

## Anonymous Referee #2

*The authors would like to thank the Referee for taking the time to comment on our manuscript. His suggestions and valuable comments gave us the opportunity to improve and clarify the manuscript. The authors' replies (in dark blue) are given below individually with the Referee's comments.*

### Comments:

1. The authors are brief in their discussion and interpretation of the model comparison. What new insights do the isotopic data add? Isotopes are introduced as 'ideal tracers of sources and sinks' (p2,l22) in natural systems, yet the presented research is focused on anthropogenic emission in winter, for which mostly the concentrations provide the information, less so the isotopes. For instance, it would be interesting to learn if the detailed isotopic data can tell us if the relative contribution of fuel/wood combustion contrast clusters representing warmer (south, cluster 1+5) or colder synoptic weather (Atlantic and continental land mass; cluster 2,6,8). Instead, the authors conclude  $^{14}\text{C}$  information is needed instead (p19l7–l13). Thus, we are shown that it is possible to have a highly detailed isotopic record of  $^{13}\text{C}$  and  $^{18}\text{O}$  of carbon dioxide for high altitude sites, but not offered a very strong argument for its relevance.

In this manuscript, we focus on the technical aspects of long-term spectroscopic measurements at a high altitude site, present a survey of the major factors influencing the analysis, and explore the capabilities of such a monitoring approach. While detailed data interpretation is not our primary goal, we also present model simulations to demonstrate the plausibility of the obtained data and to assess the relationship between atmospheric transport regimes and air composition. A more detailed analysis of the  $\text{CO}_2$  isotope data is planned for the future. It will make use of related tracers such as  $\text{CO}$ ,  $\text{COS}$ ,  $\text{O}_2/\text{N}_2$  that are also monitored online at Jungfraujoch. In the present work we highlight the benefits of high temporal resolution, which is important to reliably determine representative background values of  $\text{CO}_2$  and its isotope ratios. On hourly and diurnal time scales there is considerable variation since the site is intermittently influenced by free tropospheric and atmospheric boundary layer air masses. Only sufficiently high time resolution allows relating the observations to high-resolution transport models, regional scale ecosystem models, and regional scale flux inversions. Additionally, we expect that with the increased data quality and as the measurements are continuing, we will be able to analyze distinct events and tracer relationships depending on the weather conditions and attribute these to specific emission/uptake locations and processes. The high temporal resolution time-series contain a large amount of interesting details (e.g. diurnal variations in summer, emission specific classifications, etc.), but their interpretation require an isotope specific, regional model which has yet to be implemented. Finally, note that we do not conclude that  $^{14}\text{C}$  is needed instead of stable isotopes, but rather argue that additional tracers such as  $^{14}\text{C}$ , but also  $\text{CO}$  or  $\text{COS}$ , are complementary and contribute to our better understanding of underlying processes.

We added to the conclusions a more explicit discussion of these points to better highlight the relevance of our data.

2. With the presented study and previous experience, the authors could more prominently discuss the implications of the standards set by WMO (p3l2) for inter-laboratory compatibility when using state-of-the-art spectroscopic instruments. Can this be achieved at field sites in challenging conditions? If not, what are the options for more tuning of the setup or further technical improvements?

The goals for inter-laboratory comparability defined by the WMO ( $\pm 0.01\text{‰}$  for  $\delta^{13}\text{C}$  and  $\pm 0.05\text{‰}$  for  $\delta^{18}\text{O}$ ) are very challenging and various inter-comparison programs and round robins showed that it is extremely difficult to meet these goals, in particular for  $\delta^{18}\text{O}$  where the inter-laboratory precision is still far off the goal (WMO, 2011). This is primarily a question of calibration (how to prepare/store/analyze calibration gases, how to maintain the link to VPDB, how to maintain consistency over time, etc.) rather than of the measurement technique. We consider a detailed discussion of such general calibration issues beyond the scope of the present study. With the improvements that we have implemented, our instrument precision approaches the required level. However, under field conditions it is even more difficult to meet the WMO goals and often it becomes a trade-off between highest comparability and practical considerations such as calibration gas consumption. Nevertheless, we are convinced that instrumental improvements will further increase the long-term stability of the measurements, e.g. by further reducing the temperature sensitivity of the instrumentation. It seems thus realistic to reach the requirements set by WMO, even under field conditions at affordable cost.

3. The authors have given much thought to the schedules of calibration and measurement time. The trade-offs can lead to complications that are not discussed in much detail. For instance, a 120 s average for a reference gas is used as a benchmark for 10' averages (p12|26) and shows SDs that are higher than the reported performance. This gives a somewhat misleading image of the measurement data quality, for which the benchmark values at any time will lie between the Allan variance statistics and the variability of the 2' averages. Please explain why the measurement of reference gas was not made for a duration that is aligned with the optimal 10' averaging period for the measurements. Would periodic change from 2' to 10' make a more useful benchmark or do you advise against it (and why)?

In accordance with the Referee's suggestion we added a more detailed discussion about the trade-offs implemented in our sampling/calibration strategy. In summary, the target reference gas is actually measured for 5 min, but given the transition and flushing time of the spectrometer, we discard the first 3 min and only consider the last 2 min measurements for the evaluations. Although, the precision in this case is by a factor 2.2 lower than the ultimate instrumental precision (0.02 ‰), it is still within the uncertainty of our primary standards (0.06 ‰). Certainly, we could slightly improve the dispersion of the target reference gas data (Fig. 4), but at the expense of significantly more reference gas consumption (13 min vs. 5 min), which would involve their replacement every 5 instead of 12 months. We consider that the effort to produce, transport and reference the calibration gases, and to assure continuity in the measurements outweighs the potential 0.02‰ improvement.

Minor comments:

1. p7|1: 'traceability of the isotope data remains a challenge' what exactly do you mean?

We mean that it is difficult to remain consistent over time and to maintain high reproducibility and comparability at remote sites, because traceability of calibration gases (i.e. the link from VPDB- $\text{CO}_2$  to JRAS (Jena Reference Air Sets) to our primary gases to our secondary gases) as well as drifts in primary and secondary calibration gases need to be periodically reassessed. We adapted the relevant paragraph.

2. p7|7: Manufacturer of the carbon sources gases?

PanGas AG (Switzerland)  $\text{CO}_2$  4.5 (10L) from carbonate sources.  
Messer Schweiz AG  $\text{CO}_2$  4.5 (10L)

The information was added to the text.

3. p8|7: are these causes in order of importance?

The order of importance can differ. There are times when the temperature was relatively stable, but the laser frequency shifted or times with stable laser parameters but large temperature fluctuations due to problems with the air conditioning in the lab. We clarified this in the text.

4. p13|14: Would regression (slope,  $R^2$ ) not be a better indication of how these estimates correlate than their mean difference?

The mean difference alone would indeed not be a good indication of the correlation; however, we also indicated the standard deviation of the difference, which gives a measure of how well the estimates agree. The regression slope and  $R^2$  would be another option to characterize the correlation. However, the mean difference and the residuals of the regression are highly autocorrelated and therefore also a regression analysis needs to be interpreted with care. The regression results (slope and  $R^2$ ) are  $1.054 \pm 0.001$ ,  $R^2=0.98$ ,  $1.037 \pm 0.001$ ,  $R^2=0.97$  and  $1.024 \pm 0.001$ ,  $R^2=0.98$  for  $\text{CO}_2$ ,  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$ , respectively. Here, the standard errors of the slope are most likely underestimated due to the strong autocorrelation. Still, we added the slope and  $R^2$  values in the text for completeness.

5. p18,|7: What defines remote? A remote location could still be one with a laboratory environment, stable mains power and temperature regulated by heating (for readers not familiar with the site). The challenge seems to have been the high-altitude, low pressure conditions.

The Referee correctly points to the major elements of what we consider as a remote site. Although the research station at the high-Alpine mountain site is a well-equipped laboratory (<http://www.ifjungo.ch/jungfrauoch/infrastructure/sphinx.html>) the low pressure and the strong temperature variations are evidently the most challenging parameters. Furthermore, the available space for instrumentation is also a limiting factor, which strongly influences the amount and size of consumables (gas cylinders) and peripherals. Another important aspect of the remoteness relies in its unique geographical location, being far enough away from direct pollution sources, close to the free-troposphere, but still in the middle of continental Europe.

In the online version of our AMTD manuscript we did not use the term “remote” at the specified location.

6. Title: ‘Exemplary results’ The data are extraordinary, but is the qualification ‘the best of its kind’ justified in the title? Do you mean ‘case study’ or ‘result highlights’?

With “exemplary” we meant “as an example” and we did not intend to imply “best of its kind”. We have changed the title (see next comment).

7. Title: ‘Jungfrauoch’ may not be known well enough to be read as synonym for ‘high elevation’ site.

We changed the title to: “Tracking isotopic signatures of  $\text{CO}_2$  at the high altitude site Jungfrauoch with laser spectroscopy: analytical improvements and representative results”

8. Fig.1. The temperature stabilization in the diagram suggests heating of the sample/gas inlet tubes. The text mentions stabilization of the optics (p617). Please clarify.

Both statements are correct. The optical module is temperature stabilized, while the gas stream (reference/sample) is preheated to match the optical module temperature when entering into the multipass cell. This strategy proved to be very efficient to minimize temperature perturbations and thus maintain a stable operating temperature of the optics.

9. Fig.2. This was determined at the Jungfraujoch site? Not mentioned in the text.

Yes, all the performance characterization measurements were done on site. We added this remark to the text.

10. Fig.7. Please add statistics to these means

The error bars (both the standard error as well as the 95% confidence interval derived from bootstrap resampling) are smaller than the plot symbols and therefore not shown in Figure 7. We added this clarification to the figure caption.

11. Offset correction: "one-point (offset) calibration, drift cylinder, reference gas, drift standard, cylinder B or C" Please check if this can be simplified. Several terms are used for the same thing and it is not always clear from the text (maybe cited works) if a correction is made to the carbon dioxide concentrations only, or also the stable isotope ratios.

We have simplified the description of our calibration gases (also in accordance with the comment of Referee #1). In particular, we avoid using the terms "cylinder B", "cylinder C".