

## In response to Anonymous Referee #2 comments from November the 9th, 2018.

*This paper describes the performances of an analyser of the major greenhouse gases in air on board of an aircraft. Details are provided on the analyser hardware, the analytical software and the calibration method, followed by an evaluation of the performances by comparison with other analysers present during the same flight. While the analyser itself is not new and was already described in a previous paper (McManus 2011), in this work the number of analysed compounds was extended, the in situ performances were looked at more deeply, and the calibration method was improved. The paper is generally well written, well-structured, clear, and provides lots of details on the instruments and methods. However the section on performance evaluation needs some more work, both in its content and format. I therefore recommend a minor revision before the paper can be published in AMT.*

### 10 Comments on the terminology

- Units to be written in plain (not italic) format - The format to display a value with its unit is "value-space-unit" for example "204m" on page 4 should be written "204 m"

- "mixing ratio" to be replaced by "amount fraction", expressed in mol mol<sup>-1</sup> (nmol mol<sup>-1</sup> for ppb,  $\mu$ mol mol<sup>-1</sup> for ppm).

- Names of molecules to be written in plain (not italic) format.

### 15 - Allan deviation seems to be confused with Allan variance. When values are reported in the same unit as the concentrations, this should be a deviation. Please check the correct usage over the document

Dear Referee,

Thank you very much for the detailed and very helpful comments and for the time spent on reading and reviewing this manuscript. We greatly appreciate it. The comments made on terminology were of great help and have been implemented in a revised version of the manuscript. We will directly follow up on the specific comments.

### Specific comments by section:

#### 1. Section 2.1: the text describes two sealed cells containing CH<sub>4</sub> and N<sub>2</sub>O. Where are they on Figure 1? Please indicate the purity of the gas and its pressure.

25 There is only a single sealed cell containing CH<sub>4</sub> and N<sub>2</sub>O. Its position has been marked in a revised version of Fig. 1. The gas inside the cell has an approximate pressure of 3500 Pa. The gas does not need to be pure. As the laser scans over the absorption features of CH<sub>4</sub> and N<sub>2</sub>O the laser can be spectrally referenced to the relevant molecular absorption lines, which is the single purpose of the sealed cell.

#### 2. Section 2.3: please provide more information on the calibration mixtures. In particular NOAA standards are all identified within NOAA database and you could just provide their reference to allow the users looking at all values measured by NOAA. At least please indicate the nominal amount fractions, their uncertainties, and the isotopic composition for CO<sub>2</sub>. This last value is of importance as you noticed a bias between the CO<sub>2</sub> amount fractions measured with your instrument and those measured by the PICARRO.

30 We included the requested details on the used NOAA standards for CH<sub>4</sub> and CO<sub>2</sub>. However, we have to note, that we used working standards of synthetic nature from *Air Liquide* due to the large amount of needed calibration gas. Usually these are produced with CO<sub>2</sub> from natural gas & oil combustion processes. We determined the CH<sub>4</sub> and CO<sub>2</sub> values of each working standard gas cylinder using a Picarro G-1301m. This has the drawback that we do not know the isotopic composition of our working standards. The reason why we did not send our working standards to a central lab is because the influence of the isotopic composition had been considered negligible at this stage (Chen et al., 2010). It was only in late summer 2018, that we found out (during development of JFIT) that the instrument was using a <sup>13</sup>C<sup>16</sup>O<sub>2</sub> line to derive ambient CO<sub>2</sub>. We assume the large bias originating from differences in isotopic composition in our working standards relative to the natural terrestrial abundances.

40 "[...] The cylinders have been cross-calibrated against NOAA standards and are thus traceable to World Meteorological Organization (WMO) standards for CH<sub>4</sub> (Cert.-Nr. CB11361, WMO X2004A for CH<sub>4</sub> (Dlugokencky et al., 2005)) [...]"

45

5 "[...] In this study we used working standards of synthetic nature from *Air Liquide*. Usually these are produced with CO<sub>2</sub> from natural gas & oil combustion processes. We determined the CH<sub>4</sub> and CO<sub>2</sub> values of each working standard gas cylinder using a NOAA-anchored (Cert.-Nr. CB11361) Picarro G-1301m. This has the drawback that we do not know the isotopic composition of our working standards as its impact had been considered negligible, e.g. (Chen et al., 2010). We only found out during development of JFIT, that the instrument is using a <sup>13</sup>C<sup>16</sup>O<sub>2</sub> line to derive ambient CO<sub>2</sub>. We estimate the required isotopic composition of such a CO<sub>2</sub> to be 98.447% primary isotopologue and 1.079% secondary isotopologue or  $\delta^{13}\text{C} = -19.6\text{‰}$  which seems reasonable according to B. Coplen et al. (2002). Since we are reporting retrieved mole fractions relative to the WMO scale, only the working standard reproducibility contributes to the total uncertainty of CH<sub>4</sub>. Uncertainty on CO<sub>2</sub> is difficult to assess here because of the unknown isotopic composition in our working standards. [...]"

- 10
3. **Section 4, ground-based performance: the reported Allan deviations seem a bit large. Compared to McManus 2011 on CO<sub>2</sub> for example, a factor 10 is noted. Please consider revising the statement that “values are in good agreement with the values reported by Aerodyne” and/or provides further support. Is there an effect of the calibration system described in 2.3, which is said to be used to check the stability and the linearity?**

15 The lasers frequency reported in McManus 2011 differs from the emitted frequency reported in this publication. Therefore it is not possible to directly compare with the values reported in McManus 2011. Here we are referring to the specifications reported by Aerodyne Inc. for this particular instrument. Nevertheless we removed this sentence as it is not necessary at this point. Although the calibration system has a strong influence on the instruments accuracy, we see no variation in precision after carefully looking at signal changes before and after calibration versus similar intervals between calibrations.

- 20
4. **Section 5: while the traceability of measurements with the QCL is clear (calibration with NOAA standards), nothing is indicated regarding the PICARRO. This is needed to fully understand the origin of biases. It seems that an anchored to NOAA is assumed, but this deserves further details (which standards? How many calibration steps? Isotopic composition?). When both instruments are compared, it would be more useful to express the difference in amount fraction, both in the text and in the graphs. This should then be compared with their uncertainties, not taking into account common sources of uncertainties such as NOAA uncertainty if all amount fractions are expressed on the same scale. Going beyond this, some consideration on how this compares with the Data Quality Objectives set by WMO would be of interest. The treatment of the constant bias found between the AERODYNE and the PICARRO analysers needs to be improved. Are both analysers calibrated directly with NOAA standards? How different are the calibration gases? It would be valuable to estimate the bias one could expect from the isotopic difference, as done for example in the paper of Chen et al. (Atmos. Meas. Tech., 3, 375–386, 2010), and compare with the observed bias. Indeed, an observed bias of 10  $\mu\text{mol mol}^{-1}$  seems very large.**

25 The PICARRO instrument is anchored to NOAA. CO<sub>2</sub> is WMO X2007, CH<sub>4</sub> is WMO X2004A, CO is WMO X2014A. It is calibrated hourly during flights using a fixed standard and weekly using a three-point calibration with high, low and target calibration standards. The corresponding references have been implemented in a revised version of the manuscript. Concerning the second part of this comment: it is important to know the dynamic range that is covered when looking at differences between instruments. Vanishing differences at vanishing dynamic range do not tell the whole story about instrument performance. We included both, the differences (as histograms in Fig. 13) and the absolute values (dynamic range) in Fig. 10 and 11. The origin of the biases is not yet fully understood. It was suggested that water vapor correction could have an impact on this. The reason for this assumption is that the calibration standards are always dry, whereas sampled air is not dried before entering the sample cell. Correlation plots however show no significant influence of water vapor on the residuals between the dry-air-sampling Picarro and the QCLS. It is therefore very unlikely that the water vapor correction is the source of the large bias in CO<sub>2</sub>. Instead we identified the difference in isotopic composition of the calibration standard versus sampled atmospheric air as the most probable cause. Chen et al. (Atmos. Meas. Tech., 3, 375–386, 2010) estimated the influence for a Picarro greenhouse gas analyzer measuring the primary CO<sub>2</sub> isotopologue. It is commonly assumed that the influence of isotopic composition is on the order of 0.1 ppm. Using the <sup>13</sup>C<sup>16</sup>O<sub>2</sub> line at 2227.604 cm<sup>-1</sup> via HITRAN-based direct absorption spectroscopy, we estimate a much larger influence, that could well explain the bias encountered (see above). We estimated the required isotopic composition of such a CO<sub>2</sub> to be 98.447%

primary isotopologue and 1.079% secondary isotopologue ( $^{13}\text{C}^{16}\text{O}_2$ ). We included this in the text. It is therefore one of the major findings of this study, that knowledge on isotopic composition of the calibration standards is of paramount importance when using the mentioned absorption line.

"[...] In situ  $\text{CH}_4$ ,  $\text{CO}_2$ , and  $\text{CO}$  were measured using a PICARRO G2401-m cavity ring-down spectrometer, and in situ  $\text{CO}_2$ ,  $\text{CH}_4$ , and  $\text{H}_2\text{O}(g)$  were measured using a PICARRO G2301-m cavity ring-down analyzer. Both PICARRO instruments are anchored to WMO X2007 for  $\text{CO}_2$  (Zhao and Tans, 2006), WMO X2004A for  $\text{CH}_4$  (Dlugokencky et al., 2005) and WMO X2014A for  $\text{CO}$  (Baer et al., 2002). [...]"

"[...] In this study we used working standards of synthetic nature from *Air Liquide*. Usually these are produced with  $\text{CO}_2$  from natural gas & oil combustion processes. We determined the  $\text{CH}_4$  and  $\text{CO}_2$  values of each working standard gas cylinder using a NOAA-anchored (Cert.-Nr. CB11361) Picarro G-1301m. This has the drawback that we do not know the isotopic composition of our working standards as its impact had been considered negligible, e.g. (Chen et al., 2010). We only found out during development of JFIT, that the instrument is using a  $^{13}\text{C}^{16}\text{O}_2$  line to derive ambient  $\text{CO}_2$ . We estimate the required isotopic composition of such a  $\text{CO}_2$  to be 98.447% primary isotopologue and 1.079% secondary isotopologue or  $\delta^{13}\text{C} = -19.6\text{‰}$  which seems reasonable according to B. Coplen et al. (2002). Since we are reporting retrieved mole fractions relative to the WMO scale, only the working standard reproducibility contributes to the total uncertainty of  $\text{CH}_4$ . Uncertainty on  $\text{CO}_2$  is difficult to assess here because of the unknown isotopic composition in our working standards. [...]"

5. **Section 5, uncertainties: it is not so common to see combined uncertainties considered in such measurements, and the effort of the authors is certainly valuable. However some consideration on how these values compare with other instruments would be required. Is the calibration procedure specific to this instrument? Does this imply a larger uncertainty than for others? Would you say this instrument has comparable precisions than others?**

The calibration procedure described herein is not instrument specific. It could be applied to other in situ instruments as well. Regarding the accuracy involved, there is always a trade-off between measurement time and accuracy: Increasing the number of calibration cycles improves achievable accuracy at the cost of observation time. It does not imply a larger uncertainty than others, as we do not use the online calibration mixing with the MFCs when taking data (Online mixing would add the uncertainty on the mass flow controllers on top). The instrument described herein is unique in that it offers many simultaneously observed species. It may be possible to find instruments showing better precision figures measuring a single or two species but we seriously doubt, that any other instrument with those many species sampled simultaneously will show better precision figures. Furthermore, as described in the text we estimate the uncertainty on calibration sequence evaluation with  $2\sigma$ , which is again a worst-case assumption. Unfortunately we had a numerical error in the first version of the manuscript and the values listed in Tab. 3 were not double the precision. We corrected this in the revised version of the manuscript. We further included a short sentence on precision comparison with the available PICARRO instrument:

"[...] Precision (uncertainty) figures given in Tab. 3 can be compared to  $2s-1\sigma$  PICARRO G2401-m airborne precision (uncertainty) estimates based on ambient measurements at stable conditions of 0.3 (2) ppb, 0.02 (0.1) ppm and 2.0 (5) ppb for  $\text{CH}_4$ ,  $\text{CO}_2$  and  $\text{CO}$ , respectively. [...]"

6. **Section 5, discussion on instruments precisions: Allan deviations (not variance) were measured before the flight and during the flight. It is not very clear how those values compare. One would expect the lowest values during ground-based measurements, presumably recorded on gas mixtures with constant flow rate and pressure. During the flight, other sources of instabilities can increase the noise of the instrument. However some of the values appear to be lower during the flight (above the planetary boundary layer only). This would need some further explanation.**

Allan variances were not measured during flights. The in-flight precision values are instead based on ambient measurements at stable conditions. Ambient variability can thus not be completely ruled out. Meaning, we are looking at the worst case scenario here. Thank you for pointing us towards this mistake on the ground-based precision values. The values reported for ground based operation were based on an older version of the retrieval software. We have corrected this in a revised version of the manuscript.

"[...] Typical in-flight precision figures based on ambient measurements at stable conditions for both regimes [...]"

7. **Conclusions: the advantages and drawbacks of the aerodyne instrument could be better highlighted. The large number of species analysed together is certainly an interesting feature, but it seems to come with increased noise compared to CRDS analysers. Is that really the case or is this a wrong impression coming from an increased in-flight noise which could impact other analysers as well?**

Here, we do not want to compare the two instruments against each other. Instead our goal is to demonstrate the suitability of the described instrument, given the calibration approach and post-processing described herein, for airborne observation with the ultimate goal of inferring local to regional fluxes. The instrument has advantages and drawbacks when directly compared to CRDS analyzers. One of the drawbacks is the reduced absorption path length and the resulting lower precision, as well as the large amount of calibration gas necessary for 10 % of the measurement time. A big advantage is the simultaneous measurement of all targeted species. There is practically no dead time in between measurements, which is especially useful in close vicinity to sources and/or for young weakly dispersed (spatially narrow) plumes. This instrument sees everything, while there is a certain chance with sequentially probing instruments of missing a narrow plume or not getting the peaks right. This is further described in Sect. 5. The large number of observed species is another big advantage that can be used for source attribution.

#### Line-by-line comments:

1. **Page 3, Line 5: you may clarify that “DLR” in the title is the name of the laboratory owning the spectrometer.**

The relevant sentence has been changed to include a definition of DLR:

"[...] The spectrometer system used here builds upon the *Dual Laser Trace Gas Monitor*, a commercial tunable IR laser diode absorption spectrometer (TILDAS) available from *AERODYNE RESEARCH INC., Billerica, USA*, acquired by *Deutsches Zentrum für Luft- und Raumfahrt (DLR)* in late 2016. [...]"

2. **Page 4, Line 25: you may keep SLPM for the flow rate, but indicate the value in mL min<sup>-1</sup> as well**

We don't really see the benefit of reporting flow rates in mL min<sup>-1</sup>, but 23 SLPM would yield 23000 mL/min at standard conditions ( $p = 101325 \text{ Pa}$ ,  $T = 273.15 \text{ K}$ ). We thought about converting SLPM to SI units  $1 \text{ SLPM} = 1.68875 \frac{\text{Pa m}^3}{\text{s}}$  but we omitted this, because we assumed SLPM to be a commonly used unit for in situ measurements.

3. **Page 5, Line 19: why the use of “cross-calibrated” rather than “calibrated”? Does it involve a particular method?**

Here, we want to express the fact, that we calibrate our working standards using a Picarro G1301-m to NOAA standards as described above. This is what we refer to with "cross-calibration".

4. **Page 5, Line 32: the entire sentence may be rewritten to express more clearly that no dilution was introduced at this stage, which is why you do not need to take into account an uncertainty on the flow rate measurements.**

The relevant text portion has been rephrased to:

"[...] The online mixing feature is not used for in-flight calibration. Hence, no dilution of the calibration standard with zero air is introduced during flights and the uncertainty on the flow rate measurements can be omitted. Online mixing (relevant for linearity checks) adds the uncertainty of the controlled mass flow on top of the gas cylinder uncertainties. [...]"

5. **Page 10, Line 14-15: what is meant by “not accurately constrained”? There is certainly an issue with the difference in isotopic composition between the sample and the calibration gas, and this aspect deserves a better treatment in the paper. However at this point you are describing the fit of the spectra, and the statement about constraining the isotopic composition of the sample is unclear. Does this mean constraining the fit? The fit window?**

We agree that this information is not needed at this point for describing the spectral fit. It is dealt with in Section 5. We thus removed the complete sentence.

6. **Page 12, Line 14 “excluding absolute error”. Do you mean uncalibrated or expressing the precision only? Line 15: “values reported by Aerodyne”. Which paper? McManus 2011?**

Here, we state the 1-sigma precision and measurement frequency only. The absolute error stated by the pressure transducers manufacturer is stated with  $\pm 0.5$  hPa. The relevant text portion has been rephrased to:

"[...] The volumetric flow rate stabilized at 23SLPM for a sample cell pressure regulated at  $50.0 \pm 0.5$  hPa (0.2 hPa precision @ 5 Hz). [...]"

- 5 7. **Page 18, Line 2-3: “we were not able to reproduce...” seems a rather negative introduction for a positive result, as everything was made to be insensitive to the cabin pressure. Consider rephrasing.**

The relevant text portion has been rephrased to:

10 "[...] A severe cabin pressure dependence in excess of  $0.3 \text{ ppb hPa}^{-1}$  in  $\text{CH}_4$  mixing ratio has been previously reported for airborne TILDAS instrumentation (Pitt et al., 2016). This instrumentation however physically differs from the one reported here. It is not possible to accurately compare the dependencies of one instrument relative to another since many factors/quantities involved are instrument-specific, e.g. the open-path length, the positioning and properties of optical elements, like windows and mirrors, the stiffness and thermal expansion coefficients of employed optical stands, etc.. We were nevertheless able to effectively minimize cabin pressure dependencies during operation of the QCLS instrument aboard the C130 using the calibration strategy from Sect. 2.3. [...]"

15 **Comments on figures:**

1. **Figure 7 it is not clear if the amount fractions are provided after calibration or not. The legend seems to indicate calibrated values, but the y-axis in the right plot indicates "Methane RAW [ppt]" which would mean raw values before calibration. Please clarify.**

20 The depicted methane amount fractions are indeed raw signals before calibration. A synthetic calibration gas has been mixed from zero and calibration gases using the described calibration system, in order to verify the linearity of retrieved amount fractions.

2. **Figures 10 and 11: it is too uneasy to compare both analysers on the plots. Differences would be more interesting, as the paper does not include any consideration on the amount fractions of the gases.**

25 It is important to know the dynamic range that is covered when looking at differences between instruments. Vanishing differences at vanishing dynamic range do not tell the whole story about instrument performance. We included both, the differences (as histograms in Fig. 13) and the absolute values (dynamic range) in Fig. 10 and 11.

3. **Figure 12: y-axis of the right plot is the methane amount fraction. Use a symbol and unit such as " $x_{\text{CH}_4}$  / (nmol mol<sup>-1</sup>)" and indicate in the legend " $x_{\text{CH}_4}$  is the methane amount fraction".**

$x_{\text{CH}_4}$  is commonly used for total column measurements. We therefore refrain from changing the axis label here.