Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

Atmospheric Measurement Techniques Discussions

Calculating uncertainty for the RICE ice core continuous flow analysis water isotope record

Elizabeth D. Keller[1], W. Troy Baisden[1*], Nancy A. N. Bertler[1,2], B. Daniel Emanuelsson[1,2], Silvia Canessa[1], Andy Phillips[1]

[1]National Isotope Centre, GNS Science, Lower Hutt, New Zealand
[2]Antarctic Research Centre, Victoria University of Wellington, Wellington, New Zealand
*now at: University of Waikato, Hamilton, New Zealand

Correspondence to: Elizabeth D. Keller (l.keller@gns.cri.nz)

Abstract. We describe a novel approach to the calibration and uncertainty estimation of a high-resolution continuous flow analysis (CFA) water isotope (δ^2H, $\delta^{18}O$) record from the Roosevelt Island Climate Evolution (RICE) Antarctic ice core. Our method establishes robust uncertainty estimates for CFA δ^2H and $\delta^{18}O$ measurements, comparable to those reported for discrete sample δ^2H and $\delta^{18}O$ analysis. Data were calibrated using a time-weighted two-point linear calibration with two standards measured both before and after continuously melting three or four meters of ice core. The error at each data point was calculated as the quadrature sum of three factors: Allan variance, scatter over the averaging interval, and general calibration accuracy. Final mean total error for the entire record is δ^2H = 0.74 ‰ and $\delta^{18}O$ = 0.21 ‰. The quality over the length of the dataset is variable, likely due to a combination of poorer ice quality at lower depths, interruptions in the CFA measurements due to ice breaks and equipment failure, the build-up over time of residual drill fluid, and leaks or valve degradation in the system. Despite the somewhat uneven system performance, this represents a significant achievement in precision of high-resolution CFA water isotope measurement.

1 Introduction

Stable water isotopes (δ^2H, $\delta^{18}O$) are a fundamental part of ice core studies. They are particularly important as a temperature proxy (Dansgaard, 1964; Epstein et al., 1963) and are a key component in establishing the age-depth scale and chronology of ice cores (NGRIP Members, 2004; Vinther et al., 2006; Winstrup et al., in review). They also provide other information about climate, including accumulation rates, precipitation source region, atmospheric circulation and air mass transport, and sea ice extent (e.g. Küttel et al., 2012; Sinclair et al., 2013; Steig et al., 2013; Bertler et al., in review; Emanuelsson et al., in review). Historically, water isotopes from ice cores were analysed as a set of discrete water samples using isotope ratio mass spectrometry (Dansgaard, 1964). Recent advances in laser absorption spectrometry have allowed continuous flow analysis (CFA) to become common in ice core studies and are an essential measurement technique for obtaining high-resolution climate records (e.g. Kaufmann et al., 2008; Gkinis et al., 2011; Kurita et al., 2012; Emanuelsson et al., 2015; Jones et al., 2017). However, the multiple system components and continuous nature of CFA poses challenges for calibration and uncertainty

estimation. Because of the size and resolution of CFA ice core datasets and the relatively new application of laser spectroscopy to ice cores, few established methods exist for calculating point-by-point uncertainty throughout measurement. Building on previous studies (e.g. Gkinis et al., 2011; Kurita et al., 2012; Emanuelsson et al., 2015), we have developed a systematic approach to calibration and error calculation that allows for unique uncertainty estimates at each data point in a CFA water isotope record. In this study, we report our methodology for the calibration and calculation of uncertainty and demonstrate the application of the method on the Roosevelt Island Climate Evolution (RICE) ice core δ^2H and $\delta^{18}O$ dataset.

The RICE collaboration retrieved a 760 m ice core from the north-eastern edge of the Ross Ice Shelf over Roosevelt Island in Antarctica (79.39° S, 161.46° W, 550 m a.s.l) during the austral summer 2011-12 and 2012-13 field seasons (Bertler et al., in review). The RICE ice core provides a valuable record of a high snow accumulation site in coastal West Antarctica with annual or sub-annual resolution at the upper depths, representing the late Holocene. The climate reconstruction at the RICE site for the last 2,700 years using the CFA water isotope record is available in a separate publication (Bertler et al., in review). Aside from the value in the methodology itself, this manuscript provides confidence in the precision of the RICE dataset and the climatic interpretation on annual and sub-annual time scales. This method can be applied to other high-resolution CFA ice core water isotope records in the future, and may be suitable for other continuous water isotope measurement applications.

This paper is structured as follows: In Sect. 2, we give an overview of our data processing and data quality control. We detail our methods for calibrating the isotope data and calculating the uncertainty for each data point. Section 3 contains the resulting estimates for each component of the total error of our dataset and an analysis of the different sources of error. We conclude in Sect. 4 with a summary and recommendations for future CFA measurement campaigns.

2 Methods

The abundance of the rare isotope in a sample is conventionally reported in delta notation, defined as:

$$\boldsymbol{\delta} = \left(\frac{\boldsymbol{R_{sample}}}{\boldsymbol{R_{standard}}} - \mathbf{1}\right) * \mathbf{1000}\ ‰ \qquad (1)$$

where $R = {}^{18}O/{}^{16}O$ or ${}^{2}H/{}^{1}H$ (Coplen, 1996). Results in this manuscript are reported in δ per mil ‰, normalized to the international standard Vienna Standard Mean Ocean Water / Standard Light Antarctic Precipitation (VSMOW/SLAP) scale (Gonfiantini, 1978).

2.1 Melting and data processing

Cores were melted and processed at the Ice Core Laboratory at the GNS National Isotope Centre in Lower Hutt, New Zealand. There were two separate melting campaigns, one in June-July 2013 in which the top 500 m were melted, and the other in June-July 2014 in which the remaining 260 m (500-760 m) were melted (Pyne et al., in review). There were several important differences between the two years in the CFA setup (Emanuelsson et al., 2015; Pyne et al., in review), which necessitated that the data from each melting campaign be processed separately. These differences are noted where they are relevant to the

calibration and uncertainty calculations; some factors are calculated individually for each melting campaign and applied only to the data from that campaign.

The ice was cut into 1 m segments and melted at a controlled rate. The melting setup is based on Bigler et al. (2011) and is discussed in more detail in Emanuelsson et al. (2015), Pyne et al. (in review) and Winstrup et al. (in review). Briefly, the cores were placed vertically on a gold-coated copper melting plate and were allowed to melt continuously under gravitational pull. The water from the clean, inner part of the core was drawn from the centre of the melthead and pumped to various instruments for CFA and discrete samples for major ion and trace element analyses. The water from the outer part of the core was saved in vials for discrete stable and radioactive isotope analysis. Either three or four 1 m core segments were stacked on top of each other and melted without interruption (referred to here as a "stack"). At least one calibration cycle of three water standards was run between each stack. An optical encoder that rested on top of the core stack recorded the vertical distance displacement as the core melted. This displacement was translated into depth in millimetres, and along with the melting rate and other system information was written to a log file every second using LabVIEW software (National Instruments). These log files were used to align all CFA instrument data to the depth scale. Processing of the raw data files was performed using a graphical user interface (GUI) and a semi-automated script in Matlab (Matlab Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States). Further details of data processing and depth alignment are available in Pyne et al. (in review).

Water isotope values (δ^2H, $\delta^{18}O$) were measured using CFA with a water vapour isotope analyser (WVIA) using Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS; Baer et al., 2002) and a modified Water Vapor Isotopic Standard Source (WVISS) calibration unit (manufactured by Los Gatos Research (LGR)). This system is described in detail in Emanuelsson et al. (2015). Data was recorded at 2 Hz (0.5 s) for the first 500 m and at 1 Hz for the remaining 260 m (1.0 s), which generated over 5 million raw data points requiring processing.

2.2 Data quality control

We applied several basic selection criteria to identify and eliminate poor-quality data from the raw δ^2H and $\delta^{18}O$ dataset. The two main reasons for data removal were: 1. Changes in the water vapour concentration (H_2O ppm) in the LGR analyser; and 2. The finite response time of the analyser and the transitional period when switching between water standards (i.e. from the calibration cycle) and RICE ice core samples (which by design had very different isotopic values). In addition, some gaps were introduced as a result of the cutting and cleaning of the core into segments and natural breaks in the ice that occurred during the drilling, recovery and handling process (Pyne et al., in review). The depths at which the breaks occurred were recorded and the depth alignment was adjusted accordingly.

The isotope ratio is dependent on water vapour concentration in the analyser (Sturm and Knohl, 2010; Kurita et al., 2012). To minimize the need to correct the data for this, the concentration in the analyser was kept as close to 20,000 ppm as possible. This value was monitored and recorded at the same frequency as the isotope data. For the most part this ratio was stable, but fluctuations and sudden changes did sometimes occur (for example, when air bubbles passed through the line). We removed data when the difference between the H_2O ppm moving average over the short-term system response time of ~60 s and over a

longer-term, stable time of ~200 s was greater than the standard deviation of the short-term average (Emanuelsson et al., 2015): $|avg_s - avg_l| > \sigma_s$. In addition, data were removed if the water vapour concentration fell below 15,000 ppm for an extended period. Figure 1 shows a typical day of raw data, including both RICE ice core stacks and calibration cycles. Data marked in red were removed using these criteria. The majority of these points occur during the switch from one water standard to another in the calibration cycle and do not affect the data from the ice core itself. The percentage of data removed using these criteria was 0.4 % of the total.

It was also necessary to remove some data points at the beginning and end of every stack during the transition period between the Milli Q (18.2 MΩ) laboratory water standard and ice core. This transition is illustrated in Fig. 2. The Milli Q standard is composed of local de-ionised water and has an isotopic value much greater than the RICE ice core (Table 1). Milli Q was run immediately before and after each stack, and there is a period of instrumental adjustment and mixing when switching between them due to memory effects and the finite response time of the spectrometer (see Emanuelsson et al. (2015) for a full discussion). To ensure that the data is not influenced by the mixing at the beginning and end of the stack while including as much data as possible, we calculate the numerical derivative (or the rate of change) between consecutive δ^2H data points during the transition until the derivative falls below a threshold; all points prior are then excluded. The same process is performed at the end of the stack in reverse. The threshold was found empirically and is different in 2013 and 2014 because of the difference in the response times of the two setups and the precision of the data. Data was inspected manually for cases where the algorithm was inadequate. Approximately 2-5 cm of the beginning and end of every stack were removed using this condition. These appear as gaps in the depth of the final dataset. There were also a few occasions when melting was interrupted due to equipment failure, and Milli Q was run through the system until melting could resume; these periods were removed using the same procedure. A typical stack showing a portion of data removed is shown in Fig. 2 (δ^2H vs. depth). The fraction of total data removed was 5.4 %. This resulted in short data gaps of 5-10 cm every three or four meters.

The entire dataset was manually inspected for any other regions of poor quality, and points that visibly fell outside the normal range or were affected by known instrument problems were removed. This only applied to a few isolated sections of data and was a very small portion (< 0.1 %) of the total.

2.3 Calibration

It is necessary in laser spectroscopy to normalize the isotopic values to the VSMOW/SLAP scale and to correct for instrumental drift. To accomplish this, we used a 2-point linear calibration method (Kurita et al., 2012). Before and after each ice core stack, we ran calibration sequences consisting of four laboratory water standards: Milli Q, Working Standard 1 (WS1), RICE snow (RICE), and US-International Trans-Antarctic Scientific Expedition West Antarctic snow (ITASE). An example of a calibration cycle is shown in Fig. 3. Accepted values for these standards as measured on the VSMOW/SLAP scale are in Table 1. The accepted values for the 2013 calibrations were determined using discrete laser absorption spectroscopy measurements on an Isotope Water Analyzer (IWA) 35EP system. In 2014, our instrument was upgraded with a second laser to IWA-45EP, and the 2014 calibrations utilize values from standards measured continuously with this system. This gives a more accurate

measure of the "true" value of the standards under the same conditions as the melting campaign. We were not able to re-measure the standards using the 2013 CFA setup after the setup was modified for the 2014 campaign, so we use the 2013 discrete measurements in the 2013 calibrations. We expect that our calibration will be more accurate for the 2014 melting campaign than for the 2013 campaign because our 2014 accepted values will more accurately reflect the actual laboratory melting conditions.

The standards used for the 2-point linear calibration, RICE and ITASE, have accepted values which form an upper and lower bound, respectively, for the majority of the ice core isotopic values (the ice core values occasionally fall above the RICE standard value). The third water standard (WS1) served as a quality control to enable us to check and quantify the accuracy of the calibration. Each standard was run continuously for approximately 10 minutes (but varied between 8-15 min over the course of the melting campaigns), of which the first and last 100-200 s was discarded to ensure only the middle, stable portion of the measurement was used for calibrations. Around 300 s of data were averaged to arrive at the mean value of the measurement.

Frequent measurements of calibration standards are necessary to correct isotopic measurements for instrumental drift over time. At least one cycle of all three standards was run between stacks, and in many cases, there were several cycles. Melting a stack of three or four cores took around 2-2.5 hours, so the measurement at the mid-point of a stack (the points furthest from a calibration) is about 1-1.5 hours from the nearest calibration. While this is longer than would be ideal for isotope laser spectroscopy, the stability of other elements of the CFA system (in particular, continuous flow methane measurements) required long uninterrupted periods of melting. $\delta^{18}O$ is typically more affected by drift than is $\delta^{2}H$. Drift can be worsened by experimental conditions such as drill fluid contamination and leaks in the system as the analyte proceeds toward the vacuum in the laser cavity. We have quantified the error introduced by the amount of time between calibrations with the Allan deviation, discussed in Sect. 2.4.1.

We calculated normalization and slope corrections for each stack using the closest set of standard measurements both before and after the stack. The normalization correction is the measured mean of the RICE standard. The slope correction is the ratio of the "true" difference between RICE and ITASE water standards and the actual difference measured:

$$\boldsymbol{slope_i} = \frac{\boldsymbol{RICE_{true} - ITASE_{true}}}{\boldsymbol{RICE_i - ITASE_i}} \quad (2)$$

Where $RICE_{true}$ and $ITASE_{true}$ are the accepted standard values and $RICE_i$ and $ITASE_i$ are the *ith* measured value of the standards RICE and ITASE, respectively. We apply this linear correction to each data point, weighting the factors calculated from the calibrations before and after the stack by the time difference between the data point and the calibration:

$$\boldsymbol{\delta_{corrected}} = [(\boldsymbol{\delta - RICE_1}) * \boldsymbol{slope_1 + RICE_{true}}] * (\boldsymbol{1 - t}) * \boldsymbol{step} + [(\boldsymbol{\delta - RICE_2}) * \boldsymbol{slope_2 + RICE_{true}}] * \boldsymbol{t} * \boldsymbol{step} \quad (3)$$

where δ is the uncalibrated raw $\delta^{2}H$ or $\delta^{18}O$ data, $RICE_1$ and $RICE_2$ are the measured values of the RICE standard before and after the stack, respectively, t is the time of δ measurement, $step = (t_2 - t_1)^{-1}$, t_1 = starting time of $RICE_1$ standard measurement before the stack, and t_2 = ending time of $RICE_2$ standard measurement after the stack. Our calibration procedure was validated by comparison to discrete measurements in Emanuelsson et al. (2015).

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

Atmospheric Measurement Techniques Discussions

2.4 Uncertainty calculation

We identified three main sources of error and uncertainty in our measurements: (i) the Allan variance (a measure of instrumental stability and precision), (ii) the scatter or noise in the data over the averaging time, and (iii) a general calibration error relating to the overall accuracy of our calibration method. We calculate each one separately and add them in quadrature to arrive at the total uncertainty estimate:

$$\epsilon_{total} = \sqrt{\epsilon_{allan}^2 + \epsilon_{scatter}^2 + \epsilon_{calib}^2} \quad (4)$$

Each data point in the final record is assigned a unique error value. A detailed explanation of each source of uncertainty follows.

2.4.1 Allan variance

The Allan variance σ_{allan}^2, or two-sample frequency variance (Allan, 1966), is often used as a measure of signal stability and instrumental precision in laser spectroscopy (Werle, 2011; Aemisegger et al., 2012). In the context of CFA isotope measurements, it is a measure of how much instrumental drift accumulates over a specified period. It is defined by:

$$\sigma_{allan}^2(\tau_n) = \frac{1}{2n}\sum_{j=1}^{n}(\delta_{j+1} - \delta_j)^2 \quad (5)$$

where τ_n is the averaging time and δ_j and δ_{j+1} are the mean values of adjacent time intervals *j* and *j+1*. The Allan deviation is the square root of the variance, σ_{allan}.

To calculate the Allan deviation of our system, we used measurements of the Milli Q standard, run continuously for 24-48 hours. On a log-log plot of the Allan deviation vs. averaging time (τ), there is a minimum at the averaging time where the precision is highest; before this point, at very short averaging times, instrumental noise affects the signal, and after, at longer averaging times, the effects of instrumental drift can be seen (Fig. 4). Thus, at times longer than the precision limit, the Allan deviation provides an estimate of the error due to instrumental drift as a function of the time difference between the measurement and the nearest calibration. For our system to stay under the precision limit of 1.0 ‰ and 0.1 ‰ for δ^2H and $\delta^{18}O$, respectively (and to permit analysis with deuterium excess, $d = \delta^2H - 8 * \delta^{18}O$), a calibration cycle to correct for drift should occur at least every ~1 hr during ice core measurements (Emanuelsson et al., 2015). However, as noted above, system limitations prevented us from running calibrations as frequently as would have been optimal.

We plot the mean σ_{allan} at all averaging times τ on a log-log scale (done separately for 2013 and 2014) and perform a linear regression on the curve at averaging times greater than the minimum (Fig. 4). The equation of the fit gives what we refer to as the "Allan error":

$$\log \epsilon_{allan} = a * \log t + b \quad (6)$$

$$\epsilon_{allan} = t^a * e^b \quad (7)$$

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

where t is the time difference between the data point and the calibration (as measured from the start of the measurement of the RICE standard), and a and b are constants determined from the linear regression fit. This error factor is calculated for each data point as a function of t. Because we calibrated using standards measured both before and after each stack, there are two factors at each point that are combined with a time-weighted average. Allan error vs. depth for a typical portion of the record is shown in Fig. 4. Each parabola corresponds to a stack, with the local maximum occurring at the point furthest away in time from the nearest calibrations.

2.4.2 Scatter (analytical uncertainty)

A second error derives from the scatter or noise in the signal over our averaging interval (15 s). To quantify this analytical uncertainty, we calculate the standard deviation for every 15-s time interval contained in each measurement of the RICE standard using a moving window and average over the duration of the measurement:

$$\epsilon_{scatter} = \boldsymbol{mean}\left(\sigma_i/\sqrt{n_i}\right),\ \boldsymbol{i = 1 \ldots N} \quad (8)$$

where σ_i is the standard deviation, N is the total number of intervals, and n_i is the number of data points in the *ith* interval (n = ~30 in 2013 and ~15 in 2014). Again, because the RICE standard was measured both before and after each stack, we calculate $\epsilon_{scatter}$ for both measurements and combine them using a time-weighted average. Scatter error vs. depth for a typical portion of the core is shown in Fig. 5. Note this error is linear with time but is discontinuous at the points where a stack begins and ends.

2.4.3 Calibration accuracy

Finally, we calculate a general error from our calibration procedure as a measure of the accuracy of the calibration, ε_{calib}. This is meant to incorporate uncertainty from a variety of (unspecified) sources, and thus we expect it to be a relatively large source of error. To calculate this factor, we apply the calibration formula using the RICE and ITASE standards (Eqs. (2) and (3)) to the third quality-control standard, WS1, measured in the same cycle. The error is defined as the difference between the measured, corrected value and the accepted value of the standard. An example is shown in Fig. 6. We calculated this difference for all calibration cycles containing measurements of RICE, ITASE and WS1 of sufficient quality (there were 221 such calibration cycles in 2013 and 318 in 2014). Separate error estimates for the 2013 and 2014 melting campaigns were calculated as the mean of the error from all calibrations in each respective year and then applied to the data points from that campaign.

3 Results and Discussion

Total error vs. depth for the whole record is shown in Fig. 7 and summarized in Table 2. The mean total error for all data points is 0.74 ‰ (δ^2H) and 0.21 ‰ (δ^{18}O). Separated by melting campaign, mean total error in 2013 is 0.85 ‰ (δ^2H) and 0.22 ‰ (δ^{18}O) and in 2014 is 0.44 ‰ (δ^2H) and 0.19 ‰ (δ^{18}O). The total error reduces sharply at a depth of 500 m due to the switch between 2013 and 2014 setups and the greatly reduced calibration error in 2014. However, we observe a larger variability in

the error in the 2014 data. This is mainly a result of the highly variable amount of noise in the measurements, which is discussed below.

The mean Allan error for all data is 0.12 ‰ for δ^2H and 0.14 ‰ for $\delta^{18}O$. Calculated separately by melting campaign, the mean errors are 0.13 ‰ (δ^2H) and 0.16 ‰ ($\delta^{18}O$) in 2013 and 0.083 ‰ (δ^2H) and 0.11 ‰ ($\delta^{18}O$) in 2014. As expected, the Allan error peaks at the points in the middle of the stack furthest from a calibration (Fig. 4b). It is both absolutely and proportionally larger for $\delta^{18}O$, as $\delta^{18}O$ is typically more affected by drift.

The amount of scatter in the data varies considerably over the length of the record, particularly in 2014. The average scatter error over the whole record is 0.29 ‰ (δ^2H) and 0.10 ‰ ($\delta^{18}O$). Separated by melting campaign, the average errors are 0.26 ‰ (δ^2H) and 0.093 ‰ ($\delta^{18}O$) in 2013, and 0.37 ‰ (δ^2H) and 0.13 ‰ ($\delta^{18}O$) in 2014. The average scatter is larger overall in 2014, although during the periods of best instrumental performance the scatter was smaller than at any point in 2013. The instrument performance was highly variable in 2014, much more so than 2013. The standard deviation of $\epsilon_{scatter}$ is 0.11 ‰ (δ^2H) and 0.045 ‰ ($\delta^{18}O$) in 2014, as opposed to 0.026 ‰ (δ^2H) and 0.012 ‰ ($\delta^{18}O$) in 2013.

Among the three error factors, the general calibration error is the largest contributor to the total error in 2013: ε_{calib} (δ^2H) = 0.80 ‰ and ε_{calib} ($\delta^{18}O$) = 0.12 ‰. However, this error is greatly reduced for 2014: ε_{calib} (δ^2H) = 0.22 ‰ and ε_{calib} ($\delta^{18}O$) = 0.078 ‰, reflecting the improved accuracy of the "true" standard values. The accepted value of WS1 is well outside the range of the RICE ice core and is much greater than the RICE and ITASE standards, so one explanation for the relatively large error is that our drift correction is poorly adapted to this upper range. Ideally, we would use a quality-control standard that falls between the values of our two calibration standards, RICE and ITASE. In addition, we were not able to measure the standards against VSMOW/SLAP using the 2013 CFA setup, which would provide a better comparison between measured and accepted values. The 2013 ε_{calib} is thus likely to be a very conservative estimate of the error.

The scatter error dominates the total error in 2014. The magnitude of this error was highly variable from day to day, and thus the total error also varied considerably. There were some periods in which the instrument performed exceptionally well. During these periods, total error was as low as 0.3 ‰ (δ^2H) and 0.1 ‰ ($\delta^{18}O$). These represent the high end of the system capability. However, for much of the 2014 melting campaign the total error was closer to the average of 0.44 ‰ (δ^2H) and 0.19 ‰ ($\delta^{18}O$). The overall system performance became more variable in 2014. There are three main possible reasons for the large variations in performance. They are: 1) response to breaks in ice and associated bubbles; 2) performance degradation due to unexpected levels of drill fluid in the melt stream (contained in microfractures in the ice); 3) leaks or valve degradation in the laser spectrometer, which operates under vacuum. In addition to the different setup and gradual build-up of drill fluid in the instrument over time, the ice itself was of poorer quality at deeper depths (in the brittle ice zone at depths below 500 m; Pyne et al., in review), containing more breaks that caused interruptions in the CFA measurements. Although we have only anecdotal evidence, the more frequent stopping and restarting of the system in 2014 seemed to introduce more noise into the measurements.

Because the campaign was conducted to operate many measurement systems simultaneously, as is characteristic of ice core CFA campaigns, it was typically not possible to conduct comprehensive performance tests and systematic evaluations during

the one day of down time in each week-long, seven-day cycle. As a result, the precise sources of performance deterioration were difficult to isolate. Our method for calculating uncertainty is designed to reflect the changing day-to-day conditions without the need to attribute the exact source of error.

4 Summary and conclusions

We have described a systematic approach to the data processing and calibration for the RICE CFA stable water isotope record and presented a novel methodology to calculate uncertainty estimates for each data point derived from three factors: Allan deviation, scatter, and calibration accuracy. The mean total error for all data points is 0.74 ‰ (δ^2H) and 0.21 ‰ (δ^{18}O). Mean total error in 2013 is 0.85 ‰ (δ^2H) and 0.22 ‰ (δ^{18}O) and in 2014 is 0.44 ‰ (δ^2H) and 0.19 ‰ (δ^{18}O). This represents a significant achievement in precision of high-resolution CFA water isotope measurement. The system performed exceptionally well during some time intervals in 2014, demonstrating the high capability of the system, even though this was not sustained. The variability in quality could be due to poor ice quality, interruptions in the CFA measurements, the build-up of residual drill fluid in the instrument, and / or leaks and valve degradation. Most likely it is a combination of all of these factors.

The more accurate measurement of our laboratory water standards for the 2014 melting campaign enabled us to reduce the uncertainty considerably for the data at depths greater than 500 m. More generally, a reduction in the uncertainty in the system could be achieved through more rapid calibration cycles, enabling both the insertion of calibration during "stacks" and more rapid troubleshooting to isolate causes of degraded performance.

Acknowledgements

Funding for this project was provided by the New Zealand Ministry of Business, Innovation, and Employment Grants through Victoria University of Wellington (RDF-VUW-1103, 15-VUW-131) and GNS Science (540GCT32, 540GCT12), and Antarctica New Zealand (K049). We are indebted to everyone from the 2013 and 2014 RICE core processing teams. We would like to thank the Mechanical and Electronic Workshops of GNS Science for technical support during the RICE core progressing campaigns. This work is a contribution to the Roosevelt Island Climate Evolution (RICE) Program, funded by national contributions from New Zealand, Australia, Denmark, Germany, Italy, China, Sweden, UK and USA. The main logistic support was provided by Antarctica New Zealand and the US Antarctic Program.

References

Aemisegger, F., Sturm, P., Graf, P., Sodemann, H., Pfahl, S., Knohl, A. and Wernli, H.: Measuring variations of [delta] 18O and [delta] 2H in atmospheric water vapour using two commercial laser-based spectrometers: an instrument characterisation study, Atmos. Meas. Tech., 5, 1491, 2012.

Allan, D. W.: Statistics of atomic frequency standards in the near-infrared, Proc. IEEE, 54, 221–231, 1966.

Baer, D. S., Paul, J. B., Gupta, M. and O'Keefe, A.: Sensitive absorption measurements in the near-infrared region using off-axis integrated-cavity-output spectroscopy. Appl. Phys. B-Lasers O, 75, 261-265, 2002

Bertler, N. A. N., Conway, H., Dahl-Jensen, D., Emanuelsson, D. B., Winstrup, M., Vallelonga, P. T., Lee, J. E., Brook, E. J., Severinghaus, J. P., Fudge, T. J., Keller, E. D., Baisden, W. T., Hindmarsh, R. C. A., Neff, P. D., Blunier, T., Edwards, R., Mayewski, P. A., Kipfstuhl, S., Buizert, C., Canessa, S., Dadic, R., Kjær, H. A., Kurbatov, A., Zhang, D., Waddington, E. D., Baccolo, G., Beers, T., Brightley, H. J., Carter, L., Clemens-Sewall, D., Ciobanu, V. G., Delmonte, B., Eling, L., Ellis, A. A., Ganesh, S., Golledge, N. R., Haines, S. A., Handley, M., Hawley, R. L., Hogan, C. M., Johnson, K. M., Korotkikh, E., Lowry, D. P., Mandeno, D., McKay, R. M., Menking, J. A., Naish, T. R., Noerling, C., Ollive, A., Orsi, A., Proemse, B. C., Pyne, A. R., Pyne, R. L., Renwick, J., Scherer, R. P., Semper, S., Simonsen, M., Sneed, S. B., Steig, E. J., Tuohy, A., Ulayottil Venugopal, A., Valero-Delgado, F., Venkatesh, J., Wang, F., Wang, S., Winski, D. A., Winton, V. H. L., Whiteford, A., Xiao, C., Yang, J., and Zhang, X.: The Ross Sea Dipole – Temperature, Snow Accumulation and Sea Ice Variability in the Ross Sea Region, Antarctica, over the Past 2,700 Years, Clim. Past Discuss., https://doi.org/10.5194/cp-2017-95, in review.

Bigler, M., Svensson, A., Kettner, E., Vallelonga, P., Nielsen, M. E., and Steffensen, J. P.: Optimization of high-resolution continuous flow analysis for transient climate signals in ice cores, Environ. Sci. Technol., 45, 4483-4489, 2011.

Coplen, T. B.: New guidelines for reporting stable hydrogen, carbon, and oxygen isotope-ratio 18 data, Geochim. Cosmochim. Acta, 60(17), 3359–3360, doi:10.1016/0016-7037(96)00263-3, 1996.

Dansgaard, W.: Stable isotopes in precipitation, Tellus, 16: 436–468, doi:10.1111/j.2153-3490.1964.tb00181.x, 1964.

Emanuelsson, B. D., Baisden, W. T., Bertler, N. A. N., Keller, E. D., and Gkinis, V.: High-resolution continuous flow analysis setup for water isotopic measurement from ice cores using laser spectroscopy, Atmos. Meas. Tech., 8, 2869-2883, doi:10.5194/amt-8-2869-2015, 2015.

Emanuelsson, B. D., Bertler, N. A. N., Neff, P. D., Renwick, J. A., Markle, B. R., Baisden, W. T., and Keller, E. D.: The role of Amundsen–Bellingshausen Sea anticyclonic circulation in forcing marine air intrusions into West Antarctica, Clim. Dyn., in review.

Epstein, S., Sharp, R. P., and Goddard, I.: Oxygen-isotope ratios in Antarctic snow, firn, and ice, J. Geol., 698-720, 1963.

Gkinis, V., Popp, T. J., Blunier, T., Bigler, M., Schüpbach, S., Kettner, E., and Johnsen, S. J.: Water isotopic ratios from a continuously melted ice core sample, Atmos. Meas. Tech., 4, 2531–2542, doi:10.5194/amt-4-2531-2011, 2011.

Gonfiantini, R.: Standards for stable isotope measurements in natural compounds, Nature, 271, 534-536, 1978.

Jones, T. R., White, J. W., Steig, E. J., Vaughn, B. H., Morris, V., Gkinis, V., Markle, B. R. and Schoenemann, S. W.: Improved methodologies for continuous-flow analysis of stable water isotopes in ice cores, Atmos. Meas. Tech., 10, 617-632, 2017.

Kaufmann, P. R., Federer, U., Hutterli, M. A., Bigler, M., Schüpbach, S., Ruth, U., Schmitt, J. and Stocker, T. F.: An Improved Continuous Flow Analysis System for High-Resolution Field Measurements on Ice Cores, Environ. Sci. Technol., 42, 8044-8050, doi:10.1021/es8007722, 2008.

Kurita, N., Newman, B. D., Araguas-Araguas, L. J., and Aggarwal, P.: Evaluation of continuous water vapour δD and δ18O measurements by off-axis integrated cavity output spectroscopy, Atmos. Meas. Tech., 5, 2069–2080, doi:10.5194/amt-5-2069-2012, 2012.

Küttel, M., Steig, E.J., Ding, Q., Monaghan, A.J., and Battisti, D.S.: Seasonal climate information preserved in West Antarctic ice core water isotopes: relationships to temperature, large-scale circulation, and sea ice, Clim Dyn., 39, 1841-1857, 2012.

NGRIP Members: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, Nature, 431(7005), 147-151, 2004.

Pyne, R. L., Keller, E. D., Canessa, S., Bertler, N. A. N., Pyne, A. R., Mandeno, D., Vallelonga, P., Semper, S., Kjær, H. A., Hutchinson, E., and Baisden, W. T.: A Novel Approach to Process Brittle Ice for Water Isotope Continuous Flow Analysis, J. Glaciol., in review.

Sinclair, K. E., Bertler, N. A. N., Trompetter, W. J., and Baisden, W. T.: Seasonality of airmass pathways to coastal Antarctica: ramifications for interpreting high-resolution ice core records, J. Climate, 26, 2065–2076, 2013.

Steig, E. J., Ding, Q., White, J. W. C., Kuttel, M., Rupper, S. B., Neumann, T. A., Neff, P. D., Gallant, A. J., Mayewski, P. A., Taylor, K. C., and Hoffmann, G..: Recent climate and ice-sheet changes in West Antarctica compared with the past 2,000 years, Nature Geosci, 6, 372-375. doi:10.1038/ngeo1778, 2013.

Sturm, P. and Knohl, A.: Water vapor δ2H and δ18O measurements using off-axis integrated cavity output spectroscopy, Atmos. Meas. Tech., 3, 67–77, doi:10.5194/amt-3-67-2010, 2010.

Vinther, B. M., Clausen, H. B., Johnsen, S. J., Rasmussen, S. O., Andersen, K. K., Buchardt, S. L., Dahl-Jensen, D., Seierstad, I. K., Siggaard-Andersen, M. L., Steffensen, J. P., and Svensson, A.: A synchronized dating of three Greenland ice cores throughout the Holocene, J. Geophys. Res., 111, D13102, doi:10.1029/2005JD006921, 2006.

Werle, P.: Accuracy and precision of laser spectrometers for trace gas sensing in the presence of optical fringes and atmospheric turbulence, Appl. Phys. B-Lasers O., 102, 313–329, 2011.

Winstrup, M., Vallelonga, P., Kjær, H. A., Fudge, T. J., Lee, J. E., Riis, M. H., Edwards, R., Bertler, N. A. N., Blunier, T., Brook, E. J., Buizert, C., Ciobanu, G., Conway, H., Dahl-Jensen, D., Ellis, A., Emanuelsson, B. D., Keller, E. D., Kurbatov, A., Mayewski, P., Neff, P. D., Pyne, R., Simonsen, M. F., Svensson, A., Tuohy, A., Waddington, E., and Wheatley, S.: A 2700-year annual timescale and accumulation history for an ice core from Roosevelt Island, West Antarctica, Clim. Past Discuss., https://doi.org/10.5194/cp-2017-101, in review.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

Atmospheric Measurement Techniques Discussions

Figures

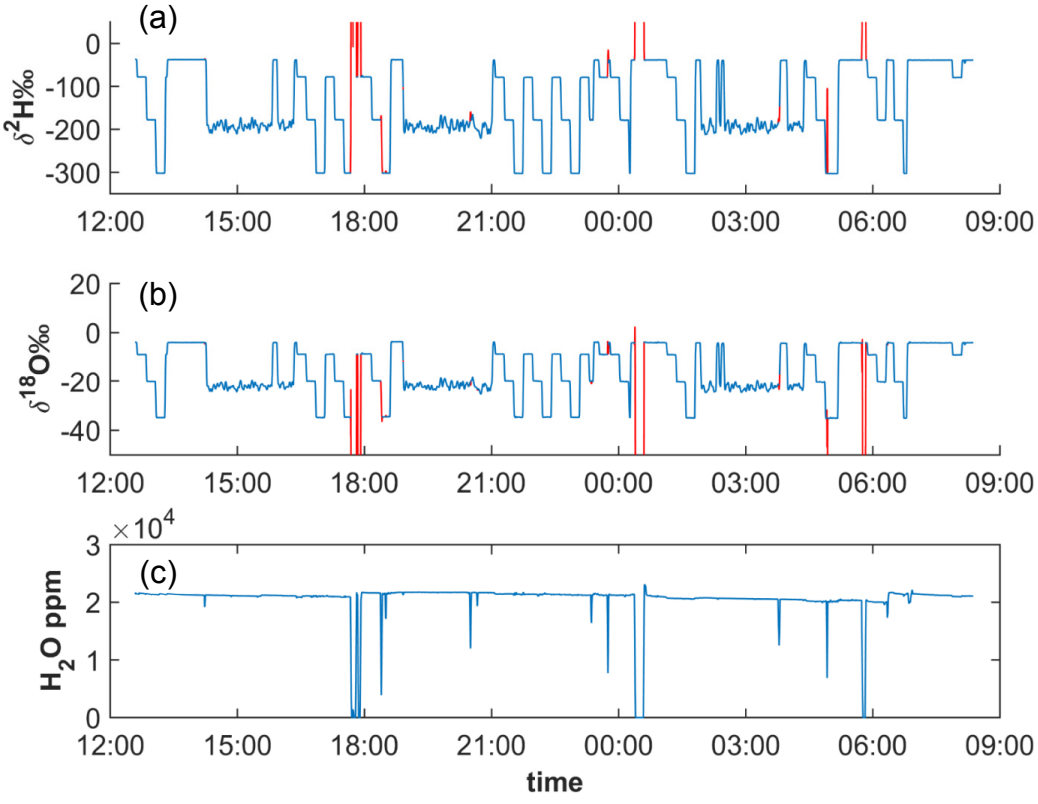

Figure 1: An example of the raw data from a full day of ice melting and calibration cycles (2-3 July 2014): (a) δ^2H, (b) δ^{18}O, and (c) water vapour mixing ratio. Isotope data that were removed because of water concentration anomalies are marked in red in (a) and (b) panels.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

Atmospheric Measurement Techniques Discussions

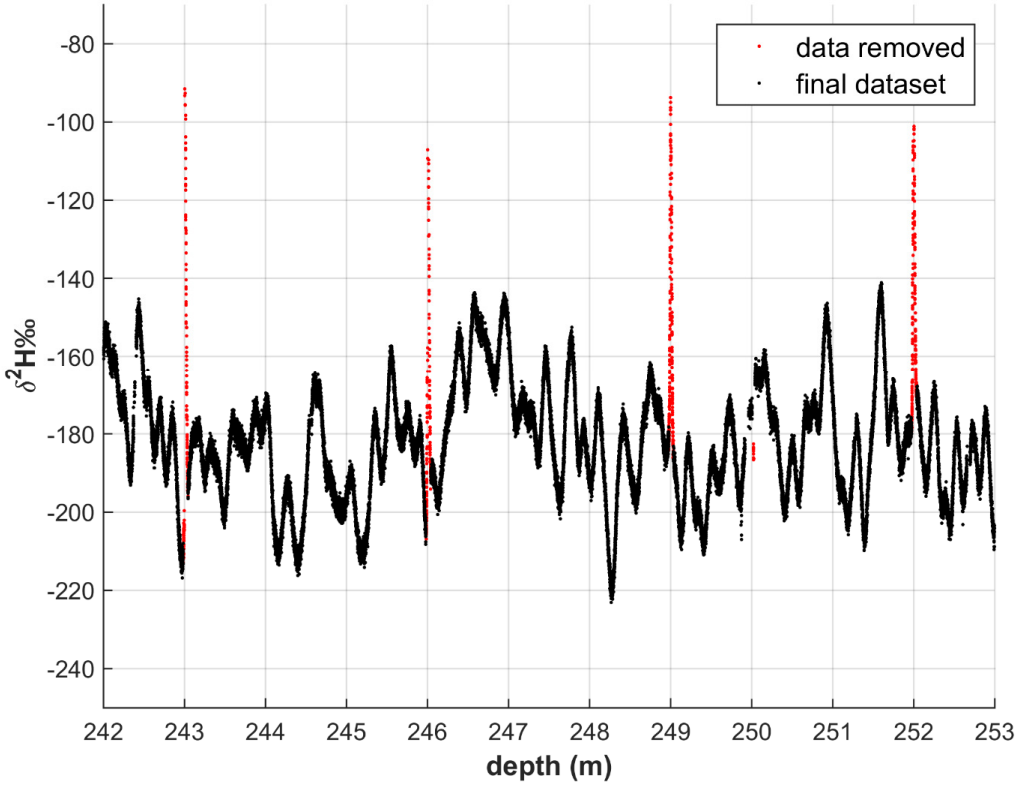

Figure 2: A selected example section of δ^2H vs. depth. The data marked in red represent the transitions between the Milli Q standard and ice core at the boundaries of each 3-metre stack. These data points (and other poor quality data) were removed from the final dataset.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

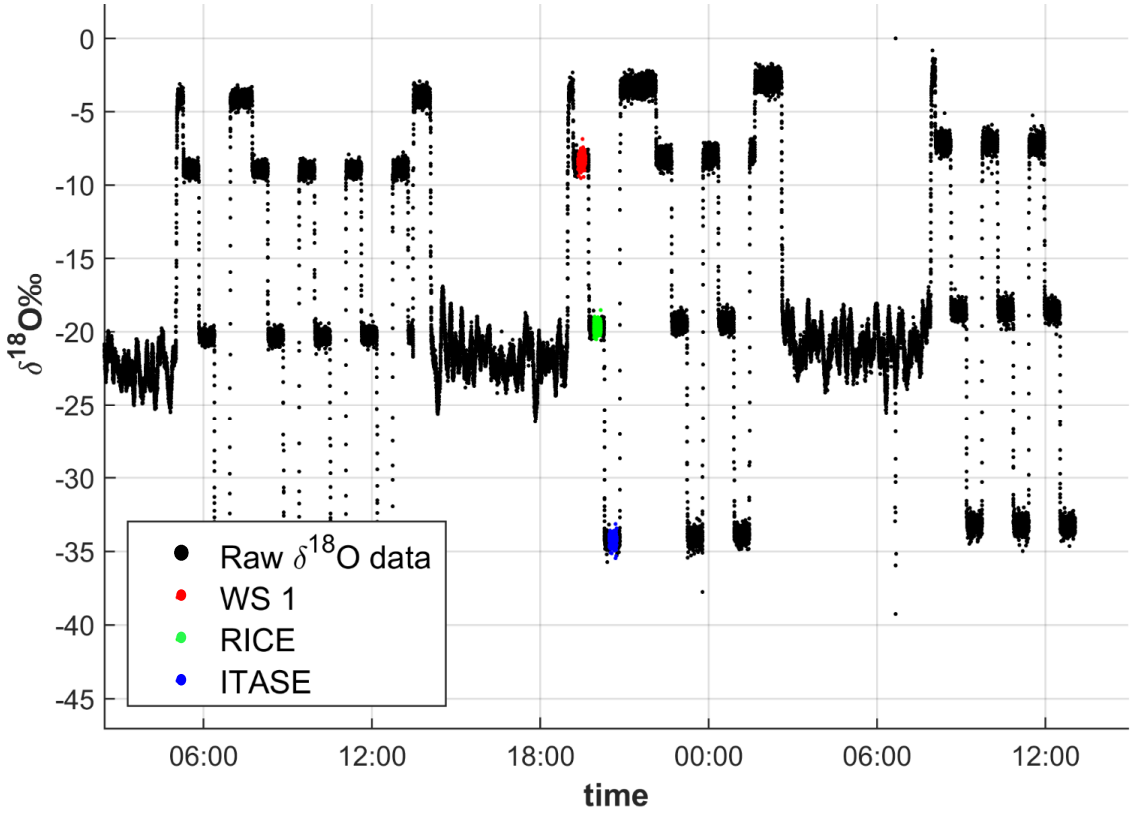

Figure 3: Time vs. raw δ^{18}O (uncalibrated) for one day of melting (3 July 2014). An example of one calibration cycle of three water standards run between ice core stacks are marked in colour: WS1 (red), RICE (green), and ITASE (blue).

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

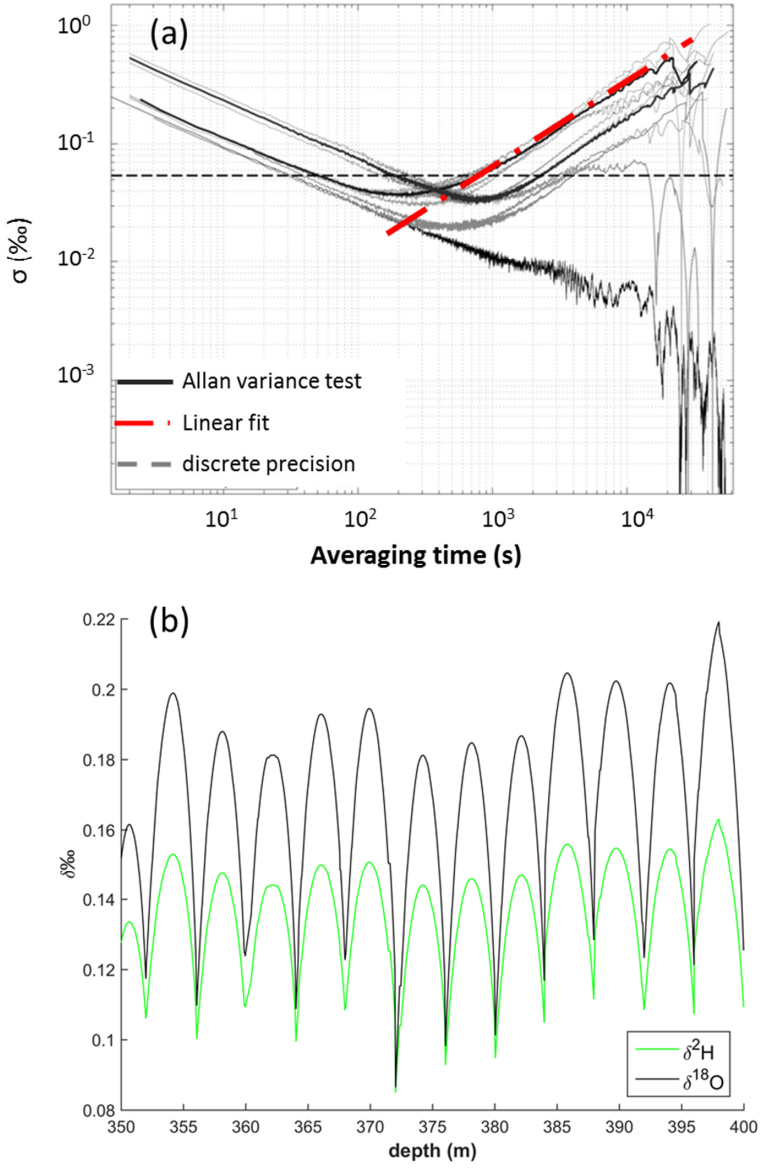

Figure 4: Allan error. (a) Log-log plot of example allan deviation test runs, with linear fit overlaid on one curve. The equation of fit is used to calculate the allan error, as explained in section 3. Average precision of 2013 discrete measurements is indicated with a dashed grey line. (b) Allan error vs. depth for a selected section of ice core. Each parabola corresponds roughly to a stack.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

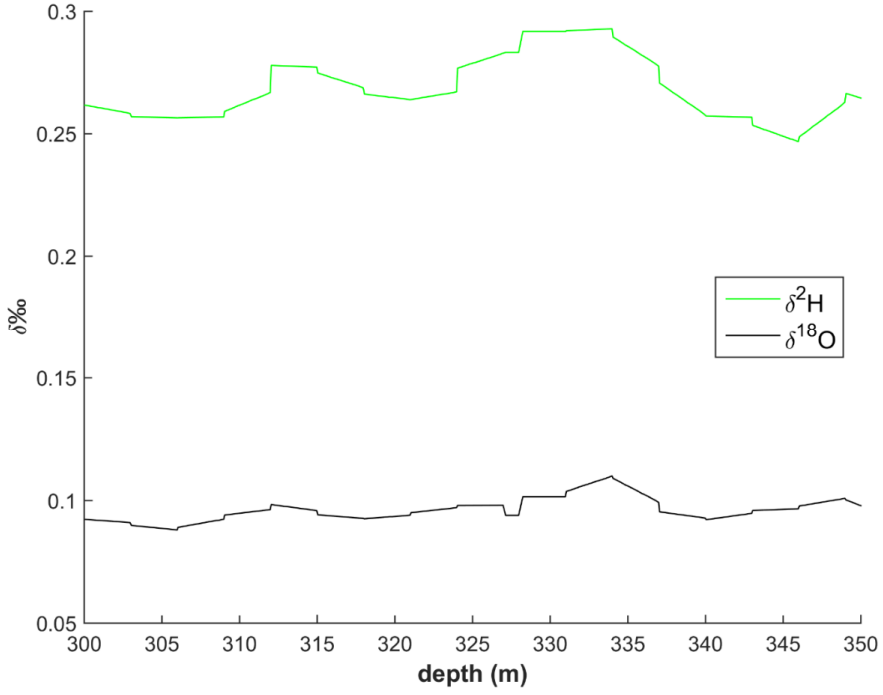

Figure 5: Scatter (analytical) error vs. depth for a selected section of ice core.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

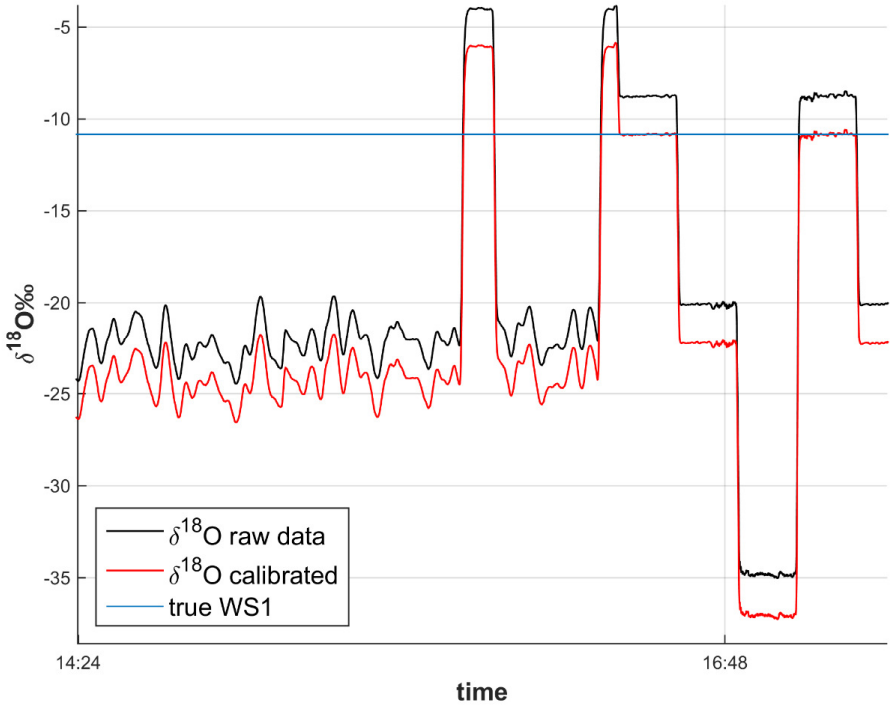

Figure 6: Representative $\delta^{18}O$ calibration of ice core stack and WS1, using RICE and ITASE standards from the same cycle, 15-second moving average vs. time (measured on 2 Jul 2014). The difference between the "true" value of WS1 (blue) and the calibrated measured value of WS1 (red) is the calibration error. The error that was applied to the CFA dataset is the average difference of all WS1 calibration measurements during the melting campaign.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

Atmospheric Measurement Techniques Discussions

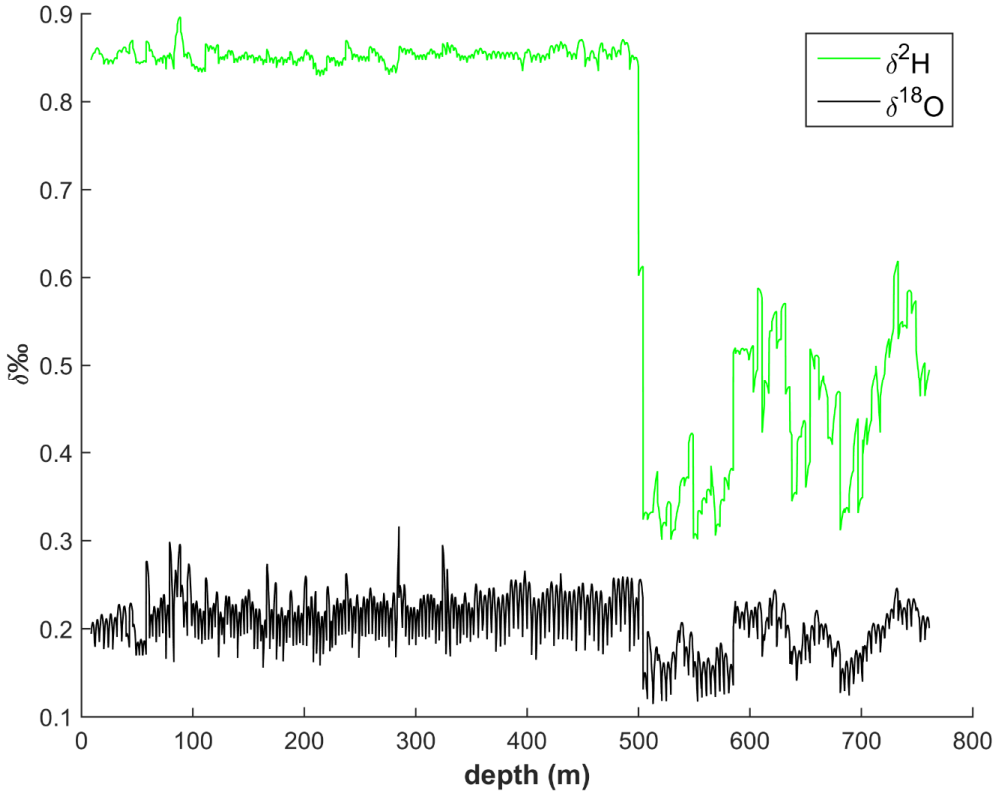

Figure 7: Total error vs. depth. The melting campaign in 2013 paused at 500 m, and melting was resumed in 2014 with a modified setup. The reduced calibration error in 2014 is responsible for the large step down in total error.

5

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-387
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 9 January 2018

Atmospheric Measurement Techniques Discussions

Tables

Table 1: Accepted values (VSMOW/SLAP scale) for water standards used for calibrations, in per mil (‰).

	$\delta^{18}O$		$\delta^{2}H$	
Water standard	**2013** ‰	**2014** ‰	**2013** ‰	**2014** ‰
Milli Q	-5.89	n/a	-34.85	n/a
WS1	-10.84	-10.83	-74.15	-74.85
RICE	-22.54	-22.27	-175.02	-173.06
ITASE	-37.39	-36.91	-299.66	-295.49

Table 2: Summary of error estimates, in per mil (‰).

	$\delta^{18}O$			$\delta^{2}H$		
Error factor	**2013** +/- ‰	**2014** +/- ‰	**Combined** +/- ‰	**2013** +/- ‰	**2014** +/- ‰	**Combined** +/- ‰
Allan	0.16	0.11	0.14	0.13	0.083	0.12
Scatter	0.093	0.13	0.10	0.26	0.37	0.29
Calibration	0.12	0.078	n/a	0.80	0.22	n/a
Total	**0.22**	**0.19**	**0.21**	**0.85**	**0.43**	**0.76**