Title: Elucidating multipollutant exposure across a complex metropolitan area by systematic deployment of a mobile laboratory

Ilan Levy, Cristian Mihele, Gang Lu, Julie Narayan, Nathan Hilker and Jeffrey R. Brook

We thank the reviewers for the careful review and valuable comments. We are submitting the revised manuscript after carefully considering and addressing both reviews.

We have made extensive changes to the manuscript after noting that both reviewers were unclear of the main goals of the paper and the context of the results we present. Clearly, in an attempt to keep a relatively long paper from getting longer we did not include as much background and motivation for the work in the first submission as was necessary. We therefore have completely re-written the Introduction to hopefully made the specific objectives and novel aspects of the paper clearer. Additionally, as the first reviewer noted, the paper did not correctly recognize the current exposure literature to provide sufficient context for our study. This reviewer also commented that some naïve statements were made about exposure. Thus, the Introduction section was expanded to better recognize the exposure issues of focus in this paper and the past literature of relevance. We now have clarified that, building from limitations of previous work, we are addressing the questions of (1) given its known limitations, how well can a mobile lab be deployed to offer detailed, representative annual average information on spatial gradients that can inform exposure model evaluation and development? and (2) what are the characteristics of average multipollutant, multi-scale spatial patterns and exposures, that a large mobile lab like CRUISER is uniquely able to explore?

The second reviewer expressed a concern that the paper did not offer sufficient novel results, or at least that such novel results were not clearly highlighted, especially given that a key set of results were published elsewhere. In addition, to clarifying the two questions we state above, which have not previously been explored in a rigorous manner, we have sharpened the results by taking out unnecessary text and adding a new section about near-highway gradients. This new section is in response to the second reviewer's interest in more focus on the information previously covered in Section 3.3 and his/her request for more novel results of broad relevance to other researchers. While with our clarified objectives we expect that relevance to others will be clearer, the new insight we offer on near-highway gradients is, we believe, a unique contribution on this topic given the wider range of days CRUISER measurements were taken, the larger distance away that was considered and the number of pollutants measured simultaneously. Also, to make these unique data useful to others we have provided a new table (SM-A1) that provides 'scale-up factors'. These can potentially be applied to currently available urban background monitoring data (in Montreal and elsewhere) to estimate average concentrations closer to busy highways similar to that studied in Montreal. We also reduced the Discussion section to further sharpen the paper and have completely re-written the Conclusions to highlight what about the most significant contributions have come from our mobile lab study in Montreal. We feel that these changes have significantly improved the paper and hope that the reviewers will agree.

Our point-by-point responses to the reviewers are given in two separate documents where we highlight each of their key concerns. is given below in blue text.

Dear authors,

You reported on multiple atmospheric pollutant measurements across the city of Montreal using a mobile laboratory. The material is original and of sufficient new interest. The subject and scope of the paper clearly fall within the scope of ACP.

After balancing the ups and downs of the manuscript, however, my opinion on this manuscript is still equivocal.

On the one hand, you provide a very extensive body of new data, showcasing multi- parameter mobile measurements across an urban atmosphere. The data set collected by the authors is, to my knowledge, unusually comprehensive and unique. I am not aware of any similar data set that has been collected across an urban landscape. The data was processed with care, and is certainly useful to assess the various aspects of complex pollutant exposure across a wide stretch of urban landscape. To sum up, the experimental material that has been generated in this work is a real achievement.

On the other hand, I see several hurdles that need to overcome in order to make this an ACP publication.

1) Overall, the paper remains rather descriptive. It lacks a certain scientific rigor, i.e. in the way of condensing the descriptive results into something that is of a wider relevance to the scientific community.

Part of the problem is that the results gained from in Sects. 3 and 4 do not really make it into the Conclusions section.

Another part of the problem is that the results sections themselves are not well balanced. Sect. 3.1, for example, is rather lengthy but yields only some modest results. Sect. 3.2 is interesting, but the discussion remains too much on a descriptive and anecdotal level. I like Sect. 3.3 very much, and that section, for example, could be further expanded.

In addition, I need to state that you did already extract a major part of their campaign results, notably the air pollutant cross-correlations, into another journal article (Levy et al., 2013). Spreading results over several journal articles can, in general, be a good idea. In the current situation, however, I feel that "picking out the cherry" of the data analysis did not particularly help this ACPD article. This article looks a bit like a loose collection of the less intuitive and less attractive results remaining from the entire work in Montreal.

> We have taken these comments into consideration and made the necessary changes to the manuscript. In its revised version, the objectives are clarified pointing out the importance of the questions we sought to answer in this paper. Firstly, was to evaluate a methodology for using a mobile measurement campaign for a quantitative assessment of representative long term average concentrations at many urban locations so as to inform questions of multipollutant chronic exposure patterns for possible future use in exposure model development. Secondly, was to compare multipollutant average concentrations among a variety of locations within the city offering some new insights regarding spatial patterns and of multiple pollutants gradients, with the more diverse and quantitative measurements afforded by use of the mobile lab. These are now described more clearly in the manuscript and our response to the other reviewer so that readers will ultimately be able come away with more concrete and hopefully more-widely applicable conclusions from the paper.

We have also revised sections 3.1-3, and added another sub-section with further results (sections 3.4 "near road gradients of multipollutants"). We now present new evidence about the distance of effect of emissions from a major highway and show that some pollutants (NO_2, CO and HOA) remain elevated above background up to 700 meters from the road, a result that provides further evidence for why health effects are often associated with NO_2 and hence why NO_2 continues to be a good overall indictor or proxy for intra-urban exposure patterns. We believe these changes also make the paper more relevant and interesting to the scientific community.
With regard to the EHP publication (Levy et al., 2013), given the richness of the data collected and the extent of the analysis performed, even two journal papers are hardly enough to cover all the results we have to show (hence the length of this paper). Our intention was not to put the best results in Levy et al., 2013 and the rest in this ACP paper, but rather to choose a distinct "theme" for each paper. While the EHP paper is focused on multi-pollutant statistical correlations, here we focus on the representativeness of mobile measurements for assessing long-term spatial patterns relevant to chronic exposure using our deployment approach and give some quantitative evaluation of the neighbourhood scale variability of air pollutants among different, common micro-environments, now also including near roadway gradients. While this variability is well known as general knowledge, it is yet poorly quantified, especially when comparing multiple pollutants. Certainly, we hope that others in the scientific community interested in learning more about multipollutant concentration patterns across cities in the context of chronic exposures will read both papers as they are intended to be complementary.

2) The Introduction part gave a fuzzy impression to me. It lacks overview and clarity. I will give some detailed suggestions to improve this below.

The Introduction was re-written following the suggestions below and according to points we raise above and in the response to the other reviewer.

3) The Conclusions section contains very little in terms of novel results. This is a major point that needs to be fixed.

The Conclusions section was completely rewritten except for the last paragraph. The key results are highlighted and readers are better able to appreciate these results given our revisions to the Introduction.

4) There are certain technical issues (processing methods) that need to be clarified. I see a particular problem with the wind speed correction to the PM measurements.

We addressed the issues of UFP and the speed corrections in the manuscript and in our detailed answers below.

5) Presentation of the geographical setting and the data can be improved.

We have made changes to Figure 1 as suggested below. We trust it is now much improved.

As a whole, publication in ACP seems possible once these (major) issues have been addressed.

Specific comments

1) Overall compilation and organization of results

As stated above, the paper provides many descriptive results, but little in terms of condensed results that may be remembered by a reader. The major statement lasting in my memory, after reading the article for quite a while, is that pollution gradients and time development in the urban atmosphere are overwhelmingly complex. This is something that a lot of people know or could have guessed already.

> Our goal in this paper was not only to demonstrate how complex urban air pollution is, which is an unavoidable outcome, but to also present an alternative method for compiling measurements taken by a mobile lab into a useful dataset that can provide insight into multipollutant chronic exposure thus helping to inform epidemiological studies. This information is useful in its own right and we hope that this is more evident in the revised manuscript. The information in this paper also lays the foundation for subsequent analyses of the data such as we have published in *Environmental Health Perspectives* and *Environmental Manager* and that we continue to work on for additional publications focusing on exposure model development. To extract useful information from our unique dataset, for example, in section 3.2 we not only show that a toluene hot spot exists near a paint and body shop, we also give a quantitative evidence that such a hot spot in a residential area has a similar effect on its vicinity as the large petrochemical industrial sources have on residential areas in a larger radius around them.
>
> To expand the quantitative information our dataset can provide in section 3.4 we now report on the rates of decay of several air pollutants away from a highway. These quantitative relations are hoped to aid in defining zones of potential impact from busy roadways, which is an important question in N.A. monitoring (see Brook et al., 2013 and related articles in that issue) and in undertaking and/or interpreting epidemiological studies exploring the health effects near-roadway exposure to traffic air pollution (TRAP). This analysis of the near roadway gradient of chronic exposures, which is based upon more pollutants collected over a wider range of conditions than most if not all past near-road gradient studies, also helps provide additional insight into the meaning of using NO_2, which has been found in many studies to be associated with adverse health effects, as an indicator of TRAP.
>
> Another aspect of the results shown in this study that is not common in other studies is the simultaneous measurement of many pollutants or derived pollutant-related parameters. This allows us to compare, for example, how different pollutants have different decay distances from a highway in section 3.4. Although near-road gradients were shown in other studies (e.g., Karner et al., 2010 to which we compare our results), here the comparison between the different pollutants is more reliable given that all pollutants are measured under exactly the same conditions (i.e., season, time of day, traffic intensity, etc.). They are also more applicable for understanding chronic exposures in a neighbourhood adjacent to a northeastern North American highway given the wider range of days and times of measurements, the greater distance explored and the inclusion of multiple pollutants measured with state-of-the-art instrumentation.

Section 3.1 is rather long, essentially establishing a certain representativeness of the mobile measurements towards a long-term average. I am wondering whether this could be condensed.

> We have shortened this section considerably (from 1870 to 592 words) to make it easier for the readers to follow, while keeping the main findings. Following the other reviewer's suggestions, some of the text was moved to the Methods section and some to the Supplemental Material along with Figure 2.

The results in Fig. 2 are not really impressive. While I clearly understand that from an experimental point of view, there are many reasons why the data clouds are scattered, I am not sure how this can really "prove" that the mobile measurements conducted during ca. 30 days are representative for the entire year.

> This plot was not presented to test if the ~30 days was representative. That evaluation was shown in Fig. 3 and Tables 2. Fig. 2 was provided to show that those subsequent evaluations could be undertaken because there was sufficient agreement between the concurrent measurements (CRUISER vs. VdM), although differences were found depending upon pollutant. We discuss reasons for these differences because they are largely explainable and subsequently these differences in concurrent measurements need to be considered when CRUISER's averages, which are hypothesized to represent longer term conditions, are compared to the actual long term conditions (Fig. 3 and Tables 2) which can only be obtained from the available monitoring network data.
> Figure 2, Fig. 3 and Table 2 also show that with our deployment approach of CRUISER it was possible to detect correctly the VdM AQ site with the highest NO_x values. They also show that the approach provides long term NOx patterns that are reasonably reliable (Fig. 3 and Tables 2 provide uncertainties as best as can be assessed). Overall, these results imply that CRUISER's measurements are able to expose areas with typically high vs. low pollution levels with a known amount of confidence. Uncovering these patterns, which no one has yet attempted to do with a mobile lab, is relevant for epidemiological studies examining chronic exposure patterns, in which the magnitude of the spatial gradients are equally, if not more important than the absolute pollutant levels.

Section 3.2 present extensive information on the spatial variability encountered, but much in the form of single anecdotes. It would be helpful to compile a systematic list, what pollutant profiles are associated with which kind of pollution sources (point sources and/or line sources).

> We have chosen to cut some of the more anecdotal information in this section and leave only the main characteristics, i.e., (1) the fact that NO_2 and particle number concentration show high concentrations near the roads while (2) SO_2 and Benzene are high near the petrochemical industry, and (3) still there are differences between these last two (SO_2 and Benzene) that stem from them being emitted from different activities/sources within this industrial complex.
> As for compiling a comprehensive list of pollutants that are associated with different kinds of sources, we feel this is beyond the scope and space of this present paper. However, with respect to the two major source areas we explored (i.e., traffic vs. petrochemical industry), we now briefly include such a list of at the end of section 3.2:
> *"Levy et al (2014) reported on the spatial correlations among all the parameters measured across Montreal highlighting which pollutants tend to co-vary and which are associated with different sources. From Fig. 3 and Levy et al. (2014) it can be seen that some mixes of pollutants show affinity to roads (BC, CO, HOA, NO, NOx, OM, PM1, PM2.5, PM10, PNC) and others are related to industry (SO2, Toluene, Xylene). "*

Section 3.3 is interesting because it highlights the problem of how to experimentally determine the spatial pollution gradients. I wish this section to be expanded a little, yielding a bolder statement on which pollutants are more and less difficult to capture with respect to spatio-temporal variability.

> It is not clear what "bolder statement" can be derived from the results in Section 3.3 that the reviewer would like to see. Any such statement will need to rely on quantitative measures, however we do not think the results we present in Figure 4 are sufficient for

this purpose. They are meant more as a demonstration of what can be achieved with a mobile lab and how a single AQ monitoring site might over/under estimate chronic exposures. Nonetheless, we have added a complementary section on the change in levels away from the main highway in the study area, which has led to some 'bold' statements about how far away levels can remain elevated. We have also revised the discussion and conclusions to better highlight what was significant about the data shown.

2) Introduction

This article deals with the relevance of ambient particles upon human health. Therefore I am missing some passages which, at least briefly, touch on the following matters:

The relevant dose affecting human health is best approximated by what we call "personal exposure". In certain cases, exposure to outdoor pollutants might make up a minor part of total personal exposure only. Much of this personal exposure by indoor exposure.

Regulatory issues. In your paper you present data on a lot of atmospheric parameters. Which of them are of regulatory concern, and why? Which of them are the currently more problematic pollutants, i.e. those whose ambient parameters exceed legal threshold values often in urban and/or hotspot areas?

Particle epidemiology. I see that you make, with your work, a move towards the determination of long-term exposure. This is perfectly right. Nevertheless, the Introduction would greatly benefit from an brief outline of what type of epidemiological studies exist that could make use of chronic exposure data.

The Introduction section was revised to address these issues. The importance of indoor exposures and personal activity patterns is now mentioned in the third paragraph:
"It should be noted that other factors also take part in determining exposure such as indoor-outdoor infiltration and individual time activity patterns (Monn, 2001). These factors can obscure how well outdoor spatial contrasts represent actual exposure leading to apparent modification of the effect of ambient concentrations (e.g., Janssen et al., 2002)."

The regulatory pollutants of concern are often only proxies to a specific source of emission. This is described in the second paragraph:
"In studies where exposure to a single pollutant is used for assessing associations with health outcomes, it is often acknowledged that this pollutant is a proxy for a specific source or more-complex mixture of concern. For example, NO2, ultrafine particles (particles with a dynamic range of >0- 100 nm) and black carbon (BC) are often considered as proxy pollutants for traffic related air pollution (TRAP) or fossil fuel combustion in general (e.g., Brook et al., 2007; Bukowiecki et al., 2003; Crouse et al., 2010; Janssen et al., 2011). SO2 is thought of as a proxy for heavy industry such as emissions from petroleum refineries (e.g., Smargiassi et al., 2009) and benzene could be related to both these sources in some cities (Wheeler et al., 2008) or mainly to one in others (Levy et al., 2014). …"

The types of epidemiological studies that use chronic exposure estimates are now mentioned in the first paragraph of the Introduction:
"Long-term or chronic exposure to air pollution has been shown in many epidemiological studies of different types, such as cohort studies, case-control studies and

cross sectional studies, to be associated with adverse human health outcomes (Ren and Tong, 2008). "

We chose not to describe these methods in detail trying to keep the focus on the main objectives and reduce the length of the manuscript.

p. 31588, l. 9: "These primary pollutants react among themselves...", unspecific - you speak about reactions of pollutants without having mentioned a single one explicitly.

We have added a specific example to the text:
"*Many of these directly emitted primary pollutants are subsequently involved in reactions in the atmosphere to produce secondary pollutants (e.g., NO+O3 to form NO2; nitrogen oxides (NOx) and volatile organic compounds (VOC) interacting to form ozone in the presence of sunlight; secondary organic aerosols derived from oxidation of VOCs).* "

l. 11: "The quantity formed depends on ambient conditions (e.g., photochemistry...)" . Would you call photochemistry an "ambient condition" ? Be more specific here, stating ambient conditions that might foster photochemical processes.

By "photochemistry" we meant solar radiation. The text was changed accordingly.

l. 17ff: This paragraph is confusing. You start with a mention of AQ monitoring, then move to LUR modelling, back to monitoring data interpolation before you speak about multipollutant numerical models. Right after that statement you turn to "saturation campaigns" again. According to my impression, the Introduction would become much clearer to a reader if you started, for example, with a clear statement on the health relevance of ambient pollutants, then move over to state-of-the art experimental observations, then to spatial data interpolation, and lastly to exposure modelling (LUR vs. numerical dispersion models).
The scientific community has hopes that AQ models will do a good job in the future to assess spatial variability, because this seems the only way to provide a spatially complete picture of pollutant exposure. Also, your data shows great potential to serve as a reference data set for the validation of pollution models. Therefore, as a reader I would appreciate some more information on the current state and capabilities numerical pollution modelling here as well.

We revised this section in the Introduction to make it shorter and clearer. We now describe the strengths and weaknesses not only of numerical models but of all other approaches, including AQ monitoring, intensive measurement campaigns, spatial interpolation, dispersion models, numerical models and empirical models (LUR). We believe it is important to provide the readers a brief overview of the methods used to assign exposure estimates in epidemiological studies, so they can evaluate the contribution of this study to the field.
With respect to future use of the data to validate models, we do plan to use the data to develop LUR models for several pollutants. This application will be unique in the extent of temporal coverage of the data (34 sampling days taken at different seasons), spatial coverage (sampling a range of land uses and different road types) and number of pollutants.

It is, of course, neither possible nor necessary to cite all previous relevant works. Nevertheless, I feel that this paper omits some of the more relevant contributions to the topic of spatial variability of urban pollutants, particularly concerning PM, particle number and size distributions. Therefore, I am suggesting you to consider the following list of studies for possible inclusion in the references.

Avery et al.: Estimating error in using ambient PM2.5 concentrations as proxies for personal exposures: a review, Epidemiology, 21, 215 -223,2010
Costabile et al.: Spatio-temporal variability and principal components of the particle number size distribution in an urban atmosphere, Atmos. Chem. Phys., 9, 31633195,2009.
Freiman et al.: Urban-scale variability of ambient particulate matter attributes, Atmos. Environ. 40, 5670-5684.2006
Hudda et al.: Inter-community variability in total particle number concentrations in the eastern Los Angeles air basin, Atmos. Chem. Phys., 10, 113852010 ,11399.
Pinto et al.: Spatial variability of PM2.5 in urban areas in the United States, J. Air Waste Managem. Assoc. 54, 4402004 ,449.
Pirjola et al.: Dispersion of particles and trace gases nearby a city highway: mobile laboratory measurements in Finland, Atmos. Environ., 40, 8672006 ,879.

We revised the Introduction and it now includes additional references, though not all, as there are already over 50 references cited.

3) Conclusions section

In its current state, the Conclusions section contains only little in terms of concrete and really novel results. The current text appears like a confusing mixture of study results and rather general statements. I have read this section many times, and arrived at extracting the following three significant conclusions:

p. 31609, l. 6ff: A lot of variability in pollutant concentrations was observed, both spatially, and also among pollutants.

p. 31609, l. 11ff: Mobile measurements collected in particular places during ca. 20 visits per year agree with corresponding annual averages. From this it is concluded that the mobile measurements, although recorded only during limited periods of time, can be used with confidence to approximate longer-term averages.

p. 31610, l. 5ff: More studies to characterize the spatial variability of air pollutants are needed, also using new observation technologies.

The body of these conclusions is not particularly impressive, many of them being common knowledge. In their current state, these conclusions are not sufficient to warrant an ACP publication.

Therefore, in my opinion, the Conclusions section needs to be completely rewritten. I suggest to start by building up some very concrete results from this study, for example mentioning distinct features of the spatial particle distribution, the profiles of particle sources, certain peculiarities found with individual pollutants, and highlighting the ups and downs when sampling from a driving car. See also my suggestions related to the Results sections above. In the second half of the section, these results can then be interpreted in the framework of the more general knowledge, and lead into recommendations of what could be done in the future.

As suggested, the Conclusions section was completely rewritten. Following our clarification of the paper's objectives we describe above, we now open with the limitations of mobile measurement campaigns and our approach for applying mobile measurements for chronic exposures, then describe some of the main findings in the paper in the context of the gaps in current knowledge and what the unique contributions are from our paper. We then conclude with a general statement about the role mobile

measurement campaigns might play in characterizing urban scale patterns in average air pollutant concentrations, developing exposure estimates from such data, along with other complementary methods like physical and empirical models and remote sensing data.

4) Technical issues

Major points

p. 31590, l. 14, caption of Fig. 4 and others: : " annual season" ," annual mean pollution levels" : Calling the total campaign average " annual" average is misleading and should be avoided. I truly accept that you wish to generate from their data something approaching the annual average but as a matter of fact, even the three campaigns cover only part of the entire year (in fact, less than 5% of the time), and nomenclature should in any case respect that situation. We all know how changeable, for instance, PM concentrations can be over the course of a year as a function of synoptic weather. Better candidates might be "campaign average", or "total average".

We agree that the term "annual" is misleading and replaced it by "campaign average" throughout the text and figures.

p. 31590, l. 23ff and Appendix A2: Wind speed correction to the PM measurements

First, the physical units are missing for the factors given on p. 31611, l. 24-25. Therefore I am unable to judge the magnitude of the effect.

The units are the inverse of the speed, i.e., h km^{-1}.

Second, I find the concept of this correction quite adventurous! One basic reason is that for its derivation, you use PM data collected up to 500 m apart from the corresponding stop site. This bears some serious contradiction: One the one hand, you are not getting tired in this manuscript stating over and over again how variable airborne pollutants are in this urban environment. At the same time you seem to be happy to pick values as far as half a kilometer from a stop site, possibly recorded half an hour earlier or later, and consider them as equivalent to those recorded at the stop site, with the exception of being recorded at a given driving speed. If I accepted your method as is, you would need to agree that your whole experiment should not warrant the discussion of spatial parameter variations at a resolution of higher than 500 meters at all.

We agree that attempting to consider how the mobile lab's speed may influence inlet collection efficiency of particles by the main size fractions PM_{10}, $PM_{2.5}$ and $PM_{1.0}$ is adventurous. As a result, most, if not all, groups undertaking mobile PM sampling tend to ignore it and this may be acceptable for fine particle measurements. However, despite being adventurous and no simple way to do it (i.e., wind tunnel work with the mobile lab and generating known particle concentrations is laborious and costly), we felt it was important to assess whether our mobile measurement approach leads to a low bias in coarse particle collection, especially given the slow inlet flow for the GRIMM instrument. Correcting for this bias, if found to exist, was assumed to enable our mobile measurements to more-confidently report on urban PM_{10} and coarse particles, a size fraction where 'spatial data' are very limited.
While the general convention is that sampling of fine particles (2.5 µm and less) are not highly sensitive to inlet configuration as they will follow the gas stream lines into the

inlet and also won't experience significant line losses with bends in the sampling line, assuming no static losses, the coarse particles have a greater potential of not being captured by the inlet and being lost in the inlet. Thus we felt it was necessary assess this potential and hypothesized this could be done given the large number of measurements we have obtained under a wide variety of conditions, including away from large local sources, and based upon the assumption that there is reasonable potential for spatial homogeneity in $PM_{1.0}$, $PM_{2.5}$ and PM_{10} over the "neighbourhood scale".

A large number of PM studies in urban areas have consistently shown that PM mass is less spatially variable and that there is a higher correlation among monitoring sites compared to locally-emitted primary pollutants such as CO, NO, and certain PM constituents (e.g., BC) or sizes (e.g., ultrafines). Thus, our assumption about "neighbourhood scale" homogeneity is not unreasonable. We agree that this assumption is not likely to be valid under all circumstances and is counter to a main theme in the paper. However by first obtaining a local, fixed measurement at an urban background monitoring location and then restricting the mobile measurements to compare with it to those obtained relatively close spatially and temporally we hypothesized that if indeed there was a systematic under-sampling, a large number of pairs of measurements, generated using the same instrument as we describe below, has the potential to uncover a pattern despite the high potential for a large amount of scatter.

Another reason is that the relative particle losses depend on the particle size distribution, which may vary from time to time. At one time you may have more coarse particles compared to fine particles present in the atmosphere, and vice versa. In these two cases, the relative particle mass losses will necessarily be different even if you encountered the same wind speed. The latter reason implies that the current correction can, in the best case, correct for some crude overall effect, and may be highly invalid in single cases.

Indeed because our approach to assess if there is an effect depends upon a pattern emerging from the expected large scatter arising, which would be due to many factors (e.g., spatial heterogeneity, instrument noise, differing orientation of the inlet relative to wind), the correction is expected to be relatively crude. However, if a significant and physically intuitive relationship was detected then we felt this correction would be valid and important to apply (i.e., as opposed to ignoring like other mobile work has tended to do). We agree that there would likely be errors introduced for single cases/measurements, but for correcting the mean, this approach can be expected to reduce the bias in average PM_{10}, $PM_{2.5}$ and $PM_{1.0}$ values that we focus on in this paper. Note that the main form of the observations we interpret in this paper are derived from averages of many measurements sorted by location as opposed to from individual 6 sec measurements.

On the basis of the information currently available, I would strictly decline the usability of this correction. At the same time, the magnitude of the correction might not be overly high. I would be willing to reassess the issue if you provided, for example, a graph of the data cloud upon which the correction was established.

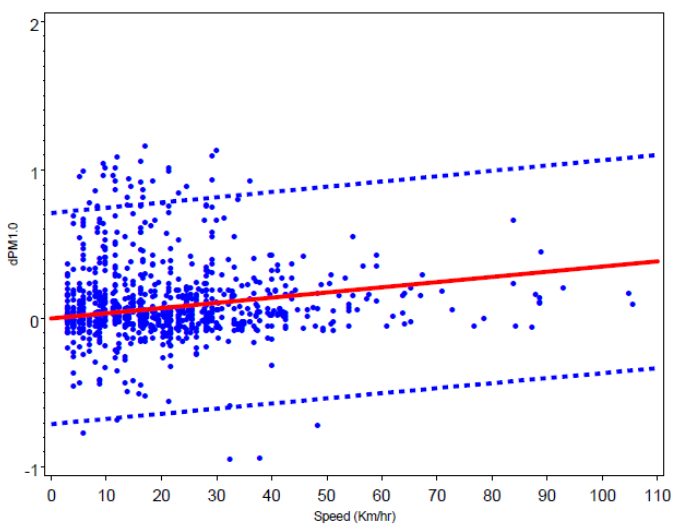
We have added this information in the supplemental material where we discuss the speed correction along with a more detailed description of the methodology used. Below are plots of the data for $PM_{1.0}$, $PM_{2.5}$ and PM_{10}. The fixed site average $PM_{1.0}$, $PM_{2.5}$ and PM_{10} ('ambient') are obtained with the mobile lab parked at an urban background location, typically a government monitoring site. All six second PM measurements taken while

approaching or leaving this fixed site (i.e., within 30 min of arriving or leaving the site) that are also within 500m of the fixed site are identified and matched with the current speed of the mobile lab. Any mobile measurement associated with NO>5 ppb is removed as this may be influenced by a very local combustion source (i.e., nearby vehicles) and is not appropriate for comparing with the recent 'ambient' values. Similarly, measurements with zero speed were excluded to avoid impact from the mobile lab's own plume or vehicle idling ahead of the lab in this analysis.

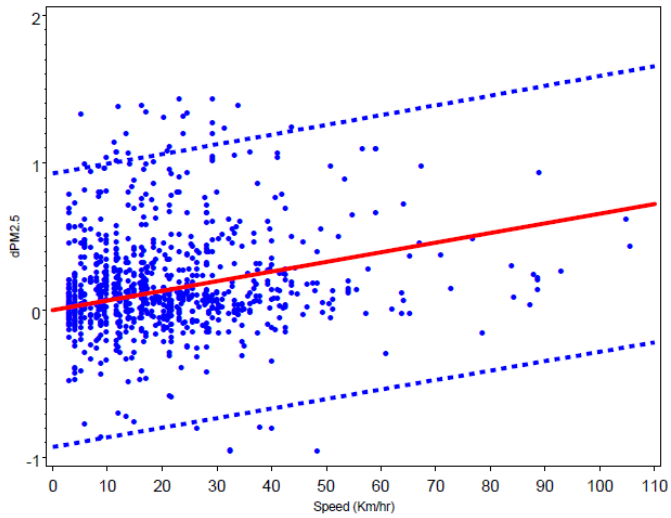
Using the remaining mobile measurements the ratio of the corresponding 'ambient' (i.e., within 500 m and 30 min) and each of the six second mobile measurement values is determined and a value of 1.0 is subtracted to compute percent increase in 'ambient' PM relative to the mobile value. This percent increase is then plotted versus the corresponding speed. As the example figures below show there is a considerable amount of scatter, which is not surprising given spatial heterogeneity and measurement noise in the GRIMM instrument when operating at 6 sec resolution. This scatter strongly supports the statement from the reviewer that we cannot use the correction for individual measurements as the confidence in the individual correction value is low and will depend upon several factors such as the details of the size distribution. However, in adjusting the mean mass, which is our intent, these corrections are more-appropriate given the statistical significant of the slopes.

As expected, as speed approaches zero the percent increase approaches zero (y-intercept) and as speed increases this percent increases (i.e., stationary PM>mobile PM). The lines on the plots are a linear fit to this relationship and the slope provides a speed-dependent correction factor. The slopes for the PM10, PM2.5 and PM1.0 relationships increased with the upper size cut for the PM measure, which provides confidence that the pattern observed is related to the effect of speed on particle collection efficiency of the inlet.

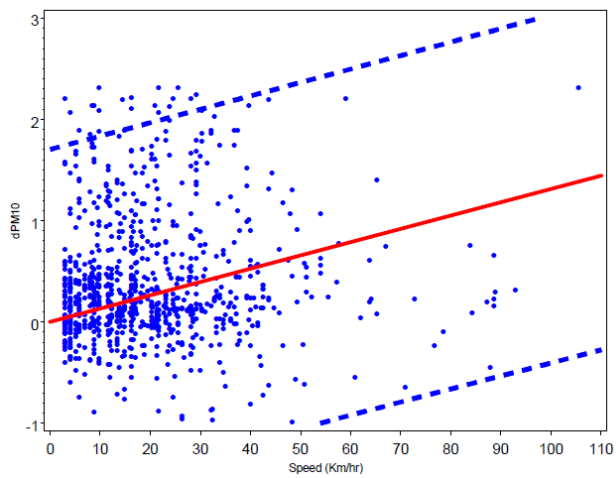
The slopes we obtained were 13.12×10^{-3} h km^{-1} for PM10, 6.53×10^{-3} h km^{-1} for PM2.5 and 3.49×10^{-3} h km^{-1} for PM1, all with p-value < 0.0001. The standard errors calculated were 0.89×10^{-3}, 0.47×10^{-3} and 0.37×10^{-3} h km^{-1}, respectively.



Scatter plot of the measurement difference vs. vehicle speed for PM1.0



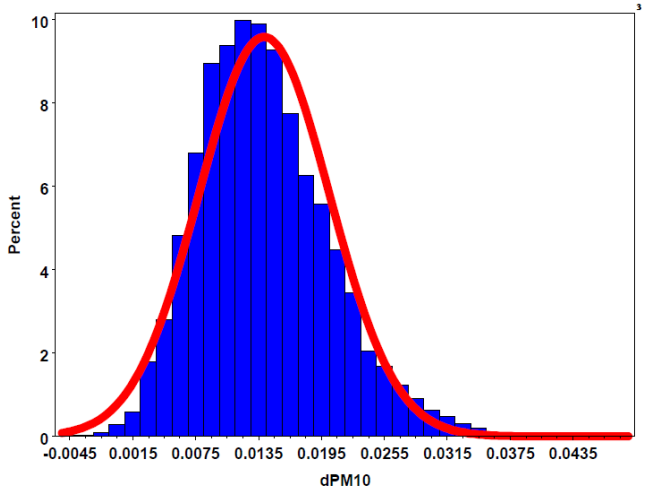
Scatter plot of the measurement difference vs. vehicle speed for $PM_{2.5}$



Scatter plot of the measurement difference vs. vehicle speed for PM_{10}

The three figures above represent one random portion of the data (i.e., are an example). To assess the sensitivity of the slope to what portion of the data are used and hence to obtain a measure of the confidence in the correction factor we repeated this regression 10,000 times using separate, randomly selected data portions using a sample size determined to be necessary to yield statistical significance in each slope estimate. The slope we applied to correct that data was derived from the median value among these 10,000 regressions. The figure below shows the distribution of the slopes obtained for PM10. We point out that virtually all of the regressions resulted in a positive slope for PM10, PM2.5 and PM1.0, suggesting that the speed effect is indeed real. These distributions were close to being normally distributed and their standard deviation thus provides an estimate of the uncertainty in the slope or the speed-dependent correction factor. As stated, this correction is uncertain for each individual 6 sec measurement. However, it is expected to appropriately account of vehicle motion in determining the mean PM values by location thus reducing the impact of vehicle speed on apparent differences in GRIMM PM measurements among locations. We believe that we have developed a novel approach to addressing the speed effect (demonstrated to exist based

upon the slopes), which can be an issue when PM measurement instruments that use a slow inlet speed are deployed on mobile platforms.



Distribution of slopes among 10,000 regressions for estimating the PM10 speed-dependent correction factor. Mean slope (dPM10)=0.0140; median slope=0.0134; standard deviation=0.0062.

Moreover, I would expect a critical assessment of the method, such as an uncertainty involved, and a discussion of the consequences on the interpretation of the measurement results, for example, in the shape of error bars that depend on wind speed.

See figure above and the corresponding text regarding our multiple estimates of the speed correction method. We have added a detailed discussion to the description of the speed correction regarding the uncertainties and limitations of the method. We now present the scatterplots above in the SM for each PM fraction, the difference between the corrected and uncorrected values per speed range and a table with the statistical parameters of mean, median and standard deviations.

Table 1, "UFP": It is not acceptable to call a total particle number concentration measurement "UFP", even if some publications might have done so. As you know, the condensation particle counter measures not only UFPs (D < 100 nm) but also other particles up to a maximum diameter defined by the inlet system. I accept the view that total particle number is often dominated by the number of UFPs but still, UFP number and total particle number are not the same! Therefore, please use "total particle number concentration", Ntot or a similar term throughout the entire manuscript.

We changed UFP to Particle Number Concentration (PNC) throughout the text and in the figures.

Minor points:

p. 31586, l. 6 and five more occurrences in the text: "saturation monitoring campaign". I do not completely grasp the meaning of "saturation" in this expression. I just speculate that you mean that somebody measures as long as needed to determine some representative state of the atmosphere.

However, experience tells that the state of the atmosphere can be very variable when looking at pollutant concentrations. It can, in fact, take years in the atmosphere to determine a climatologically representative average. Therefore, I suggest to replace “saturation monitoring campaign” by the less controversial “intensive monitoring campaign” unless explained in more detail why there should be any “saturation”.

The term “saturation” is used here to indicate a large number of monitoring instruments/sites over a relatively small area. For better clarity we have changed the term with “intensive”, as suggested, but point out that “saturation monitoring” is, in North America, at least, a relatively well-recognized term.

p. 31590, l. 14: “ three week deployment” . Please note the periods of the measurements explicitly by date.

The start and end dates of each period were added to the text.

p. 31590, l. 17: “23 different species”, replace by “23 different parameters”

We’ve changed the text as requested.

p. 31591, l. 17: “)1) East Montreal; (2) Central and West Montreal” . I was trying hard to identify in Fig. 1c which routes you mean. In my eyes, the Island of Montreal seems actually aligned much in a direction from the south to the north. After combining some thoughts I guess that the area in Fig. 1b is what you mean by East Montreal. But this seems to be neither noted in the map, nor in the figure caption, nor in the description of the study area (Sect. 2.1). Therefore, please add labels to the maps where “East”, “Central” and “West” Montreal are. See also below my suggestion to improve Fig. 1.

The terms “East Montreal” etc. are common among Canadians. Following this comment and the comments below we have revised Figure 1 and added label to East and West Montreal.

p. 31592, l. 6: “Figure 1c shows that a large majority of the sampling was conducted in residential areas...” . Sorry, but this cannot be grasped from Fig. 1c. It could be easily seen if there were a topographic map and a concentration map side by side showing exactly the same area. See also below my suggestion to improve Fig. 1.

We have made the suggested changes to Figure 1 and it now addresses these concerns.

Table 1, “Response time”: Do you mean time resolution? Please clarify. In environmental technology “response time” is usually associated with the delay time of a particular sensor to react (“respond”) to a change in sample concentration, emphasizing potential limitations in the physical measurement process. “Time resolution” would simply refer to the frequency at which data points are collected.

True. The term “response time” is not correct. It was changed to “Time resolution”

Fig. 2: Changing “VdM measurements” to “VdM fixed site measurements” would make a reader grasp immediately that these are the fixed site measurements.

“VdM measurements” was changed to “VdM AQ sites measurements”.

5) Presentation of the geographical setting and data.

Figure 1: This Figure plate is nicely compiled, but confusing in some aspects.

Fig. 1a and 1b are inconsistent. They use quite different color shades (red) to mark roads. Second, Fig. 1a features black as a color code for certain roads that neither appears in the legend, nor in Fig. 1b. This is quite annoying to a reader and should be fixed.

Fig. 1b contains a corrupted road label "25 & 40". Please fix this, or delete it.

The legend is confusing in that " CRUISER's stops sites" and " AQ site" appear below the headline "NPRI emission sources" . These need to be filed under a dedicated headline, such as " Experimental points of interest" or similar.

I guess that the terms " East" ," Central" and " West" Montreal refer so specific urban areas. I found these descriptions in the text, but was unable to locate these areas on the map. Hence, please add corresponding labels to the maps.

There is another issue I discovered: On the one hand, you put a focus on North Montreal in that you zoom into that region (Fig. 1b). On the other hand, this interest is not reflected in a corresponding separate map for particle concentrations. Fig. 1c does this job for Fig. 1a, but it would be insightful to see North Montreal as a dedicated concentration map covering the same area and displayed at the same scale. This would certainly improve the visual orientation on the pairs of maps (geography/concentrations). I like the maps, but had real trouble associating the driving patterns in Fig. 1c with the geographical features displayed in Fig. 1a.

To improve the situation for the reader I suggest the following: Do assemble two Figures, each containing a topographic and a concentration map side by side. The first Figure would cover the area of Fig. 1c, which is the area of interest for the exposure measurements, or maybe a bit larger. The topographic map would be much like Fig.1a, the concentration map much like Fig. 1c. This new Figure would provide the overall situation for the entire campaign. Try to make sure that both maps cover exactly the same area. This is important for visual orientation!

A second Figure would provide the situation for the focus area in north Montreal, containing a topographic map similar to Fig. 1b and a concentration map (new) side by side. The advantage over the large-scale map would be the identification of industrial point sources, which are already omitted from Fig. 1a. I believe that the production of two separate Figures would be of excellent service to the readership.

To save some space, the overview map showing Montreal on the continental scale could be dropped. Knowledge of the location of Montreal might be widespread enough.

> Figure 1 was changed to address all the concerns raised. It now holds four panels, covering the topography (i.e., land use, emission sources, roads and locations of CRUISER's and AQ monitoring sites) and concentration of measurements (measurements /km), each in a large overview covering the entire island and a close-up to the "east" part of the island. A fifth panel presents the neighborhood of Anjou in greater detail, showing the location of the cross-section used for analyzing near road gradients in the newly added Section 3.4. Labels were added marking the "East" and "West" parts of the island.

References:

Brook RB, I Levy, C Mihele, 2013: From near-road to urban background: lessons learned from mobile lab monitoring. EM Magazine, July 2013

Levy, I., Mihele, C., Lu, G., Narayan, J., and Brook, J. R.: Evaluating Multipollutant Exposure and Urban Air Quality: Pollutant Interrelationships, Neighborhood Variability, and Nitrogen Dioxide as a Proxy Pollutant. Environ. Health. Perspect. http://ehp.niehs.nih.gov/1306518/, 2013.