Determining source impacts near roadways using wind regression and organic source markers

David A. Olson,*† Ram Vedantham, † Gary A. Norris, † Steven G. Brown, ‡§ Paul Roberts‡

for submission to Atmospheric Environment

†National Exposure Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, North Carolina 27711
‡Sonoma Technology, Inc., Petaluma, California 94954
§Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado 80523

*Corresponding author. Tel: +1-919-541-3190; fax: +1-919-541-0905; e-mail: olson.david@epa.gov (D.A. Olson)

Keywords: source apportionment, motor vehicles, hopanes, polycyclic aromatic hydrocarbons (PAHs)

Abstract

Concentrations of 13 organic source markers (10 polycyclic aromatic hydrocarbons and 3 hopanes) are reported from time-integrated samples (24-hr and sub-daily) collected near a highway in Las Vegas, NV. Sample selection for assessing source impacts from the roadway was completed using the wind regression model Air Pollution Transport to Receptor model (EPA APTR 1.0). The model uses a kernel smoothing method for estimating source sectors (sector apportionment) of chemicals across wind speeds and wind directions. The model was applied
using semi-continuous (5-min averaging time) pollutant data (black carbon (BC), CO, NO₂, and NOₓ) and meteorological data. Using simple screening criteria to identify source impacts (>30% sector apportionment from the roadway and errors in the estimated sector apportionment <30%), sector apportionment results were consistent with organic source marker concentrations representative of motor vehicle exhaust (e.g., benzo(g,h,i)perylene and hopane). Results demonstrated the use of APTR to identify source-impacted time intervals when compared with filter samples analyzed for organic source markers.

1. Introduction

Numerous studies have shown an association between roadway traffic and health effects (McConnell et al., 2006; Gauderman et al., 2005; Heinrich et al., 2005; Peters et al., 2004; Janssen et al., 2003; Brauer et al., 2002; Buckeridge et al., 2002; Brunekreef et al., 1997). Although studies have reported concentrations of bulk components such as carbon monoxide (CO), nitrogen oxides (NOₓ), and particulate matter (PM) size fractions in near roadway environments, well-established relationships between PM mass (e.g., PM₂.₅ and PM₂.₅₋₁₀) and health endpoints are unlikely to consistently reflect roadway exposures given its relatively minor effect on PM mass levels. For example, Zhu et al. (2006) reported an approximately five-fold decrease in near roadway particle number concentrations (30–300 m) in Los Angeles during daytime, while PM₂.₅ and PM₂.₅₋₁₀ concentrations varied by only a few percent for the same sampling locations (Zhu et al., 2002).

Improved understanding of specific PM size fractions and/or chemical components is needed to better establish links between roadway emissions and health endpoints. Source tests can provide detailed chemical information for the vehicles and operating conditions tested. For example, exhaust from gasoline-powered motor vehicles (Schauer et al., 2002) and medium-duty
trucks (Schauer et al., 1999) have been reported to consist of >50% organic and elemental carbon (% of PM$_{2.5}$ mass). Numerous particle-phase organic compounds have been identified and quantified (e.g., polycyclic aromatic hydrocarbons, alkanes, alkenes, hopanes, and steranes), and hopanes have been widely used in source apportionment studies (e.g., Schauer and Cass, 2000). Because source tests are limited to a small subset of the vehicle population, roadway tunnel measurements (McGaughey et al., 2004; Fraser et al., 2003; El-Fadel and Hashisho, 2001) have also been used to represent a composite vehicle fleet.

Although source tests can provide valuable information on chemical emissions, detailed PM speciation is not always possible at the microenvironmental level due to mass limitations. Traditionally, source apportionment studies have used longer sampling composites (e.g., monthly or even quarterly) to represent general trends in an urban airshed. Lower time resolution results in less variability in source signatures among different samples, potentially obscuring source impacts. The effect of sampling duration on resolving source types has been described previously for factor analysis methods (Lioy et al., 1989). Recent studies have reported organic concentrations at the lower mass concentrations typically seen at the microenvironmental level (Brinkman et al., 2009; Olson et al., 2008), but source apportionment efforts with lower PM concentrations will likely depend on a much larger percentage of samples that are near or below analytical detection limits.

The aim of this paper is to use wind regression results from semi-continuous pollutant and meteorological data to evaluate source impacts near roadways. The wind regression model Air Pollution Transport to Receptor (EPA APTR 1.0) was recently developed to estimate sector apportionment (% of a given pollutant from a specific wind speed and wind direction). Time-integrated filter samples were collected near a highway in Las Vegas, NV, and resulting organic
source marker concentrations were then compared to the sector apportionment results to test the validity of using wind regression to assess source impacts.

2. Methods

2.1. Site Description

Semi-continuous and time-integrated measurements were collected at Fyfe elementary school, which is adjacent to a highway in Las Vegas, NV (Figure 1). U.S. Highway 95 (US-95) is a limited access highway consisting of five lanes each direction and carrying an average traffic volume of approximately 175,000 vehicles/day during the sampling period. Both semi-continuous and time-integrated samples were collected approximately 18 m from the soundwall from January 5–28, 2008.

Semi-continuous data (5-min averaging time) were collected for the following pollutants: black carbon (BC) using an Aethalometer; CO using a Thermo-Electron Model 48i carbon monoxide analyzer; and NO\textsubscript{2} and NO\textsubscript{x} using a Thermo-Electric Model 42i Chemiluminescence NO/NO\textsubscript{2}/NO\textsubscript{x} monitor. In addition, meteorological measurements (wind speed, wind direction, and wind direction variability sigma theta) were collected using a RM Young AQ 5305-L monitor.

Time-integrated filter samples were collected daily from 5:00 to 9:00 AM, 9:00 to 11:00 AM, 11:00 AM to 5:00 PM, 5:00 PM to 5:00 AM Pacific Standard Time (PST); 24-h samples were collected daily starting at 5:00 AM. The intent of this sampling strategy was to enable inclusion of both sub-daily samples (consistent with commuting patterns) and daily samples (consistent with epidemiological studies and routine monitoring) for subsequent analyses. Based on results using wind regression analysis (described below), a subset of time-integrated filter samples (n = 27) were selected for subsequent laboratory analysis for organic source markers.
Traffic data were collected along US-95 at the Torrey Pines overcrossing (approximately 1 mile from Fyfe Elementary) using a Wavetronix SmartSensor HD model 125. The Wavetronix detectors use microwave radar to detect vehicles entering and exiting a detection zone, with a 1-minute resolution of speed and volume by lane; 1-minute data were aggregated to hourly values across all lanes. Data were internally consistent throughout the January 5–28 sampling period, e.g., with no sudden shifts in traffic counts that could indicate an instrument malfunction.

Additional measurements were also collected at Fyfe and three other nearby schools and are discussed elsewhere (Vedantham et al., 2011).

2.2. Sample Collection

PM$_{2.5}$ samples were collected at a flow rate of 68 m$^3$/hr using a high-volume sampler with a PM$_{2.5}$ cyclone inlet. Quartz filters (8” x 10”, Pall Life Sciences, TISSUQUARTZ 2500QAT-UP) were pre-baked in a muffle furnace (550 °C) for 12 hours. All filter samples were stored in a laboratory freezer (-80 °C) after sample collection.

2.3. Analytical Methods

Quartz filters were extracted in a solvent mixture of hexane, methanol, and dichloromethane (1:1:1) (GC$^2$ grade, Burdick and Jackson) using a Dionex ASE 200 Accelerated Solvent Extractor. Chromatographic grade nitrogen was used to concentrate samples to a final volume of 300 µL. A gas chromatograph (GC, Hewlett-Packard 6890N) and inert mass selective detector (MSD, HP 5973N) was used and chromatographic separation was completed using a 30-m, 0.25-mm i.d. with 0.25-µm film thickness capillary column (DB-5MS, Part No. 122-5532, J&W Scientific). The GC was operated as follows: injector temperature of 315 °C; column flow of 1 mL/min; pressure pulse of 25 psi for 0.5 minutes; initial GC oven temperature initially set at
35 °C, ramped at 20 °C/min until reaching 160 °C, ramped at 2 °C/min until reaching 315 °C. The MSD was operated under selective ion monitoring (SIM) mode. Recovery, precision, limits of detection and quantitation, blank levels, calibration linearity, and agreement with certified reference materials for compounds analyzed in this study are given elsewhere (Turlington et al., 2010).

External calibration curves (5–500 pg/µL) using authentic standards (Absolute Standards, Accustandards, and Chiron) were determined for all analytes. Deuterated surrogate standards (50–200 pg/µL) (n-C26-d54, n-C30-d62, n-C36-d74, chrysene-d12, benz(b)fluoranthene-d12, indeno(1,2,3-cd)pyrene-d12 and coronene-d12) and deuterated internal standards (1000 pg/µL) (n-C20-d42, n-C25-d52, n-C28-d58, n-C32-d66, benz(a)anthracene-d12, benzo(e)pyrene-d12, dibenzo(a,h)anthracene-d14 and dibenzo(a,i)pyrene-d14) were used for all samples.

2.4. Wind Regression Analysis

Wind regression analysis has been recently developed as a means of using semi-continuous (e.g., sub-hourly time resolution) meteorological and pollutant data to estimate the percent of a given pollutant originating from a specific wind sector (hereafter referred to as sector apportionment). All sector apportionment data used in this study were grouped using the same sampling intervals as the time-integrated filter samples. For this paper, model results (S in Equation 3 below) were developed using the same sampling periods as filter samples.

Wind regression analysis was performed using the Air Pollution Transport to Receptor model (EPA APTR 1.0) (U.S. Environmental Protection Agency, 2009), which is described in detail elsewhere (Vedantham et al., 2011; Henry et al., 2009). Briefly, this analysis used the local scale wind analysis with the Sustained Wind Incidence Method (SWIM). The SWIM is similar to Non-parametric Wind Regression (NWR) as described by Henry et al. (Henry et al., 2009), a
kernel smoothing method for apportioning mean pollutant concentrations using highly time-resolved meteorological and concentration data. Using wind speed, wind direction and ambient concentration, the method provides a likelihood type output. The SWIM for estimating the expected concentration of a pollutant for each wind direction and wind speed pair \((\theta, u)\) is given by:

\[
E(C|\theta, u) = \frac{\sum_{i=1}^{N} K_1 \left( \frac{\theta - W_i}{\sigma} \right) K_2 \left( \frac{u - U_i}{h} \right) C_i \ W_i}{\sum_{i=1}^{N} K_1 \left( \frac{\theta - W_i}{\sigma} \right) K_2 \left( \frac{u - U_i}{h} \right)}
\]

where \(C_i\), \(U_i\), and \(W_i\) are the observed concentration of a particular pollutant, resultant wind speed and directional standard deviation, respectively, for the \(i\)-th observation at time \(t_i\); \(N\) is the total number of observations; \(K_1\) and \(K_2\) are smoothing kernels; \(\theta\) is the wind direction; \(u\) is the wind speed; and \(\sigma\) and \(h\) are smoothing parameters for wind direction and wind speeds, respectively. Smoothing parameters have been set at values of 20º and 2 m/s for this study.

The dimensionless weighting term is defined in the SWIM method by the following:

\[
W_i = \frac{C_i U_i}{\max(C_i U_i)} \overline{(\sigma \theta_i)} \sigma \theta_i
\]

The first part of Equation 2 describes wind flux while the second part describes the effect of sustained wind. This weighting is intended to be proportional to wind fluxes \((C_i U_i)\) and inversely proportional to variation in wind direction \((\sigma \theta_i)\). The term \(\overline{(\sigma \theta_i)}\) was calculated using the median of standard deviations in wind direction, a reasonable assumption given that a Gaussian type distribution was observed for the wind direction data in this study.

The conditional probability of pollutant concentration (Equation 1) is then weighted by wind frequency using a joint probability of wind speed and wind direction, resulting in the
following expression the mean value of the pollutant concentration associated with winds from
the sector defined by the intervals $U$ and $\theta$ (hereafter referred to as sector apportionment):

$$ S(\Theta, U) = \int_{U} \int_{\Theta} f(\theta, u) E(C|\theta, u) \, d\theta \, du $$

(3)

where the joint probability of wind speed and wind direction ($f$) is calculated using a kernel
density estimate:

$$ f(\theta, u) = \frac{1}{N\sigma_h} \sum_{i=1}^{N} K_1 \left( \frac{\theta - W_i}{\sigma} \right) K_2 \left( \frac{u - U_i}{h} \right) $$

(4)

The uncertainty is the sector apportionment $S(\Theta, U)$ is estimated as:

$$ \text{var}(S(\Theta, U)) = \sum_{\theta_i, \theta_j} \sum_{u_i, u_j} s(\theta_k, u_j) \, s(\theta_m, u_n) \exp \left( -\frac{1}{2} \left( \frac{\theta_k - \theta_m}{\sigma} \right)^2 \right) $$

(5)

Additional details on smoothing functions, weighting procedures, and error estimation are
given elsewhere (Vedantham et al., 2011; Henry et al., 2009).

2.5. Data Analysis

Data processing and descriptive statistics were performed using SAS v.9.1 (SAS Institute, Cary, NC). In addition to APTR wind regression results, meteorological data (wind speed, wind direction) and semi-continuous (5-min) chemical data were each evaluated for data quality. Statistical distributions were used to describe these data for all sampling intervals used in this study, including inter-quartile range, differences between 95th and 5th percentiles, and differences between 90th and 10th percentiles. In addition, the root mean square error associated with exponential smoothing (Brown, 1961; Brown and Meyer, 1961) of each sample interval was used as an indicator of time series discontinuities. These analyses were completed so that
integrated samples selected for laboratory analysis represented a range of wind speeds and wind directions.

3. Results and Discussion

3.1. Assessing Source Impacts using the APTR Wind Regression Model.

Example model results using APTR are shown in Figure 2, where estimated concentrations of CO are given as a function of wind direction. The center of the circle indicates the receptor location and the radial length of each sector represents the estimated contribution of CO from that range of wind directions (i.e., \( E(C|\theta) \) in Equation 1, integrated across all wind speeds from that particular range of wind directions). Figure 2 illustrates two different sector apportionment results, one that is more uniformly distributed and one that is predominately from the southwest. Because US-95 runs in a predominately east-west direction near the sampling location, sectors used throughout this study were separated by cardinal wind directions (0°, 90°, 180°, 270°). In other words, source sectors were estimated using Equation 3, integrated across all wind speeds, and integrated from wind directions 0-90°, 90-180°, 180-270°, and 270-360°.

Similar results were seen from the other semi-continuous pollutants measured during this study (BC, NO\(_2\), and NO\(_x\)), where sector apportionment results were well-correlated between pollutants (\( r^2 > 0.9 \) between all compounds, see Figures S1-S3, Supporting Information).

An objective to this paper is determining whether APTR results can be used to predict sampling intervals that are more indicative of source impacts from the roadway. An example using all samples that were analyzed for organic source markers is given in Figure 3, where sector apportionment concentrations for CO is given as a function of benzo(g,h,i)perylene concentration. Previous research has shown that benzo(g,h,i)perylene is an indicator of traffic contributions of PAHs (Nielsen, 1996) and along with indeno(1,2,3-cd)pyrene and hopanes have
a strong influence on vehicle exhaust source apportionment (Fujita et al., 2007). Concentrations of all organic source markers measured in this study (10 PAHs and 3 hopanes), average BC, and average CO are listed in Table 1. The weak relationships between high-time resolution data and time-integrated samples (e.g., \( r^2 < 0.2 \) between hopane and average BC and between hopane and average CO) illustrates the need for better approaches to identify source impacts from time-integrated samples. Figure 3 shows results for all samples analyzed for organic markers as well as for samples classified as source-impacted samples. An initial designation for source-impacted samples was constrained to samples having >30% CO sector apportionment from the south and errors in the estimated sector apportionment <30%. An example of a source-impacted sample is given in Figure 3b. Although these criteria for source impacts were prescribed in this example, Figure 3 clearly shows an improved indication of source impacts using the APTR sector apportionment results. Also noteworthy is that even for source-impacted samples with lower sector CO concentrations (e.g., <50% CO from the south), these samples have correspondingly lower concentrations of benzo(g,h,i)perylene. In other words, source identification using APTR is not limited to samples where predicted concentrations are exclusively from a single sector, meaning that the approach can be used for typical time-integrated samples that may consist of a variety of wind speeds and wind directions. In addition, error estimates \( (2\sigma/\mu) \) from the source-impacted samples (average error = 20%) were less than those for the remaining samples (48%). Utilizing higher frequency data to better resolve time-integrated samples have also been shown using multiple time resolution models (e.g., Zhou et al., 2004; Ogulei et al., 2005).

This example is extended for all sampling periods included in this study as shown in Figure 4. The BC sector from the south (calculated using APTR) and the average BC concentration is shown by sampling date. Sector apportionment results for BC range from 0 to 95%, although for most samples days results are between 30 and 60%. The poor relationship
shown in Figure 4 between sector apportionment and average BC further illustrates that concentrations alone are not necessarily indicative of near-source impacts.

Further evidence validating the use of APTR to attribute source impacts is shown in Figure 5 as a function of total traffic counts averaged over each sampling period. Figure 5a shows total hopane concentration as a function of traffic counts, while Figure 5b shows the predicted hopane concentration from the south (i.e., sector contribution of hopane predicted using APTR). Data in Figure 5 were limited to samples collected during morning rush hour (9:00-11:00 AM) as they were expected to be more indicative of commuter highway traffic. While a detailed statistical analysis of the data shown in Figure 5 is not warranted given the smaller number of samples, results in Figure 5b clearly indicate that sector apportionment results for a commonly-used organic source marker (hopane) is well-correlated with traffic data ($r^2 = 0.72$).


The difference in error estimates between source and non-source samples of CO suggests that APTR model diagnostics could be a useful screening tool for source effects. Among semi-continuous wind speed, wind direction, and CO data used in this study, wind speed appeared to be the best indicator of sector apportionment uncertainty. This pattern is shown graphically in Figure 6, where error percent estimates from the south sector are plotted as a function of median wind speed for all sampling intervals where filter samples were collected. The highest estimated errors were associated with lower median speeds. In fact, all error estimates >30% occurred between median wind speeds of approximately 0.7–1.1 m/s. Similar patterns were seen for statistical distribution parameters (interquartile range, difference between 95th and 5th percentile, etc.), where the highest errors were associated with the lowest wind speeds.
This pattern where lower wind speeds have the highest relative error is especially pertinent as previous research (e.g., Charron and Harrison, 2005) has reported higher PM$_{2.5}$ concentrations at lower wind speeds. A pattern of decreasing BC with increasing wind speed has also been noted previously for urban areas (e.g., Wang et al., 2011). Thus, quality of the sector apportionment estimate and not just the highest mass level is an important consideration for assessing source impacts. This pattern may also be a result of higher relative variability often present at lower wind speeds. For example, the average relative error in wind speeds ($2\sigma/\mu$) using the sample intervals in this study was 89% for median wind speeds <1.5 m/s, while all remaining samples had an average error of 69%.

4. Conclusions

Detailed chemical information is often needed to gain better understanding of PM sources in an urban airshed. While numerous chemical marker species have been proposed for various sources of PM (e.g., motor vehicle, wood smoke), measurement of those same chemicals at the microenvironmental level can be difficult owing to greater uncertainty associated with lower mass concentrations. Increased sample mass achieved with coarser time resolution can minimize this limitation, but may result in the inability to resolve transient sources that vary at finer time scales (e.g., motor vehicles). The results in this study demonstrated the use of the wind regression model APTR to identify source-impacted time intervals when compared with filter samples analyzed for organic source markers using GC-MS. Given both the time and cost constraints of completing detailed PM speciation for organic source markers, focused laboratory analysis of source-impacted samples should be a consideration for future study design. Thus, identifying samples most indicative of source impacts using highly time-resolved data can be an effective approach in bridging the gap from source profiles to microenvironmental samples.
Acknowledgments

The authors would like to thank Nevada Department of Transportation (NDOT) for funding the black carbon measurements and monitoring shelter at Fyfe as part of the US-95 Settlement Agreement. This study could not have been completed without extensive cooperation of the Clark County School District; we especially recognize Paul Gerner as well as the staff and teachers of the District and individual schools. David Vaughn and Joey Landreneau (STI) assisted with sampler setup and Jeff Collett and Taehyoung Lee (Colorado State University) assisted with sample preparation and storage.

Disclaimer

The United States Environmental Protection Agency funded and managed the research described here under contract no. 68-W-04-005 to Lockheed Martin. It has been subjected to Agency’s administrative review and approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Supporting information available

Information on sector apportionment between CO and BC, CO and NO₂, CO and NOₓ, organic source marker concentrations for samples analyzed in this study. This material is available free of charge via the Internet at http://www.sciencedirect.com.

References


Charron, A., Harrison, R.M., 2005. Fine (PM_{2.5}) and coarse (PM_{2.5-10}) particulate matter on a heavily trafficked London highway: Sources and processes. Environmental Science and Technology 39, 7768–7776.


McConnell, R., Berhane, K., Yao, L., Jerrett, M., Lurmann, F., Gilliland, F., Kunzli, N.,
childhood asthma. Environmental Health Perspectives 114, 766–772.

Houston tunnel during the Texas Air Quality Study 2000. Atmospheric Environment 38,
3363–3372.

Nielsen, T., 1996. Traffic contribution of polycyclic aromatic hydrocarbons in the center of a
large city. Atmospheric Environment 30, 3481–3490.

for multiple time resolved species: The Baltimore supersite. Atmospheric Environment
39, 3751–3762.

Indoor and outdoor concentrations of organic and inorganic molecular markers: Source
apportionment of PM$_{2.5}$ using low-volume samples. Atmospheric Environment 42, 1742–
1751.

Peters, A., von Klot, S., Heier, M., Trentinaglia, I., Hormann, A., Wichmann, H.E., Lowel, H.,
2004. Exposure to traffic and the onset of myocardial infarction. New England Journal of
Medicine 351, 1721–1730.

phase air pollutants using organic compounds as tracers. Environmental Science and
Technology 34, 1821–1832.


Figure captions

Figure 1. Map of sampling site.

Figure 2. Example SWIM model results showing (a) CO sector apportionment with contributions from multiple directions (time period January 11–17) and (b) CO sector apportionment mainly from the southwest indicating a source-impacted sample (24-hr sample on January 6).

Figure 3. Sector apportionment of CO from the south of as a function of benzo(g,h,i)perylene concentration for (a) all organic source marker samples and (b) all source-impacted samples.

Figure 4. Sector apportionment of BC from the south and average BC concentration as a function of sample date; all data in time series are for a 24-h sampling period.

Figure 5. Hopane concentrations as a function of average hourly traffic counts for (a) total hopane concentration and (b) predicted hopane concentration from the south. Data were restricted to morning rush hour samples.

Figure 6. Percent error in CO sector apportionment from the south (2σ/µ) as a function of median wind speed. Percent errors and median wind speeds are both plotted as five-point moving averages.