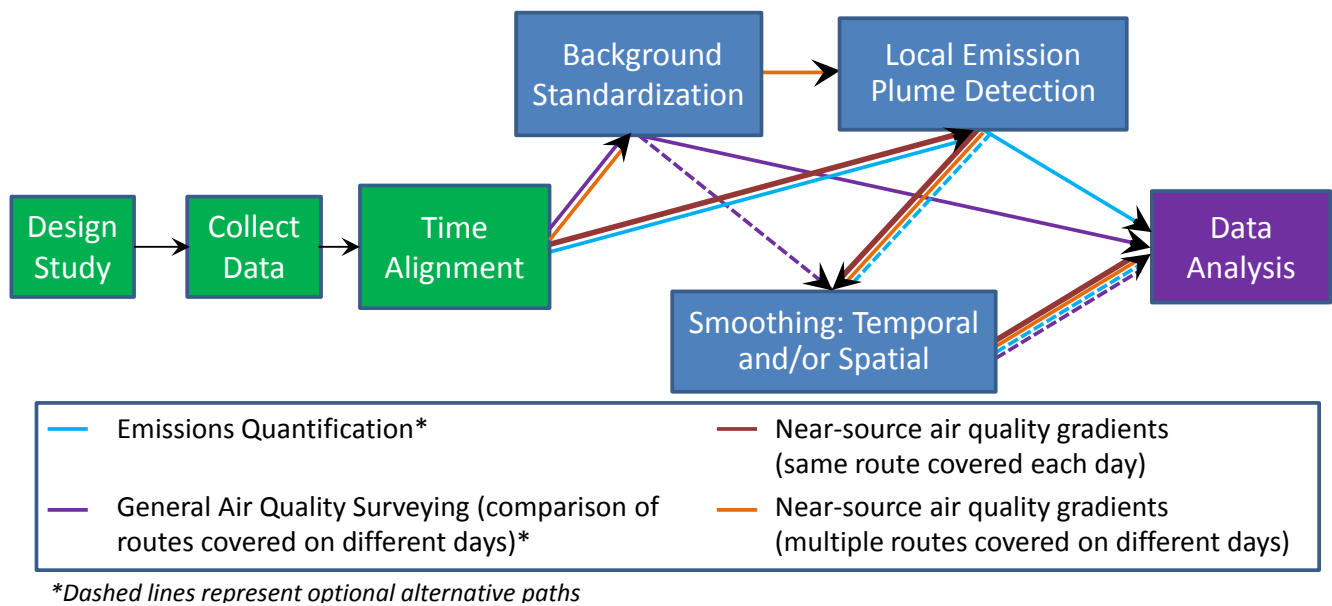
*Acknowledgements.* This research would not have been possible without the careful field measurements conducted by ARCADIS employee Parikshit Deshmukh under contract EP-C-09-027. The authors are also grateful for the research support provided by a number of EPA staff in the Office of Research and Development, particularly Richard Shores, Bill Mitchell, and Robert Wright.

#### References

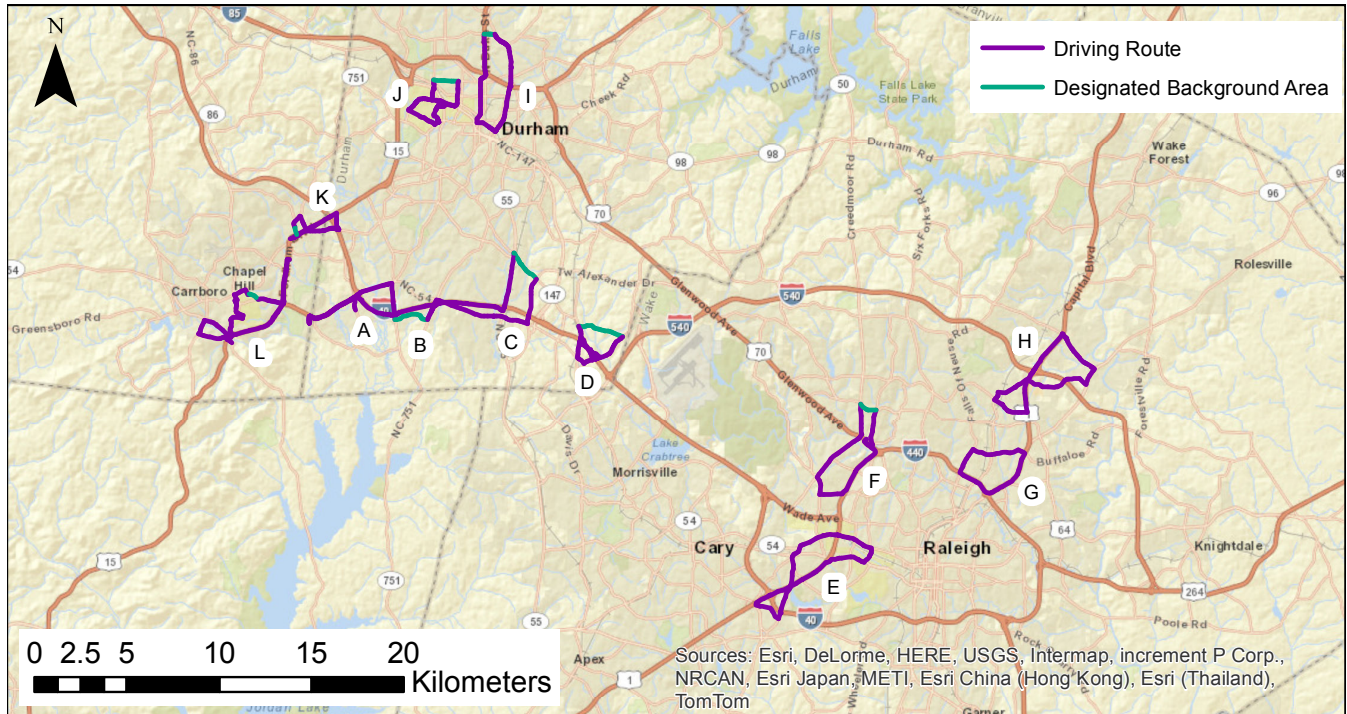
- Adams, M. D., DeLuca, P. F., Corr, D., and Kanaroglou, P. S.: Mobile Air Monitoring: Measuring Change in Air Quality in the City of Hamilton, 2005-2010, *Social Indicators Research*, 108, 351–364, 2012.
- Arku, R. E., Vallarino, J., Dionisio, K. L., Willis, R., Choi, H., Wilson, J. G., Hemphill, C., Agyei-Mensah, S., Spengler, J. D., and Ezzati, M.: Characterizing air pollution in two low-income neighborhoods in Accra, Ghana, *Science of The Total Environment*, 402, 217–231, 2008.
- Baldauf, R., Thoma, E., Khlystov, A., Isakov, V., Bowker, G., Long, T., and Snow, R.: Impacts of noise barriers on near-road air quality, *Atmospheric Environment*, 42, 7502–7507, 2008.
- Bukowiecki, N., Dommen, J., Prevot, A., Richter, R., Weingartner, E., and Baltensperger, U.: A mobile pollutant measurement laboratory—measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution, *Atmospheric Environment*, 36, 5569–5579, 2002.
- Carslaw, D. C. and Ropkins, K.: openair — An R package for air quality data analysis, *Environmental Modelling & Software*, 27–28, 52–61, 2012.
- Choi, W., He, M., Barbesant, V., Kozawa, K. H., Mara, S., Winer, A. M., and Paulson, S. E.: Prevalence of wide area impacts downwind of freeways under pre-sunrise stable atmospheric conditions, *Atmospheric Environment*, 2012.
- Dionisio, K. L., Rooney, M. S., Arku, R. E., Friedman, A. B., Hughes, A. F., Vallarino, J., Agyei-Mensah, S., Spengler, J. D., and Ezzati, M.: Within-neighborhood patterns and sources of particle pollution: mobile monitoring and geographic information system analysis in four communities in Accra, Ghana, *Environ Health Perspect*, 118, 607, 2010.
- Drewnick, F., Böttger, T., Weiden-Reinmüller, S.-L., Zorn, S., Klimach, T., Schneider, J., and Borrmann, S.: Design of a mobile aerosol research laboratory and data processing tools for effective stationary and mobile field measurements, *Atmospheric Measurement Techniques Discussions*, 5, 2273–2313, 2012.
- Durant, J., Ash, C., Wood, E., Herndon, S., Jayne, J., Knighton, W., Canagaratna, M., Trull, J., Brugge, D., Zamore, W., et al.: Short-term variation in near-highway air pollutant gradients on a winter morning, *Atmospheric Chemistry and Physics Discussions*, 10, 5599–5626, 2010.
- ESRI: ArcGIS Desktop: Release 10, Environmental Systems Research Institute, Redlands, CA, USA, 2011.
- Farrell, P., Culling, D., and Leifer, I.: Transcontinental Methane Measurements: Part 1. A Mobile Surface Platform for Source Investigations, *Atmospheric Environment*, 2013.
- Hagler, G., Baldauf, R., Thoma, E., Long, T., Snow, R., Kinsey, J., Oudejans, L., and Gullett, B.: Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants, *Atmospheric Environment*, 43, 1229–1234, 2009.
- Hagler, G., Yelverton, T., Vedantham, R., Hansen, A., and Turner, J.: Post-processing method to reduce noise while preserving high time resolution in aethalometer real-time black carbon data, *Aerosol and Air Quality Research*, 11, 539–546, 2011.
- Hagler, G. S. W., Thoma, E. D., and Baldauf, R. W.: High-Resolution Mobile Monitoring of Carbon Monoxide and Ultrafine Particle Concentrations in a Near-Road Environment, *J Air Waste Manag Assoc*, 60, 328–336, 2010.
- Hagler, G. S. W., Lin, M.-Y., Khlystov, A., Baldauf, R. W., Isakov, V., Faircloth, J., and Jackson, L. E.: Field investigation of roadside vegetative and structural barrier impact on near-road ultrafine particle concentrations under a variety of wind conditions, *Science of The Total Environment*, 419, 7–15, 2012.
- Hu, S., Paulson, S. E., Fruin, S., Kozawa, K., Mara, S., and Winer, A. M.: Observation of elevated air pollutant concentrations in a residential neighborhood of Los Angeles California using a mobile platform, *Atmospheric Environment*, 51, 311–319, 2012.
- Karner, A. A., Eisinger, D. S., and Niemeier, D. A.: Near-roadway air quality: synthesizing the findings from real-world data, *Environmental science & technology*, 44, 5334–5344, 2010.
- Kimbrough, S., Baldauf, R. W., Hagler, G. S., Shores, R. C., Mitchell, W., Whitaker, D. A., Croghan, C. W., and Vallero, D. A.: Long-term continuous measurement of near-road air pollution in Las Vegas: seasonal variability in traffic emissions impact on local air quality, *Air Quality, Atmosphere & Health*, pp. 1–11, 2013.
- Kolb, C. E., Herndon, S. C., McManus, J. B., Shorter, J. H., Zahniser, M. S., Nelson, D. D., Jayne, J. T., Canagaratna, M. R., and Worsnop, D. R.: Mobile laboratory with rapid response instruments for real-time measurements of urban and regional trace gas and particulate distributions and emission source characteristics, *Environmental Science & Technology*, 38, 5694–5703, 2004.
- Kozawa, K. H., Fruin, S. A., and Winer, A. M.: Near-road air pollution impacts of goods movement in communities adjacent to the Ports of Los Angeles and Long Beach, *Atmospheric Environment*, 43, 2960–2970, 2009.
- Massoli, P., Fortner, E. C., Canagaratna, M. R., Williams, L. R., Zhang, Q., Sun, Y., Schwab, J. J., Trimborn, A., Onasch, T. B., Demerjian, K. L., et al.: Pollution gradients and chemical char-



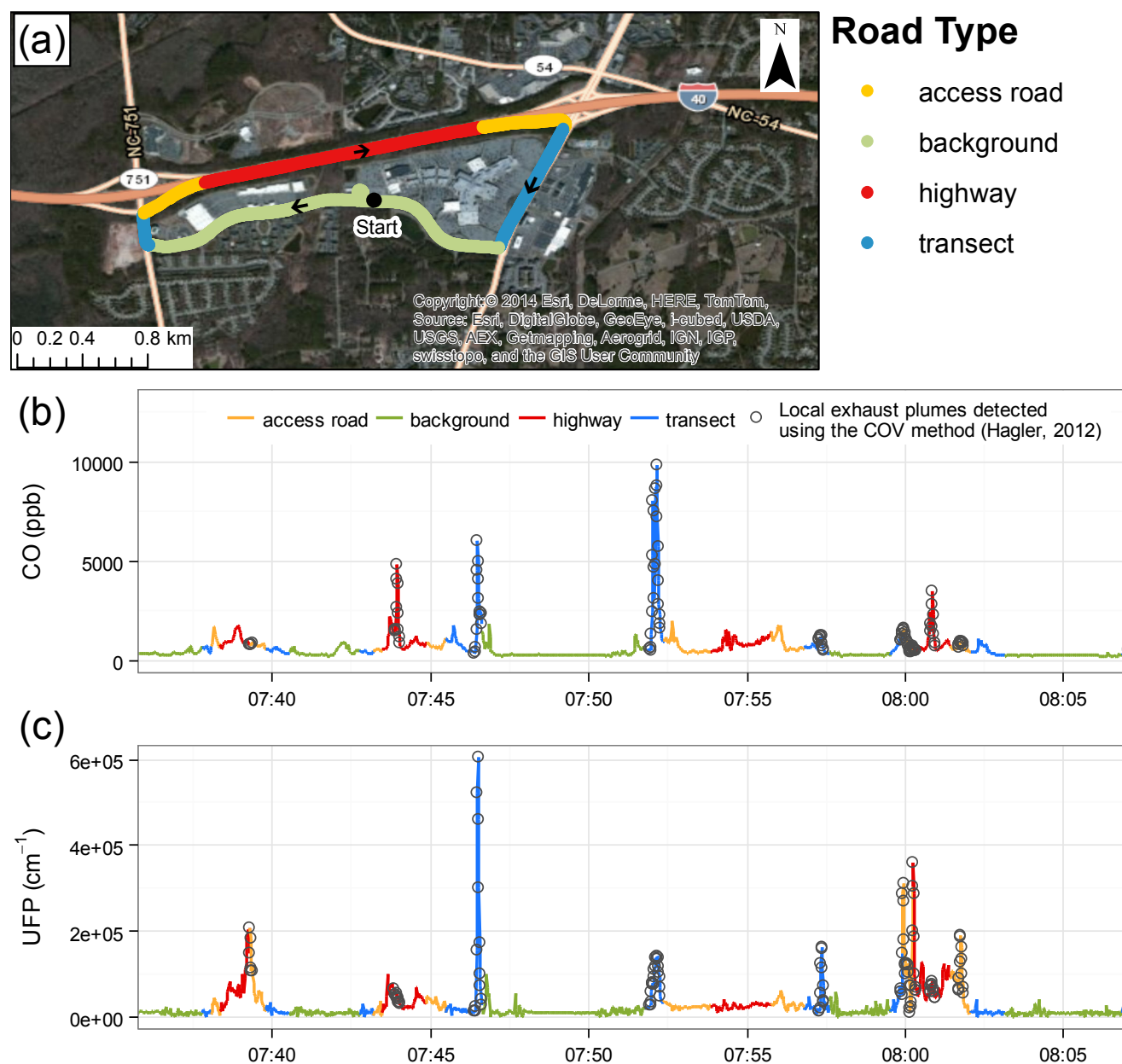
- acterization of particulate matter from vehicular traffic near major roadways: results from the 2009 Queens College Air Quality study in NYC, *Aerosol Science and Technology*, 46, 1201–1218, 2012.
- MATLAB: version 7.14.0 (R2010a), The MathWorks Inc., Natick, Massachusetts, 2012.
- Padró-Martínez, L. T., Patton, A. P., Trull, J. B., Zamore, W., Brugge, D., and Durant, J. L.: Mobile monitoring of particle number concentration and other traffic-related air pollutants in a near-highway neighborhood over the course of a year, *Atmospheric Environment*, 2012.
- Park, S. S., Kozawa, K., Fruin, S., Mara, S., Hsu, Y.-K., Jakober, C., Winer, A., and Herner, J.: Emission factors for high-emitting vehicles based on on-road measurements of individual vehicle exhaust with a mobile measurement platform, *J Air Waste Manag Assoc*, 61, 1046–1056, 2011.
- Pirjola, L., Lähde, T., Niemi, J., Kousa, A., Rönkkö, T., Karjalainen, P., Keskinen, J., Frey, A., and Hillamo, R.: Spatial and temporal characterization of traffic emissions in urban microenvironments with a mobile laboratory, *Atmospheric Environment*, 2012.
- Pétron, G., Frost, G., Miller, B. R., Hirsch, A. I., Montzka, S. A., Karion, A., Trainer, M., Sweeney, C., Andrews, A. E., and Miller, L.: Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study, *Journal of Geophysical Research: Atmospheres* (1984–2012), 117, 2012.
- R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/>, ISBN 3-900051-07-0, 2012.
- Rooney, M. S., Arku, R. E., Dionisio, K. L., Paciorek, C., Friedman, A. B., Carmichael, H., Zhou, Z., Hughes, A. F., Vallarino, J., and Agyei-Mensah, S.: Spatial and temporal patterns of particulate matter sources and pollution in four communities in Accra, Ghana, *Science of the Total Environment*, 435, 107–114, 2012.
- Snyder, E. G., Watkins, T., Solomon, P., Thoma, E., Williams, R., Hagler, G., Shelow, D., Hindin, D., Kilaru, V., and Preuss, P.: The Changing Paradigm of Air Pollution Monitoring, *Environmental Science & Technology*, 2013.
- Van Poppel, M., Peters, J., and Bleux, N.: Methodology for setup and data processing of mobile air quality measurements to assess the spatial variability of concentrations in urban environments, *Environmental Pollution*, 2013.
- Wallace, J., Corr, D., Deluca, P., Kanaroglou, P., and McCarry, B.: Mobile monitoring of air pollution in cities: the case of Hamilton, Ontario, Canada, *Journal of Environmental Monitoring*, 11, 998–1003, 2009.
- Wang, X., Westerdahl, D., Chen, L. C., Wu, Y., Hao, J., Pan, X., Guo, X., and Zhang, K. M.: Evaluating the air quality impacts of the 2008 Beijing Olympic Games: On-road emission factors and black carbon profiles, *Atmospheric Environment*, 43, 4535–4543, 2009.
- Wang, X., Westerdahl, D., Wu, Y., Pan, X., and Zhang, K. M.: On-road emission factor distributions of individual diesel vehicles in and around Beijing, China, *Atmospheric Environment*, 45, 503–513, 2011.
- Wang, X., Westerdahl, D., Hu, J., Wu, Y., Yin, H., Pan, X., and Max Zhang, K.: On-road diesel vehicle emission factors for nitrogen oxides and black carbon in two Chinese cities, *Atmospheric Environment*, 46, 45–55, 2012.
- Weijers, E., Khlystov, A., Kos, G., and Erismann, J.: Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit, *Atmospheric Environment*, 38, 2993–3002, 2004.
- Westerdahl, D., Fruin, S., Sax, T., Fine, P. M., and Sioutas, C.: Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles, *Atmospheric Environment*, 39, 3597–3610, 2005.
- Westerdahl, D., Wang, X., Pan, X., and Zhang, K. M.: Characterization of on-road vehicle emission factors and microenvironmental air quality in Beijing, China, *Atmospheric Environment*, 43, 697–705, 2009.
- Wickham, H.: *ggplot2: elegant graphics for data analysis*, Springer New York, <http://had.co.nz/ggplot2/book>, 2009.
- Wood, S. N.: Thin-plate regression splines, *Journal of the Royal Statistical Society (B)*, 65, 95–114, 2003.
- Yamartino, R.: A comparison of several “single-pass” estimators of the standard deviation of wind direction, *Journal of Climate and Applied Meteorology*, 23, 1362–1366, 1984.
- Zwack, L. M., Paciorek, C. J., Spengler, J. D., and Levy, J. I.: Characterizing local traffic contributions to particulate air pollution in street canyons using mobile monitoring techniques, *Atmospheric Environment*, 45, 2507–2514, 2011a.
- Zwack, L. M., Paciorek, C. J., Spengler, J. D., and Levy, J. I.: Modeling spatial patterns of traffic-related air pollutants in complex urban terrain, *Environmental health perspectives*, 119, 852, 2011b.



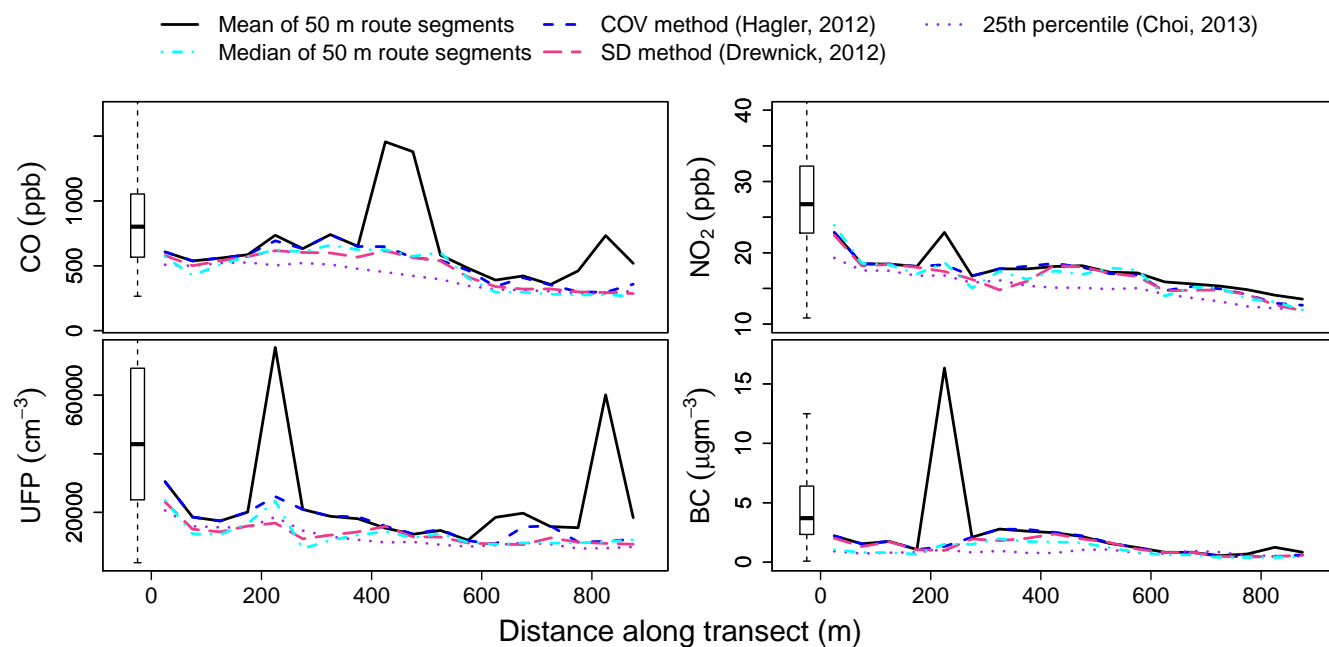
**Fig. 1.** Mobile data processing steps.



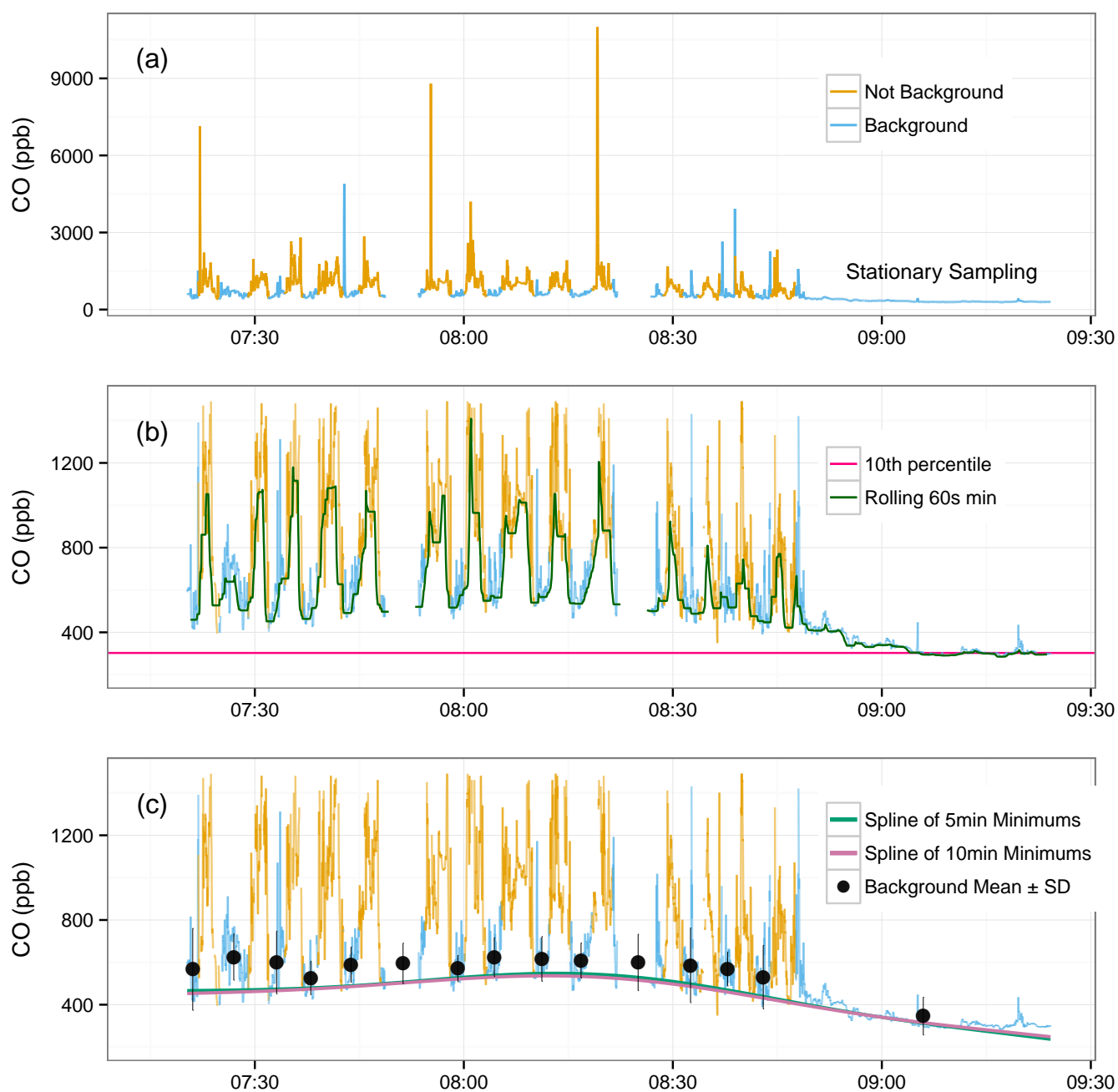
**Fig. 2.** Map of all routes and designated background areas. Background areas were designated for eight of the twelve routes that had identifiable low traffic roads distant from known sources. Routes are labeled A-L.



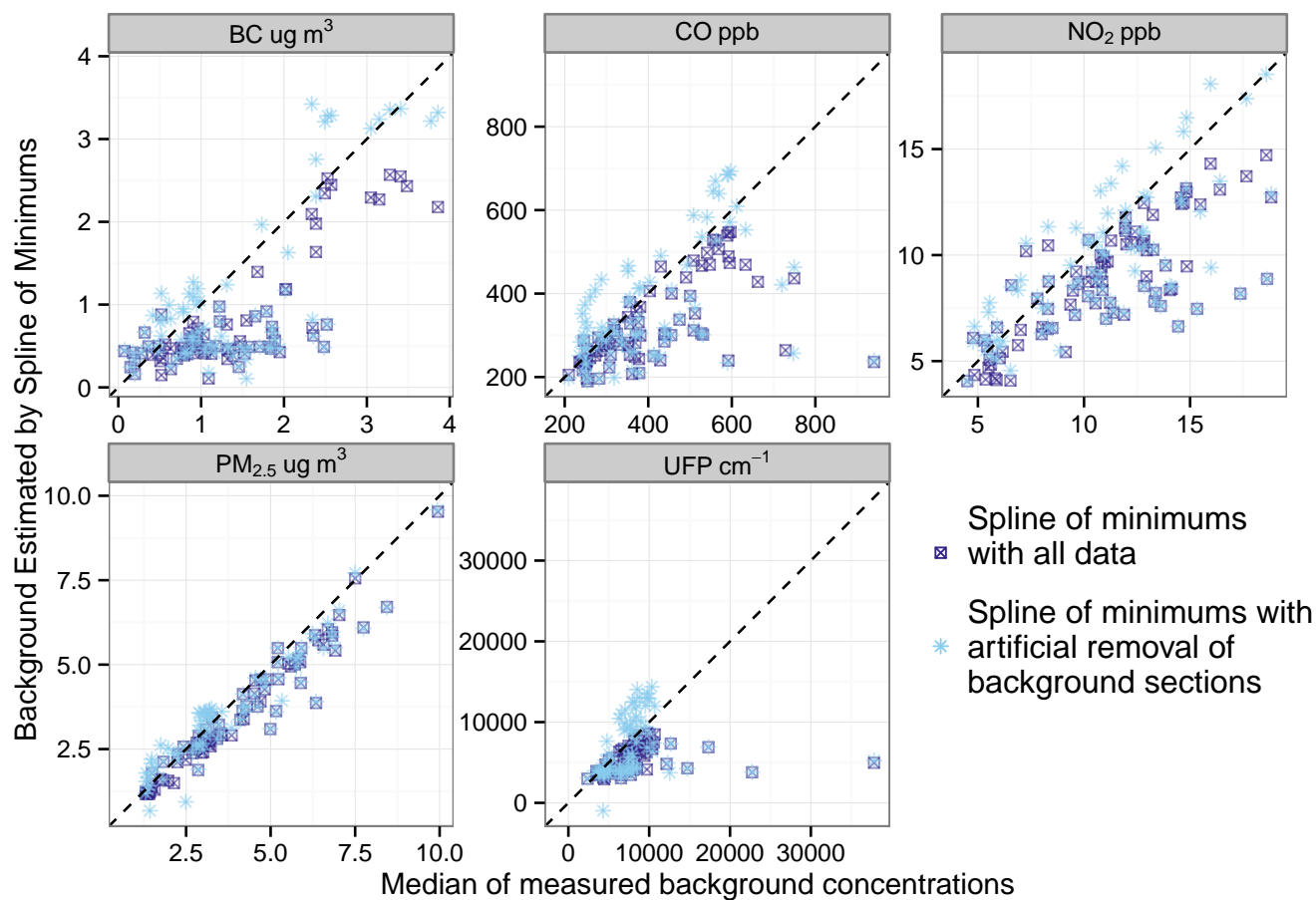
**Fig. 3.** Map of Route B used to compare methods detecting local exhaust plumes and smoothing techniques, (a) measured CO concentrations (b) measured UFP concentrations ( $\text{cm}^{-3}$ ) during a portion of a sampling run. Circles represent local exhaust plumes identified using the COV method (Hagler et al., 2012).



**Fig. 4.** Comparison of the effect of methods of removing the influence of exhaust plumes on transect gradients of CO (a), UFP (b), BC (c), and NO<sub>2</sub> (d). Lines represent 50 m means (except for the line which represents the medians) of measurements from the entire run (12 laps). Boxplots represent unfiltered concentrations measured on the highway.

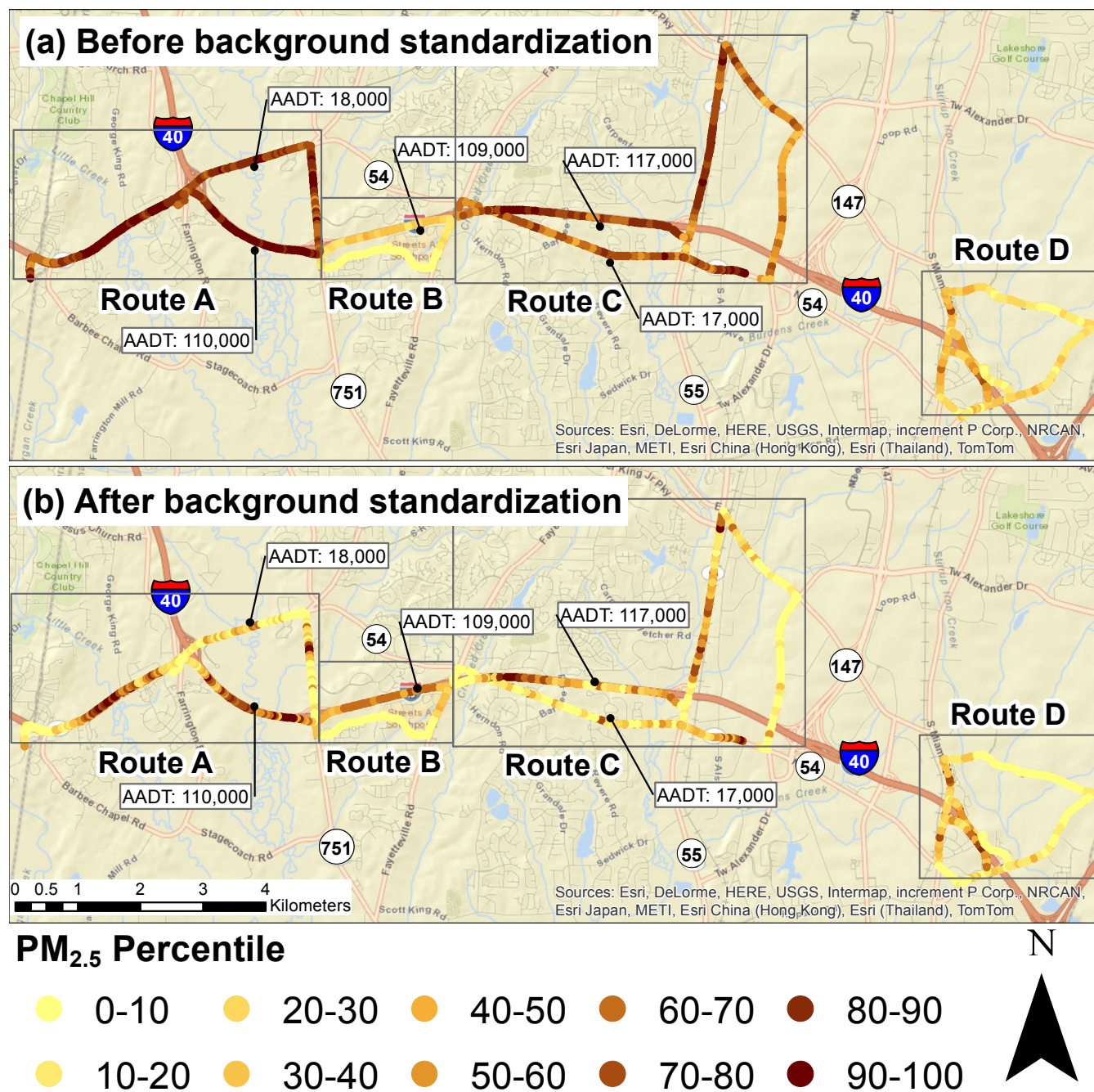


**Fig. 5.** Background estimation methods: a) example time series; b) time series methods: running minimum and 10th percentile; c) location-based method (mean and standard deviation of background areas) and time series methods: spline of 5 min minimums, and spline of 10 min minimums. Gaps in the time series are due to quality control checks. Legend definitions are consistent across panels. The limits of the y-axis of b) and c) have been reduced in to more clearly display the baseline.

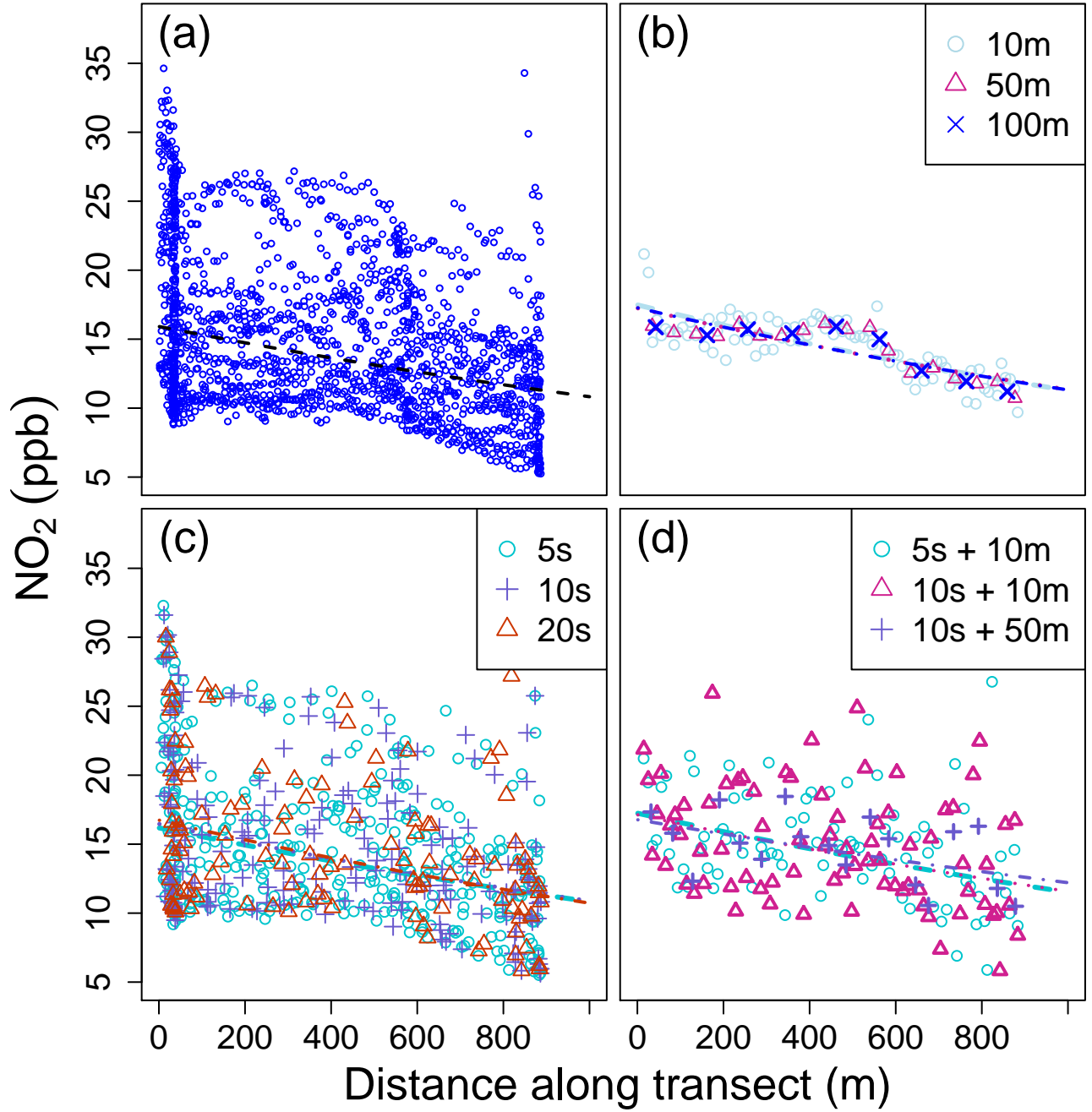


**Fig. 6.** Comparison of location-based (median of measured background concentrations) and time-series based (spline of minimums) background estimates. Time-series based estimates were calculated using all of the data and after the measurements in areas designated as background had been removed. Dashed black lines represent  $y=x$ .





**Fig. 7.** Spatial distribution of PM<sub>2.5</sub> before (a) and after (b) background standardization. Points represent median 50 m values from 8 sampling runs with each route measured on 2 days. Points are colored by PM<sub>2.5</sub> percentile.



**Fig. 8.** Comparison of the effect of temporal and spatial smoothing on  $\text{NO}_2$  measurements collected on the 900 m transect of Route B, shown in Fig. 3, by distance from the highway: a) raw data; b) data after spatial smoothing by calculating mean concentrations by 10 m, 50 m, and 100 m route segments; c) data after temporal smoothing by calculating discrete 5 s, 10 s, and 20 s averages; d) data after combination of temporal and spatial smoothing. The model  $\text{NO}_2 = m \times \log(\text{distance}) + b$  was fit for each smoothing case. The regression lines are plotted with the respective data and the color corresponds with the points used.



**Table 1.** Mobile Monitoring Example Applications

Category	Example Investigations	Measurement Platform	Data Processing Steps Applied	References
Emissions Quantification	Determine and compare emissions factors from vehicles under various driving conditions	Electric vehicle	Local exhaust plume detection, temporal smoothing	Park et al. (2011)
	Evaluate change in emissions factors after traffic intervention	Vehicle	Local exhaust plume detection, background standardization, temporal smoothing	Wang et al. (2009)
	Hydrocarbon emissions characterization	Vehicle	Local exhaust plume detection	Pétron et al. (2012)
Near-source air quality gradients and mitigation strategy evaluation	Roadside barrier impacts	Electric vehicle	Local exhaust plume detection, background standardization, spatial smoothing	Hagler et al. (2012)
	Near-road gradients	Electric vehicle	Time alignment optimization, local exhaust plume detection, background standardization, spatial smoothing	Kozawa et al. (2009); Choi et al. (2012)
	Assess contribution of traffic in street canyons to concentration above background	Backpack	Background standardization, spatial smoothing	Zwack et al. (2011a,b)
	Characterize spatial and temporal variation of near-road gradients	Recreational Vehicle	Temporal and spatial smoothing	Padró-Martínez et al. (2012)
General air quality surveying	Change in air quality in City of Hamilton, 2005-2010	Van	Background standardization, temporal smoothing	Adams et al. (2012); Wallace et al. (2009)
	Characterizing pollution in low-income neighborhoods in Ghana	Handheld	Background standardization, spatial smoothing	Arku et al. (2008); Dionisio et al. (2010)
	Spatial variability of urban air quality	Bicycle	Background standardization, spatial smoothing	Van Poppel et al. (2013)
	Characterize exposure zones	Electric vehicle	Local exhaust plume detection	Hu et al. (2012)

**Table 2.** Mobile Data Processing Methods

Category	Method Description	References
Background Estimation	Designation of background zone	Hagler et al. (2012); Van Poppel et al. (2013)
	Average of fixed monitoring sites	Arku et al. (2008); Dionisio et al. (2010)
	1 min or 5 min 5th percentile	Bukowiecki et al. (2002)
	Assume all of the measurements lower than the most frequent measurement are background	Kolb et al. (2004)
	Include a smooth function of time over each sampling shift as a term in a linear regression	Zwack et al. (2011b,a)
Local Exhaust Plume Detection	Calculate standard deviation of measurements below the median ( $\sigma_b$ ). Flag any measurement more than $3\sigma_b$ greater than the previous measurement. Flag all measurements $> 3\sigma_b + \sqrt{n} \times \sigma_b$ from the last non-flagged measurement, where n is the number of points since the last non-flagged measurement. (UFPs <sup>a</sup> and CO <sub>2</sub> <sup>b</sup> )	Drewnick et al. (2012)
	Modified 5 s running coefficient of variance, with maximum value of 2. (UFP <sup>a</sup> )	Hagler et al. (2012)
	Smoothed rolling minimum (CO <sub>2</sub> <sup>b</sup> , NO <sup>c</sup> )	Kolb et al. (2004)
	Rolling 25th percentile (UFPs <sup>a</sup> , NO <sup>c</sup> , PB-PAH <sup>d</sup> , CO <sup>f</sup> , PM <sub>2.5</sub> <sup>g</sup> )	Choi et al. (2012)
	Video records were checked at times when pollution concentrations peaked at $\geq 2$ observed background concentrations. (BC <sup>e</sup> , NO <sup>c</sup> , UFP <sup>a</sup> , PB-PAH <sup>d</sup> )	Kozawa et al. (2009)

<sup>a</sup>ultrafine particles ( $\leq 100$  nm), <sup>b</sup>carbon dioxide, <sup>c</sup>nitrogen monoxide, <sup>d</sup>particle-bound polycyclic aromatic hydrocarbon, <sup>e</sup>black carbon, <sup>f</sup>carbon monoxide, <sup>g</sup>particulate matter ( $\leq 2.5$   $\mu$ m)

**Table 3.** Summary comparison of pollutant background concentration and variation.

	Mean of Background	Within Run SD <sup>a,b</sup>	Between Run SD <sup>c</sup>	Contribution of Background to Total <sup>d</sup>
BC ( $\mu\text{g m}^{-3}$ )	0.48	0.01 - 0.50	0.40	15%
UFP ( $\text{cm}^{-3}$ )	4990	20 - 2620	1570	26%
CO ( <i>ppb</i> )	298	3 - 57	75	41%
NO <sub>2</sub> ( <i>ppb</i> )	8.8	0.1 - 2.2	2.4	57%
PM <sub>10</sub> ( $\mu\text{g m}^{-3}$ )	6.1	0.1 - 2.9	2.0	60%
PM <sub>2.5</sub> ( $\mu\text{g m}^{-3}$ )	3.7	0.1 - 2.6	1.6	68%

<sup>a</sup>standard deviation, <sup>b</sup>calculated by first calculating the SD of the estimated background for each run and then taking the range of those values. , <sup>c</sup>calculated by determining the mean background value for each run and determining the standard deviation of the resulting 24 mean background values, <sup>d</sup>mean of estimated background for all 24 runs divided by mean measured concentration of all runs multiplied by 100

**Table 4.** Effect of temporal and spatial smoothing on pollutant Spearman correlation coefficients.

	Filtered Raw Data	Temporal Smoothing		Spatial Smoothing		
		5 s	10 s	10 m	50 m	100 m
	N=8386	N=1801	N=921	N=529	N=105	N=52
NO <sub>2</sub> ( <i>ppb</i> ) and PM <sub>2.5</sub> ( $\mu\text{g m}^{-3}$ )	0.07	0.10	0.12	0.77	0.85	0.86
NO <sub>2</sub> ( <i>ppb</i> ) and BC ( $\mu\text{g m}^{-3}$ )	0.11	0.12	0.12	0.76	0.81	0.80
CO ( <i>ppb</i> ) and NO <sub>2</sub> ( <i>ppb</i> )	0.16	0.18	0.18	0.81	0.85	0.85
UFP ( $\text{cm}^{-3}$ ) and PM <sub>2.5</sub> ( $\mu\text{g m}^{-3}$ )	0.41	0.53	0.57	0.83	0.89	0.89
UFP ( $\text{cm}^{-3}$ ) and BC ( $\mu\text{g m}^{-3}$ )	0.47	0.53	0.58	0.80	0.85	0.86
BC ( $\mu\text{g m}^{-3}$ ) and PM <sub>2.5</sub> ( $\mu\text{g m}^{-3}$ )	0.54	0.69	0.74	0.77	0.82	0.84
NO <sub>2</sub> ( <i>ppb</i> ) and UFP ( $\text{cm}^{-3}$ )	0.56	0.60	0.62	0.88	0.87	0.87
CO ( <i>ppb</i> ) and UFP ( $\text{cm}^{-3}$ )	0.58	0.63	0.65	0.88	0.92	0.92
PM <sub>2.5</sub> ( $\mu\text{g m}^{-3}$ ) and CO <sub>2</sub> ( <i>ppm</i> )	0.60	0.72	0.75	0.80	0.84	0.86
CO ( <i>ppb</i> ) and PM <sub>2.5</sub> ( $\mu\text{g m}^{-3}$ )	0.61	0.73	0.76	0.84	0.86	0.86
CO ( <i>ppb</i> ) and BC ( $\mu\text{g m}^{-3}$ )	0.69	0.74	0.76	0.81	0.85	0.87