

Response to the Reviewer's Comments

We thank the reviewer for their comments and below is a response to the reviewer's general comments, which addresses the major points that they have highlighted. We have also provided responses to the specific comments and highlighted changes in the manuscript.

The time delay is discussed in the specific comments section (lines 82-84) and this was tested using a lighter, a 10 second time delay for the UGGA was applied and no time delay was observed for the SIFT-MS. Wind conditions were not investigated as the number of routes that were performed should cover a range of meteorological conditions as they were performed over multiple days.

Further discussion on the calibration approach has been added to provide clarity on the approach and the reasoning behind it. A detailed response is included below (lines 154-158).

In the paper we have now expanded on the results and analysis section by developing a new method that is complementary to the existing analysis in the paper. This analysis is now in the paper as an additional section (Section 3.4) and involves investigating the toluene to benzene (T/B) ratio using quantile regression modelling. This has been used alongside a gaussian kernel smoother to show the T/B ratio along the distance of a road of interest to reveal intermittent sources that would be missed through only spatially averaging compounds.

Further analysis will be carried out to continue the characterisation of different emission sources in York and further development of the new method of quantile regression modelling will be carried out, which will also assist with determination of dominant emission sources. C₂-alkylbenzenes and C₃-alkylbenzenes may not all be required to identify gasoline emissions, but both of these classes of compounds were in the calibration gas, so it was useful to measure these externally calibrated compounds to determine the success of the new measurement technique. It is worth noting that the selected compounds can be changed to suit different applications, so this will be done for future measurements. Ratios of toluene-to-benzene are investigated in the results section and are compared to other studies to help distinguish different emissions sources. The results section discusses the use of toluene-to-benzene to distinguish between emissions and evaporative/solvent emissions along certain sections in York. Other sources such as hairdressers and dry cleaners were not revealed by spatial mapping. However with further development of the quantile regression analysis, it is likely that other intermittent sources will be revealed.

We hope that these responses are satisfactory and that incorporation of a further discussion of the calibration approach and further development of the results section has addressed the points raised by the reviewer.

Reviewer 2

This study presents the development of a mobile analytical platform equipped with two analyzers: an Ultraportable Greenhouse Gas Analyzer (UGGA) and a Selected-Ion Flow-Tube Mass Spectrometer (SIFT-MS) measuring carbon dioxide, methane, several VOCs and other trace gases. The authors describe the different techniques used by these instruments and give some details about the calibration of the SIFT-MS. This mobile platform was deployed in the city of York, UK, where it completed a total of 31 surveys of the same route over a 10-day period. The authors present here some preliminary results of these measurements.

Overall, this study has a lot of potential and could be a valuable contribution to the literature as it is important to monitor pollutant emissions and their evolution in urban areas. Mobile measurements of VOCs are very useful especially to characterize and identify different types of sources within a city. However, I was expecting a deeper analysis of the measurements. The development of such a mobile platform represents a lot of work but it is enough to justify a paper. The authors managed to collect an impressive amount of data (especially during a pandemic period), the paper would really benefit from developing the results section, I recommend addressing the following major points before publishing it:

It is not specified in the article if the authors took into account the time delay between the sampling inlet at the front of the van and the moment it reaches the instruments. This time delay can induce a shift of the measurements due to the motion of the vehicle and impact the aggregation of the measurements into the 30 metres segment. It is important to give this information when presenting mobile measurements. Did the authors look at the wind conditions during the measurements? If the wind

was coming from one direction on given day and on the opposite direction on another day, it will likely have an impact on the composition of the sampled air, especially when measurements are performed close to known sources.

I would have liked more discussions about the calibration approach. First, I would make it clear at the beginning of Section 3.1 that 2 or 3 different calibration approaches were used for different components (AGCU and equation 1 plus calibration of UGGA?). Second, I would explain why the authors chose this seven steps approach? I understand that due to time constraint, it is not possible to measure each step for too long but 3 minutes seem a bit short. Why not measure each step only once but twice as long? Why did the authors only used the right-hand steps of the post-drive calibration? What is the point of the pre-drive calibration? Did the authors observe any drift of the measurements between the first and the last day of the campaign? Third, I would have also been interested in a comparison of the two calibration methods. The authors could have applied both approaches to benzene measurements for example, and shown how the two sets of calibrated measurements compare.

Reorganize Section 3.2: having only one subsection does not make much sense. I would either add a title to the first part of Section 3.2 or remove “3.2.1 Spatial correlation mapping”. Also, I do not understand why the authors presented the distribution maps of only 2 components, I would have liked to see more, especially species that are correlated (benzene, toluene, alkylbenzenes...). I was a bit disappointed at the end of this section when the authors teased the application of the correlation approach to smaller scales for a future study, I was expecting to see that in this paper... In the end, I am not sure I am convinced of the benefit of measuring all of these components at the same time. Here are a list of questions I got after reading this paper: What is the dominant source of VOCs of York? Do we really need C2-alkylbenzenes and C3-alkylbenzenes measurements to identify emissions from gasoline evaporation? What are the ratios of these different components for known sources? How do they compare with ratios found in other studies for the same type of sources? Can we differentiate emissions from gasoline fuel and diesel with one of the measured component? What about the other sources mentioned in the paper (dry cleaners, hairdressers...), were you able to detect any emissions from them?

- 1.1 Line 1: *“The importance of emissions source types...”: the authors should specify what type of sources they are referring to (VOCs? GHG? atmospheric pollutants?).*

This sentence has been changed accordingly.

“Over the last two decades, the importance of emissions source types of atmospheric pollutants...”

- 1.2 Lines 23-24: *Be consistent with the notation throughout the text: choose between “ozone” and “O₃” and stick to it. Same for “secondary organic aerosol”/SOA, “limit of detection”/LOD...*

Suggestion implemented.

- 1.3 Lines 29-30: *I would specify that you are talking about static measurements at monitoring sites here. This is developed later but it is not very clear in this sentence*

Suggestion implemented.

- 1.4 Line 50: *Replace “targeted” by “targeting”?*

This sentence has now been removed.

- 1.5 Line 83: *“1/2 inch”: shouldn't the authors use SI units?*

SI units have now been included.

- 1.6 Lines 82-84: *How long does it take for the sampled air to go from the inlet in front of the car to the instrument? Is this delay taken into account and corrected on your maps? If not, it will have an impact when you bin your data into 30 meters segments.*

The time delay for the instruments was investigated by using a lighter at the front of the sample line and observing the time delay in an increase in butane (SIFT-MS) and CO₂ (UGGA). The SIFT-MS gave an instant response to the lighter test so no offset had to be applied. The time delay of the UGGA was 10 seconds so this offset was applied to the measurements of the UGGA. Since the measurements, the pump of the UGGA has been changed, so any future measurement will have a faster response.

1.7 Line 86: *Is the wind corrected for the motion of the platform?*

The wind was not corrected for the motion of the platform and therefore was not used in the study. The text now reads:

“Real-time location of the WASP is recorded by a Garmin GPS 18x PC and a measurement of wind speed and direction is measured using a Gill 2D Ultrasonic Wind Sensor (which was not used in this study)...”

1.8 Section 2.2: *The authors give the precision of the UGGA for methane and carbon dioxide but they do not talk about the precision of the SIFT-MS. What is the precision of the SIFT-MS for the different species? Does it correspond to the limit of detection described in Section 3? What is the limit of detection of the UGGA for methane and carbon dioxide?*

The precision of the SIFT-MS measurements has been added to Table 3 and was calculated using 2 times the standard deviation (σ) of measurements made when sampling from zero air or nitrogen gas.

1.9 Line 120: *How many species does the SIFT-MS actually measures? This is very confusing: Table 2 and Section 2.2.1 state that it measures 13 components but Table 1 states 14 species.*

The SIFT-MS measures 13 during the measurement period, there was a mistake in table 1 and we thank the reviewer for noticing this.

1.10 Lines 123-125: *I am not sure I correctly understand what is a cycle? Does it correspond to the measurement with one reagent or to the succession of the three reagents?*

To avoid confusion to the reader, the term cycle time has been removed and acquisition rate added. The text now reads:

“ To maximise spatial data density during mobile measurements, the instrument acquisition rate was minimised with only a single product ion monitored for each compound. Therefore the sampling method used during measurements has an acquisition rate of 2.5 seconds with a 90 ms ion dwell time.”

1.11 Lines 134-135: *The precision of an instrument should always be stated over a time period.*

The precision time stated in the text is over 1 second and the time period has been included in the text.

1.12 Lines 154-158: *Why do you only use the right-hand steps of the post-drive for the calibration? Did you test several calibration approach with the AGCU? I wonder why you implemented these seven consecutive steps of three minutes instead of only doing four steps but with longer measurements. Also, each step of the calibration last 3 minutes, did you use all the data within the 3 minutes or did you get rid of the first measurements that are usually influenced by the transition? Why is the agreement between pre- and post-drive calibration not as good for ethanol and methanol?*

The calibration set up consisted of a heated palladium catalyst (380°C) to generate clean zero diluent gas from ambient air coupled with a gas blender to deliver controlled gas concentrations to the SIFT-MS. As this calibration set up had to be switched off at the end of each day and restarted the following morning (for safety and practical reasons) there had to be adequate time for the catalyst and calibration system internal surfaces to condition with the target VOCs on a daily basis. This is why only the right-hand steps of the post drive calibration were used. Practically these produced the most agreeable results.

We tested several calibration profiles, but the main problem is the time constraints involved. In order to get the SIFT-MS operational (which has to be switched on and off each day), stable and calibrated pre and post drive and then spend several hours performing mobile measurements on a daily basis there are considerable time constraints, hence the calibration protocol has to be concise. 4 calibration points were chosen as this is considered to be the minimum requirement for accurate instrument calibration. 2 x 3 minute calibration steps were chosen to assess both the reproducibility within the calibration and across pre- and post-drive calibrations. In our opinion this is a minimum requirement for confidence in any calibration.

We did not use the first two minutes of the calibration, this was regarded as an “equilibration period” at that step. The average of the last minute of each step was used for the applied calibration. The poorer agreement between the pre and post calibrations for methanol and ethanol is unclear. However, in our experience both methanol and ethanol residence times in PFA sample tubing are influenced by the humidity of the gas. We believe that the zero air generator (which provides VOC free air from ambient air whilst maintaining its humidity) has had insufficient time to warm up before the pre-drive calibrations. Additional text has been added to the calibration discussion to explain this.

1.13 Line 161: *Delete “instrument” in “The daily instrument ICF was derived. . .”.*

Suggestion implemented.

1.14 Line 162: *Add a space between “2” and “ppm”.*

Suggestion implemented.

1.15 Lines 175-178: *It would have probably be more intuitive to test the effect of the movement and vibration of the van by comparing measurements when the platform is in motion and when it is parked.*

We feel that testing the effect of the movement and vibration of the van on the SIFT-MS whilst driving is a sufficient test as this gave us the same conditions as when the mobile measurements were carried out.

1.16 Line 178: *Why is there no mention of the calibration procedure for the UGGA? Did you also calibrate this analyzer before and after each drive with a multi-point calibration approach? This should be discussed here.*

A quantitative measurement of mixing ratio could be determined directly from the UGGA as the measured absorption spectra is recorded and combined with measured gas temperature and pressure in the cell, effective path length and known line strength. Therefore, the UGGA could be used without performing external calibration for each drive. The UGGA was externally calibrated before and after the measurement period using external gas cylinders. A sentence has been added to the text.

“A quantitative measurement of mixing ratio could be determined directly from the UGGA without the need for calibration during each drive, but the UGGA was calibrated with external gas cylinders before and after the measurement period.”

1.17 Lines 188-190: *I do not understand this sentence: why does having more data points in bins than in your limit of detection test confirms the confidence of the SIFT-MS measurements?*

This sentence has been removed from the text.

1.18 Figures 1 and 2: *I would combine these two figures.*

We think that the figures work better separately in their respective sections for the WASP and the SIFT-MS.

1.19 Figure 6: *Change the legend of the benzene plot to better emphasize that the cps are represented by the blue dots (and only the dots) and the mixing ratio are represented by the pink line (without dots).*

Suggestion implemented.

1.20 Figure 8: *Replace "No2" and "Hono" by "NO2" and "HONO", respectively.*

Suggestion implemented

1.21 Figure A1: *Replace "No2" by "NO2".*

Suggestion implemented and "Hono" has been changed to "HONO" as suggested for Figure 8.

1.22 Table 1: *The second column most likely indicates the measurements frequency or resolution rather than the response time of the instrument.*

The header on the second column has now been changed to time resolution.