Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-112-AC2, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Traffic-related air pollution near roadways: discerning local impacts from background" *by* Nathan Hilker et al.

Nathan Hilker et al.

nathan.hilker@mail.utoronto.ca

Received and published: 1 July 2019

Response to Anonymous Referee 2

The authors are grateful for the useful comments and criticisms. Referee comments are shown in *italicized* text below, with author's replies and actions following each comment.

1) The first paragraph in Introduction about exposure is not that relevant to the rest of the paper, so would better focus on the topic of traffic-related pollutants close to the road.

Reply: Being that this is a methodology-focussed paper, and the primary interest is in better understanding traffic's contribution to traffic-related pollutant concentration in the near-road environment, we tend to agree with this comment. While the motivation for better understanding this is in part exposure-driven, it is not the main topic of the paper.

Action: The introduction will be altered to reduce emphasis on exposure/health effects and more on the methodology of characterizing near-road air pollution.

2) In L174, why the calculated average difference is expected to converge the true average difference between sites? Are the differences between sites normally distributed? Is this convergence non-trivial and not simply a property of averages or central limit theory?

Reply: Perhaps the wording in the manuscript is unnecessary and/or confusing in this section. The concept of random sampling and error theory has been addressed by others in the context of air quality monitoring (Xu et al., 2007), where the amount of data needed for a sample mean to converge to a "true" mean has been understood. The point to this statement was to imply a trivial convergence: as the number of samples increases so does the certainty in the mean of the difference (as a result of more variability due to seasonal effects, meteorology, etc.), similar to central limit theory as is pointed out.

Action: Wording of the manuscript in this section will be altered for greater clarity.

3) In L190, since the authors have realized the downwind and upwind scenarios may encompass different time frames and may influence the results, why not do some *statistical tests? It seems important to the final outcomes.*

Reply: Thank you for pointing out this lapse in analysis. Indeed, if downwind periods occur largely at night time, for example, then its average will be biased low. This is a relatively straight forward analysis and will be implemented in a revised version of the manuscript. As an example, refer to Fig. 1 in this document. At NR-TOR-1 downwind data are mostly uniformly distributed w.r.t. hour of day. Upwind conditions, however, are more likely to occur in the afternoon compared with the morning, meaning the upwind average will be defined by afternoon pollutant concentrations more so than morning concentration.
This is a potential issue as certain times of day will influence the mean values more so than others. One means of addressing this is to randomly sample an equivalent number of hours for each hour of day and compare the resulting distribution with the case in which all data is used to see if they are significantly different. As a proof of concept, observe the distributions of UFP concentrations at NR-TOR-1 in Figs. 2-3, which were generated by randomly sampling an equivalent number of points from each hour of the day (this number was determined from the minimum values in Fig. 1). The resulting downwind and upwind averages were 5.64E+4 $\pm$ 240 [cm-3] and 1.46E+4 $\pm$ 260 [cm-3] ($\pm$ 1σ), respectively (compare with values in Table 3 of the manuscript: 5.70E+4 and 1.53E+4, respectively). More rigorous statistical analyses will involve tests on whether these bootstrapped populations are significantly different than those reported in Table 3 of the manuscript.

Action: Additional analyses will be performed to create diurnal trends in frequency of downwind/upwind hours, as well as include rose wind plots for each near-road site (suggested also in response to RC1). Further, statistical tests will be performed between data sampled from a uniform distribution w.r.t. time of day vs. all data used.

4) In L238-274, the method3 was not explained properly.

Reply: While the algorithm has been described in detail already in Wang et al. (2018), we agree that a more mathematical description of the algorithm is necessary.

Action: The manuscript will be edited to include a more mathematically rigorous definition of the background-subtraction algorithm detailed in this section.

Firstly, ‘Time-series analysis’ seems too general and may not be a good subtitle here and in the rest part of the MS. It often implies decomposition and forecasting.

Reply: Agreed. Time-series analysis does seem too vague for this section, especially considering only one algorithm is really explored for signal deconvolution. Perhaps more appropriate is “baseline estimation” or “moving minimum” as you have suggested.

Action: The subtitle of this section will be changed to be more descriptive.

Secondly, the authors talked about the frequency of signals very often in the first two paragraphs (L238-254) and allude to the wavelet decomposition algorithm used by Sabaliauskas (2014) as similar to their method. But I think this is not quite right and misleading. What I expected after the description is a frequency analysis, but method 3 is approximately a ‘moving minimum’ baseline algorithm. As an example of signal processing and a spatial frequency domain in the road-environment can be seen in Xing and Brimblecombe (2019). Although wavelet analysis can also be used to exact baselines as shown by Liland et al. (2010), the underlying theory is different. There have been many baseline algorithms in Liland et al review (2010), method3 doesn’t seem more accurate although it may be efficient. Besides, could the authors validate the extent to which the baselines derived using this algorithm represent the *background?*

Reply: Thank you for referencing these two articles. The Liland et al. (2010) review highlights many different algorithms intended for background subtraction of spectroscopic data that could be applied to pollutant time series for a similar effect.
We agree that too much emphasis is placed on frequency analysis and, in particular, the discrete wavelet transform (which is not explored in this paper but is described as an analogue for what is instead used). While the notion of correlation between different spatial scales and signal frequencies is still relevant, we agree that the wording as it is does seem to imply that the results would contain some frequency-domain analysis, which was not the intention (these results would certainly be interesting but do not necessarily align with the scope of this manuscript).
In response to the last point, a validation of how well the derived baseline represents the urban background was performed already between NR-TOR-2 and BG-TOR-2 by Wang et al. (2018). We admit that this validation may need to be extended to NR-TOR-1 and NR-VAN, however. The intent behind comparing methods 1 and 3 was to validate that the differences as estimated using the background derived using method 3 was indeed similar to the differences measured between near-road and urban background stations pairs.

Action: Broader references to background-subtraction algorithms utilized by other atmospheric scientists will be included in the introduction, with emphasis being shifted away from frequency-domain analysis. Furthermore, the supplementary will be updated to include a table comparing the backgrounds determined by methods 1 and 3.

Thirdly, many details about method3 were not shown in the paper but presented in Wang et al., 2018. I understand this is a method in the published paper, but since this is a journal about measurement techniques, I think more details should be provided, especially the setting of the time window. Wang et al. (2018) used 8h, but

is it appropriate here since a new station near a highway (NR-TOR-1) is added in this MS? As mentioned in L243-244, characteristics of emission sources determine the frequency of signals. Thus should the time window for a station near highway may need to be different from that near streets, intersections or bus stops, which have their own frequency components (probably higher)? In addition, would different pollutant species require a different setting of window, especially a secondary pollutant such as ozone?

Reply: In response to the first point made regarding choice of time window and different receptors, while different near-road environments will inevitably affect higher frequency signals in a pollutant time-series (due to closer source proximities) as mentioned, the choice of time window was intended to be more of a reflection of the spatial scale differentiation between what is considered "background" and "local". As a first-order approximation, consider a primary pollutant affected only by physical dispersion. If this rate of dispersion is proportional to wind speed, then the pollutant's length of influence would in some way be proportional to:

$$d \approx u \cdot t$$

where 'u' is wind speed and 't' is time since emission. Then, for a wind speed of 1.0 [m/s] and a time of 8 [hrs], for example, the pollutant's range of influence would be approximately 30 [km]. Thus, an argument could be made that utilizing a time window of 8 [hrs] in the background-subtraction algorithm is effectively distinguishing between emissions from sources within approximately 30 [km] of the receptor (those originating from nearest roadways will have the greatest influence on the signal) and those from outside of 30 [km]. Of course, this is a gross approximation and is likely not physically accurate, but it emphasizes the spatiotemporal relationship between signal frequency (choice of time window, which is related to signal cutoff frequency) and source distance. Moreover, given that these measurements were made within urban

regions with relative homogeneous distributions of roads, averaging the background over a smaller or larger spatial area, should not make much difference.
Regarding the second point made (time window and different pollutants): you have raised an important issue, and we believe further analysis is necessary to support the time windows used in this study. A sensitivity analysis showing the distribution of measured urban background concentrations vs. the distribution of derived backgrounds as a function of time window, pollutant, and site would better support the time windows used in this study.

Action: A sensitivity analysis will be performed for each near-road/urban background station pair to compare measured urban background concentrations with derived baseline concentrations as a function of time window. This will further support the time windows used in this study and will determine whether different time windows for different pollutants is justified.

5) In L380-384, I don't understand why method1 and method3 are better as they both provide lower difference values. In my opinion, method3 has more disadvantages than method2, because the outcomes highly depend on the choice of time window and it's hard to determine if the baselines represent a real background.

Reply: The choice of wording here is perhaps inappropriate then, as the intention was not to claim one method being "better" than the other, but to highlight the advantages and disadvantages of each along with the fundamental differences between them. For example, methods 1 and 3 may be more appropriate for understanding traffic's influence to a 24 hour-averaged exposure from an epidemiological perspective (as all meteorological conditions are considered), whereas method2 may be better for extracting data whose impact from local traffic is greatest for use in fleet-averaged emission factor calculations, for example.

Action: This section will be reworded to include greater detail of the advantages and disadvantages of each method while retaining a tone of neutrality.

As I understand it, method2 only used part of the data and clearly gave the largest difference between roadside and background concentrations. While the other two methods used the data even when the roadside stations experience background concentrations (e.g. under upwind conditions). Literally, the output from method2 over-predicts average local concentrations (L131 in supplementary information). If the aim of this MS is determining the averaged concentration difference, method2 should be revised, otherwise, the difference between method2 and method13 is just caused by the difference in the methods.

Reply: We fully agree that what method 2 is measuring is inherently different from the other two methods. We will revise the text to better emphasise this important point. However, quantifying this difference is still of importance so that others might better understand the extent to which they differ. Moreover, method 2 provides an upper limit of the impact of the road on exposure. Section S3 proposes an alternative methodology for utilizing meteorological data that falls in line better with methods 1 and 3.

6) L410-412, it seems the increase of pollutant concentrations under downwind conditions compared to upwind conditions is a main finding in this MS (as also mentioned in abstract). Could the authors provide the factors for each station and pollutant species? Theoretically, this factor should be a function of distance between source and receptor, wind speed, eddy diffusivity etc. Is it possible to add some tests about this?

Reply: In the manuscript we have chosen to express this as the ratio of the local portion of the upwind and downwind concentration to the average value (Figure 4). This inherently makes more physical sense to use than directly comparing the ratio of the downwind to upwind concentrations, given that they both contain a "background" that is not related to the road. Values can be calculated and reported for each site and pollutant as a supplement to Figure 4, which just shows an agglomerated average for all species. Hypothetically, if these primary pollutants disperse similarly in the near-road regime and are not significantly impacted by secondary processes in the time it takes for them to be detected, then these curves should be similar between species. We agree that differences in dispersion between gas and particle-phase pollutants, and post-tailpipe transformation (e.g. UFP dynamics), for example, may lead to differences between pollutants.
The reason these trends were analysed with respect to normalized local concentrations was so that they would be invariant with respect to source strength. I.e., the area above and below unity for each curve are equivalent thanks to the property:

$$\int_0^N \left(\frac{x(\theta)}{\bar{x}} - 1 \right) d\theta = 0$$

However, as you have mentioned, the shape will be impacted by distance of receptor to source, wind speed, eddy diffusivity, receptor height, atmospheric stability, etc. While we agree that the siting of these near-road stations along with meteorological conditions will have a theoretical impact on these data, it is out of the scope of this manuscript (the focus of which is a comparison of background subtraction methodologies) to attempt to model these results in a theoretical manner.

Action: Graphs similar to Figure 4 for each pollutant will be added to the SI, along with a table summarizing downwind/upwind ratios.

7) In L484, why is method3 accurate and robust? Is it because the outputs agreed with those derived from method1?

Reply: It is deemed accurate because it agrees with those values derived from method 1 which is the closest estimate to a real background. In terms of robustness, it appeared to be applicable across all near-road monitoring locations and data did not need to be filtered by meteorology, for example. How robust this algorithm is, exactly, will be available following the aforementioned time window sensitivity analysis.

Action: See above response regarding time window sensitivity analysis.

References:

Wang, J. M., Jeong, C-H., Hilker, N., Shairsingh, K. K., Healy, R. M., Sofowote, U., Debosz, J., Su, Y., McGaughey, M., Doerksen, G., Munoz, T., White, L., Herod, D., and Evans, G. J.: Near-Road Air Pollutant Measurements: Accounting for Inter-Site Variability Using Emission Factors, Environ. Sci. Technol., 52, 9495-9504, doi:10.1021/acs.est.8b01914, 2018.

Xu, X., Brook, J. R., and Guo, Y.: A Statistical Assessment of Saturation and Mobile Sampling Strategies to Estimate Long-Term Average Concentrations across Urban Areast, J. Air Waste. Manag. Assoc., 57, 1396-1406, doi:10.3155/1047-3289.57.11.1396, 2007.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-112, 2019.

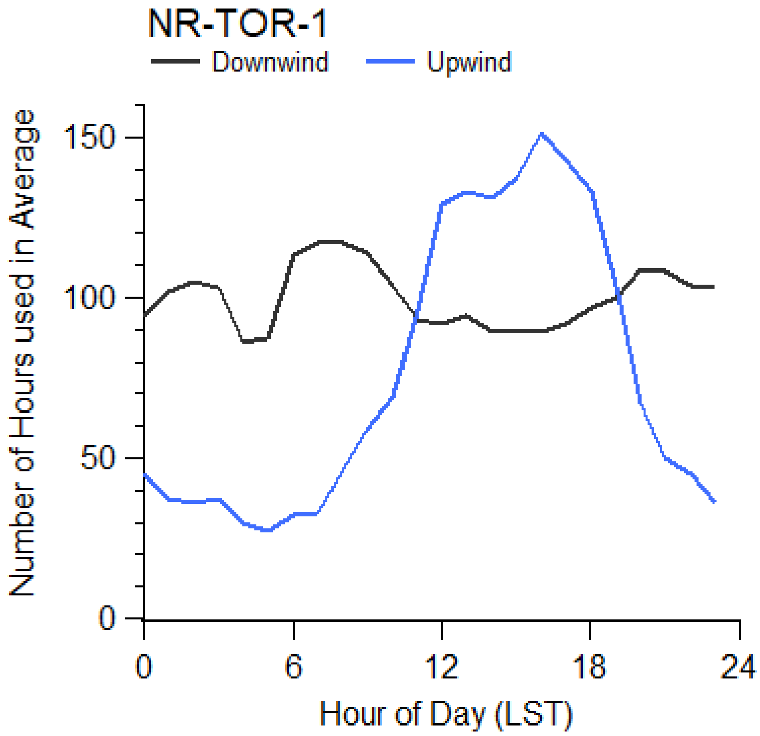

Fig. 1. Number of hours in which data were sampled downwind and upwind of Highway 401 at NR-TOR-1, aggregated by hour of day.

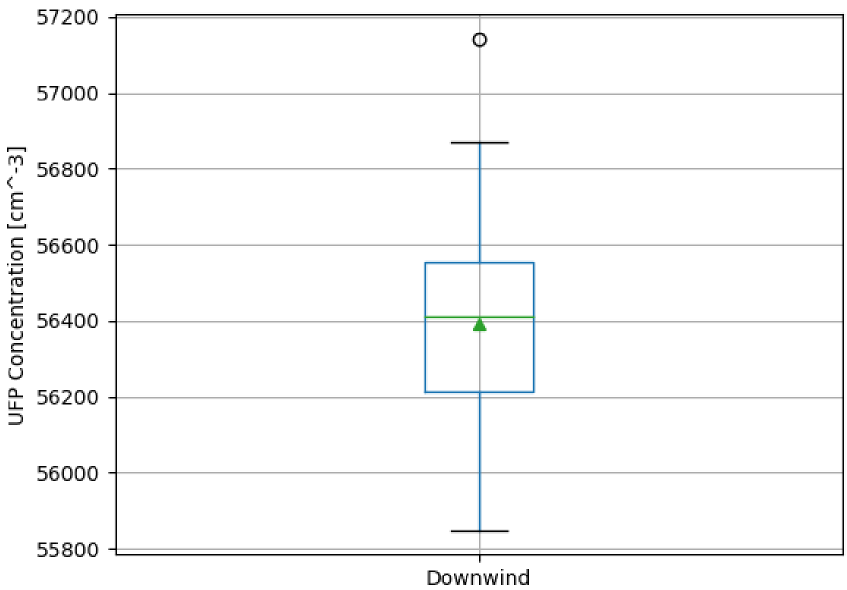

Fig. 2. Distribution of downwind UFP concentrations at NR-TOR-1, generated by bootstrapping (N = 100) an equivalent number of hours from each hour of day.

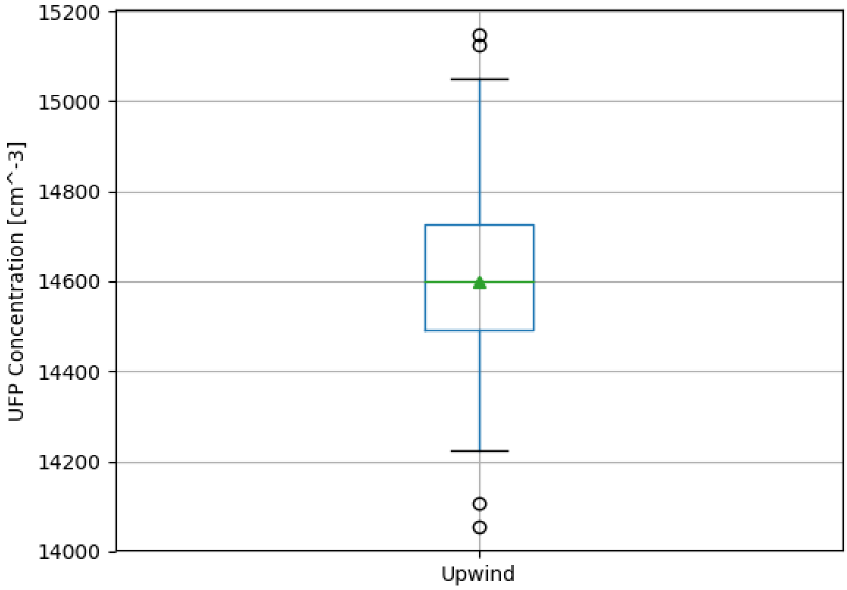

Fig. 3. Distribution of upwind UFP concentrations at NR-TOR-1, generated by bootstrapping (N = 100) an equivalent number of hours from each hour of day.