Review of amt-2021-314

"A tracer release experiment to investigate uncertainties in drone-based emission quantification for methane point sources" by Morales et al.

This paper presents a set of experiments using controlled methane releases to validate the mass balance approach to flux quantification when applied to UAV-based sampling of methane. The work rigorously characterises sources of uncertainty, bias, and error with respect to interpolation methods, instrumentation, and treatment of atmospheric transport. As such, this represents a useful contribution to UAV-based flux quantification methodology, and provides valuable guidance to others attempting similar emission measurements. However, I have a few major concerns regarding the validity of the methane measurements presented. The work appears to present methane mole fractions below the global atmospheric background mole fraction (of roughly 1.9 ppm; see e.g. Lan et al., Phil. Trans., 2021). Furthermore, calibration of the three instruments used to measure methane is not mentioned. I support publication of this work once these, and the additional comments and suggestions outlined below, have been implemented.

General comments

- I recommend swapping the phrase 'tracer release' for something more appropriate such as 'controlled release', especially in the title. So-called tracer techniques rely on the concurrent measurement of the target gas (in this case methane), and a tracer gas, and are generally referred to as tracer ratio, tracer dispersion or tracer release methods in the literature (see e.g. Mønster et al., Waste Management, 2014; Yacovitch et al., Elem. Sci. Anth., 2017 and various references in the authors' own introduction). Such an experiment, involving a tracer gas, was not performed here. Validation of plume mapping and flux methodology using a source with a known emission rate (as done in this work) is usually referred to as a controlled release experiment (see e.g. Thorpe et al., Remote Sensing of the Environment, 2016; Heltzel et al., Environments, 2020; Shah et al., Atmos. Meas. Tech., 2020).
- Drone and UAV are used interchangeably throughout the text. It should be made more clear, somewhere early on, that the two refer to the same thing and one term (probably UAV) used consistently.
- Figure text is often very difficult to read without considerably zooming in. All figure text needs to be much larger.
- Calibration of the three instruments, with respect to measurements of methane, is not mentioned at all in the text. Was calibration to a methane standard (World Meteorological Organization) performed at all, and if so what was the calibration procedure? If no calibration was performed, this throws the validity of the results into question.
- The OTM-33A method usually uses instrumented vehicles and mobile sampling to quantify a flux, although examples do exist in the literature of stationary measurements (e.g. Foster-Wittig et al., Atmospheric Environment, 2015; Shaw et al., JAPCA Waste Management, 2020). Could the authors' comment on the possibility of applying the mobile OTM-33A method to the UAV-based measurements (has this possibility been explored)?
- The plots appear to show CH_4 mole fractions of 0 ppm (Figure 3 and 5). Unless these plots are actually showing ΔCH_4 (CH_4 – background), this is impossible. The tropospheric

background mole fraction of methane is roughly 1.9 ppm (e.g. Lan et al., Phil. Trans., 2021). Could the authors explain these results?

- Abbreviations: Abbreviations for the six methods (CK, OK, PW, LW and combinations thereof) are used inconsistently. It would be useful to the reader for them to be introduced more distinctly in the methods section and then used consistently throughout the results and in figures/tables.
- Did the authors consider measuring wind speed and wind direction in situ from the drone by attaching an anemometer? Concerning wind speed estimation/interpolation, would there be any improvement using a combination of methods two and three (LW and PW)? Further, in Section 4.3, I would recommend definitively stating the abbreviated definitions used to refer to the three approaches later in the text (for example "Proj. wind" Table 2, or "PW" later).
- How representative is the range of controlled release rates used here (0.2 – 0.7 g s^{-1}) of true emissions from oil and gas facilities, or other methane sources? I would expect real emissions to have a much greater range, and that the release rates used are at the lower end of that range. Are the authors' conclusions (for example, on wind speed, wind direction, and distance from plume limits) therefore only applicable to the controlled release rates used in this work, or are they equally applicable to emissions tens, or hundreds, of times stronger? If this is not the case, then the conclusions should be caveated by stating that these results are for a limited range of emission strengths.

Specific comments

- L13: It is not clear to me what "stretched by 7 s and 0.06 seconds for every second of QCLAS measurement, respectively" is referring to here. This phenomenon is better explained in Section 5.3 and the authors should consider amending the abstract text to avoid confusion.
- L105: For comparison, it would be useful to include the instrument measurement precision for the Picarro CRDS (as mentioned for the QCLAS system on L83).
- L129: Could "not too strong winds" be quantified here e.g. greater than X m s^{-1}? I also assume this was due to the limitations of the UAV system used? Explaining the reason behind this limitation would be useful for guiding others.
- Figure 3: The wind rose is exceptionally small and doesn't add much information to the figure in its current form. The wind rose might be better viewed in a separate panel, adjacent to the top-down view of the CH_4 data. The wind rose is also not mentioned in the figure caption.
- Figure 3: Is the orange line showing the source-transect distance? This is not clear and should be made clear in the figure caption.
- Figure 3: The figure may benefit from an additional arrow illustrating the average wind direction for this flight (which ties in with the wind rose).
- L195: As for the Picarro instrument, it would be useful to include instrument characterisation (measurement precision etc.) here for the LI-COR instrument.
- L219: Extra "to" in "matching the timestamp of the anemometer to the to GPS location".

- L230: For clarification, were background CH_4 mole fractions measured upwind of the emission source, or from either side of the emission plume? It might be useful to present the measured background mole fraction value(s) and uncertainty somewhere.
- L237: Missing the word “to” between “due” and “the”.
- L271: Two identical references on this line – Tadić et al., 2015.
- L285: Missing the word “one” between “only” and “cluster”.
- Figure 5: Could the caption include which instrument was used for the methane measurements shown?
- Figure 5: The amount of panels here makes readability particularly difficult. I would recommend splitting into two separate figures: Fig. 5a as a single figure, and Fig. 5b, 5c, and 5d as a single figure. It may also be useful to have a direct side-by-side comparison of actual in situ measured CH_4 (showing sparse spatial distribution on the vertical plane) alongside the ‘predicted measured’ Krigged CH_4.
- Fig. 5b: As in general comments above, here CH_4 mole fraction is in a range of -0.06 to +0.06 ppm. These values are impossible (especially the negative values) in the troposphere. Could the authors explain these results?
- I would recommend moving equations 17 through 23 (and surrounding text) to a relevant section(s) in the methods section, as this is more Methodology than Results.
- Table 2: This table is difficult to read due to the sheer amount of values. The information is much better visualised in a plot such as Figure 6. I would consider moving the full table to the Supplement, and only including the overall results (NMAE, Bias, RMSE) for all six methods in the main manuscript.
- Table 2: Abbreviations for the six methods (e.g. CKPW) are used throughout the text but not in this table.
- Figure 8: The caption should probably mention that these are residuals in flux estimates.
- Section 5.2.2: Could this section refer to Figure 8 as well, and the comparison of different meteorological regimes?
- Table 4: Would it be useful to present the NMAE, bias, and RMSE in this table, as done in comparisons of the AirCore with the QCLAS results (Table 3), and for the comparison of the six drone-based methods (Table 2)?
- The link in the reference for US EPA 2014 goes to a page which states that “Emissions Measurement Center has Moved” – the link might need to be corrected.
- Figure S6: Would it be worth showing the plume constructed from the QCLAS data too, for comparison?

Lan et al., Phil. Trans., 2021 https://doi.org/10.1098/rsta.2020.0440
Mønster et al., Waste Management, 2014 http://dx.doi.org/10.1016/j.wasman.2014.03.025
Yacovitch et al., Elem. Sci. Anth., 2017 https://doi.org/10.1525/elementa.251

Thorpe et al., Remote Sensing of the Environment, 2016 https://doi.org/10.1016/j.rse.2016.03.032
Heltzel et al., Environments, 2020 https://doi.org/10.3390/environments7090065
Shah et al., Atmos. Meas. Tech., 2020 https://doi.org/10.5194/amt-13-1467-2020
Foster-Wittig et al., Atmospheric Environment, 2015 http://dx.doi.org/10.1016/j.atmosenv.2015.05.042
Shaw et al., JAPCA Waste Management, 2020 http://dx.doi.org/10.1080/10962247.2020.1811800