Reply to Reviewer 2 comments

We would like to thank Reviewer 2 for his/her supportive and interesting comments. Unfortunately, these comments were made based on the first version of the manuscript, after we made thorough reorganization after suggestions by the Editor. It also makes it difficult understanding where these comments are located but we tried all our best. We provided here detailed explanations/comments/modification. For clarity, we kept the reviewer´s comments in red and our replies in Black colors.

Review of manuscript "High-precision atmospheric oxygen measurement comparisons between a newly built CRDS analyzer (Picarro G-2207) and existing measurement techniques" by Tesfaye A. Berhanu et al. submitted to Atmospheric Measurement Techniques General comments:

This paper from A. Tesfaye et al. presents the principle, the method and experimental tests conducted on a new CRDS analyzer dedicated to high precision oxygen measurements in the atmosphere and possibly additional measurement of isotopic content of O2. The first in-situ monitoring results obtained with this instrument are also presented and compared to other current existing measurements technics running in parallel. In the introduction, the authors remind us about the scientific context and the scientific interest to measure O2 mixing ratio in the atmosphere, in the framework of carbon cycle budget and natural/anthropogenic source/sinks attributions, due to the strong link between the oxygen variability and the carbon cycle (combustion and respiration reactions). Then they highlight the analytical challenge to obtain a high precision measurement of O2 due to the very low level of atmospheric variability and they then shortly review the existing measuring technics currently avail- able and the main experimental difficulties associated. In the second section, (Materials and methods), part 2.1, there is a description and discussion of the analyser design principles and characterizations (p4-14). The authors first describe the general instrument principle and design (including associated program modules), then explain the best conditions to be met for an ideal high precision measurement of molecular oxygen and finally constrains linked with an operational deployable field instrument. This provides them justification for the technical and methodological choices made such as spectroscopic model used, water vapour measurement and correction considerations, O2 measurement method design as well as O2 isotopic content measurement. On my opinion this section is a bit too long (about 1/3 of the full article) and also sometime a bit difficult to follow as a non-specialist of spectroscopy. I would suggest to shorten and simplify a bit this section if possible so that it can be more easy to follow. In the case it is not possible to shorten it I would recommend to modify the title of the article to better take into account this section which is anyway useful and interesting (but at the moment reading the title, I would expect the work to focus more on instrumental atmospheric data inter-comparison than technical and spectral analysis).

We are aware that the spectroscopy section comprises significant part in this manuscript. However, we would like to keep these sections in the manuscript in line with Reviewer 1, who requested as much detail as possible about spectroscopy. We understood the need for reflecting this in the title of the manuscript but our main focus is still the intercomparison study between these analyzers.

The second subsection (part 2.2 and following) presents the instrumental tests and evaluation conducted in the laboratory at Picarro, at the University of Bern and in the field in Switzerland (two sites, Jungfraujoch and Beromünster). Experimental set up and conditions as well as methodologies adopted for the tests are presented in these subsections. The last section (section 3) presents and

discuss the results of the different laboratory tests and in-situ monitoring. I would suggest to re-organize a bit this section with the previous one. I think this would be easier for the reader to follow if the test results (in section3) were merged together with the description of the tests procedure (in section 2). So I would merge 2.2.1 with 3.1.1, 3.2.2 with 3.1.2 (actually labelled 3.2 but should be 3.2.2) and also 2.3.3 and 3.2.3 (water correction). I would then keep all the in-situ parts together in section 3.

We believe there is a small misunderstanding to which version of the manuscript these comments are provided. While we submitted our manuscript for the first time indeed we have these sections separated. But based on the recommendations of the Editor we have modified these sections similar to the comments given above. For example section 2.2.1 does not exist in the final version of the manuscript published here but rather merged to section 3.

In the last subsection, the authors present some results for test conducted with the analyser on the isotopic mode. The paper ends with a last concluding section. 2 One general comment and concern of this paper is the reporting unit used for O2 concentration all over the article. The authors used either the ppm unit (most of the time) or the per meg unit (also depending on the instrument used). As there are inter-comparison results used here to validate the new instrument but presented with a mix unit data, it is not easy to follow and to fully compare all data sets as well as precision of the different methods and instruments. Even though there is currently no official international unit to report O2/N2 mixing ratio, and also no Central Calibration Laboratory, there were recommendations given in the last WMO GAW report (report n□242) to report the O2/N2 mixing ratio in per meg units and also if possible to report it on the Scripps Institution of Oceanography (SIO). I would then suggest to make a choice of unit (preferentially per meg) all over the paper and present all the results in a uniform way. When necessary there is a relationship that might be used to express changes in O2/N2 ratio and equivalent changes in O2 mole fraction (Keeling et al, 1998; WMO GAW report n□ 142). Having all number on the same unit would greatly help in the data comparison sections and table 1 (for example) except if this is not applicable. My general feeling about the paper is good, it is generally well written and most of the time clear. I would recommend this article for publication in AMT after revision, as this is a quite interesting new method to measure O2 with a great potential for atmospheric monitoring. Nevertheless, I would highly recommend to take into account the remarks and suggestions raised in the present review in order to straighten and improve the present manuscript. In particular, some additional calculation of mean values and standard deviations would help better evaluate the performances of the instrument compared to current ones.

We used the units ppm and per meg to reflect the actual measured values by the specific analyzers as we believe it will be best to keep the reporting as closely connected as possible to what we directly measure, which is optical absorption. By reporting what we actually measure, our reported values also show most honestly whatever is missing from the picture, such as the dilution due to unmeasured carbon dioxide.

In times where conversion is needed for comparison purpose, we provided the measurements from different analyzers in ppm unit for example Figures 15 &16.

Regarding Table 1, as it is clearly shown, we have provided all the oxygen values in per meg units and CO2 values in ppm.

Specific comments:

Abstract

Line 21: May need to precise that the given short term precision given here refers to the O2 mixing ratio (not to the isotopic ratio).

We have now specified this by modifying this sentence as "…standard error of one-minute O_2 mixing ratio measurements …."

Line 21-23: In this sentence the authors state that the long term stability of the instrument is excellent and prevent high frequency calibration to assess an overall uncertainty of <5 ppm. The recommended calibration frequency is every 12h. With regards to my knowledge and own experience, paramagnetic technics only recall 24h calibration frequency to achieve similar precision. So I would suggest to moderate a bit this sentence, especially the beginning "In contrast to the currently existing techniques".

This statement is partly correct. Indeed, a full calibration for the paramagnetic technique is also made every 18 hour but a frequent 18 minute offset correction is applied since the drift rate of a paramagnetic cell is immense despite a thorough control of the pressure and temperature. In our sentence we refer in particular to the short-term drift that is much better than for corresponding techniques and therefore, we would like to keep this sentence as it is.

The section two of the paper "material and methods" which is the longest part of the paper is not really mentioned in the abstract. May be a few more words should be added in the abstract to remind the reader about the work described in this later section.

We have now added a line to reflect this point in the abstract as follows:

"Here we present detailed description of the analyzer and its operating principles as well as comprehensive laboratory and field studies for a newly developed high-precision oxygen mixing ratio and isotopic composition analyzer (Picarro G-2207) that is based on cavity ring-down spectroscopy (CRDS)".

Introduction:

The introduction section is not labelled as for the other sections (it should be section 1).

It is now labelled as section 1

Line 34-36: I would suggest to update the CO2 mixing ratio to the one of year 2017 (around 405 ppm).

Modified to 405.0 ppm

Line 47-51: There are also WMO/GAW precision recommendations and guidelines for O2/N2 ratio, as describe in the last GAW report (GAW report n ◦ 242, table 1 and p42-44).

We have now added this information on Line 70 as:

Note that the GAW recommendation for the measurement precision of O_2/N_2 is 2 per meg.

Line 55 and 56: Gas chromatography => gas chromatography

Corrected as suggested

Lien 57: As far as I know the techniques described in the previous sentence are not really commercially available. The sensors or detectors can be delivered by commercial companies but cannot be used directly to monitor O2 concentration. There is a need to "customize" these 3 detectors to build a monitoring instrument reaching the precision goal needed for atmospheric monitoring. This is most of the time done by the laboratories themselves!

We agree with the reviewer here. Instruments capable of making O2 measurements of the well-mixed atmosphere at anything close to the precision needed for the scientific goals of the atmospheric community are neither commercially available nor widely used. These are custom-built analyzers that require a great deal of expertise to set up and run them, and to interpret the results properly.

We have removed part of the sentence "commercially available" and the beginning of this paragraph now reads as:

"Currently there are several techniques mostly custom built that can measure…."

Line 61: I would add the following words at the end of the sentence: ". . . of the analysis method especially for continuous monitoring".

We have now added "…especially for continuous monitoring" to this section.

Material and methods:

Analyser design principles:

Line 79: Please define DFB

DFB signifies Distributed Feedback and this term used now in the manuscript instead of DFB.

Line 99-100: What is the typical range of variation for noise between the different instruments?

As far as noise-equivalent absorption goes, that varies by as much as a factor of two between instruments.

Line 102: This is the first time that the Per meg unit it used in the paper (i.e. ppm is used most of the time). As already stated, it would be better to choose and harmonized the unit all over the paper.

We have now defined the per meg unit with equation as follows in the manuscript:

Note that the variations in atmospheric O2 is expressed in units of per meg due to its small variations with respect to a large background and to account for dilution effects from CO2 or any other gas of relevant amount change, which is expressed as:

$$\delta\left(\frac{O_2}{N_2}\right)(per\ meg) = \left(\frac{(\frac{O_2}{N_2})_{sample}}{(\frac{O_2}{N_2})_{reference}} - 1\right).10^6 \qquad (1)$$

Line 116: I'm not an expert in spectroscopy, and the formalism used here to describe the absorption band is a bit unclear for me and a non-specialist. I don't know if there is a way of clarifying or simplifying this another way?

We provided the details of transitions we measure with the quantum numbers of the states measured. However, the main concept here, for a non-spectroscopist, is that we measure a single, isolated absorption line in the 1.27 micron band.

Line 143-145: This sentence is not clear, I suppose there is a verb missing: "In addition, the optical power in the ringdown cavity IS set by the ring-down detector threshold, which . . ."

This section is now excluded from the manuscript as suggested by Reviewer 1

Line 159-160: This sentence is not clear: I think it should be "It stands out that the residuals that are largely an odd. . ."

We clarified this sentence as:

"It stands out that the residuals are largely odd in detuning from the line center…"

Line 185: Can the authors argue why they consider the dependence of Z on O2 too small to be significant?

The measurements show that any variation is at most comparable to the error bars, so we do not consider it a significant effect.

Line 196: please correct "for in measurements".

The word "in" is now removed

Line 224: Can the author explicit what they mean by 161 isotopologue of water. This is absolutely not clear for a non-specialist in spectroscopy.

This sentence is now modified as follows:

The main features are the Q13Q13 line from trace contamination of oxygen in the sample and several lines that arise from normal water ($^{1}H_{2}^{16}O$, AFGL abbreviation 161) and deuterated water ($^{1}H^{2}H^{16}O$, AFGL abbreviation 162, also abbreviated HDO).

Line 277: Please define "scm".

It is now corrected as "sccm" meaning Standard Cubic Centimeter per Minute

Line 300: I would suggest to change ". . . adds additional .." to ". . . adds more. . ."

Now modified accordingly

Line 308-309: I'm a bit surprised that the instrument is providing a dry mole fraction for O2 using the water dilution experiment as it is stated by the authors above in the manuscript (line 273-274) "more works need to be done to investigate the water vapour correction of the oxygen measurement". My feeling is that the present day correction is still not fully satisfying and should be used with caution! There is also no direct explicit correction equation given in the text nor explanation on how this correction is implemented (or to they use the directly the linear function given on figure 7?).

Complete validation of the water vapor correction has not yet been performed, and is beyond the scope of this paper.

As for the numerical details, the linear fit in Figure 7 determines a linear relationship between optical absorption and water fraction, and the correction to oxygen is just the usual dilution correction.

Line 327-330: Taking into account the low precision of the analyser for isotopic content as stated by the authors, is there still an interest to measure them within the context of environmental studies? Could the authors give us a few example and/or possible application of O2 isotope measurement in the environment that could be achieve with that instrument?

We are not quite sure whether we understood the reviewer's comment correctly. Therefore, we refer to the two issues we can think of. First, regarding the degraded precision of the concentration measurement in the isotope mode. This is due to choosing a weaker main oxygen line to be closer to the minor isotope line selected and a significantly reduced number of ring-downs for the main oxygen line to favor the precision of the minor isotope line. Here, a further optimization depending on the users' needs is possible. Secondly, the interference on the isotope ratio itself on breath air is not yet understood. Further measurements are required to see which substance or substances are responsible for this interference. We would like to mention, though, that measurements on the compressed air composition led to a good agreement. Therefore, we can think of the following applications in the field of environmental research. Biological applications relevant for the climate and environmental research, i.e. photosynthesis/respiration processes close to the plants or even using leaf chambers. Analysis of vertical profile air samples taken by means of an AirCore is another application. First measurements have been taken in 2018. Here, stratospheric-tropospheric differences can be a focus. Many other process studies can be thought of where the oxygen is involved, e.g. combustion processes, electrolysis where incompleteness of the process will lead to isotope anomalies.

Another important application is in isotopic tracer experiments, in which either isotopically labeled carbon dioxide or water can be introduced into a closed plant system to understand better the photosynthesis. The isotope labeling can be performed at levels where the signals are greater than the errors in the measurement.

Laboratory tests at Picarro, Santa Clara:

This section and following subsections should be relabelled, 2.2; 2.2.1; 2.2.2 etc. . .

These sections have been modified and merged to section 3 of the manuscript as suggested by the Editor

Line 346: Please define sccm.

Defined above

Laboratory measurements at the University of Bern:

Line 351-355: Could you please add a reference describing the Bern O2 analytic measurement systems (Both for The Fuel cell system and for the Mass Spectrometer) if available.

We have now added the following reference:

Sturm, P., M. Leuenberger, F.L. Valentino, B. Lehmann, and B. Ihly, Measurements of CO_2, its stable isotopes, O_2/N_2, and ^{222}Rn at Bern, Switzerland, *Atmospheric Chemistry and Physics*, *6*, 1991-2004, 2006.

Line 356: Could you please give us a bit more details about the "pressure controlling unit": What is it, What kind of flow meter? (short description or reference).

The pressure control system includes an electronic controller (Type 250E, MKS) which maintains a pressure difference of zero (precise to 0.005 mbar) across the pressure transducer (Baratron 223, MKS) by adjusting the waste flow using the nearby solenoid control valve (Type 248, MKS)

Line 367-377: I'm a bit surprised that there was no direct measurement made to the IRMS without the tee junction. To my knowledge the IRMS is the only one instrument that can provide a very high precision O2/N2 measurement and should be seen as the reference instrument. So I would have made the test in three steps, first with the Tee measuring on both instruments, then directly on the IRMS without the tee which would have given a reference value and then directly to the Picarro. All this at the different splitting ratios. Is there a reason why the direct measurement to the IRMS was not done?

Actually, there is no specific reason why a direct measurement was not conducted on the IRMS. However, as we conducted comparison of the CRDS analyzer and the IRMS by directly measuring multiple standard gases (See Table 1), we believe it can provide an excellent estimate of how the CRDS measurements are comparable to the IRMS.

Line 380: please replace "case b" by "case ii".

Corrected to Case ii

Line 378-389: The conclusion of this section are a bit disappointing as none of the results are shown and only one value is given (without uncertainty). Would it be possible to show the results of the tests? Previous studies by A. Manning have shown that the tee junction effect could be relevant at the level of precision that we are looking for atmospheric O2 monitoring. The impact given here (0.5ppm) is already more than half of the global precision stated for the instrument (<1 ppm line 23). So, if the instrument is to be commercialized, I would deeply recommend to go deeper into that question and firmly establish the conditions of use of a Tee junction or not.

We agree with the reviewer that it is important to further investigate the tee junction influence on the O2 measurements. During this test period, we have tried to test different scenarios of splitting ratios effect. Unfortunately, our observations are inconclusive which is mostly attributed to the temperature effect observed while decanting a cylinder at high flow. Note that the CRDS analyzer takes about 45ml/min and if we would like to go for a splitting ratio of 1:100, we need to decant the cylinder at a flow of 4.5 L/min, which led to cooling effect at the cylinder gauge. As A. Manning has also shown that temperature plays a major role in fractionating oxygen. Meanwhile, the analyser is commercially available and we ask users to make their own tests or use split ratios if needed in the range where we document the values in the manuscript.

Line 403: How was established this correction function? What is the link with the test from figure 7? See also comment for line 308-309.

Already replied above

Line 424: Add a reference to JFJ measurements and set up.

We have now added the requested reference:

Schibig, M. F., Steinbacher, M., Buchmann, B., van der Laan-Luijkx, I. T., van der Laan, S., Ranjan, S. and Leuenberger, M. C,: Comparison of continuous in situ CO2 observations at Jungfraujoch using two different measurement techniques, Atmospheric Measurement Techniques , 8, 57-68, 10.5194/amt-8-57-2015, 2015.

Line 465: Please remove "is avoided" at the end of the sentence. Can the authors give us more precision about what they call "preconditioned"?

The word "is avoided" is now removed

Preconditioning is a standard procedure at our lab for all flask samples prior to using them for sampling. A dedicated vacuum line was used to pump these flasks to vacuum, then flush multiple times with dry air and fill them to 1 atm with this air prior to sampling.

Line 463-469: I have one question regarding this evaluation. Why do the authors use glass flasks? Why not connecting directly the CO2 free cylinder to the analytical device? This would avoid potential contamination during flask filling.

Its simply because it is easier to for experimental set up for example as we were placing these flasks before and after a water trap, which cannot be easily done with the cylinder.

Results and Discussions: See general comments for re-organisation proposition.

As we mentioned in the sections above, these comments were made on the first version of the MS and these reorganizations have already been applied.

Line 483-490: Looking at figure 9, it doesn't seem to me that there is a real drift. For me a drift would show a smooth continuous tendency to increase or decrease in the values. Here what I see is more something like a large variability on the measurement, I see an anti-correlation between O2 values et Y parameter and to a lesser degree a correlation between peak height and O2 as well. So I don't really understand what the authors mean by drift here. Could you please clarify. Are there also some ideas to eliminate or identify those small drift as stated on line 490?

It should be noted here that the measurement that the instrument reports is not Gaussian in nature, such that the Allan std. deviation does not decrease according to the square root of averaging time for long times (> 1 hour).

To clarify more this paragraph and in agreement with the reviewer's comment, we have now replaced the word 'drift' with 'error' as "the residual error of the analyzer…." And "Possible sources of error…..".

Line 511: What is a very good agreement? Can the authors give us an estimate of the mean difference (on a comparable unit for example?). This would help to evaluate the accuracy of the

new instrument and see how well it meets the WMO recommendation or not. Same goes for table 1 which is difficult to use because of the different units for the different instruments!

We intentionally converted the O2-CRDS values from ppm to per meg units to make them comparable. As can be seen from table 1 the three Scripps cylinders (ST3-ST5) IRMS UBern measurements are in agreement with the assigned values by Scripps to 0.6 ± 3.7 per meg, the O2-CRDS measurements for those 5.6 ± 17.8 per meg and the O2-Parameagnetic measurements show a comparison of 20.6 ± 26.8 per meg. A similar agreement is obtained between the three methods when including the cylinders ST-1 and ST-2 prepared by UBern. The picture is different for the ST-6 and ST-7 for known reasons as explained in the manuscript.

Line 517-519: This is absolutely needed if the final goal is to get high precision O2 measurement. There is no need for high precision CO2 measurement but this dilution effect is to be taken into account as already done with present day "homemade" monitoring systems, even for atmospheric monitoring purpose.

Here we are referring to a cylinder with 2700 ppm CO2! As we mentioned in multiple sentences including the conclusion section, it will be important to have a parallel CO2 measurement (or the possibility to have a second laser for CO2, at least in the future) to account for dilution effect. As a side note any kind of gas addition to an ambient air will lead to a gas dilution effect. This even includes using compressed air by gas filling company or self-made ambient air compression when there small leak in the compression line, which could alter the gas mixture. Which could, for instance, lead a change in Ar/N2 or O2/N2. Even more care should be taken when using artificially compressed air-like gas mixtures. Here a proper determination of the gas components needs to be done.

Line 521: Please change "The measurement precision of the Picarro G-2207 measurement was calculated. . ." to "The measurement precision of the Picarro G-2207 measurement was calculated. ."

This comment is not clear. But we changed the text as:

"The measurement precision of the CRDS analyzer was calculated. . ."

Line 521-526: What about the precision of the reference instrument that should be the IRMS?

See answer to the question above (line 511…)

Line 527- 536: I fully agree that it is difficult to compare the graph as there is this problem of unit already highlighted in this review. I'm not sure how significant is the small difference in the correlation coefficient calculated here.

The fact that the mean offset as well as the standard deviation of the measurements of ST-1 to ST5, as given above in the answer to question line 511…., is larger for the paramagnetic cell (at least for the instrument at UBern) than for the CRDS instrument, can easily clarify our statement in this line.

Line 541-542: I agree that the drift at the beginning could be linked with unstable condition after unpacking but the drift remains all over the measuring period and usually Picarro are stable within 3-5 hours after starting measurements. Line 551-553: Did the manufacturer find the cause of this drift. Was there any significant change in the hardware or software configuration of the initial instrument?

One possibility for the cause in the drift was an optical amplifier in the system, which produced a significant amount of broadband light. This light could fill the cavity (albeit with a low coupling coefficient), and would ring down with a different (and generally much faster) time constant that the baseline loss of the cavity. However, the ringdown time on the peak of the oxygen line is just 10 microseconds, such that the broadband light might have distorted the single exponential decay of the central laser frequency, leading to the observed drift in the oxygen signal. We were however not able to confirm this hypothesis. There is no optical amplifier in the present design of the product.

Line 558-562: Did the authors also made water measurement comparisons between Licor and Picaro on wet air conditions?

We have made all the LICOR measurements using wet samples. However, we did not make an absolute comparison between the LICOR and CRDS analyzers for two reasons:

- First the LICOR water measurements are not calibrated (it of course could be done but firstly this would be outside of the scope of this publication and secondly we generally dry the ambient air)

- Second our focus in this manuscript was not comparing the water measurements by both devices

But as it can be seen in Figure 14, the water measurements from both analyzers for dried and non-dried air show similar behavior with matching water peaks

Line 564-565: This sentence is very confusing. Please reword as follow: ". . . in O2 measurements in both cases. (Figures 15c & 15d) shows in case. . ."

This section is now rephrased as follows:

"The water correction test was conducted by measuring dried ambient air (Figure 15a) into both analyzers as well as allowing non-dried air to the CRDS analyzer only (Figure 15b) and comparing the difference in O_2 measurements in both cases. Figures 15c & 15d show the water contents of dried ambient air measured in both analyzers (note that the CRDS uses its in-built water correction function)."

Line 573-577: Can the authors provide some more precise numbers such as for example mean values and standard deviation calculated from data shown in figure 15c and 15d. This would greatly help quantify the variability and give a comparison element with regards to the given instrumental precision.

For 15 c, mean = 1.85, Standard deviation=6.8

For 15 d, mean = -10.4, Standard deviation=14.6

Line 581-583: I disagree with this statement, there are several sections of the paper dealing with the water correction factor. There was a choice stated in the paper and made by the manufacturer to enable wet air measurement, so the water correction is a key issue if the instrument is going to be sold soon and to assess high precision measurement. I'm convinced the correction factor is not easy task to handle and the results presented here are not sufficient to close the problem and give a final solution but this has to be further investigated.

We fully agree with the reviewer that further and even more detailed and extended water correction analyses have to be performed, but we do not agree that it should be part of this publication. We note that for the most of the instruments on the market further improvement of correction functions are found over time. This will certainly also be the case here.

Line 585-593: How are the Picarro data calibrated (based on the in-situ calibration cylinders that have been measured also I suppose)? Could the author quantify a bit more precisely the "very good agreement" like for example providing the mean and the standard deviation of the data for both analysers over the full period. For me, based on the figure 16, it seems that there is a little offset between both instrument (paramagnetic a bit lower) and that there is a slight higher variability for the Picarro instrument compared to the other one but it is difficult to assess with only the figure.

Yes, the Picarro data is calibrated using the standard cylinders measured in-situ.

This question is a bit unclear. These are ambient air measurements with natural variability (on top of the variability from the analyzers) and we do not see the point of providing the mean and sd of these measurements.

If what the reviewer is implying here is for the difference, the calculated mean is -0.33 ppm and a standard error of 0.11 ppm.

Line 595-607: I understand that the isotopic mode is not well suited for ambient O2 concentration measurement but what about the isotopic values? Any comment about those?

Yes, this in an interesting question. The isotope values of ambient air after calibration using internal standards corresponds to expectations. The short-term (second) variations are large but the standard deviations of 5 minute means corresponds to about 0.3 permil.

Line 641: Can the authors provide a table with the individual values for each flasks and instrument so that we can really compare the results and evaluate the precision and repeatability of the measurements on each instrument? Can we add mean values and standard deviation for the three replicates?

We do not think it is relevant to include these values in the manuscript as we already stated that the isotope measurements from the CRDS analyzer needs a closer look but the plots in Figure 18 already gives a clear idea about the above mentioned topics.

Line 644: I think the authors mean ". . .of water and CO2 in addition. . ."

Corrected accordingly

Line 653-655: Would the authors then recommend using the isotopic mode of the instrument at the moment (at least for atmospheric monitoring on atmospheric range) or still need some work to improve it and be sure it is reliable? (at least for atmospheric monitoring on atmospheric range)?

This analyzer cannot measure natural isotopic variations. But, it can be used in tracer experiments where artificially enriched isotopes are used to study various biological processes such as photosynthesis. However, the manufacturer will continue to work on improving it.

Conclusion:

Line 672-677: I feel that the conclusion driven here are a bit optimistic. It is stated several time in the paper that there is still work to do on this question. I would suggest to reword a little bit that conclusion in that way.

We have now added the following sentence recommending for further tests about the water correction.

"However, we believe that it is important to further investigate this issue and identify an improved water correction strategy."

Line 680-681: Here also I would like to see the data with mean values and standard deviation before drawing such an optimistic conclusion (see comments in the previous sections). I think this conclusion also lack a more general statement about the future applications of this instrument and possible improvement (especially for the isotopic mode).

We have now added the following sentence at the end of this paragraph:

"However, we believe that this analyzer can be used for tracer experiments where artificially enriched isotopes are used to study biological processes such as photosynthesis in plants using isotopically labelled CO_2 and H_2O".

Figure 1: The parameter τ is not define neither in the legend of the figure nor in the text. I wonder why the measurements presented here are made at 333Hpa and not at 340 hPa which is the nominal working pressure of the instrument (see line 86)

Tau is just the averaging time for the Allan variance. All measurements with this analyzer were made at 255 Torr, which are same as 340 hPa.

Figure 9: There are strange value above each of the three upper graphs (+2.1028e5 on the upper one). What do they mean?

This is a notation to indicate an offset of the y-axis. In other words, the oxygen fraction reported in the top panel of what is now Figure 8 varies by about 12 ppm about a value of 2.1028e-5.

Line 773: correct "shown to show".

Corrected as "show"