

## ***Interactive comment on “Calculating uncertainty for the RICE ice core continuous flow analysis water isotope record” by Elizabeth D. Keller et al.***

**Anonymous Referee #2**

Received and published: 5 April 2018

This paper describes in some detail a method of calibrating, and uncertainty estimation, of stable water isotopes in ice cores measured by a continuous flow technique. It is described as a novel approach, but I am not sure that this claim is substantiated given that measurement of gases, chemistry and water isotopes in ice cores is now rather common, and all use somewhat similar methods of calibration, drift correction and error analysis. Perhaps it is true that the uncertainty calculation for water isotope has not been published in such a detailed manner before, and this is where the novelty lies. On the whole, the paper is well written: it reads clearly, and is logically set out, and it is interesting. I am aware that an earlier reviewer has made substantial comments about each of the error estimates, and the method of combining the errors for an overall error estimate. While I largely agree with the details of this earlier review of the manuscript,

C1

I do not find it negates the value of the paper to this journal – the discussion of how to characterise the differing contribution to the error of the measurement will be helpful to many working in the field of continuous measurements who have not yet thought through the error contributions quite as deeply as the authors here. Having worked with water isotope instruments for many years, I cannot help feeling that the total mean error estimate recorded in the abstract is qualitatively about as expected, and refining the method of obtaining it will likely not change the figures substantially. And, for example, a small change to the total error in  $\delta\text{dD}$  of 0.74 per mille would not substantially alter the (qualitative) interpretation of the sample data, which shows a range of about 70 per mille in figure 2, though to be fair small changes in the individual isotope pair errors has a larger impact on the error on the calculated  $\text{d}\text{x}\text{s}$ . But for sure, the method by which the total error was obtained is the whole point of the paper, so I am not recommending acceptance without taking into account the criticism of my colleague with the prior review, which I am not about to repeat here in my comments, but offer some further points to address. (One of the downsides of open reviewing is that it is hard to be completely independent.) Specific comments: Given that this is a paper describing the calibration and error estimate, I really want to know how the local standards themselves (four are mentioned on page 4: MilliQ, RICE, WS1, ITASE) were calibrated, and how often. P4/L30 describes the local standards as the accepted value (and sometimes in the manuscript as the ‘true values’ – when used as a check standard). Several points here. One: how were these accepted values obtained? Details of the calibration of the local standards against the primary Vienna standards (or intermediates) are needed here. Two: they are described as on the VSMOW/SLAP scale at various points in the manuscript, but if I were to be pedantic, I wonder if they were actually calibrated against VSMOW-2 and VSLAP-2 since the original Vienna standards were exhausted about 10 years ago. I wonder whether the local standards here are calibrated against the primary Vienna standards, or through intermediate secondary standards – this level of detail is needed in a paper focussed on calibration. Three: additional errors creep in when producing and maintaining the local isotope standards – each calibration step

C2

removed from the primary Vienna standard introduces a new accuracy and precision error. This error is not included here at all (it isn't there in equation 4 – the calibration error here, as described in section 2.4.3, means an internal error in calibrating the system, rather than the error of the actual standard). As an example, I've just checked my primary Vienna (SMOW-2 and SLAP-2), and they have uncertainties of +/- 0.02 on  $\delta^{18}\text{O}$ , and +/- 0.3 on  $\delta\text{D}$ . My commercial standards claim uncertainties of +/- 0.2 on  $\delta^{18}\text{O}$ , and +/- 1.8 on  $\delta\text{D}$ . Four: there is something rather worrying about the 'accepted value' of the local standards in table 1 which is not explained adequately. Why do the standards change from one year to the next? Ideally, once a local standard is prepared in bulk, and calibrated against a primary standard, the local standard is then sealed in aliquots (preferably heat sealed glass ampoules), stored refrigerated, and only opened when used. The paper mentions that the isotope values depend on the water volume in the cavity (section 2.2). This is handled by removing suspect data where the water volume has drifted away from 20,000 ppm, using a short-term averaging method, and whenever the volume fell below 15,000 ppm. Just to be clear: does this remove any need to calibrate the isotope value with the water volume; is the cutting of suspect data sufficient? Figure 3 and P8, L15-19. There is something really rather odd here. The manuscript (P8, L15-19) claims that the drift correction might be 'poorly adapted to the upper range', and that 'Ideally, we would use a quality-control standard that falls between the two calibration standards, RICE and ITASE.' Yet, there are four local standards available: RICE and ITASE, plus MilliQ and WS1. I see no reason why the 'calibration standards' could not have been WS1 and ITASE (or even MilliQ and ITASE) which would have encompassed the whole range of the samples (figure 3 shows that samples fall above the RICE standard for example), and given one (or two) check standards that are within standard range. Figure 3. Could the legend box be moved off the actual data? Currently, figure 3 amply demonstrates the system drift – all standards are drifting to lighter isotope ratio. But the legend box obliterates the early ITASE standard. Figure 4. The upper panel is poorly produced, with no explanation about the different grey curves, or why one Allan plot does not curve upwards. The

C3

dashed 'discrete precision' line is grey in the legend, and black in the figure. The lower panel shows the Allan error parabola for each stack – it would clarify what is demonstrated here to explain that the lowest part of the dip is the start/end of the stack of ice, and the point at which the calibration is carried out. Abstract and P8, L25-30, Summary. Have you any evidence that drill fluid is influencing the error? What fluid is being used? What is its vapour pressure? Could it actually build up in the instrument, or would it simply evaporate at the low cavity pressure and high temperature and be swept out with the waste gas? Are system leaks and valve degradation really affecting the data? What evidence is there, or is this speculation? (I appreciate the text does say 'could affect. ...'). Following this last point, figures 5 and 7 don't help me understand the scatter and total error over the full record. The scatter error is greater in 2014, due to poor core quality, build up of drill fluid and system leaks – yes? But the overall error is better in 2014 (figure 7). What is surely needed here to help clarify this point is a scatter error figure that encompasses the whole core, rather than the short 2013 section in figure 5. Perhaps best of all would be a single figure that for the whole core demonstrates the total error and the three component errors – is this possible?

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-387, 2018.

C4