Atmos. Meas. Tech. Discuss., 5, C659–C667, 2012
www.atmos-meas-tech-discuss.net/5/C659/2012/

Atmospheric
Measurement
Techniques
Discussions

Interactive comment on "Measuring variations of δ^{18}O and δ^{2}H in atmospheric water vapour using laser spectroscopy: an instrument characterisation study" *by* F. Aemisegger et al.

Anonymous Referee #2

Received and published: 11 April 2012

The manuscript presents an extensive intercomparison of four different commercial water vapor isotope analyzers from two different suppliers. This study is potentially interesting to the user community of such laser-based analyzers. However, I must say that I am surprised that such studies are still being carried out at the individual laboratory level. It appears to be high time that the community, including the manufacturers, gets together to agree on appropriate measurement protocols. In the past two years a fair number of publications have appeared that attempted to contribute to this aim by describing one or more measurement and calibration strategies for one or more different analyzers. These studies all have in common that they are carried out by end users of the instruments who not necessarily have a thorough understanding of the operating principles of the devices (evidenced by too often erroneous ‘explanations’ of the spectroscopic basis for the isotope ratio measurements or other inaccuracies). The current manuscript appears to be a positive exception to this observation (containing as it does only minor inaccuracies), but still suffers from a related, frustrating, shortcoming: Since the inner workings of the instrument are largely hidden from the end-user, it is practically impossible to obtain a detailed understanding of the physical reasons for the observed effects, such as the water concentration dependence of the isotope ratio. I share referee 1’s opinion that overall the article is well written, but would greatly benefit from more concise formulation and presentation. Specifically, the inclusion and semi-separate discussion of preliminary results obtained with two newer versions of the two principal instruments is at times a bit confusing. It may be an option to move the discussion of these newer and still incomplete results to a dedicated section. The strict separation between the experimental (3. . .) and discussion (4. . .) sections is perhaps commendable from a principled point of view, but personally I prefer to have my cash on the barrel. Discussing the implications of the measurement immediately, avoids forcing the reader to go back and forth between the two sections. The discussion section itself can then serve to collect the different observations in a coherent view and to formulate recommendations. This will also reduce the need for duplication in the text and allow for a significant shortening of the text.

Specific comments and suggestions (page and line numbers refer to print version of manuscript):

P. 1597 L. 1: The title could be more specific by explicitly mentioning the commercial nature of the laser-based spectrometers and the comparative nature of the study: “. . . using two commercial laser-based spectrometers”.

P. 1600 L. 14: The references to Baer 2002 and Crosson 2008 are inappropriate here, as they do not refer to water (vapor) isotope analyzers. I suggest referring instead to the company websites.

P. 1601 L. 23: replace “permil” by “relative” and “approved” by “accepted”.

P. 1602 L. 4-9: The formulation is imprecise: VSMOW is not just a reference standard, it is the internationally accepted primary standard defining the scale zero. R_standard is not a reference standard but rather the relevant isotopic (atomic, not molecular . . .) ratio of the primary standard VSMOW. SLAP is also a primary standard, used in combination with VSMOW to define the isotope scale (according to the IAEA guide lines), but it is not a normalization factor. . .

P. 1602 L. 12-18: The term “atomic mass spectrometry” is inappropriate here, as in Isotope Ratio Mass Spectrometry (IRMS) the measurement is made on the molecular species. “Excitation energy levels” is a meaningless term when used without further qualification. What is different here is the rotational-vibrational energy level structure of the different isotopic molecules, leading to isotope characteristic transition frequencies in the near-infrared region of the spectrum. For completeness, a reference to reviews of the two methods could be given, for example the relevant chapters in the Handbook of Stable Isotope Analytical Techniques by De Groot.

P. 1602 L. 26: The reason that your laboratory has so far been working exclusively with the LGR analyzer is not a good reason to consider it the benchmark system. If anything, the “Golden Standard” in isotope ratio measurements is still IRMS, however imperfect (especially for water vapor measurements). In any case, both the LGR and Picarro systems are benchmarked independently from each other against a series of working standards, and there is also no need to express the performance of one in terms of that of the other.

P. 1603 L. 6-10: Concerning the comment by referee 1: The cavity can be considered “high finesse” whether used in an on-axis or off-axis configuration. I have been informed that the finesse is defined for one family of transverse modes, not simply as the distance in frequency between neighboring transmission peaks (see, e.g., Ch. 14 of Lasers by A.E. Siegman). Thus even if the number of transverse mode families is increased drastically, the finesse of the cavity does not decrease, and neither does the width of the transmission profile. As the number of excited transverse mode families increases to the extent that neighboring peaks start to overlap within the width of the laser profile, the cavity becomes resonant irrespective of the precise wavelength of the laser. This is facilitated by the use of astigmatic (not astigmatized!) mirrors.

P. 1603 L. 13: When one is not thoroughly familiar with the operating principles of the techniques, it may be better to refer to the appropriate literature. However, the reference given here (Iannone 2009) is not a primary reference. The appropriate reference in this particular case appears to be Paul, Lapson, and Anderson, Appl. Opt. 40 (2001) 4904 (but note that they, and their figure 1 in particular, suggest the interpretation of the FSR as given by referee 1).

P. 1603 L. 21: I do not see why the pressure and temperature cannot be precisely regulated in the case of a fast gas exchange rate. It seems more likely that the Picarro systems operate with a low gas flow rate for rather different reasons; for example to avoid turbulence, which would induce noise on the spectra by its modulation of the refractive index and/or simply out of a desire to use a small, low power pump.

P. 1604 L. 19: Lis et al. (2008) did not invent, or were the first to demonstrate, the combination of a liquid water autosampler and a laser-based water isotopic analyzer. This is again not a primary reference. At the very least, include “see, e.g., “ in the reference.

P. 1604 L. 23: For completeness include Lee et al. (2005) in the list of references.

P. 1604 L. 24: replace “specific” by, e.g., “controlled” or “well-known”.

P. 1604 L. 28: “. . . are precisely regulated or measured”. BTW: knowledge of the pressure is only required for an estimate of the water mixing ratio, not for the vapor phase isotope ratios.

P. 1605 L. 15-20: The Picarro standards delivery module (SDM) introduces a continuous stream of water into the vaporizer. Evaporation occurs at the meniscus of the water surface at the end of the needle. There is no droplet formation!

P. 1606 L. 2: Note that Figure 1(a) gives the wrong liquid water flow rate.

P. 1606 L. 21: Given that the LGR WVISS allows for only one standard delivery at a time, it is at least curious that this device was used instead of the two standards delivery module of Picarro, which would allow for changing the second standard (bag) for another during the delivery of the first standard. What characteristics made you decide to use the WVISS? How much time does the changeover take? How long does it take for the WVISS to purge the delivery lines and was this time included in the two minutes of discarded data at the beginning of each cycle, or does it still need to be added to this? Ten times 10 minutes represents only about 20 percent of the total of 8 hours used for each calibration run. What happened in between?

P. 1606 L. 26: If I understand correctly, the so-called "drift" standard was not used to remove instrumental drift from the measured data, since these were after all known (working) standard values. Rather, its use was to purge the WVISS water delivery lines. If this is indeed the case, I suggest using a different name for this purge liquid.

P. 1607 L. 8: Why would you use different calibration materials for the laser and IRMS measurements?? I cannot think of any good reason to do so. Nor can I think of any good reason to use the precious VSMOW and SLAP primary standards directly in the laser instruments, as this goes directly against all IAEA guidelines, and makes no sense considering the use of WS11 and WS12 for the IRMS analyses. It practically invalidates the intercomparison of laser and IRMS, at least outside the interval spanned by WS11 and WS12.

P. 1608 L. 2 and L. 22: Why do you revert to the Picarro SDM for the measurements of section 3.3 and than back again to the LGR WVISS in section 3.4? In the manuscript I found no better argument than the nice alternation between the two. But what if you had decided to present the isotope stability measurement before that of the water

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

concentration dependency...? I think that most readers, like myself, are rather curious to learn what motivated your choice of calibration unit for each specific task. In fact, it would be very interesting to compare the performance of the calibration units, in addition to that of the analyzers themselves. Without some kind of deliberation, this choice appears completely random and illogical. After all, unless all measurements are made with the "best" calibration unit, the only logical choice would be to test the analyzer with the calibration unit of the same manufacturer, which is very likely the most common configuration found with other end-users.

P. 1609 L.12 (Figure 8): It would be most instructive to present the Allan plots of Figure 8 together with the time series data used for its calculation. In fact, for anybody who has generated such figures him- or herself it is evident that the shape of the curve depends, sometimes significantly, on the exact beginning and end point of the time series, as well as its exact length. The curves tend to change from measurement to measurement, which is not captured in the given plots, as they lack an error estimate. In the absence of a proper, analytical error estimate the best one can do is to calculate the Allan deviation for a linear instead of logarithmic time step. Although in the latter case neighboring data points will be correlated, at least the spread in the Allan deviation values becomes visible. Including a correct estimate of the error on the determination of the (minimum) Allan deviation will have repercussions for the "optimal" numbers reported in Table 6, which are without a doubt too optimistic. In fact, a hint of this is given by the sharp and very fast fluctuations seen in figure 9. Some of the fast and large changes in the "bias correction" value appear to occur on a time scale comparable to that of the optimum averaging times reported in Table 6. If such a sharp transition were included in the time series used for the calculation of Figure 8, the reported short-term precision would be rather worse. How do I interpret the results of Figure 9 in light of the two incidental calibration curves of Table 4? The latter shows that the raw d2H_VSMOW values reported by the L1115-i differ by some 11 per mil between the two calibration runs. This is about 10 times larger than the largest excursion seen in Figure 9! For SLAP the situation is another factor of 5 worse still, given the large differences in normalization slopes. Based on the results of Table 4, the 3-day calibration curve of Figure 9(a) could be expected to be dramatically worse than that of Figure 9(b). That this is not the case may indicate a problem with the calibration performed for Table 4. It seems plausible that due to its slow response, the Picarro analyzer is more prone to memory effects. If this is the case (you subscribe to this in section 4.1, P 1613, L. 18), the measurement strategy should have been adapted to take this effect into account and the conclusion would also be that the Picarro analyzer cannot be used to faithfully capture fast (< 1 min) isotopic ratio variations.

P. 1609 L. 16: best or optimal precision is minimum precision, not maximum precision.

P. 1609 L. 23 (and P. 1622 L. 25): I cannot agree with the statement that the Allan variance analysis is not a useful tool for very long averaging times. All depends on the time series, and especially its length, used to perform the Allan variance. Beyond the optimal averaging time the Allan plot generally shows the onset of instrumental drift, which provides extremely useful information! In fact, in order to assure the most precise, as well as most accurate measurements over longer periods of time, it is essential that the complete calibration procedure (i.e., both sample and reference material (working standard) measurements, and any necessary purging or cleaning of the delivery lines and analyzer) needs to be completed within the optimal averaging time determined by an Allan variance analysis. This was also the consensus of the EGU Workshops on laser-based isotope ratio methods, organized by Kerstel, Gianfrani, and Werle in 2009 and 2011. If the actual averaging time extends beyond the optimal value, the Allan plot provides an estimate of the increased standard deviation of the measurement due to the effect of drift. I would argue that this is highly useful information. For example, the d2H Allan plots predict by extrapolation that the precision of data averaged over approximately one day (∼10ˆ5 sec) will be of the order of 0.3 per mil, due to instrumental drift. This is indeed in agreement with the general trend observed in Figure 9. However, it is also clear that the Allan plots of Fig. 8 present by far too optimistic values of the achievable precision. If these plots were representative of the overall performance of the instruments under all laboratory conditions encountered during the tests of Figure 9, than one would not expect to see such large and rather rapid (some appear to occur on the time scale of 0.5 hour) changes in the bias corrections of Figure 9. This stresses the importance of including an error estimate in the Allan plots, preferably determined by repeated determination of the Allan plot using different time series.

Figure 9: It should be pointed out that the flat appearance of the L1115-i 3-day bias correction curve is purely accidental. If the calibrations had been carried out on days 3, 6, and 9.5 (instead of on days 4, 7.5, and 11) the curve of the L1115-i would have shown a peaked appearance, whereas that of the WVIA would have appeared flat. Both instruments appear to exhibit roughly the same long-term behavior.

P. 1610 L. 22: The lag time for the L1115-i instrument is given in the text as 88 sec, whereas the figure shows a time lag of approx. 130 sec. Also, the lag times are about equal for the Picarro and LGR instruments for the same length of inlet tubing. This is not what I would expect based on the 20 times higher pumping speed of the latter instrument.

P. 1616 L. 15: The measurement precision decreases (= improves!) with increasing mixing ratio.

P. 1616 L. 24: I don't understand the reference to Heliker (2010). The spectral features are reported in HITRAN, whereas the first water isotopic ratio measurements in the spectral ranges used in these commercial instruments were reported by Kerstel et al. (Spectroc. Acta 2002) and Gianfrani et al. (Opt. Ex. 2003).

P. 1617 L. 25: It should be noted that the uncertainties in the parameters of Eq. (6) and (7) (presumably corresponding to the shaded areas around the curves in Figure 7) propagate into the uncertainties of the isotopic ratio measurements. These uncertainties are certainly not negligible and ought to be quantified. The paper by Rambo et al. (2011) shows that these amount corrections, in their case for the WVIA, are not stable and vary significantly and in a random manner on an hour-to-hour basis.

P. 1618 L. 22: Please note that a hydrocarbon trap typically does not remove methane, which is know to be an important interfering species, especially at low water concentration. See, e.g., Hendry et al. (Anal. Chem. 2011).

P. 1620 L. 22: The important point to note is that for such mesoscale meteorological measurements, one may thus assume that the natural variations are larger than the instrument precision, and larger than the instrumental drift over the averaging time period.

P. 1620 L. 25: For standard Eddy covariance measurements, a measurement rate of typically 20 Hz is required. Even though the update frequency of the analyzers may be of the order of 2 Hz, their actual response time does not even come close to this number because of pumping speed and memory effect constraints (see Table 7). In fact, in the case of the L1115-i analyzer, even the 5-s averaging time is not useful, as the analyzer is not able to capture variations at this time scale. With either of these analyzers, only relaxed Eddy accumulation provides a realistic option for flux measurements, but is cumbersome to implement because of its requirement of conditional sampling.

P. 1625 L. 22 and 28: How are these accuracy numbers to be reconciled with the accuracy statements made on P. 1623 L. 5 and 8? I believe this merits a thoughtful discussion, especially considering that WS6 and 7 do not span an extremely large isotope range.

P. 1626 L.7: When comparing the time series obtained by two different instruments with rather different time responses, it is crucial to first put the time series on the same time sampling axis. That this was indeed done, is only mentioned in the caption of Figure 13, but given its importance should also be mentioned in the text. Have the measurements also been weighted with the water mixing ratio, in order to account for the over-representation of the high mixing ratio values in the L1115 time series due to its longer response time (see, e.g., Iannone et al. J. Geophys. Res. 2010)?

Interactive comment on Atmos. Meas. Tech. Discuss., 5, 1597, 2012.