

General.

This is a very thorough paper outlining a novel method for dealing with one of the main challenges associated with using laser-based spectroscopy on very small samples- namely the concentration dependence of isotopes. It is a significant body of work that provides a framework for evaluating performance challenges associated with calibrating CRDS systems under a variety of conditions to make the most precise and accurate measurements. The authors provide several applications, including the measurements of inclusions in stalagmites which provides a very concrete example of how this methodology could transform the current methods used for the analysis of fluid inclusions. This paper has significant impacts on the future design and implementation of critical measurements of water vapor isotopes both in the lab and in the field.

Laser-based instruments have provided an enormous step change in our ability to make important measurements in all aspects of the hydrosphere and the atmosphere. But our ability to thoroughly calibrate and account for a number of artifacts, including a mole fraction or concentration dependence for the isotopes; sample carry over; and interaction with surfaces to which the sample is exposed along the measurement route, has remained a challenge. This paper addresses a number of those.

The authors have shown the details of applying existing technology from another field (inkjet printers). While this technology has appeared in previous publications – notably the 2008 paper by St. Clair et al. which is referenced by this paper, this remains to be an important next step in validation and proof of concept for its application with isotopic calibration on small sample sizes. It goes well beyond the scope and purpose of St.Clair (2008).

Key findings include how the micro-dropper device is able to produce a water background of constant isotopic compositions that can be used to obtain accurate and precise $\delta^{18}\text{O}$ and $\delta^2\text{H}$ measurements over a variety of concentrations, controlled by parameters of the Dispenser Heads. The inclusion of the comparison to the Standard Delivery Module (SDM, Picarro Inc, Sunnyvale, USA) is a good benchmark and will no doubt be of interest to readers.

The examples of the fluid inclusions would appear to provide validation of methodology that would stand to be an improvement over current methods of measuring fluid inclusions. While a comparison with samples analyzed with mass spectrometer methods would have been useful, the validation provided by the analysis of standards introduced via capillary pipets is convincing.

Specific

Line 51: In the introduction, a discussion of similar techniques used to introduce and vaporize liquid water and highlights those used for CFA analyses (Gkinis et al., 2010). To be complete, consider also referencing Jones et al. ,2017 which employs a nebulizer, that solved some of the challenges found in the WVISS system outlined in Aemisegger et al., (2012).

Line 99: A long with Figure 1, a picture of at least the apparatus in general and/or the critical components (eg. The microdrop dispenser heads; evaporation chamber) would greatly help.

Line 103: With regard to the use of the Microdrop, dispenser head (DH), JM St. Clair et al., 2008 (Reference on line 575) uses the Microdrop part MD-K-130-010, while the authors chose to use Microdrop GmbH, Germany, Part Nr. MK-K-130. It could be useful to explain the choice of microdrop dispensers, and whether a heated or non-headed head was used and why.

Line 150: Clarify “To this end, we first calibrate the raw mixing 150 ratio signal with a calibration curve obtained from a dewpoint hygrometer.” Consider adding “To this end, we first calibrate the raw mixing 150 ratio signal *measured by the Picarro* with a calibration curve obtained from a dewpoint hygrometer.” Also, it would be helpful to name the brand/model dewpoint hygrometer.

Line 154: Performance characteristics for DH1 and DH2 are mentioned, but not DH3. Consider mentioning why these two were selected for testing, and not DH3. Perhaps because these two were the ones chosen for the system, based on performance? Also, can operating frequencies and dispersion parameters be specified, or are they DH dependent, and constitute unique operating parameters to be determined for each unit?

In the section on the application of this method, it would be useful to have comparisons of this method vs. the traditional extraction/ mass spec method as validation of the method.

Line 332: A significant side bar finding of this study is the statement that “ We find that different analyzers require different polynomials for fitting in the $\Delta\delta$ and the $\ln(x)$ direction. This may come as a surprise to some adopters of CRDS systems for isotopic measurement and is an important finding.

Line 370: To my knowledge, Affolter et al., (2014) does not mention memory (or sample carry over) explicitly. So, it could be appropriate to address memory in this paper more directly? Also, does sample carry over affect $\delta^{18}\text{O}$ differently than δD in this system? Deuterium excess would be an interesting parameter to include if possible.

Line 440: Long term sustainability and reliability are important qualities of any system like this. Consider expanding upon the statement “While in general working reliably over several hours at a time, these problems are important to be aware of, and one needs to search for remedies in future development of the calibration system”. This is an important step forward. Knowing how long this system can be expected to function reasonably well without manual intervention is useful. I would encourage authors to speculate on areas where reliability might be improved to potentially lengthen the time it can run unattended.

Line 502. It is stated that “Prominently, Synflex tubing has been shown to render $\delta^2\text{H}$ measurements meaningless”. This (rather important) conclusion does not appear to be

supported by evidence provided in this manuscript. If this has been shown elsewhere, a reference would help.

Appendix A: Crusher application details: Figures A1 and A2 are referred to in the main text, so I am not sure why they are in the appendix, unless there are space limitations? No details beyond figures are found here. Consider either moving appendix figures to the main body of the text or changing to “Additional crusher application figures”? Editorial call.

Some small Typos or wording:

Line 17: needs period at end of sentence.

Line 268 & 276: Not sure which version AMT prefers: English version of ‘artefact’ vs. artifact. Editor’s choice?

Line 322: 20’000 should be 20,000.

Line 349: og -> of

Figures

Figure 1. This schematic explains the set up adequately, but it would help tremendously to add a figure contains images of the micro-drop heads, and/or the evaporation chamber, as this appears to be the heart of the innovation here and deserves some focus.

Figure 2. These images are very helpful. Consider adding to the caption how the images were obtained or with what optical imaging system.

Figure 3. See comment for line 154. Consider adding a note here on why just 2 of the 3 heads were tested. Perhaps 3 were evaluated and two selected for the vaporizer. In 5 b) there are many fewer points above about 10 Hz for DH1. An explanation would help.

Figure 4. Both a) and b) show bimodal distribution. According to the text, the day-to-day standard deviations are shown in red and blue. Grey histograms are for performance over the entire time period. The difference is not large (~9 or 10 μ) but groupings are distinctly different. An explanation of these groupings might help. Perhaps it is due to the physics of discrete droplet formation that the volume (or diameter) of droplets is not a continuum, and is constrained by the droplet formation process itself? This might be obvious to those familiar with the DH technology, but a sentence of explanation would help those who are not.

Figure 5. Allen deviation plots are very helpful, as is the comparison with the SDM.

Figure 8. This is an important figure. Consider adding a panel o

Fig. 11: This is an excellent and key figure, showing the comparison of correction functions of Weng et al. (2020). The caption refers to dashed grey lines in the figure – but there are no

dashed lines as mentioned as far as I can tell, but there are possibly blue and black lines? (I am thinking the caption may have been created with a previous version of the plots?) Either make the colors more distinct, or make one set of lines dashed, as the caption reads.

Figure 13. Excellent figure. IF I understand this method correctly, peak values are integrated between the green markers, which show the start and the end of the peak. For both $\delta^{18}\text{O}$ and δD there is a significant drop (lighter isotopically) just as the peak in mixing ratio arrives in the analyzer. Presumably, integration between the green markers (including this dip) yields the correct value for water standards introduced with capillary tubes? Maybe I missed something about the calibration? Based on the performance of the analyzer as shown in figure 9, It would seem that integration of the sample might be more stable later? Maybe sentence in the text or caption could address this.

Figure 14. Image of glass pipets is helpful. Presumably the glass pipets are introduced in the crusher unit, if so, this could be mentioned in the caption.

Figure A1. Caption could be improved by listing the names of the internal (or inhouse, secondary) standards and their assigned values with \pm uncertainty, and explanation of the analysis method used to assign their values to the VSMOW-SLAP scale. Perhaps adding "note y-axis scales differ" would also help.

1. Does the paper address relevant scientific questions within the scope of AMT? **YES**
2. Does the paper present novel concepts, ideas, tools, or data? **YES**
3. Are substantial conclusions reached? **YES**
4. Are the scientific methods and assumptions valid and clearly outlined? **YES**
5. Are the results sufficient to support the interpretations and conclusions? **YES**
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? **YES, with some caveats as noted in comments**
7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? **YES, mostly except where noted.**
8. Does the title clearly reflect the contents of the paper? **YES**
9. Does the abstract provide a concise and complete summary? **YES**
10. Is the overall presentation well-structured and clear? **YES**
11. Is the language fluent and precise? **YES**
12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? **YES**
13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? **YES, minor edits are recommended, but nothing major.**
14. Are the number and quality of references appropriate? **Generally, YES**
15. Is the amount and quality of supplementary material appropriate? **YES**