S1 Stable water isotope measurement

In the following, the calibration protocol including the isotope-humidity dependency of the two SWI time series SWI-13 and SWI-8 are documented in detail.

S1.1 SWI-13

S1.1.1 Calibration protocol

The L2130-i raw measurements were calibrated using the SDM calibration runs to correct for the humidity dependent isotope bias and the drift of the instrument during the cruise using a similar procedure as described in Aemisegger et al. (2012). Calibration runs with two liquid standards (MP [$-11.42\,‰$, $-82.05\,‰$] and GRIP [$-34.59\,‰$, $-267.22\,‰$], for [δ^2H, $\delta^{18}O$], respectively) were performed. Ambient air was dried by pumping it through a replaceable 5 Å molecular sieve with indicating drierite (Agilent, volume of $400\,cm^3$) and used as carrier gas in the calibration system. The residual humidity in the system after 5 min was 50-100 ppmv for dry runs in the laboratory and is expected to be 200-300 ppmv in the tropics at high ambient humidity. The two standards were sampled approximately every second week. The liquid sample analysis for $\delta^{18}O$ and δ^2H performed according to routine procedures at FARLAB (UiB) showed that the variations in the liquid standards during the cruise are smaller than the standard deviation of the SDM calibration runs.

The two liquid standards were measured once per day at different hours of the day for 15 min at a mixing ratio of 12'000 ppmv. Calibration runs had to fulfil three criteria during a time window of 5 min: 1) the standard deviation of the 1 Hz signals of δ^2H and $\delta^{18}O$ is below $1.0\,‰$ and $0.3\,‰$, respectively, 2) the relative standard deviation of the water vapour mixing ratio w is $\leq 5\,\%$, and 3) no significant trend in δ^2H occurs in the 5 min time window. A total of 225 successful calibration runs were conducted during ACE. There are successful calibration runs available at least every second/third day for MP/GRIP, respectively, except for leg 0 and one occasion each during legs 1 and 2 (see also supplementary Fig. S2). For each successful calibration run, the mean values used for calibration of the raw data were calculated for the last 5 min window that fulfils the above criteria to minimise memory effects. The data post-processing procedure adopted for SWI-13 includes the following three steps:

1) Humidity dependent isotope bias correction: Isotope measurements using laser spectroscopy show a bias that depends on the humidity of the measured air sample [referred to as isotope-humidity dependency (e.g. Schmidt et al., 2010; Aemisegger et al., 2012; Steen-Larsen et al., 2013)]. The isotopic composition of known standards at different mixing ratios was measured in the lab after ACE. The raw $\delta^{18}O$ and δ^2H time series were corrected with the isotope-humidity dependency correction terms $\mathcal{H}_{\delta^{18}O}$ and $\mathcal{H}_{\delta^2H}$, respectively, which are the best fit curves using the method of least squares to the measured standard samples (see also Section S1.1.2). The uncertainty of the correction curve is estimated with fitted curves to the measurements ± 1 standard deviation. In this study, we observe a slightly stronger isotope-humidity dependency for very low humidity compared to the isotope-humidity correction curve presented in the study by Sodemann et al. (2017), which characterised the same instrument during a campaign in 2013. All measurements below 12'000 ppmv were corrected for the isotope-humidity dependency. Above 12'000 ppmv, the effect was considered negligible.

2) Drift correction: The humidity-corrected SWI time series was then further post-processed to account for the instrument's drift using a two-point slope correction and normalisation to VSMOW2-SLAP2 according to the IAEA recommendations. A 10-day running mean of the calibration runs was calculated to correct each measurement point using the corresponding running mean standard values. If the isotope-humidity dependency corrected δ^2H was larger than the value of the MP standard, the slope correction was done using a mean over all calibration runs of the corresponding leg for each standard. This was done because successful calibrations with GRIP were rare in the tropics during ACE due to problems with the SDM syringe pump. Because the instrument's drift is small during ACE, a mean value of the calibration runs for each leg should adequately
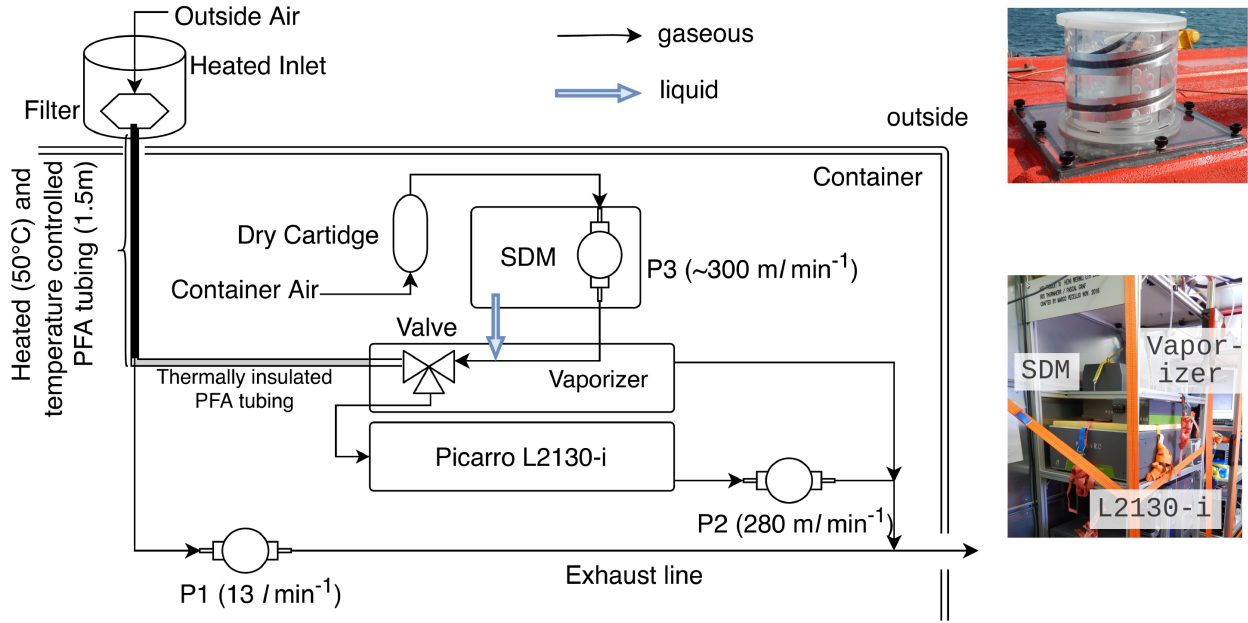
Figure S1: Flow diagramm of measurement setup of SWI-13. SDM is the standard delivery module from Picarro. P1 and P2 are KNF pumps for L2130-i and for flushing the inlet, respectively. P3 is the internal pump of the SDM. The top right photograph shows the heated, 3-layered inlet which was mounted on the container. The bottom right photograph shows the measurement tower in the container with the SDM, the vaporizer and L2130-i.

approximate the standard value. In total, 17% of all data are corrected using a mean standard value over the whole leg, which affects the tropics and subtropics on legs 0 and 4 and the measurements close to South Africa on legs 1 and 3.

2) Water vapour mixing ratio: For the calibration of the water vapour mixing ratio with a dew point generator (LI-COR LI 610), the following correction curve was derived using linear interpolation between w_{raw} as measured by the L2130-i system and expected w from the regulated dew point of the LI-COR:

$w_{cal} = a \cdot w_{raw} + b$ with $a = 0.734$ and $b = -1102.37$ ppmv.

S1.1.2 Isotope-humidity correction curves

The isotope-humidity correction curves were derived using least-square fits to the standard measurements at different water vapour mixing ratio using L2130-i with an SDM. L2130-i has two operational modes: a normal mode with a flow rate of $50 \, \text{m}\ell \, \text{min}^{-1}$ and a flight mode, which has an increased flow rate of $300 \, \text{m}\ell \, \text{min}^{-1}$. The isotope-humidity dependency shows different properties for the two operational modes of L2130-i. At normal mode flow rate, the isotope-humidity correction curve depends on the isotopic composition of the standard sample (supplementary Fig. S5b). For more depleted standard compositions, larger deviations from the reference isotopic composition are recorded at low humidities, whereas the deviations nearly vanish for the most enriched standard composition (MP: [-11.42 ‰, -82.05 ‰], for [δ^2H, δ^{18}O], respectively). This isotope-dependency of the isotope-humidity correction curve is only seen for δ^2H in normal flow mode. An isotope-dependency of the isotope-humidity correction curve was also reported by Bonne et al. (2019) for δ^2H (see supplementary figure 1 in their publication). In flight mode, the isotope-dependency is negligible as the isotope-humidity correction curves vary within the standard deviation of the measurements. The cause of these different isotope-humidity dependency properties for different operational modes of L2130-i is not evident. There could be several causes such as memory effects in the tubings and cavity of the laser spectrometer or background

effects due to the chemical composition of the dry air used for the calibration runs (Aemisegger et al., 2012; Johnson and Rella, 2017). For the calibration of the ACE SWI-13 data, the isotope-dependency of the isotope-humidity correction curves is neglected because the instrument was operated in flight mode. The following isotope-humidity correction curves ($\mathcal{H}_1$ in supplementary Fig. S4) were used for SWI-13:

$$\delta^{18}\text{O}_{corr} = \delta^{18}\text{O}_{raw} + \mathcal{H}_{\delta^{18}\text{O}}(w_{raw}) \quad \text{with} \quad \mathcal{H}_{\delta^{18}\text{O}}(w_{raw}) = -a(w_{raw}^b - 12000^b), \tag{1}$$

$$\delta^2\text{H}_{corr} = \delta^2\text{H}_{raw} + \mathcal{H}_{\delta^2\text{H}}(w_{raw}) \quad \text{with} \quad \mathcal{H}_{\delta^2\text{H}}(w_{raw}) = -c(w_{raw}^d - 12000^d) \tag{2}$$

where w_{raw} is the uncalibrated water vapour mixing ratio in ppmv and $a = 1.369\text{e}{+}04$, $b = -1.078$, $c = 5.033\text{e}{+}04$ and $d = -1.156$.

S1.2 SWI-8

S1.2.1 Calibration protocol

For the calibration of SWI-8, calibration runs were conducted with three liquid standards (Bermuda [$-0.25\,‰$, $2.1\,‰$], NEEM [$-33.5\,‰$, $257.1\,‰$] and CPH MilliQ [$-8.9\,‰$, $-61.7\,‰$] for [$\delta^2\text{H}$, $\delta^{18}\text{O}$], respectively). The main standard was CPH MilliQ which was used for the instrument drift calibration. Samples of CPH MilliQ were collected every two weeks to check for possible storage effects. Sample analysis after the expedition showed that the variation of the isotopic composition of the standard samples did not exceed the instrument's precision. Ambient air was dried by pumping it through a drierite (Agilent) drying unit and used as carrier gas in the calibration systems. Calibrations were started when the humidity in the cavity was less than $200\,\text{ppmv}$.

The laser spectrometers were calibrated once a day with two alternating types of calibration: (1) Instrument's drift calibration was performed by measuring each of the three standards for $30\,\text{min}$ at a mixing ratio of $15000\,\text{ppmv}$. For further calculations, we used the results of the last $10\,\text{min}$ to avoid memory effects. (2) For the isotope-humidity dependency correction, we measured the CPH MilliQ standard at different humidities from 3000 to $30000\,\text{ppmv}$. At every level the standard was measured for $12\,\text{min}$, from which we took only the last $10\,\text{min}$ for further calculations. Best fits using the method of least squares are applied to these calibration runs to correct for the isotope-humidity dependency of the SWI. The following isotope-humidity corrections ($\mathcal{H}_3$ and $\mathcal{H}_4$ in supplementary Fig. S4) are applied:

$$\delta^{18}\text{O}_{corr} = \delta^{18}\text{O}_{raw} + \mathcal{H}_{\delta^{18}\text{O}}(w_{raw}) \quad \text{with} \quad \mathcal{H}_{\delta^{18}\text{O}}(w_{raw}) = a \cdot \exp(\frac{1}{x^{1.08}}) - b, \tag{3}$$

$$\delta^2\text{H}_{corr} = \delta^2\text{H}_{raw} + \mathcal{H}_{\delta^2\text{H}}(w_{raw}) \quad \text{with} \quad \mathcal{H}_{\delta^2\text{H}}(w_{raw}) = c \cdot \exp(\frac{1}{x^{1.08}}) - d \tag{4}$$

where w_{raw} is the uncalibrated water vapour mixing ratio in ppmv and $a = -4294.0$, $b = 4294.1326$, $c = -54830.0$, $d = -54830.6937$ for L2120 and $a = -16036.0$, $b = -16036.4954$, $c = -46986.0$, $d = -46986.4514$ for L2130-i.
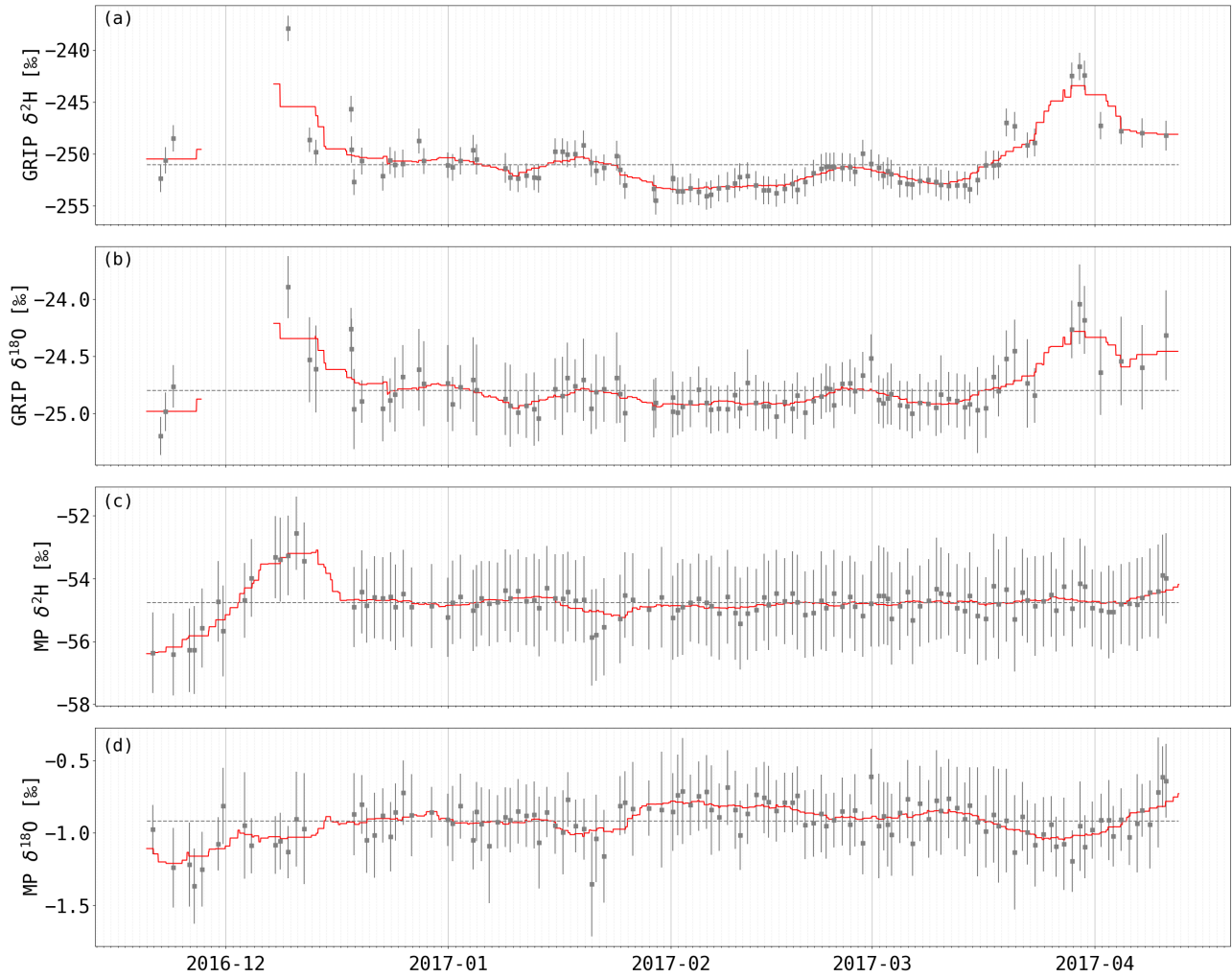
Figure S2: Timeplots of successful mean calibration runs with standard deviation of MP and GRIP for SWI-13 during ACE. The red line represents the 10-day running mean of the calibration runs. The instrument was not running on 13:00 - 23:30 UTC 1 Dec 2016 and from 08:00 UTC 14 Dec 2016 - 09:00 UTC 17 Dec 2016 during leg 0.
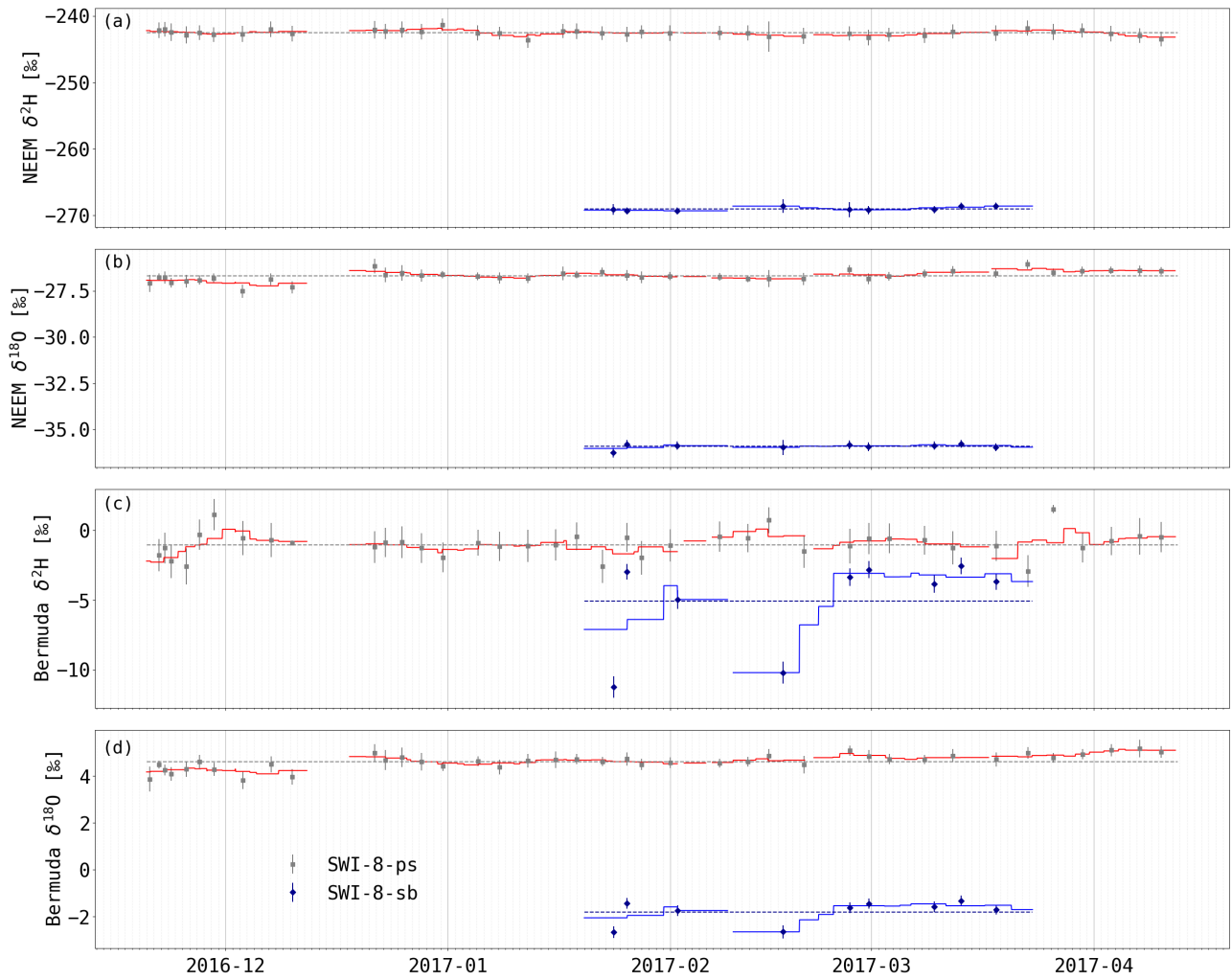
Figure S3: Timeplots of successful mean calibration runs with standard deviation of NEEM and Bermuda for SWI-8-ps and SWI-8-sb during ACE. The red line represents the 10-day running mean of the calibration runs for SWI-8-ps, the blue line, the 14-day running mean for SWI-8-sb.
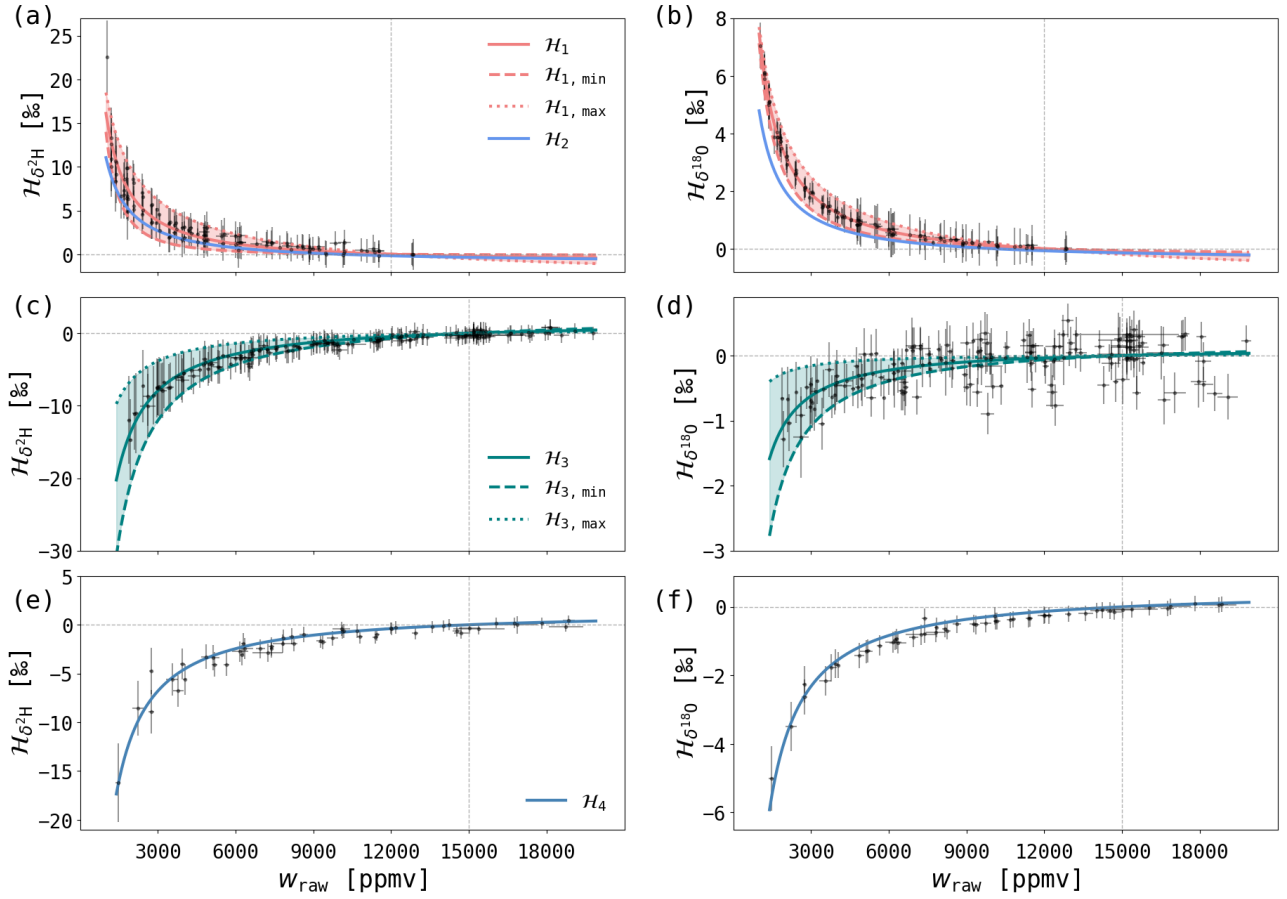
Figure S4: Isotope-humidity dependency correction curves $\mathcal{H}_{\delta^2 H}$ and $\mathcal{H}_{\delta^{18} O}$ for L2130-i (SWI-13) **(a,b)**, L2120 (SWI-8-ps) **(c,d)**, and L2130-i (SWI-8-sb) **(e,f)**. $\mathcal{H}_1$, $\mathcal{H}_3$ and $\mathcal{H}_4$ are the best fit curves to the measured standards of L2130-i (SWI-13), L2120 (SWI-8-ps) and L2130-i (SWI-8-sb), respectively. For L2130-i (SWI-13) and L2120 (SWI-8-ps), the uncertainty of the best fit is shown with maximum $(\mathcal{H}_{1,\max}, \mathcal{H}_{3,\max})$ and minimum $(\mathcal{H}_{1,\min}, \mathcal{H}_{3,\min})$ curves representing the best fits to the measurements +/- 1 standard deviation. $\mathcal{H}_2$ is the isotope-humidity correction curve for L2130-i (SWI-13) from Sodemann et al. (2017).
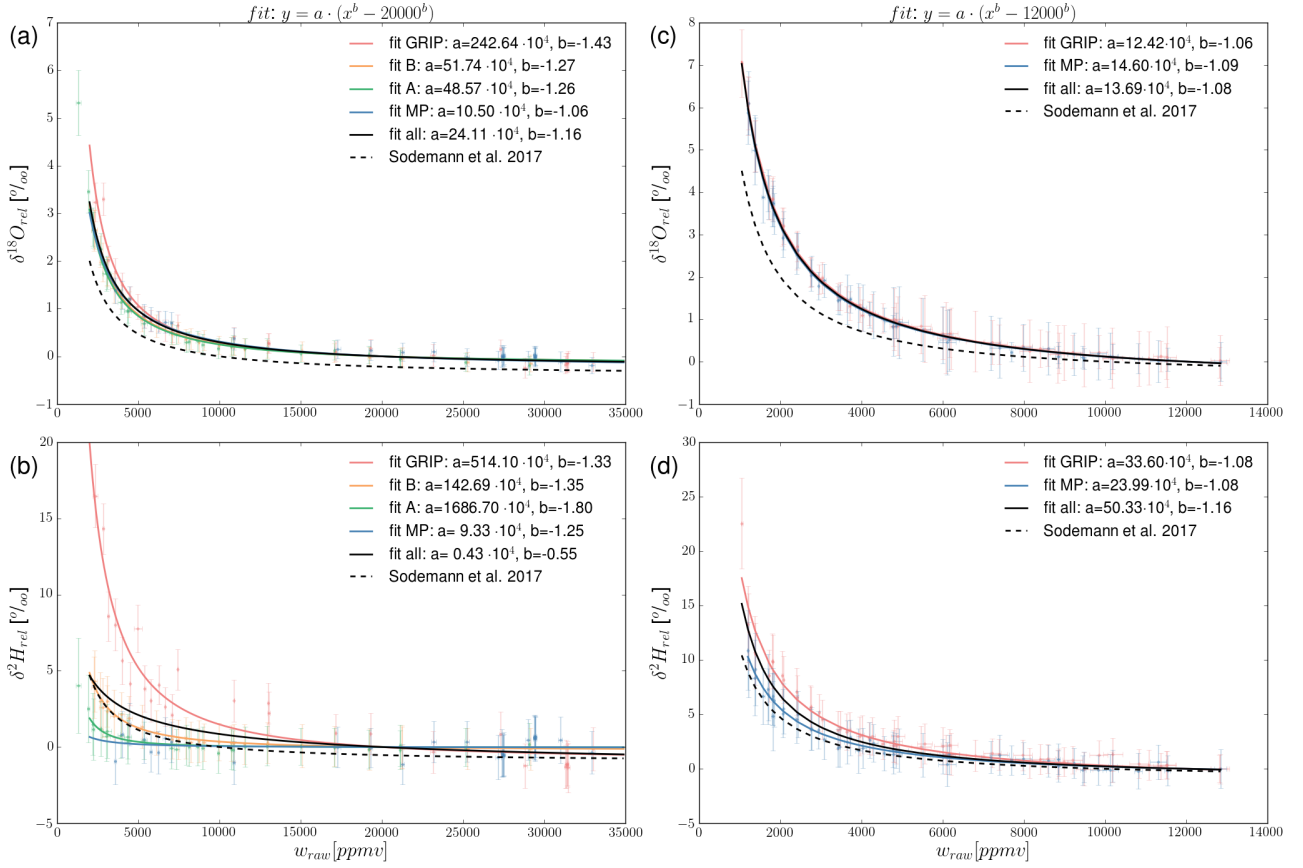
Figure S5: Isotope-humidity correction curves in normal **(a,c)** and flight **(b,d)** mode for L2310-i (SWI-13). $\delta^{18}O_{rel}$ and δ^2H_{rel} are $\delta^{18}O$ and δ^2H, respectively, relative to the mean isotopic composition of the standard runs at humidities above 12000 ppmv. The calibration runs are shown with standard deviations and are coloured by the different standards. The correction curves are the best fits of $y = a \cdot (x^b - 12000^b)$ to the calibration runs with a and b for the corresponding values given in the plot legends. The dashed line is the isotope-humidity correction curve for L2130-i from Sodemann et al. (2017). The isotopic composition (in ‰) of the standards are the following: GRIP (-34.59, -267.33), B (-18.39, -140.07), A (-14.34, -107.81), MP (-11.42, -82.05), for $\delta^{18}O$ and δ^2H respectively.
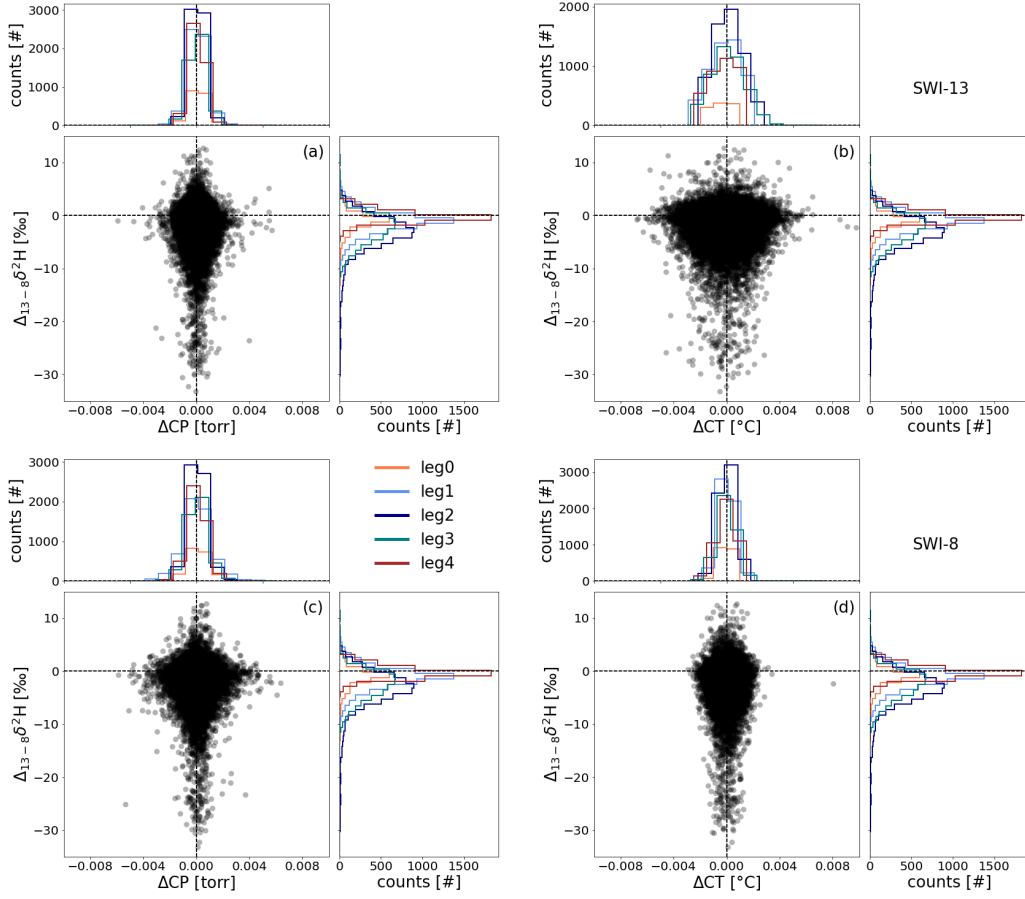
Figure S6: Scatter plots of vertical differences in δ^2H between SWI-13 and SWI-8 ($\Delta_{13-8}\delta^2$H) versus the deviations ΔCP **(a,c)** and ΔCT **(b,d)** from the regulated value of the cavity pressure (CP) and cavity temperature (CT), respectively, are shown for SWI-13 and SWI-8 are shown. Here, 5-minute averages are used. Histograms for each leg (coloured) are shown for the variable in each scatter plot.

S2 Variations in cavity properties

Since δ^2H and δ^{18}O show a similar picture, the following analysis focuses on δ^2H. Figure S6 shows the distributions of $\Delta_{13-8}\delta^2$H against deviations from the regulated value of the CP (Fig. S6a,c) and CT (Fig. S6b,d) of SWI-13 and SWI-8-ps, respectively, with additional histograms of the variables separately for each leg (coloured). The shift towards negative values in y-direction (i.e. $\Delta_{13-8}\delta^2$H) illustrates the systematic difference between SWI-13 and SWI-8-ps. A similar shift is not visible in x-direction for the cavity properties. These measurements are distributed symmetrically around the regulated values of CP and CT for all legs and both instruments and variations in CP and CT do not show any correlation with $\Delta_{13-8}\delta^2$H.

S3 Calibration versions

In Fig. S7 the difference between SWI-8-ps v1 and SWI-13 v1 (Δ_{8-13}) versus the difference between selected calibration versions and version 1 of SWI-13 (Δ_{cal}) is shown. As a proxy to compare Δ_{8-13} and Δ_{cal}, a weighted median for Δ_{8-13} [m_{8-13}] and Δ_{cal} [m_{cal}] is defined:

$$m_{8-13} = \text{median}(\Delta_{8-13}(t)), \tag{5}$$

$$m_{cal} = \text{median}(\Delta_{cal}(t)) \tag{6}$$

$\forall\, t$ within the analysed time period with $\Delta_{cal}(t) \neq 0$. m_{8-13} is larger than m_{cal} for all versions. This means that the vertical differences between SWI-8 and SWI-13 are larger than the uncertainties associated with variations in the calibration protocol. Only for d of v2, points tend to fall on the 1:1-line for very depleted water vapour mixing ratios. In other words, for very low humidities the observed difference in SWI with height could be partially diminished by not applying any isotope-humidity correction. Measurements with low humidity need to be interpreted carefully as their value depends strongly on the humidity correction function which shows a steep slope and increased uncertainty for low humidities. Even though, low humidities lie on the 1:1-line for d in v2, m_{8-13} is 2 times larger than m_{ver2} for d (see Fig. S7g). The uncertainty of the fitted humidity correction curve $\mathcal{H}_1$, visualised with version 3, is small compared to Δ_{8-13}. Δ_{ver2} stays below the standard deviation for most of the measurement points. Version 5 shows only small differences compared to version 1 and affects only points with small Δ_{8-13}.

The ratio of $\frac{m_{cal}}{m_{8-13}}$ is used to estimate how much of the vertical differences Δ_{8-13} can be explained by uncertainties in the calibration procedure. The changes due to different calibration versions amounts for 10-30% of Δ_{8-13} for most versions. The scenario of no humidity calibration shows a higher ratio for d of 82%. The difference between calibration versions using $\mathcal{H}_1$ or $\mathcal{H}_{1,\text{min}}$ explains 15%, 10% and 26% of the vertical differences for δ^{18}O, δ^2H, and d, respectively.

The same analysis was done for SWI-8-ps (Fig. S8). The picture is similar as for SWI-13 with larger m_{8-13} than m_{cal}, except for δ^2H and version 2 without isotopic-humidity correction (Fig. S8d). For this version, vertical differences are of similar order as the differences to the final calibration version. Note, that this is only the case for δ^2H, whereas δ^{18}O and d have smaller or opposite changes between the calibration version compared to the vertical differences. The difference between calibration versions using $\mathcal{H}_3$ or $\mathcal{H}_{3,\text{min}}$ explains 38%, 65% and 13% of the vertical differences for δ^{18}O, δ^2H, and d, respectively. This version comparison underlines that the largest uncertainties are introduced by the isotope-humidity dependency correction and shows that, based on our best knowledge, Δ_{8-13} cannot be fully explained by uncertainties in the calibration procedure.
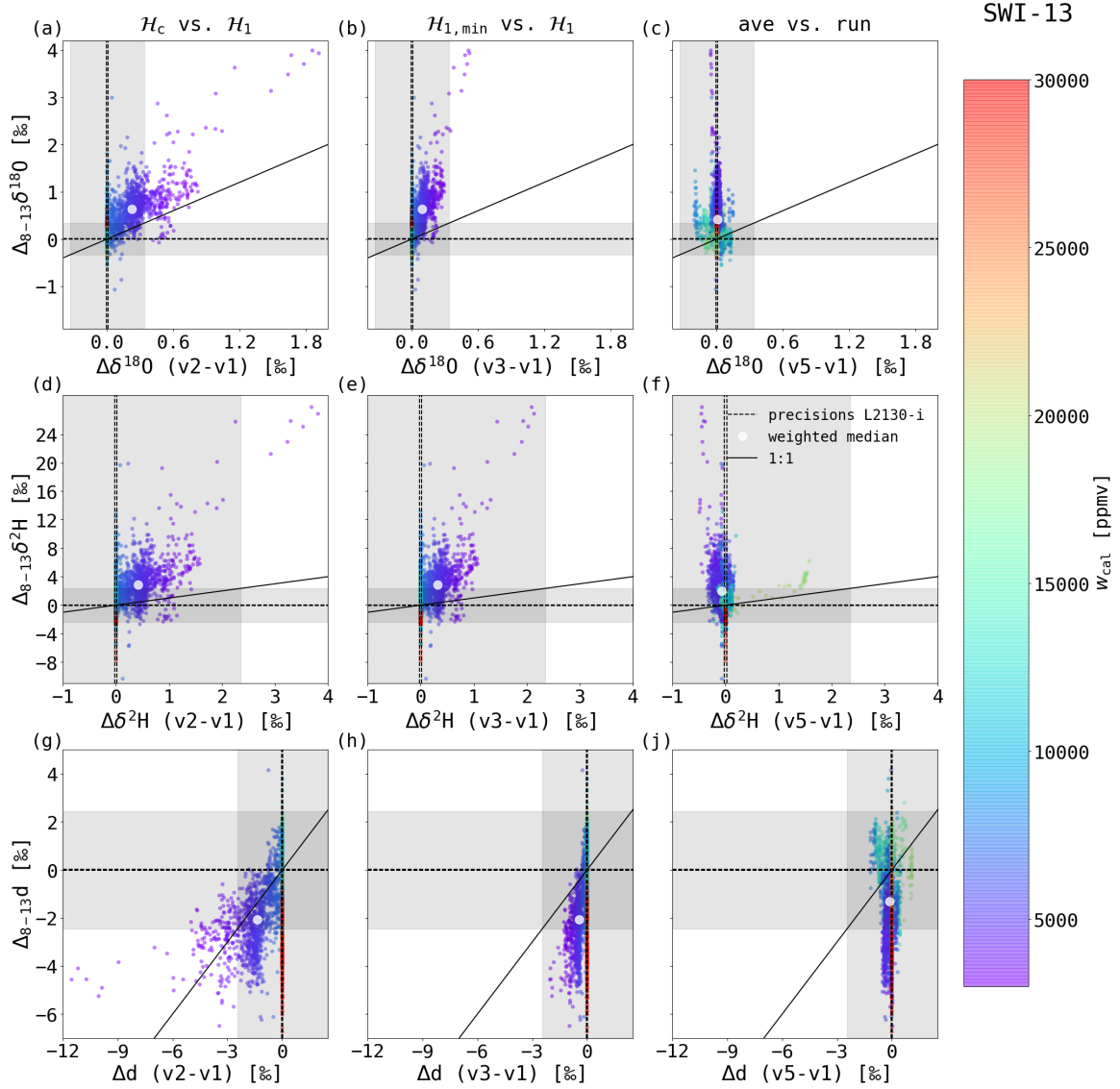
Figure S7: Scatter plots showing different calibration versions of SWI-13: 1-hourly data of vertical differences between SWI-8-ps and SWI-13 $\Delta_{8-13}\delta^{18}O$ **(a-c)**, $\Delta_{8-13}\delta^{2}H$ **(d-f)** and $\Delta_{8-13}d$ **(g-j)** are shown versus calibration versions v2 (**[a,d,g]**, $\mathcal{H}_{c}$, i.e. correction term is zero), v3 (**[b,e,h]**, $\mathcal{H}_{1,\ min}$) and v5 (**[c,f,j]**, average calibration runs) relative to the reference version v1 ($\mathcal{H}_{1}$, run). Dashed lines indicate the precision of L2130-i (from Aemisegger et al., 2012). The solid lines show 1:1 lines. Shaded areas lie within the mean 1-hourly standard deviation of the reference version over all legs. The weighted median (white points) represents the median of the vertical differences m_{8-13} and version differences m_{cal} for all points with $\Delta_{ver} \neq 0$ (see text for details).
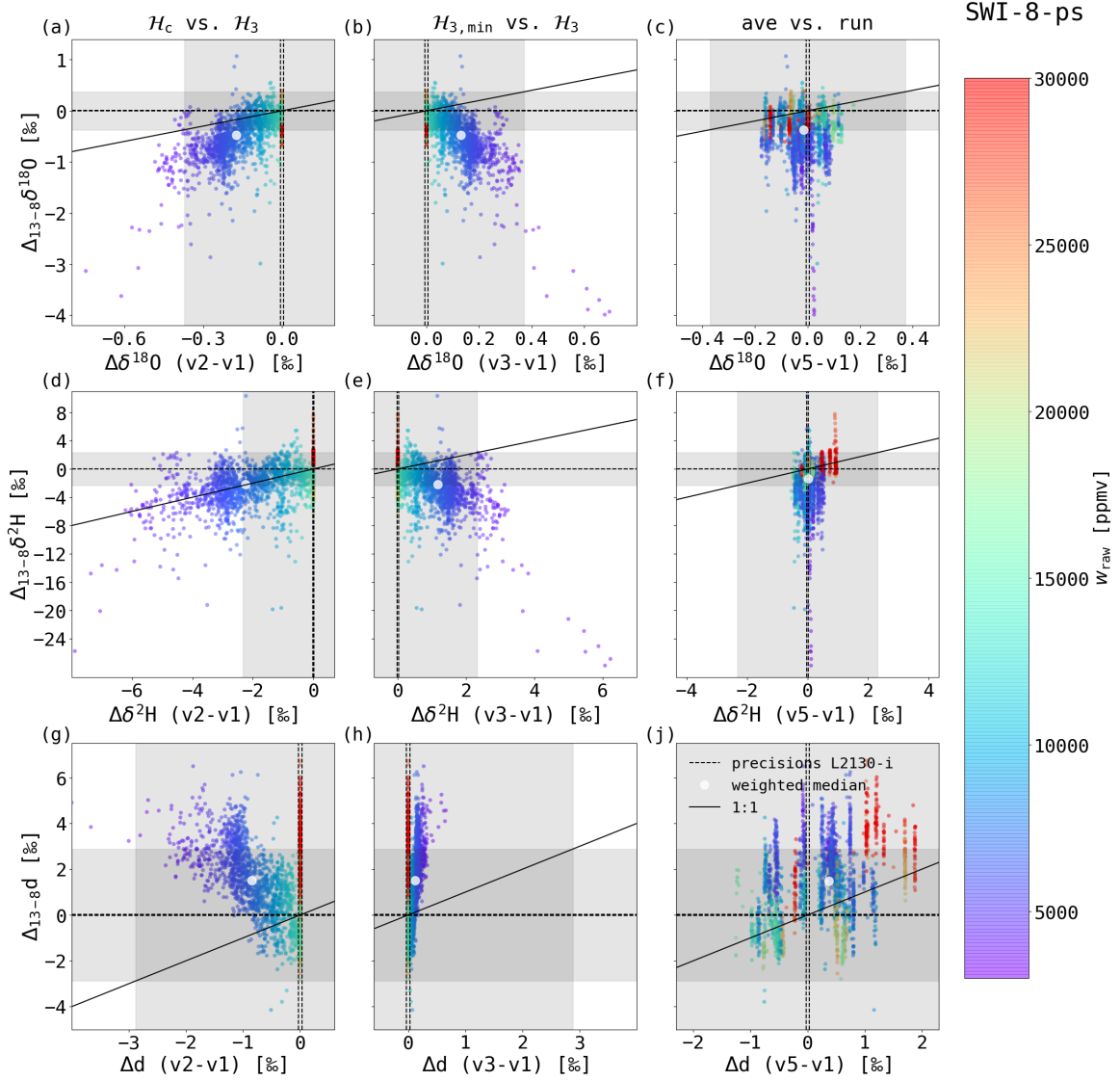
Figure S8: Scatter plots showing different calibration versions of SWI-8-ps: 1-hourly data of vertical differences $\Delta_{13-8}\delta^{18}O$ **(a-c)**, $\Delta_{13-8}\delta^2H$ **(d-f)** and $\Delta_{13-8}d$ **(g-j)** are shown versus calibration versions v2 ([**a,d,g**], $\mathcal{H}_c$, i.e. correction term is zero), v3 ([**b,e,h**], $\mathcal{H}_{3,min}$) and v5 ([**c,f,j**], average calibration runs) relative to the reference version v1 ($\mathcal{H}_3$). $\mathcal{H}_{3,\,min}$ and $\mathcal{H}_3$ refer to the isotope-humidity dependency curves as described in Fig. S4. For the vertical differences, SWI-13 v1 and SWI-8-ps v1 are used. Dashed lines indicate the precision of L2130-i (from Aemisegger et al., 2012). The solid lines show 1:1 lines. Shaded areas lie within the mean 1-hourly standard deviation over all legs. The weighted median (white points) represents the median of the vertical differences m_{13-8} and version differences m_{cal} for all points with $\Delta_{ver} \neq 0$ (see text for details).

S4   Measurement techniques: Wave age and wind speed

S4.1   Wave age

The wave age during ACE was derived from the WaMoS-II measurements. WaMoS-II records and stores consecutive series of 32 radar images per minute. During ACE, the radar operated with a range of $\sim$2.5 km and sampled a total area of 20 km^2 around the ship. Standard image processing techniques based on Fourier transforms are used to extract the wave energy spectrum $E(\omega, \theta)$, where ω is the angular frequency and θ the direction of propagation from 32 images (i.e. one spectrum every minute). An average wave spectrum is calculated every 20 min. N-th moments (m_N) of the spectrum are used to compute the significant wave height $(4 \cdot m_o^{1/2})$, mean periods and group velocity. Wave age is calculated from the derived wave group velocity and the wind speed.

S4.2   Wind speed correction

Wind speed and direction were measured with two 2D-sonic anemometers (models: WS425 and WMT702), which were mounted on vertical poles on the uppermost side arms of the main mast. The anemometers are approximately 7.5 m above the monkey island and 30.5 m above average sea level. The relative wind speed and direction are recorded at a rate of 0.33 Hz together with the ship's heading (compass) by the automated weather station. Averages of the relative wind speed were calculated over 5 min intervals.

The bias caused by airflow distortion was estimated by comparing the observed relative wind speed with the expected relative wind speed based on ECMWF analysis data. The expected relative wind speed was calculated using the following steps: (1) The wind at 30 m a.s.l. [U30, V30] was estimated from the wind at 10 m a.s.l. assuming a logarithmic wind profile modified by the universal stability function (Fairall et al., 2003; Edson et al., 2013). For this, the Monin-Obukov length scale was estimated from the ECMWF surface fluxes. (2) [U30, V30] was interpolated to 5 min resolution and used to calculate the expected relative wind speed based on the ship's heading and velocity by inverting the equations from Smith et al. (1999). The average ratio of observed and expected wind speed was evaluated as function of the relative wind direction and used to scale the observed wind speeds. True wind speed and direction were calculated using the ship's heading (compass) and velocity (GPS) following Smith et al. (1999). Subsequently, true wind speed at 10 m a.s.l. was estimated using the previously used logarithmic wind profile. A manuscript describing this analysis in detail is in preparation.

References

Aemisegger, F., Sturm, P., Graf, P., Sodemann, H., Pfahl, S., Knohl, A., and Wernli, H.: Measuring variations of δ^{18}O and δ^{2}H in atmospheric water vapour using two commercial laser-based spectrometers: an instrument characterisation study, Atmos. Meas. Tech., 5, 1491–1511, https://doi.org/10.5194/amt-5-1491-2012, 2012.

Bonne, J.-L., Behrens, M., Meyer, H., Kipfstuhl, S., Rabe, B., Schönicke, L., Steen-Larsen, H. C., and Werner, M.: Resolving the controls of water vapour isotopes in the Atlantic sector, Nat. Commun., 10, 1632, https://doi.org/10.1038/s41467-019-09242-6, 2019.

Edson, J. B., Jampana, V., Weller, R. A., Bigorre, S. P., Plueddemann, A. J., Fairall, C. W., Miller, S. D., Mahrt, L., Vickers, D., and Hersbach, H.: On the exchange of momentum over the open ocean, J. Phys. Oceanogr., 43, 1589–1610, https://doi.org/10.1175/JPO-D-12-0173.1, 2013.

Fairall, C. W., Bradley, E. F., Hare, J. E., Grachev, A. A., and Edson, J. B.: Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm, J. Clim., 16, 571–591, https://doi.org/10.1175/1520-0442(2003)016¡0571:BPOASF¿2.0.CO;2, 2003.

Johnson, J. E. and Rella, C. W.: Effects of variation in background mixing ratios of N2, O2, and Ar on the measurement of δ^{18}O-H_2O and δ^{2}H–H_2O values by cavity ring-down spectroscopy, Atmos. Meas. Tech., 10, 3073–3091, https://doi.org/10.5194/amt-10-3073-2017, 2017.

Schmidt, M., Maseyk, K., Lett, C., Biron, P., Richard, P., Bariac, T., and Seibt, U.: Concentration effects on laserbased δ^{18}O and δ^{2}H measurements and implications for the calibration of vapour measurements with liquid standards, Rapid Commun. Mass Spectrom., 24, 3553–3561, https://doi.org/10.1002/rcm.4813, 2010.

Smith, S. R., Bourassa, M. A., and Sharp, R. J.: Establishing more truth in true winds, J. Atmos. Ocean. Technol., 16, 939–952, https://doi.org/10.1175/1520-0426(1999)016<0939:EMTITW>2.0.CO;2, 1999.

Sodemann, H., Aemisegger, F., Pfahl, S., Bitter, M., Corsmeier, U., Feuerle, T., Graf, P., Hankers, R., Hsiao, G., Schulz, H., Wieser, A., and Wernli, H.: The stable isotopic composition of water vapour above Corsica during the HyMeX SOP1 campaign: insight into vertical mixing processes from lower-tropospheric survey flights, Atmos. Chem. Phys., 17, 6125–6151, https://doi.org/10.5194/acp-17-6125-2017, 2017.

Steen-Larsen, H. C., Johnsen, S. J., Masson-Delmotte, V., Stenni, B., Risi, C., Sodemann, H., Balslev-Clausen, D., Blunier, T., Dahl-Jensen, D., Ellehøj, M. D., Falourd, S., Grindsted, A., Gkinis, V., Jouzel, J., Popp, T., Sheldon, S., Simonsen, S. B., Sjolte, J., Steffensen, J. P., Sperlich, P., Sveinbjörnsdóttir, A. E., Vinther, B. M., and White, J. W. C.: Continuous monitoring of summer surface water vapor isotopic composition above the Greenland Ice Sheet, Atmos. Chem. Phys., 13, 4815–4828, https://doi.org/10.5194/acp-13-4815-2013, 2013.