Dear editor,

We would like to thank the two reviewers for their comments. We have been working toward a new version of the manuscript taking their respective comments into account. We include and number the reviewers' comments in black. The comments from Reviewer 1 are numbered from 1 to 14, and those from Reviewer 2 from 15 to 36. The referees' comments have been addressed individually, as requested by the journal. When both reviewers address the same topic, we indicate the corresponding reference number(s).
Our responses are in blue, and the modifications to the manuscript in red in this response file.

On behalf of all the co-authors,
Agnese Petteni

Reviewer #1 (Remarks to the Author):

Review of "An International Intercomparison of Continuous Flow Analysis (CFA) Systems for High Resolution Water Isotope Measurements in Ice Cores" by Agnese Petteni et al.

General comments:
The submitted manuscript presents an intercomparison of coupled CFA-CRDS systems for high resolution water isotope analysis in ice cores, involving two French laboratories (LSCE, IGE) and one Italian laboratory (ISP-UNIVE). The CFA is now an established technique to analyze ice cores, but has inherent mixing limits. Thus, the use of the power spectral density (PSD) approach to quantify the analysis resolution is of strong interest for the climate reconstruction based on ice cores. Moreover, the study provides practical recommendations to optimize the CFA-CRDS setups, making this work relevant for colleagues in ice core analysis. I would like to see this work published, but only after taking care of the following revisions:

Several parts of the manuscript shall be re-written
- Paragraphs are sometimes pretty heavy to read, due to too long sentences
- I had the feeling that some messages are repeated
- Information are mixed within sections (Intro, Method, Results).

The manuscript has been partly re-written, refer to the specific comments for more details.

1) Despite the (surprising) small differences with the other systems' results, the very low humidity level of the ISP-UNIVE CFA may most likely be source of the (1) increased measurement noise shown by the Allan deviation test, and especially (2) a bias in the measurement as water vapor mixing ratios below <15000ppm leads to non-linearity in the instrument response (Gkinis, 2010). I agree that the instruments used in this study give better performances than the one used back then, which likely explains why the results are that good. Nevertheless, Fig 3 clearly shows a significant trend towards higher value (especially for dD and thus also dexcess), and I would be very surprised if the measurement stays the same at 20000ppm… Thus, to me the conclusions using these data are questionable in the current state.

We agree that the ideal measurement range for the Picarro instrument is around 20,000 ppmv, rather than 10 to 14,000 ppmv conditions. We were obliged to work at this level at ISP-UNIVE, as the water flux provided at the vaporizer did not allow us to reach higher constant humidity levels. And indeed, since modern Picarro instruments do not show strong non-linearity in the instrument response down to 10,000 ppmv now (for instance Aemisegger et al, 2012 compared a L1115i to a L2130i and really showcase the improvement compared to 2010), a big part of the bias is removed, but still, the precision of the instrument is going to be lower at the humidity levels we studied compared to 20,000 ppmv. In the updated manuscript, we will provide Allan variance plots at different humidity levels to illustrate which part of the lower performances of the Venice setup can be explained by the lower humidity.

The new figure will include similar data to the figure below, but for longer time intervals:

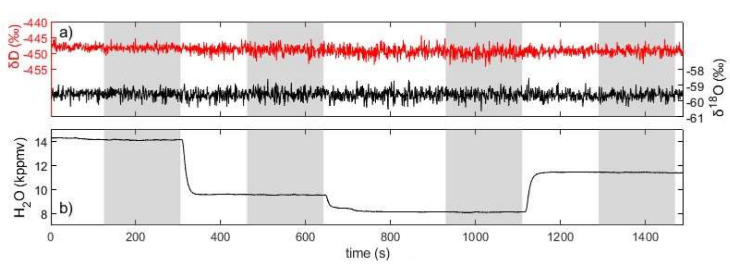

Fig. R1. d18O and dD continuous analysed (top) at different humidity levels (bottom)

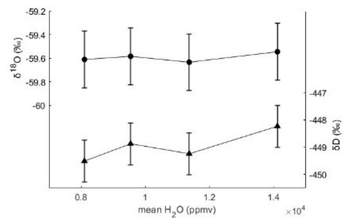

Fig. 3. Mean δ18O and δD values for the 3-minute intervals selected at humidity level of 8,000, 9,500, 11,500 and 14,000 ppmv in the ISP-UNIVE setup (Fig. R1 - grey intervals). The confidence levels are defined as the Allan Deviation for integration time of 1 s (±1σ).

This will also allow us to assess the potential trend toward 20,000 ppmv, extending Fig. 3 (above right) to cover higher humidity levels. Based on these new results, we will more accurately determine whether a correction of the ISP-UNIVE record is necessary.

2) A first question would be what did you choose such low range for? If the effect of humidity was a goal of the study, then comparing the results of low (8-14000ppm) and recommended levels (17-22000ppm) would have indeed been relevant, but only if made on the same instrument.

The referee highlighted an important point. The level of 10,000-14,000 ppmv at Venice wasn't proposed as a target for comparing results at different humidity levels. Instead, it represents the maximum humidity level achievable with the setup proposed by ISP-UNIVE system, balancing melting speed, the collection of discrete samples via the Fraction Collector, online measurement and the need to supply a constant water volume to the CRDS instrument. The ISP-UNIVE system was a novel setup coupled with CRDS and will be further optimized based on the insights provided in this study to enhance future continuous isotopic measurements. As support to the reliability of the results presented by Venice, we are going to include a Section which include the impact of humidity levels based on the same instrument (as mentioned above).

This part will be revised in **Line 121-140**, as addressed in the response to Referee 2 (see comment 21)).

3) A second question is why did you decide not to apply any correction (i.e. humidity level calibration) on this particular dataset? If possible, a discussion regarding differences "with and without" correction can help the reliability.

We decided not to apply any correction to the data because, within the 10,000-14,000 ppmv range, the max observed variation - 0.1‰ for $\delta^{18}O$ and 1.2‰ for δD - is comparable to the confidence interval associated with 1-second integration times (~time of acquisition of Picarro instrument), as determined from the Allan Variance analysis. For this range of humidity, the deviation is such small that we are not able to separate it from the drift of the instrument. In addition, we highlight that both standards and samples were measured at the same humidity levels.
Applying a correction introduces additional uncertainties into the data, which would not be advantageous given the minimal correction that would result for this specific humidity range.

4) The study focuses on 2 French and 1 Italian laboratories, making "International" in the title a bit misleading. I would recommend a term like "Interlaboratory comparison" to better reflect the study.

Title: "Interlaboratory Comparison of Continuous Flow Analysis (CFA) Systems for High-Resolution Water Isotope Measurements in Ice Cores"

Specific comments:
1/ Introduction:

5) The authors present the CFA technique with respect to the traditional discrete sampling technique by focusing only on the reduced time of analysis. However, another aspect of CFA

is related to minimizing contamination, both by avoiding touching the sample, and the inner / outer channels at the melt-head. As all systems involved do not analyze only water isotopes, I assume that the melt-heads used feature both channels. Thus, this should be introduced.

We agree with this point.

Line 51-59: "This would take more than two years of discrete analysis if conducted with a single CRDS instrument. In contrast, with CFA-CRDS, operating at a melt rate of 2.5-3 cm min^{-1}, the same analysis can be completed in roughly three months, with the capability to process up to 10 meters of ice core per day. Additionally, the CFA offers the great advantage of providing - in parallel to the line for isotopic analysis a non-contaminated innermost melt water flow for further analysis.
The innermost melt water flow is used for direct measurement, such as chemical analysis (trace elements, heavy metals, biomass burning tracers, etc.), and insoluble particle volume and distribution, and is simultaneously collected as discrete aliquots, greatly reducing the need for decontamination procedures in clean room. Despite its advantages, CFA-CRDS faces some technical limitations for isotopic analysis, one being the mixing of water molecules within the system, leading to signal smoothing (Gkinis et al., 2011)"

6) Lines 59-62: Dataset synchro and depth assignment... this comes out of the blue at this point. To me, these information belong to the method/data processing paragraph.

Line 59-63: "Additionally, measurement noise – referring to random fluctuations in the instrument' signal output – further limits the effective resolution, restricting the ability to retrieve meaningful climatic signals at high frequencies~~. Accurately determining the delay time between the melt-head and CRDS signals is critical for converting the timescale to a depth scale. Additionally, issues during core melting, such as temporary blockages or stick collapses, introduce uncertainties when assigning depth to the isotopic profile.~~"

Section 2.5

Line 192-216: "The isotopic raw data from the Picarro analyser, provided at acquisition time of ~1-s, are calibrated to the V-SMOW/SLAP and then are post-processed. The post-processing includes (i) converting the time to depth scale, (ii) filtering data affected by memory effects and artifacts and (iii) custom block averaged the data at resolution of 0.5 cm.
(i) The depth scale for the isotopic record is built through two common computational steps across all three laboratories. First, the timescale at MH is converted to depth using the measured ice stick lengths and the continuous recording of the encoder position. The encoder (located at the top of the ice sticks) records the melting of the cores at a frequency of ~1 Hz. Second, the arrival time at the CRDS, conductivity cell, and fractionation collectors is calculated based on the peristaltic pump flow rates. At ISP-UNIVE and IGE, the MH-to-CRDS delay is estimated through preliminary tests and corrected for pump rate changes. At LSCE, the MH-to-CRDS time for liquid phase is calculated using the continuously recorded flow rates and the volumes associated with each component of the CFA setup. Additionally, the delay time for gas phase transport to the CRDS is estimated based on isotopic steps from the SV, under standard operating conditions (specifically, N_2 and water flow rates). This gas phase delay is assumed to remain constant throughout the duration of the CFA run.
Conductivity profiles plotted on the common depth scale help validated the processing. For LSCE, validation is based on the effective synchronization of three conductivity profiles. At IGE, where a single device is located prior the isotopic analyser, validation relies on aligning conductivity peaks – that correspond to transitions between individual ice sticks - with logged depths. Lastly, adjustments may be required between the actual stick length and the depth logged in the field, particularly for fragile and crumbling firn cores. Temporary blockages or stick collapses can introduce uncertainties when assigning depth to the isotopic profile. Whenever possible, such events should be documented and considered during the interpretation of the depth profiles.
(ii) Data at the beginning and end of CFA runs that are affected by mixing with pre- and post-circulated water are manually removed. At ISP-UNIVE, humidity drops – well below the typical work condition – are manually selected using a MATLAB graphical user interface and substituted by linearly interpolated values. (iii) Eventually, data are custom block-averaged at resolution of 0.5 cm."

2/ Method:

7) Line 104: Which Picarro is used where? Line 99 states Venice uses a l2140-i for discrete samples: Fig 2 shows a l2130-i for Venice CFA; Appendix A1 says the LSCE has a l2130-i; Appendix A2 says IGE has a l2140-i. Please introduce more and be consistent.

The previous version of the text did not clearly indicate that two different Picarro models were used at ISP-UNIVE for discrete (L2140-i) and continuous (L2130-i) measurements. We have revised the text as follows:

Line 100-106: "The discrete analysis were conducted at ISP-UNIVE using a Picarro L2140-i. The standards TD (Talos Dome) and AP1 (Antarctic Plateau 1) were used for calibration, while two vials of DCS (Dome C Snow) were analysed as controls. The accuracy of offline measurements was determined as the mean difference between control and true values of the STDs controls, with uncertainty represented by their SD. This yielding an accuracy of -0.01‰ for $\delta^{18}O$, -0.07‰ for δD, and -0.02‰ for d-excess, with corresponding uncertainties of ±0.07‰, ±0.4‰, and ±0.4‰, respectively.

The three CFA systems are coupled with CRDS Picarro-brand isotopic analysers: the L2130-i model at ISP-UNIVE and LSCE, and the L2140-i model at IGE."

8) Lines 138-139*: "The isotopic line includes an additional filter (identical to the LSCE one) and a conductivity device prior the analyser, maintaining a humidity levels between 17,000-21,000 ppmv."*

A filter and a conductivity device are used to maintain humidity levels...??!! please rephrase.

The entire section has been rewritten taking into account the comments from Referees 1 and 2 (see Comment 21).

9) ***Impact of humidity level:***

- See general comment.
- The data shown in Fig3 are results, thus should be presented in the Results section.

Continuous Measurement Noise:

Similarly, the results (Table 3, Fig 4) should be presented in the Results section.

We thank the referee for raising this point and we move Section 2.4.1 and 2.4.2 in the Results section.

10) ***Data Processing***

- Long paragraph (lines 191-215) with heavy sentences, not easy to read. Please rewrite it.

As described above, Section 2.5 has been rewritten

11) Complementary to the Yaxis isotopic data, having information regarding the X-axis (melt-rates plus standard deviation) should be given (in results section or at least in Appendix).

We agree with the Referee, and we believe that providing an indication regarding the X-axis could be a valuable addition. The revised version of the manuscript will include it for the Venice results, as an illustrative example in the Appendix.

12) Line 264-265: the quadrature difference was already mentioned, line 187.

Taken into account. This Section has been rewritten following the additional comments of Referee 2 (see comment 26)).

13) Lines 290-295: The author states that ISP-UNIVE-CFA data were not removed during data processing with intervals during which humidity level decreased down to 1200ppmv and 7850ppmv, respectively. But earlier (line 210), the authors stated removing data below 8000ppmv. This is either very strange and inconsistent, or the corresponding paragraph needs to be rephrased.

We have rewritten Section 3.2 as follows, incorporating the suggestions of Referee1 and Referee2:

Line 285-307: "We present the continuous δ18O and d-excess records from ISP-UNIVE, LSCE and IGE in comparison with the discrete profiles (Fig. 7). For both ISP-UNIVE and IGE, the top 0.50 m of the first bag was removed due to complications encountered during the initial melting phase of the firn cores. These issues were related to the collapse of firn sticks at the melt head and the intrusion of

particles into the distribution line, which required cleaning with UPW. As a result, we show the PALEO2 data from the 12.5-16 m depth section. The CFA data, provided at 0.5 cm resolution after post-processing, was block-averaged at lower resolution by matching the discrete depth intervals (Fig. 7 a and c; original data in Appendix C, Fig. C1). The differences between the averaged continuous and discrete data are analysed using histograms of the differences at each depth point (Fig. 7. e and f), and statistical significance is assessed using a Kruskal–Wallis non-parametric ANOVA test. Differences with $p < 0.05$ are considered statistically significant. Overall, the variability in ice core δ18O records, primarily at the decimetric scale, is comparable between the three CFA profiles and the discrete sampling, showing no statistically difference. In detail, the ISP-UNIVE-CFA shows difference of -0.01±0.26‰ (mean ± 1σ). Two data sections exhibit larger differences to the discrete profile, corresponding to depths of 13.85-13.96 m and 15.62-15.85 m (Fig. 7a, orange areas). The first interval involves ~15 cm of data removed and interpolated due to a humidity drop to 1,200 ppmv. The second interval shows a humidity fluctuation to 7,850 ppmv, slightly below the typical working conditions. This section has been retained. For the LSCE-CFA, the mean difference for δ18O is slightly higher but remains non-significant, with a SD within the instrument's error (0.13±0.18‰ for δ18O, Tab. 5). The IGE-CFA results show a difference of -0.06±0.24‰ compared to the discrete data. The SD, similar to that of ISP-UNIVE but larger than that of LSCE, is primarily attributed to small depth scale shifts (Fig. 7b). For d-excess, the ISP-UNIVE and IGE show statistically difference from the discrete data of -0.78±0.64‰ and 0.88±0.48‰, respectively. These discrepancies are mostly attributed to calibration, as shown by the reduced difference after applying a calibration correction that aligns the mean difference to zero (Fig. 7d). In contrast, the LSCE record shows a non-significant difference of 0.03±0.55‰.
Overall, the good agreement between CFA and discrete record for both δ18O and d-excess suggests that the LSCE data processing is the most reliable among the three CFA setups, which is why the entire PALEO2 core was analysed at LSCE (Sect. 4)."

14) Sentences like (line 295, 296: "*Although rare, such events occurred in the novel coupling of the CFACRDS at ISP-UNIVE, where achieving a highly consistent flow rate was challenging during initial analysis")* are not relevant for the reader.

Taken into account

Additional points:

Line 17: "instrument' signal output" → "instrument's signal output"

Taken into account

Line 29: low accumulation rates on the EAP of 20-50mm weq.yr-1. The accumulation rate at Kohnen station (DML plateau) is of 75 mm w.e.yr-1 (Wesche and others, 2016). Please correct.

Taken into account

Line 55: extend of this effect → extent of this effect

Taken into account

Line 88: remove δ2H

Taken into account

Line 130: "prior it enters" → "before it enters"

Taken into account

Line 205: synchronization instead of superposition

Taken into account

Line 344: Please delete sentences like "*We confirm that the primary advantage of the continuous method is its time efficiency"*

Taken into account

Line 398: TO clean the lines

Taken into account