

Dear Editor,

We have compiled our point-by-point response to comments made by both reviewers below. We have also included a marked-up version of the modified manuscript, addressing all of the reviewer's concerns.

Yours faithfully,

Adil Shah

Response to reviewer 1

We thank the reviewer for taking the time to review our manuscript and for some very useful and constructive comments and suggestions. We have responded (red text) to each comment (black italic text) in turn below. We have used these comments and suggestions to update and improve a revised version of the manuscript.

Shah et al. employ two different UAV platforms to quantify known sources of CH₄ during a series of release experiments. The deployment of a lighter prototype Microportable Greenhouse Gas Analyzer (pMGGA) is new and may be interesting to other potential users as well. The authors have done a reasonable job to characterise the sensor in the laboratory; however, testing of the sensor in a harsh environment with varying temperature and pressure is missing. The near-field Gaussian plume inversion methodology presented in previous work was applied to the release experiments, and the estimated fluxes and associated uncertainties were compared with the known sources. The paper is well structured and well written, and can be published at AMT after taking into account the following comments.

We agree that the importance of temperature and pressure variability were not well described. We hope we have now strengthened the manuscript by addressing this concern. We have added a new table (Table 5), which shows that variability in cell temperature for the MGGA and pMGGA was no more than 1.3° C and 2.7° C, respectively, within each UAV flight, compared to a variability of 2.9° C noted during our 19-hour MGGA Allan variance characterisation. During the MGGA Allan variance test, there was no discernible systematic effect of either cell temperature or cell pressure on measured methane mole fraction (see section S1 of the supplement, newly added, which now describes this). With regard to pressure specifically, all measurements were made at ambient atmospheric surface pressure, which changes only very slightly up to 50 m in height and between flight surveys. Furthermore, the pMGGA was closely controlled to a fixed cell pressure. Regarding temperature and pressure conditions during sampling, we would argue that they were not “harsh” (see Table 5) and that we should have made this clearer in the original manuscript. During UAV sampling, recorded cell temperature was on average 23° C across all UAV flights compared to calibration temperatures for the MGGA and pMGGA of (31.4±0.7)° C and (24.6±0.1)° C, respectively. We thank the reviewer for raising this point. On a wider point, a truly harsh environment, such as volcanic caldera sampling, would (we agree) likely require further characterisation in extreme temperatures, but this is out of scope of this work.

General comments:

1. A weak point of the calibration in Section 3.2 is that the temperature and pressure dependence of CH₄ measurements by both the MGGA and the pMGGA is not characterised, which may be potentially much larger than the gain factor uncertainty and the offset uncertainty. In the case that the field characterisation was not performed, why not characterize it in the laboratory?

This is a very useful and valid point. To test the effect of changes in cell temperature, we measured a test gain factor of 0.9979 (at 12.7° C lower than the main calibration) and for cell pressure, we measured a test gain factor of 0.9967 (at 37.2 mbar lower than the main calibration), in the MGGA (noting, as above, that cell pressure is fixed and controlled for the pMGGA). Neither test revealed a discernible change in MGGA gain factor within uncertainty (this point is now made in the final paragraph of section 2.3). To put this into context, if sampling 12.7° C above the cell temperature conditions of the main calibration, mole fraction enhancements (above the background) of at least 2.3 ppm would be required for the temperature effects on gain factor to be larger than the instrumental noise (characterised by

the 10 Hz Allan deviation). Such an enhancement is relatively large compared with typical ambient sampling of plume emissions at a distance from a source.

In addition to these specific tests, we utilised sampling from the 19-hour MGGA Allan variance test, where variability in cell temperature and pressure was characterised. This is also now described in the supplement. Though the cell pressure and temperature variability ranges were limited to a “reasonable” range of typical ambient conditions, these parameters showed no obvious methane mole fraction correlation, with Pearson’s correlation coefficients of -0.3835 and 0.4849, respectively. Unfortunately, housekeeping data were not recorded by the pMGGA during the Allan variance test, but in principle, any correlation should be similar as the same spectroscopic fitting is used in both instruments. These results are now summarised in the final paragraph of section 2.1.

2. What’s the reason behind the exponential decay of H₂O with CH₄ mole fraction? Is it due to line interferences? It is difficult for readers to judge when the wavelengths of H₂O, CH₄ are not given. Is there an interference between CO₂ and H₂O as well? Notice that the exponential fits in Figures S3&S4 are based on very limited data points. What’s the air matrix of the 100 ppm CH₄ cylinder? Could the dependence of H₂O measurements be caused by other species?

We realise that the choice of exponential fit is arbitrary, but we chose this as it provided a suitable fit the available data points. However we also now note that that the WMO calibration scale only extends up to 5 ppm of methane, so attempting to model the water baseline any higher than this is problematic and potentially meaningless. We have therefore changed our approach and we now instead use a linear fit across the methane mole fraction range.

There is a difference between the water and methane absorption peaks of 0.2 nm which is now stated in paragraph 1 of section 2.1. This line separation is sufficient that we do not expect spectral line overlap to have a significant impact here, as the reviewer rightly points out.

All cylinders used during the water correction contained a synthetic air mix that included ambient levels of argon. While including argon in the mix is important for the calibrations (as it ensures the dry air broadening for the cylinder measurements is representative of dry air broadening for the sample measurement) we do not expect the air matrix to affect the water baseline because this represents the retrieved water vapour mole fraction for a dry sample. Therefore changes in the water baseline reflect changes in the baseline offset, not the line broadening, for this absorption line.

3. Regarding the uncertainties of the estimated fluxes σ_F , what are the fractional contributions due to individual components? This information may help reduce the uncertainties in future measurements.

This is a very useful idea. The σ_F uncertainties were derived from flux density uncertainties, which were in turn derived by combining individual statistical uncertainties in quadrature. Therefore fractional uncertainties (that sum up to form σ_F) cannot be derived. However it is possible to derive σ_F for individual uncertainty components, assuming other uncertainties to be zero. This revealed that uncertainty in wind speed was the dominant source of uncertainty. This conclusion is now given in the paragraph 3 of section 4, with individual σ_F results given in the supplement.

Detailed comments:

L92-93: The flow rate needs to be given when the e-folding time is discussed. Alternatively, the e-folding volume can be provided.

This is a very useful comment. The flow rate through both instruments is now given in Table 1.

L97-98: The unit should be ppb instead of ppm.

Thank you for spotting this. This has now been corrected.

L159-162: should make it clear that +0.27% and +1.8% are the differences between with and without the water vapour corrections, instead of an increase of measurement accuracy.

The suggested change has now been made.

L207: Equation 10 should use the molar density of dry air since CH₄ is given in dry mole fraction.

The reviewer raises a very valuable point here. We realise that we have thus far used wet methane mole fraction when this should have been converted into dry mole fraction. This requires the humidity to be taken into account when calculating molar density. We have renamed the “molar density of air” as “molar density of dry air” in paragraph 2 of section 2.4 to reflect this. We are thankful to the reviewer for raising this.

L234-239: what was the nominal flow rate? Was the flow rate recorded? What fraction of measurements on average were omitted from each flight?

Although the flow rate was not recorded, the average volumetric flow rate through the tubing can be calculated to be $(110 \pm 10) \text{ cm}^3 \text{ s}^{-1}$ from the volume of the tube and the measured lag time, which was determined as the time between a spike (cough) at the inlet and the instrument response. This has been added to paragraph 3 of section 3.3 alongside the measured flow rate through the instrument of $(27.90 \pm 0.05) \text{ cm}^3 \text{ s}^{-1}$.

16% of all sampling conducted by UAV1 was omitted due to kinks in the tubing. This now stated paragraph 3 of section 3.3.

L239-240: what was the flow rate through the pMGGA?

The flow rate was measured to be $(5.08 \pm 0.02) \text{ cm}^3 \text{ s}^{-1}$. This has been added to Table 1 and to paragraph 3 of section 3.3.

L326-330: Comparing T1.1 with T1.2 in Figure 3, I expect that larger emissions would be quantified for T1.1 and with larger uncertainties, however, the results showed the opposite. Why is that? Where are the centers of the plumes found?

The reviewer makes an interesting observation. While it may appear as though survey T1.1 should derive higher flux as there were more plume intersections, on close examination of where T1.2 intersected the plume near the ground, methane mole fraction enhancements in T1.2 were actually far higher than in T1.1. This can be seen more clearly in the time series in Figure 2, where mole fractions of up to 12.5 ppm were recorded for T1.1, whereas up to

25.3 ppm was recorded for T1.2. Both plumes were horizontally centred near the centre of the sampling plane, with T1.1 centred at -3 m and T1.2 centred at -5 cm from the centre line, according to the NGI method. The importance of taking into account the magnitude of flux density peaks, as well as the number of plume intersections is now described in a new paragraph (paragraph 4) of section 4.

L335-341: It looks that the crosswind distance is not sufficient to cover the plume, especially for UAV1. Why were the transects of UAV1 not centered? The current sampling tends to miss the center of the plume.

There may be some misunderstanding here, as there may be some ambiguity in the description of our flux quantification methodology, which we hope we have now improved. The sampling shown in Figure 3 and Figure 4 may appear as if the time-averaged plume was not centred. This was because the time-invariant plume was narrow (when nearer the source) and appeared episodically due to coincidental intersection with the UAV. When nearer to the source (approximately 50 m away), the time-invariant plume can be conceived to have had less time to disperse compared to sampling approximately 100 m away. Therefore it may appear as though the time-averaged plume was not centred in our sampling plane. We now explain this in the final paragraph of section 3.3. However, in reality, where the time-averaged plume is centred on our arbitrary sampling plane is irrelevant to the flux calculation, so long as the time-averaged plume morphology can be characterised (as is the case here).

During our sampling campaign, we ensured that we flew on a downwind sampling plane, with the UAV approximately centred downwind of the source using wind direction measurements. The flight strategy was decided on each day based on wind forecasts and on-site surface wind measurements to optimise downwind sampling of the turbulently advected emission plume. This point is now made in paragraph 2 of section 3.1 and its implications are discussed in paragraph 4 of section 3.3.

Response to reviewer 2

We thank the reviewer for taking the time to review our manuscript and for some very useful and constructive comments and suggestions. We have responded (red text) to each comment (black italic text) in turn below. We have used these comments and suggestions to update and improve a revised version of the manuscript.

Review

Testing the near-field Gaussian plume inversion flux quantification technique using unmanned aerial vehicle sampling

Adil Shah et al.

The authors use a near-field Gaussian plume inversion (NGI) technique to determine the strength of methane emission sources. In that context, they describe, characterize and deploy two commercial near-IR off-axis integrated cavity output spectroscopy (ICOS) spectrometers from ABB/Los Gatos Research. One instrument is ground-based and samples through a tube from the UAV, while the other, more lightweight instrument is deployed onboard the UAV. The authors rightly claim that more reliable techniques are needed to determine emissions from diffuse methane sources. The paper has two main technical sections: chapter 2 "Methane instrumentation and calibration" and chapter 3 "Method Testing". After struggling hard with chapter 2, I decided not to continue the review, and I suggest that the paper must be fundamentally rewritten, with respect to both language and content, before publication in AMT.

We regret that the reviewer has identified a number of issues with section 2. We have worked hard to address all of the specific points raised by the reviewer by making significant modifications to section 2 of the revised manuscript. We hope that this improves its readability.

Chapter 2 has four subsections:

2.1 Instrument inter-comparison (20 lines) Jonas, please check # lines

2.2 Water vapour correction (61 lines)

2.3 Calibration (28 lines)

2.4 Methane enhancement and uncertainty (20 lines)

Section 2.1 d contains very little actual intercomparison, and thus the title is misleading. The inter-comparison consists mainly of one table containing the manufacturer's specifications and the key numbers of one Allan Variance plot (based on laboratory measurements), which is given in SI. In the context of this paper, which attempts to use the instruments under harsh field conditions, much work could and should have been done to provide useful, real-world characterization under field conditions, beyond what the manufacturer reports. This may include (but should not be limited to) performance characteristics and comparison (see section title) of long-term stability, stability of calibration factors, temperature and pressure dependence, sensitivity to vibration (UAV), linearity, selectivity, (measured) response time, concentration range, etc.

We agree that the title of this section may be misleading. We have therefore renamed the section title, "Instrumental overview", to reflect the more general scope of the comparisons

made in this section. As the reviewer rightly points out, the Allan variance plots are useful and have now been moved to the main manuscript.

Unfortunately, we lacked the facilities to replicate some UAV sampling conditions, such as UAV vibration, in the laboratory. However, we can see no reason why vibrations this should affect measurements, especially for the MGGA which was stationary on the ground

We disagree that we sampled in harsh conditions for this work. We sampled in the bottom 50 m of the planetary boundary layer, where pressure and temperature changes are relatively small compared to laboratory conditions, especially in temperate summer months. To better convey this, the sampling conditions have now been plotted alongside the calibration conditions, for both instruments, in the supplement. Furthermore, as suggested, we now discuss the very small potential effect of temperature and pressure variations on measured methane mole fraction for the MGGA in the final paragraph of section 2.1, which showed no discernible impact on measured methane mole fraction with changes in the ambient range of cell temperature and cell pressure during the Allan variance test.

We believe that the temporal stability of calibration factors was already captured in section 2.4, where the standard deviation in offset and gain factor is given over a long calibration period in the laboratory. However the reviewer rightly points out that we failed to capture the uncertainty due to linearity which we have now determined and included in the revised manuscript in the final paragraph of section 2.4.

The measured response time depends on the length of the inlet, as air takes time to travel through the external and internal tubing. This was measured in the field. Therefore these details are provided in paragraph 3 of section 3.1 where it is more relevant to the specific instrumental set-up.

Section 2.2 is the core part of the instrumental section, which is reflected by its length. It covers an in-depth description of water vapor correction but sadly contains a number of fundamental flaws. Some examples are below, but this list is not exhaustive:

- The largest correction to obtain dry CH_4 values ($[X]^{dry}$), needed for flux calculations, is water vapor. While the authors do many tests and calculations (see below) to get this value, they never show how accurate the H_2O measurements of the instruments are. They likely have the information (based on the measurements with a dew point generator), but the corresponding performance is never shown. For such a detailed water vapor study, it would be key to deduce this effect first and then discuss (and correct) the remaining influencing parameters in a second step.

The reviewer rightly points out that we do not accurately derive water vapour mole fraction in the cell. As we conduct an empirical water correction, which by its nature is based on reported methane mole fraction, it is not necessary to derive water mole fraction accurately, based on the conclusions of Rella *et al* (2013). However this empirical correction requires the water vapour mole fraction measurement to be stable and independent of methane mole fraction. As it was observed that the water baseline was in fact affected by methane mole fraction, our water baseline corrected for this systematic effect. In order to emphasise the fact that our water correction is purely empirical and is not based on spectroscopic first principles, we have renamed the section “Empirical water vapour correction” and we have redefined v as an “empirical water correction factor”. To be clear, we do not attempt to characterise the suitability of these instruments for measurement of water vapour concentration, which would

require a separate evaluation focussed on water vapour with a different (high accuracy) water reference instrument. This may constitute future work, but the scope of this paper is on methane measurement only, with water vapour as a “problem” to be accounted for.

- *Figure S3 shows the effect of CH₄ on apparent water, due to cross sensitivity of the analyzers. Equation 1 has no physical meaning or motivation, but it produces an exponential curve, which fits the data points nicely. If there is no clear motivation, such a weak set of data may as well be described by a linear interpolation (see S3). Furthermore, applying the same parameter (ω) for the other analyzer, which shows clearly different absolute and relative behavior (S4), is not justified.*

The reviewer is correct to point out that the exponential baseline fit was arbitrary and was used to fit the data well. On reflection, we realise that the WMO calibration scale only extends up to 5 ppm, so any water baseline correction above 5 ppm is meaningless, without an equally good test for linearity up to a higher mole fraction using certified standards. To this end, we agree with the reviewer and have now applied a linear regression to the baseline from both instruments, over the range of the WMO scale.

As we observed the water baseline to behave non-linearly at 100 ppm methane mole fraction in the MGGA, we have also applied a linear fit including the additional 100 ppm data. This allowed us to assess the change in mole fraction measurements with the 100 ppm data included in the baseline to be, on average, $(0.02 \pm 0.03)\%$ higher for all MGGA sampling recorded from UAV1. This analysis is now presented in the supplement.

- *Somewhat similar is the issue in S5 and S7. It is not surprising to find a (nearly linear) dependence between v and water vapor. It is also not surprising that the residuals become a bit smaller if using a second order polynomial. However, since there is no information on the long-term stability of this system (e.g. repeatability and reproducibility), these fit characteristics are of very little use (a larger order polynomial would give even smaller residuals).*

The choice of a second order polynomial fit was based on polynomial water corrections used in previous work, using similar spectroscopic techniques (O’Shea *et al.* (2013), for example). Furthermore, O’Shea *et al.* (2013) (who used the same off-axis ICOS technique but a different instrument) found that v remains consistent over time with negligible temporal drift. This valuable point has now been made in the final paragraph of section 2.2.

- *The authors call v the water correction factor. This, again, is misleading because it contains normalization to dry conditions, which is a standard procedure, and corrections for non-perfect spectroscopy (which is specific to their analyzers).*

We realise that this may be misleading and have therefore renamed it an “empirical water correction factor”.

- *Based on these badly justified fits and corrections, the authors deduce a water uncertainty factor (σ_v). However, σ_v is the standard deviation of the mean of the residuals in S6/7. This parameter carries information about the quality of the fit of the polynomial, but no information about the quality of the water correction for a specific measurement at a given CH₄ and H₂O mole fraction.*

The reviewer rightly points out that our uncertainty is based on the quality of our polynomial fit. We simply cannot derive an uncertainty on something that is already uncertain. We have to assume that the uncertainty is represented by the goodness of the fit we can produce from empirical data. We have used the same approach as Rella *et al.* (2013) and O'Shea *et al.* (2013), as guidance of optimal practice. We have therefore clarified this in the final paragraph of section 2.2.

Sections 2.3 and 2.4 describe the calibration procedure and derive measurement uncertainty.

- Calibrated enhancement in methane mass density (E) is calculated based on parameters that are obtained in the above experimental and fit procedures. The procedures themselves are correct (but see some remarks above and below). The authors then state that their approach is especially useful because it can be used to determine measurement uncertainty (equation 11). However, this uncertainty calculation is based on laboratory measurements and badly motivated fit procedures. It does not contain non-linearity nor any changes (drifts) in time, e.g. of G and v . Furthermore, it does not include any field (in)stability considerations, nor the fact that one parameter is just assumed to be the same in MGGA and pMGGA. A realistic uncertainty budget would be highly welcome, but the approach chosen here is useless.

We thank the reviewer for highlighting the fact that we had not previously accounted for non-linearity. This has now been quantified up to 5 ppm by introducing a new non-linearity uncertainty parameter (σ_L), which has been incorporated into the enhancement uncertainty, as suggested.

With regards to drifts, the variability in G was already included in the enhancement uncertainty, represented by σ_G . As σ_G was derived from changes in G over time, we believe that it implicitly captures temporal drifts. We have now made this clear in the second paragraph of section 2.4. Additionally, we assume here that v does not drift, based on previous work using a similar instrument described by O'Shea *et al.* (2013).

Some less fundamental comments are below, illustrating that the whole paper should be strongly re-vised:

We have addressed all of the reviewer's comments and we have significantly revised the manuscript to take these comments into account.

Abstract

Page 1 – Line 12

Many parameters contribute to the uncertainty. Are they large or small, and is the contribution due to the "maybe" poorly quantified sources important or neglectable?

This sentence simply motivates the background to the paper. We are not tackling the global methane budget problem head on in the work. The choice of the words "may be" is due to the large uncertainties within the source and sink terms in global methane budget and the widely recognised difficulty in attributing the cause of these uncertainties. Section 1 expands on this and cites relevant papers that discuss the global methane problem. To make this clearer for the purposes of our manuscript and the abstract, we have added the words "potentially" and "many", so the sentence reads, "Methane emission fluxes from many facility-scale sources

may be poorly quantified, potentially leading to uncertainties in the global methane budget”. We hope this clarifies our point of view.

Page 1 – Line 14

This is not a NEW near-field Gaussian plume inversion (NGI) technique – it has been published before by the same authors.

The word “new” has been removed. We instead claim that this is the “first test” of the NGI method using “unbiased sampling”.

Page 1 – Line 20

Modified with respect to what?

This is a good point. It was not previously clear what this was with respect to. The sentence has been rephrased to make it clear that the water correction was specially adapted for the instrument. The sentence now reads, “a water vapour correction factor, specifically calculated for the instrument, was applied and is described here in detail”.

Page 1 - Line 23

Simplify sentence, e.g. the uncertainty was between ... and ...

We appreciate that this sentence can be simplified. However the quoted values are not uncertainties but rather uncertainty bounds as a fraction of the controlled emission flux. We have rephrased the sentence to simplify it, whilst maintaining the correct meaning.

Page 1 – Line 25

The term "range" implies that it is the range between 17 and 218 %, but this is not meant here.

The term “These highly conservative uncertainty ranges” has been replaced by “This range of highly conservative uncertainty bounds”. This is now less ambiguous.

Page 1 – Line 27

"flux approaches": check language

The term “flux approaches” has been replaced by “flux quantification approaches”.

Page 1 – Line 27

Replace "may perform well" by "may be a valuable alternative"

This replacement has been made.

Page 1 – Line 28

"applied to UAV sampling" should read "combined with UAV sampling"

This replacement has been made.

Introduction

Page 2 – Line 53

precision should read accuracy or precision and accuracy

The suggested change has been made.

Page 2 – Line 54

have not yet

This replacement has been made.

Page 2 – Line 60

This suggests that controlled release is the only acceptable method. Make statement more general, e.g. the method was not validated, and its uncertainty was not quantified (e.g. with controlled release experiments).

Although some methods can be mathematically correct on paper and conceptually sound, we stand by the idea that any method should be tested by sampling emissions of an accurate known flux rate to prove its utility. However we acknowledge the reviewer’s point that a controlled release from a gas cylinder is not necessarily required and other approaches could be adopted, so long as a flux is known. We have rephrased the sentence so it reads, “However this method was not tested for UAV sampling with an accurately known (controlled) methane flux rate”, to make it more general and allow for testing using any known flux.

Page 2 – Line 62

Make it clear that "Shah et al" is not any other researchers but the same group, e.g. our previous study...

The suggested change has been made.

Page 2 – Line 65

Language: the sampling does not develop the technique. e.g.: ... sampling ... used in combination with ...

The downwind sampling used during this previous study was used to develop our flux quantification technique. The technique did not exist before the sampling campaign. The measurements we acquired showed us that a new flux quantification approach was required, as previous approaches failed. We have now added this point to the sentence, to make it clearer that the new approach was a consequence of the failure of other approaches. We have also clarified that the “data-set” from the previous sampling campaign was used to develop the new method.

Methane Instrumentation and Calibration

Page 3 – Line 97

σ_{AV} is a very unusual parameter. Contains very little information because it depends on averaging time. "Allan deviation uncertainty factor" is confusing.

We realise that the Allan deviation at the maximum sampling frequency (σ_{AV}) is not so useful for direct instrument comparison, which is why we also provide the Allan deviation at the same frequency (both 1 Hz and 0.1 Hz) in Table 1. For our purposes, as both instruments sampled at their maximum sampling frequency (or minimum averaging time) and these were different for each instrument, σ_{AV} is used to quantify instrumental noise (for the instrumental averaging time used during sampling) and serves as a component of the total mole fraction enhancement uncertainty. We have renamed σ_{AV} as "sampling noise uncertainty" to make this less ambiguous.

Page 3 – Line 108

"These three effects have a net effect of decreasing $[X]_0$ in both instruments". You probably mean "decreasing $[X]$ ". Furthermore, this is only useful information if you state before which effects have a positive/negative impact. Whether the sum is a net negative effect depends on the conditions (usually dilution dominates, so the statement is not wrong).

We feel there may be some misunderstanding here in terms of the effect of water vapour on $[X]$. $[X]$ is a true mole fraction measurement (*i.e.* the true dry mole fraction in the atmosphere), independent of the instrument, whereas $[X]_0$ is an instrumental reading which can be different to the true value because it may be affected by conditions such as atmospheric water vapour and is not calibrated.

Regarding isolating the individual water vapour effects on $[X]_0$, we do not deem it empirically possible as what the instrument observes is decreasing $[X]_0$ with increasing water mole fraction. We have described all three effects in the manuscript to provide a background and motivation for our correction procedure. We believe it is unnecessary to fully characterise each of the three individual effects if the combined effect can be adequately characterised empirically. We have now clarified that we observe consistent decrease in $[X]_0$ at a range of methane and water mole fractions.

Page 3 – Line 110

If the pressure broadening coefficient should account for effects two AND three, then state so explicitly.

We have now clarified this in the first paragraph of section 2.2 that spectral overlap is expected to be an issue. Therefore we have now made it clear that the algorithm only corrects for pressure broadening in paragraph 2 of section 2.2.

Page 4 – Line 122

Has a calibration for water vapor been performed, and what are the results?

This is addressed in our response to the comment regarding Section 2.2 above. In summary, we do not characterise this instrument for water vapour measurement.

Page 4 – Line 132

"more points" means: 1.9, 2.1, 5.0 and ca. 104 ppm. It does not make sense to derive parameters of a (complicated) exponential function if there is no physical concept that supports this function.

We agree and have replaced it with a linear regression, over the range of the WMO scale (up to 5 ppm).

This H₂O baseline is very large; your data in S3 suggests that 100 ppm (dry) CH₄ results in ca. -5500 apparent ppm H₂O. Is this analyzer malfunctioning and should be returned to the manufacturer, or is it within the specifications for water vapor measurements? Using the MGGA H₂O baseline function "as-is" for another analyzer does not make sense, especially since the experimental values for the pMGGA (S4) are significantly different from those of the MGGA, indicating that the two functions may well not be the same. Alternatively, one may just acknowledge this fact and do a corresponding uncertainty budget. In the context of the large uncertainty of flux estimation with NGI, this may be appropriate. However, it would fundamentally change the arguments and line of thought in this paper.

We do not evaluate the performance of this instrument for water vapour measurement, which is beyond the scope of the manuscript. In this work, we are interested in the effect of water vapour on our target gas, methane. Water vapour influence is therefore something we simply seek to account for. However, the instrument specifications suggest that the instrument is designed to sample up to 0.03 mol_{water} mol⁻¹, with a manufacturer specified accuracy of 0.2 mmol_{water} mole⁻¹ at 1 Hz. The objective of characterising a water baseline was to ensure that the subsequent water correction was independent of the effects of methane in the cavity.

Regarding our choice of baseline, we agree with the reviewer and now use fully independent linear baselines for each individual instrument.

Page 4 – Line 135

"Having established a well characterised water baseline..."

This baseline may be well characterized. However, it corresponds to a very large correction, which can-not easily be redone in the field. Therefore, it is necessary to measure the longterm stability of the baseline and instrument-to-instrument variability. Otherwise the "well characterized" is not helpful.

This is a very useful point. We are confident that that the water baseline is relatively stable. To test this, we also conducted an Allan variance test on the dry water baseline for both instruments, at 2 ppm of methane. This revealed that the Allan variance of the MGGA and pMGGA (at the averaging time used to derive the baseline) was $\pm 16 \cdot 10^{-6}$ mol_{water} mol⁻¹ and $\pm 27 \cdot 10^{-6}$ mol_{water} mol⁻¹, respectively. This information is now added to paragraph 4 of section 2.2. These water baseline values are far smaller than water mole fractions sampled during the water correction procedure of up to $20\,000 \cdot 10^{-6}$ mol_{water} mol⁻¹, also at 2 ppm methane.

Page 4 – Line 136

What was the mole fraction of CH₄ in this cylinder?

The methane mole fraction was 2.205 ppm for the MGGA and 2.183 ppm for the pMGGA (different cylinders were used). This is now stated in paragraph 4 of section 2.2.

The H₂O baseline has a huge dependence on CH₄. Nevertheless, this experiment was only done at one CH₄ concentration only. This may not be representative and add uncertainty to the retrieval of the "real-world" measurements.

The concentration of methane in the gas cylinder was 2.2 ppm, which is similar to the current atmospheric background of approximately 1.9 ppm (depending on the time of year). Thus the water correction was performed in the most real-world conditions possible, to minimise uncertainty.

Furthermore, pressure broadening line shape changes and dilution affect the instrument methane gain factor, with increasing water mole fraction. There is no physical (instrumental) reason for these effects to manifest themselves as an effect on the instrumental methane offset. Therefore, a water correction derived at one methane mole fraction should be equally valid at other methane mole fractions.

Page 4 – Line 150

The difference between wet and dry corresponds to humidity, thus typically 1-3 %. This is not "almost equal". If it has little impact on the result, then this is because the retrieved apparent H₂O does not very strongly depend on CH₄ under the conditions used to test the simplification.

We accept that the phrase “almost equal to” may be misleading. We have therefore rephrased this as “close to”. We have also now clarified that the water correction has little overall impact on dry mole fractions in the penultimate paragraph of section 2.2, in our simplified test, where it is referred to as a “simple example”.

Page 5 – Line 156-157

The uncertainty of water measurements cannot be derived using "the water correction residual (R) from Eq. (2)". The calculated value is not the uncertainty of water measurements, nor the uncertainty after water correction. It is the standard deviation of the mean of the residuals. This parameter carries information about the quality of the fit of the polynomial, but no information about the quality of the water correction for a specific measurement at a given CH₄ and H₂O mole fraction.

The reviewer correctly points out that we do not present an uncertainty in the water correction factor, but instead an uncertainty in our fit. This was also addressed in our response above, where this was raised. We now make this clear in the final paragraph of section 2.2. Furthermore, the water “uncertainty factor” has been renamed the water “fitting uncertainty factor” to reflect this.

Page 5 – Line 160

"For example our correction was used to increase 160 the MGA measurement accuracy of [X]₀ (at 2 ppm) by +0.27%, at a humidity of 0.001 mol water mol⁻¹, and by +1.8%, at a humidity of 0.01 mol water mol⁻¹". This is a meaningless statement. It does not consider that any correct measurement includes the correction of the dilution effect. You compare apples with rotten pears.

This sentence may have been poorly written and misleading. The purpose of this sentence is to convey that we can improve the measured mole fraction, *i.e.* it more closely agrees with

the true dry mole fraction. We have replaced the phrase “increase the MGGA measurement accuracy of $[X]_0$ ” with “increase $[X]_0$ ”. This reduces ambiguity, to make it clear what the correction achieves.

Calibration

Page 5 – General Comments on Section/Calibration

Useful:

The alternating measurement of two standards. This gives the repeatability of measurements under laboratory conditions. A simple average and standard deviation would characterize this. Alternatively $(x - \text{average } x)$ vs. N (in bins) as a figure would illustrate that the values show a normal distribution.

Not useful:

The concept of uncertainty in C and G because it neglects the fact that the linearity in response is most frequently the key factor determining the uncertainty of the measurement.

Needed:

Determination of linearity.

The reviewer raises a valuable point regarding quantifying uncertainty due to non-linearity. During our main calibrations, only two standards were used (2 ppm and 5 ppm), as the World Meteorological Organisation Greenhouse Gas Calibration Scale does not extend higher than 5 ppm and reliable alternative standards cannot readily be obtained. Therefore, we have now tested the linearity within the calibration range, by sampling 5 gases with the MGGA. This showed that the residual uncertainty between measured and certified $[X]$ was on average 2.3 ppb. As we no longer have access to the pMGGA, we assumed similar linearity in both instruments as they use identical spectroscopic techniques. This test is described in section S4 of the supplement. The non-linearity uncertainty has been incorporated into the overall enhancement uncertainty.

Page 5 – Line 168-169

This does not make sense. Water vapor in reference gas cylinders should be very low if filled dry (which I assume was done here). Even if filled wet, then 60 bar/20°C should be significantly below 500 ppm. Thus, H_2O from a pressurized bottle (if treated correctly) is so low that it has very little impact on the measurements. An additional trap (as used here) may well lead to water contamination if one is not very careful.

We used a water trap as these cylinders were synthetic blends, prepared to include ambient levels of argon to provide an appropriate air matrix. This may cause a small amount of water to be present in the cylinders.

Page 6 – Line 195-196

C deviates by about 12 ppb from 0. Whether that is good or bad (and under what conditions) depends on the application. However, to say that this is “almost equal to 0” for an analyzer with an Allan Minimum below 1 ppb is bizarre.

This is a valid point. 12 ppb is indeed much larger than a 3 ppb Allan deviation (at 10 Hz) for the MGGA. However when compared to the atmospheric methane background of approximately 1800 ppb, 12 ppb is very small. The sentence now includes this qualification, to clarify our point of view.

Page 6 – Line 195

v is ca. 0.98 at 1 % H₂O. This is not "almost equal to 1". Furthermore, v in S6 contains all water-related corrections, including normalization to dry conditions. Stating that this is "almost equal to 1" does not make sense, because the correction to dry conditions should be done in any case.

We recognise the reviewer's point. We have therefore removed any reference to v from this sentence.

Additional remarks

"A lightweight wind sensor (FT205EV, FT Technologies Limited) was mounted on-board UAV1, on a carbon fibre pole 305 mm above the plane of the propellers. It recorded wind speed and direction at 4 Hz. This data was used to model the change in wind speed with height above ground level (z)."

- Was this sensor validated?

The wind sensor was used to derive a wind-height profile, for cross comparison against stationary ground-based wind measurements. The key need here was a linear response to wind speed changes. This was tested by comparing wind speed at 3.3 m from the UAV sensor to a stationary 3.3 m anemometer. A linear fit intersecting the origin agreed with all points, within uncertainty, implying linear response of the UAV wind sensor. These results are now presented in section S7 of the supplement.

- Are the measurements independent of the propellers? How was this validated?

The effect of air disturbance on the measured wind field due to rotating propellers was minimised as the wind sensor was positioned 0.3 m above the plane of the propellers. Unfortunately, we did not have required facilities (such as a wind tunnel) to fully test that this was the case. However each UAV transect was carefully planned to minimise this effect by rotating the UAV orientation at each turning point. This would effectively cancel out any potential residual net wind vector distorting the wind field due to the rotation of the propellers. This is now stated in section S7 of the supplement.

- How reliable were the direction and speed when the drone was moving? How was this determined?

The reviewer is correct to suggest that the UAV wind measurement may have been affected due to the movement of the UAV. This impacts the measured wind vector component parallel to the direction of movement of the UAV. Therefore only the wind component perpendicular to the orientation of the UAV sampling plane was used. The motion of the UAV has no impact on the wind component perpendicular to the plane of movement of the UAV. It is the perpendicular wind component that is used to construct the wind profile used to derive flux. This rationale is now explicitly stated in paragraph 1 of section S7 in the supplement.

One analyzer was using an inlet just underneath the drone. The corresponding sample is representative of the air above the drone because of downwash. Was this considered? How?

This is an astute comment. The reviewer is right to point out that mixing due to downwash could occur and that what is sampled is air from a subtly higher elevation to the position of the inlet. The inlet was 31 cm below the plane of the propellers. We have not (and cannot usefully) account for this small height difference in the calculation of flux as the air sampled is likely mixed in a thin (31 cm) layer between the plane of the rotors and the plane of sampling.

Flux of one drone --> the other just sampled further away.

We do not understand what the comment above is suggesting.

The Allan Variance plot should be in main section of the paper. Its v-shape is relevant for interpreting the measurements during flight (i.e. 10 – 30 minutes, typically).

Both plots have now been moved to the main manuscript, as suggested.

Modified manuscript

Testing the near-field Gaussian plume inversion flux quantification technique using unmanned aerial vehicle sampling

Adil Shah¹, Joseph R. Pitt¹, Hugo Ricketts^{1,2}, J. Brian Leen³, Paul I. Williams^{1,2},
Khristopher Kabbabe⁴, Martin W. Gallagher¹, Grant Allen¹

5 ¹Centre for Atmospheric Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

²National Centre for Atmospheric Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

³ABB – Los Gatos Research, 3055 Orchard Drive, San Jose, CA 95134, California, United States of America

10 ⁴School of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

Correspondence to: Adil Shah (adil.shah@manchester.ac.uk)

Abstract. Methane emission fluxes from many facility-scale sources may be poorly quantified, potentially leading to uncertainties in the global methane budget. Accurate atmospheric measurement based flux quantification is urgently required to address this. This paper describes the first test (using unbiased sampling) of a ~~new~~-near-field Gaussian plume inversion (NGI) technique, suitable for facility-scale flux quantification, using a controlled release of methane gas. Two unmanned aerial vehicle (UAV) platforms were used to perform 22 flight surveys downwind of a point-source methane gas release-~~of methane gas~~ from a regulated and flow-metered cylinder. One UAV was tethered to an instrument on the ground, while the other UAV carried an on-board high-precision prototype instrument, both of which used the same near-infrared laser technology. The performance of these instruments from UAV sampling is described. Both instruments were calibrated using certified standards, to account for variability in the instrumental gain factor. Furthermore, a water vapour correction factor, specifically calculated for the instrument, was modified approach to correcting for the effect of water vapour applied and is described here in detail. The NGI technique was used to derive emission fluxes for each UAV flight survey. We found good agreement of most NGI fluxes with the known controlled emission flux, within uncertainty, verifying the flux quantification methodology. The lower and upper NGI flux uncertainty bounds were, on average, $17\% \pm 10(1\sigma)\%$ and $227\% \pm 98(1\sigma)\%$ of the controlled emission flux, respectively and the upper NGI flux uncertainty bound was, on average, $218\% \pm 100(1\sigma)\%$ of the controlled emission flux. This range of ~~ese~~ highly conservative uncertainty ranges-bounds incorporate factors including the variability in the position of the time-invariant plume and ~~the~~-potential for under-sampling. While these average uncertainties are large compared to methods such as tracer dispersion, we suggest that UAV sampling can be highly complementary to a toolkit of flux quantification approaches and may perform well be a valuable alternative in situations where site access for tracer release is problematic. We see a tracer release applied-combined with UAV sampling as an effective combination-approach in future flux quantification studies. Successful flux quantification using this-the UAV sampling methodology described here demonstrates its future utility in identifying and quantifying emissions from methane sources such as oil and gas extraction infrastructure facilities, livestock agriculture and landfill sites, where site access may be difficult.

35 1 Introduction

The global methane budget is subject to significant uncertainties (Kirschke et al., 2013; Saunio et al., 2016b; Nisbet et al., 2019), particularly from inventory uncertainty in facility scale sources such as landfill sites (Scheutz et al., 2009), herds of cattle (Blaxter and Clapperton, 1965) and oil and gas extraction infrastructure (Brantley et al., 2014), which collectively contribute significantly to global methane emissions (Dlugokencky et al., 2011; Saunio et al., 2016a). These uncertainties can be reduced through the accurate source identification and subsequent quantification of methane emission fluxes using

top-down (atmospheric measurements based) methods, in order to validate bottom-up (inventory based) emission flux estimates (Lowry et al., 2001; Nisbet and Weiss, 2010; Allen, 2016; Desjardins et al., 2018).

45 Accurate top-down flux quantification from facility scale sources requires a combination of wind vector measurements along with in situ measurements of atmospheric methane mole fraction (Dlugokencky et al., 1994; Rigby et al., 2017). Facility-scale emission fluxes can be derived from near-field sampling (less than 500 m from the source), which may be acquired from an unmanned aerial vehicle (UAV) platform (Gottwald and Tedders, 1985). UAVs are cheap, versatile and relatively easy to use (Villa et al., 2016), compared to large manned aircraft (Illingworth et al., 2014; Lehmann et al., 2016). They can fly near to source and can be directed automatically using waypoints, to enable even and unbiased spatial sampling
50 (Greatwood et al., 2017; Feitz et al., 2018). There are three principal approaches for measuring methane mole fraction from a UAV in situ: on-board air samples can be collected for subsequent analysis (Chang et al., 2016; Greatwood et al., 2017; Andersen et al., 2018), air can be pumped through a long tube to a sensor on the ground for analysis (Broisy et al., 2017; Wolf et al., 2017; Shah et al., 2019) or air can be analysed live using a sensor mounted on-board the UAV (Berman et al., 2012; Khan et al., 2012; Nathan et al., 2015; Golston et al., 2017). Yet, a key limitation to accurate source identification and
55 flux quantification is the precision and accuracy of methane mole fraction measurements (Hodgkinson and Tatam, 2013). Miniaturised sensors suitable for UAV sampling are emerging (Villa *et al.*, 2016), but high precision lightweight in situ sensors, featuring superior techniques, such as off-axis integrated cavity output spectroscopy, have ~~previously failed to not~~ yet materialised.

60 Some studies have also used UAV remote sensing measurements to derive emission fluxes (Golston et al., 2018; Yang et al., 2018). However, to our knowledge, only Nathan et al. (2015) have derived methane emission fluxes using UAV in situ measurements. In that study, a UAV with an on-board in-situ low precision sensor (± 0.1 ppm at 1 Hz) flew in orbits around a gas compressor station, using mass balance box modelling, with geospatial kriging for interpolation, to derive the emission flux. However this method was not tested for UAV sampling ~~using with an accurate (known) (controlled) methane flux~~
65 ~~rate release of methane gas, beforehand~~. It is crucial that novel flux quantification techniques are tested by sampling a ~~controlled-known~~ flux ~~release~~, prior to investigating unknown emission sources (Desjardins et al., 2018; Feitz et al., 2018). ~~Our previous study, Shah et al. (2019) was were~~ the first ~~to~~ test of an in situ flux quantification technique using UAV sampling downwind of a controlled methane release. ~~In that study, where~~ a UAV was connected to a high precision methane analyser on the ground using 150 m of tubing (Shah et al., 2019). ~~A data-set of Two~~ two-dimensional downwind sampling
70 measurements, on a vertical flux plane, was used to develop the near-field Gaussian plume inversion (NGI) technique for flux quantification, as other flux quantification approaches failed (Shah et al., 2019). Fully manual UAV piloting was employed in this previous study to actively pursue the position of the time-invariant emission plume on the sampling plane, using mid-flight knowledge of its position. This resulted in calculated emission fluxes that were significantly positively biased compared to known emission fluxes; this represents a source of vulnerability in fully manual UAV sampling, which
75 we address in this work.

Here we test the application of the NGI method with unbiased UAV sampling of controlled methane emission sources, by flying two UAVs downwind of the release. In this work, the causes of positive flux bias reported in Shah et al. (2019) were addressed in our sampling strategy, by flying a UAV without prior knowledge of the position of the time-invariant emission
80 plume. One UAV was connected to a commercially available instrument on the ground and the other carried a lighter prototype on-board instrument (sect. 3). Both instruments were characterised and calibrated, with the effects of cell pressure and cell temperature also assessed (sect. 2). Our ~~modified~~ approach to water vapour correction is also outlined in sect. 2.

Sampled data was then used to derive NGI flux uncertainty ranges (sect. 4) for each of 22 flight surveys. In sect. 5 the success of the NGI method is assessed overall and its sampling constraints are summarised.

85 2 Methane instrumentation and calibration

2.1 Instrumental ~~overview~~ ~~inter-comparison~~

Two instruments were used to derive dry atmospheric methane mole fraction ($[X]$) measurements during UAV sampling. $[X]$ is given in units of parts-per-million (ppm) throughout this paper, which are defined here as the number of moles of methane per million moles of dry air ($10^{-6} \cdot \text{mol}_{\text{methane}} \text{mol}^{-1}$), with parts-per-million (ppb) defined as the number of moles of methane per billion moles of dry air ($10^{-9} \cdot \text{mol}_{\text{methane}} \text{mol}^{-1}$). In this section, the ABB Los Gatos Research, Inc. Micro-portable Greenhouse Gas Analyzer (MGGA) and a lighter prototype MGGA (pMGGA), designed for UAV use, are compared and characterised to assess their performance. The technical specifications of both instruments are compared in Table 1. Both instruments use off-axis integrated cavity output spectroscopy (ICOS) to derive simultaneous measurements of methane, carbon dioxide and water mole fraction, from the absorption of a near-IR (~~1650-1651~~ nm) laser, with the water and methane absorption peaks separated by 0.2 nm. The pMGGA uses an additional laser (1603 nm) to measure carbon dioxide mole fraction more accurately. Off-axis ICOS techniques reflect a tuneable laser between two mirrors in a high-finesse optical cavity, to obtain high-precision mole fraction measurements (see Paul et al. (2001) and Baer et al. (2002) for further details on off-axis ICOS).

100 The e-folding time of the high-finesse cavity in both sensors was measured here by fitting an exponential decay function to the transition from a high to low mole fraction standard gas (see Table 1 for results, with sensor flow rate also given). This represents the time taken for 63.2% of the contents of the high-finesse cavity to be replaced. The Allan variance ~~precision~~ of each sensor was also derived at various integration times, by measuring $[X]$ from a dry cylinder of compressed air for at least 17 hours of continuous sampling (see Fig. S1 and Fig. S2 for Allan variance plots), by sampling a dry gas standard continuously (17 hours and 23 minutes for the MGGA and 38 hours and 30 minutes for the pMGGA). The 1 Hz and 0.1 Hz Allan deviation for both instruments is given in Table 1. The Allan deviation sampling noise uncertainty factor ($\sigma_{AV}\sigma_{fl}$), defined here, used within the total mole fraction enhancement uncertainty (discussed in section 2.4), represents the Allan deviation at the maximum sampling frequency. $\sigma_{AV}\sigma_{fl}$ for the MGGA and pMGGA are 2.71 ppm-ppb (at 10 Hz) and 5.44 ppm-ppb (at 5 Hz), respectively, ~~with the Allan deviation at 1 Hz and 0.1 Hz given in Table 1~~. The optimum Allan variance integration time was also assessed for each sensor ((20 ± 3) s for the MGGA and (70 ± 10) s for the pMGGA); this represents the maximum sampling time before instrumental drift begins to dominate over instrumental noise. During the MGGA Allan variance test, cell temperature (which varied between 24.9° C and 27.8° C) and cell pressure (which varied between 1.0093 bar and 1.0128 bar) were also recorded to assess their correlation with $[X]$ in ppm (see Fig. S1 and Fig. S2). Correlation of both cell temperature and cell pressure was poor, with Pearson correlation coefficients of -0.4849 and -0.3835, respectively, and linear gradients of $-0.0022 \text{ ppm}^\circ \text{C}^{-1}$ and $-0.0022 \text{ ppm mbar}^{-1}$, respectively. Thus over a limited cell pressure and cell temperature range, there was no definitive correlation with $[X]$ for the MGGA under typical laboratory conditions.

2.2 ~~Water~~ Empirical water vapour correction

120 Raw wet methane mole fraction measurements ($[X]_0$) recorded by each instrument were corrected for ~~the~~ influence of atmospheric water vapour, on mole fraction retrievals. Water vapour influences dry measurements of $[X]$ for three main

reasons (Karion et al., 2013; O'Shea et al., 2013; Rella et al., 2013). First and most significantly, dilution effects occur, where the bulk presence of water reduces the quantity of methane in the cavity at a given pressure. Second, strong, broad infrared absorption bands of water can ~~lead to interference~~ with the absorption spectrum of methane, though this effect is thought to be small in this case as the spectral lines are well separated in the spectral sampling region of these instruments. Third, pressure broadening ~~of can alter the shape of the methane spectral absorption band~~ absorption peaks can occur, where due to collisional interaction between water and methane molecules changes the shape of the methane spectral absorption band, compared to pressure broadening in the absence of water in the cavity. ~~These~~ The combined impact of pressure broadening absorption band changes and dilution has ~~three effects have~~ a net effect of decreasing $[X]_0$ in both instruments, based on laboratory observations at a range of methane and water mole fractions, under typical lower boundary layer conditions.

To account for ~~these effects~~ pressure broadening absorption band changes, both the MGGA and pMGGA use an internal retrieval algorithm to derive methane mole fractions, which includes empirically ~~derived~~ estimates of the effect of pressure broadening coefficient in the presence as a function of varying water vapour mole fraction. The instruments then output both raw dry mole fraction measurements, which have additionally been corrected for the effect of mole fraction dilution by water vapour, and raw wet methane mole fraction measurements ($[X]_0$) which have not been corrected for ~~this dilution effect~~ (but are still calculated using the same empirically derived pressure broadening ~~coefficient~~ correction, as a function of water mole fraction). ~~A typical pressure broadening coefficients~~ correction (as a function of water mole fraction) are determined by the manufacturer based on experiments conducted with a sample batch of instruments, yielding an average correction values which are then applied to all instruments. However, because ~~these derived coefficients~~ correction convolves pressure broadening absorption band changes ~~line broadening~~ due to water vapour with pressure broadening absorption band changes ~~line broadening~~ due to instrument factors, there is some variability from unit to unit. Therefore, to obtain a more accurate correction for the influence of water vapour on the individual instruments used here, we apply a further empirical post-processing correction factor to the $[X]_0$ measurements (without the dilution correction) reported by the instruments, using reported measurements of water mole fraction ($[H_2O]$). Although $[H_2O]$ measurements reported by the instruments may not be an accurate representation of the true water mole fraction in the cavity, they are sufficient for an empirical correction on $[X]_0$, provided that $[H_2O]$ does not drift and is independent of dry uncalibrated methane mole fraction ($[X]_0^{dry}$). Therefore $[H_2O]$ was not calibrated against standards and an instrumental reported value was used for this empirical correction.

For the water correction to be valid, $[H_2O]$ should be independent of $[X]_0^{dry}$. However both instruments reported a small but non-zero $[H_2O]$ value when sampling dry air, which decreased with increasing $[X]_0^{dry}$. Therefore ~~before~~ a water correction could be applied, a $[H_2O]$ baseline ($[H_2O]_0$) was derived up to 5 ppm, which represents the upper limit of the World Meteorological Organisation Greenhouse Gas Scale (WMO-X2004A) for methane, as both instruments reported small, but non-zero values of $[H_2O]$, when sampling dry air. Gas from two cylinders with different methane compositions (1.901 ppm and 5.049 ppm) was dried by passing it through a dry ice water trap (a stainless-steel coil immersed in solid carbon dioxide pellets) before being sampled by both the MGGA and the pMGGA. Dry air from ~~two an~~ additional cylinders (-104 ppm and 2.167 ppm) was also sampled by the MGGA. Each gas was sampled a minimum of 11 times for 4-minute periods, from which 1-minute averages were taken. $[H_2O]_0$ ~~was observed to decay exponentially~~ decreased with ~~dry uncalibrated methane mole fraction ($[X]_0^{dry}$)~~, given by Eq. (1), where a is the water baseline offset, and b is the water baseline coefficient ~~and w is the water baseline decay factor.~~ The ~~plotted~~ data used to fit $[H_2O]_0$ is given plotted in Fig. S3 and Fig. S4.

$$(1) \quad [H_2O]_0 = a + \left(b \cdot [X]_0^{dry} e^{\frac{[X]_0^{dry}}{w}} \right)$$

165 a and b and w for both instruments are given in Table 2. w for the pMGGGA was assumed to be the same as w for the MGGGA. The same w value was used for the pMGGGA as gas was sampled by the MGGGA at more $[X]_0^{dry}$ points, resulting in an improved fit. The effect of changes in the water baseline on $[X]$ beyond the 5 ppm range was also tested up to $[X]_0^{dry}$ of approximately 100 ppm (see SI for details).

170 $[H_2O]_0$ is assumed here to be relatively constant over time. To test this, an Allan variance test was conducted on $[H_2O]_0$ for both instruments (see Fig. S5 and Fig. S6), using the same Allan variance data-set described in the previous section. This revealed a water baseline Allan deviation precision for the MGGGA and pMGGGA of $\pm 16 \cdot 10^{-6} \text{ mol}_{\text{water}} \text{ mol}^{-1}$ and $\pm 27 \cdot 10^{-6} \text{ mol}_{\text{water}} \text{ mol}^{-1}$, respectively, using a 1-minute integration time (the averaging time used for each $[H_2O]_0$ point). These 1-minute Allan deviation averages are small compared to the water vapour content of typical tropospheric air, suggesting that $[H_2O]_0$ remains relatively stable. Having established a stable and well characterised water baseline, a post-processing empirical water correction factor (v) was derived by sampling gas from a single cylinder (2.205 ppm for the MGGGA and 2.183 ppm for the pMGGGA), which was humidified to 9 fixed dew points (from 0 °C to 18 °C), using a dew point generator (LI-610, LI-COR, Inc.), following a similar experimental set-up used by O'Shea et al. (2013). The humidified gas was first sampled dry (to measure $[X]_0^{dry}$), by passing it through the dry ice water trap, and then sampled wet (to measure $[X]_0$ as a function of $[H_2O]$). An example of sampled $[X]_0$ and $[H_2O]$ measurements, used to calculate each data point, is given in Fig. S6S8. A single gas standard was deemed sufficient for this test as both dilution and pressure broadening absorption band changes affect the gain factor on methane mole fraction measurements (i.e. they do not affect the instrumental methane offset). Thus this water correction is assumed to be independent of $[X]_0^{dry}$ and solely dependent on the amount of water in the cavity.

185 $[X]_0$ is then corrected by dividing it by v , as v is effectively the ratio between $[X]_0$ and $[X]_0^{dry}$, as a function of $[H_2O]$. The ratio of $[X]_0$ to $[X]_0^{dry}$ was plotted against $([H_2O] - [H_2O]_0)$, where $[H_2O]_0$ was the water baseline measured during dry sampling (see Fig. S6S9 and Fig. S7S10). Subtracting the baseline in this analysis minimised the effects of $[X]_0^{dry}$ on $[H_2O]$. A quadratic fit was applied to both curves, with the intercept forced to unity. The first order coefficient (α) and second order coefficient (β) of the quadratic fit, given in Table 2, were then be used to derive v using Eq. (2), as a function of $[H_2O]$.

190 (2)
$$v = 1 + (\alpha \cdot ([H_2O] - [H_2O]_0)) + (\beta \cdot ([H_2O] - [H_2O]_0)^2)$$

As $[H_2O]_0$ in Eq. (2) is typically unknown, $[H_2O]_0$ defined in Eq. (1) can be substituted into Eq. (2), to yield Eq. (3).

(3)
$$v = 1 + \left(\alpha \cdot \left([H_2O] - a - \left(b \cdot [X]_0^{dry} e^{-\frac{[X]_0^{dry}}{w}} \right) \right) \right) + \left(\beta \cdot \left([H_2O] - a - \left(b \cdot [X]_0^{dry} e^{-\frac{[X]_0^{dry}}{w}} \right) \right)^2 \right)$$

As $[X]_0^{dry}$ in Eq. (3) is also unknown, an approximation that $[X]_0^{dry}$ is almost equal to $[X]_0$, in typical tropospheric humidity conditions, can be used. Thus Eq. (3) can be rewritten in terms of $[X]_0$ and $[H_2O]$, using Eq. (4).

195 (4)
$$v \approx 1 + \left(\alpha \cdot \left([H_2O] - a - \left(b \cdot [X]_0 e^{-\frac{[X]_0}{w}} \right) \right) \right) + \left(\beta \cdot \left([H_2O] - a - \left(b \cdot [X]_0 e^{-\frac{[X]_0}{w}} \right) \right)^2 \right)$$

To check on the above assumption, as a simple example (for the MGGGA), when $[H_2O]$ is 0.01 $\text{mol}_{\text{water}} \text{ mol}^{-1}$ and $[X]_0^{dry}$ is 5 ppm, Eq. (3) yields v of 0.98089 whereas Eq. (4) yields a similar value for v of 0.9808598092, for the MGGGA. This small v change supports the use of Eq. (4) as an alternative to Eq. (3), by confirming that $[X]_0^{dry}$ is close to $[X]_0$ in this simple example.

200 The fit given by Eq. (4) relies on a reliable water baseline. If the MGGGA sampled 5 ppm of dry methane and without a baseline correction, v would be 1.0020, thus representing a methane mole fraction reduction of 0.0098 ppm (at 5 ppm). However, as Eq. (4) acts to remove this small uncertainty, the residual uncertainty would be very small. In addition, The the

205 uncertainty in ~~our empirical~~the water correction ~~fit~~ was quantified using ~~each~~the water correction residual (R) from Eq. (2), to derive a water ~~fitting~~ uncertainty factor (σ_v) for each instrument (see Table 2), ~~using Eq. (5). This σ_v uncertainty is the standard deviation of the mean of the residuals and quantifies the quality of our applied water correction fits, where N is the total number of residuals.~~

$$(5) \quad -\sigma_v = \left(\frac{\sum(R^2)}{N} \right)^{\frac{1}{2}}$$

210 ~~N is the total number of residuals. Our water correction approach (given by Eq. (2)) is analogous to the approach of previous work using the same spectroscopic technique (O'Shea et al., 2013); this previous work found that the water correction is stable and does not drift. Thus an uncertainty in our water correction fit was deemed to be sufficient to characterise uncertainty empirically. To summarise, this is a purely empirical correction specific to the instruments Using the above to correct for the effects of water vapour in the measurement cavity, valid for the tested water mole fraction range of up to 0.016 mol_{water} mol⁻¹ we can increase the accuracy of $[X]_0$ measurements. For example our correction was used would to increase ~~the~~ MGGA $[X]_0^{dry}$ measurement ~~accuracy of $[X]_0$ (at 2 ppm) by +0.27%, at a humidity of 0.001 mol_{water} mol⁻¹, and by +1.8%, at a humidity of 0.01 mol_{water} mol⁻¹, thus improving measurement accuracy.~~~~

215

2.3 Calibration

220 In order to convert $[X]_0$ into $[X]$, both instruments were calibrated by sampling ~~gas from two cylinders: one contained~~ a low standard methane mole fraction ($[X]_{low}$) of 1.901 ppm and ~~the other contained~~ a high standard methane mole fraction ($[X]_{high}$) of 5.049 ppm, ~~both of which were certified WMO-X2004A standards. The composition of both cylinders was certified such that they were referenceable to the World Meteorological Organisation Greenhouse Gas Scale (WMO-X2004A).~~ Each gas was sampled intermittently for 4-minute periods of continuous sampling. The dry ice water trap was used throughout each calibration ~~as an extra precaution~~, to ensure dry gas entered the sensor cavities.

225 One-minute averages from each 4-minute sampling period were taken to derive one value of low $[X]_0^{dry}$ ($[X]_0^{dry}_{low}$) and one value of high $[X]_0^{dry}$ ($[X]_0^{dry}_{high}$) representative for each 8-minute period. The time increment between each $[X]_0^{dry}_{low}$ and $[X]_0^{dry}_{high}$ value was then interpolated from 8 minutes to 4 minutes, such that every measured value of $[X]_0^{dry}_{low}$ had a corresponding interpolated value of $[X]_0^{dry}_{high}$ and vice-versa. Individual measured and interpolated ~~values of $[X]_0^{dry}_{low}$ and $[X]_0^{dry}_{high}$ values~~ for both instruments are plotted in Fig. ~~S8-S11~~ and Fig. ~~S9-S12~~.

230

~~These measured and interpolated averages were used to calculate an average gain factor (G) and gain factor uncertainty (σ_G), from the average and standard deviation, respectively, using of a set of at least 24 individual gain factors, calculated using Eq. (6) (Pitt et al., 2016).~~

$$(6) \quad \text{gain factor} = \frac{[X]_{high} - [X]_{low}}{[X]_0^{dry}_{high} - [X]_0^{dry}_{low}}$$

235 The average offset (C) and offset uncertainty (σ_C) was calculated by taking the average and standard deviation, respectively, of individual offsets, calculated using Eq. (7) and Eq. (8) (Pitt et al., 2016).

$$(7) \quad \text{low offset} = [X]_{low} - (G \cdot [X]_0^{dry}_{low})$$

$$(8) \quad \text{high offset} = [X]_{high} - (G \cdot [X]_0^{dry}_{high})$$

240 G , σ_G , C and σ_C for both instruments are given in Table 3. ~~During these calibrations, the cell temperature of the MGGA and pMGGA were (31.4±0.7)° C and (24.6±0.1)° C, respectively, and the cell pressure of the MGGA and pMGGA were (1005.9±0.2) mbar and (614.30±0.01) mbar, respectively.~~

245 A key advantage of this calibration procedure is that ~~the~~ uncertainty in G is well quantified up to $[X]_{\text{high}}$ ~~(for assuming stable cell temperature and cell pressure) and can be incorporated in the measurement uncertainty.~~ Cell temperature and cell pressure both effect spectral fitting parameters and may consequently have an impact on G , though this effect would be smaller for the pMGGA which is pressure controlled. We believe cell temperature effects on G to be small: the MGGA was also calibrated at $(44.08 \pm 0.02)^\circ \text{C}$, yielding a gain factor of 0.9979 (see SI for details). In addition, the MGGA was calibrated at $(968.7 \pm 0.3) \text{ mbar}$, yielding a gain factor of 0.9967 (see SI for details). These short-term (test) gain factors are both similar to G (from the main calibration) of 0.9970 ± 0.0002 . Furthermore, there was no discernible correlation for both cell temperature and cell pressure during the Allan variance test for the MGGA (see above), which suggests that G is negligibly insensitive to these parameters, over the limited environmental range for the duration of the Allan variance test. Although the pMGGA was not tested in this way, we assume similar behaviour due to identical spectroscopic techniques. Nevertheless, ~~separate~~ ~~Separate~~ in-field calibrations would be preferable to enhance measurement accuracy, by characterising the effect of ~~by accounting for~~ variability in cell temperature and cell pressure on G . However there are logistical challenges with in-field calibrations, such as the need for calibration gases and the time required to perform calibrations in dynamic atmospheric temperature and pressure conditions. The laboratory calibrations described here required at least three hours of sampling ~~in order~~ to characterise ~~the~~ variability in G : this may be impractical in field conditions. ~~Therefore the calibration coefficients presented here are useful as they account for variability in G under laboratory conditions.~~

250

255

260

2.4 Methane enhancement and uncertainty

The calibration ~~and water vapour correction~~ procedures described above show that G is almost equal to 1; and C is almost equal to 0, relative to the atmospheric methane background, for both instruments (see Table 3) ~~and v is almost equal to 1 (at 5 ppm methane and a very high $0.01 \text{ mol}_{\text{water}} \text{ mol}^{-1}$ humidity) for both instruments.~~ This means that both instruments record raw $[X]_0$ measurements with very little systematic error, even when uncalibrated. Thus for most methane measurement purposes, $[X]_0$ may not need to be corrected. However in this work, G ~~and v were~~ was applied to $[X]_0$ for ~~optimal~~ instrumental improved accuracy.

265

$[X]$ can be calculated in ppm using Eq. (9).

$$270 \quad (9) \quad [X] = \left(\frac{G}{v} \cdot [X]_0 \right) + C$$

However, during flux calculation, the enhancement in methane mass density (E), in kg m^{-3} , above some background is required and was calculated here using Eq. (10). The background methane mole fraction ($[X]_b$) and corresponding background uncertainty (σ_b) can be calculated from a subset of $[X]_0$ measurements, which can be acquired from out-of-plume sampling (see SSect. 3). The molar density of dry air (ρ) and the uncertainty in ρ (σ_ρ), in units of dry mol m^{-3} , can be derived from pressure ~~and~~ temperature and humidity measurements. The molar mass of methane (M) is fixed at $0.01604 \text{ kg mol}_{\text{methane}}^{-1}$.

275

$$280 \quad (10) \quad E = \left(\frac{[X]_0}{v} [X] - [X]_b \right) \cdot \frac{G}{v} \cdot \rho \cdot M$$

To calculate the ~~The~~ uncertainty in E (σ_E), the linearity in the instrument response was characterised up to 5 ppm (i.e. the extent of the WMO-X2004A scale). This was achieved by characterising the MGGA response to five certified WMO-X2004A standards. A linear fit was then applied to measured $[X]$, with residuals used to derive an uncertainty due to non-linearity (σ_L) of $\pm 2.3 \text{ ppb}$ (see SI for further details). The linearity of the MGGA and pMGGA was assumed to be the same as they use identical spectroscopic techniques. σ_E can then be calculated by combining σ_b with the precision and accuracy uncertainty components of $[X]$, using Eq. (11). Precision is characterised by σ_{Av} ~~σ_L~~ and accuracy is characterised by $\sigma_{L \pm} \sigma_G$

280

285 and σ_v terms. σ_G also incorporates the effects of drifts, as it was derived from a prolonged sampling period over which drifts could develop.

$$(11) \quad \sigma_E = E \cdot \left(\left(\left(\sigma_{Avn}^2 + \sigma_L^2 + \left(\frac{\sigma_b \cdot G}{v} \right) \sigma_b^2 \right) \cdot \left(\frac{\rho \cdot M}{E} \right)^2 \right) + \left(\frac{\sigma_G}{G} \right)^2 + \left(\frac{\sigma_v}{v} \right)^2 + \left(\frac{\sigma_\rho}{\rho} \right)^2 \right)^{\frac{1}{2}}$$

Although M remains constant, v in Eq. (11) changes as a function of $[X]_0$ and $[H_2O]$, for each value of E . σ_C is not required in Eq. (11) as the offset cancels out in Eq. (10), when substituting in Eq. (9). This is an important advantage of calculating using E rather than $[X]$, during-in the flux analysis used in the following section.

3 Method Testing

3.1 Experimental description

295 A UAV sampling methodology for source identification and flux quantification was tested in two fields adjacent to a natural gas extraction facility in Little Plumpton (near Wesham), Lancashire, United Kingdom (+53.78785° N, -2.94758° E), prior to any drilling or hydraulic fracturing, over five sampling days in August and September 2018. A map of the field site is given in Fig. 34. The two adjacent grass fields, in which all UAV sampling took place, belong to a fully operational dairy farm. Methane was released from within the operating site at one of two controlled flux rates (F_0), from 0.25 m above ground level (see SI for controlled release details). F_0 was undisclosed during flux analysis, prior to the comparisons shown later in this paper, allowing for blind method testing.

300 Two adapted DJI Spreading Wings S1000+ octocopter UAVs (labelled UAV1 and UAV2) were used to sample the methane plume on a downwind vertical plane, roughly perpendicular to mean wind direction (see Table 4 for UAV details). The location of the UAVs in relation to the controlled release and their sampling paths was decided on each day based on public wind forecasts and on-site wind measurements, to horizontally centre (as best as possible) each UAV flight track downwind of the controlled release. ~~$[X]$ measurements of $[X]$~~ from both platforms are given in Fig. 42. UAV1 was operated using pre-programmed waypoints and ascended diagonally. Each UAV1 flight survey was composed of two parts: one flight to the right of the source (projected onto the sampling plane, perpendicular to mean wind direction) and one to the left. Meanwhile each UAV2 flight survey was composed of a single flight, to perform horizontal transects, with each transect at a roughly fixed height, up to approximately 100 m laterally away from the take-off position. 7 surveys were conducted by UAV1 (labelled, T1.1 – T1.7) and 15 surveys were conducted by UAV2 (labelled, T2.1 – T2.15). Individual flight survey details are given in Table S1 and Table S2.

315 UAV1 was connected to the MGGA on the ground, using 150 m of perfluoroalkoxy (PFA) tubing (4.76 mm inner diameter; 6.35 mm outer diameter). Air was pulled through the tubing using a small pump (NMS 030.1.2 DC-B 12V, KNF Neuberger UK Ltd), from which the MGGA subsampled. The sampling lag time between air entering the UAV air inlet ~~at the end of the tube on the UAV~~ and air entering the MGGA cavity was measured to be 25 s, with an average volumetric flow rate through the tube of $(110 \pm 10) \text{ cm}^3 \text{ s}^{-1}$ and a flow rate through the instrument (at ambient pressure) of $(27.90 \pm 0.05) \text{ cm}^3 \text{ s}^{-1}$. Both the MGGA and the pump were powered by a 12 V lead-acid battery. As the tether connected to UAV1 occasionally kinked during flight, blocking air through the tube, ~~there were periods 16% of all $[X]_0$ sampling from UAV1 was discarded that were omitted from each flight~~ (such periods were identified and recorded in the field from the flow of air to the pump). The pMGGA was mounted on board UAV2, beneath the centre frame. The sampling lag time between air entering the external air inlet and air entering the pMGGA cavity was measured to be 2 s, with a flow rate through the instrument of $(5.08 \pm 0.02) \text{ cm}^3 \text{ s}^{-1}$. The pMGGA was powered using the on-board 22.2 V UAV2 battery. Both the MGGA and pMGGA

325 transmitted live, real-time mole fraction measurements wirelessly, to a tablet computer on the ground. Satellite geolocation
was recorded by the pMGGA, on-board UAV2, simultaneous with every $[X]_0$ measurement. Satellite geolocation was
330 recorded on UAV1 by a separate on-board computer, sampling at 1 Hz. Aerial ~~plots of~~ UAV flight tracks are given in Fig.
[S10-S14](#) for UAV1 and Fig. [S11-S15](#) for UAV2.

A lightweight wind sensor (FT205EV, FT Technologies Limited) was mounted on-board UAV1, on a carbon fibre pole
330 305 mm above the plane of the propellers ([see SI for further details and testing](#)) to minimise any potential distortion of the
[wind field due to air disturbance from the rotating propellers](#) (Zhou et al., 2018). It recorded wind speed and direction at
4 Hz. ~~These measurements were~~ ~~is data was~~ used to model the change in wind speed with height above ground level (z). A
two-dimensional stationary sonic anemometer (WS500-UMB Smart Weather Sensor, G. Lufft Mess- und Regeltechnik
GmbH) was also situated on the southern boundary of the operating site (see Fig. [43](#)), (3.30±0.03) m above ground level.
335 This provided wind speed, wind direction, [relative humidity](#), temperature and pressure measurements every minute. Wind
measurements from both sensors were combined to derive the average absolute wind speed as a function of z ($WS(z)$), for the
duration of each flight survey. This is described in detail in the SI.

3.3 Flux density measurements

Each [individual](#) UAV1 survey resulted in (9±1(1 σ)) minutes of useable $[X]_0$ measurements and each UAV2 survey resulted
340 in (8±1(1 σ)) minutes of useable $[X]_0$ measurements (see Table S1 and Table S2 for individual sampling periods). This data
was prepared for flux quantification by carrying out the following steps. The $[X]_0$ timestamp from both instruments was
corrected to account for lag time. 1 Hz satellite geolocation from UAV1 was interpolated to the 10 Hz $[X]_0$ frequency of the
MGGA. [\[X\]₀ was converted into \[X\] using Eq. \(9\)](#). E was calculated with $[X]_0$ measurements from both instruments using Eq.
(10). $[X]_b$ was derived by fitting a log-normal distribution to all recorded $[X]_0$ measurements from each flight survey, using
345 the method described by Shah et al. (2019) [in our previous study](#). This background was derived from a histogram of all
useable $[X]_0$ measurements acquired during each flight experiment; a log-normal fit can usually be applied to the lowest $[X]_0$
measurements in the histogram, which represent out-of-plume sampling. The peak of the log-normal fit to these lowest $[X]_0$
measurements was taken to be $[X]_b$. ρ was derived using average temperature ~~and~~, pressure [and relative humidity](#) recorded
~~by at~~ the stationary anemometer for the duration of each flight survey, with the standard deviation in temperature ~~and~~,
350 pressure [and humidity](#) used to derive σ_ρ .

Satellite-derived altitude was corrected to obtain the height of the air inlet above ground level, by taking into account take-
off altitude and the height of the air inlet when on the ground. This step ensures that the data represent the true point of
sampling. After converting longitude and latitude from degrees into meters, metric longitude and latitude were projected
355 onto a plane perpendicular to and a plane parallel to mean wind direction, [respectively](#). Mean wind direction was derived
from the stationary anemometer for the duration of each flight survey. The coordinate projection procedure is described in
further detail by Shah et al. (2019).

In order to calculate flux, flux density, q , (in $\text{kg s}^{-1} \text{m}^{-2}$) was derived. To achieve this, each geospatially mapped E
360 measurement was combined with $WS(z)$, using Eq. (12).

$$(12) \quad q = E \cdot WS(z)$$

Geospatially mapped q , on a plane perpendicular to mean wind direction, for each flight survey, is [given-plotted](#) in Fig. [53](#)
365 for UAV1 and in Fig. [64](#) for UAV2. Measurements of $[X]$ (see Fig. [4 for a time series for each survey](#)) were not used in the
flux analysis, but are nevertheless of interest, as they show $[X]$ to generally reduce with z , as expected, to support
observations of [q enhancements](#) ~~in q~~ shown in Fig. [53](#) and Fig. [64](#).

Both Fig. 53 and Fig. 46, show significant background sampling (yellow data points), extending sufficiently far away from the position of the source projected onto the sampling plane (0 m), such that the narrow turbulently advecting time-invariant plume centre across each transect (typically manifested by ~~a peak mole fraction~~ q increase) had been passed. All of the UAV1 surveys in Fig. 3-5 took place from a similar distance from the source, of approximately 50 m. It is clear that during most UAV1 surveys, enhancements in q were concentrated near the ground (below 10 m) and close to the position of the source, projected onto the sampling plane (0 m). However T1.3 shows considerable enhancements in q above the ground (up to approximately 30 m), which was possibly due to a transient updraft. Meanwhile, the UAV2 flight surveys in Fig. 46, many of which took place approximately 100 m from the source, show large enhancements in q across the flux plane, up to approximately 15 m above the ground. Enhancements of q in Fig. 64 can also be seen at a much greater lateral distance from the source, projected onto the sampling plane. This is likely a consequence of many UAV2 flight surveys sampling at a greater distance from the source than UAV1 flight surveys, which gave the time-invariant plume more time to ~~spread~~ spread and disperse. On the other hand, UAV1 flight surveys, which took place nearer to the source, show that UAV1 intersected the time-invariant plume less often. Thus, it may appear that the UAV flight track was not centred downwind of the source, when in practice erratic variations in the position of the time-invariant plume-centre made it appear this way, as the time-invariant plume did not have time to disperse.

3.4 Flux quantification

Calculated ~~flux density~~ (q) from each flight survey was used to derive an emission flux (in units of kg s^{-1}) using the near-field Gaussian plume inversion (NGI) flux quantification technique (see Shah et al. (2019)). In principle, the NGI method accounts for turbulent variations in wind using Gaussian statistics. The method also takes into account sampling on a slightly offset sampling plane (compared to the plane perpendicular to mean wind direction) by introducing a third dimension to the traditional two-dimensional Gaussian plume model. The NGI method uses a least-squares approach to compare measured and modelled values of q . Residuals in q are minimised to output model parameters, which include an initial flux estimate (F_e).

Full details of the NGI method can be found in our previous study in Shah et al. (2019). We provide a brief overview here. The size of the time-averaged plume is assumed to increase linearly with distance from the source, by assuming q to decrease according to the inverse square law with distance (an assumption which is valid over short distances). Therefore instead of using constant crosswind and vertical dispersion terms, these terms are allowed to increase with distance from the source, with both terms being fixed at a one metre distance. The crosswind dispersion term (at 1 m) is characterised using measurements of q , rather than assumptions of atmospheric stability, as these assumptions are valid for time-averaged plumes characterised by dispersion, rather than turbulent advection. In addition, the centre of the time-averaged plume in the crosswind direction is derived from measurements of q , as the precise position of the source may be unknown. The vertical dispersion term (at 1 m) and F_e can then be acquired by inverting modelled values of q , derived by minimising residuals, as described above.

A measurement flux uncertainty (σ_F) is calculated by combining the uncertainties in individual ~~$E_T(z)$~~ and ~~$WS(z)$~~ $WS(z)$ values. A lower uncertainty bound (σ^-) is calculated using residuals between modelled and measured q values. An upper uncertainty bound (σ^+) is calculated by incorporating σ^- with the potential effects of negative flux bias due to under-sampling, using a random walk simulation. The simulation is repeated 180 times for each flight survey. In each simulation, a static Gaussian plume (simulating a prescribed arbitrary target flux) is sampled across three dimensions, where sampling is constrained to the spatial limits of UAV sampling and is limited to the UAV sampling duration. The NGI method is used to derive a flux

410 from these random walk simulations. The average fractional target flux underestimation from these simulations can be incorporated into σ^+ . Random walk flux underestimation occurs due to limited spatial sampling coverage (i.e. sampling gaps) and limited spatial sampling extent. This simulation step therefore gives an important indication of the systematic error due to potential under-sampling. All F_e , σ^- , σ^+ and σ_F values for each flight survey are given in Table S5.

4 Flux results and discussion

415 Calculated NGI emission fluxes were compared to the known (controlled) emission fluxes, using the ratio between the NGI flux uncertainty range and F_0 (see Fig. 75). As this was a blind flux analysis, F_0 was not revealed to the analysis team ~~researchers~~ prior to calculating the NGI flux uncertainty range. Fig. 75 shows that the NGI flux uncertainty range agrees well with F_0 , for most flight surveys. Only ~~four~~ three surveys (T2.1, T1.1, ~~T1.3~~ and T1.7) had a flux uncertainty range that fell short of F_0 . Although no flux uncertainty range exceeded F_0 , T2.3 spanned a large flux range, much of which fell above F_0 . Flux underestimation may be explained using the plots shown in Fig. 53 and Fig. 64, which demonstrate the following: a limited sampling duration made it possible to almost entirely avoid the emission plume, thus resulting in low flux results; 420 similarly, some flights intersected the emission plume multiple times resulting in flux overestimation in cases, although large NGI uncertainty ranges can conservatively account for this effect. Therefore it is clear that the F_e value obtained using the NGI method must not be taken at face value and the full NGI flux uncertainty range must be considered. Furthermore, the flux ranges in Fig. 75 represent uncertainty bounds of one standard deviation; it is statistically realistic to expect some discrepancy between F_0 and the NGI flux uncertainty range.

425 The flux uncertainty ranges given in Fig. 5-7 are asymmetric, although the magnitude of this asymmetry was different for flight experiments conducted by the different UAVs. σ^+ was $(0.33 \pm 0.1314(1\sigma))$ times larger than σ^- for UAV2 but was only $(0.0608 \pm 0.03(1\sigma))$ times larger for UAV1. This is because UAV2 sampled further from the source, on average, on a similar sized sampling plane to UAV1. As UAV2 was further from the emission source, the ~~instantaneous~~ time-invariant plume had a greater likelihood of extending beyond the sampling plane and being missed (beyond the horizontal edges of the sampling plane), due to spatially limited sampling extent. This potential loss of in-plume sampling may have otherwise contributed towards the overall flux, thus enhancing σ^+ . Therefore σ^+ is comparatively larger than σ^- for flights conducted by UAV2.

435 The suitability of our experimental sampling methodology can be assessed by quantifying σ_F as a fraction of F_e , which was on average $(\pm 45 \pm 8)\%$. To assess the dominant sources of σ_F , the contribution of $WS(z)$ and E uncertainty components towards it were analysed (see SI for details and results). As σ_F is derived by combining individual components in quadrature, this analysis was conducted by assuming other uncertainties to be zero. The test showed that if wind speed was the only source of uncertainty, it would on average result in $(\pm 90 \pm 8)\%$ of σ_F , therefore representing a dominant source of uncertainty. The standard deviation variability in cell temperature and cell pressure within each flight survey (see Table 5) was, on average, far smaller than maximal cell temperature changes (2.9°C) and cell pressure changes (3.5 mbar) observed during the MGGGA Allan variance test. The average cell temperature and cell pressure during each flight survey was also derived (see Fig. S18 and Fig. S19) with averages given in Table 5. The values in Table 5 are not dissimilar to conditions during calibrations (plotted in Fig. S18 and Fig. S19). As there was no discernible correlation between $[X]$ and cell temperature and cell pressure from the MGGGA Allan variance test and considering dominance of winds contributing towards σ_F , one can 440 assume that variation in cell temperature and cell pressure had negligible net effect on σ_F . Furthermore, the (poorly correlated) temperature trend from the MGGGA Allan variance test reveals a maximum uncertainty of 20 ppb for the MGGGA and 14 ppb for the pMGGGA (derived from the maximum difference between average calibration cell temperature and average UAV sampling cell temperature). These uncertainty values are far smaller than the average enhancement uncertainty

450 (expressed as a dry mole fraction) within each flight survey of (55±47) ppb (see Table S7 for individual values), though
further laboratory testing would be needed to better characterise these effects.

455 It is important to recognise the magnitude of the NGI uncertainty ranges in Fig. 57, relative to F_0 , which are due to the difficulties in inverting sparse spatial sampling to derive an emission flux, following the NGI method. These uncertainties reflect the limited sampling duration of sampling and the effects of variability in wind. While we fully acknowledge that flux uncertainty ranges in Fig. 5-7 are large, we believe that the true value of the NGI method with UAV sampling is to derive snap-shot rapid flux estimates at low cost, with an order-of-magnitude level precision, for subsequent flux investigation using more precise flux quantification techniques. Although longer periods of sampling in each flight survey may reduce the uncertainties in Fig. 57, this is practically difficult with limited UAV battery life, with little additional benefit. Tethered power or multiple UAV flights may alternatively be used, as was the case with UAV1, but wind conditions can quickly
460 change when sampling for prolonged periods with too many lengthy intervals between flights.

465 Some flux results (T1.1 for example) intersected the time-invariant plume more often than others (T1.2 for example) but resulted in a lower NGI flux range. On closer inspection of the mole fraction time series given in Fig. 4, flight surveys such as T1.2 sampled higher mole fraction enhancements (and hence q), than T1.1. However, as the time-invariant plume may have largely been centred near to the ground, it can be more difficult to distinguish from a simple plot of the flux density UAV flight track. The comparative magnitude of mole fraction enhancements is clear, on examination of the mole fraction time series. Thus it is important to take into account both the number of plume intersections and the magnitude of q during each plume intersection, when assessing NGI flux results.

470 In order to assess whether multiple flight surveys could be used effectively to capture the known controlled emission flux, within uncertainty, the upper and lower NGI uncertainty bounds were averaged for all surveys (see penultimate row of Fig. 57). The average lower NGI flux uncertainty bound as a fraction of F_0 (\bar{F}_-) was $0.2 \pm 0.1(1\sigma)$ and the average upper NGI flux uncertainty bound as a fraction of F_0 (\bar{F}_+) was $2 \pm 1(1\sigma)$, for all surveys. Thus F_0 (i.e. 1 in Fig. 57) falls comfortably within the average NGI flux uncertainty range, over 22 independent flight surveys. \bar{F}_- and \bar{F}_+ were also calculated for surveys
475 conducted by UAV1 and UAV2, separately. These separate \bar{F}_- and \bar{F}_+ values for each UAV also comfortably overlap with the \bar{F}_- and \bar{F}_+ values for all surveys combined. This suggests that the sampling strategies employed by UAV1 and UAV2 were both capable of deriving the known emission flux, with a similar degree of both lower and upper uncertainty. The percentage standard error in \bar{F}_- and \bar{F}_+ , over all 22 flight surveys, was 1312% and 109%, respectively. The large standard errors in \bar{F}_- and \bar{F}_+ may be reduced with more surveys, in order to better constrain the NGI flux uncertainty range. However
480 more precise flux estimates can be obtained using other approaches such as ~~with the~~ tracer dispersion methods. Although we recognise that the \bar{F}_- and \bar{F}_+ uncertainty averages are large, we emphasise that our methodology has been adapted for rapid flux analysis, rather than precise flux estimates for inventory publication.

485 The ability of the NGI method to calculate a target emission flux was further assessed by calculating the central flux estimate as a fraction of F_0 (F_c) for each flight survey, using Eq. (13). F_c is distinct from F_e (as a fraction of F_0), in that F_c finds the centre of an asymmetric flux uncertainty, whereas F_e is an initial flux estimate calculated using the NGI method, which does not take into account the potential effects of under-sampling, which may result in a potential negative flux bias.

$$(13) \quad F_c = \frac{F_e + \left(\frac{\sigma_+ - \sigma_-}{2}\right)}{F_0}$$

490 The mean of F_c ($\overline{F_c}$) and the mean standard error in $\overline{F_c}$ for the 22 surveys (see bottom row of Fig. 57) treats each survey as an independent quantification of the flux, with no weighting for sampling time (as flight times were broadly similar). This clearly demonstrates the improvement in flux accuracy (for a constant source) that can be obtained with greater sampling time or repeated flights, as expected. $\overline{F_c}$ was also calculated for surveys conducted by UAV1 and UAV2 separately: these separate $\overline{F_c}$ values both overlap with the combined $\overline{F_c}$ value for all flight surveys (within one standard deviation); there is no discernible difference in the NGI flux result results obtained by either UAV. This suggests that both UAV sampling
495 strategies were equally capable of delivering the same emission flux estimate, by taking the average of multiple flight surveys.

The overlap of the standard deviation in $\overline{F_c}$ (shown in Fig. 57) with the known emission flux (i.e. 1 in Fig. 57) also suggests that there was no apparent flux bias (within uncertainty) in this study. This indicates that we have successfully overcome the
500 causes of positive biases reported [in our previous study \(Shah et al., 2019\)](#) by [Shah et al. \(2019\)](#). Shah et al. (2019) sampled downwind of a controlled emission source and actively pursued the [time-invariant](#) emission plume (projected onto the sampling plane) using mid-flight knowledge of its position, inferred by releasing smoke grenades during flight surveys. However in this [current](#) work, manual sampling was avoided by either flying UAV1 using pre-programmed waypoints or by flying UAV2 using lateral transects in course-lock. Both [of the](#) approaches [presented here](#) successfully avoided biased
505 sampling.

To conclude, UAV sampling can be used to practically derive unbiased snap-shot emission fluxes with the NGI method, with an order-of-magnitude precision, by sampling on a plane perpendicular to wind direction from at least approximately 50 m away from the source. Although typical flux uncertainties were high, NGI UAV fluxes serve as an important tool for snap-shot source identification and flux quantification. Our UAV methodology fills an important gap between cheap leak
510 detection techniques (such as infrared cameras), which do not provide fluxes, and reliable flux quantification techniques (such as the tracer dispersion method), which require expensive instrumentation and may be more difficult to organise. For example, tracer methods can be problematic in cases where site access for tracer release is impossible or in cases where the plume may be lofted. The UAV methodology we describe is highly suitable for [regulatory](#) leak detection and source
515 isolation, ~~for regulatory leak detection~~, with the added capability to gauge the severity of flux leaks, for subsequent investigation using other approaches. We anticipate a combination of UAV sampling with a tracer release, where both a target gas (in this case methane) and a proxy tracer can be measured simultaneously downwind, taking advantage of vertical sampling enabled by UAVs, as a powerful future toolkit for precise [facility-scale](#) flux quantification.

5 Conclusions

520 Two UAVs were used to test the near-field Gaussian plume inversion technique for flux quantification. One UAV was connected to the MGGGA on the ground using a tether, while the other carried a new ABB pMGGGA prototype instrument on-board. Both instruments measured atmospheric methane mole fraction, which was calibrated and corrected for the influence of water vapour, following laboratory testing.

525 The flux approach was tested for 22 UAV flight surveys, by deriving fluxes from a controlled release of methane gas. This yielded successful results, with [18-19](#) out of 22 fluxes falling within the UAV-derived flux uncertainty range. This demonstrates that the near-field Gaussian plume inversion methodology used here could be used to derive emission fluxes from UAV sampling of plumes from facility-scale (point) sources, where such sources are relatively invariant over the period of such UAV sampling. The lower flux uncertainty bound was, on average, $17\% \pm 10(1\sigma)\%$ of the controlled emission

530 | flux and the upper flux uncertainty bound was, on average, $218227\% \pm 10098(1\sigma)\%$ of the controlled emission flux. Thus the known emission flux was comfortably encapsulated by the UAV flux results, within uncertainty.

A key advantage of the methodology used here is the ability to sample downwind of sources to obtain off-site mole fraction measurements. Such sampling allows for independent and portable studies of methane emissions without the need for heavy
535 | infrastructure, special permissions, runway access or prior notification. We conclude that the near-field Gaussian plume inversion flux quantification method can be used confidently in future with UAV sampling to derive snap-shot methane emission fluxes from relatively constant facility-scale sources such as oil and gas extraction infrastructure, livestock agriculture and landfill sites. An exciting future application may be the incorporation of UAV sampling within the tracer release method, where simultaneous measurement of a target gas and a proxy tracer can take advantage of vertical sampling
540 | enabled by UAVs. This avoids the limitation of current mobile vehicle sampling which cannot sample lofted plumes. Together, this may represent a powerful future toolkit for precise and efficient flux quantification.

Author contribution

AS, JRP, HR, JBL, PIW and GA carried out the field experiments. All authors designed the field experiments. AS and JRP carried out and designed the laboratory experiments. AS, JRP and JBL developed the sensor characterisation procedures.
545 | JBL provided access to the prototype sensor. AS wrote the manuscript. GA edited the manuscript. All authors contributed towards the manuscript.

Acknowledgments

Adil Shah's PhD is funded by the Natural Environment Research Council (NERC), grant reference NE/L002469/1, and is supported through a CASE partnership with the Environment Agency. NERC also provided contributions, in kind, through
550 | EQUIPT4RISK, grant reference NE/R01809X/1. We thank ABB – Los Gatos Research for loaning us the pMGGA. We also thank the farmer for allowing us to operate on his land and Cuadrilla Resources Ltd. for facilitating access.

References

- Allen, G.: Rebalancing the global methane budget, *Nature*, 538, pp.46-48, 2016.
- 555 | Andersen, T., Scheeren, B., Peters, W., and Chen, H.: A UAV-based active AirCore system for measurements of greenhouse gases, *Atmos Meas Tech*, 11, pp.2683-2699, 10.5194/amt-11-2683-2018, 2018.
- Baer, D. S., Paul, J. B., Gupta, M., and O'Keefe, A.: Sensitive absorption measurements in the near-infrared region using off-axis integrated-cavity-output spectroscopy, *Appl Phys B-Lasers O*, 75, pp.261-265, 10.1007/s00340-002-0971-z, 2002.
- Berman, E. S. F., Fladeland, M., Liem, J., Kolyer, R., and Gupta, M.: Greenhouse gas analyzer for measurements of carbon dioxide, methane, and water vapor aboard an unmanned aerial vehicle, *Sensors and Actuators B: Chemical*, 169, pp.128-135, 10.1016/j.snb.2012.04.036, 2012.
- 560 | Blaxter, K. L., and Clapperton, J. L.: Prediction of the amount of methane produced by ruminants, *British Journal of Nutrition*, 19, pp.511-522, 10.1079/BJN19650046, 1965.
- Brantley, H. L., Thoma, E. D., Squier, W. C., Guven, B. B., and Lyon, D.: Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements, *Environ Sci Technol*, 48, pp.14508-14515, 10.1021/es503070q, 2014.

- 565 Brody, C., Krampf, K., Zeeman, M., Wolf, B., Junkermann, W., Schäfer, K., Emeis, S., and Kunstmann, H.: Simultaneous multicopter-based air sampling and sensing of meteorological variables, *Atmos Meas Tech*, 10, pp.2773-2784, 10.5194/amt-10-2773-2017, 2017.
- Chang, C. C., Wang, J. L., Chang, C. Y., Liang, M. C., and Lin, M. R.: Development of a multicopter-carried whole air sampling apparatus and its applications in environmental studies, *Chemosphere*, 144, pp.484-492, 10.1016/j.chemosphere.2015.08.028, 2016.
- 570 Desjardins, R. L., Worth, D. E., Pattey, E., VanderZaag, A., Srinivasan, R., Mauder, M., Worthy, D., Sweeney, C., and Metzger, S.: The challenge of reconciling bottom-up agricultural methane emissions inventories with top-down measurements, *Agr Forest Meteorol*, 248, pp.48-59, 10.1016/j.agrformet.2017.09.003, 2018.
- 575 Dlugokencky, E. J., Steele, L. P., Lang, P. M., and Masarie, K. A.: The growth rate and distribution of atmospheric methane, *J Geophys Res-Atmos*, 99, pp.17021-17043, 10.1029/94jd01245, 1994.
- Dlugokencky, E. J., Nisbet, E. G., Fisher, R., and Lowry, D.: Global atmospheric methane: budget, changes and dangers, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369, pp.2058-2072, 10.1098/rsta.2010.0341, 2011.
- 580 Feitz, A., Schroder, I., Phillips, F., Coates, T., Negandhi, K., Day, S., Luhar, A., Bhatia, S., Edwards, G., Hrabar, S., Hernandez, E., Wood, B., Naylor, T., Kennedy, M., Hamilton, M., Hatch, M., Malos, J., Kochanek, M., Reid, P., Wilson, J., Deutscher, N., Zegelin, S., Vincent, R., White, S., Ong, C., George, S., Maas, P., Towner, S., Wokker, N., and Griffith, D.: The Ginninderra CH₄ and CO₂ release experiment: An evaluation of gas detection and quantification techniques, *Int J Greenh Gas Con*, 70, pp.202-224, 10.1016/j.ijggc.2017.11.018, 2018.
- 585 Golston, L. M., Tao, L., Brody, C., Schäfer, K., Wolf, B., McSpiritt, J., Buchholz, B., Caulton, D. R., Pan, D., Zondlo, M. A., Yoel, D., Kunstmann, H., and McGregor, M.: Lightweight mid-infrared methane sensor for unmanned aerial systems, *Appl Phys B-Lasers O*, 123, article number: 170, 10.1007/S00340-017-6735-6, 2017.
- Golston, L. M., Aubut, N. F., Frish, M. B., Yang, S. T., Talbot, R. W., Gretencord, C., McSpiritt, J., and Zondlo, M. A.: Natural Gas Fugitive Leak Detection Using an Unmanned Aerial Vehicle: Localization and Quantification of Emission Rate, *Atmosphere-Basel*, 9, article number: 333, 10.3390/Atmos9090333, 2018.
- 590 Gottwald, T. R., and Tedders, W. L.: A Spore and Pollen Trap for Use on Aerial Remotely Piloted Vehicles, *Phytopathology*, 75, pp.801-807, 10.1094/Phyto-75-801, 1985.
- Greatwood, C., Richardson, T. S., Freer, J., Thomas, R. M., MacKenzie, A. R., Brownlow, R., Lowry, D., Fisher, R. E., and Nisbet, E. G.: Atmospheric Sampling on Ascension Island Using Multirotor UAVs, *Sensors-Basel*, 17, article number: 1189, 10.3390/S17061189, 2017.
- 595 Hodgkinson, J., and Tatam, R. P.: Optical gas sensing: a review, *Meas Sci Technol*, 24, article number: 012004, 10.1088/0957-0233/24/1/012004, 2013.
- Illingworth, S., Allen, G., Percival, C., Hollingsworth, P., Gallagher, M., Ricketts, H., Hayes, H., Ładosz, P., Crawley, D., and Roberts, G.: Measurement of boundary layer ozone concentrations on-board a Skywalker unmanned aerial vehicle, *Atmos Sci Lett*, 15, pp.252-258, 10.1002/asl2.496, 2014.
- 600 Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, A., Kofler, J., Neff, D., and Tans, P.: Long-term greenhouse gas measurements from aircraft, *Atmos Meas Tech*, 6, pp.511-526, 10.5194/amt-6-511-2013, 2013.
- Khan, A., Schaefer, D., Tao, L., Miller, D. J., Sun, K., Zondlo, M. A., Harrison, W. A., Roscoe, B., and Lary, D. J.: Low Power Greenhouse Gas Sensors for Unmanned Aerial Vehicles, *Remote Sens-Basel*, 4, pp.1355-1368, 10.3390/rs4051355, 2012.
- 605 Kirschke, S., Bousquet, P., Ciais, P., Saunoy, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J. F., Langenfelds, R. L., Le Quere, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R., Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, *Nat Geosci*, 6, pp.813-823, 10.1038/NNGEO1955, 2013.
- 610 Lehmann, J. R. K., Münchberger, W., Knoth, C., Blodau, C., Nieberding, F., Prinz, T., Pancotto, V. A., and Kleinebecker, T.: High-Resolution Classification of South Patagonian Peat Bog Microforms Reveals Potential Gaps in Up-Scaled CH₄

- 615 Fluxes by use of Unmanned Aerial System (UAS) and CIR Imagery, *Remote Sens-Basel*, 8, article number: 173, 10.3390/rs8030173, 2016.
- Lowry, D., Holmes, C. W., Rata, N. D., O'Brien, P., and Nisbet, E. G.: London methane emissions: Use of diurnal changes in concentration and $\delta^{13}\text{C}$ to identify urban sources and verify inventories, *J Geophys Res-Atmos*, 106, pp.7427-7448, 10.1029/2000jd900601, 2001.
- 620 Nathan, B. J., Golston, L. M., O'Brien, A. S., Ross, K., Harrison, W. A., Tao, L., Lary, D. J., Johnson, D. R., Covington, A. N., Clark, N. N., and Zondlo, M. A.: Near-Field Characterization of Methane Emission Variability from a Compressor Station Using a Model Aircraft, *Environ Sci Technol*, 49, pp.7896-7903, 10.1021/acs.est.5b00705, 2015.
- Nisbet, E., and Weiss, R.: Top-Down Versus Bottom-Up, *Science*, 328, pp.1241-1243, 10.1126/science.1189936, 2010.
- 625 Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Lund Myhre, C., Platt, M., Allen, G., Bousquet, P., Brownlow, R., Cain, M., France, J. L., Hermansen, O., Hossaini, R., Jones, A. E., Levin, I., Manning, A. C., Myhre, G., Pyle, J. A., Vaughn, B. H., Warwick, N. J., and White, J. W. C.: Very Strong Atmospheric Methane Growth in the 4 Years 2014-2017: Implications for the Paris Agreement, *Global Biogeochem Cy*, 33, pp.318-342, 10.1029/2018GB006009, 2019.
- 630 O'Shea, S. J., Bauguutte, S. J. B., Gallagher, M. W., Lowry, D., and Percival, C. J.: Development of a cavity-enhanced absorption spectrometer for airborne measurements of CH₄ and CO₂, *Atmos Meas Tech*, 6, pp.1095-1109, 10.5194/amt-6-1095-2013, 2013.
- Paul, J. B., Lapson, L., and Anderson, J. G.: Ultrasensitive absorption spectroscopy with a high-finesse optical cavity and off-axis alignment, *Appl Optics*, 40, pp.4904-4910, 10.1364/Ao.40.004904, 2001.
- 635 Pitt, J. R., Le Breton, M., Allen, G., Percival, C. J., Gallagher, M. W., Bauguutte, S. J. B., O'Shea, S. J., Muller, J. B. A., Zahniser, M. S., Pyle, J., and Palmer, P. I.: The development and evaluation of airborne in situ N₂O and CH₄ sampling using a quantum cascade laser absorption spectrometer (QCLAS), *Atmos Meas Tech*, 9, pp.63-77, 10.5194/amt-9-63-2016, 2016.
- Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, *Atmos Meas Tech*, 6, pp.837-860, 10.5194/amt-6-837-2013, 2013.
- 640 Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A. L., Manning, A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Muhle, J., Weiss, R. F., Fraser, P. J., Steele, L. P., Krummel, P. B., McCulloch, A., and Park, S.: Role of atmospheric oxidation in recent methane growth, *P Natl Acad Sci USA*, 114, pp.5373-5377, 10.1073/pnas.1616426114, 2017.
- 645 Saunois, M., Bousquet, P., Poulter, B., Pregon, A., Ciais, P., Canadell, J. G., Dlugokencky, E. J., Etiope, G., Bastviken, D., Houweling, S., Janssens-Maenhout, G., Tubiello, F. N., Castaldi, S., Jackson, R. B., Alexe, M., Arora, V. K., Beerling, D. J., Bergamaschi, P., Blake, D. R., Brailsford, G., Brovkin, V., Bruhwiler, L., Crevoisier, C., Crill, P., Covey, K., Curry, C., Frankenberg, C., Gedney, N., Höglund-Isaksson, L., Ishizawa, M., Ito, A., Joos, F., Kim, H. S., Kleinen, T., Krummel, P., Lamarque, J. F., Langenfelds, R., Locatelli, R., Machida, T., Maksyutov, S., McDonald, K. C., Marshall, J., Melton, J. R., Morino, I., Naik, V., O'Doherty, S., Parmentier, F. J. W., Patra, P. K., Peng, C., Peng, S., Peters, G. P., Pison, I., Prigent, C., Prinn, R., Ramonet, M., Riley, W. J., Saito, M., Santini, M., Schroeder, R., Simpson, I. J., Spahni, R., Steele, P., Takizawa, A., Thornton, B. F., Tian, H., Tohjima, Y., Viovy, N., Voulgarakis, A., van Weele, M., van der Werf, G. R., Weiss, R., Wiedinmyer, C., Wilton, D. J., Wiltshire, A., Worthy, D., Wunch, D., Xu, X., Yoshida, Y., Zhang, B., Zhang, Z., and Zhu, Q.: The global methane budget 2000–2012, *Earth Syst Sci Data*, 8, pp.697-751, 10.5194/essd-8-697-2016, 2016a.
- 650 Saunois, M., Jackson, R. B., Bousquet, P., Poulter, B., and Canadell, J. G.: The growing role of methane in anthropogenic climate change, *Environ Res Lett*, 11, article number: 120207, 2016b.
- 655 Scheutz, C., Kjeldsen, P., Bogner, J. E., De Visscher, A., Gebert, J., Hilger, H. A., Huber-Humer, M., and Spokas, K.: Microbial methane oxidation processes and technologies for mitigation of landfill gas emissions, *Waste Manage Res*, 27, pp.409-455, 10.1177/0734242X09339325, 2009.
- 660 Shah, A., Allen, G., Pitt, J. R., Ricketts, H., Williams, P. I., Kabbabe, K., Hollingsworth, P., Helmore, J., Robinson, R., Finlayson, A., Rees-White, T. C., Beaven, R. P., Scheutz, C., and Bourn, M.: A Near-Field Gaussian Plume Inversion Flux Quantification Method, Applied to Unmanned Aerial Vehicle Sampling, *Atmosphere-Basel*, 10, article number: 396, 10.3390/atmos10070396, 2019.

Villa, T. F., Gonzalez, F., Miljevic, B., Ristovski, Z. D., and Morawska, L.: An Overview of Small Unmanned Aerial Vehicles for Air Quality Measurements: Present Applications and Future Perspectives, *Sensors-Basel*, 16, article number: 1072, 10.3390/S16071072, 2016.

- 665 Wolf, B., Chwala, C., Fersch, B., Garvelmann, J., Junkermann, W., Zeeman, M. J., Angerer, A., Adler, B., Beck, C., Brosy, C., Brugger, P., Emeis, S., Dannenmann, M., De Roo, F., Diaz-Pines, E., Haas, E., Hagen, M., Hajsek, I., Jacobeit, J., Jagdhuber, T., Kalthoff, N., Kiese, R., Kunstmann, H., Kosak, O., Krieg, R., Malchow, C., Mauder, M., Merz, R., Notarnicola, C., Philipp, A., Reif, W., Reineke, S., Rödiger, T., Ruehr, N., Schäfer, K., Schrön, M., Senatore, A., Shupe, H., Völksch, I., Wanninger, C., Zacharias, S., and Schmid, H. P.: The SCALEX Campaign Scale-Crossing Land Surface and Boundary Layer Processes in the TERENO-preAlpine Observatory, *B Am Meteorol Soc*, 98, pp.1217-1234, 10.1175/Bams-D-15-00277.1, 2017.

Yang, S. T., Talbot, R. W., Frish, M. B., Golston, L. M., Aubut, N. F., Zondlo, M. A., Gretencord, C., and McSpurr, J.: Natural Gas Fugitive Leak Detection Using an Unmanned Aerial Vehicle: Measurement System Description and Mass Balance Approach, *Atmosphere-Basel*, 9, article number: 383, 10.3390/Atmos9100383, 2018.

- 675 Zhou, S. D., Peng, S. L., Wang, M., Shen, A., and Liu, Z. H.: The Characteristics and Contributing Factors of Air Pollution in Nanjing: A Case Study Based on an Unmanned Aerial Vehicle Experiment and Multiple Datasets, *Atmosphere-Basel*, 9, article number: 343, 10.3390/Atmos9090343, 2018.

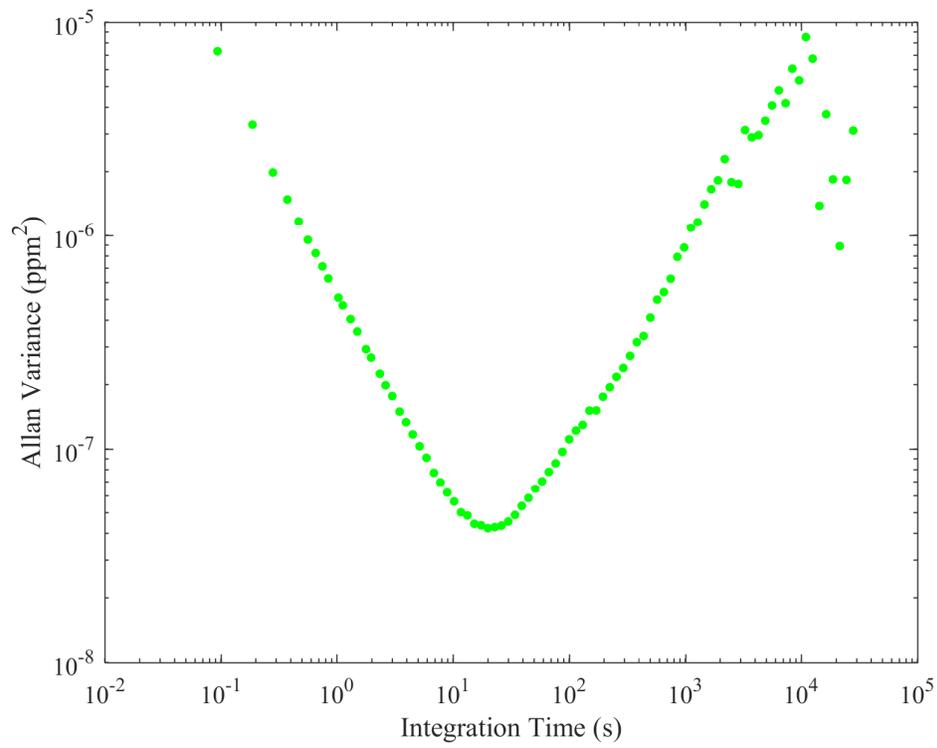
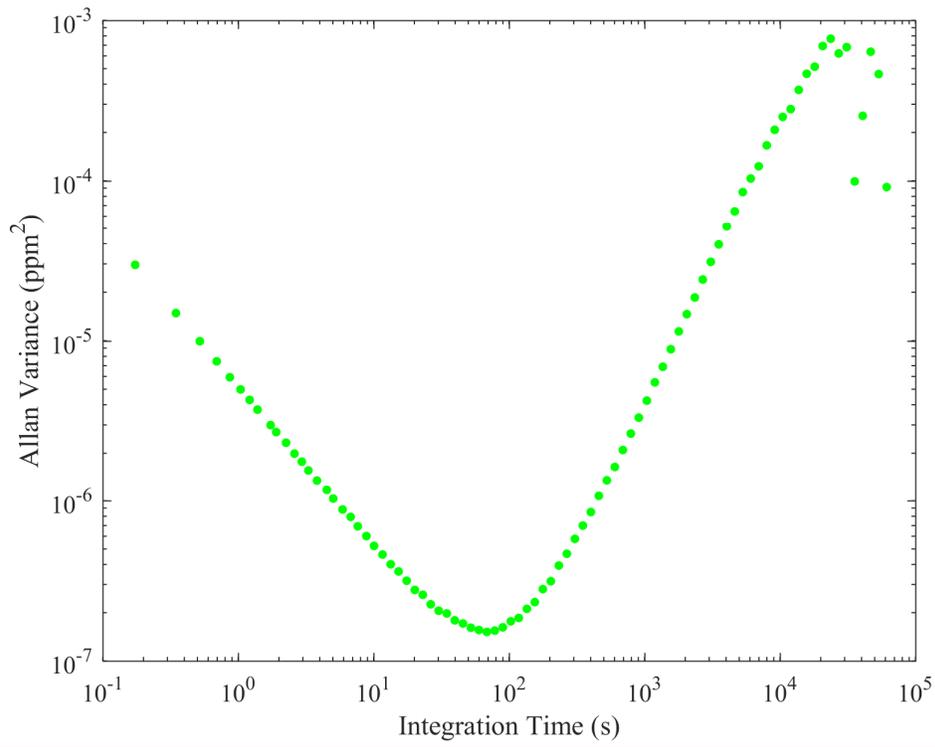
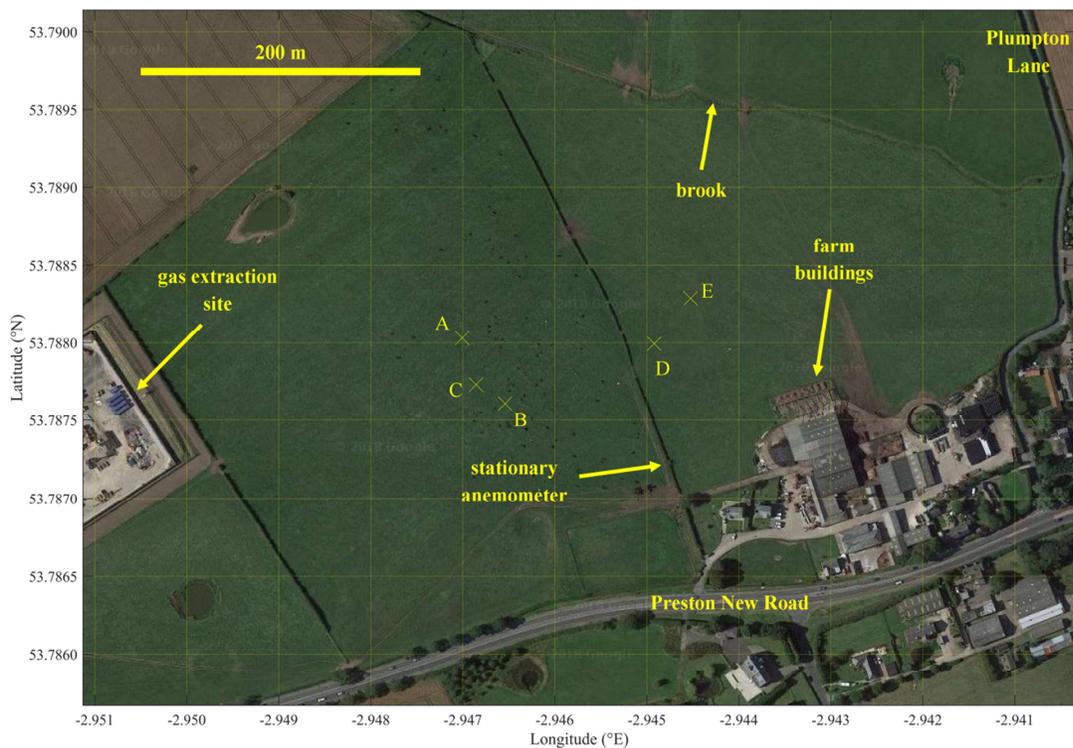


Figure 1. Allan variance for the MGGA plotted against integration time on logarithmic axes.



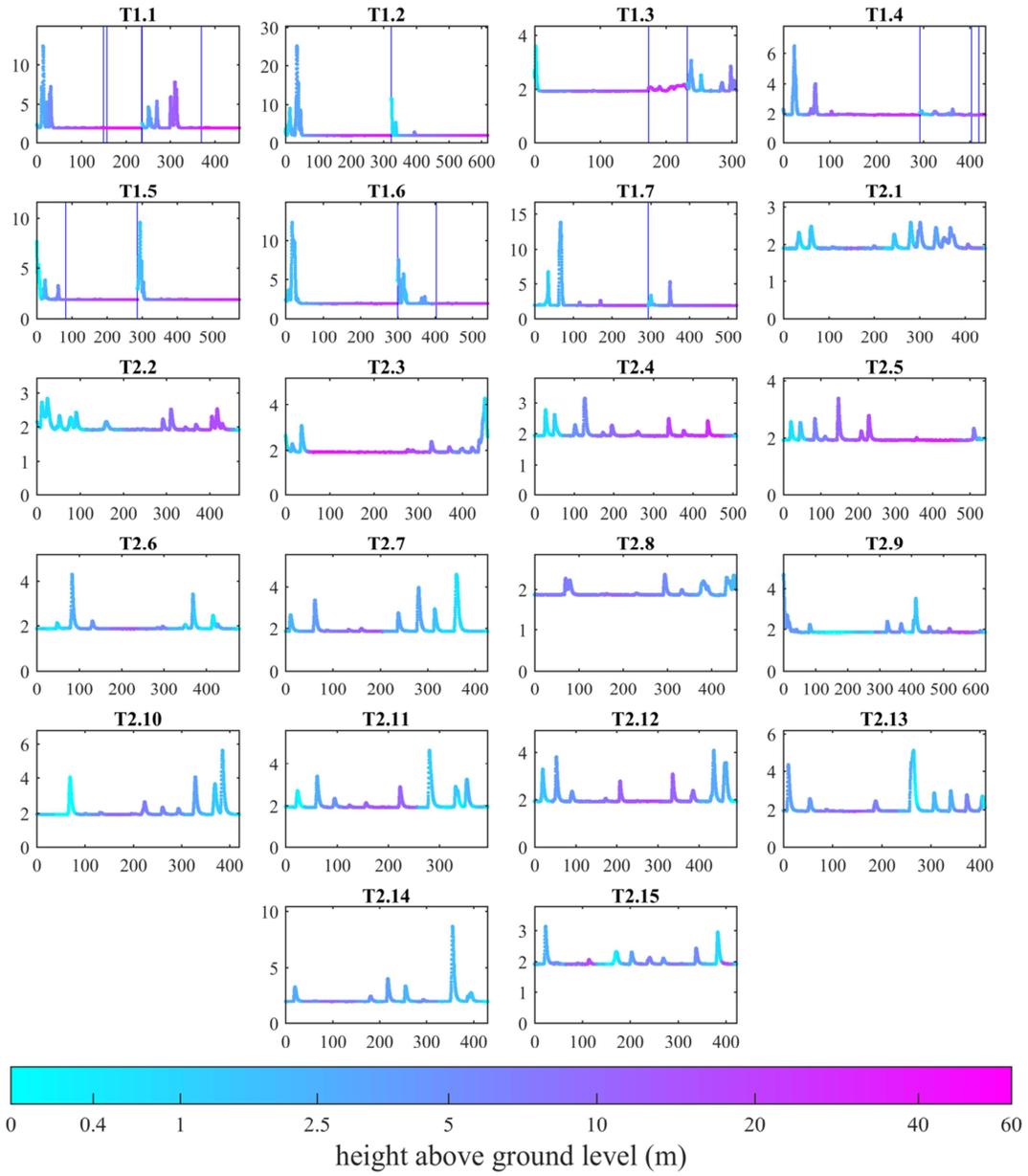
685

Figure 2. Allan variance for the pMGGA plotted against integration time on logarithmic axes.

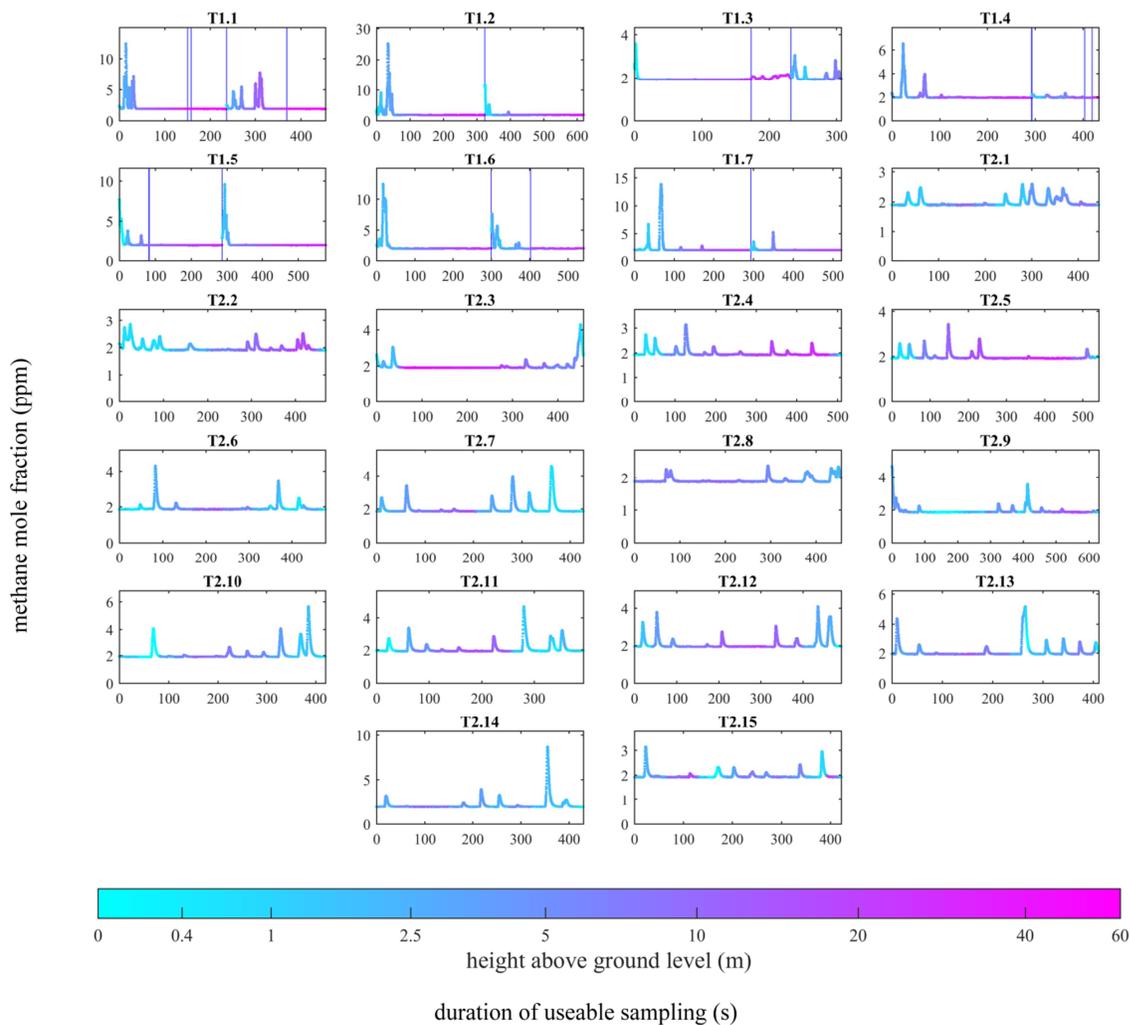


690 **Figure 3. The two fields used for UAV sampling. The map extends 0.71 km horizontally and 0.50 km vertically. The controlled release points are marked by labelled crosses (see Table S3 for details). The background image is taken from Google Maps (imagery (2017): DigitalGlobe, GetMapping plc, Infoterra Ltd & Bluesky, The GeoInformation Group).**

methane mole fraction (ppm)

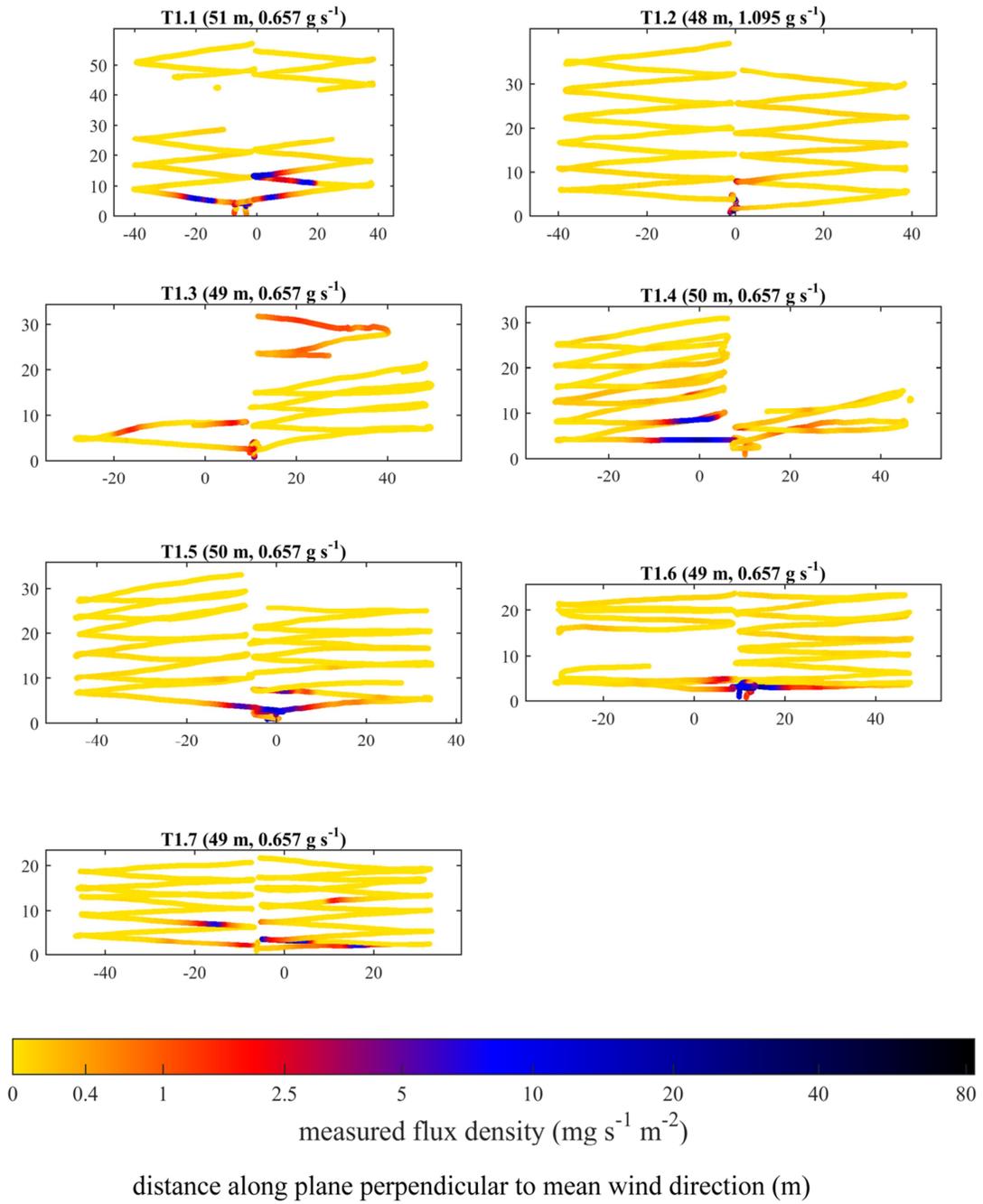


duration of useable sampling (s)



695 **Figure 4. [X] measurements acquired by the MGGA and the pMGGA, as a function of sampling duration, for each flight survey, with sampling height above ground level also plotted (coloured dots). A logarithmic colour legend has been used. Vertical blue lines indicate an interruption in continuous sampling.**

height above ground level (m)



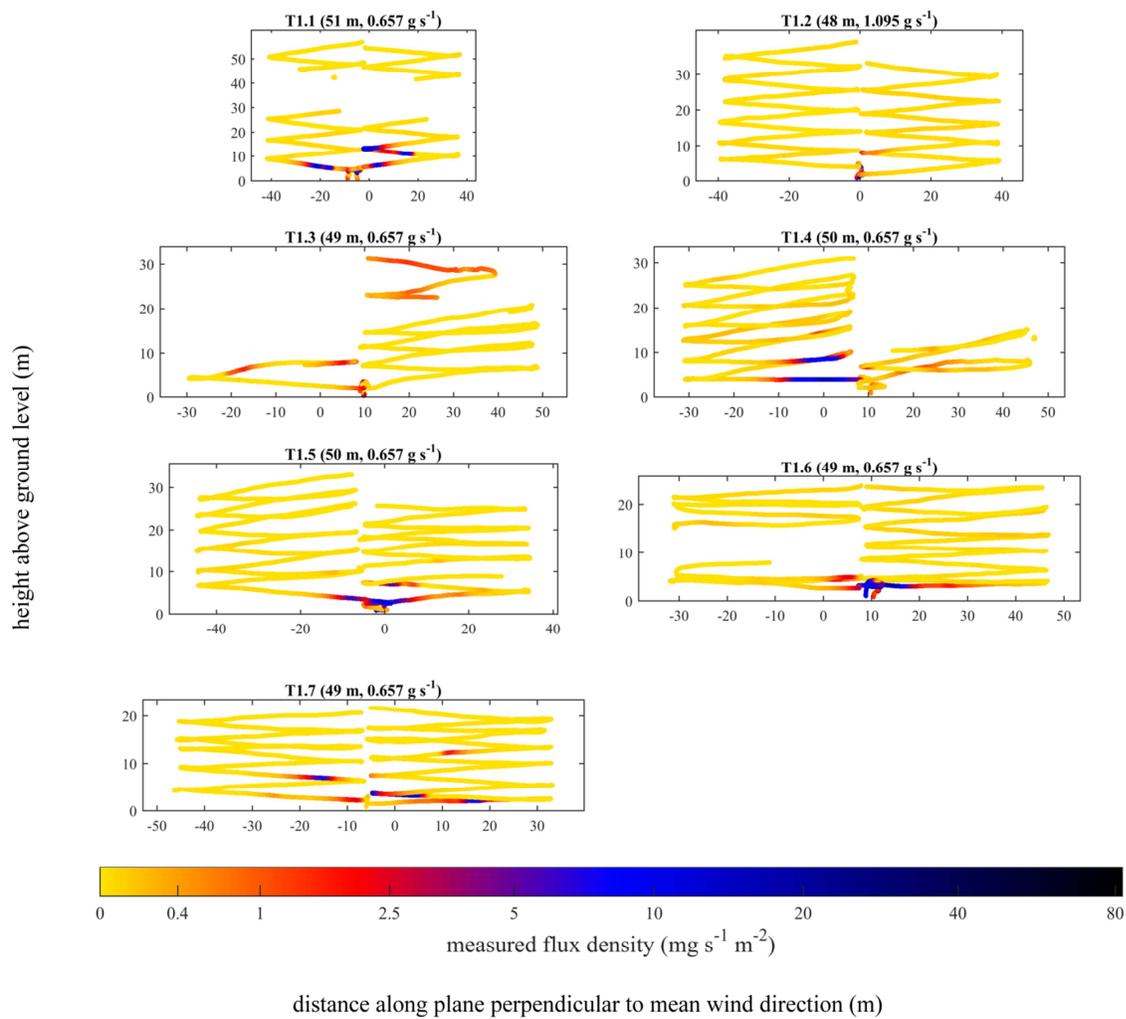
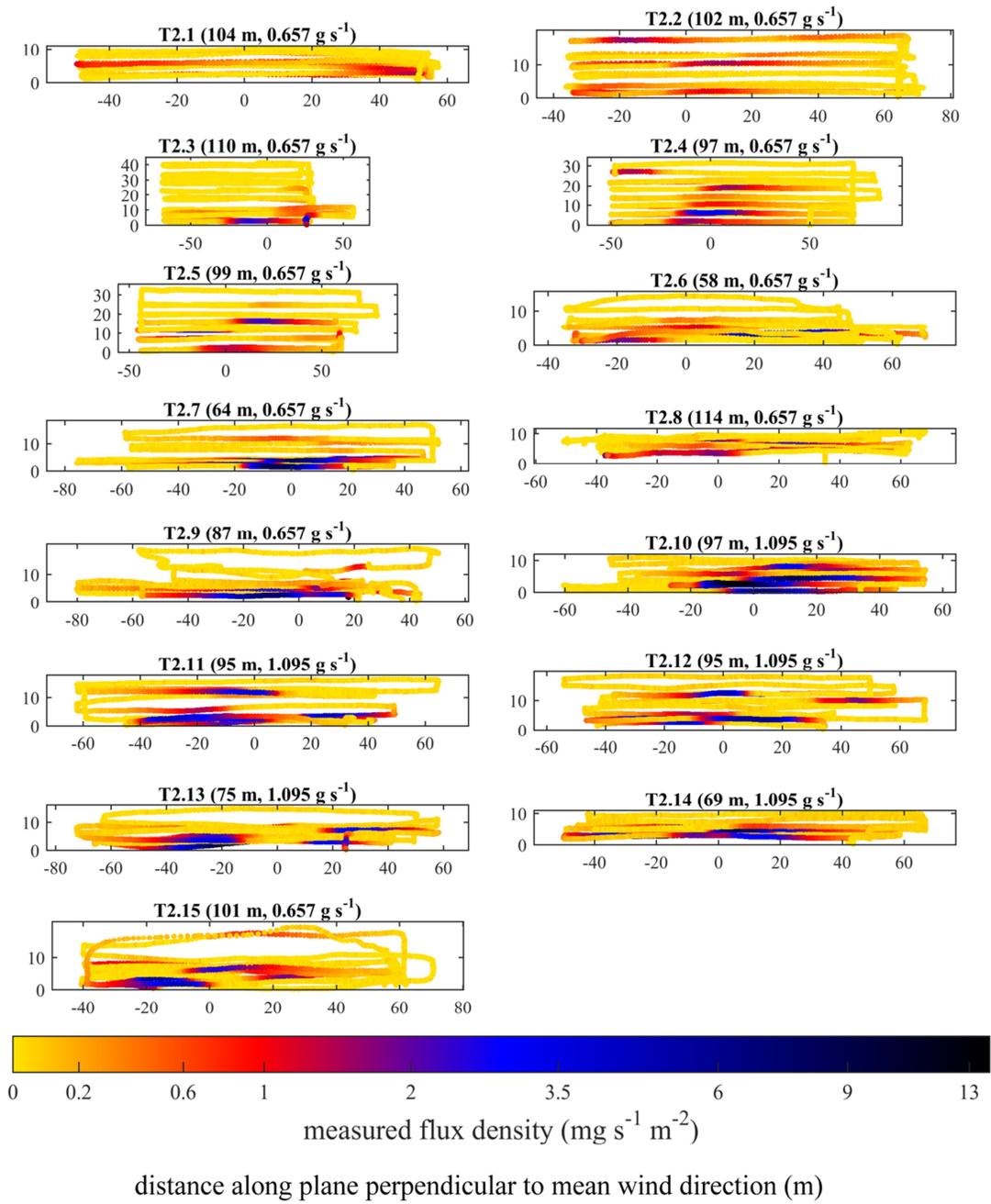
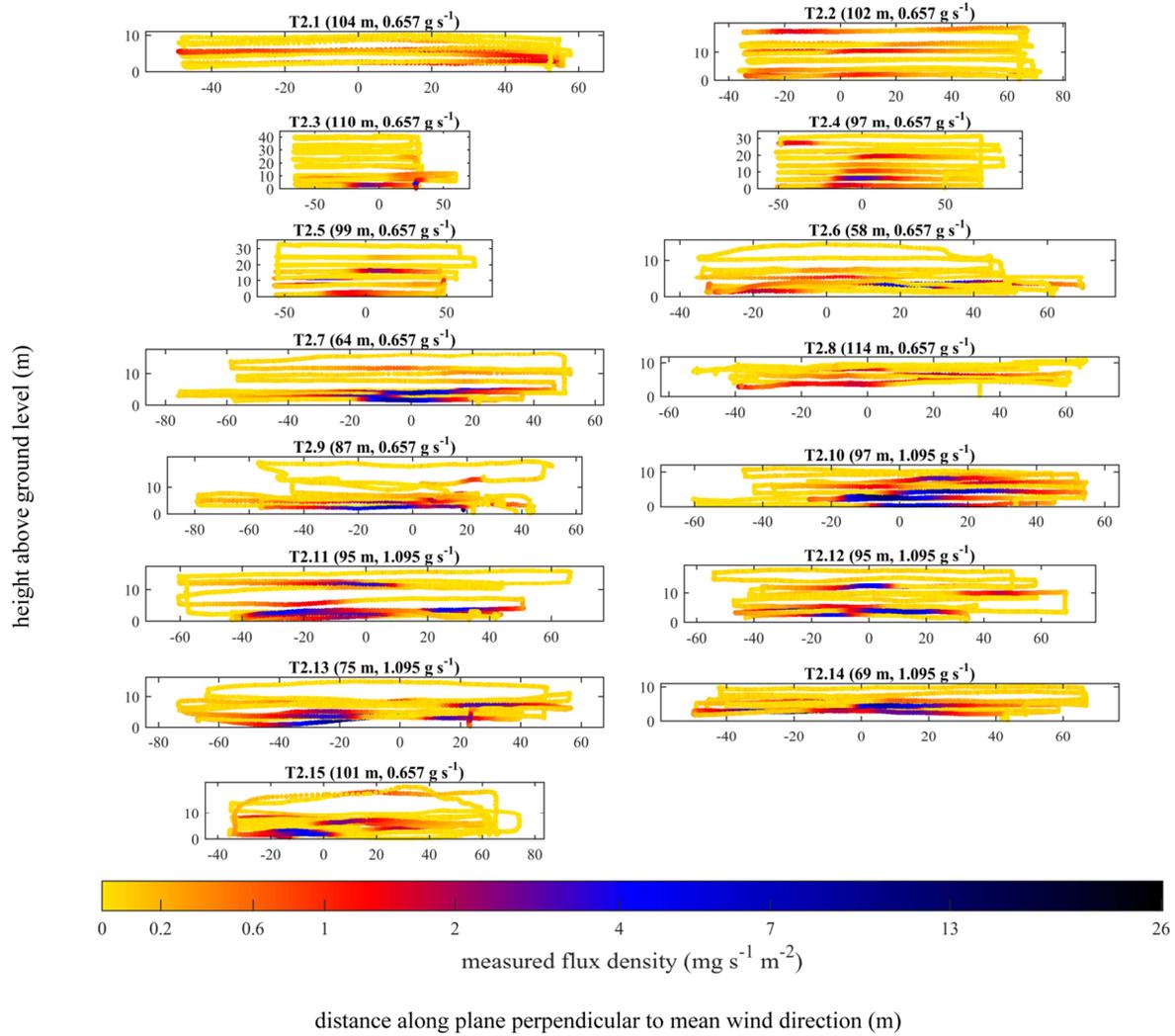


Figure 5. UAV1 flight tracks (coloured dots), with the colour corresponding to q . Periods in which the tubing **inlet** kinked have been removed. A logarithmic colour legend has been used. The position of the source projected on the plane perpendicular to mean wind direction has been set to a reference of 0 m. The controlled emission flux and the parallel distance of the sampling plane from the source (weighted to the position of q enhancements) are given in brackets.

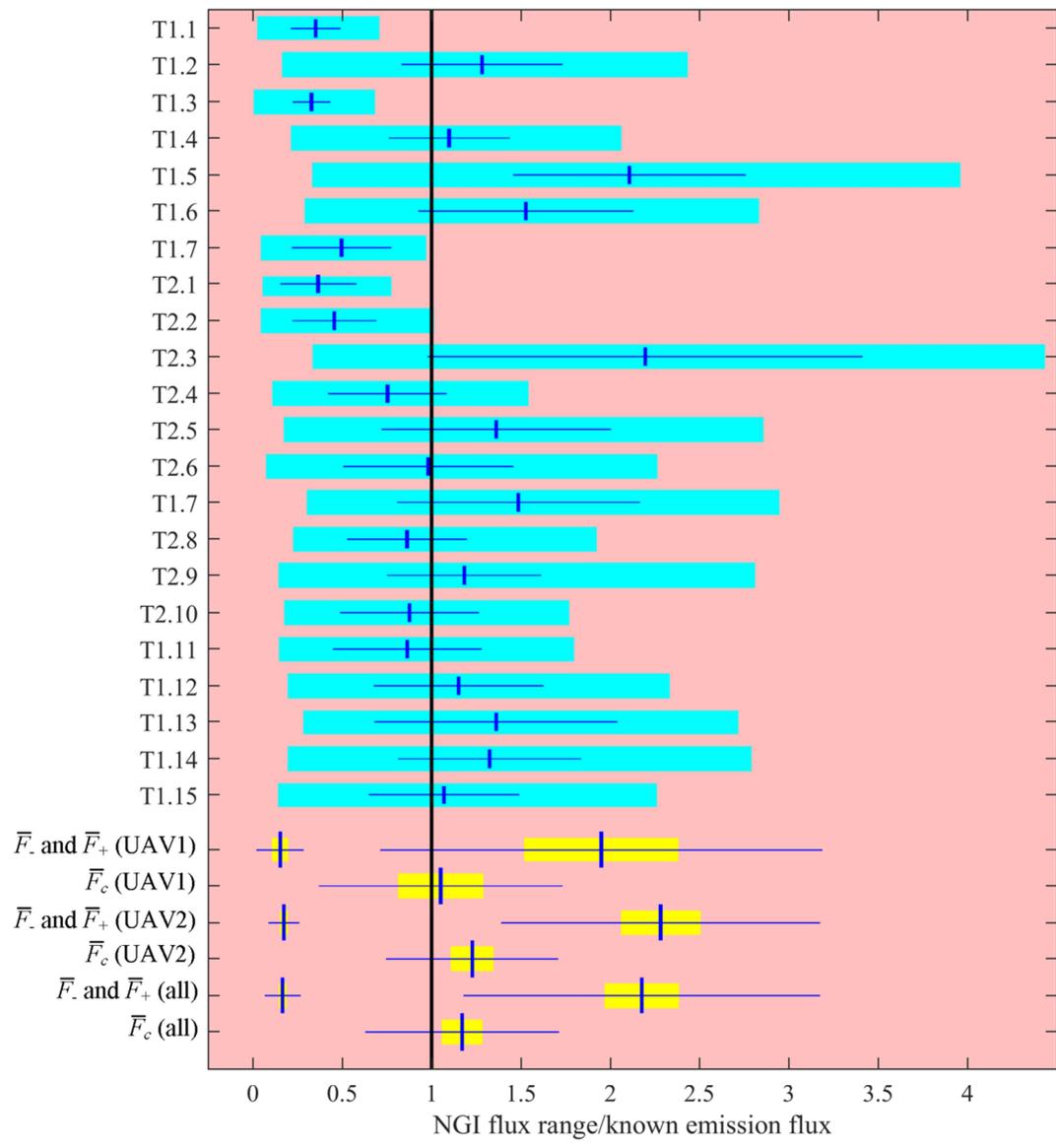
705

height above ground level (m)





710 **Figure 6.** UAV2 flight tracks (coloured dots), with the colour corresponding to q . The position of the source projected on the plane perpendicular to mean wind direction has been set to a reference of 0 m. The controlled emission flux and the parallel distance of the sampling plane from the source (weighted to the position of q enhancements) are given in brackets.



715

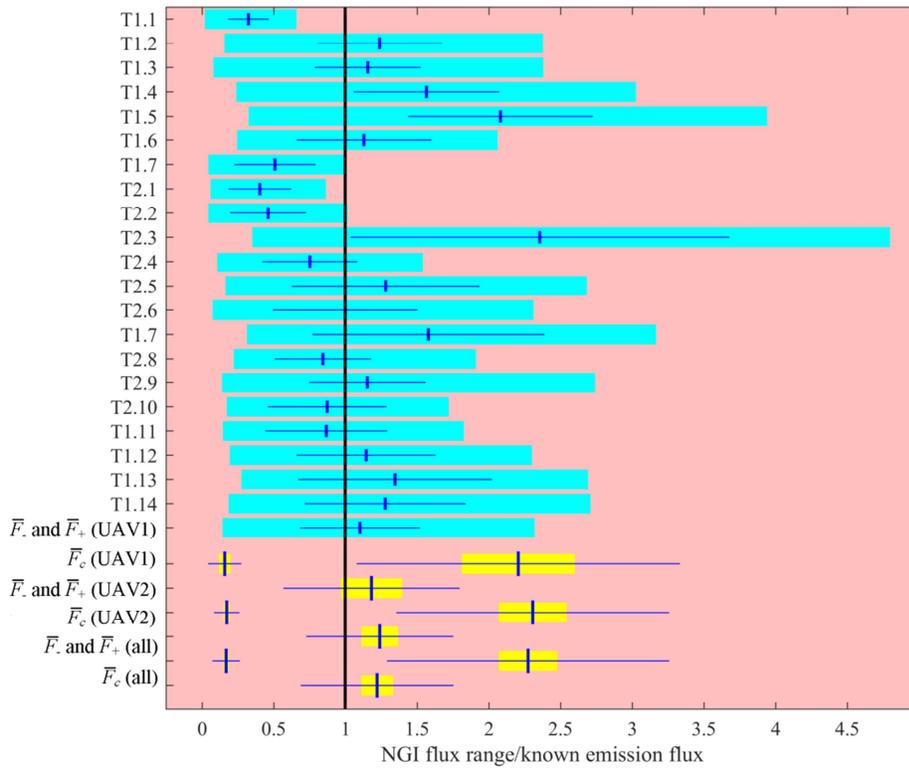


Figure 7. NGI flux uncertainty range (thick cyan bars), for each method testing flight survey, as a fraction of F_0 . The σ_F uncertainty range (horizontal blue lines) is given on either side of F_e (vertical blue lines). \bar{F}_c and \bar{F}_- and \bar{F}_+ averages (vertical blue lines) are plotted for UAV1, UAV2 and for all flight surveys. Standard deviation uncertainty ranges (horizontal blue lines) and standard error uncertainty ranges (thick yellow bars) are given on either side of \bar{F}_c , \bar{F}_- and \bar{F}_+ values.

720

	MGGA	pMGGA
Mass	4.8 kg	3.4 kg
Length	0.35 m	0.33 m
Width	0.30 m	0.20 m
Depth	0.15 m	0.13 m
Power consumption	35 W	32 W
Operating DC voltage	10 V – 30 V	10 V – 28 V
Cell pressure	atmospheric	pressure controlled to 0.61 bar
E-folding time	(1.6±0.2) s	(3.0±0.1) s
<u>Volumetric flow rate</u>	(27.90±0.05) cm³ s⁻¹	(5.08±0.02) cm³ s⁻¹
Maximum sampling frequency	10 Hz	5 Hz
σ_{AVQ}	2.71 ppb	5.44 ppb
1 Hz Allan deviation	0.71 ppb	2.2 ppb
0.1 Hz Allan deviation	0.24 ppb	0.72 ppb
Optimum integration time	(20±3) s	(70±10) s

Table 1: General properties of the MGGA and the pMGGA.

	MGGA	pMGGA
<i>a</i>	-0.00557 -0.000312 mol _{water} mol ⁻¹	-0.000503 +0.000195 mol _{water} mol ⁻¹
<i>b</i>	+0.00530 -0.000193 mol _{water} mol ⁻¹ ppm ⁻¹	+0.000704 -0.0000257 mol _{water} mol ⁻¹ ppm ⁻¹
<i>w</i>	23.6 ppm	23.6 ppm
<i>α</i>	-1.556 mol mol _{water} ⁻¹	-1.640 mol mol _{water} ⁻¹
<i>β</i>	-12.25 mol ² mol _{water} ⁻²	-1.208 mol ² mol _{water} ⁻²
<i>σ_v</i>	0.0004253	0.0002613

725 | **Table 2: Water correction coefficients for the MGGA and pMGGA, required to obtain ν using Eq. (4) and σ_v .**

	MGGA	pMGGA
$G \pm \sigma_G$	0.9970 \pm 0.00023	0.9869 \pm 0.00028
$C \pm \sigma_C$	(+0.0132 \pm 0.0020) ppm	(-0.0019 \pm 0.0015) ppm

Table 3: Calibration coefficients for the MGGA and pMGGA.

	UAV1	UAV2
Flights per survey	2	1
Distance of sampling plane from source	48-47 m – 51-50 m	64 m – 104-114 m
Take-off and landing	Manual	Manual
Flight control	Waypoints	Manual (course lock)
Average velocity across the sampling plane	$(1.5 \pm 0.1) \text{ m s}^{-1}$	$(2.8 \pm 0.6) \text{ m s}^{-1}$
Payload	PFA tubing and inlet, wind sensor	pMGGA
Height of plane of propellers	0.540 m	0.680 m
Height of air inlet	0.845 m	0.370 m

Table 4: A comparison between UAV1 and UAV2.

730

	<u>UAV1</u>	<u>UAV2</u>
<u>Instrument</u>	<u>MGGA</u>	<u>pMGGA</u>
<u>Average cell temperature standard deviation within flight surveys</u>	<u>(±0.16±0.09)° C</u>	<u>(±0.28±0.28)° C</u>
<u>Average cell pressure standard deviation within flight surveys</u>	<u>(±1.15±0.86) mbar</u>	<u>(±0.40±0.01) mbar</u>
<u>Average cell temperature mean across flight surveys</u>	<u>(25±2)° C</u>	<u>(22±4)° C</u>
<u>Average cell pressure mean across flight surveys</u>	<u>(1025.6±5.4) mbar</u>	<u>(614.4±0.1) mbar</u>

Table 5: Average cell temperature and cell pressure standard deviation variability within each UAV flight survey, recorded by the MGGA and the pMGGA. The average cell temperature and cell pressure mean is also given.