We thank referee #1 for the useful comments to improve the manuscript. We have gone through each comment carefully and answer them below.

This paper by Laitinen et al. presents two important datasets of long-term atmospheric CO2 and CH4 observations in northern Finland. Although they follow the same general guidelines defined by the WMO for in-situ GHG observations, including the adoption of the same calibration scale, these datasets were produced at the Pallas site using two completely independent observing systems (i.e. different sampling systems, different calibration standards, different data flagging procedures). The manuscript mainly focuses on the comparison of the two datasets (one produced within the GAW/WMO framework and the other within the ICOS-RI) to show the high consistency between them and to briefly discuss the reasons for the small observed deviations.

The longer CO2 and CH4 dataset (i.e. the GAW/WMO one) was also compared with the widely used NOAA MBL data product to put it into a global context.

The topic of comparison and attribution of deviations in contiguous measurement sites is an important and, I would say, emerging issue (see e.g. https://doi.org/10.1088/1748-9326/abe74a, https://doi.org/10.5194/amt-11-1599-2018, https://doi.org/10.5194/amt-16-2399-2023), so I appreciate the appearance of this paper.

Having said that. I recommend publication and my comments can be seen as mostly minor or technical suggestions/recommendations.

Specific comments

Throughout the paper: I think the nomenclature "mixing ratio" should be changed to "mole fraction" as recommended by WMO.

Changed the use of ”mixing ratio” to ”mole fraction”

An interesting point is that the GAW and ICOS data sets are produced using two different sampling inlets. I suggest that the author also add a figure or map to Figure 1 showing the different locations of the two inlets.

Added a schematic figure A1 to Appendix to show the spatial distances of the two inlets.

Line 86. “Due to its remote…representative for unpolluted air”. Is there a reference to quote?

Citation **(Hatakka et al, 2003)** added

Line 88. For the reader not familiar with ICOS, can you briefly explain the difference between the different classes of ICOS stations (Class 1 vs. Class 2)?

Added further explanation in the Introduction chapter:

Within the ICOS ATC, the stations are classified into Class 1 and Class 2 stations. The requirements for the Class 2 stations are continuous measurements of CO_2 and CH_4 complemented by basic meteorological parameters: Air temperature, relative humidity, wind speed and direction as well as atmospheric pressure. The Class 1 stations are required, in addition to the requirements of the Class 2, to have continuous CO and boundary layer height measurements as well as to operate the ICOS flask sampler (described more in detail in Levin et al, 2020). Furthermore, the stations are classified to three types based on their location: continental stations targeting mainly continental air-masses, coastal stations targeting mainly

marine air-masses and mountain stations targeting mainly free troposphere during night (ICOS RI, 2020).

Section 2.2: For the ICOS instrument, you can report the instrumental uncertainty as indicated by the continuous measurement repeatability and the short - long term repeatability. Are these (or similar) indications available for the GAW instrument?

Removed Figure 4 showing the LT of ICOS instrument and moved the information to Figures 7 and 13 with target cylinders of GAW instument as well, where also the ST of the ICOS instrument is presented

Line 120: How did you obtain the correction factors to correct for residual water vapour interference? Did you do the "droplet test"? How often? This should also be reported for the GAW instrument (line 150).

Added explanation for the water vapor correction assesment to sections 2.2.1 and 2.2.2

For the ICOS analyzer, the correction coefficients are determined by the ATC during the initial instrument test by first measuring a dry gas stream from a cylinder, and then humidifying the stream for 20 minutes a step with 0.25 v% steps from 0.5 v% to 2v %, 2.5v % and 3 v%. The coefficients for CO_2 and CH_4 are then determined with the following equation:

$C_w/C_d = 1 + aH + bH^2$ (1)

Where C_w is the measured wet mole fraction, C_d is the dry mole fraction (measured when H = 0), H is the measured water vapor concentration and a and b are the correction factors.

And

Similarly to the ICOS analyzer, the instrument specific water vapor correction factors are determined as well. The approach used by the FMI is similar to that of ATC; a dry gas stream is humidified using a self-build instrument, ranging from 0 to 3.5% (Aaltonen et al, 2016). The coefficient are then calculated using Eq 1.

And

These coefficients are evaluated during the ICOS audit, as well as approximately once per year by the station PI, and updated if deemed necessary by the ATC.

Section 2.2.3. I agree with the comments by Tonatiuh Guillermo Nuñez Ramire that flask results should be shown, especially the flask vs. ambient air comparison. Or at least the authors should refer to them in some way.

This is a good comment, and we agree that the flask data inclusion could bring some additional insights. However, for the scope of this paper we wanted to focus on the CRDS-based systems and their comparison and feel like adding the flask data to the manuscript would widen the scope of the paper too much. Moreover, the flask data is unlikely to reveal additional information about the offsets, considering their small magnitudes. Instead, we decided to leave section discussing the flask measurements out of the paper completely.

Removed Flask sampling section, moved "Auxiliary measurements" to 2.3

Line 186: "...last and first days of the year". Can you be more precise?

Added further clarification:

... by taking the difference of the values of the **last day (31.12) and the first day (01.01) of the given year. For example, the GR of 2020 would be calculated by taking the difference of the daily trend values of 31.12.2020 and 01.01.2020.**

Figure 6/12: Since there is an apparent impact of the implementation of Nafion on the ICOS instrument, have you tried to analyse DCO2 (and DCH4) as a function of ambient water vapour values to better quantify possible impacts? This would also help to better attribute the differences for CH4 highlighted in lines 336 - 368 .

Added to section 4.1:

There exists a negative correlation between DCO_2 and the water vapor concentration of about -0.06 ppm/v% (intercept 0.05 v%, $p < 0.05$) when both of the instruments measure wet air.

And to section 4.2:

As for CO_2, there exists a negative correlation between DCH_4 and the water vapor concentration of about -0.94 ppb/v% (intercept 1.34 v%, $p < 0.05$) when both of the instruments measure wet air.

Lines 290 - 292: A "large" spread is also visible for the comparison of ICOS and GAW instruments. As stated by the authors the observed differences are mostly within the WMO compatibility goal but I'm wondering what happen if you subset the data as done for the trend analyses (i.e. only keeping data with standard deviation less than 0.5 ppm). This would in some way minimize the impact of the different inlet locations by catching conditions characterized by low CO2 variability.

For the long-term comparison between the ICOS and GAW instruments, the same wind speed and STD based filtering has been applied as for the trend analysis. This is explained in section 3.3.

Lines 298-305 and 362 - 366: These paragraphs have been added to place the results of the intercomparison between ICOS and the GAW instrument at Pallas in a broader context. It would be interesting to add information from the other intercomparisons carried out in other WCC-EMPA audits.

Added information on other audits to the end of chapters 4.1.3

In (Zellweger et al., 2016) an overview of results of GAW audits performed in Danum Valley (DMV), Cape Verde (CPVO), Mace Head (MHD) as well as an earlier audit at Pallas, are presented. The audits were performed in 2012 (PAL & CBO) and 2013 (DMV & MHD). The comparison for CO2 was made in PAL, DMV and CVO. The measurement methods were non-dispersive infrared (NDIR) in PAL and DMV, and off-axis integrated cavity output spectroscopy (OA-ICOS) in CVO, compared against the traveling CRDS instrument. Results of the audits show a median deviation (1h aggregation, ± standard deviation) of 0.08 ± 0.03 ppm at PAL, 0.03 ± 0.21 ppm at DMV, 0.06 ppm ± 0.06 at CVO. For Pallas, a comparison of the local CRDS against the traveling CRDS is presented in (Rella et al., 2013). For CO2, the mean deviation was -0.025 ± 0.034 ppm.

And 4.2.3

Earlier GAW audits in CVO and MHD for CH4 are presented in (Zellweger et al., 2016). In CVO the same instrument (OA-ICOS) was used as for CO2. In MHD the measurement method was GC. The results show a mean deviation (± standard deviation) of -0.61 ± 0.32 ppb at CVO and 0.22 ± 3.59 ppb and MHD. Comparison of the traveling CRDS against a local CRDS at Pallas in 2012 is presented in (Rella et al., 2013), with a mean deviation of -0.032 ± 0.367 ppb

Line 336-338: How can you say "better" performance? Can it not be "different performance"? Or do you have some quantitative test to show that the water vapour correction on the GAW analyser is more accurate?

We mainly base this claim on the fact that the agreement between the two instruments is improved when the sample air of the ICOS analyzer is dried. Since no change is made for the GAW analyzer, this indicates that the poorer agreement prior is caused by the water correction of the ICOS analyzer, while the GAW analyzer performs well even against the dried ICOS analyzer.

Line 388 – 391: "The largest differences were observed between the GAW and WCC systems in both CO2 and CH4, however the largest spread (CI range) in the differences was between ICOS and GAW for CO2 and between GAW and the ICOS Mobile Lab for CH4. This is expected, as the GAW system is the only one measuring from its own inlet and all the other systems are connected to the same inlet". A comparable spread for CO2 (Table 2: 0.17 ppm) was also observed for WCC and GAW vs. ICOS Mobile Lab.

We have added a remark of this:

The largest differences were observed between the GAW and WCC systems in both CO_2 and CH_4, however the largest spread (CI range) in the differences was between ICOS and GAW **and between WCC and ICOS Mobile Lab** for CO_2 and between GAW and the ICOS Mobile Lab for CH_4. This is **partly expected**, as the GAW system is the only one measuring from its own inlet

Line 394: "This is likely due to the filtering of the data based on the wind speed, assuring well mixed air". I think you should apply this filtering to the audit as well and see if the deviations decrease (see also my previous comment about lines 290 – 292).

Tested the filtering of the audit data for windspeed; we generally notice a decrease in the spreads of the deviations. Added discussion to sections 4.1

When the data are filtered for wind speed and hourly standard deviation, as done for the analysis for trend and long-term differences, the spreads between GAW and ICOS Mobile Laboratory decreases to 0.14 ppm, between GAW and WCC to 0.09 ppm and between ICOS and GAW to 0.12 ppm, indicating that the different inlet location affects the comparison when the air is not well mixed. The difference between WCC and ICOS Mobile Laboratory decreases as well to 0.15 ppm, but decrease is observed for comparison of ICOS and WCC.

4.2

Filtering the audit data for $CH_4$for wind speed and hourly standard deviation leads to similar results as for CO_2, decreasing the spread between GAW and ICOS Mobile Laboratory to 1.40 ppb, between ICOS and WCC to 0.88 ppb and between ICOS and GAW to 1.29 ppb. Between WCC and ICOS Mobile Laboratory the spread is only reduced to 1.31 ppb and between GAW

and WCC to 0.98 ppb, and no significant difference in the spread is noticed between ICOS and ICOS Mobile Laboratory.

and 5.

When filtering the audit data for the wind speed and hourly standard deviation as well, the spreads between the GAW instrument and the rest generally decrease.

Technical comments

Data citation: the authors only provide data availability at the end of the manuscript. I recommend that the analysed dataset can be fully cited in the main manuscript, since doi are available for both GAW and ICOS datasets.

Added citation for GAW and ICOS datasets to the main manuscript

Line 255, but also in the rest of the manuscript: I would suggest that CI can be expressed as e.g. [-8.0, 11.8] ppm.

Changed to [-8.0, 11.8] ppm, and similarly in the rest of the manuscript

Line 262: "(7 bottom) -> (Fig. 7 bottom). The same in line 264. In addition, I suggest that Figure 7 can be moved to Figure 6, as it is referenced earlier in the manuscript.

As suggested by other reviewer, changed the sentence and referenced to (Fig 7 (a)), and changed accordingly.

Swapped figures 6 and 7.

Line 271: "before"

changed to "before" without capitalization

Line 291: I would suggest to be consistent and continue to use "ICOS Mobile Laboratory" or "ICOS Mobile Lab".

Changed "Icos Mobile Lab" to "ICOS Mobile Laboratory"

I suggest moving Figure 9 and Figure 14 to Supplementary Material.

Moved to supplementary

Sections 4.1 and 4.2: I would split these two sections into two/tree sub-sections as the following topics are considered: 1) Comparison of Pallas measurements with NOAA PBL dataset, 2) Long-term comparison of GAW and ICOS measurements at Pallas, 3) Combined GAW and ICOS audit.

Sections 4.1 and 4.2 split into three sub-sections

Figures 7 and 11: Please describe in the caption what the error bars represent.

Added description: […] **points are hourly (a) and monthly (b) means with associated standard deviations**

Line 373: "global development" -> "global tendencies"

Changed as suggested

Added figures:

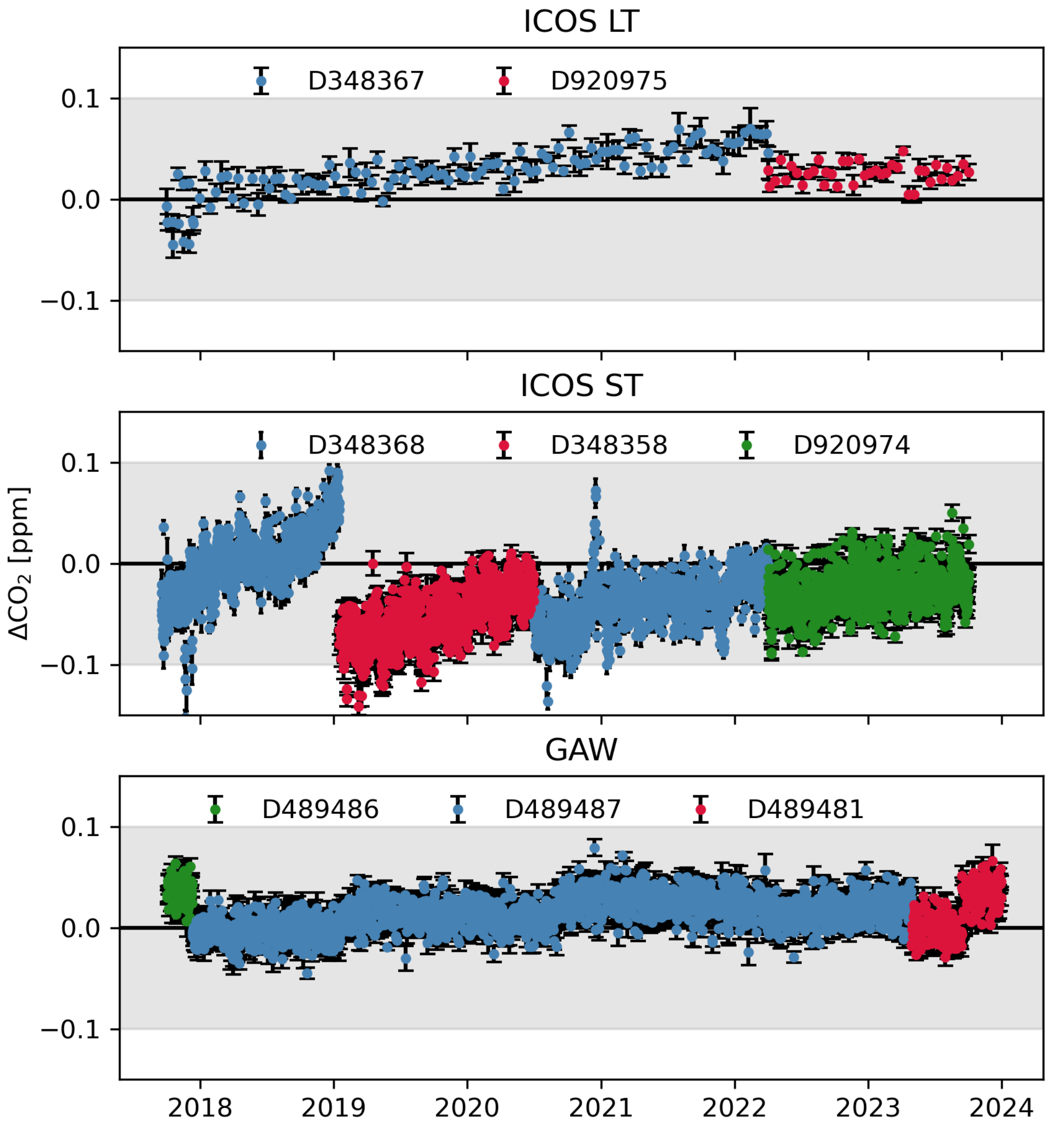

Figure 7. CO2 target measurements of ICOS LT, ICOS ST and GAW target cylinders over the whole comparison period. Different cylinders used are marked with distinct colors. The data are given as means of each sequence with the associated standard deviation.

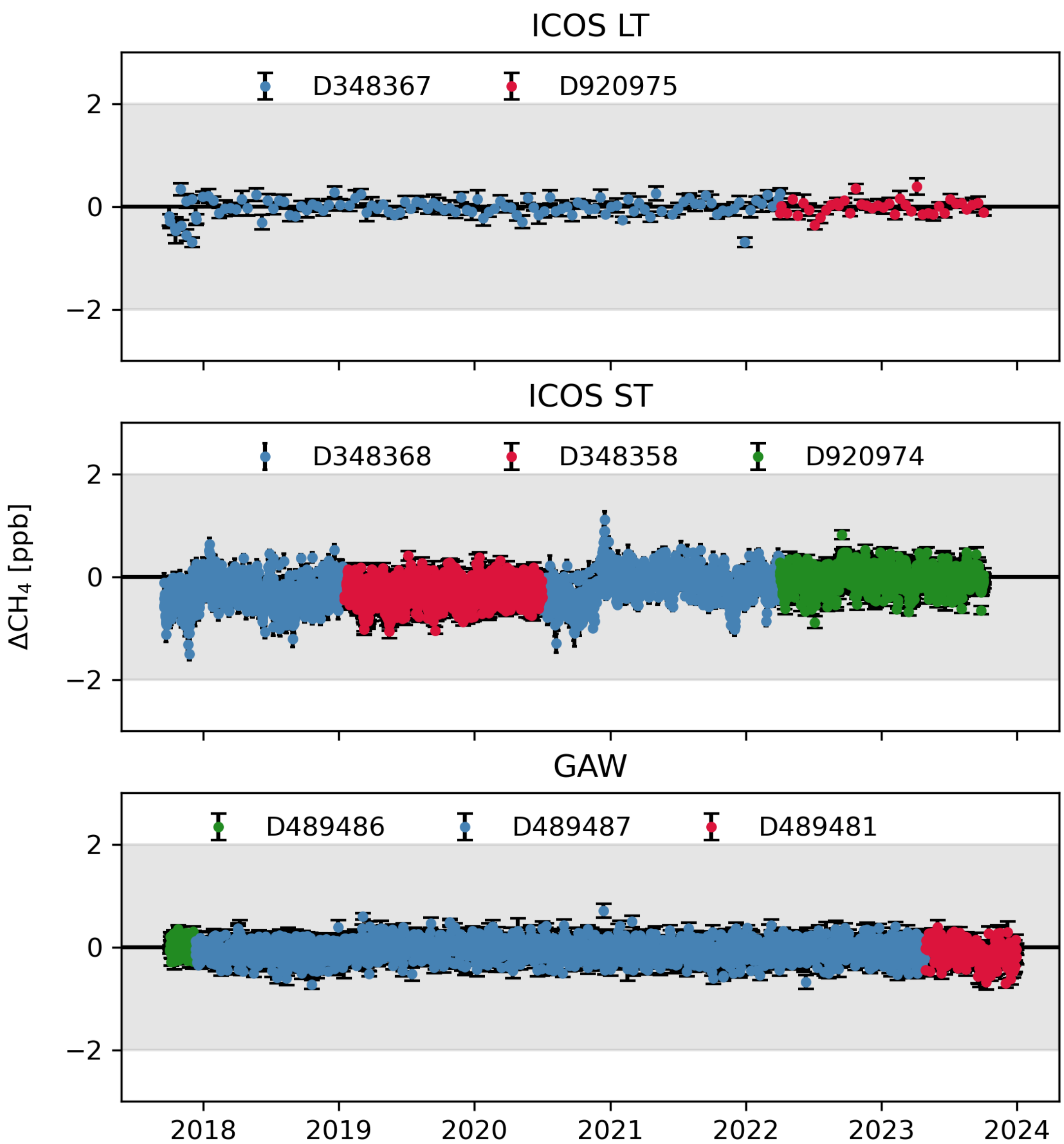

Figure 13. CH4 target measurements of ICOS LT, ICOS ST and GAW target cylinders over the whole comparison period. Different cylinders used are marked with distinct colors. The data are given as means of each sequence with the associated standard deviation

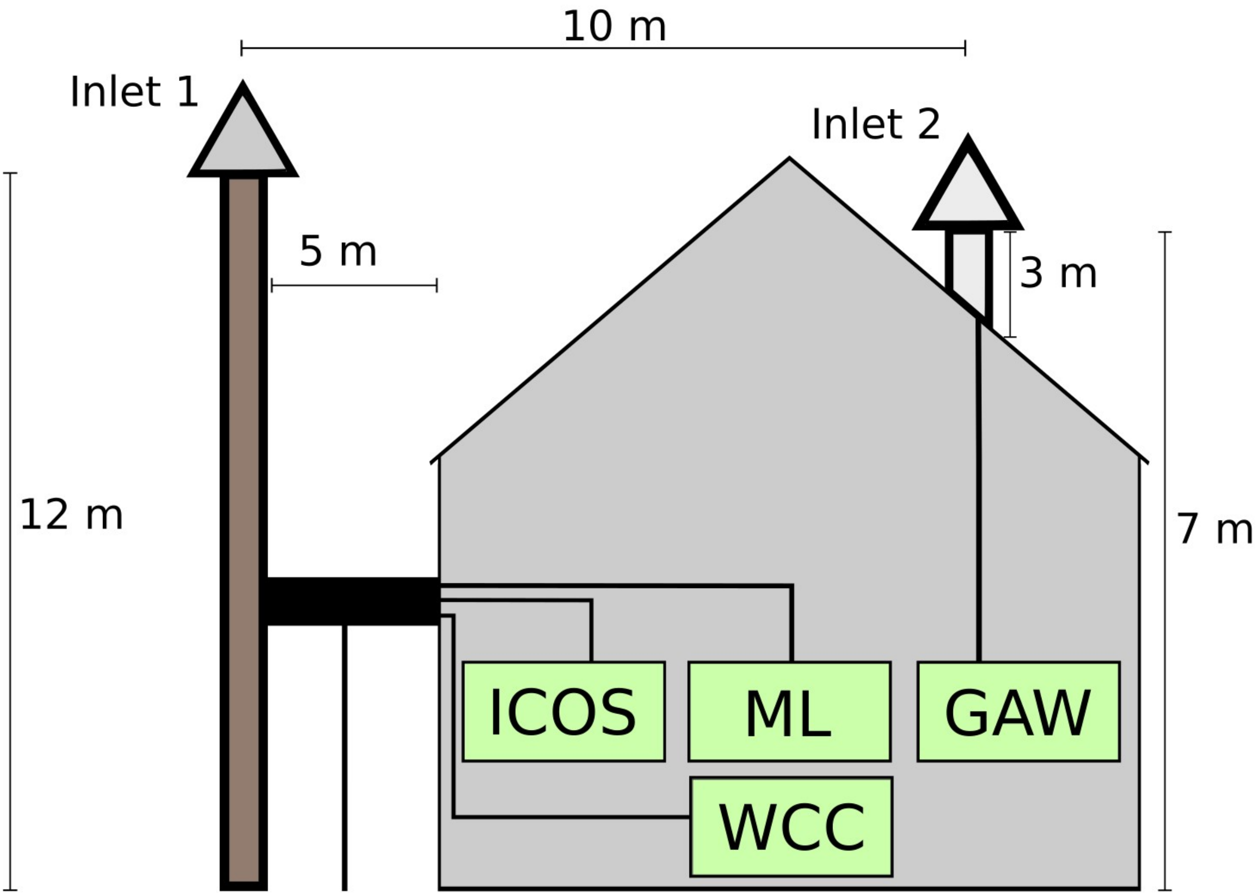

Figure A1. Schematics of the measurement setup during the audit at Pallas. ML refers to the ICOS Mobile Laboratory instrument