

Response to final comments Associate Editor Tim Arnold on manuscript Steur et al., 2021

We want to thank Tim Arnold for his comments on the manuscript, as well his comments for earlier versions of the manuscript. Please see below a point by point answer on the comments. In grey the comments, in black the responses.

Sincerely,

Farilde Steur

**Point by point answers to comments:**

L10 - 'methods' should read 'calibration methods'

Changed as suggested.

L11 - I agree with the referee report regarding the phrase 'optimal conditions'. This implies that the steps to reach these conditions are well know, but I don't think this is the case. Instead of 'optimal..' suggest 'During one specific measurement period..'

Line 12: “During one specific measurement period the precision and accuracy of the quality control tank reach WMO compatibility requirements, being 0.01‰ for  $\delta^{13}\text{C}$  and 0.05‰ for  $\delta^{18}\text{O}$ , respectively. “

~L70 - there is reference to chapters - please correct

Line 68: “In this paper we present the performance, in terms of precisions and accuracy, of an Aerodyne dual laser optical spectrometer (CW-IC-TILDAS-D) in use since September 2017, for the simultaneous measurement of  $\delta^{13}\text{C}$ ,  $\delta^{18}\text{O}$  and  $\delta^{17}\text{O}$  of atm- $\text{CO}_2$ , 70 which we refer to as “Stable Isotopes of  $\text{CO}_2$  Absorption Spectrometer” (SICAS). The instrument performance over time is discussed, followed by an analysis of the  $\text{CO}_2$  mole fraction dependency of the instrument. We report  $\text{CO}_2$  mole fractions in  $\mu\text{mol/mol}$ , also referred to as ppm. The actual ways of performing a calibrated measurement using either individual isotopologue measurements or isotope ratios is discussed and whole air measurement results of both calibration methods are evaluated for their compatibility with IRMS stable isotope measurements. Conclusively, the usefulness of the triple oxygen 75 isotope measurements for capturing signals of atmospheric  $\text{CO}_2$  sources and sinks is evaluated. “

Eq1: I don't really understand this equation.  $M_{\text{wg}}(t)$  is calculated and  $M_{\text{s}}(t)$  is measured. But  $M_{\text{s}}(t)_{\text{dc}}$  can't be the ratio of these?

$M_{\text{s}}(t)_{\text{dc}}$  has been changed to  $I_{\text{s}}(t)$  as suggested by D. Griffith, to clarify that the  $I_{\text{s}}(t)$  is calculated from the  $M_{\text{s}}(t)$  and  $M_{\text{wg}}(t)$ .  $M_{\text{s}}(t)$  and  $M_{\text{wg}}(t)$  are both measured values, while  $I_{\text{s}}(t)$  is the calculated ratio from those two values.

L156-157: 'The effectiveness of this drift correction method was tested for the measured

isotope ratios, as for calibration of the measured isotopologue abundances to delta values, isotope ratios will always be used.' I can't make sense of this sentence - please revise.

Line 158: "The effectiveness of this drift correction method was tested for the measured isotope ratios only. Although one of the tested calibration methods uses isotopologue abundances for the initial calibration, the isotope composition is expressed 160 as a delta value and will therefore eventually be calculated using isotope ratios (section 3.3.3). The precision of the isotope ratios will therefore always determine the measurement precision."

Figure 5: Can you be clear what is actually being calculated in the residual - the use of per mil on the y axis is confusing if relating to amount fraction difference. Also 'darkblue' should be 'dark blue' etc.

Caption figure 5: "Figure 5. Residuals (expressed in ‰ relative to the measured amount fraction) of the linear fit of the rare isotopologue abundancies as a function of the X626 and the quadratic fit on the residuals. From top to bottom: Experiment 1, experiment 2 and experiment 3. The colours red, dark blue and light blue are used for the isotopologues 636, 628 and 627 respectively...."

Figure 6: In Figure 5 amount fractions go above 1000 ppm so why is that range not also obvious in Figure 6?

Not all points of experiment 1 are shown in figure 6 as this reduces the readability of the figure since the range of the other experiments is much lower.

285-295: mole or molar fraction? - be consistent throughout text

Changed to mole fraction throughout the text.

Figure 7's caption: H or M measurement period? Perhaps relate back to the text section where this is explained. Reader needs to be able to relate captions to the figures easily, or at least know where to find missing info.

Changed caption of figure 7: "...Colour of the points indicates whether the measurements were performed in a High quality (green), Medium quality (black) or Low quality (red) measurement period (see section 4.1 for definitions).

L458: I don't see 0.005 in the table - should these values correspond to those in the table?

The numbers in the text were not updated to the latest version of the table. This has been changed and now the numbers correspond to the table.

L583: 'We developed two different calibration methods based..' I think the methods are

developed elsewhere, implementation here is new. Perhaps 'we implemented two different calibration methods'

Changed as suggested.

Response to final comments D. Griffith on manuscript Steur et al., 2021

We want to thank D. Griffith for his final comments on the manuscript. The comments were used to adjust the manuscript to the final version as uploaded now. His thorough reading and critical review helped us to improve the manuscript significantly.

Please find the point by point answers below. In grey the points made by D. Griffith, in black the response.

Sincerely,

Farilde Steur

#### Point by point answers to comments D. Griffith:

L8: The references are also calibrated for mole fractions, which is required for the IM. Please specify that here.

Line 8: "Calibration with the ratio method and isotopologue method is based on three different assigned whole air references calibrated on the VPBD and the WMO 2007 scale for their stable isotope compositions and their CO<sub>2</sub> mole fractions, respectively. "

L28: I am not convinced that the differences in quoted repeatabilities between the two methods IM and RM are really statistically significant. It would be better to avoid value judgements about which method is "better" and simply state the repeatabilities as observed under the measurement conditions.

Line 25: "The ratio method shows residuals ranging from 0.06 to 0.08‰ and from 0.06 to 0.1‰ for the  $\delta^{17}\text{O}$  and  $\Delta^{17}\text{O}$  results, respectively. The isotopologue method shows residuals ranging from 0.04 to 0.1‰ and from 0.05 and 0.13‰ for the  $\delta^{17}\text{O}$  and  $\Delta^{17}\text{O}$  results, respectively."

Are they really the "optimal" conditions (L11)?

Line 12: "During one specific measurement period the precision and accuracy of the quality control tank reach WMO compatibility requirements, being 0.01‰ for  $\delta^{13}\text{C}$  and 0.05‰ for  $\delta^{18}\text{O}$ , respectively. "

L58: FTIR has also been developed recently for this purpose, especially at BIPM as is referenced in this paper. Perhaps reword this as "Optical (infrared) spectroscopy now offers this possibility following strong developments in recent years in FTIR and especially for the laser light Sources, ..."

Line 58: “Optical (infrared) spectroscopy now offers this possibility following strong developments in recent years in FTIR and 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).”

Figure 2. From the plots it is clear that the linewidths in the calculated (fitted) spectrum do not match those measured, leading to the typical “second derivative” shaped residuals (and a position error in the case of the 636 line). Is this due to the accuracy of the Hitran widths, or the linewidth model employed in TDLWintel? Could the Aerodyne authors comment, perhaps in the Figure caption. It would be desirable to reassure the reader that the resultant errors in retrieved amounts are systematic and will calibrate out under identical measurement conditions, but they may lead to dependence on such conditions (eg. temperature, pressure) if they vary.

Added in caption of figure 2: “The residuals show systematic deviations at the line positions. These deviations are primarily due to the use of the Voigt lineshape function in the spectral fitting model, rather than a more complex lineshape function such as Hartmann- Tran. Careful analysis has shown that the use of the more convenient Voigt lineshape function does not add noise, drift or calibration error as implemented in the isotope analyser.”

L125: strictly speaking, the r18 and r17 isotope ratios differ from the 628/626 and 627/626 ratios by a factor of 2.

Added as a footnote to line 126: “Note that the r628 and r627 differ strictly speaking from the isotope ratios (r18 and r17) by a factor 2.”

Figure 4: I cannot understand the relationship of the third column in the table (labelled std dev (‰)) to the Y axis on each figure, ppm<sup>2</sup> or ‰<sup>2</sup>. I would expect this to be the square root of Allan deviation at the minimum, in units of ppm (upper plots) or ‰ (lower plots). These are sqrt(AV) at a particular averaging time, not strictly the same as standard deviations. But the numbers do not line up:

For example for 626 in the top left plot, the minimum AV is ~20 ppm<sup>2</sup> at 16 s, so sqrt(AV<sub>min</sub>) = 4 ppm. The given value in the plot table is 0.01 ‰. 4 ppm seems rather high (1% or 10‰ of 400 ppm) for the laser measurement.

Please explain and/or correct the numbers in the table

This was indeed incorrect due to use of the incorrect unit on the y-axis. The numbers of the y-axis were in ppb, while the axis title says ppm. I changed the numbers to the correct (ppm) unit, and now the numbers in the table correspond to the numbers in the graph.

Eq 1. It is confusing to use the same symbol (M) for the measured amounts (RHS) and the ratio to the WG measurement (LHS) – the former has the units of an amount of gas, the latter is dimensionless. Could you be clear in the text by using a different symbol for the

LHS, say M', or R for ratio, or I for index and referring to the appropriate quantity in the text strictly according to which is being used (for example L 235 and following).

The ratio has now been changed to the symbol I as suggested, and this has been adjusted for the use of it in the rest of the text.

L159 et seq: As stated in the text, the std dev of the corrected isotope ratio should be lower than that for the uncorrected ratio if the drift correction is effective, by an amount that depends on the drift, but only after taking into account that the random error in the ratio (M') will be  $\sqrt{2}$  times larger than in either of the individual M values (so if there is no drift, the std dev of M' would be 1.4 times the std dev of M). But the statement line 160-164 that the std dev should not get larger with n is not correct – if the correction procedure were perfect the std dev should decrease by  $\sqrt{2}$  or a factor of 1.4 from n=5 to n=10 measurements. In fact it increases in every case, so the drift correction is not perfect over that timescale. Why?

We show here the standard deviation, not the standard error, so it will decrease not with a factor of 1.4 from n=5 to n=10. However, it is right that in the case of a perfect drift correction, the standard deviation would decrease. This implies the drift correction is effective, but not perfect. This has been changed in the text. Line 165: "It is expected that, if the drift correction is effective, the standard deviations of the uncorrected values are higher than the standard deviations of the corrected values. The drift correction is effective as the standard deviations of the corrected values are always lower than of the uncorrected values. Although the drift correction procedure is not perfect, as we see a small increase of the standard deviation between n=5 and n=10 between 0.005 and 0.012‰, we can still conclude that the drift correction will result in a better repeatability of the isotope ratios."

Further, in Table 1 caption, are these really RELATIVE standard deviations (in ‰ of ‰), or are they actual standard deviations of the isotope ratios in ‰? This was changed from the original version. Please be very wary of using ‰ for a relative value of 2 quantities in isotope work because it is so easily confused with a delta value itself, also quoted in ‰.

These are relative standard deviations, which is now explicitly written in the caption.

L174, L1276: should this be adsorption, not absorption?

Yes, this was changed to adsorption throughout the text.

L188: please refer to comment on Figure 2 above where such a lack of fit is evident in the linewidths.

We now refer to the comments in the caption of figure 2, line 194: "Capturing the true absorption spectrum is very complicated, due to among others line broadening effects of the various components of the air, far wing overlap of distant but strong absorptions, temperature and pressure variability and the choice of lineshape function (see figure 2 and caption).

L197: should this read “when eq 3 is brought into eq. 2...”?

Yes, this has been adjusted.

L198: ... for either or both of those ...

Line 203: “When equations 3 is brought into 2 for either or both of the rare and the abundant isotopologue mole fraction, and  $\beta$  is non-zero for one of those...”

L236: This should read either “The CO<sub>2</sub> mole fraction is calculated by multiplying  $M'_{626}(t)_{dc}$  by the known  $M_{626}(t)_{WG}$  of the working gas” or more strictly “The CO<sub>2</sub>-626 mole fraction is calculated by multiplying  $M'_{626}(t)_{dc}$  by the known  $M_{626}(t)_{WG}$  of the working gas” (These will be the same for the same isotopic composition).

Line 234: “The 626 mole fraction is calculated by multiplying  $I_{626}(t)_{dc}$  by the known 626 mole fraction of the working gas. “

L258: the meaning is not clear – what quantity is “The  $WG(t)$ ”

Line 264: “The  $r^*_{WG(t)}$  is calculated using the same method as  $M_{WG(t)}$  is calculated in equation 1. “

L263: “As measurements were not conducted on CO<sub>2</sub> of similar isotope composition” Not clear, of similar composition to what?

Line 269: As the CO<sub>2</sub> used for the different experiments was not of similar isotope composition, the  $\delta_{636}$  measurements in figure 6 were normalized such, that at the CO<sub>2</sub> mole fraction of 400 ppm all ratios are 1.

L330: The Picarro analysis is based on 626, not whole CO<sub>2</sub> – there is an inherent assumption that the 626/totalCO<sub>2</sub> ratio is constant. This is a potential source of error if the isotopic composition of different reference gases varies significantly, but from Table 5 the variation in isotopic composition is not significant for this error (Griffith 2018). However the narrow range of  $\delta_{13C}$  across the reference gases does not provide a wide span for calibration using the RM – for IM it does not matter. See L390.

Line 342: “The PICARRO analysis is based on the 626 isotopologue mole fraction, not on whole CO<sub>2</sub>. This is a potential source of error if the isotope composition of different reference gases varies significantly. As the isotope compositions of the used reference gases are close (see table 5), the variation is not significant for this error (Griffith, 2018). “

The comment on the isotope range of the reference gases used for the RM is correct, but this is covered in section 3.3.

L370, 371: Tans, Croswell and Thoning (2017) also successfully implemented the IM and should perhaps be referenced here – it is used routinely at NOAA in the generation of reference gases.

Is added to the references.

L584: For the benefit of those reading only the abstract, intro and conclusions, please spell out Ratio Method and Isotopologue method here.

Is spelled out now.