February 19, 2021

Response to E.. Flores's comments for the manuscript "Simultaneous measurement of $\delta^{13}C$, $\delta^{18}O$ and $\delta^{17}O$ of atmospheric CO_2 - Performance assessment of a dual-laser absorption spectrometer

We want to thank Edgar Flores for reading the manuscript and for the comments he made which helped us improving the manuscript, and especially in defining the uncertainties of the measurements.

In this document we will address all the points that were raised. We use **bold text** for repeating the points of the reviewer, normal text for our answers, and *italics text* for changes made in the manuscript. Page and rule numbers apply to the new version of the manuscript.

Technical comments

1) Page 4, line 99: Please provide details on the type of materials, treatments on lines and valves to transfer the gas to the SICAS instrument. Is the gas cell under constant gas flow? If yes please provides the flow rate. We extended this section to provide more details. *Page 4 line 99: " The gas inlet system, depicted in figure 3, is designed to measure discrete air samples in static mode, such that one can quickly switch between measurements of different samples. The system consists of Swagelok stainless steel tubing and connections and pneumatic valves. There are two inlet ports (11 and 14) which are connected to the sample cross at the heart of system (from now on indicated as inlet volume), where a sample is collected at the target pressure of 200± 0.25 mbar before it is connected to the optical cell. One of the inlet ports (11) is connected to a 1/8" VICI multivalve (Valco Instruments) with 15 p otential positions for flask samples or cylinders. The cylinders depicted in figure 3 will be defined in section 2.2 and 3.2. When the VICI valve switches from position, the volume between port 10 and 9 is flushed 7 times with the sample gas to prevent memory effects due to the dead volume of the VICI valve. "*

2) Page 5, line 91: Describe in detail the flushing procedure to avoid cross contamination providing evidence that it works (supplementary information). The flushing procedure has been described in the section that is showed after comment 1. We extended the section on cross-contamination slightly and provided the results of the analysis in the Appendix. *Page 8 line 164:"Cross-contamination, being the dilution of a* *small volume of the working gas in the sample aliquot that is being measured, and vice versa, as described for a Dual-Inlet IRMS in Meijer et al. (2000), will occur in the SICAS due to the continuous switching between sample and machine working gas. If cross-contamination is not corrected for DI-IRMS measurements inaccuracies can occur when samples of a highly deviating isotope composition are measured. On the SICAS only atmospheric samples are measured that are of very similar isotope values. The* CO_2 *mole fraction of the samples can deviate quite strongly from the machine working gas, so effects of cross-contamination will have an influence on the* CO_2 *mole fraction in the optical cell. From experimental data we quantified the fraction of the preceding sample that affects a sample measurement to be max 0.01%. A sensitivity analysis was performed using this fraction and showed that this is such a small amount that scale effects due to cross-contamination are well below the precisions found in this study (for a detailed description of the analysis, see Appendix E). If samples of* CO_2 *concentrations outside the range of atmospheric samples are measured it will be essential to also take into account the surface adsorption effects of the aluminum cell which is known to absorb* CO_2 *Leuenberger et al., 2015).* CO_2 *absorption in the cell of the SICAS was clearly visible as a drop of measured* CO_2 *concentration when an atmospheric sample was let into the cell right after the cell was flushed with a* CO_2 *free flush gas (hence stripped from* CO_2 *molecules sticking to the cell surface). "*

3) Page 5, line 97: Please explain how you took into account the changes in the internal pressure of the cylinders and their potential fractionation effects? Instrument performance The internal pressure of the cylinders has been well above 50 bar during the whole study period shown in this paper. This made us confident that no measures had to be taken to prevent fractionation due to cylinder depletion, as this is known to show pronounced effects below 4 bar. Moisture inside the cylinder is known to cause instability in the isotope composition of CO_2 in the cylinder (Socki et al., 2020). All cylinders were evacuated overnight before filling and were filled with dried air. While working with low sample consumption and at high internal cylinder pressure, we did not evaluate potential fractionation effects thoroughly. For future use it would be interesting to evaluate the potential fractionation effects due to storing cylinders in vertical position, and consider to store long-lasting reference cylinders in horizontal position. Although essential for long-term measurement stability, we consider cylinder treatment out of the scope of the paper, which is mainly on evaluating the measurement performance of our laser absorption spectrometer and the calibration method for determining the triple stable isotope composition.

4) Page 6, line 110: "machine working gas" is not defined. In Figure 3. The term Ref/Working gas is used. Is it the same? The term machine working gas is replaced by working gas, which is defined in section 2.2. In section 3.2 the other cylinder names are defined.

5) Page 6, line 116 is stated: "The precision became significantly worse for all species but isotopologue 627 in the time period between September 2017 and July 2019 due to a gradual but significant decrease (of about 50And in Page 6, line 127 is stated: "Hence, we were able to clean the mirrors and retrieve..Questions: o When exactly were the mirrors cleaned? o Which measures were affected? o What was the short and long term effect on the measurements with a timeline instrument response? o What was the effect of introducing moisture and ethanol for the cleaning of the mirrors, short and long term effect? The cleaning procedure took place at the 31^{st} of October 2019, meaning that the measurements presented in the results and discussion section took all place after the procedure and we cannot relate any measurement instabilities from those long-term results to the change in laser signal. Note that that all results presented in section 2 and 3 are from the period before the cleaning procedure. The measurements did not improve due to the cleaning procedure, indicating other issues played a role, maybe already before the cleaning procedure, or as a result of the cleaning and following realigment procedure. We added this extra information in the text. This section was added to the manuscript to provide the reader with extra information on potential reasons for mirror contamination, being most likely the cause for the observed decrease in laser intensity. Presenting the effects of contamination and finding the best way to clean the mirrors is, however, not our aim. *Page 7, line 144: "This procedure, performed at the 31^{st} of October in 2019, deviates from the recommended mirror cleaning instructions in which it is advised to use ethanol only to clean the mirrors. The additional use of distilled water was in our case necessary since the precipitated aerosols were not dissolved in ethanol and were therefore not removed when we used ethanol only. Despite the increase of the laser signal due to the cleaning procedure, precisions did not improve as a consequence of it. This indicates that other, still unidentified, issues played a role in the decrease of measurement precision."*

6) Page 9, line 199: the three experiments performed over the last two years means: before or after the cleaning? All three experiments presented in this section were performed before the cleaning procedure.

7) Cross contamination and drift are only considered as uncer-

tainties components on the instrument performance. However the instrument was used to measure cylinders and flask that could have important differences in matrix composition. What was the pressure broadening effect on the CO2 measurements? All our cylinders and flasks, also the ones that were prepared by dilution of CO_2 with CO_2 free air, are natural air mixtures. We do not use any synthetic reference or sample gases, except for the samples presented in table 4, which shows the results of an experiment to test possible effects of differences in matrix composition. The composition of natural air is very stable, except for trace gases as CO_2, CH_4 and N_2O, etc.. Differences in CO_2 mole fractions are taken into account in the calibration schemes. We typically collect our reference gases at near background, or slightly enhanced trace gas concentrations. Therefore we consider the other trace gases of such low mole fractions that we consider the air matrix to have a negligible influence on our isotope measurements.

8) What is the real contribution of various components of the air, temperature and pressure variability into the instrument? Uncertainty analysis is missing. We show measurement stability over a whole measurement sequence (12 hours), and we show the effect of measuring a working gas alternately with every sample gas measurement. We realise that more insight in the repeatability and accuracy, as well as uncertainty effects of our reference gases used will help the reader getting insight in the performances nad the potential of this instruments. We therefore added section 4.2 in which we show how we derive the combined uncertainty of our measurements. Although insight in the real contributions of parameters that were given here by the reviewer are of great importance as this would help in the development of an instrument of higher measurement quality, we want to present the measurement and calibration quality as users of the instrument instead of developers. We therefore think that presenting the measurement errors and repeatability is enough to give readers an idea about the potential of the measurement device for the application for atmospheric isotope research. *Page 20, line 454: "A combined uncertainty consisting of measurement uncertainties and scale uncertainties is calculated for the sample measurements. Measurement uncertainties include the standard error of the sample measurement, the repeatability of all (usually four) measurements of the quality control gas throughout the measurement sequence, and the residual of the mean of the quality control gas measurements from the assigned value. The measurement uncertainties will therefore vary with each measurement/measurement sequence. We observe a high repeatability of all sequences included in the analysis of figure 7 (8 in total); with standard errors ranging between 0.005 and 0.03‰ and a mean of 0.014‰ for* $\delta^{13}C$*, and standard er-*

rors ranging between 0.011 and 0.04‰ and a mean of 0.012‰ for $\delta^{18}O$*, for both methods. The residuals in these sequences show a higher contribution to the combined uncertainty and a small difference between the two calibration methods. The absolute residuals of the RM range between 0.0009 and 0.07‰ with a mean of 0.026‰ for* $\delta^{13}C$*, and between 0.007 and 0.06‰ with a mean of 0.04‰ for* $\delta^{18}O$*. For the IM the residuals range between 0.002 and 0.05‰ with a mean of 0.023‰ for* $\delta^{13}C$*, and between 0.012 and 0.05‰ with a mean of 0.03‰ for* $\delta^{18}O$*. Hence, the RM shows slightly higher contributions to the combined uncertainty as a result of the accuracy of the quality control gas measurements. The scale uncertainties, which are fixed for all measurement sequences in which the working gas, low reference and high reference are used as calibration gases, were simulated using the Monte Carlo method. Input values were generated by choosing random numbers of normal distribution with the assigned value and uncertainty as in table5 being the mean and the standard deviation around the mean, respectively. As the RM and IM follow different calibration schemes, the Monte Carlo simulations are discussed separately; for the RM the scale uncertainties of the assigned delta values result in an uncertainty in the calculated residuals which are quadratically fitted against the measured* CO_2 *mole fraction. The average uncertainties in the calibrated delta values of the 5 simulations are 0.03 and 0.05‰ for* $\delta^{13}C$*, and* $\delta^{18}O$*, respectively. Besides the uncertainties introduced by the scale uncertainties of the delta values, the calibrated measurements of the IM are also affected by the scale uncertainties of the* CO_2 *mole fractions. Both the uncertainties in the delta values and in the* CO_2 *mole fractions affect the calculated assigned isotopologue abundances, which are quadratically fitted against the measured isotopologue abundances. The uncertainties in the assigned delta values result in average uncertainties of 0.03 and 0.06‰* $for\delta^{13}C$ *and* $\delta^{18}O$*, respectively. The uncertainties in the assigned* CO_2 *mole fractions result in uncertainties of 0.005 and 0.018‰* $for\delta^{13}C$ *and* $\delta^{18}O$*, respectively, and are small compared to the uncertainties of the assigned delta values. Reducing the combined uncertainty of the* $\delta^{13}C$ *and* $\delta^{18}O$ *measurements of the SICAS will be most effective by determining the isotope composition of the reference gases with a lower uncertainty on the VPDB-*CO_2 *scale.*

9) Page 8, line 160 is stated: "A sensitivity analysis was performed and showed that this is such a small amount that scale effects due to cross-contamination are well below the precisions found in this study" Where is the sensitivity analysis? This an important effect and evidence should be shown to sustain this statement. Please see Appendix E and experimental data.

10) How the uncertainty of the method is constrained by the uncertainty in the reference values of the CO2 mole fractions in the calibration standards? Please see the added section 4.2, as already showed in point 8.

11) What is the effect of diluting pure CO2? Uncertainty in the mole fractions related to this? What is the contribution of the loss of CO2 on the wetted surfaces and the emptying of the flasks? We did not measure the absolute CO_2 values of the diluted samples, as we use in the analysis the measured CO_2 mole fractions. The dilution process was done, keeping in mind the range of CO_2 mole fractions that we would like to test, not trying to produce samples of a CO_2 mole fraction with high accuracy. Flasks that were used for the CMFD experiments were dry, we made sure to evacuate the glass system that was used for the preparation of the samples for at least one night before starting the procedure. What the uncertainties are exactly is hard to say, and would only be possible to check by doing a so called closed loop experiment: diluting a well-known pure CO_2 reference gas to atmospheric concentrations, extract it again and measure it to check whether any fractionation occured. We did not include this work in this paper, as this is still ongoing work.

12) Page 10, line 224: "..the known CO2 mole fraction of the working gas..": "mole fraction" is the quantity referred to in the paper. The proper unit would be mumol/mol rather than ppm. We added an extra sentence to the introduction to clarify that we express the CO_2 mole fraction in $\mu mol/mol$, also referred to as ppm. *Page 3 line 70: We report* CO_2 *mole fractions in* $\mu mol/mol$*, also referred to as ppm.*

13) Figure 5 shows the experiment 1,2 and 3 that correspond to different measurements carried out in 2017, 2018 and 2019. The authors stated in page 6, line 116 that "the precision became significantly worse for all species but isotopologue 627 in the time period between September 2017 and July 2019 due to a gradual but significant decrease (of about 50 percent) in the measured laser intensity over that period". Why do the error bars in 2017, 2018 and 2019 measurements look the same for the mole fraction range 400 to 460 ppm? The statement applies to figure 4, in which it is clearly visible that measurement precision of a single gas measurement gets worse over time. The figure the reviewer is referring to is figure 5, in which the mean values of drift corrected measurement series are shown. Here, repeatability of the iterations of the measurements will introduce uncertainty as well. Apparently, the reduction of the quality of a single gas measurement, which was probably caused due to reduction of the laser intensity, is of less importance than the repeatability of the measurements.

14) Page 12, Figure 5: Why there no uncertainty bars in x-axis? Where are the confidence bands that could support the statements? Typical measurement uncertainties of CO_2 concentrations of the SICAS are ∼0.2 ppm. These uncertainties would fall within the size of the marker and would therefore not be visible. The fits that are showed here have the aim to show the non-linearities of the rare isotopologue abundances as a function of the abundant isotopologue, not to quantify these relations exactly.

15) Page 14, line 269: is stated "In our lab CO2 in air samples of the same isotope composition but deviating CO2 mole fractions are prepared manually, introducing again uncertainties, and doing these experiments regularly is therefore labor- and time intensive". However no uncertainty budget (assessment) considering all the uncertainty contributors for the manual preparation is shown. As a minimum, it is likely that air composition affects CO2 measurements and this is recognized latter in the text, line 291. We used natural air, only scrubbed from CO_2 using Ascarite for the preparation. We assume therefore that the effect of air composition will be negligible, see also my answer on point 8.

16) Page 14, line 275: the reference cylinders are mentioned for the first time in this section but those are not identified (serial number) neither their composition (air matrix). The same issue with the air samples used. We don't see the need for providing the serial numbers as we identify the tanks by name, not by serial number. As we already elaborated in point 8 all our reference cylinder contain dried natural air.

17) Page 14, line 285: A brief description of the gaseous reference materials is given in this section 3.2 with additional information page 15, line 302 and Appendix B. Nevertheless it is essential to have as much as information as possible on the reference materials for allowing the readers to reproduce such experiments. All the reference materials that we use are dried natural air mixtures. Please see also our answer on point 7.

18) Page 14, line 287: The author specify that "two tanks that are specifically used for CMFD corrections. These latter two consist of a high mole fraction reference tank (HR) and a low mole fraction reference tank (LR) covering a great part of the CO2 mole fraction

range occurring in atmospheric samples" Such range is: 342.8 to 423.77 ppm. However in Page 9, line 195 the author state: The SICAS is designed for the measurement of atmospheric samples of which the relevant range of CO2 mole fractions is 370 – 500 ppm, and experiments were therefore for the most part conducted in this range. The author can only warrant that the CMFD corrections of the instrument following the method proposed is valid for the range 342.8 to 423.77 ppm, not higher since there is no evidence for that. We agree with this statement. In the modified manuscript measurements that are done outside the range are left out of the analysis with the argument that extrapolation of calibration curves should at all times be avoided. One of our main conclusion is that extending the range of the reference cylinders will make the instrument better suitable for measuring the isotope composition of the whole range of atmospheric CO_2 mole fraction samples. We sentence in which the range of 370-500 ppm is mentioned applies to the experiments on determining the CMFD of our instrument, in which no calibration, only drift correction, is applied. The sentence has been rephrased to make this clear. *Page 10 line 211:"The SICAS is designed for the measurement of atmospheric samples of which the relevant range of* CO_2 *mole fractions is* $\sim$*370 – 500 ppm, and CMFD experiments were therefore for the most part conducted in this range."*

19) Page 15, line 317: It will be important for the robustness of this paper to list, and assign an experimental uncertainty, to each of the uncertainty sources cited in this section "small leakages or other gas handling problems might be introduced". As per reference gas five flasks were sampled, gas handling problems and small leakages will appear in the combined uncertainty which is given in table 5. As uncertainties are rather small, we don't think gas handling problems would have caused major problems. We are mainly interested in the total uncertainty of our measurements, as the overall uncertainty is acceptable we are less interested in the exact contribution of uncertainty of the filling process.

20) Page 15, line 304 . The claimed standard uncertainties on the gas tanks that were produced in-house from dry compressed natural air, "HR 423.770.01 ppm" and "LR 342.810.01ppm", are very unlikely (considering that Y = y U). During the CCQM K120.a international comparison only one laboratory submitted comparable uncertainties (NMIJ) but considering potential effects of adsorption of a proportionof the molecules onto the internal surface of a cylinder and valve a limit on the uncertainties claimed by par-

ticipants contributing to the reference values on this comparison was fixed to 0.095 mol mol-1 meaning that any uncertainty claim smaller to this value was replaced by 0.095 mol mol-1 to calculate reference values. The claimed standard uncertainties from the gas tanks in this work are even half of the NOAA (real air mixtures) submitted uncertainty for the CCQMâAR K120.a international comparison. These values were indeed incorrect. The 0.01 ppm uncertainty applies to the measurement uncertainty, while the scale uncertainty was not considered, while very important for determining the calibration uncertainty. We expanded the text in section 3.2 and changed the uncertainties to the correct values after this comment of the reviewer. *Page 15 line 330: The* CO_2 *mole fraction of the tanks was measured on a PICARRO G2401 gas mole fraction analyzer and calibrated using in-house working standards, linked to the WMO 2007 scale for* CO_2 *with a suite of of four primary standards provided by the Earth System Research Laboratory (ESRL) of the National Oceanic and Atmosphere Administration (NOAA). The uncertainty of the WMO 2007 scale was estimated to be 0.07* $\mu mol/mol^{-1}$*. The typical measurement precision of the PICARRO G2401 measurements is 0.01* μmol^{-1} *resulting in a combined uncertainty of 0.07* $\mu mol/mol^{-1}$ *for the assigned* CO_2 *mole fraction values of the calibration tanks, while difference between the two cylinders is known with a much lower uncertainty.*

21) Page 16, Table 5. define what are CO2 err, 13C st. err., 18O st. err. and their units. It appears very unlikely that calibrated reference materials of CO2 in air are provided with an uncertainty equal to 0.01 mol mol-1. Therefore CO2 err is most likely a precision. Please justify why the precision is the only uncertainty contributor that maters by proper considerations on the goal of the experiment. We agree, see our comments after point 20.

22) Page 17, line 368: The statement "The mole fraction (X) of the four most abundant isotopologues of a measured CO2 sample are determined using two references gases with known CO2 mole fractions and isotope compositions" must be completed by including: - A phrase stating that the measurements are only valid for the range of 342.81 ppm to 424.52 ppm (according to Table 5, page 16) since the standards must bracket each of the three expected isotopologue mole fractions in the samples; We agree, and added therefore the following phrase: *Page 19 line 416:"The* CO_2 *mole fractions are ideally chosen such that normally occurring* CO_2 *mole fractions in atmospheric air are bracketed by the two reference gases. The low and high reference gases* *cover the range between 324.81 and 424.52 ppm, meaning that this method is only valid for samples within that range of* CO_2 *concentrations. "*

- an uncertainty budget including at least two components, the first related to the repeatability of the measurement results (MPI-Jena), and the second related to the stability and homogeneity of the isotope ratio values in different standards containing nominally the same CO2 gas; The first component is now included in the manuscript, see section 4.2. The second component is very important for long-term measurement stability, however, the reference gases have not been in use for such a long time that we have sufficient information to answer this question. We will, of course, do a re-calibration of our reference gases in the near future, and we will be able to address this issue accordingly.

- It is also crucial when proposing a calibration procedure to consider the uncertainty of the calibration standards. The uncertainties on the reference isotopologue mole fraction values need to be first estimated, and then used in an uncertainty budget reflecting a two point calibration process for each isotopologue. We added section 4.2, in which we consider these uncertainties. See also our answer on point 8.

23) The term "reference gases" is not clear into the document. It is first mentioned in - Page 17, line 369; "The CO2 mole fractions are chosen such that normally occurring CO2 mole fractions in atmospheric air are bracketed by the two reference gases" - in Page 17, line 348: ": We developed a calibration method based on the idea that including the measurement of two reference gases covering the CO2 range of the measured samples (in our case LR and HR) enables the correction of the measured isotope ratios: ",.. then; - In Page 17, line 371: "..Due to the broad range of CO2 mole fractions that are covered by the reference gases, measurement of both working standards will enable the calculation of the (linear) relation of the measured mole fraction (Xm) and the Xa,.." .. then in; - Page 24,line 506: "..we used natural air as reference gases (or air mixtures close to natural air)" - but then in Page 5, Figure 3: Ref 1 and Ref 2 appears without being defined elsewhere in the document and the Ref/Working gases term as well. So are Ref 1 and Ref 2, the reference gases? Same as listed in Table 5 as LR and HR (Page 16)? Which gases are the Ref/Working gases? Please define; reference gases, measured samples, reference cylinders, calibration gas mixtures, ïnC ask samples, HR cylinder, LR

cylinder, Ref orking gases and machine working gas. We now define the term reference gas as: *Page 14, line 291: In the daily procedure of the SICAS there are at least two CO_2-in-air reference gases (in short reference gases), high pressurized cylinders containing gas of known isotope composition and CO_2 mole fraction..."* The term working gas is defined earlier as: *Page 6 line 129: "The cylinder used for drift correction which we define as the working gas contains natural air of which the isotope composition and the CO_2 concentration is known."*

Form these definitions one can conclude that the working gas is a reference gas.

We decided to not use the abbreviations WH, LR and HR in the text for clarity. Instead we use the, earlier explained term working gas, and low reference and high reference. Low and high reference are defined as: *Page 15 304: "two tanks containing a high mole fraction reference gas and a low mole fraction reference gas, from now defined as the high reference and the low reference, which can thus be used for CMFD corrections. The high and low reference cover a great part of the CO_2 mole fraction range occurring in atmospheric samples."*

Also the terms used in figure 3 were changed to the terms defined above.

24) Page 17, line 370: Is Appendix 1 or Appendix A? Changed to Appendix B.

25) Page 18, Table 6. No units? Has been added, as well as in table 7,8 and 9.

26) Page 19, line 407: when was the sausage series 90-94 measured in 2020 (month exactly?)*Page 22 line 490:"SICAS measurements took place in the period from December 2019 to April 2020* **As described in Page 20, line 434. There was an important long term effect of the aliquot storing during 3 to 20 months for 18O measurements and this could strongly influence the conclusions of this intercomparison.** Because of the observed effect of drifting oxygen isotope values cannot compared the $\delta^{18}O$ measurements, and only conclusions are drawn for the $\delta^{13}C$ measurements as no observed drift was observed for the $\delta^{13}C$ values.

27) Page 20, Table 7. No units? Why comparing the difference SICAS-MPI with NOAA-MPI? Which technique was used by NOAA? Units are added to table 7. Comparing the difference SICAS-MPI with NOAA-MPI has the aim to put our isotope measurement results into context of the performance of other labs. As NOAA is known to have a good reputation in stable isotope measurements of atmospheric samples, we think showing both comparisons: SICAS-MPI and NOAA-MPI gives a good idea on how we are doing in comparison with expected results from the stable isotope community.

28) Page 22, line 440: figure 8 shows results outside the mole fraction range validated in this work (343-425ppm). No conclusions can be stated for measurements outside the range. We agree, the results out of the range are kept out of the analysis with NOAA, and the figure has been adjusted so it is clear to the reader which measurements are out of the range.

29) Page 24, line 491: The author state "In this study we show that WMO compatibility goals can be reached with our Aerodyne dual-laser absorption spectrometer for stable isotope measurements of atm-CO2 in dry whole air samples if the instrumental conditions are optimal and there is no uncertainty induced because of gas handling procedures (flask sampling for instance)" but unfortunately no uncertainty budget was shown in the paper to underpin this statement. We agree that this statement only applies to the measurement precision and accuracy of our quality control gas measurements, and not to the combined uncertainty. We therefore do not claim to reach WMO compatibility goals but give values for all components of the uncertainty analysis in the conclusion. *Page 27, 592: In optimal measurement conditions, precisions and accuracies of <0.01 and <0.05‰ for* $\delta^{13}C$ *and* $\delta^{18}O$ *are reached for measurements of the quality control tank for both calibration methods. The combined uncertainty of the measurements includes also the repeatability of the four quality control gas measurements throughout the measurement sequence, with mean values of 0.014 and 0.012‰. The last components in the combined uncertainty calculation are caused by scale uncertainties of the reference gases used for the sample calibration, which are 0.03 and 0.05‰ for* $\delta^{13}C$ *and* $\delta^{18}O$ *of the RM, respectively and 0.03 and 0.06‰ for* $\delta^{13}C$ *and* $\delta^{18}O$ *of the IM, respectively. "*

30) Page 24, 496: The author state "Non-linear dependencies on the CO2 mole fraction occur for measured isotopologue abundances but are insignificant in the typical ambient CO2 mole fraction range" This is relatively true if compared to the compatibility goals. We do not state this anymore, as we see that doing a quadratic fit improves the accuracy of the results for both methods.

31) Page 24, line 510: The author state "From studying the re-

sults of the QC we conclude that precisions are significantly better for the RM, while measurement stability is very similar, both for the 13C and the 18O measurements.". With the current version of the paper, there is no evidence to show that the isotopologue method was applied correctly, using proper reference materials and proper measurement sequences which would ensure frequent drift corrections and two points calibration for each isotopologue. If the isotopologue method is kept after revision, and if the same observations are made, consider explaining why the two methods provide different precisions. The IM is revised after comments of D. Griffith, and we now see similar precision results as for the RM, as one would indeed expect. Only the precision calculation was adjusted. We think the application of the IM was already correct, as we apply continuous drift corrections and all isotopologues were calibrated with a 2-point calibration curve. We now adjusted that to a 3-point, quadratic calibration curve which improved the measurement results of the IM more.

Sincerely,

Farilde Steur

References:

Richard Socki, Matt Matthew, James McHale, Jun Sonobe, Megumi Isaji, and Tracey Jacksier: Enhanced stability of Stable Isotopic Gases, ACS Omega, 5(29), 17926-17930, 10.1021/acsomega.0c00839, 2020