

Answers to referee #1:

We thank the referee for the recommendations and valuable comments that helped to improve the quality of the paper, as well as the compatibility of the data we present. Below our reactions on the specific comments can be found in the **bold and italic text**.

Specific comments:

Suggest that the authors change " λ " in Equation 2 to " θ ". Equation 2 refers to specific kinetic processes with unique isotope fractionation factors. Such physical variables are commonly designated as theta " θ " values in the literature to distinguish them from the slope " λ_{RL} " in Equation 3, which is an arbitrary number.

We agree with this suggestion and changed the λ to θ in equation 2, and, for consistency, changed the R_{IMAU} to θ_{IMAU} in section 2.3.

Strongly suggest the authors recalculate and report the $\Delta^{17}\text{O}$ values using a $\lambda_{RL} = 0.528$ instead of 0.5229. First, water triple oxygen papers use $\lambda_{RL} = 0.528$, and since the composition of air CO_2 is closely linked to water compositions, it is reasonable to use the same λ_{RL} . Second, the triple oxygen isotope community is now adopting $\lambda_{RL} = 0.528$ as a consensus value, independent of the field of study and materials analyzed; see Miller & Pack (2021). Using the consensus value of 0.528 will make comparing the presented data with existing and future literature easier.

We now use $\lambda_{RL} = 0.528$ in all the reported $\Delta^{17}\text{O}$ values, as this will indeed make comparing the data with literature easier. It has to be noted that calibration of CO_2 and water will now give a $\Delta^{17}\text{O}$ of -0.21 ‰ instead of 0.

In line 53, the authors state that O_2 - CO_2 exchange currently provides the highest measurement precision triple O data. It may be worth noting that multiple papers in recent years have demonstrated sub-10 ppm precision for CO_2 measurements using laser spectroscopy (e.g., Bajnai et al., 2023; Hare et al., 2022; Perdue et al., 2022; Stoltmann et al., 2017).

We thank the reviewer for the suggested references. We added these techniques and references and changed the text to (line 52): "These measurements can therefore only be done by measuring ion fragments, requiring a higher mass resolution and a very high sensitivity IRMS system, or by O_2 - CO_2 exchange, a sample preparation procedure that is very labor intensive (Adnew et al., 2019; Mahata et al., 2013). The last method mentioned is at this moment acquiring a precision higher than 10 per meg for measurements of $\Delta^{17}\text{O}$ (Adnew et al., 2019; Liang et al., 2023). Laser absorption spectroscopy measurements of $\Delta^{17}\text{O}$ (next to $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$) on pure CO_2 (Stoltmann et al., 2017) and directly on CO_2 -in-air (Steur et al., 2021; Hare et al., 2022; Perdue et al.,

2022; Bajnai et al., 2023) now reach precisions close to, or higher than the IRMS measurements. ”

To the paragraph starting with line 130: It may be worth noting that Perdue et al. (2022) doesn't observe a shift in $\Delta^{17}\text{O}$ values related to the $p\text{CO}_2$ mismatch between the sample and reference (see their Fig. 8), whereas Bajnai et al. (2023) does, but they correct it by precisely matching the $p\text{CO}_2$ of the reference to the sample (see Fig. 4). It seems that the mismatch in $p\text{CO}_2$ between the sample and reference is the largest source of uncertainty in the presented data. As an outlook, could the authors discuss how to make their measurements more precise?

We have attempted to provide the requested outlook, and new text was added to elaborate on the influence of the CO_2 amount fraction on the measurement uncertainty:

Line 138 “The calibration method used for a sample measurement depends on the CO_2 amount fraction of the sample relative to the references. The uncertainty introduced by the calibration is highly dependent on the difference, in CO_2 amount fraction, of a sample from the closest reference, as well as the difference between the references (Steur, 2023). We calibrate with the reference cylinders only, instead of having an on-line mixing facility where the reference and sample CO_2 amount fraction can be matched (Perdue et al., 2022; Bajnai et al., 2023). Therefore, samples that fall outside the range of the CO_2 amount fraction that is covered by our reference cylinders will have higher uncertainties.”

In line 171 we added the sentence: “Extending the CO_2 amount fraction range of our reference cylinders will improve the measurement precision of samples with elevated CO_2 amount fractions, as well as extend the range of CO_2 amount fractions that can be shown in the results. A way to prevent that a high number of reference cylinders has to be included at all times, is to make the selection of references more dynamic. As sample measurements are always alternated with a working gas measurement, it is possible to do a 1-point calibration immediately after a sample is measured. In this way it will be possible to select the ideal set of references to calibrate the samples based on the CO_2 amount fractions derived from the 1-point calibration. This would save reference gas, as well as measurement time of a measurement series.”

Bajnai et al. (2023) observed the dependence of the TILDAS- $\Delta^{17}\text{O}$ data on the measurement temperature. While the reference bracketing method used in the presented dataset likely addressed such temperature variations, the authors could further increase the credibility of their data by discussing this effect in the paper.

The reference bracketing method should indeed correct for instrumental drift caused by temperature variations. Although the temperature effect specifically has not been studied for the SICAS, we do provide analysis of stability of the measurements over a period of 2 years in Steur (2023). The standard deviations found in this analysis are included in the combined uncertainty of all our measurements. We therefore do not separately address the temperature effect in this manuscript.

In lines 267 and 336, the authors argue that CO₂-enriched signals are due to the contribution of fossil fuel emissions. In this case, one would expect to see correlations between $\delta^{13}\text{C}$, $\Delta^{17}\text{O}$, and $p\text{CO}_2$. Can the authors underline their statements with data and possibly additional figures?

Based on this request, we have evaluated the available data in more detail and decided to change our text to weaken the statement. The $\delta^{13}\text{C}$ and CO₂ amount fraction logically correlate which is also clear from figure 4, but a Keeling plot (not shown) did not distinguish natural from anthropogenic sources in this region, as signatures are too similar. We added a new figure in the Appendix showing the $\Delta^{17}\text{O}$ summer and winter values of Lutjewad plotted against $1/\text{CO}_2$ and $\delta^{13}\text{C}$. We see no correlation in these plots, and we therefore make no conclusions on the exact CO₂ sources in the Lutjewad record. The statements were adjusted in line 278:

“The Lutjewad flasks, although sampled at noon with the aim to sample well-mixed tropospheric air, occasionally show large positive deviations from the background curve, especially in winter, of up to +47 $\mu\text{mol/mol}$ in December 2017. The CO₂ enriched signals are most probably due to local and regional sources of CO₂, either natural or anthropogenic, that occur on the continent. We therefore expect to see more deviations from the seasonal cycles of stable isotope values induced by the more continental influence at the Lutjewad record when compared to the Mace Head record.”

And in line 350: “The most important difference between the Lutjewad and Mace Head $\Delta^{17}\text{O}$ records is the presence of more depleted values in the Lutjewad record, with the lowest value being -0.43 ‰ in the summer of 2022. CO₂ equilibrated with water following λ_{RL} will have an $\Delta^{17}\text{O}$ of -0.21 ‰. In summer, leaf water gets enriched in oxygen isotopes, and depleted in $\Delta^{17}\text{O}$ as the result of high rates of evapotranspiration (Landais et al., 2006). Due to the active biosphere during summer, CO₂ and leaf water will equilibrate and the depleted $\Delta^{17}\text{O}$ signal will be translated to the CO₂ (Adnew et al., 2023). We estimated that this could result in $\Delta^{17}\text{O}$ values being up to 0.1 ‰ more depleted, when assuming the minimum θ of 0.516 for evapotranspiration (Landais et al., 2006), and considering the range of $\delta^{18}\text{O}$ values that were measured in our Lutjewad record. For the full estimation we refer to Appendix E. $\Delta^{17}\text{O}$ values up to -0.31 ‰ can be explained by this process. CO₂ emitted from combustion processes has very negative $\Delta^{17}\text{O}$ values (Laskar et al., 2016; Horváth et al., 2012). All points that have lower $\Delta^{17}\text{O}$ than -0.3 ‰, and are sampled during winter/spring, have more depleted $\delta^{13}\text{C}$ values and more enriched CO₂ values than would be expected from the seasonal trends. This indicates that local CO₂ emission sources are the reason for the more depleted $\Delta^{17}\text{O}$ values in winter. Samples that are very enriched in CO₂ amount fractions are not shown here, as that results in very high measurement uncertainties. This could be the reason that a correlation of $\Delta^{17}\text{O}$ and CO₂ amount fractions does not appear in figure D1. A few points show depletions lower than -0.31 ‰ without CO₂ amount fraction enrichments, and remain for now unexplained.”

In the paragraph starting with line 310, the authors argue that they should be able to resolve a 130 ppm annual variation in $\Delta^{17}\text{O}$, as observed by Hoffman *et al.* (2017). However, their argument that their uncertainty of ± 100 ppm is lower than 130 ppm is misleading and needs to be revised. Instead, the authors should take into account the signal-to-noise ratio and the number of measurements to determine what cyclic signal can be resolved in their time series.

The uncertainty of our measurements is (on average) 70 per meg. The 100 per meg is the range in which the majority of the points fall. We agree with the reviewer that this is a confusing statement, and we changed it to the following:

Line 320: “ $\Delta^{17}\text{O}$ measurements from the Lutjewad and Mace Head stations are presented in figure 6. The total range in the Lutjewad and the Mace Head record is 0.5 and 0.2 ‰, respectively, with an average combined uncertainty of the measurements of 0.07 ‰ for both records.”

In line 315, the authors write that the amplitude of the seasonal $\Delta^{17}\text{O}$ signal in Göttingen is larger due to a stronger biosphere signal. This is an important statement in comparing the presented record with existing data and thus should be expanded upon. Would the 3-D model used in this paper be able to reproduce the 130 ppm signal observed by Hoffman *et al.* (2017)?

We thank the reviewer for this suggestion. We now compared the amplitude of the model simulation from Hoffman et al. (2017) for the Göttingen location and the model simulation of Lutjewad, conducted with the model described in Koren et al. (2019). The results were added as the following text:

Line 330: “The amplitude of the seasonality that was determined from the Göttingen $\Delta^{17}\text{O}$ record is (0.13 ± 0.02) ‰. If such a seasonality would be present in the Lutjewad and Mace Head record, we would expect to see it, as this signal is higher than the average combined uncertainty of the SICAS measurements. It can be, that due to the more continental location, the amplitude of the $\Delta^{17}\text{O}$ seasonality is higher at the Göttingen site reflecting a stronger biosphere signal. A model simulation of the Göttingen location shows an amplitude of 0.045 ‰ (Hofmann et al., 2017), while the amplitude of the simulation at the Lutjewad location, shown as the black line in figure 6 is close to 0.025 ‰. The model used in the Hofmann paper (2017) is an earlier version of the model used in this study (Koren et al., 2019), so the results should be well comparable. The higher amplitude for the simulation of the Göttingen location confirms the hypothesis of a higher $\Delta^{17}\text{O}$ seasonality due to the more continental location in comparison with Lutjewad. It is unlikely that a lower seasonal signal than observed at the Göttingen location would be detected by the SICAS measurements considering their average combined uncertainties”

The following changes are suggested for Figures 4, 5, and 6: The range of the top and bottom plots should be the same, which will help the reader make visual comparisons easily. The measurement locations should be written above the curves

and not on the vertical axis label. The coloring of the $\Delta p\text{CO}_2$ should be changed to a diverging, color-blind-friendly color scale.

We thank the reviewer for these suggestions that we all applied to the plots.

Please note the following suggestions for improving the visibility of Figure 7: The vertical year-markers should be made thinner so that they don't clash with the data and error bars. The red trend should be plotted accurately without any shift by 0.08‰ to avoid confusion. Moreover, the horizontal axis grids, similar to those in Excel-made figures, are unnecessary for any plots.

The changes suggested here were also applied to the plot, and the colors of the lines in all plots were changed to a color blind friendly palette.

Suggest adding Carlstad & Boering (2023) to the list of references in line 15.

The reference was added to the list.

The sentence in line 437, "A better precision...", is without precedence in the text. The authors may consider either expanding on it or removing it.

This sentence was removed.

Correct the spelling in line 87: "continues".

The spelling was corrected.

References cited in this