February 19, 2021

Response to D. Griffith'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 David Griffith for thoroughly reading the manuscript and giving important comments. We used his comments to adjust the manuscript, which changed, and we think improved, considerably thanks to his constructive review.

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.

General comments

The reviewer suggests to add a literature overview of existing work on calibration methods as they relate to the present work. An introduction in both calibration methods is added in section 3.3, including most of the work that was suggested by the reviewer. *page 17 line 355: "Two different calibration strategies are discussed in this section. The calibration strategies are based on the two main approaches for calibration of isotope measurements, as also described by Griffith et al. (2012) and, more recently by Griffith (2018), being (1) determine the isotopologue ratios, and calibrate those, taking the introduced CMFD into account, from now on defined as the ratio method (RM), and (2) first calibrate the absolute isotopologue mole fractions individually and then calculate the isotopologue ratios, from now on defined as the isotopologue method (IM). We give a brief introduction of the two calibration methods, as described in literature and we describe the measurement procedure that is used for both calibration methods. This section ends with a detailed description of both methods as applied for the SICAS measurements. The RM, being very similar to calibration strategies applied by isotope measurements using DI-IRMS (Meijer, 2009), is usually based on reference gases covering delta values of a range which is similar to the range of the measured samples. Determination of the CMFD can be done by measuring different tanks of varying CO_2 mole fractions or by dynamical dilution of pure CO_2 with CO_2 free air (Braden-Behrens ewt*

al., 2017; Sturm et al., 2012; Griffith et al., 2012; McManus et al., 2015; Tuzson2008, again covering the CO_2 mole fraction range of the measured samples. The IM has the advantage that there is no need to take into account the introduced CMFD (Griffith, 2018). As all isotopologues are calibrated independently, it is only necessary to use reference gases covering the range of isotopologue abundances as occurring in the samples. This can be realised by using reference gases containing CO_2 of similar isotope composition but varying CO_2 mole fractions as described in Griffith (2018) and successfully implemented in Griffith et al. (2012), Flores et al. (2017) and Wehr et al. (2013). The range of delta values that is measured in samples of atmospheric background air is limited (range in unpolluted troposphere is -9.5 to -7.5‰ en -2 to +2‰ for $\delta^{13}C$ and $\delta^{18}O$, respectively (WMO, 2016), hence this also applies to the range of delta values that should be covered by the reference gases when applying the RM. We decided therefore to use the same reference gases to test both calibration methods, varying mainly in CO_2 mole fraction (342.81-424.52 $\mu mol/mol^{-1}$).”

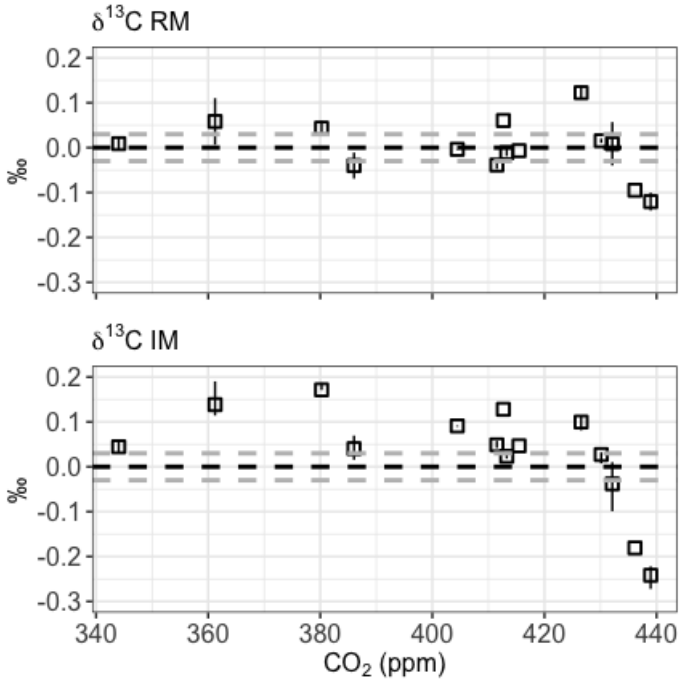
The reviewer asks why the accuracy and precision as presented is significantly lower for IM than for RM. We agree that intuitively the precision should be similar for the RM and IM. We realised that the difference that was shown in the original manuscript was caused by the way the standard deviation of the measurements were calculated. The ratio method is more successful in reaching high precisions because using ratios has the advantage that drifts, that occur at the same rate and direction for both isotopologue abundance measurements, are cancelled out. The standard deviation of a ratio measurement will therefore be, in most cases, more stable than the measurement of individual isotopologue abundances. In the original manuscript the standard deviation of the delta values derived with the isotopologue method were calculated by:

$$\sqrt{sd_{626}^2 + sd_{rare}^2} \tag{1}$$

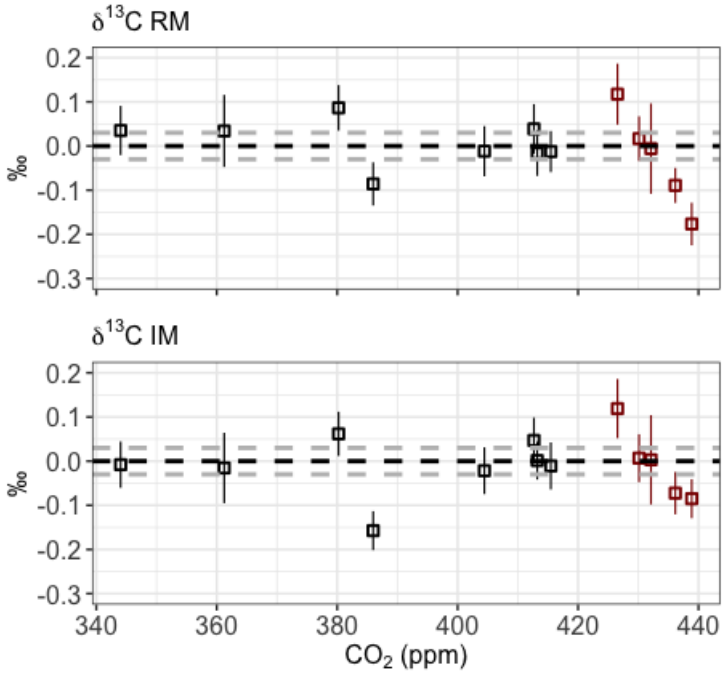
With sd_{626} being the standard deviation the 626 isotopologue measurement and sd_{rare} being the standard deviation of the rare isotopologue measurement. This seemed to be logical as all isotopologues are calibrated individually, and we treated the uncertainties of those isotopologues as different components when calculating the measurement precision. The reviewer points out in his report, as in the comment for L152: ”It is only to be expected that ratios will show lower Allan Variance than individual isotopologues because some sources of noise and drift in the ratioed quantities are correlated and cancel. When isotopologue amounts are ratioed later in a calibration calculation, as in the IM, these correlated noise sources will also cancel. This

should be made clear." This statement is true, and we changed the calculation of the measurement precision accordingly. There are 9 iterations of a sample measurement conducted, and the mean of those 9 measurements will eventually be expressed as a delta value. Initially, we used the mean of the individually calibrated isotopologues for the calculation of one delta value. Now, we use all 9 measurements to calculate 9 delta values, and we calculate the mean of those 9 delta values. The measurement error can therefore be derived by calculating the standard deviation of those 9 delta values. This improves the measurement precision significantly, resulting in similar measurement precisions for $\delta^{13}C$ and $\delta^{18}O$ for the RM and IM (see table 6 in the modified manuscript). (This is no surprise, really, as now the difference between the IM and RM data treatment is much smaller than in our original implementation of IM.) For the $\delta^{17}O$ results we even observe a better measurement precision than the RM. We argue that this difference is possible by a low correlation between the 627 and 626 measurement results. In that case, the IM shows to perform better due to a more effective outlier analysis. All calculation steps for both calibration methods are added to the manuscript in Appendix C.

The accuracy results from the intercomparison with MPI-BGC results showed to be worse for the IM in the original manuscript. This is, as the reviewer suggests, due to the non-linearity which was dealt with for the RM, but not for the IM. In the first part of the manuscript, our study on the CO_2 dependency showed non-linearities of both the rare isotopologue abundances as a function of the measured abundance of the abundant isotopologue, and the measured isotopologue ratio as a function of the measured CO_2 mole fraction. These studies were mainly focussed on getting an idea about the CO_2 dependency of the isotope ratios and we applied these findings of non-linearity for determining the CO_2 correction for the RM. Although a study of the non-linearities of measured isotopologues as a function of the true isotopologue abundance has not been performed (CO_2 samples that were used for the dependency experiments were not determined for their CO_2 mole fraction), this will occur as the result of 'imperfect spectroscopy and spectral fitting', as stated by the reviewer. We therefore adjusted the determination of the calibration curves for the IM by using the hypothetical measurement of the working gas, which will always have a drift corrected value of 1, for doing a quadratic fit instead of a linear fit using only the low and high reference measurements. This reduced the residuals of the SICAS measurement using the IM and the MPI-BGC results considerably.



Results intercomparison in the original manuscript



Results intercomparison in the adjusted manuscript

The reviewer thereby suggests that for the RM real-time empirical concentration dependency is determined, while for the IM merely one two-point calibration curve was determined per isotopologue. Actually, we use the mean of the isotopologue/delta measurement of the reference gases over a whole measurement sequence for determining the calibration curves of both the RM and the IM. We think this approach is more practical than a real-time calibration as a deviating measurement of a reference gas will not affect the sample results as this measurement is usually filtered out by the outlier analysis. We added all calculations applied for both methods in Appendix C for clarification. We stick to this approach in the modified manuscript.

Technical comments

Abstract: In the interest of readability, I think best editorial practice is to avoid abbreviations in the abstract, and to introduce them at first use in the main text. I agree with reviewer 1 that there is some over-use of abbreviations and acronyms. Depending on the response to the general comments above, the key conclusions from the study may change. Abbreviations are now avoided in the abstract. We reduced the use of abbreviations in the text by writing out the full names of the calibration tanks as we agree this contributes to the readability of the text. The conclusions have changed as a result of the general comments that were given by the reviewer as described above. *Main conclusion that were adjusted in the manuscript, page 1, line 9: "Measurements of the quality control tank show that the measurement precision and accuracy of both methods is of similar quality for $\delta^{13}C$ and $\delta^{18}O$ measurements. In optimal measurement conditions the precision and accuracy of the quality control tank reach WMO compatibility requirements, being 0.01‰ for $\delta^{13}C$ and 0.05‰ for $\delta^{18}O$, respectively. Uncertainty contributions of the scale uncertainties of the reference gases add another 0.03 and 0.05‰ to the combined uncertainty of the sample measurements. Hence, reaching WMO compatibility for sample measurements on the SICAS requires reduction of the scale uncertainty of the reference gases used for calibration. An inter-comparison of flask samples over a wide range of CO_2 mole fractions has been conducted with the Max Planck Institute for Biogeochemistry resulting in a mean residual of 0.01 and -0.01‰ and a standard deviation of 0.05 and 0.07‰ for the $\delta^{13}C$ measurements calibrated using the ratio method and the isotopologue method, respectively. The $\delta^{18}O$ could not be compared due to depletion of the $\delta^{18}O$ signal in our sample flasks because of too long storage times. Finally, we evaluate the potential of our $\Delta^{17}O$ measurements as a tracer for gross primary production by vegetation through photosynthesis. Here, a measurement precision of <0.01‰*

would be a prerequisite for capturing seasonal variations in the $\Delta^{17}O$ signal. The isotopologue method performs better for the measurement precision of the $\delta^{17}O$ and $\Delta^{17}O$ with standard errors not exceeding 0.02‰, showing that the IM is close to reaching the high precision requirement for capturing seasonal trends in the $\Delta^{17}O$ measurements. The accuracy results show consequently too enriched results for both the $\delta^{17}O$ and $\Delta^{17}O$ measurements for both methods. The ratio method shows residuals ranging from 0.06 to 0.08‰ and from 0.06 to 0.1‰ for the $\delta^{17}O$ and $\Delta^{17}O$ results, respectively. The isotopologue method shows residuals ranging from 0.04 to 0.1‰ and from 0.05 and 0.13‰ for the $\delta^{17}O$ and $\Delta^{17}O$ results, respectively. Direct determination of the $\delta^{17}O$ of all reference gases would improve the accuracy of the $\delta^{17}O$, and thereby of the $\Delta^{17}O$ measurements."

L27: Low variability is due to the large size of the carbon reservoir and the long lifetime of CO2 in the atmosphere, not the high mole fraction. (Note: mole fraction is strictly the correct term, not mixing ratio) *Page 2, line 35: "Due to the large size of the carbon reservoir of the atmosphere and the long lifetime of CO_2 in the atmosphere, the effects of sources and sinks on the atmospheric composition are heavily diluted."*

L31: Can this reference to WMO 2016 be updated to the latest (from 2019 meeting). The reference is updated to Crotwell et al., 2020, in the whole text

L49: Optical methods to which this paper applies include non-laser methods such as FTIR and even NDIR. Remove (laser), it is in any case appropriately mentioned later in the sentence. *Page 3 line 58: "Optical spectroscopy now offers this possibility following strong developments in recent years especially for the laser light sources, to perform isotopologue measurements showing precisions close to, or even surpassing IRMS measurements (Tuzson et al., 2008; Vogel et al., 2013; McManus et al., 2015). "*

L65: replace "consists (among others of)" with "includes" *Page 3 line 77: "The optical bench as depicted in figure 1 includes the two lasers, several mirrors to combine and deflect the laser beams, the optical cell and two detectors. "*

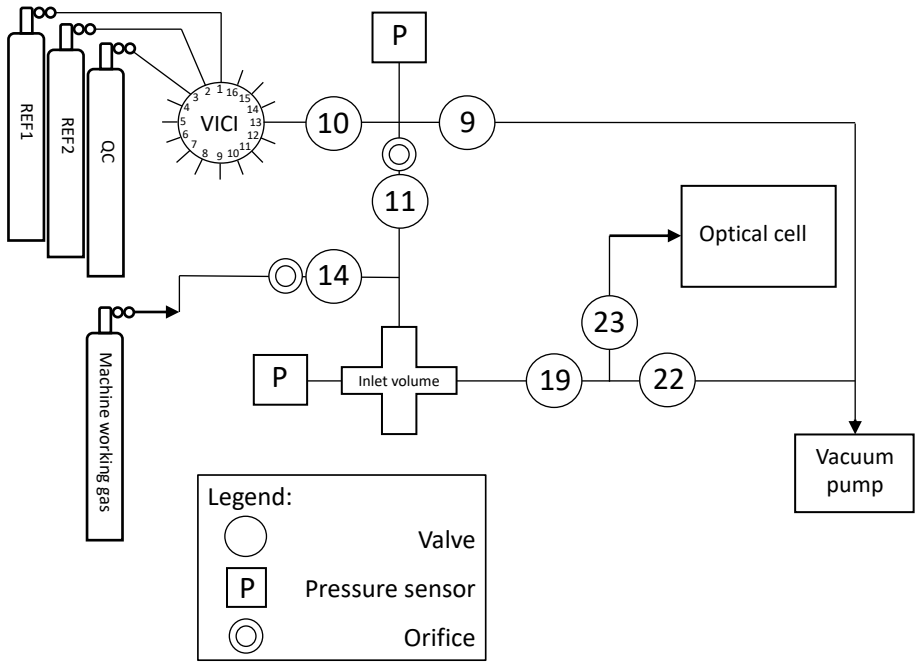
L68: In contrast to Reviewer 1 I am comfortable with the Hitran notation – it is well established (in optical spectroscopy at least) and quicker to type and read than full molecular labels, especially in equations. We stick to the HITRAN notation in the paper, as suggested here.

L69: (with a sweep frequency of...) Add sweep to avoid potential confusion with optical frequency. *Page 3 line 81:"Application of a small current ramp causes small frequency variations so the lasers are swept (with a sweep frequency of 1.7kHz) over a spectral range in which ro-vibrational transitions of the isotopologues occur with similar optical depths (Tuzson et al., 2008)."*

L74: Please specify that this is the path inside the cell (the outside path is dealt with in the next paragraph, but it is ambiguous here) *Page 3 line 86:"The total path length of the laser light in the optical cell is 36 meters."*

L79: How accurate is the temperature control? This is relevant in assessing the causes of drift which are important to the whole calibration strategies. Likewise pressure control (see line 201). We determined the maximum fluctuations of the cell temperature over a measurement period of 12 hours at normal circumstances. The sentence was adjusted additionally, as it stated that the the chiller controlled the temperature inside and outside the housing, which is incorrect, it only controls the temperature inside the housing. *Page 3 line 90: "The temperature within the housing is controlled using a re-circulating liquid chiller set at a temperature of 20°C to keep the temperature in the cell stable. Within a measurement sequence (12 hours) the temperature does normally not fluctuate more than 0.05°C."*

Figure 3: requires a full caption with all labels and abbreviations spelt out. It is quite difficult to follow the text on p5 with the current level of detail provided. L91 – where is the inlet valve, for example? The figure has been simplified by removing elements that are not relevant for this study. A legend has been added for clarity and the caption of figure 3 has been extended as below:

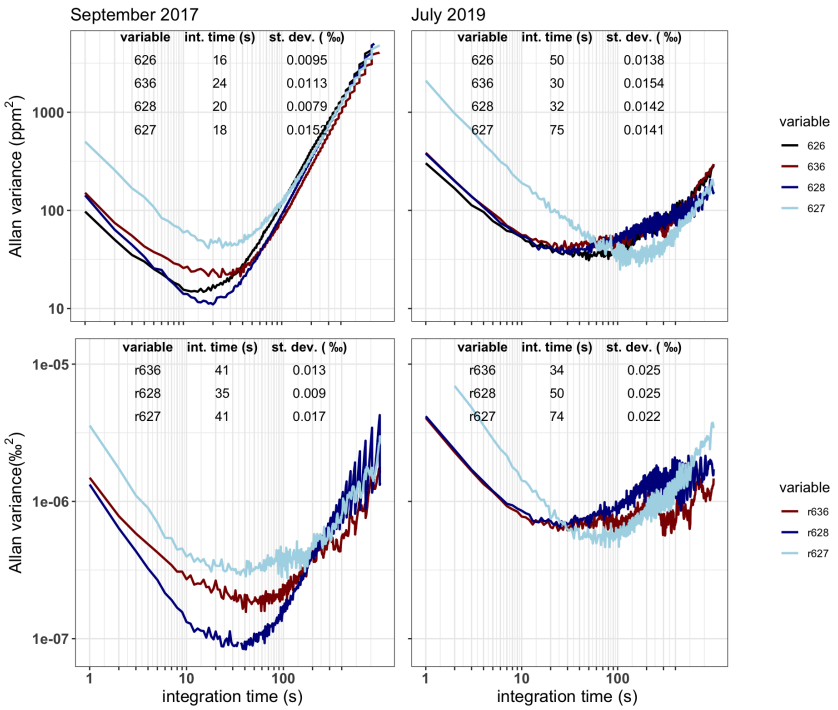


Gas inlet system of the SICAS with one VICI multivalve inlet port, connected to three high pressure natural air tanks and 12 free ports for samples. The includes an extra inlet port for the machine working gas, also a high pressure natural air tank.

L98: mbar not mBar, and please specify this is gauge pressure, not absolute. It is absolute pressure, now indicated in the text and updated to more recent user values. We did not see any influence of the sample pressure, except when too low or too high (very slow filling times and low repeatability of the cell pressure, respectively) so this range is suitable. *Page 6 line 112:"For the cylinders, single stage pressure regulators are in use (Rotarex, model SMT SI220), set at an outlet pressure of 600-1000 mbar (absolute)."*

L109 and Fig 4. What are the units of the Y axes of plots? For the upper plots they should be ppm2 but I cannot relate this to the mimimum figures given in the in-plot tables. The lower plots are presumably $\%^2$. Also the abbreviation "St. dev"; these are not standard deviations, they are the minimum Allan deviations (ie sqrt(AllanVariance)), which is strictly not quite the same thing. Finally, unless the amounts for the isotopologues are scaled by their natural abundances, the Allan variances in ppm2 will have quite different absolute values. The units of the y-axis of the upper figures are indeed ppm^2, now added in the axis title, and the units of the y-axis of the

lower figures were in the original manuscript the absolute deviation of the measured ratio. We adjusted this to $‰^2$ for clarity by multiplying the values by 1000. The adjusted figure is shown below.



The Allan variance as a function of the integration time in seconds for a single gas measurement plotted for both the measured isotopologue abdundance (top) and the isotope ratios (bottom) at September 2017 (left) and July 2019 (right). The best achieved precisions and corresponding integration times are shown as a table in the plots.

L143: please replace "n" with "number of measurements"

L143: should "lower" read "higher"? *Page 6 line 159: "It is expected that, if the drift correction is effective, the standard deviation does not get worse with a higher number of measurements, and that the standard deviations of the uncorrected values are higher than the corrected values."*

L129: what is "demi-water" Demi-water is replaced by ultrapure water.

L152: This is not a fair comparison, or is a misleading statement. It is only to be expected that ratios will show lower Allan Variance than individual isotopologues because some sources of noise and drift in the ratioed quantities are correlated and cancel. When

isotopologue amounts are ratioed later in a calibration calculation, as in the IM, these correlated noise sources will also cancel. This should be made clear. We realise this, and therefore we decided to only show the results of the ratios, as in the end, all amounts are ratioed later.

Table 1: These are relative standard deviations in ‰ please add "relative" to the caption and make clear in the case of the ratios that it is not the delta value itself but the relative std dev of it. This has been adjusted in the table and in the caption, as shown below:

All st. dev.	n=5		n=10	
in ‰	uncor	cor	uncor	cor
r636	0.036	0.020	0.055	0.025
r628	0.046	0.021	0.104	0.029
r627	0.060	0.018	0.177	0.031

Relative standard deviations for n=5 and n=10 of uncorrected (uncor) and corrected (cor) isototope ratio sample measurements. Sample measurements were always bracketed by measurements of the working gas. Standard deviations of the uncorrected measurements only use the sample measurements, standard deviations of the corrected measurements use drift corrected (equation 1) sample measurements using the working gas measurements.

L180, eq. 3. Although this equation is only used for illustrative purposes, it is confused by the true calibration equation (10). The difference is between XCO2 and Xa. XCO2 is itself a function of X* (or Xm) through the relevant isotope ratios so the coefficients a here and c in Eq 10 are different. It would be clearer to use the same equation as (10) here. In the adjusted manuscript there is not used a linear fit for the calibration of the isotopologue measurements, but a quadratic fit. We assume, therefore, that the difference between the calibration equation in equation 8 and the illustrative example in equation 3 is clear. To avoid double use of a and b, the a and b in equation 3 were adjusted to the greek symbols α and β.

L214: Griffith2018 showed that a pseudo linear + inverse relationship to CO2 amount is both theoretically expected and fits data in that paper in practice. The quadratic fit is only an approximation to this. *Page 10 line 226 "Griffith (2018) showed that a combination of a linear and inverse relationship to the CO_2 mole fraction is theoretically expected, and this relationship fitted the data used in his study in practice. As we expect to have a relation of the measured delta values and the CO_2 mole fraction which is close to linear, we use a quadratic relation which approxi-*

mates this expected theoretical relation closely. "

Section 3.2.1 and Fig 5. I have difficulty following this section and figure. It may be better titled "Spectroscopic non-linearities of isotope measurements". If the "residuals" are the differences between a linear regression of X* against X626, they should be in units of ppm, not in Fig 5. I do not understand the sentence on L 223 "The CO2 mole fraction..." The title has been changed as suggested. The residuals are calculated as the difference between a linear regression of X_*, expressed as $M_{S(t)dc}$ for drift correction, against X_626. $M_{S(t)dc}$ is in fact a ratio, which makes it, in our opinion logical express the residuals in permil and not in absolute values. The sentence L223 in the original manuscript has been rephrased to: *Page 10, line 237: "The CO_2 mole fraction is calculated by multiplying $M_{626(t)dc}$ by the known CO_2 mole fraction of the working gas."*

L249: There should be no apostrophe in "deltas". An apostrophe is used to indicate possessive, not plural. There are many examples through the MS. This has been adjusted throughout the text.

Fig 6. See comment for L214, it would be interesting to see these data plotted against XCO2 and 1/XCO2. Does this improve the comparisons in Table 3? The description of this table is also difficult to follow in the text. Data were plotted agains XCO2 and 1/XCO2, showing very similar. See the table below where the new fit results were added as fit inv:

all values in ‰		δ^{636}	δ^{628}	δ^{627}
4*exp.1 (404-1025ppm)	lin	0.871	0.120	0.376
	q	0.072	0.142	0.100
	fit lin	0.141	0.090	0.169
	fit q	0.034	0.092	0.078
	fit inv	0.063	0.090	0.092
4*exp. 2 (313-484ppm)	lin	0.095	0.181	0.095
	q	0.054	0.164	0.097
	fit lin	0.086	0.175	0.093
	fit q	0.049	0.155	0.093
	fit inv	0.054	0.017	0.095
4*exp. 3 (426-522ppm)	lin	0.075	0.186	0.048
	q	0.084	0.162	0.032
	fit lin	0.093	0.191	0.037
	fit q	0.082	0.161	0.028
	fit inv	0.085	0.164	0.029

The XCO2 and 1/XCO2 fit does not show significant improvements, so we decided that the quadratic fit is a good approximation for the theoretically expected relationship of the delta as a function of the XCO2. *Page 14 line 274: "The theoretically expected combination of a linear and inverse relationship as described in Griffith (2018) showed very similar results as the quadratic fit results, so we consider the quadratic fit to be a good approximation of the theoretically expected relationship."*

L245: Please clarify the sentence "From now on we use the with an isotopologue superscript. . . ." and/or give an example. It is quite unclear to me. Has been removed, was unnecessary.

L256-283: This section is important in the context of general comments above on correcting for stability, non linearity and drift. It points out that the calibration changes with time, but my understanding of the analysis is that this is compensated in RM but not IM. This might be clarified in dealing with earlier comments on the IM in 3.3.2.

3.3.2 Isotopologue method. Please see earlier general comments on this section. This does not have the same level of detail as 3.3.1 and if I understand it correctly the IM does not include the equivalent of the interleaved comparison with a reference gas R, or any correction for changes in the coefficients c and d in Eq 10. These changes are the root of any calibration drift in both RM and IM . As this section is written they appear to be corrected in RM but not IM.

Calibration changes over time are dealt with for both the RM and the IM. This has been clarified in the text after earlier comments.

Sincerely,

Farilde Steur