Atmos. Meas. Tech. Discuss., 8, C776–C780, 2015
www.atmos-meas-tech-discuss.net/8/C776/2015/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Plume-based analysis of vehicle fleet air pollutant emissions and the contribution from high emitters" *by* J. M. Wang et al.

Anonymous Referee #2

Received and published: 21 April 2015

Wang and co-authors use a highly-time resolved measurement system located in a ground-floor laboratory in a building near a busy Toronto road to measure traffic-related plumes and derive fuel-based emission factors. Sampling periods are spread through the four seasons, beginning in December 2013 and running to September 2014. Plume detection was based on the rate of change of CO2 concentration. Over this period, the team identified over 150,000 plumes. Of these, 100,000 were of sufficient length (10 s) and size (integrated CO2 area) for determination of the fuel-based emission factors for particle number (> 7 nm), CO, NO, NOx, black carbon, methanol, benzene, toluene, ethylbenzene, and xylenes. The technique recovers a distribution of the emission fac-

tors of the in-use vehicle fleet in Toronto. The plume detection rate (assuming sampling was conducted continuously for 3300 hours – the issue of downtime is not discussed) was about 45 plumes per hour.

1. The paper requires major revisions as I believe it is not in enough depth with respect to the methodological details and sensitivity of results to data analysis assumptions. And, for an article in a journal focusing on metrology technique, the paper has too much in terms of results and policy relevant statements. I believe a stronger emphasis on the technique and data analysis will serve the measurement technique and research group well in the long term. The results can be placed in a more focused article, with methodological details in the AMT paper. For example, at lines 113-117, the authors make clear that they have a strong interest in showing the applications of the method. I feel the method itself needs more exploration, and that the paper would be improved by more in depth methodological evaluation, and moving some of the applications (particularly the evaluation of the local emissions testing program) to a more policy relevant journal.

This reviewer believes the measurements are an important contribution, and that (because of the large number of plumes) a valuable quantification of fleet EF variability. The methodological and data analysis rigor that I believe are needed (and describe below) are not meant to be a barrier to publication, but rather to add value to the measurements so they can be widely used by the metrology, modeling, and policy communities.

2. At the paragraph starting at line 137, it is not clear which of these are from government statistics on the fleet, and which are from observations at the monitoring location. For the observations at the monitoring location, the method is not clear. If the method is video analysis, is daytime included as well as nighttime?

3. The supplement has much less treatment of error, precision, uncertainty, and sensitivity analysis than I expected. These concepts have very limited discussion in the

article itself. This seems a critical deficiency in a measurement-focused journal.

4. The paper mentions validation of the PTR-MS using standard gases and refers the reader to Wang et al. (2005). Does a 2005 paper really quantify the errors in a measurement taken in 2013/2014? Can the uncertainty and detection limit of the instrument be stated, if not in the paper, in the supplement?

5. Is there any intercomparison information or electrometer calibration available to quantify the CPC accuracy?

6. The acronym ADL is used in the supplement without being spelled out.

7. In the supplemental section, the integral of the CO2 time series is reported as having units of ppmv. However, the time integral of a CO2 time series will have dimensions of mixing ratio seconds.

8. At line 195, how was 10 seconds and 5 ppm selected as the threshold. How sensitive are the results to this threshold? I believe a sensitivity analysis is needed, with the results recalculated for alternate settings (e.g. 12 seconds and 4 ppm; 8 seconds and 6 ppm, etc.). Alternately, can the error or uncertainty embodied in the arbitrary selection of 10 sec or 5 ppm be quantified and then propagated through the analysis by traditional error propagation / error analysis techniques.

9. What is the accuracy or uncertainty associated with the selection of the baseline? (line 207)

10. What is the formula for the effective sensitivity (line 210 and table 1)

11. The time resolution in table 1 – is that the reporting frequency of the instrument? Another important temporal variable in the method is the degree of temporal "smearing" that occurs; in other words if a sharp step change is applied at the inlet of the monitoring system, how sharp is the reported signal from each of the instruments?

12. At line 224, QA measures are mentioned, but it is not clear what the QA measures

are. From figure S2, I gather that the QA is to test if the peak is sufficiently large. This falls short of what most readers think of in terms of quality assurance – which usually involves some combination of tests relative to independent challenges such as replicate measures, a test of monitor response to a calibration standard, zero check, leak check, etc.

13. The validation methods mentioned at line 234 are important, but this is the first mention of them in the paper.

14. What is "capture efficiency" as used at line 253

15. I find the report of the experiments in the paragraph starting at line 253 incomplete and confusing. Is there a diagram of this test, and clear explanation of what is being controlled, what is being measured / calculated, etc?

16. At line 264, it seems there is confusion between a high concentration event, and a high emission factor.

17. At line 276, what constitutes a "well-tuned" engine for the purposes of this study? This is, as used, an unnecessarily vague term.

18. I do not see the purpose of Figure 1e.

19. The results from testing against the direct injection engine are discussed but not shown graphically or quantitatively. Line 246. In the current manuscript, it is stated that a test was done, and that the result was within 14% of a value in the literature. This is to be accepted as method validation. I believe significantly more detail is needed to make this into something that lends credibility to the measurement technique. Comparing to only one emission factor, and not discussing its representativeness of accuracy / QA – lessens the impact of this section. Not presenting any tabular or graphical data from the drive by plumes also lessons the impact. I recommend this be strengthened or deleted. And since comparison to a different vehicle is always needed (I don't believe there are dynamometer tests and drive by tests in Toronto of the same vehicle), removal of this

is probably prudent.

20. The size cut of the CPC (and thus of the particle number emission factor) needs to be mentioned more prominently, since the intercomparison of particle number emission factors is only relevant when the size cut of the measurement is known. It should appear at line 180 (methods section) and in all tables and figures than include PN measurement or PN emission factor.

21. How much downtime occurred during each measurement campaign? If there was significant downtime, was the remaining active sampling time concentrated in particular time periods such as certain days of the week or nighttime periods?

22. In the supplemental section, it can be seen that the CO plumes start before the CO2 plumes. However, it seems the plume start time is determined by the CO2 time series. The CO instrument seems to be smearing out the CO over time. If this is not accounted for the CO emission factors will be biased low.

This is a key issue in the method (possibility for temporal mismatch and temporal broadening of the peaks across different instruments). It needs to be clearly explained/explored in the methods and results sections – possibly with support in the supplemental section. Alternate methods such as allowing different plume start and end times for different pollutants need to be considered. Sensitivity analysis of the results to the algorithm start and end time selection method is also needed.

Minor comment 23. Line 143 “observation is observed”

Interactive comment on Atmos. Meas. Tech. Discuss., 8, 2881, 2015.