Authors Response Document:

This document contains:

1. Replies to the reviewer's comments and critique.

2. List of major changes to manuscript.

3. Line by Line document comparison with previous version.

1. Reply to Reviewers:

We would like to thank the reviewers for their useful and constructive comments regarding the manuscript and have made changes where possible to satisfy their questions and concerns. The original comments from the reviewers are shown in red, with the authors response following in black.

The title has been changed to reflect the request made by Jorg Hacker to describe the aircraft as medium sized.

Anonymous Referee #1

Received and published: 31 July 2020

The Authors describe the development of an airborne measurement platform for the quantification and source attribution of methane from offshore oil and gas operations. The instruments, their airborne deployment and the techniques for data analysis are not really new, but the manuscript would provide a useful reference in future publications that use the data from this platform. I agree that is a worthwhile objective. Overall, the paper is very straightforward and can be published after accounting for the following comments.

Line 56: Suggest "used" instead of "trialled". The latter suggests that previous work should be regarded as somewhat preliminary, but I believe that airborne determination of methane fluxes is quite a mature method by now.

We Agree – this has now been changed in the manuscript.

Line 105: A 5-second delay between air entering the inlet and reaching the instruments seems quite long. What is the volume of the sampling manifold and what is the pump speed?

The 5 second delay was measured for the uGGA by spiking the aircraft inlet tip using breath CO_2, and monitoring the concentration response of the spectrometer. The lagtime was also confirmed by an analysis of the various inlet section showed in Figure 2, their dead volumes and plug-flow flush rates, namely: 36 cm^3 for the 52 cm long ½" od SS rearward facing inlet flushed at ~15000 cm^3/min (inlet pump), and 16 cm^3 for the 110 cm long ¼" od Synflex transfer line flushed at ~ 200 cm^3/min (uGGA internal pump), combining to ~5 sec.

For the Picarro spectrometer, which external pump maintains a sample flowrate of 5000 cm^3/min, the combined lagtimes of its three inlet sections can be estimated at 0.14 + 0.14 + 0.16 ~0.4 sec.

We admit that our quoted 5 sec is an overestimate for the Picarro CRDS. As the lagtime varies for each instrument, this sentence has been removed as it is not especially helpful to the reader.

Lines 125-135: I suggest adding a table with the different instruments, parameters measured, measurement precisions and time responses.

A table with measured precisions, response and lag times has been collated for the Appendix and replaced some of the description in the text.

Lines 153-154: What mixing ratio would be required for in-flight calibrations and what was available?

There was no discernible ethane in the cylinder mixes and therefore not possible to complete in-flight calibrations. This is now clarified in the manuscript text.

Lines 160-165: Is ethane reported as mixing ratios in dry air (as you presumably do for methane)?

The TILDAS outputs the dry mixing ratio as it monitors a water line itself that is used to calculate, and correct the mixing ratio. The TILDAS ethane measurement description has this information added to the manuscript.

Lines 166-174: I re-read papers from two other groups that have used the same TILDAS instrument for airborne measurements of ethane and they seem to have overcome this issue (Smith et al. 2015; Peischl et al. 2018). For example, Smith et al. say that "in-flight drift varied within the instrument precision during a typical research flight" and Peischl et al. gave a "variability of in-flight standard retrieval, ±0.7%". These papers should be cited and discussed in this context. How is the airborne deployment of the TILDAS instrument different between the present study and Smith et al. and Peischl et al.? Note that even in a pressurized aircraft, the cabin pressure can still show considerable variability after take-off.

It is true that there are several published papers containing data from TILDAS instruments. However, the two papers cited here by the reviewer focus on the estimated fluxes rather than a detailed analysis of instrument performance. Therefore, we don't believe it would be appropriate to cite them as evidence that the cabin pressure sensitivity has been overcome in these cases, because such evidence is not provided. The relevance of the in-flight standard variability from Peischl et al. (2018) to this issue is not clear, because the paper does not state over what range of altitudes these standard measurements were taken. The study by Smith et al. (2015) states that the instrument stability was evaluated based on the difference between free troposphere measurements at the beginning and end of a flight. This accounts for temporal drift but not necessarily sensitivity to cabin pressure changes (depending on whether the data comparison was conducted over the same altitude range).

To our knowledge, there are four instrument-focussed papers reporting the performance of Aerodyne TILDAS instruments on aircraft (Santoni et al., 2014; Pitt et al., 2016; Gvakharia et al., 2018; Kostinek et al., 2019). Of these, the three most recent papers all report measurement sensitivity to cabin pressure. Gvakharia et al. (2018) and Kostinek et al. (2019) use frequent measurements of zero air and/or calibration gas at very high flow rates to correct for this issue while maintaining a reasonably high duty cycle. These measurements are performed every 2 minutes by Gvakharia et al. (2018) and every 5-10 minutes by Kostinek et al. (2019). Both

studies demonstrate that at these frequencies even the rapid drift during vertical profiles can be adequately accounted for. On the other hand, Floerchinger et al. (2019) made zero measurements every 10 minutes, and still observed a strong sensitivity to cabin pressure changes (C. Floerchinger, personal communication). This wasn't reported in the manuscript because the vertical profile data was not required for the study. Peischl et al. (2018) and Smith et al. (2015) made zero measurements every 15 minutes but do not provide data on stability during vertical profiles (again, it doesn't appear that this data was used in the mass balance flux calculations).

Santoni et al. (2014) do not explicitly demonstrate the measurement sensitivity (or lack thereof) to cabin pressure changes. One of the two instruments they describe (the QCLS-CO2) is housed within a hermetically sealed pressure vessel, which presumably solves this issue. Developing a similar customised hardware fix would be one option to improve performance on the BAS Twin Otter going forward. Alternatively, if payload constraints allow, it may prove simpler to implement a frequent, fast calibration system of the sort used by Gvakharia et al. (2018) and Kostinek et al. (2019). However, as discussed in the manuscript, as long as each in-plume measurement is referenced against a background measurement at the same altitude the calculated enhancements are not impacted by this cabin pressure sensitivity. Therefore, this issue does not have a serious impact on the results presented in this study. We have added references to Santoni et al. (2014) and Kostinek et al. (2019) to the manuscript.

Section 2.5.1 and 2.5.2: When collecting whole air or bag samples in narrow plumes near sources, the exact timing of the sample delay, and open and close times is important to get the best correlation with the in-situ measurements. How well are these known for the instrumentation described here? Fill times are typically a function of altitude.

The fill times are very rapid (of the order of a few seconds) with this configuration for the bag samples and the Whole Air Sample collection benefitted from the upgrade to the continuous through-flow for the second campaign. As samples are taken at low altitude passes within the boundary layer, the problem of altitude adjusted fill times was not a problem, although this is a noted problem on high altitude flying on aircraft such as FAAM (France, Cain et al. 2016).

Lines 220-223: It is not clear if these limitations pertain to the study by Gorchov Negron or to the present manuscript.

This is clarified in the text.

Lines 226-238: How far from point sources were the downwind transects typically?

Transects varied between 1 and 10 km downwind of the source rig locations. In some cases, for example where multiple potential sources were in the same upwind direction, the maximum distance downwind may have been greater than 10 km – analysis of the attribution of point sources will be dealt with in future work. This has been added to the text.

Figures 4 and 5: Perhaps you can add the flight tracks or general location of the flights to the map.

General non-specific area locations have been added to the figures

Lines 262-264: Part of the reason that methane shows so little structure in Figure 5 is that it is near the global background. If there are no nearby sources, methane will be perfectly constant regardless of how stable or well-mixed the boundary layer is. Do you have a better example where methane is enhanced more and mixed evenly across the boundary layer?

A new example from the available data has been chosen to reflect this more clearly in Figure 5.

Section 4.2: Showing that the slow-response instrument is insufficient to separate plumes and determine plume shape is not particularly new or surprising. Does this instrument provide other strengths to justify being part of the payload? For example, is the slow-response instrument more stable and accurate, and allow for important cross calibration opportunities with the fast-response instruments?

It serves as a relatively lightweight and cost effective redundancy in this configuration. Although with minor modification to the set-up it could have been run in a faster mode, the LGR uGGA would still not have matched the fast Picarro instrument for performance. It was therefore kept as a simple tool for cross-checking and redundancy. The lack of detail in the plume measured by the slower instrument can be clearly seen in Figure 6, but it still serves as a useful cross-checking dataset.

Figure 6: Please provide a clearer legend. I found it difficult to decide what is what from the caption and the axis labels.

Figure 6 has been updated with a new set of legends for the figure.

Equation (1): I found this confusing. Why do you use an average methane enhancement in a plume when you have the time response that allow you to integrate fluxes across a plume (with fluxes in every bin calculated by Eq. 1)? The approach described here relies on a normal distribution of methane in the plumes. Is that true? In addition, this equation yields the flux in units of moles per seconds per meter altitude. This still needs to be integrated across altitude for a meaningful flux number (in moles per second) that can be compared with emissions estimates, but that last step is not included in Equation (1).

The equation employed for mass balance here does use average values for methane evaluated across the horizontal width of the plume. However, as this mean is evaluated over an explicit plume width (Δx), the flux calculated is exactly equal to a piecewise (i.e. smaller dx) integration over discrete horizontal distances. There is no requirement for the plume to be Gaussian, or for a normal distribution of measured concentrations when in the plume. The only assumption being made here is that the aircraft speed and heading remain constant as it passes through the plume – a good assumption in the case of narrow single-facility plumes like the ones presented here.

Also, the reviewer is correct that the equation is missing the vertical height of the plume. The equation has been amended to include an additional "Δz" term, relating to the vertical extent of the plume.

Line 323: "vertical extent of the plume" instead of "vertical resolution"?

We agree, this reads better as "vertical extent of the plume".

Lines 340-341: Have you tried to calculate cross correlations between the CH4 and C2H6 measurements to determine the difference in delay times between the two measurements?

This was attempted, but the results were not suitable due to the variability in the different turnover time (e-fold response) of the instruments and significantly affected the ratios. We expect the integration method to be more accurate.

Section 5.3: Why not simply show a Keeling plot? The discussion of why such plots are challenging for this application is hard to fully understand without having an example with actual data to look at. Instead, the results are presented in Figure 8 in a very indirect way. The point of this analysis appears to be that with fewer data points (when downwind sampling is less extensive) the deterioration of precision for the 13CH4 delta value is not too severe. But what are the delta values you are trying to distinguish? It might be helpful to add those as horizontal lines in Figure 8 and discuss the loss in precision in terms of those delta values. Overall, the discussion left it unclear to me whether or not the 13CH4 measurements gave useful information. This measurement is one of the more novel aspects of this work and it would be good to see the potential of the method demonstrated in more detail.

A Keeling plot is now shown along with the current Figure 8. We agree that this was a bit of an oversight not to include this data plotted in this way.

Figure 8: What are the differently colored symbols? Also, the caption repeats “source signature data” twice.

Caption has been updated to correct this and clarify.

Table A1 seems a little out of place as none of these data are used in the manuscript.

Although we agree, we would like to keep this in as it is a demonstration of measurement capability and fits with the theme of the paper.

Anonymous Referee #2

Received and published: 29 July 2020:

France et al. quantify CH4 emissions from offshore oil platforms using combinations of instruments aboard a Twin Otter aircraft. They describe the lessons learned from two years of flying downwind of these platforms. They also discuss methods of distinguishing sources of CH4 based on isotopic measurements and correlations with ethane. They find ethane:CH4 emission ratios of 0.029 in both years of flying, in line with published estimates. Their estimates of CH4 mass fluxes improved significantly when flying in 2019 in a well-mixed marine boundary layer. This paper provides a straightforward description of the project. As such, there is not much to critique. They lessons the authors learned during the two years were mostly to be expected, i.e., faster response instruments were able distinguish source locations better than slow response instruments; a well-mixed marine boundary layer was easier to measure a downwind plume than a layered, poorly-mixed marine boundary layer; etc. However, since the paper will stand as an overview of the project studying emissions from offshore platforms, and because it provides some guidelines for future projects, it is worthy of publication in this journal. I have mostly minor comments related below.

line 54, "pinpoint" seems redundant. Is there a difference between locating and pinpointing emission sources? Maybe the authors mean locate facilities that are emitting, then pinpoint where in the facility the emissions are? And I think this sentence would read better if it were presented in a hypothetical chronological order: first locate emissions, second quantify them, third validate inventories, fourth design effective mitigation.

Agree – this sentence / paragraph has been restructured in line with the suggestions here.

line 131, stating the precision of the ethane measurement in flight would be more appropriate than in the lab

It would indeed be a better metric to provide, however as the cylinders contained no discernible ethane, an in-flight precision measurement estimate is not possible.

line 315, when the authors say a "vertical run", do they mean a vertically-stacked horizontal run?

We agree, this does read better too. "vertical run" has been changed to "stacked horizontal run" through the paper.

In Figure 7, please state how far downwind the aircraft was for each of these two flights.

The distances that the aircraft were for various flights is clarified in the text. The aircraft varied in distance from emitting sources from ~1 to 10km away downwind.

line 355, I don't think the word "ideally" is necessary. There must be some variability in the source strength compared to the background in order to fit a line through the data points.

We agree, this has been changed.

Other comments:

line 253, what does NAME stand for?

The expansion for NAME (Numerical Atmospheric-dispersion Modelling Environment) has been added to the text.

Grammar suggestions:

All subsequent grammar suggestions have been resolved. Thank you for spotting these.

line 23, add comma after (SLCP)

line 93, it looks like the superscript "-1" is a different font

line 103, it looks like the second end parenthesis of the O'Shea reference is a different size?

line 176, it is unclear what "fit" means here

line 181 and elsewhere, suggest changing "in O'Shea" to "by O'Shea"

line 188, suggest "canister sampling" instead of "canisters sampling"

line 199, need ending parenthesis after Lowry reference

line 245, suggest "by Stull" instead of "in Stull"

line 255, change "decision" to "decisions"

line 319, I found this sentence a little confusing to read. I suggest instead of "between the maximal and minimal altitude transects that do not demonstrate CH4 enhancements so are outside of the plume", perhaps say "between the highest and lowest transects without CH4 enhancements, which are above and below the plume, respectively"

line 328, same strange small parenthesis in the Plant reference

line 346, suggest "by Peischl" instead of "in Peischl"

line 352, suggest “by Keeling” instead of “in Keeling”

line 384, “dramatically” is subjective. I suggest removing this word.

References for Reply to Reviewers

Floerchinger, C., McKain, K., Bonin, T., Peischl, J., Biraud, S. C., Miller, C., Ryerson, T. B., Wofsy, S. C. and Sweeney, C.: Methane emissions from oil and gas production on the North Slope of Alaska, Atmos. Environ., 218, 116985, doi:10.1016/j.atmosenv.2019.116985, 2019.

France, J. L., M. Cain, R. E. Fisher, D. Lowry, G. Allen, S. J. O'Shea, S. Illingworth, J. Pyle, N. Warwick, B. T. Jones, M. W. Gallagher, K. Bower, M. Le Breton, C. Percival, J. Muller, A. Welpott, S. Bauguitte, C. George, G. D. Hayman, A. J. Manning, C. L. Myhre, M. Lanoisellé and E. G. Nisbet (2016). Measurements of δ13C in CH4 and using particle dispersion modeling to characterize sources of Arctic methane within an air mass. Journal of Geophysical Research: Atmospheres 121(23): 14,257-214,270.

Gvakharia, A., Kort, E. A., Smith, M. L. and Conley, S.: Testing and evaluation of a new airborne system for continuous N2O, CO2 , CO, and H2O measurements: the Frequent Calibration High-performance Airborne Observation System (FCHAOS), Atmos. Meas. Tech., 11, 6059–6074, doi:10.5194/amt-11-6059-2018, 2018.

Kostinek, J., Roiger, A., Davis, K. J., Sweeney, C., DiGangi, J. P., Choi, Y., Baier, B., Hase, F., Groß, J., Eckl, M., Klausner, T. and Butz, A.: Adaptation and performance assessment of a quantum and interband cascade laser spectrometer for simultaneous airborne in situ observation of CH4, C2H6, CO2, CO and N2O, Atmos. Meas. Tech., 12, 1767–1783, doi:10.5194/amt-12-1767-2019, 2019.

Peischl, J., Eilerman, S. J., Neuman, J. A., Aikin, K. C., de Gouw, J., Gilman, J. B., Herndon, S. C., Nadkarni, R., Trainer, M., Warneke, C. and Ryerson, T. B.: Quantifying Methane and Ethane Emissions to the Atmosphere From Central and Western U.S. Oil and Natural Gas Production Regions, J. Geophys. Res. Atmos., 123, 7725–7740, doi:10.1029/2018JD028622, 2018.

Pitt, J. R., Le Breton, M., Allen, G., Percival, C. J., Gallagher, M. W., Bauguitte, 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, 63–77, doi:10.5194/amt-9-63-2016, 2016.

Santoni, G. W., Daube, B. C., Kort, E. A., Jiménez, R., Park, S., Pittman, J. V., Gottlieb, E., Xiang, B., Zahniser, M. S., Nelson, D. D., McManus, J. B., Peischl, J., Ryerson, T. B., Holloway, J. S., Andrews, A. E., Sweeney, C., Hall, B., Hintsa, E. J., Moore, F. L., Elkins, J. W., Hurst, D. F., Stephens, B. B., Bent, J. and Wofsy, S. C.: Evaluation of the airborne quantum cascade laser spectrometer (QCLS) measurements of the carbon and greenhouse gas suite – CO2, CH4, N2O, and CO – during the CalNex and HIPPO campaigns, Atmos. Meas. Tech., 7, 1509–1526, doi:10.5194/amt-7-1509-2014, 2014.

Smith, M. L., Kort, E. A., Karion, A., Sweeney, C., Herndon, S. C. and Yacovitch, T. I.: Airborne Ethane Observations in the Barnett Shale: Quantification of Ethane Flux and Attribution of Methane Emissions, Environ. Sci. Technol., 49, 8158–8166, doi:10.1021/acs.est.5b00219, 2015.

2. Major Changes to Manuscript:

There have been no structural changes to the manuscript and the flow of the document remains the same.

Title minor change to reflect aircraft classification as requested by Jorg Hacker

Section 2.3 Details clarified and new table added to appendix to allow easy reference of instruments on board, precision and response rates.

Section 2.4 Details of the TILDAS ethane calibration and known in-flight pressure problems expanded on, with suggestions for mitigation on future deployment. Reference made to extra literature for further reading.

Section 4.1 Extra details around the meteorology added. Figures 4 and 5 updated to show sampling area and synoptic conditions requested by reviewer.

Section 5.1 Equation corrected and terminology in the text clarified around the sampling strategy for the flux calculations.

Figure 6 Updated with new legend to improve ease of viewing.

Figure 8 Updated to include Keeling plot alongside error analysis.

Table A1. Added to display details of instrument performance in aircraft configuration.

3. Line by line comparison document.

Facility level measurement of off-shore oil & gas installations from a ~~small~~medium-sized airborne platform: Method development for quantification and source identification of methane emissions

James France[1,2], Prudence Bateson[3], Pamela Dominutti[4], Grant Allen[3], Stephen Andrews[4], Stephane Bauguitte[5], Max Coleman[2,10], Tom Lachlan-Cope[1], Rebecca E Fisher[2], Langwen Huang[3&,9], Anna E Jones[1], James Lee[6], David Lowry[2], Joseph Pitt[3&,8], Ruth Purvis[6], John Pyle[7], Jacob Shaw[3], Nicola Warwick[7], Alexandra Weiss[1], Shona Wilde[4], Jonathan Witherstone[1], Stuart Young[4].

[1] British Antarctic Survey, Natural Environment Research Council, Cambridge CB3 0ET, UK
[2] Department of Earth Sciences, Royal Holloway, University of London, Egham TW20 0EX, UK
[3] Department of Earth and Environmental Science, University of Manchester, Manchester, M13 9L, UK
[4] Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, Heslington, YO10 5DD, UK
[5] FAAM Airborne Laboratory, National Centre for Atmospheric Science, Cranfield, Bedfordshire, MK43 0AL, UK
[6] National Centre for Atmospheric Science, Innovation Way, University of York, York, UK
[7] National Centre for Atmospheric Science, Department of Chemistry, University of Cambridge, Cambridge, UK
[8] School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11974, USA
[9] Departement Mathematik, ETH Zurich, Rämistrasse 101, 8092 Zürich, Switzerland
[10] Department of Meteorology, University of Reading, Reading, RG6 6BB, UK.

Correspondence to: James L France (james.france@rhul.ac.uk)

Abstract.

Emissions of methane (CH_4) from offshore oil and gas installations are poorly ground-truthed and quantification relies heavily on the use of emission factors and activity data. As part of the United Nations Climate and Clean Air Coalition (UN CCAC) objective to study and reduce short-lived climate pollutants (SLCP~~)~~), a Twin Otter aircraft was used to survey CH_4 emissions from UK and Dutch offshore oil and gas installations. The aims of the surveys were to i) identify installations that are significant CH_4 emitters, ii) separate installation emissions from other emissions using carbon-isotopic fingerprinting and other chemical proxies, iii) estimate CH_4 emission rates~~,~~ and iv) improve flux estimation (and sampling) methodologies for rapid quantification of major gas leaks.

In this paper, we detail the instrument and aircraft set up for two campaigns flown in the springs of 2018 and 2019 over the southern North Sea and describe the developments made in both planning and sampling methodology ~~in order~~ to maximise the quality and value of the data collected. We present example data collected from both campaigns to demonstrate the challenges encountered during offshore surveys, focussing on the complex meteorology of the marine boundary layer, and sampling discrete plumes from an airborne platform. The uncertainties of CH_4 flux calculations from measurements under varying boundary layer conditions are considered, as well as recommendations for attribution of sources through either spot sampling

for VOCs / $\delta^{13}C_{CH4}$ or using in-situ instrumental data to determine C_2H_6-CH_4 ratios. A series of recommendations for both planning and measurement techniques for future offshore work within ~~the~~ marine boundary layers are provided.

1. Overview

Methane is a potent greenhouse gas in the atmosphere, with a global warming potential 84 times that of carbon dioxide when calculated over a 20-year period (Myhre et al., 2013). Increases in atmospheric CH_4 mixing ratios are expected to have major influences on the Earth's climate, and emission mitigation could go some way toward achieving goals laid out in the UNFCCC Paris Agreement (Nisbet et al., 2019).

Offshore oil and gas fields make up ~28% of ~~the~~ total global oil and gas production ~~worldwide~~ and are expected to be significant sources of CH_4 to the atmosphere, given that 22% of global CH_4 emissions are estimated to be from the oil and gas (O&G) sector (Saunois et al., 2016). Some emissions arise from routine operations or minor engineering failures (Zavala-Araiza et al., 2017), while others stem from large unexpected leaks, (e.g. (Conley et al., 2016; Ryerson et al., 2012)). In some O&G fields, large amounts of non-recoverable CH_4 can be flared or vented due to a number of factors. Thus, the composition of O&G emissions can be influenced by several variables, including the targeted hydrocarbon product (oil or gas), extraction techniques and gas capture infrastructure. O&G installations co-emit volatile organic compounds (VOCs) such as alkanes, alkenes and aromatics in addition to CH_4. Some of these VOCs are toxic and can have direct health impacts or, together with NO_x, can produce ozone, having an impact on the regional air quality (Edwards et al., 2013). VOC and $\delta^{13}C_{CH4}$ measurements can be utilised to fingerprint the main processes or likely location responsible for associated CH_4 emissions (Cardoso-Saldaña et al., 2019; Lee et al., 2018; Yacovitch et al., 2014a). A recent study has also demonstrated the cost-effectiveness of airborne measurements for leak detection and repair at O&G facilities relative to traditional ground-based methods (Schwietzke et al., 2019).

There is thus a need to develop reliable methodologies to locate emissions, ~~pinpoint~~determine sources ~~and design effective mitigation action~~ in sufficient detail to allow quantification of emissions and validate against publicly reported inventory emissions~~.~~ to enable the design of suitable mitigation. To date, a number of approaches have been ~~trialled~~used. Airborne measurements of both individual and clusters of facilities, along with production data, have been used to scale up to an inventory of CH_4 emissions for the U.S. Gulf of Mexico (Gorchov Negron et al., 2020). Ship based measurements of CH_4 and associated source tracers have been made in both the Gulf of Mexico (Yacovitch et al., 2020) and in the North Sea (Riddick et al., 2019). The latter reported fluxes of CH_4 from offshore O&G installations in UK waters that were derived from observations made from small boats at ~2 m above sea level. This approach has advantages in terms of cost, but the authors recognised a number of key uncertainties in their approach associated with assumptions around boundary layer conditions and a lack of 3D

information (i.e. Gaussian plume modelling and assumptions of constant wind speed). Measurements from aircraft can provide this 3D spatial information enabling better characterisation of both plume morphology and boundary layer dynamics.

Here we report a project that was designed around the use of a small-aircraft with flexible instrument payload suitable for agile deployment. Key objectives were i) to identify and quantify emissions of CH_4 from a suite of offshore gas fields within a limited geographical area, and ii) to develop methodologies that can be applied to gas fields elsewhere to assess emissions at local scales. The project was part of the United Nations Climate and Clean Air Coalition (UN CCAC) objective to ~~characterize~~characterise global CH_4 emissions from oil and gas infrastructure. Targeted observations of atmospheric CH_4 and C_2H_6, plus sampling for VOC and $\delta^{13}C_{CH4}$ analysis were made from a Twin Otter aircraft operated by the British Antarctic Survey. Two campaigns were conducted, one in April 2018 and one in April/May 2019, with a total of 10 flights (~45 hours) over the two campaigns.

The specific aims of the surveys were:

1. CH_4 surveying of facilities with a range of expected (from inventories) CH_4 emissions.
2. Resolution of types of emission from installations (such as flaring, venting, combustion and leaks) using carbon-isotopic fingerprinting and analysis of co-emitted species (including VOCs).
3. Estimation of total CH_4 emissions for the target region.
4. Improvement of flux estimation (and sampling) methodologies for rapid quantification of major gas emissions.

Here, we provide an overview of measurement platform configuration and sampling strategy during these campaigns, including instrument comparisons for hydrocarbon plume detection, spot sampling strategies for VOCs and $\delta^{13}C_{CH4}$, and flight planning to cope with complex boundary layer meteorology to allow estimation of emission fluxes. Analysis methods to determine diagnostic hydrocarbon plume characteristics such as C_2H_6-CH_4 ratios and ~~$\delta^{13}C$~~$\delta^{13}C_{CH4}$ source attribution are also discussed. A sister publication will present the estimated facility-level emissions in detail and discuss the results in a regional context.

2. Experimental

A DHC6 Twin Otter research aircraft, operated by the British Antarctic Survey, was equipped with instrumentation to measure atmospheric boundary layer parameters, including the boundary layer structure and stability, as well as a number of targeted chemical parameters. These included CH_4, CO_2, H_2O, C_2H_6 as well as whole air sampling for subsequent analysis of $\delta^{13}C_{CH4}$ and a suite of VOC's. Here we describe the aircraft capability, aircraft fit, and the instruments deployed.

2.1. Aircraft capability

The maximum range of the Twin Otter aircraft during the flight campaigns was approximately 1000 km. Although the aircraft is capable of flying up to 5000 m altitude, most of the flying was limited to below 2000 m; in regions with no minimum altitude limit, the aircraft could be flown at the practical limit of 15 m ~~(~50 ft)~~ above sea level. The instrument fit included use of a

turbulence boom, which limited the speed to a maximum of 140 kts (~70 ms^{-1}); throughout the campaigns, the target aircraft speed for surveying was 60 ms^{-1}. The aircraft was limited to a minimum safe separation distance of 200 m from any O&G production platforms.

The total weight of the aircraft on take-off is limited to 14,000 lbs (6,350 kg). Allowing for fuel and crew, this left 2,086 kg for the instrumentation. The total power available on the aircraft is 150 A at 28 V and inverters were used to provide 220 V to those instruments that required it.

Altitude and air speed were determined by static and dynamic pressure from the aircraft static ports and heated Pitot tube, logged using Honeywell HPA sensors at 5 Hz. A radar altimeter recorded the flight height at around 10 Hz. An OXTS Inertial measurement system coupled to a Trimble R7 GPS was used to determine the aircraft position and altitude. This system gives all three components of aircraft position, altitude and velocity at a rate of 50 Hz.

The chemistry inlets on the Twin Otter are similar to those fitted to the FAAM BAE 146 large atmospheric research aircraft (e.g. O'Shea et al., (2013)) and were fitted with the inlet facing to the rear (Fig. A1). A single line (¼" Synflex tubing) was taken from the inlet to a high capacity pump with the instruments branching from this line. ~~This approach was taken to minimise the delay between air entering the inlet and reaching the instruments, which, with this configuration, was ~5 seconds.~~ The aircraft was fitted out during the week before each of the two flight campaigns allowing significant changes to be made between 2018 and 2019 based on instrument performance and data from 2018 (Fig. 1).

2.2. Boundary layer physics instrumentation

A fast response temperature sensor and a nine hole NOAA BAT 'Best Air Turbulence' probe (Garman et al., 2006) ~~was~~were mounted on a boom on the front of the aircraft (see photo, Fig. A2). This instrumental set-up was chosen to reduce flow distortion effects by the aircraft. These fast response measurements of wind and temperature fluctuations were made with a frequency of 50 Hz. Garman et al. (2006) investigated the uncertainty of the wind measurements by testing a BAT probe in a wind tunnel. They assessed that the precision of the vertical wind measurements due to instrument noise was approximately ±0.03 ms^{-1}. Garman (2008) showed that an additional uncertainty in the wind data occurs when a constant up-wash correction value is used, as proposed by the model of Crawford (1996). We use the Crawford model which increases the uncertainty in the vertical wind component, w, to approximately ±0.05 ms^{-1}. We assume for the two horizontal wind components, u and v, similar high uncertainties due to aircraft movement. A detailed description of the Twin Otter turbulence instrumentation and associated data processing can be found in Weiss et al. (2011).

Ambient air temperature was observed with Goodrich Rosemount Probes, mounted on the nose of the aircraft. A non-de-iced model 102E4AL and a de-iced model 102AU1AG logged the temperature at 0.7 Hz. Atmospheric humidity was measured with a Buck 1011C cooled mirror hygrometer. The 1011C Aircraft Hygrometer is a chilled mirror optical dew point system.

The manufacturer stated ~~the~~a reading accuracy of ± 0.1 ºC in a temperature range of -40 to +50 ºC. Chamber pressure and mirror temperature were recorded at 1 Hz.

2.3. *In situ* atmospheric chemistry instrumentation

A Los Gatos Research (LGR) Ultraportable Greenhouse Gas Analyser (uGGA) was installed to measure CH_4, CO_2 and H_2O. Expected manufacturer precision for the CH_4 measurement was < 2 ppb averaged over 5 s and < 0.6 ppb over 100 s. The response time of the LGR uGGA itself (i.e. the flush time through the measurement cell) was over 10 s. To achieve higher temporal frequency data, a fast Picarro ~~G2301~~G2311-f was installed ~~that provided~~to provide measurements of CH_4, CO_2 and H_2O at ~10 Hz, with 1-σ precision ~~at 10 Hz~~ of ~~< 3 ppb~~~1ppb over 1 s for CH_4. A third greenhouse gas analyser, a LGR Ultraportable CH_4/C_2H_6 Analyser (uMEA) was used to measure CH_4 and C_2H_6. ~~The 1-σ precision of the uMEA, as stated by the manufacturer, was < 2 ppb for CH_4 and < 30 ppb for C_2H_6 at 1 s, however in~~In-house laboratory measurements suggest C_2H_6 1-σ precision at 1 s is ~17 ppb for ~~this unit~~the LGR uMEA. During the 2019 airborne campaign, atmospheric C_2H_6 was also monitored by a Tuneable Infrared Laser Direct Absorption Spectrometer (TILDAS, Aerodyne Research~~,~~ Inc~~).~~) (Yacovitch et al., 2014b)~~.~~ with expected precision of 50 ppt for C_2H_6 over 10 seconds. This instrument utilises a continuous wave laser operating in the mid-infrared region (at λ = 3.3 μm). Further description of the TILDAS instrument set-up and performance is available in the Appendices along with instrument precisions and response times in Table A1.

2.4. Calibration of *in situ* instrumentation

2.4.1. CH_4 and CO_2 calibration

In-situ CH_4 and CO_2 instruments were calibrated in-flight using a manually operated calibration deck, shown in schematic form in Fig. 2. The calibration gases consisted of a suite of WMO referenced standards with a "High", "Low" and "Target" designation. The "High" CH_4 concentration was ~2600 ppb, ~~the~~ "Low" ~1850 ppb and ~~the~~ "Target" ~2000 ppb ~~concentration~~. CO_2 concentrations were "High" ~468.5 ppm, the "Low" ~413.9 ppm and the "Target" ~423.6 ppm. The absolute values of the cylinders varied between years as they were re-filled and re-certified to the NOAA WMO-CH_4-X2004A and WMO-CO_2-X2007 scales. The calibration deck is designed so that upon the calibration valve opening, the calibration gas flow rate is sufficient to overflow the inlet. A similar approach to in-flight calibration is also applied on the NOAA WP-3D aircraft (Warneke et al., 2016). Full details of the calibration procedure ~~is~~are recorded in the Appendices. CH_4 uncertainty (1σ) ~~at 1 Hz~~ is calculated from the in-flight target gas measurements as 1.24 ppb for the Picarro ~~G2301~~G2311-f and 1.77 ppb for the uGGA, giving performance comparable with similar instrumentation on the FAAM aircraft (O'Shea et al., 2014). The excellent agreement between measured and expected values of CH_4 for the target cylinder (for the Picarro and uGGA) gives us confidence in being able to operate to high levels of accuracy with a very limited period of instrument fitting and testing. CO_2 uncertainty (1σ) at 1 Hz is calculated as 0.20 ppm for the Picarro ~~G2301~~G2311-f and 0.35 ppm for the uGGA. More details on the calibration and associated uncertainties are shown in the Appendices.

2.4.2. C_2H_6 calibration

The calibration cylinders installed on the Twin Otter during both campaigns did not contain ~~a sufficient range~~measurable amounts of C_2H_6 ~~concentrations to enable~~and therefore in-flight calibrations ~~to~~could not be performed. This represents a limitation on the accuracy and traceability of the C_2H_6 measurements during these campaigns and will be addressed for future studies using the BAS Twin Otter. The uMEA was calibrated in the laboratory post ~~-~~campaign for the 2018 campaign, and pre-and post-campaign in the laboratory for the 2019 season. The uMEA instrument cavity is not temperature stabilised, resulting in significant measurement drift during the course of operation. Corrections for C_2H_6 and CH_4 measurement drift as a function of cavity temperature were determined experimentally by analysing two calibration cylinders alternately over the course of several hours as the cavity temperature increased. These corrections were then applied to the uMEA C_2H_6 and CH_4 measurements obtained from both 2018 and 2019 flight campaigns.

The TILDAS (deployed in 2019) ~~was~~measures a water line to allow measurements to be corrected ~~for water vapour~~to dry mole using the TDLWintel software (Nelson et al., 2004) to account for changes in humidity during the flight (as discussed in Pitt et al., (2016)). ~~The instrument has a quoted precision of 50 ppt for an averaging time of 10 s.~~ The raw measured data were calibrated pre- and post-flight using two cylinders of known concentration, whose mole fractions spanned the measurement range observed during flights for C_2H_6. By assuming a linear relationship, the calibrated mole fraction corresponding to each measured TILDAS mole fraction was given by interpolating the scale between the pre- and post-flight calibration reference points.

~~Previous studies have reported the sensitivity of TILDAS systems to aircraft cabin pressure (e.g. Pitt et al., (2016) and Gvakharia et al., (2018)).~~Previous studies have reported the sensitivity of TILDAS systems to aircraft cabin pressure ((Gvakharia et al., 2018; Kostinek et al., 2019; Pitt et al., 2016)). This sensitivity means that the C_2H_6 mole fractions measured during the flight contain a systematic altitude-dependent bias. However, as cabin pressure only affects the spectroscopic baseline, the zero-offset of the measurements is affected, but not the instrument gain factor. Therefore, as long as each plume measurement is referenced to a measured background at the same altitude, this cabin pressure sensitivity does not significantly impact the calculated C_2H_6 mole fraction enhancements. As stated above, future deployments will mitigate this issue by employing in-flight calibration cylinders that are certified for C_2H_6. The potential to use a fast, frequent calibration for baseline correction as described by Gvakharia et al., (2018) and Kostinek et al., (2019) will also be investigated, although this has payload implications as it requires an extra calibration cylinder. Alternatively, the optical bench could be re-engineered to sit within a hermetically sealed pressure vessel, as described by Santoni et al.,(2014).

2.5. Spot Sampling

~~On a rapid aircraft fit, manually~~Manually triggered spot sampling provides a cost-effective and relatively simple sample collection method ~~for analysis~~to allow analyses which cannot be performed mid-flight or require specialist laboratory facilities

to gain useful levels of precision. Two discrete air-sampling systems were used during these flights to enable post-flight analysis for VOCs and $\delta^{13}C_{CH4}$.

2.5.1. Son of Whole Air Sampler (SWAS)

The Son of Whole Air Sampler (SWAS) is a new, updated version of the parent WAS system fitted to the FAAM BAE 146 large atmospheric research aircraft (e.g. as used ~~in~~by O'Shea et al., (2014)), which it is designed to supersede. The system comprises a multitude of inert Silonite-coated (Entech) stainless steel canisters, grouped together modularly in cases with up to 16 canisters per case. On-board the Twin Otter, 2 cases can be fitted allowing up to 32 canisters to be carried per flight. The theory of operation is to capture discrete air samples from outside of the aircraft and compress the sample either into 1.4 L or 2 L canisters at low pressure (40 psi) via pneumatically-actuated bellows valves (PBV, Swagelok BNVS4-C). Full details of the operation of SWAS are included in the Appendices. For the 2019 campaign, SWAS was updated with the addition of 2 L flow-through canisters making narrow plumes easier to capture due to reduced sample line lag and fill times.

SWAS ~~canisters~~canister sampling was manually triggered during the flights according to in-situ observations made by fast response instrumentation ~~such as~~of CO_2, C_2H_6 and CH_4, with the aim of capturing specific oil and gas plumes. The samples were analysed at the University of York for VOCs post-flight using a dual-channel gas chromatograph with flame ionisation detectors (Hopkins et al., 2003). Firstly, 500 mL aliquots of air are withdrawn from the sample canister and dried using a condensation finger held at -30 °C then pre-concentrated onto a multi-bed carbon adsorbent trap consisting of Carboxen 1000 and Carbotrap B (Supelco), and transferred to the GC columns (Al_2O_3, $NaSO_4$ deactivated and open tubular PLOT) in a stream of helium. Chromatogram peak identification was made by reference to a calibration gas standard (NPL30, D600145 - 2018) containing known amounts of 30 VOCs ranging from C2-C9. Compounds of interest include C_2H_6, propane, butanes, pentanes, benzene and toluene; a full list is shown in Table ~~A1~~A2.

2.5.2. Flexfoil Bag Sampling

Spot sampling for $\delta^{13}C_{CH4}$ by collecting whole air samples into Flexfoil bags (SKC Ltd) has been in use on both the FAAM BAE 146 research aircraft (e.g. (Fisher et al., 2017)) and during ground based mobile studies (e.g. (Lowry et al., 2020)) and provides a relatively cost-effective and rapid methodology for sample collection. The method does have some limitations, however, as the Flexfoil sample bags are only stable for a number of compounds (including CH_4) ~~and is not a true whole air sample.~~). Samples captured in both Flexfoil bags and SWAS were measured at Royal Holloway using ~~CF-IRMS (~~continuous flow isotope ratio mass spectrometry (CF-IRMS) (Fisher et al., 2006) and each measurement has a $\delta^{13}C_{CH4}$ uncertainty of ~0.05 ‰. Each sample is also measured for CH_4 mole fraction using cavity ring-down spectroscopy to allow direct comparison to in-flight data (Fig A3). Alternative, continuous in-flight $\delta^{13}C_{CH4}$ instrumentation currently cannot replicate the precision of laboratory sampling, and the few seconds of enhanced CH_4 that would be encountered during flight is not sufficient for averaging of continuous $\delta^{13}C_{CH4}$ data to gain a meaningful source $\delta^{13}C_{CH4}$ signature (e.g. (Rella et al., 2015)).

3 Overall ~~approach~~Approach to ~~flight planning~~Flight Planning

The majority of flights were conducted during good operating conditions i.e. ~~–~~daytime, no precipitation, clear or broken cloud, winds < ~~10ms~~10 ms^{-1} and visibility to allow flying at minimum safe altitude around the task area. Two approaches were trialled to assess CH_4 emissions from offshore gas installations: (~~1~~i) regional survey, and (~~2~~ii) specific plume sampling. The flight modes are demonstrated in Fig. 3, with the dark grey pattern showing a flight plan for regional measurements and the orange and white patterns demonstrating specific plume sampling flight patterns. Flight plans to sample specific installations were designed to capture a full range of expected emissions using the UK National Atmospheric Emissions Inventory (NAEI) as a guide.

Regional survey intentions were two-fold: firstly, to offer an identification process for emitters of interest that could specifically be targeted for plume sampling modes, and secondly, to build a picture of aggregate bulk emissions for multiple upwind platforms. This method has been successfully employed during a Gulf of Mexico airborne study (Gorchov Negron et al., 2020). However, in ~~this instance~~the work presented here, regional surveys were poor for identifying plumes (being too far downwind of platforms or not intercepting thin filament layers containing CH_4 enhancements) and attempts to aggregate bulk emissions were hindered by the often encountered complex boundary layer structure over the area, which controlled dispersion of CH_4 emissions from rigs. From the regional flight data derived in 2018, and considering the work in other offshore studies in this area (e.g. Cain et al., (2017)), the regional flight mode was determined to be of limited scientific value in the context of this project and this flight pattern was not used during the 2019 campaign.

Plume sampling flights were conducted in both 2018 and 2019. These flights involved the use of a box pattern to create both upwind and downwind transects either side of the infrastructure of interest. Upwind transects provided an understanding of other methanogenic sources (such as other installations, ships or long range transport of air masses from on-shore sources) that could interfere with observed CH_4 plumes downwind, and were conducted to be confident that plumes were solely originating from the targeted infrastructure. Vertically stacked downwind transects at a distance of 1 to 10 km away from emission sources were conducted to better capture the vertical extent of the plume in a ~~2-dimensional~~2D Lagrangian plane for CH_4 flux quantification using mass balance analysis (e.g. O'Shea et al., (2014)). The vertically stacked transects in profile, as planned from the 2019 field deployment, are demonstrated in Fig. 3. The separation between vertically stacked transects was usually ~~200 ft (~~60 m~~)~~ with a minimum absolute height of ~~150 ft (~~45 m~~)~~ above sea surface up to approximately ~~850 ft (~~260 m~~)~~ to capture the entire extent of a downwind plume. Plume dispersion was dependent on meteorology and emission type (venting, fugitive, or combustive emissions) and as such, maximal plume heights varied between individual infrastructure. Upwind transects were flown at a median height between the minimum and maximum stacked runs.

4. Assessing and addressing issues encountered during flights

A number of issues were encountered during the flights that influenced the measurements made. An initial presentation of these issues is given here, with recommendations for improvements given in Section 6 below.

4.1. Complex marine boundary layers

Boundary layer structure proved to be a ~~decisive~~important influence on observed CH_4 mixing ratios. Figure 4 shows the measured profiles of CH_4 (left hand panel) and potential temperature (right hand panel) during an off-shore flight in April 2018 along with the corresponding synoptic chart. Potential temperature was calculated as described ~~in~~by Stull (1988). The potential temperature profile demonstrates that the boundary layer structure on this day (and many other days) was partly stable stratified, showing mostly an increase in potential temperature with height, and the boundary layer showed complex layering. The prevailing meteorological situation at that time, illustrated by the synoptic chart in ~~Figure~~Fig. 4, was of a persistent anticyclonic ridge, stretching from the south-west over the British Isles and Western Europe, with associated low wind speeds and poorly defined air flow over the southern North Sea sector. The observed layering was partly also caused by residual boundary layers from previous days and nights which had not dispersed. The structure of the boundary layer in Fig. 4 clearly had an important influence on the vertical profile of CH_4, which varied and shows a complex profile with height. Due to the complexity of the boundary layer structure, it was concluded that it would be inappropriate to use a particle dispersion model such as ~~NAME~~the Numerical Atmospheric-dispersion Modelling Environment (NAME) (Jones et al., 2007) to derive a bulk regional emission estimate.

The impact of the residual layers of CH_4 enhancement make in-flight ~~decision~~decisions very challenging for two main reasons: ~~1.~~(i) The difficulty of determining which enhancements are from installations and require further investigation, especially if flying at some distance downwind from a potential source or on a regional survey pattern. ~~2.~~(ii) Emissions being actively released can become trapped in vertically thin filaments, which can be easily missed when flying stacked legs, depending on flight altitude.

In contrast, on days with a well-mixed boundary layer the CH_4 profile stays relatively constant with height, and shows increase only ~~in the surface layer~~ near a CH_4 ~~sources. Fig.~~source. Figure 5 shows an example of ~~composite~~ CH_4 and potential temperature profiles~~ flown within an hour of each other~~, in a well-mixed boundary layer during a flight in May 2019; the synoptic situation on that day was consistent with a slow-moving cyclonic south-easterly air flow. It can clearly be seen how the potential temperature and CH_4 profiles stay almost constant with height ~~above 200 m,~~ and only show structure ~~near~~when intercepting a CH_4 ~~sources~~emission at 300 to 350 m altitude. The potential temperature profile indicates neutral stratification of the boundary layer.

4.2. Instrument response times

The role of the continuous in-flight measurements is to provide the backbone of the dataset and ensure that, at a bare minimum, the flights are able to identify areas of CH_4 enhancement and inform on the likely sources of the CH_4 enhancement, hence the decision to run redundancy measurements of CH_4 utilising an LGR uGGA. Figure 6 shows ~~a~~ typical ~~suite of~~ instrument ~~response~~responses to a CH_4~~-~~ plume and it is clear that the cell turnover time of the uGGA is not sufficient to capture the fine detail of the plume. Whilst the uGGA and uMEA are capable of determining whole infrastructure mass balance and average infrastructure ethane-methane ratios, the refined understanding of the true plume is lost in these slower response instruments. This is important as the combined Picarro ~~G2301~~G2311-f and TILDAS data can detect several sources from the same installation (Fig. 6), because of their rapid measurement cell turnover. This information can be used to infer either cold venting (CH_4 & C_2H_6) or combustion from flares or generators (CO_2, CH_4 and C_2H_6) which could then be used to determine CH_4 emission factors from identified flares (Gvakharia et al., 2017).

There are a number of other implications that arise from slow measurement response. For example, in-flight spot sampling requires guidance from fast response instruments that can indicate the optimum timing to collect samples that span the plume, and thereby capture the representative chemical nature of the plume. Further, in-flight calibrations must be matched to the slowest response instrument to ensure stabilisation of the measurement of calibration gases across all instruments. ~~Use~~Although useful from a cross-checking purpose, use of slower-response instruments ~~thus induces~~can introduce additional, unwanted loss of measurement time and excessive use of calibration gases and the benefits of instrument redundancy should be carefully considered.

4.3. Spot sampling improvements between 2018 and 2019 campaigns

In-flight spot sample collection was carried out during both the 2018 and 2019 campaigns. Such sampling is challenging, and requires fast response instruments to be viewable to the operator to give the best chance of collecting samples at appropriate ~~times that span~~points across the ~~plume~~plumes. For 2019, a number of simple adaptations were introduced that significantly increased the success of capturing plumes (Fig. A3). The improvements included modified flight planning, with an increased number of passes through discovered plumes. This approach resulted in increased fuel consumption per plume, but contributed to the higher success rate of plume capture. The comprehensive update to the SWAS system, which included continuous sample through-flow allowed more precise spot sampling to be achieved.

5. Creation of data products

5.1. Methane fluxes

A methane flux can be calculated from the CH_4 mixing ratio data using mass balance techniques (e.g. (O'Shea et al., 2014; Pitt et al., 2019) in which a vertical 2D plane is defined at a fixed distance downwind of the infrastructure of interest, and sampling is conducted across the stacked transects at this distance if a plume is identified in the downwind plane. Fluxes were derived using Eq. (1):

$$Flux = \left(X_{plume} - X_{background}\right) \times n_{air} \times V \times \Delta x \text{————} \times \Delta z \quad (1)$$

where *Flux* is the bulk net flux passing through the *x-z* plane per ~~metre altitude (mol m^{-1} s^{-1}),~~unit time, n_{air} is the molar density of air (mol m^{-3}), X_{plume} is the average CH_4 mole fraction measured within the plume, and $X_{background}$ is the CH_4 mole fraction of the background. *V* is the wind component perpendicular to the flight track ~~and,~~ Δx is the plume width perpendicular to upwind-downwind~~.~~ and Δz relates to the vertical extent of the plume.

The CH_4 and CO_2 measurements from the 10 Hz response instruments were used to provide the highest accuracy in (~~1~~i) lateral plume width and (~~2~~ii) number of unique plumes identified from each individual platform. Slower response instruments would allow for flux calculations but would not be able to identify individual plumes from the same platform. This could be useful to distinguish, for example, multiple plumes from different emission processes that are spatially distinct within the same platform (e.g., a fugitive source vs. a flare). A background mixing ratio was selected to best represent the conditions observed during the flight at the specific time of survey. An average of 30 s of data either side of the plume on each run ~~was~~were used, if this was deemed appropriate with a clean upwind sampling leg. When the upwind sampling was contaminated, more caution should be taken when selecting an appropriate background so that the background value is not distorted by extraneous far-field sources.

For ~~this~~the flux analysis, a flux across each individual ~~vertical~~stacked horizontal run downwind of a plume was calculated before scaling in the vertical component. The flux was then integrated across potential minimum and maximum plume depths. Figure 7 (upper panel) represents a reduced vertical resolution of the plume where transects at intermediate altitudes through the plume were not conducted. In this case, the minimal plume depth is the narrow span captured by observation in the 45.9 – 51.9 m altitude window. The maximal plume depth is taken as the height difference between the ~~maximal~~highest and ~~minimal altitude~~lowest transects ~~that do not demonstrate~~without CH_4 enhancements ~~so~~, which are ~~outside of~~above and below the plume, respectively; this value has to be used as the maximum due to incomplete sampling of the void area seen in the upper panel of Fig. 7. In cases where the base and top of the plume were not sampled (e.g. during 2018 sampling), the lower limit was selected as the sea surface and the upper limit of the plume was selected as the atmospheric marine boundary layer. The greatest

uncertainty in bulk flux arises when the vertical ~~resolution~~ extent of the plume is not fully captured. ~~Between~~For the ~~2018 and~~ 2019 ~~portions of the~~ campaign, the flux uncertainty related to plume depth was reduced by a factor of 10 compared to the 2018 campaign (as seen in Table 1) by completing a rigorous set of stacked transects at multiple heights throughout the plume. The fluxes presented here serve to demonstrate the approach and the impact of sampling strategy and meteorological conditions on the calculation. Flux estimates for all sampled platforms will be presented in a future study, including a full treatment of component uncertainties.

5.2. Ethane-Methane ratios (C2:C1) as a source tracer

It has already been well established that continuous C_2H_6 measurements can be an excellent diagnostic tool for ascribing enhancements of co-located CH_4 and C_2H_6 to natural gas emissions in both urban areas (e.g. (Plant et al., 2019)), semi-rural areas (e.g. (Lowry et al., 2020)) and during large scale evaluations of oil and gas fields from aerial studies in the USA (e.g. (Peischl et al., 2018)), Canada (Johnson et al., 2017), and the Netherlands (Yacovitch et al., 2018). During this work, two methods were used to establish C_2H_6-CH_4 ratios (hereafter, described as C2:C1). In 2018 ~~a Los Gatos ultraportable CH_4/C_2H_6 analyser (~~the LGR uMEA~~)~~ was used to measure C_2H_6-CH_4 ratios. The benefits of such instrumentation are in its simplicity of operation and that few considerations are required for corrections or variable lags as both species are measured at the same rate and within the same optical cavity. C2:C1 can therefore be readily determined as the gradient of a linear regression between the C_2H_6 and CH_4 measurements. However, the low sensitivity to C_2H_6 (standard deviation of ~~~~>10 ppb in C_2H_6 over 10 s of background flying) only allowed emissions from two platforms to be characterised for C2:C1 ratios during the whole of the 2018 ~~flying~~campaign, and none during 2019 using the LGR uMEA method.

In 2019 the addition of the TILDAS 1 Hz C_2H_6 instrument allowed for better precision of C_2H_6 (< 1 ppb) with a faster flush time in the measurement cell. The C_2H_6 data is time matched with the 1 Hz Picarro CH_4 data set to allow C2:C1 derivation. As the instruments do not have the exact same flow rate and different cell residence times, the C2:C1 ratios were determined using the integral of each CH_4 and C_2H_6 enhancement using Gaussian peak fitting. A comparison between the 2018 flight, 2019 flight and published data derived from the same geographical area, is shown in Table 2. Although both instruments have been operated for this work without in-flight calibration or engineering solutions to address cabin pressure sensitivity issues (Gvakharia et al., 2018) due to weight and time constraints, the agreement between years and with published expected values is highly reassuring. The added value in high precision C2:C1 demonstrates that C_2H_6 is not just a tracer for matching emissions to natural gas; it can give information as to proportions of emissions from mixed sources (as previously used ~~in~~by Peischl et al., (2018)) or can be used to identify a likely emission point in a process chain depending upon enrichment or depletion of C_2H_6 relative to CH_4. The inclusion of a continuous instrument with a sub-ppb level of detection for C_2H_6 is considered vital for future work with thermogenic sources of CH_4 to allow more precise source attribution of emissions where no spot sampling has occurred.

5.3. $\delta^{13}C_{CH4}$ for CH_4 source attribution

The principal method of $\delta^{13}C_{CH4}$ source characterisation utilises the principles outlined ~~in~~by Keeling (1961) and Pataki et al., (2003), and has been well utilised since to create $\delta^{13}C_{CH4}$ databases for a plethora of known CH_4 sources (e.g. (Sherwood et al., 2017)). In order for a Keeling plot to give useful results to determine a $\delta^{13}C_{CH4}$ source signature of a CH_4 emission, the emission must have been successfully captured multiple times and ~~ideally~~ with a range of CH_4 mixing ratios (which could be achieved by passes at different distances or heights downwind of a point source). This sampling process takes time (especially on an aircraft), where the emission plume is only intercepted once per transect and time in the plume is limited so that only one spot sample can be taken whilst "in-plume". Beyond the time limitations, sampling of a range of CH_4 mixing ratios from emissions and appropriate background samples is not straightforward. Background sampling must capture the air into which emissions are released, but during flights the meteorological conditions often resulted in significant variation of CH_4 mixing ~~ratio~~ratios and $\delta^{13}C_{CH4}$ with altitude, in addition to horizontal variations. Where repeat transects were conducted at different altitudes, this made selection of appropriate background samples for Keeling plots challenging, since the background CH_4 mixing ratio and $\delta^{13}C$ varied over the different altitudes. This becomes particularly detrimental to Keeling plot validity where the range in sampled emission mixing ratios is small, since uncertainty in the background samples then becomes more important.

In Fig. 8, a sensitivity analysis is presented from one of the flights investigating the effect of reducing the number of samples on the uncertainty in the $\delta^{13}C_{CH4}$ source signature determined for a plume. In this case nine samples were collected, but over eight downwind transects and one upwind transect of a cluster of installations, which is not feasible to repeat for sampling large numbers of installations. As shown in Fig. 8, the uncertainty in the $\delta^{13}C_{CH4}$ source signatures increases only slightly with reduction in number of sampling points, with the exception of one n = 3 run where the source signature is poorly defined. A minimum of three data points can ~~be~~ therefore be sufficient for classifying a source of CH_4 emissions (such as thermogenic, microbial or pyrogenic sources), providing that the background and point samples are captured with a large enough range of CH_4 concentration, and providing that there is no mixing of sources. This will typically require collection of more than three samples, given some may miss the targeted plumes or potentially be lost during storage/processing as aforementioned. Although a two-point Keeling plot is technically possible, it is impossible to gauge the quality of the regression to be sure that only a single source has been captured.

6. Conclusions

Given the restrictions and time constraints on the science flights, important lessons for offshore oil and gas airborne measurement campaigns have been learned for rapid instrument re-fitting and agile deployment of a small aircraft for future campaigns. A key finding from this study is that offshore meteorological conditions define the ability of the flights to produce valuable data, and suitable meteorology with a well-mixed (neutral) boundary layer is critical to deriving a regional emission

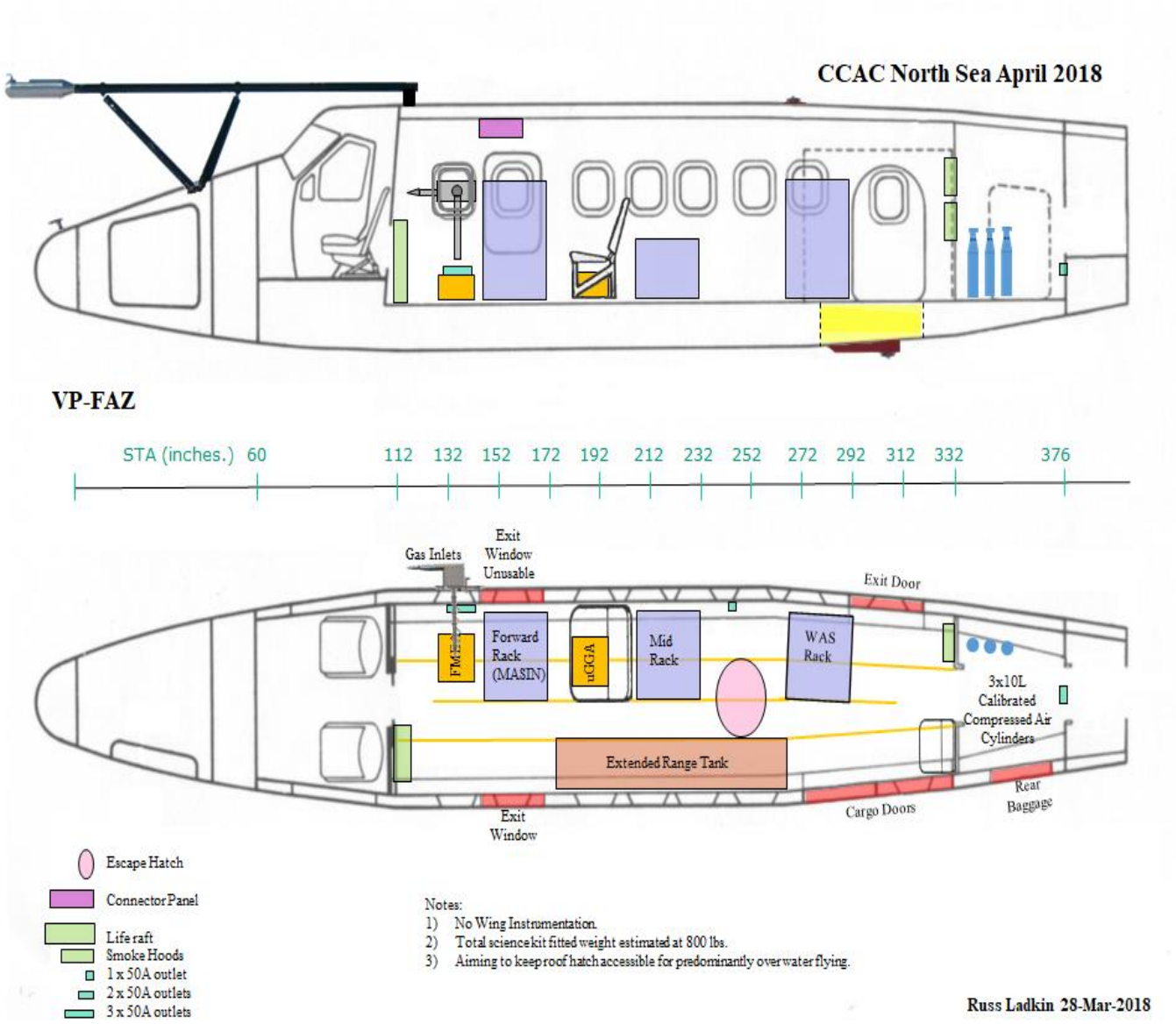
estimate through regional modelling. Flying in conditions with multiple residual boundary layers makes interpretation difficult and pin-pointing emissions especially challenging as emission plumes can easily be missed when they are trapped in thin filaments, ~~dramatically~~ increasing the uncertainties of measurement-based emission flux calculations. Although not possible for this work given aircraft scheduling, it is recommended that offshore observations are scheduled with a long window of opportunity to ensure optimal flying conditions. Predictions of the likelihood of a residual boundary layer over a coastal area could be achieved through high spatial resolution forecast models such as the UK Met Office forecast model (Milan et al., 2020). Information on the temperature structure over the previous few days using all the assimilated information, such as tephigrams and synoptic charts, would help determine the likelihood of residual boundary layers versus a simpler stratified, well mixed layer. For methods using alternative platforms such as ships or drones, ~~co-incidental~~coincidental measurements of vertical profiles must be made to capture the true nature of the emission plume in the current meteorology.

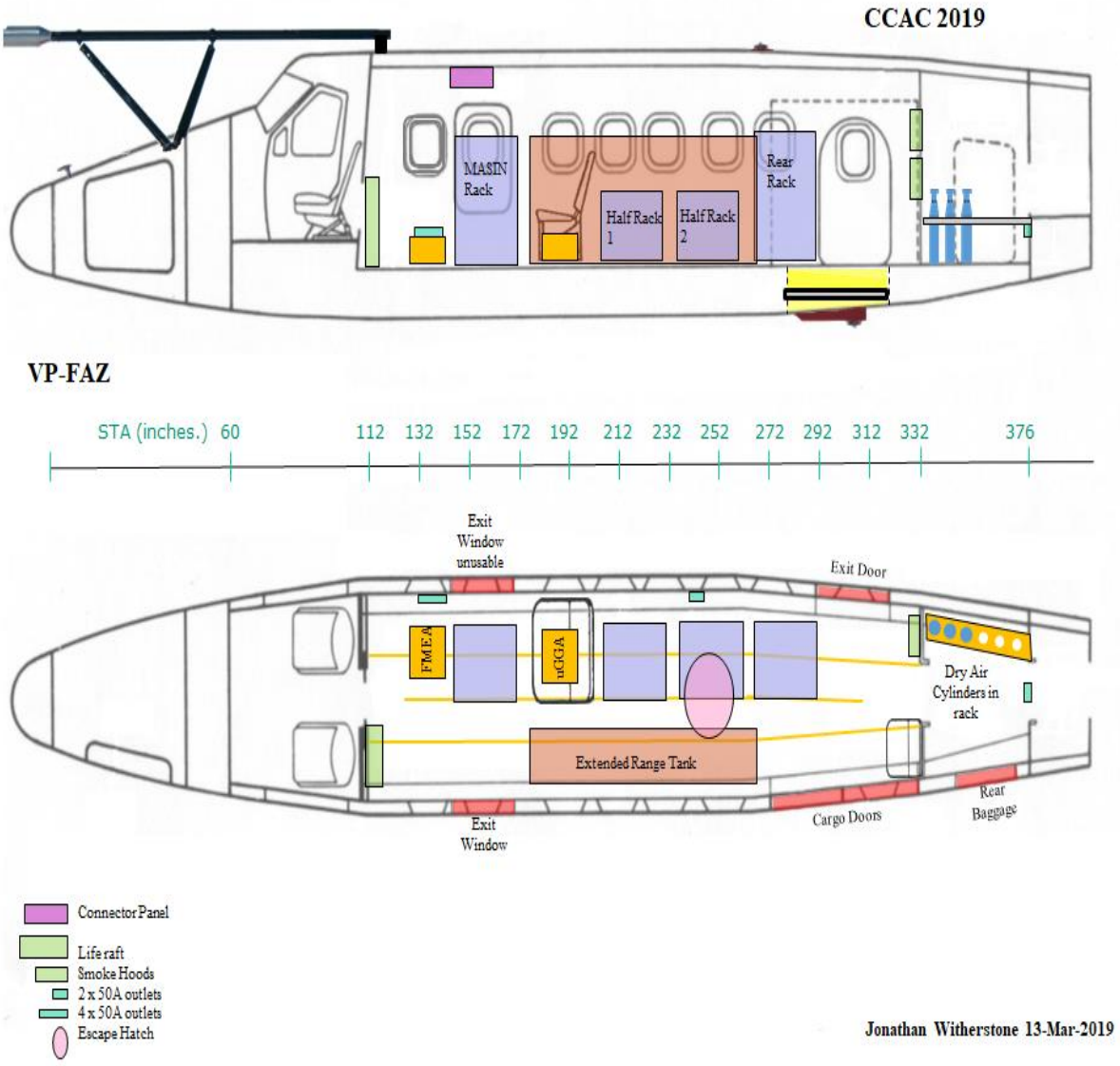
~~Unlike some~~Due to the size of the ~~larger~~ aircraft, payload restrictions and power limitations demand challenging decisions for instrument selection. We recommend deploying at least one instrument measuring CH_4 (and CO_2) at 10 Hz, allowing several plumes emitted from a single installation to be resolved (Fig. 6). Priority should next be given to a C_2H_6 ~~measurement~~instrument capable of sub ppb limit of detection at 1 Hz (or higher) in order to give certainty to the source of the CH_4 emission. Using C2:C1 appears to be the simplest method for source attribution, and is robust for distinguishing natural gas emissions, where the gas has an C_2H_6 component (Lowry et al., 2020; Plant et al., 2019). Spot sampling is challenging, payload heavy and time consuming as several passes are needed to collect enough samples (especially for $\delta^{13}C_{CH4}$ source attribution). However, results can be very informative such as the ability to distinguish between a gas leak, a geological reservoir from depth or near surface reservoir (Lee et al., 2018). The improvements to SWAS, allowing for continuous through flow, has increased the success rate of peak sampling, but still relies on accurate user triggering.

For mass balance flux calculations, an ~~instantaneous~~emission plume and the surrounding background variation in the species of interest, alongside local meteorology, must be fully resolved during the observation stage. This includes instruments with appropriate response times to fully capture the plume and identify any internal structure that may suggest a mixed source. An upwind leg must be conducted to ensure the plume and background is not contaminated by extraneous far-field sources and the plume must be significantly distinct from this background for meaningful flux calculations. The plume must be laterally and vertically resolved in the 2D plane as much as possible at a fixed distance downwind of the source. Straight and level runs must extend either side of the plume and the vertical resolution must include multiple stacked transects with an identification of the top and bottom of the plume (where feasible) to reduce uncertainty in the plume bulk net flux. Full understanding of the meteorology ~~during data collection~~ with meteorological measurement instrumentation and ~~an entire~~a complete profile to determine the marine boundary layer characteristics from the top to the surface, including determination of inversion heights, must be conducted during the flight day when appropriate radiosonde soundings are not available. The observed impact of complex boundary layer dynamics on plume dispersion also highlights an important limitation of ship-based plume

measurements, which are unable to resolve the vertical structure of the plume and therefore rely on the assumption of idealised models of plume dispersion.

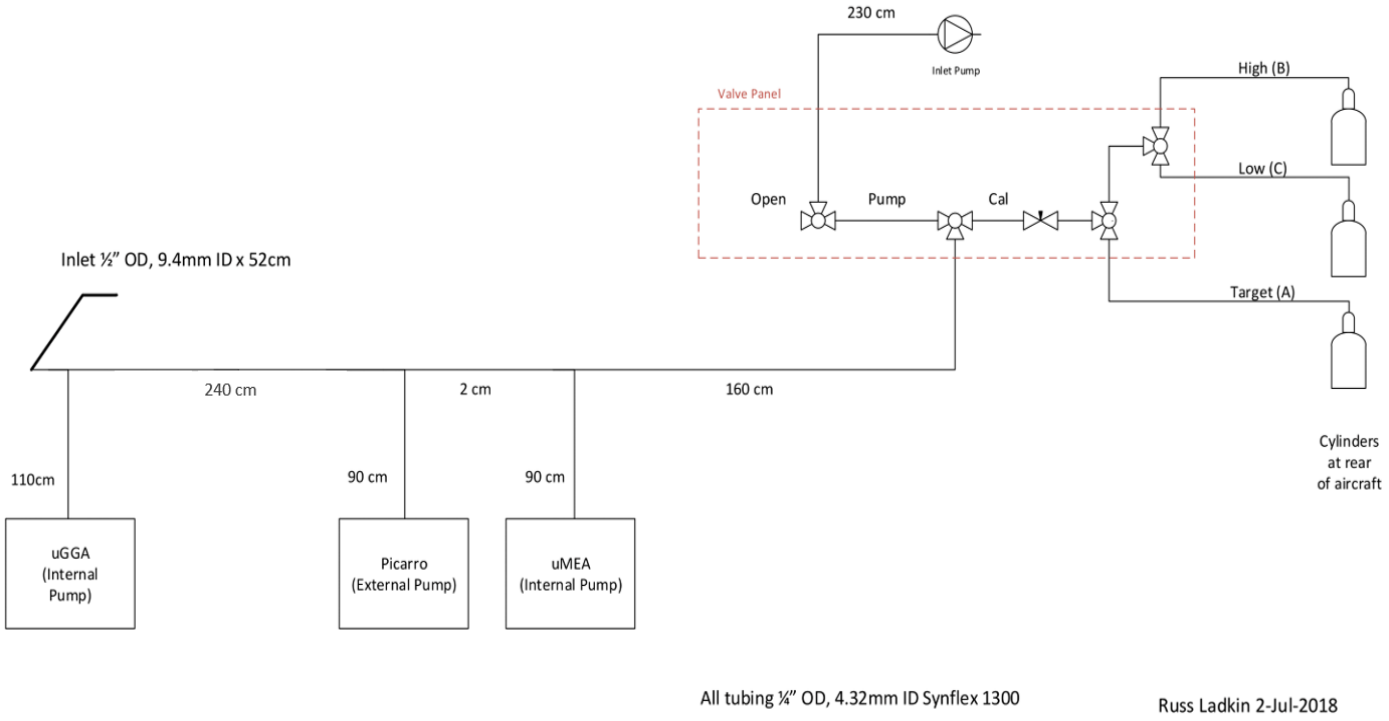
7. Figures and Tables





715

Figure 1. Instrument schematics for the Twin Otter aircraft as deployed in 2018 and 2019, detailing changes in layout and instrumentation between the two campaigns. The top panel is the 2018 fit and lower panel 2019 fit.



720

Figure 2. Layout of plumbing of the calibration system (and inlet system) for 2018 campaign.

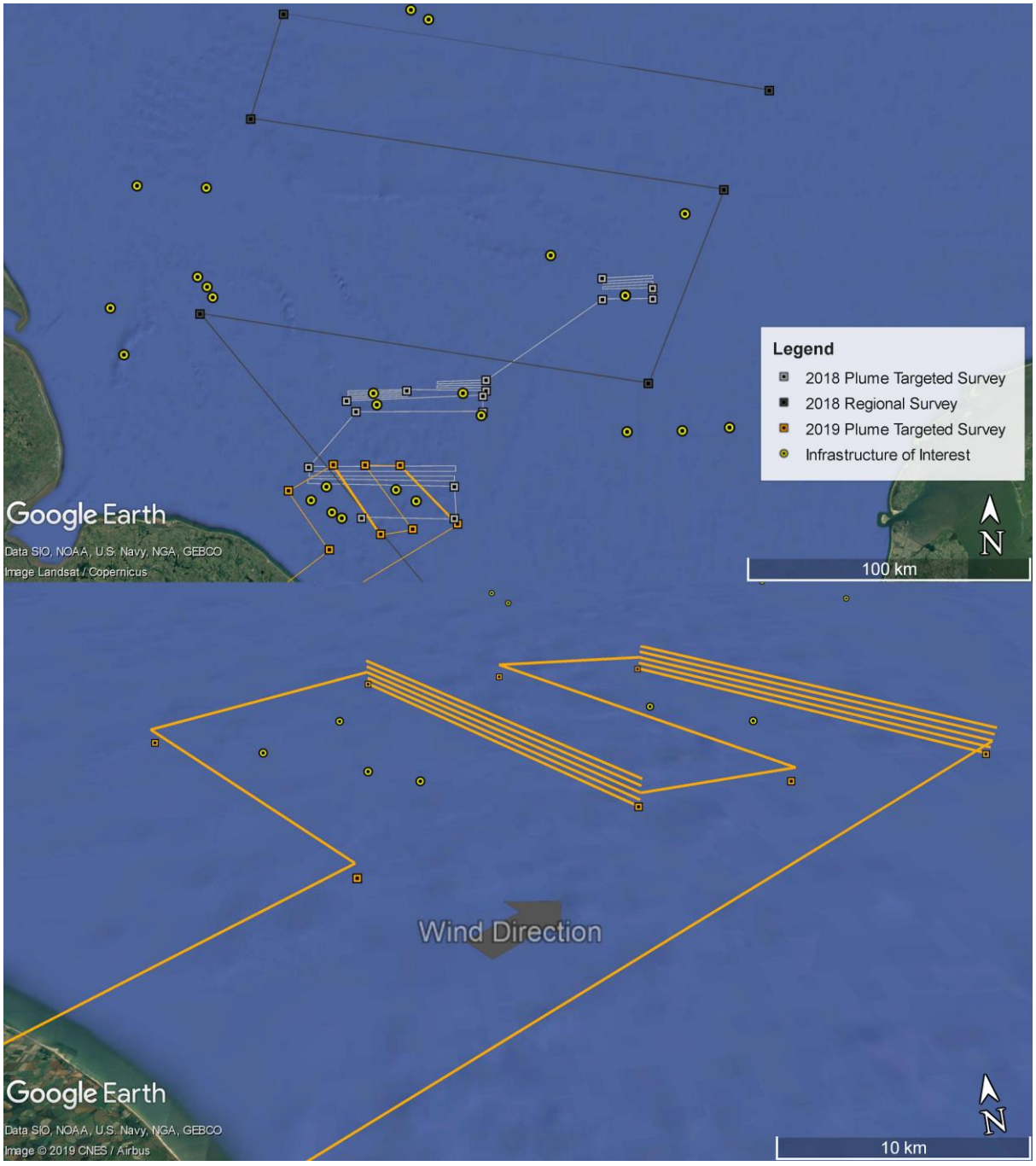


Figure 3: [~~top~~Top panel] flight patterns showing the regional and plume capture styles of flight deployed between 2018 and 2019, alongside infrastructure of interest (such as drilling rigs, gas distribution platforms or production platforms). [~~bottom~~Bottom panel] a 2019 plume sampling survey showing idealised stacked transects in the 2D plane downwind of infrastructure of interest.

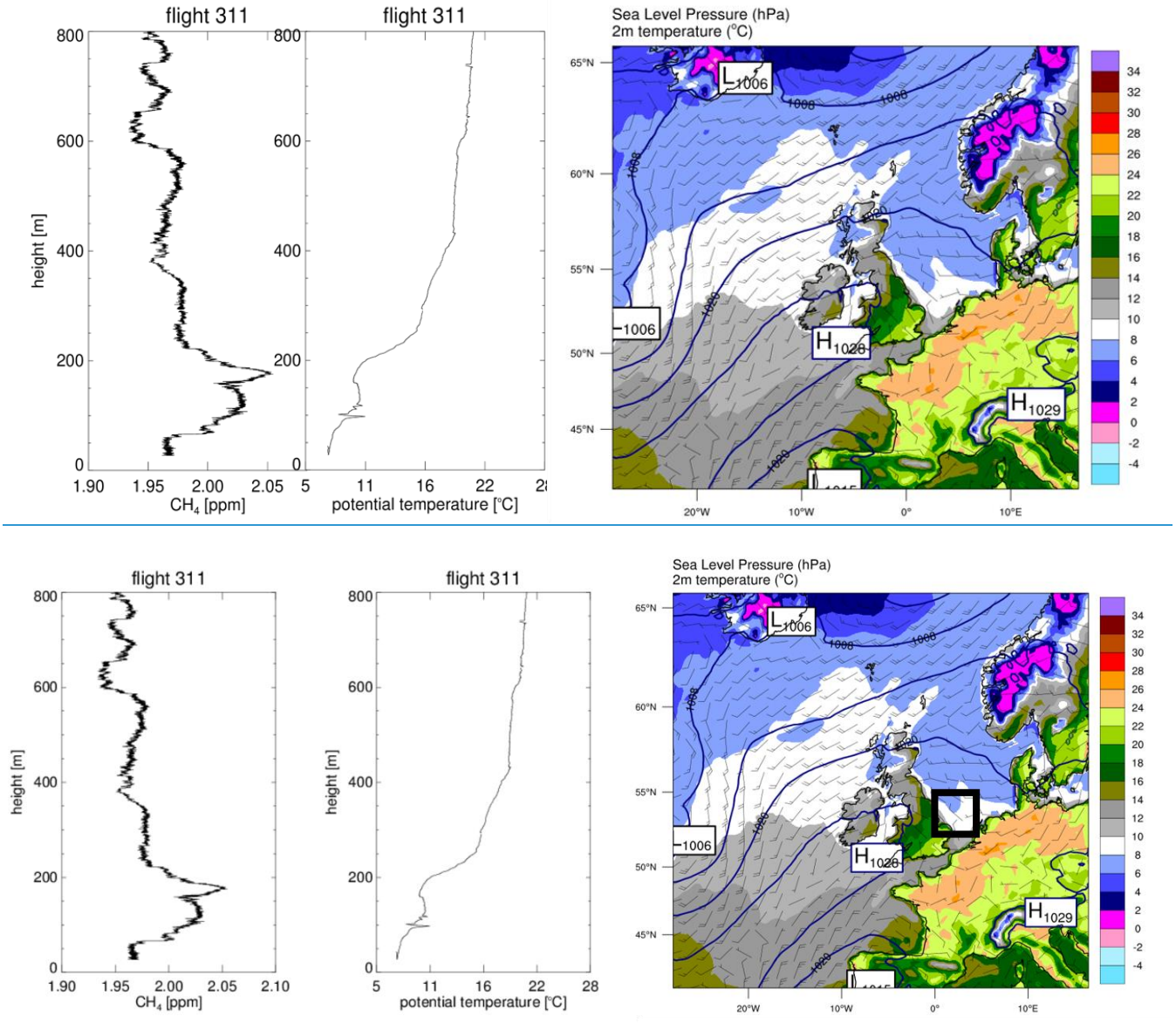
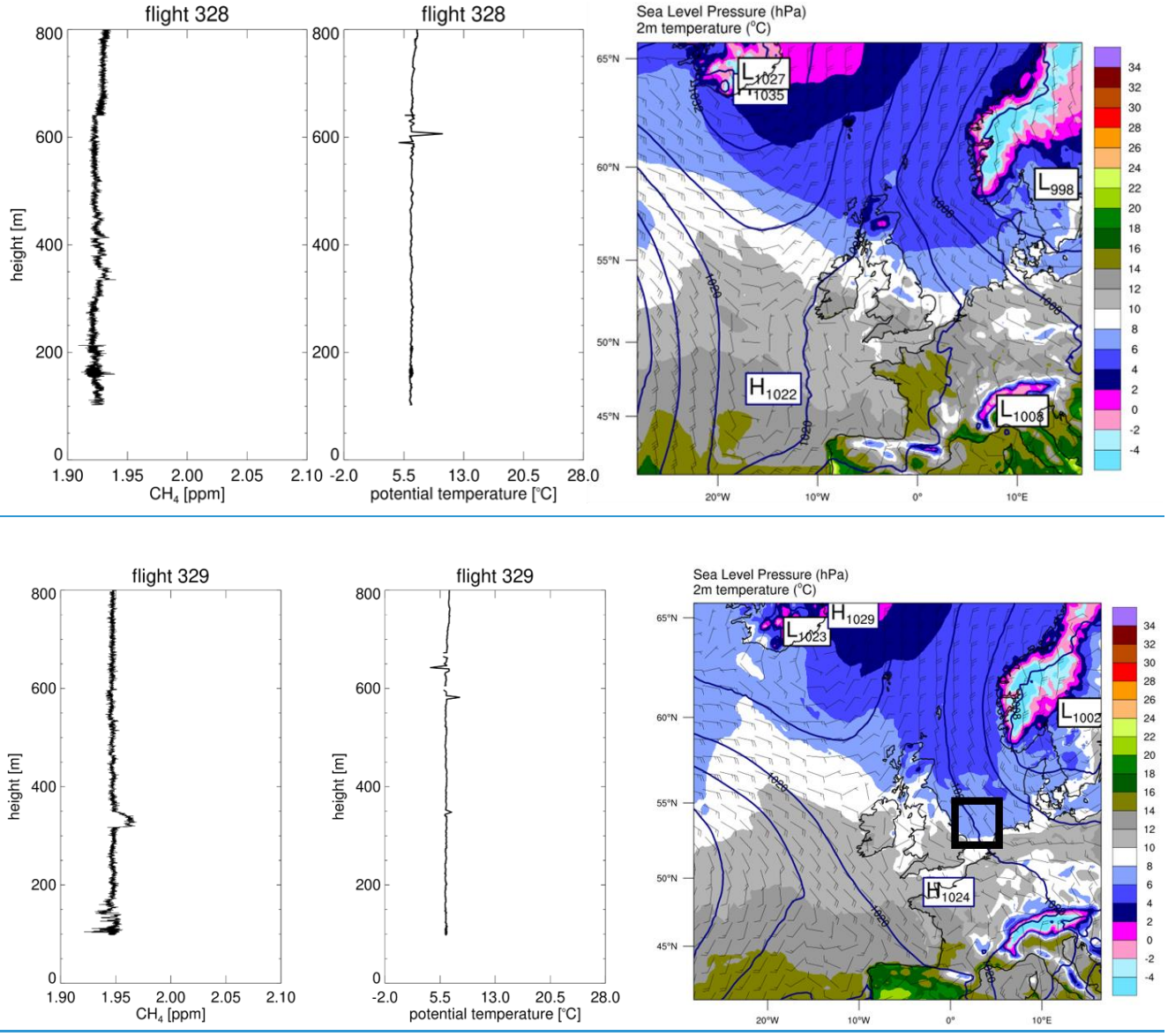


Figure 4. Example of CH_4 and potential temperature profiles showing the large amount of structure arising from residual boundary layers. The increase of the potential temperature with height shows stable stratification of the boundary layer. The synoptic chart over the eastern North Atlantic and ~~north~~ North-West Europe shows contoured sea level pressure (hPa), ~~2m~~2 m temperature (°C, right-hand side colour scale) and wind for 20/04/2018 12:00 UT, and reveals relatively low wind speeds and poorly defined air flow over the southern North Sea sector, allowing the build-up of residual boundary layers. Synoptic chart image produced by the UK National Centre for Atmospheric Science (NCAS) using Weather Research and Forecasting model WFR-ARW version 3.7.1, with a 20 km grid spacing, 51 vertical levels initialised using the NOAA Global Forecast System. NCAS Weather Research Catalogue (sci.ncas.ac.uk/nwr/pages/home~~)~~). The black rectangle approximates the survey region.



740 **Figure 5. Example of CH_4 and potential temperature profiles in a well-mixed boundary layer under neutral conditions. ~~These are composite profiles flown within an hour over water.~~ The potential temperature and CH_4 profiles stay relatively constant ~~above~~ 200 and CH_4 shows only an increase in the surface layer and when intercepting an enhancement at 300 to 350 m height. The synoptic chart for ~~03~~06/05/2019 12:00 UT shows a cyclonic south-easterly air flow over the southern North Sea sector originating from the Benelux region. The black rectangle approximates the survey region over open water.**

745

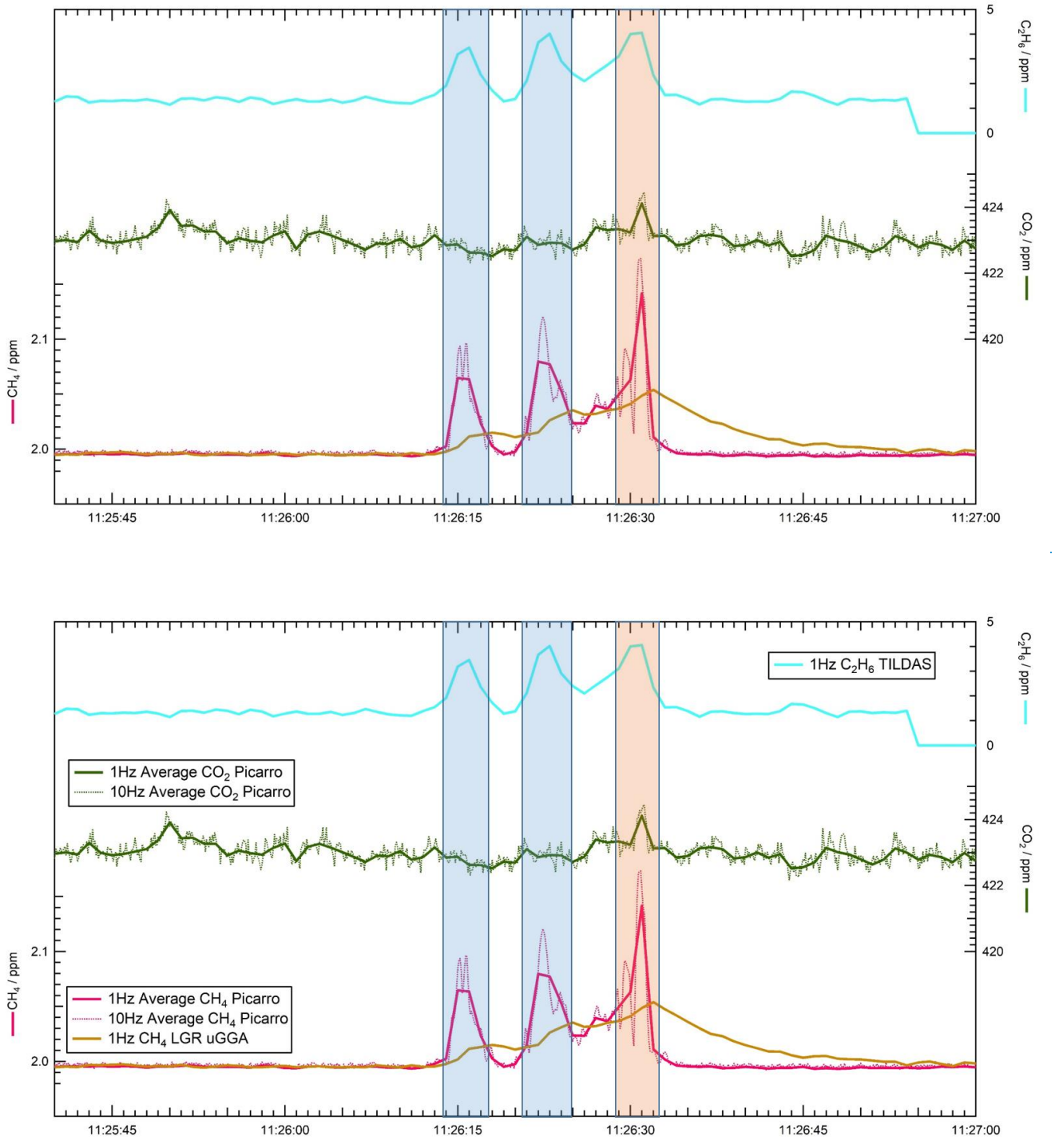

5
0
C_2H_6 / ppm
424
422
420
CO_2 / ppm
2.1
2.0
CH_4 / ppm
11:25:45
11:26:00
11:26:15
11:26:30
11:26:45
11:27:00
1Hz C_2H_6 TILDAS
1Hz Average CO_2 Picarro
10Hz Average CO_2 Picarro
1Hz Average CH_4 Picarro
10Hz Average CH_4 Picarro
1Hz CH_4 LGR uGGA

Figure 6. A cross-section of CH_4, CO_2 and C_2H_6 measurement response during one plume sample as recorded by Picarro ~~G2301~~G2311-f in ~~Pink~~pink and ~~Green~~green (10 Hz as dashed lines and downsampled to 1 Hz as solid lines), TILDAS 1 Hz in cyan and Los Gatos uGGA 1 Hz in brown. The difference ~~of~~between the uGGA and Picarro at 1 Hz arises from the slower uGGA response time due to the slower cell turnover. The blue shaded area shows enhancement in C_2H_6 and CH_4, indicating cold venting, the orange shaded area shows enhancement in C_2H_6, CH_4 and a small amount of CO_2 potentially indicating a co-located combustion source.

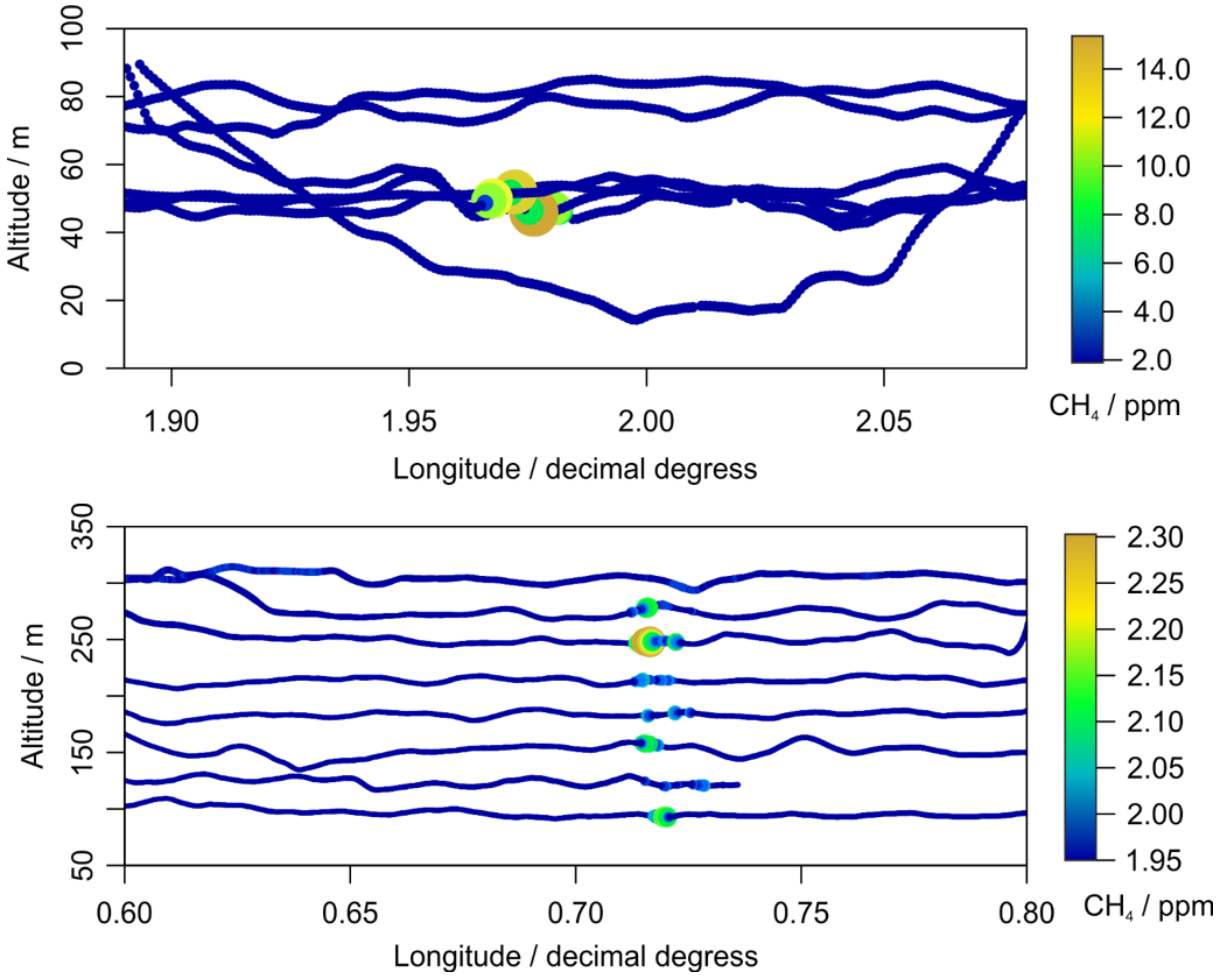
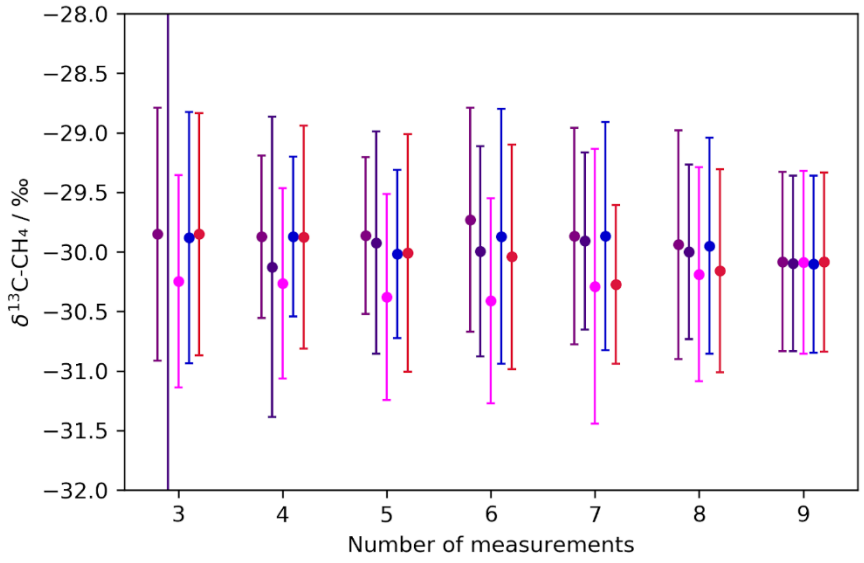


Figure 7. Plumes measured from separate installations to demonstrate the differences in strategies between 2018 and 2019. [~~upper~~Top panel] Plume sampled downwind with poorer vertical spatial resolution in the 2D plane during the 2018 portion of the campaign. CH_4 measured values are much higher due to platform activities during the survey time. [~~lower~~Bottom panel] Plume sampled downwind in 2019 with intermediate transects enabling higher vertical spatial resolution. Note the colour scale across each plot signifies different measured CH_4; the scales on the upper and lower plots are different.



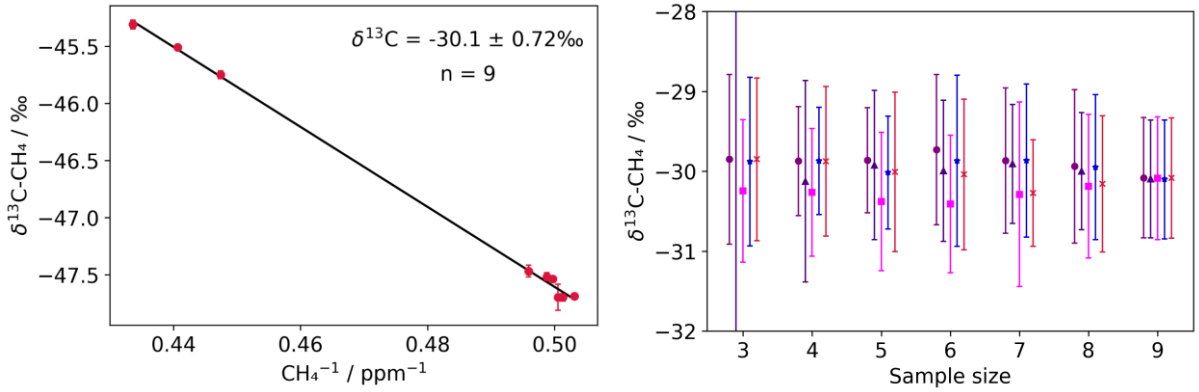


Figure 8. ~~$\delta^{13}C_{CH4}$ source signature data source signature data derived from data~~ (a) Keeling plot determined using nine samples collected around one installation ~~and assumes a~~, assumed as the single source of excess ~~methane. The coloured bars represent the~~ CH_4. (b) An illustration of the variation in $\delta^{13}C_{CH4}$ source signature and its uncertainty ~~in the intercept~~determined by Keeling plot analyses for ~~the regression with varying numbers of~~ reduced sample sizes. Each analysis represents a single monte-carlo experiment with the original data, reducing the number of data points ~~removed from the dataset~~ to ~~represent~~ the ~~data if fewer in plume samples had been collected.~~sample size indicated at random, the $\delta^{13}C_{CH4}$ source signature is then calculated with the remaining sample points. Error bars are two times standard error.

Survey Year	**CH_4 flux lower bound (kT yr^{-1})**	**CH_4 flux upper bound (kT yr^{-1})**

2018	1.83	17.9
2019	0.67	1.04

Table 1. A comparison of flux lower and upper bounds for two individual example plumes across each year of survey as scaled by the vertical resolution available. The plumes themselves are not comparable, but the method changes demonstrate the increased certainty in the final results.

	Instrument(s)	Method	C2:C1	Uncertainty
2018 flight	Los Gatos ultraportable CH_4/C_2H_6	Linear regression	0.029	± 0.014
2019 flight	TILDAS C_2H_6 & Picarro ~~G2301~~G2311-f CH_4	Plume area integration	0.029	± 0.003
Published well data			0.031	± 0.009

Table 2. Reported data for C2:C1 for a single installation surveyed during both 2018 and 2019 surveys. Well data from UK oil and gas authority report: https://dataogauthority.blob.core.windows.net/external/DataReleases/ShellExxonMobil/GeochemSNS.zip alongside measured C2:C1 for CH_4 enhancements measured during flights in the same geographic area.

Author Contribution

The manuscript was written and figures prepared by ~~JLF~~JF, PB and PD with assistance from AEJ, MC, JP, SB and ~~JTS~~JS. Experimental design and flight planning was performed by GA, JP, JDL, TLC and DL. Aircraft set-up and in-flight measurements performed by JP, PB, PD, ~~SJA~~SA, SY, ~~AIW~~AW, TLC, ~~JLF~~JF, SW, JW and SB. Laboratory measurements made by REF, RP, SW and SB. Data Processing and calibrations performed by ~~JLF~~JF, LH, PB, ~~JTS~~JS, PD and MC and SB. Modelling work by NW, ~~JAP~~JP, PB and LH.

Code / Data Availability

The data for this work will be available via request at the British Antarctic Survey Polar Data Centre.

Competing interests

The authors declare that they have no conflict of interest

Acknowledgements

This work was funded under the Climate and Clean Air Coalition (CCAC) Oil and Gas Methane Science Studies (MSS), hosted by the United Nations Environment Programme. Funding was provided by the Environmental Defense Fund, Oil and Gas Climate Initiative, European Commission, and CCAC.

References

————Cain, M., Warwick, N. J., Fisher, R. E., Lowry, D., Lanoisellé, M., Nisbet, E. G., France, J., Pitt, J., O'Shea, S., Bower, K. N., Allen, G., Illingworth, S., Manning, A. J., Bauguitte, S., Pisso, I., and Pyle, J. A.: A cautionary tale: A study of a methane enhancement over the North Sea, Journal of Geophysical Research: Atmospheres, 122, 7630-7645, 2017.

————Cardoso-Saldaña, F. J., Kimura, Y., Stanley, P., McGaughey, G., Herndon, S. C., Roscioli, J. R., Yacovitch, T. I., and Allen, D. T.: Use of Light Alkane Fingerprints in Attributing Emissions from Oil and Gas Production, Environmental Science & Technology, 53, 5483-5492, 2019.

————Conley, S., Faloona, I., Mehrotra, S., Suard, M., Lenschow, D. H., Sweeney, C., Herndon, S., Schwietzke, S., Pétron, G., Pifer, J., Kort, E. A., and Schnell, R.: Application of Gauss's theorem to quantify localized surface emissions from airborne measurements of wind and trace gases, Atmos. Meas. Tech., 10, 3345-3358, 2017.

————Conley, S., Franco, G., Faloona, I., Blake, D. R., Peischl, J., and Ryerson, T. B.: Methane emissions from the 2015 Aliso Canyon blowout in Los Angeles, CA, Science, 351, 1317-1320, 2016.

————Crawford, T. L., Dobosy, R. J., and Dumas, E. J.: Aircraft wind measurement considering lift-induced upwash, Boundary-Layer Meteorology, 80, 79-94, 1996.

————Edwards, P. M., Young, C. J., Aikin, K., deGouw, J., Dubé, W. P., Geiger, F., Gilman, J., Helmig, D., Holloway, J. S., Kercher, J., Lerner, B., Martin, R., McLaren, R., Parrish, D. D., Peischl, J., Roberts, J. M., Ryerson, T. B., Thornton, J., Warneke, C., Williams, E. J., and Brown, S. S.: Ozone photochemistry in an oil and natural gas extraction region during winter: simulations of a snow-free season in the Uintah Basin, Utah, Atmos. Chem. Phys., 13, 8955-8971, 2013.

————Fisher, R., Lowry, D., Wilkin, O., Sriskantharajah, S., and Nisbet, E. G.: High-precision, automated stable isotope analysis of atmospheric methane and carbon dioxide using continuous-flow isotope-ratio mass spectrometry, Rapid Communications in Mass Spectrometry, 20, 200-208, 2006.

————Fisher, R. E., France, J. L., Lowry, D., Lanoisellé, M., Brownlow, R., Pyle, J. A., Cain, M., Warwick, N., Skiba, U. M., Drewer, J., Dinsmore, K. J., Leeson, S. R., Bauguitte, S. J.-B., Wellpott, A., O'Shea, S. J., Allen, G., Gallagher, M. W., Pitt, J., Percival, C. J., Bower, K., George, C., Hayman, G. D., Aalto, T., Lohila, A., Aurela, M., Laurila, T., Crill, P. M., McCalley, C. K., and Nisbet, E. G.: Measurement of the 13C isotopic signature of methane emissions from northern European wetlands, Global Biogeochemical Cycles, 31, 605-623, 2017.

————Garman, K. E., Hill, K. A., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: An Airborne and Wind Tunnel Evaluation of a Wind Turbulence Measurement System for Aircraft-Based Flux Measurements, Journal of Atmospheric and Oceanic Technology, 23, 1696-1708, 2006.

————Garman, K. E., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: The Contribution of Variability of Lift-induced Upwash to the Uncertainty in Vertical Winds Determined from an Aircraft Platform, Boundary-Layer Meteorology, 126, 461-476, 2008.

————Gorchov Negron, A. M., Kort, E. A., Conley, S. A., and Smith, M. L.: Airborne Assessment of Methane Emissions from Offshore Platforms in the U.S. Gulf of Mexico, Environmental Science & Technology, 54, 5112-5120, 2020.

————Gvakharia, A., Kort, E. A., Brandt, A., Peischl, J., Ryerson, T. B., Schwarz, J. P., Smith, M. L., and Sweeney, C.: Methane, Black Carbon, and Ethane Emissions from Natural Gas Flares in the Bakken Shale, North Dakota, Environmental Science & Technology, 51, 5317-5325, 2017.

————Gvakharia, A., Kort, E. A., Smith, M. L., and Conley, S.: Testing and evaluation of a new airborne system for continuous N_2O, CO_2, CO, and H_2O measurements: the Frequent Calibration High-performance Airborne Observation System (FCHAOS), Atmos. Meas. Tech., 11, 6059-6074, 2018.

————Hopkins, J. R., Lewis, A. C., and Read, K. A.: A two-column method for long-term monitoring of non-methane hydrocarbons (NMHCs) and oxygenated volatile organic compounds (o-VOCs), Journal of Environmental Monitoring, 5, 8-13, 2003.

————Jones, A., Thomson, D., Hort, M., and Devenish, B.: The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III, Boston, MA2007, 580-589.

————Keeling, C. D.: The concentration and isotopic abundances of carbon dioxide in rural and marine air, Geochimica et Cosmochimica Acta, 24, 277-298, 1961.

———Kostinek, J., Roiger, A., Davis, K. J., Sweeney, C., DiGangi, J. P., Choi, Y., Baier, B., Hase, F., Groß, J., Eckl, M., Klausner, T., and Butz, A.: Adaptation and performance assessment of a quantum and interband cascade laser spectrometer for simultaneous airborne in situ observation of CH_4, C_2H_6, CO_2, CO and N_2O, Atmos. Meas. Tech., 12, 1767-1783, 2019.

Lee, J. D., Mobbs, S. D., Wellpott, A., Allen, G., Bauguitte, S. J. B., Burton, R. R., Camilli, R., Coe, H., Fisher, R. E., France, J. L., Gallagher, M., Hopkins, J. R., Lanoiselle, M., Lewis, A. C., Lowry, D., Nisbet, E. G., Purvis, R. M., O'Shea, S., Pyle, J. A., and Ryerson, T. B.: Flow rate and source reservoir identification from airborne chemical sampling of the uncontrolled Elgin platform gas release, Atmos. Meas. Tech., 11, 1725-1739, 2018.

———Lowry, D., Fisher, R. E., France, J. L., Coleman, M., Lanoisellé, M., Zazzeri, G., Nisbet, E. G., Shaw, J. T., Allen, G., Pitt, J., and Ward, R. S.: Environmental baseline monitoring for shale gas development in the UK: Identification and geochemical characterisation of local source emissions of methane to atmosphere, Science of The Total Environment, 708, 134600, 2020.

———Milan, M., Macpherson, B., Tubbs, R., Dow, G., Inverarity, G., Mittermaier, M., Halloran, G., Kelly, G., Li, D., Maycock, A., Payne, T., Piccolo, C., Stewart, L., and Wlasak, M.: Hourly 4D-Var in the Met Office UKV operational forecast model, Quarterly Journal of the Royal Meteorological Society, doi: 10.1002/qj.3737, 2020. 1-21, 2020.

———Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (Eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

———Nelson, D. D., McManus, B., Urbanski, S., Herndon, S., and Zahniser, M. S.: High precision measurements of atmospheric nitrous oxide and methane using thermoelectrically cooled mid-infrared quantum cascade lasers and detectors, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 60, 3325-3335, 2004.

———Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Myhre, C. L., Platt, S. 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 Biogeochemical Cycles, 33, 318-342, 2019.

———O'Shea, S. J., Allen, G., Gallagher, M. W., Bower, K., Illingworth, S. M., Muller, J. B. A., Jones, B. T., Percival, C. J., Bauguitte, S. J. B., Cain, M., Warwick, N., Quiquet, A., Skiba, U., Drewer, J., Dinsmore, K., Nisbet, E. G., Lowry, D., Fisher, R. E., France, J. L., Aurela, M., Lohila, A., Hayman, G., George, C., Clark, D. B., Manning, A. J., Friend, A. D., and Pyle, J.: Methane and carbon dioxide fluxes and their regional scalability for the European Arctic wetlands during the MAMM project in summer 2012, Atmos. Chem. Phys., 14, 13159-13174, 2014.

———O'Shea, S. J., Bauguitte, S. J. B., Gallagher, M. W., Lowry, D., and Percival, C. J.: Development of a cavity-enhanced absorption spectrometer for airborne measurements of ~~CH₄~~CH_4 and ~~CO₂~~CO_2, Atmos. Meas. Tech., 6, 1095-1109, 2013.

———Pataki, D. E., Ehleringer, J. R., Flanagan, L. B., Yakir, D., Bowling, D. R., Still, C. J., Buchmann, N., Kaplan, J. O., and Berry, J. A.: The application and interpretation of Keeling plots in terrestrial carbon cycle research, Global Biogeochemical Cycles, 17, 2003.

Peischl, J., Eilerman, S. J., Neuman, J. A., Aikin, K. C., de Gouw, J., Gilman, J. B., Herndon, S. C., Nadkarni, R., Trainer, M., Warneke, C., and Ryerson, T. B.: Quantifying Methane and Ethane Emissions to the Atmosphere From Central and Western U.S. Oil and Natural Gas Production Regions, Journal of Geophysical Research: Atmospheres, 123, 7725-7740, 2018.

Pitt, J. R., Allen, G., Bauguitte, S. J. B., Gallagher, M. W., Lee, J. D., Drysdale, W., Nelson, B., Manning, A. J., and Palmer, P. I.: Assessing London CO_2, CH_4 and CO emissions using aircraft measurements and dispersion modelling, Atmos. Chem. Phys., 19, 8931-8945, 2019.

Pitt, J. R., Le Breton, M., Allen, G., Percival, C. J., Gallagher, M. W., Bauguitte, 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 N_2O and CH_4 sampling using a quantum cascade laser absorption spectrometer (QCLAS), Atmos. Meas. Tech., 9, 63-77, 2016.

Plant, G., Kort, E. A., Floerchinger, C., Gvakharia, A., Vimont, I., and Sweeney, C.: Large Fugitive Methane Emissions From Urban Centers Along the U.S. East Coast, Geophysical Research Letters, 46, 8500-8507, 2019.

Rella, C. W., Hoffnagle, J., He, Y., and Tajima, S.: Local- and regional-scale measurements of CH_4, $\delta^{13}CH_4$, and C_2H_6 in the Uintah Basin using a mobile stable isotope analyzer, Atmos. Meas. Tech., 8, 4539-4559, 2015.

Riddick, S. N., Mauzerall, D. L., Celia, M., Harris, N. R. P., Allen, G., Pitt, J., Staunton-Sykes, J., Forster, G. L., Kang, M., Lowry, D., Nisbet, E. G., and Manning, A. J.: Methane emissions from oil and gas platforms in the North Sea, Atmos. Chem. Phys., 19, 9787-9796, 2019.

Ryerson, T. B., Camilli, R., Kessler, J. D., Kujawinski, E. B., Reddy, C. M., Valentine, D. L., Atlas, E., Blake, D. R., de Gouw, J., Meinardi, S., Parrish, D. D., Peischl, J., Seewald, J. S., and Warneke, C.: Chemical data quantify Deepwater Horizon hydrocarbon flow rate and environmental distribution, Proceedings of the National Academy of Sciences, 109, 20246-20253, 2012.

————Santoni, G. W., Daube, B. C., Kort, E. A., Jiménez, R., Park, S., Pittman, J. V., Gottlieb, E., Xiang, B., Zahniser, M. S., Nelson, D. D., McManus, J. B., Peischl, J., Ryerson, T. B., Holloway, J. S., Andrews, A. E., Sweeney, C., Hall, B., Hintsa, E. J., Moore, F. L., Elkins, J. W., Hurst, D. F., Stephens, B. B., Bent, J., and Wofsy, S. C.: Evaluation of the airborne quantum cascade laser spectrometer (QCLS) measurements of the carbon and greenhouse gas suite CO_2, CH_4, N_2O, and CO during the CalNex and HIPPO campaigns, Atmos. Meas. Tech., 7, 1509-1526, 2014.

Saunois, M., Bousquet, P., Poulter, B., Peregon, 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, 697-751, 2016.

————Schwietzke, S., Harrison, M., Lauderdale, T., Branson, K., Conley, S., George, F. C., Jordan, D., Jersey, G. R., Zhang, C., Mairs, H. L., Pétron, G., and Schnell, R. C.: Aerially guided leak detection and repair: A pilot field study for evaluating the potential of methane emission detection and cost-effectiveness, Journal of the Air & Waste Management Association, 69, 71-88, 2019.

————Sherwood, O. A., Schwietzke, S., Arling, V. A., and Etiope, G.: Global Inventory of Gas Geochemistry Data from Fossil Fuel, Microbial and Burning Sources, version 2017, Earth Syst. Sci. Data, 9, 639-656, 2017.

————Warneke, C., Trainer, M., de Gouw, J. A., Parrish, D. D., Fahey, D. W., Ravishankara, A. R., Middlebrook, A. M., Brock, C. A., Roberts, J. M., Brown, S. S., Neuman, J. A., Lerner, B. M., Lack, D., Law, D., Hübler, G., Pollack, I., Sjostedt, S., Ryerson, T. B., Gilman, J. B., Liao, J., Holloway, J., Peischl, J., Nowak, J. B., Aikin, K., Min, K. E., Washenfelder, R. A., Graus, M. G., Richardson, M., Markovic, M. Z., Wagner, N. L., Welti, A., Veres, P. R., Edwards, P., Schwarz, J. P., Gordon, T., Dube, W. P., McKeen, S., Brioude, J., Ahmadov, R., Bougiatioti, A., Lin, J. J., Nenes, A., Wolfe, G. M., Hanisco, T. F., Lee, B. H., Lopez-Hilfiker, F. D., Thornton, J. A., Keutsch, F. N., Kaiser, J., Mao, J., and Hatch, C.: Instrumentation and Measurement Strategy for the NOAA SENEX Aircraft Campaign as Part of the Southeast Atmosphere Study 2013, Atmos Meas Tech, 9, 3063-3093, 2016.

————Weiss, A. I., King, J., Lachlan-Cope, T., and Ladkin, R.: On the effective aerodynamic and scalar roughness length of Weddell Sea ice, Journal of Geophysical Research: Atmospheres, 116, 2011.

————Yacovitch, T. I., Daube, C., and Herndon, S. C.: Methane Emissions from Offshore Oil and Gas Platforms in the Gulf of Mexico, Environmental Science & Technology, 54, 3530-3538, 2020.

————Yacovitch, T. I., Herndon, S. C., Roscioli, J. R., Floerchinger, C., McGovern, R. M., Agnese, M., Pétron, G., Kofler, J., Sweeney, C., Karion, A., and Conley, S. A.: Demonstration of an ethane spectrometer for methane source identification, Environ. Sci. Technol., 48, 8028, 2014a.

————Yacovitch, T. I., Herndon, S. C., Roscioli, J. R., Floerchinger, C., McGovern, R. M., Agnese, M., Pétron, G., Kofler, J., Sweeney, C., Karion, A., Conley, S. A., Kort, E. A., Nähle, L., Fischer, M., Hildebrandt, L., Koeth, J., McManus, J. B., Nelson, D. D., Zahniser, M. S., and Kolb, C. E.: Demonstration of an Ethane Spectrometer for Methane Source Identification, Environmental Science & Technology, 48, 8028-8034, 2014b.

————Zavala-Araiza, D., Alvarez, R. A., Lyon, D. R., Allen, D. T., Marchese, A. J., Zimmerle, D. J., and Hamburg, S. P.: Super-emitters in natural gas infrastructure are caused by abnormal process conditions, Nat. Commun., 8, 14012, 2017.

Appendices

A1. TILDAS data processing and performance

The TILDAS data was processed as follows. Rapid tuning sweeps of the laser frequency (2996.8 to 2998.0 cm^{-1}) by varying the applied current result in the collection of thousands of spectra per second, which are co-averaged. The resulting averaged spectrum is processed at a rate of 1 Hz using a nonlinear least-squares fitting algorithm to determine mixing ratios within the operating software, TDLWintel (© Aerodyne). Averaging of these spectra, and the path length of 76 m achieved using a Herriott multipass cell, provide the sensitivity required for trace gas measurement. Continuously circulated fluid from the Oasis chiller unit is used as a heat sink for the thermodynamically cooled components and a flow interlock cuts power to the relevant components if the coolant flow stops. Other optical components of the instrument include a 15x Schwarzchild objective in front of each laser, a germanium etalon for measuring the laser tuning rate, a reference gas cell containing air at 25 Torr, and numerous mirrors for adjusting the laser beam alignment. During the airborne campaign the instrument was operated remotely via an Ethernet connection. The TILDAS C_2H_6 instrument accuracy has been tested against two standards containing C_2H_6 in mixing ratios of 39.79 ± 0.14 ppb and 2.08 ± 0.02 ppb (high concentration standard and target gas, respectively). As the TILDAS technique relies on highly precise alignment of the focussing and beam-alignment optics before and after the multipass measurement cell it is particularly prone to motion that applies torque to the optical bench. To remove measurement artefacts associated with this sensitivity all data collected for roll angles greater than 20 degrees has been flagged. The presence of the TILDAS in the 2019 campaign ruled out using the multiple circular pass method around a potential emission source as developed by Scientific Aviation for installation emission flux measurements (Conley et al., 2017) as there was a risk of invalidating data due to the roll angle of the plane if circling tightly around an installation.

A2. CO_2 and CH_4 Calibration

The three cylinders were sampled periodically in-flight to determine the instrument gain factor (slope) and zero-offset for each analyser. These parameters were linearly interpolated between calibrations and used to rescale the raw measured data (for further details see Pitt et al., 2016). The uncertainties associated with instrument drift and any instrument non-linearity were assessed by sampling the "target" cylinder mid-way between high-low calibrations. The raw target cylinder measurements were rescaled as per the sample data; the mean offset of these target measurements from the WMO-traceable cylinder value (and associated standard deviations) are given for the LGR uGGA and Picarro instrument and are plotted in Figure A4.

The typical duration of calibration cylinder measurements during the 2018 campaign was 45s. The Picarro ~~G2301~~G2311-f analyser had a high flow rate of ~5 SLPM resulting in rapid flushing of both the inlet tubing and sample cavity. The measured value for each calibration was taken as the average over 15 s prior to the calibration end, as this allowed sufficient time for the measured value to reach equilibrium. The uGGA and uMEA both had much lower flow rates of ~0.5 SLPM, resulting in a much longer equilibration time. Consequently, the calibration duration was not of sufficient length for the uGGA and uMEA

measurements to reach equilibrium and their calibration routine was compromised. For these instruments each calibration run was fitted to an offset exponential function in an attempt to predict the mixing ratio at which equilibration would have occurred, given an infinite amount of calibrating time. In order to improve the data quality and to reduce the post processing time, the calibration periods were run for 75 seconds per cylinder during the 2019 campaign to ensure that all instruments reached equilibrium. Target cylinders were run approximately every 1 hour of flight.

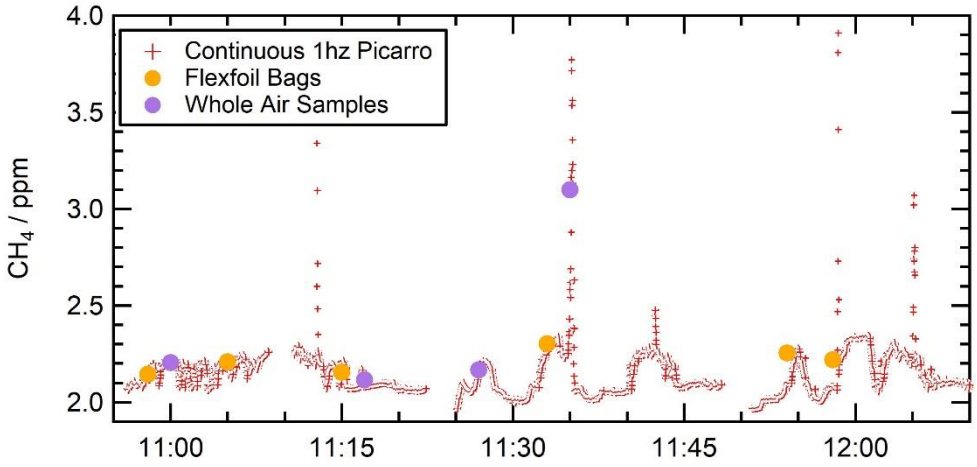
A3. SWAS Operation

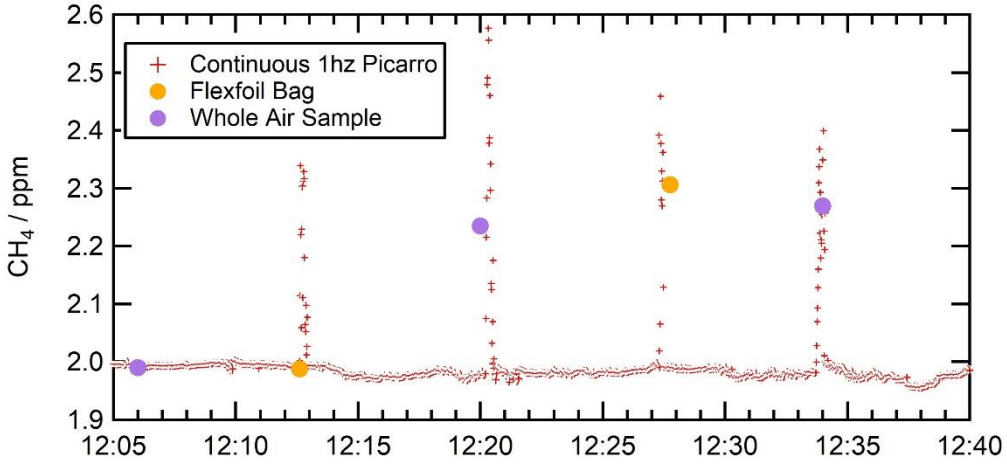
Each sample is compressed into the canisters using a modified metal bellows pump (Senior Aerospace 28823-7) capable of 150 SLPM open flow but filling the canisters at ~50 SLPM measured average integrated for ~6 and 9 seconds for the 1.4 L and 2 L canisters, respectively. Canister fill pressure is controlled electronically using a back-pressure controller (Alicat, PCR3), BPC. The BPC can maintain flow at any set point pressure (in general 40 psi), including the final fill pressure setpoint. This allows the 2 L flow through canisters to be filled, even before the operator activates the sampling, enabling air masses to be sampled through which the aircraft has already flown seconds earlier.

Bespoke software was created to allow control of the SWAS system wirelessly from any position in the aircraft using the Ethernet network. Bespoke software was also created for the analysis of the canisters once in the laboratory. The SWAS flown on the ~~1st~~2018 campaign (V1) was a prototype and was updated to the current final version (V2) to fulfil the requirements of the FAAM BAE146 and to address potential issues experienced with the prototype. V2 uses the same canisters and valves as V1 but differs slightly in the size of each case and the plumbing of gas lines. In V2, the canister and valve geometry was optimised to allow an elbow compression fitting between the valve and the canisters to be eliminated, with the valve mounted directly to the canister. This reduces the risk of leaks by 66%. The geometry also allowed the reduction in size by 1U rack unit, allowing more canisters to be fitted in the same space, improved control electronics and sample logging to ensure canister fill times were captured accurately and stored securely. V2 also saw the addition of ~~2L~~2 l flow-through canister cases to complement the 1.~~4L~~4 l to-vacuum canister cases. These allowed sample air to be flushed through the canister at a user defined pressure and makes capturing narrow plumes easier due to reduced sample line lag and fill time.

Figure A1. Photo of the rear-facing chemistry inlets on the BAS Twin Otter aircraft.

Figure A2. Photo of the BAS Twin Otter showing the turbulence boom protruding from the front of the aircraft superstructure.


Continuous 1hz Picarro
Flexfoil Bags
Whole Air Samples
CH4 / ppm
4.0
3.5
3.0
2.5
2.0
11:00
11:15
11:30
11:45
12:00


Continuous 1hz Picarro
Flexfoil Bag
Whole Air Sample
CH4 / ppm
2.6
2.5
2.4
2.3
2.2
2.1
2.0
1.9
12:05
12:10
12:15
12:20
12:25
12:30
12:35
12:40

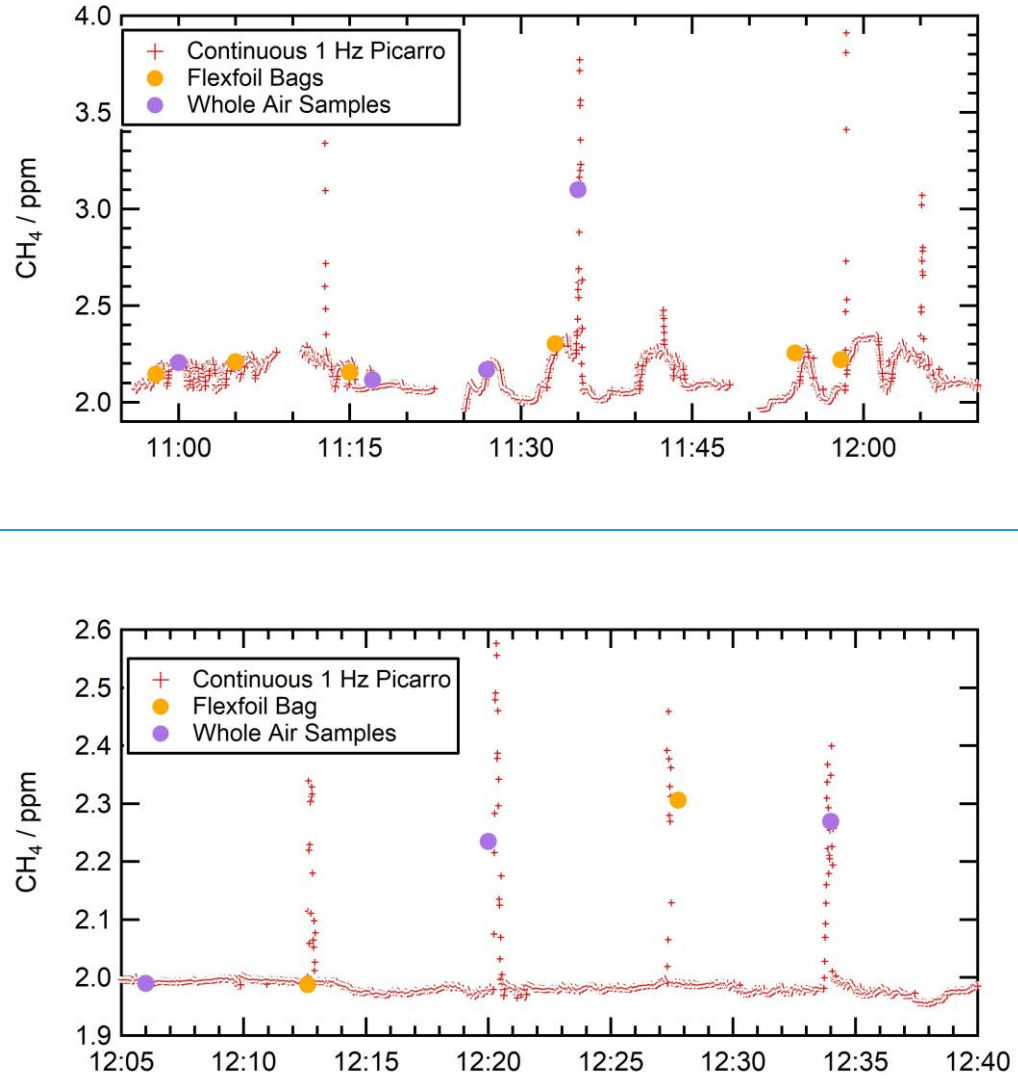


Figure A3. Examples from a 2018 flight (top panel) and a 2019 flight (lower panel) with attempted capture of CH_4 plumes in spot samples (both SWAS and Flexfoil bags). Note the improved ability to sample at the correct period to capture short-lived enhancement in both SWAS and Flexfoil samples for 2019 compared ~~for 2019~~ to 2018 thanks to flight planning and SWAS development improvements.

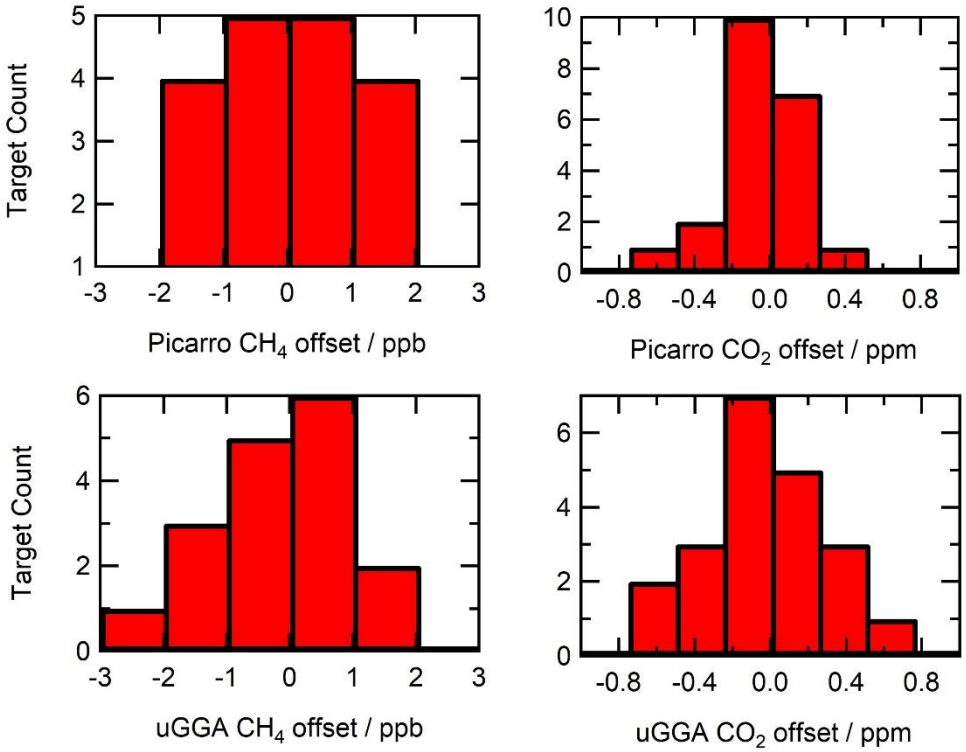


Figure A4. Target gas data from flights during 2018 for the Picarro ~~G2301~~G2311-f and Los Gatos uGGA instruments for both CO_2 and CH_4.

INSTRUMENT	MEASUREMENT SPECIES	T_{90} RESPONSE RATE	PRECISION OF PRIMARY SPECIES OF INTEREST
LGR UGGA	CH_4, CO_2	17 s	(CH_4) 1 ppb over 10 sec
PICARRO G2311-F	CH_4, CO_2	0.4 s	(CH_4) 1.2 ppb over 1 sec
LGR UMEA	C_2H_6, CH_4	17 s	*(C_2H_6) 17 ppb over 1 sec
TILDAS	C_2H_6	< 2 s	**(C_2H_6) 50 ppt over 10 sec

* measured in laboratory ** manufacturer's expected precision

Table A1. Response rates and precision for the instrument set-up on the BAS Twin Otter. All measurements were time-shifted to match the Picarro G2311-f for analysis.

Volatile organic compounds identified and quantified from SWAS samples

compound	detection limit (ppt)
ethane	4
ethene	4
propane	6
propene	2
iso-butane	1
n-butane	1
acetylene	1
trans-2-butene	2
but-1-ene	2
cis-2-butene	2
cyclopentane	2
iso-butene	2
iso-pentane	1
n-pentane	1
1,3-butadiene	2
trans-2-pentene	2
pent-1-ene	2
2,3-methylpentanes	2
n-hexane	2
isoprene	1
n-heptane	2
benzene	1
2,2,4-trimethylpentane	2
n-octane	2
toluene	1
ethylbenzene	2
m+p- xylenes	2
o-xylene	2

Table ~~A1~~A2. Summary of ~~VOC's~~VOCs measured from SWAS samples at York University.