

Supplement of:

New water fractions and transit time distributions at Plynlimon, Wales, estimated from stable water isotopes in precipitation and streamflow

Julia L. A. Knapp¹, Colin Neal², Alessandro Schlumpf³, Margaret Neal², and James W. Kirchner^{1,3,4}

¹Department of Environmental Systems Science, ETH Zurich, 8092 Zurich, Switzerland

²Centre for Ecology and Hydrology, Wallingford, OX10 8BB, UK

³Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland

⁴Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA

Correspondence to: Julia L. A. Knapp (julia.knapp@usys.ethz.ch)

Data archive

The isotope data presented here are publicly available. The data sets are attached to the accompanying manuscript as supplemental information, and the data sets will also be archived at the EnviDat data base of WSL/ETH Zurich upon acceptance of this manuscript.

The archived data include a raw data archive and quality-controlled data. The raw data archive (denoted by column headers with "RDA" at the end of each variable name) contains the raw measurements as reported by the laboratory. This raw data archive is provided for completeness of documentation, and also in the interests of full transparency concerning data editing that has been subsequently applied. The raw data archive also provides the opportunity to go back to the original data in case new information emerges suggesting that interesting real-world phenomena may have been inadvertently edited out. However, *the raw data archive should not be used for routine analyses*, because the raw data include some values that are believed to be substantially in error. Instead, routine analyses should use the edited version of the data, where to the greatest extent possible, problematic data values have been corrected or excluded.

Sample analysis and calibration

The samples were analyzed for oxygen-18 and deuterium isotope ratios using a Picarro L1102-i cavity ring-down spectrometer (Picarro, Inc., Sunnyvale, CA, USA) equipped with a Picarro V1102-i vaporization module and a PAL HTC-xt-LEAP-Pic autosampler (CTC Analytics AG, Zwingen, Switzerland). Routine calibrations used three secondary standards (mixed seawater, Fiji artesian water, and Sion drinking water), which in turn were referenced to IAEA VSMOW2, SLAP2, and GISP. Samples were analyzed in blocks of roughly 50 samples, with calibration series immediately preceding and following each block. Each calibration standard was injected twice, with only the results from the second injection being used in order to minimize carryover effects, in view of the large isotopic differences among the standards. The calibration values before and after each block were averaged together, and used to construct a three-point calibration line that was applied to all the analyses within the block.

Each analysis did not include multiple sequential injections of individual samples, as are commonly used to minimize carryover effects in CRDS analyses. Instead, each block was analyzed in sequence, with each sample injected just once. Each block consisted of only streamwater samples or only precipitation samples, to minimize carryover effects from the highly variable precipitation values to the much more stable streamwater values. The streamwater and precipitation samples within each block

were arranged in chronological order to further minimize carryover effects. Each block of samples was re-run, separated by at least two other sample blocks. Thus each sample was analyzed twice, separated by at least 100 other samples and standards, and the reported value for each sample is the average of the pair. Any pair that differed by more than 0.20 ‰ in oxygen-18 or 1.0 ‰ in deuterium was re-analyzed (again twice), and whichever of the two pairs was more consistent was averaged for the final reported value.

The likely magnitudes of memory effects from single injections were estimated from long runs of the calibration standards, and averaged 2.4 ‰ for oxygen-18 and 6.6 ‰ for deuterium (expressed as a percentage of the sample-to-sample change in isotope ratios).

Drift and reproducibility of measurements

To check for drift and reproducibility, we analyzed a single standard 124 times in a continuous 18-hour run, without calibration. Rates of drift were 0.008 ‰ per hour and 0.02 ‰ per hour for oxygen-18 and deuterium, respectively. The mean absolute deviations between pairs of sequential analyses were 0.048 ‰ and 0.17 ‰ for oxygen-18 and deuterium, respectively, and the 90th percentiles of the absolute deviations were 0.104 ‰ and 0.32 ‰ for oxygen-18 and deuterium, respectively.

For comparison, in the real-world stream water samples, the absolute deviations between pairs of sequential analyses averaged 0.161 ‰ and 0.667 ‰ for oxygen-18 and deuterium, respectively, or nearly four times the average sample-to-sample deviations among replicates. The absolute deviations between pairs of rainwater samples were larger still, averaging 1.68 ‰ and 13.3 ‰ for oxygen-18 and deuterium, respectively. These results indicate that analytical noise accounts for only a small fraction of the variability in the Plynlimon isotope time series.

Comparison of CRDS and IRMS measurements

The cavity ring-down spectroscopy (CRDS) measurement results of 41 streamwater samples were cross-checked with independent analyses by isotope ratio mass spectrometry (IRMS). The mean values and standard deviations of the oxygen-18 values for the set of samples analyzed by both CRDS and IRMS were -6.24 ± 0.14 ‰ and -6.52 ± 0.11 ‰ for measurements with the CRDS and IRMS, respectively. These analyses are plotted in Fig. S1.

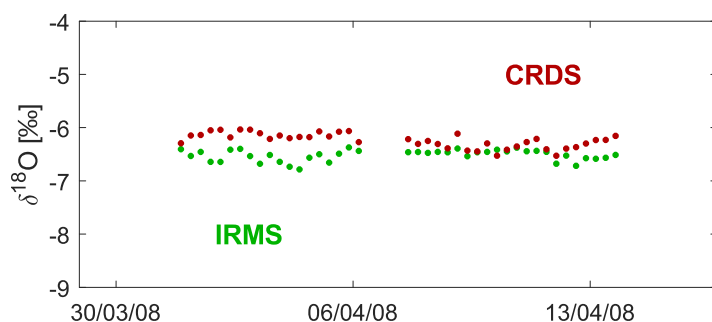


Figure S1: Comparison of measurements with IRMS (isotope ratio mass spectrometry) and CRDS (cavity ring-down spectroscopy) for 41 streamwater samples.

Effect of sample acidification

Samples were split in the laboratory, and one of the aliquots was acidified for analysis of trace metals. Usually the sample for isotope analysis was taken from the un-acidified split, but in some cases it was taken from the acidified split by mistake. This occurred with 7-hourly streamwater samples from 10 July to 4 December 2007, and with precipitation samples from 5 November to 26 November 2007. These acidified samples are noted in the database.

The isotopic consequences of sample acidification could not be quantified precisely, because we did not have acidified and un-acidified splits of the same samples to compare. We also could not create comparison samples after-the-fact, because none of the original lot of acid was still available by the time we discovered that some of the isotope samples had been acidified. Nonetheless, the isotopic behaviors of the acidified and un-acidified samples were broadly similar. Comparing the 429 acidified streamwater samples from 2007 with 482 un-acidified streamwater samples covering the same range of dates in 2008, we found that they differed in oxygen-18 by only 0.02 ‰ (the un-acidified mean was -5.927 ‰, with a standard deviation of 0.318 ‰ and standard error of 0.014 ‰, compared to the acidified mean of -5.949 ‰, with a standard deviation 0.256 ‰ and standard error of 0.012 ‰). Thus the two oxygen-18 means differed by only 1.2 times the standard error of the difference, implying either that acidifying the samples had little effect on their isotopic composition, or that this effect was masked by a similarly large offsetting difference in the stream water isotopic composition between the two years.

The isotopic difference in deuterium between the acidified samples from 2007 and the un-acidified samples from 2008, however, was somewhat larger. To estimate the effect of acidification while correcting for real-world isotopic variability between two years, we constructed a dual-isotope plot for the acidified samples from 2007 and the un-acidified samples from the same range of dates in 2008 (see Fig. S2, below). We then fitted these samples using a single slope, and a constant offset between the acidified and un-acidified samples, using least squares. The acidified samples were 0.68 ± 0.08 ‰ lighter in deuterium than the un-acidified samples. Thus the offset between the samples was much larger than its standard error. Because H and ^2H can exchange between water molecules and HNO_3 in solution but ^{16}O and ^{18}O cannot, one would expect the addition of HNO_3 to more greatly affect the deuterium of the water than the oxygen-18, consistent with what we observed by comparing the acidified and un-acidified samples. Note that the samples were analyzed by cavity ring-down spectroscopy, which measures the deuterium and oxygen-18 of the water molecules rather than of the whole solution, so the isotopic composition of the HNO_3 itself did not directly influence the measurements.

The offset of 0.68 ± 0.08 ‰ between the acidified and un-acidified samples is a small fraction of the deuterium variability in precipitation (standard deviation of 24.4 ‰). It is not a trivial fraction of the variability in streamwater deuterium, however (standard deviation of 2.2 ‰). For this reason, we corrected the deuterium measurements of all acidified samples (streamwater and precipitation) by this average offset value. We have applied this correction as a courtesy to future users of the data, not because it has any significant effect on the results reported here. The ensemble hydrograph separation technique is insensitive to constant biases in the tracer concentrations that it uses as inputs (Kirchner, 2019, Sects. 5.1 and 5.3). Because the acidified streamwater samples comprise an entire block of samples before a long data gap (too long to be bridged by any of the analyses presented here), any isotopic effect of acidification would appear as a constant bias in the inputs to the ensemble hydrograph separation and thus would not affect our results.

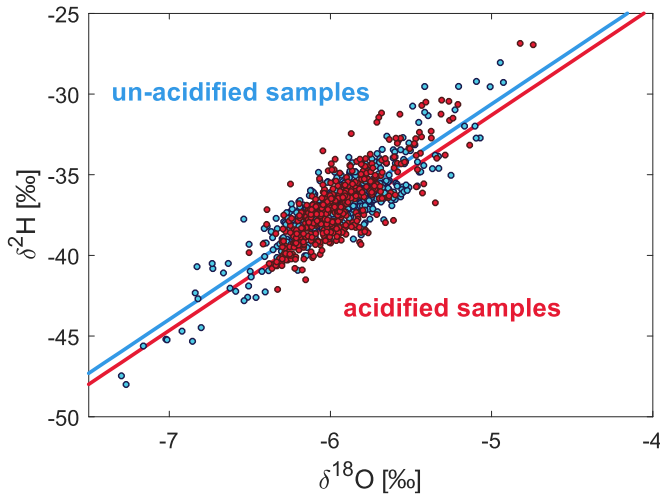


Figure S2: Comparison of acidified streamwater samples (red) and un-acidified samples from the same range of dates one year later (blue).

Sample exclusion due to evaporative fractionation during storage

We excluded one 7-hourly streamwater sample from further analysis, since its isotope ratio deviated significantly from the GMWL, suggesting evaporation during storage (see Fig. S3). In addition, some of the weekly samples showed clear evaporation trends and had to be omitted from the data set; they were isotopically heavy in both deuterium and oxygen-18, but followed a line that was much shallower than the meteoric water line, indicating evaporative fractionation. In total around one-eighth of the weekly samples had to be excluded from the data set (22/177 precipitation samples, 25/215 Lower Hafren streamwater samples, and 27/215 Tanllwyth streamwater samples). Most of these sample bottles had visually obvious head space when they were opened for analysis, despite having been completely filled at the time of original sample processing. Figure S1 compares the excluded data points to the final data sets in dual-isotope space.

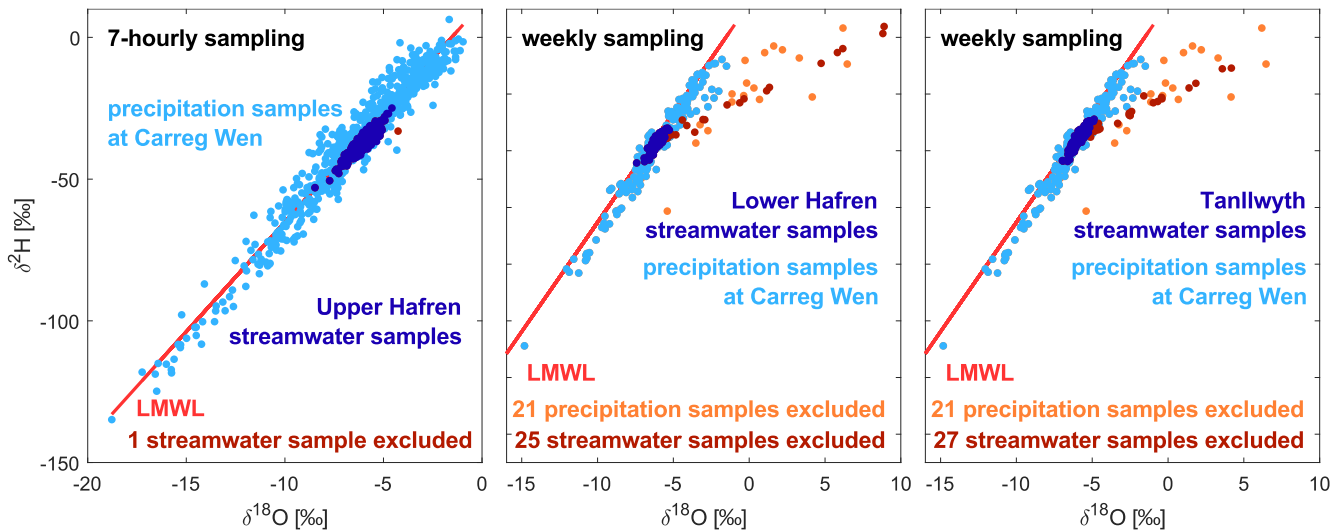


Figure S3: Dual-isotope plots of 7-hourly (left) and weekly (center and right) data. Precipitation samples are shown in lighter colors, whereas darker markers indicate streamwater samples. Samples excluded from the data set due to evaporative effects are shown in red (light red for precipitation and darker red for streamflow), while blue markers indicate data points included in the final data set. The red line defines the local meteoric water line (LMWL) derived from the 7-hourly precipitation data. In total, 74 samples had to be excluded. Excluded data points clearly show evaporation effects and are mainly found among the weekly samples.

Evaporative fractionation during storage in the autosampler carousels

7-hourly precipitation and streamwater samples were stored in open bottles in an autosampler at the field site for up to one week before the bottles were picked up and returned to the laboratory. The same is true for the weekly precipitation samples, because these are cumulative open-bottle samples. Only the weekly streamwater samples from Lower Hafren and Tanllwyth were collected as grab samples and returned to the lab directly after collection.

We examined the question of whether storage in open bottles may have led to evaporative fractionation, particularly in the warmer summer months. We observed different slopes and intercepts of the meteoric water line in summer and winter, with slightly heavier oxygen-18 and deuterium values in summer than in winter (see Fig. S4). This was however true for all samples, including the weekly streamwater samples, indicating that additional evaporative fractionation during storage was probably negligible, and thus that the observed difference in isotopic values between summer and winter represents real-world variability. This was also confirmed by the analysis of deuterium excess relative to the GMWL, which yielded similar seasonal effects for all sample types (see Fig. 3 in the main paper). If the storage effects were significant, we would also expect substantially different deuterium excess values in bottles stored for longer times inside the autosampler in the field, compared to samples collected closer to the time of pick-up. This, however, was not the case; deuterium excess did not change with storage duration in the field (see Fig. 4 in the main paper).

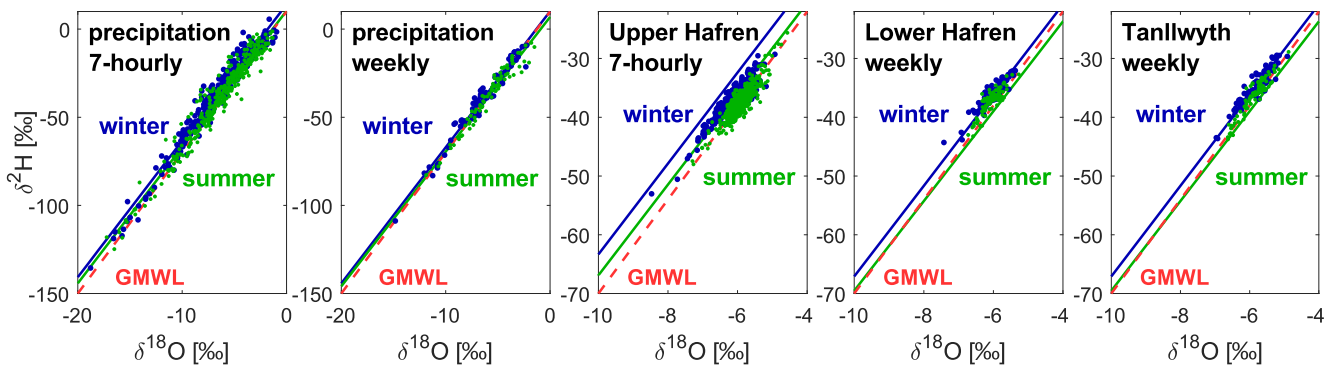


Figure S4: Separating summer and winter samples in dual-isotope plots shows that summer samples are isotopically slightly heavier than winter samples, and winter samples lie farther above the global meteoric water line. Weekly streamwater grab samples at Lower Hafren and Tanllwyth, which were not subject to evaporative fractionation in the field, exhibit this same pattern, suggesting that the 7-hourly samples did not undergo significant fractionation after collection.

Dry deposition correction for Plynlimon chloride data

Dry deposition of chloride artificially increases the variability of the precipitation input signal and can thus increase the damping observed between catchment input and output chloride concentrations. Samples affected by dry deposition need to be excluded from the calculation of new water fractions to avoid this bias. We filtered out dry deposition effects in the 7-hourly chloride data only. Dry deposition effects are likely to also be important in the weekly data, but much harder to detect and remove, because most of the weekly precipitation samples are likely to be affected.

This filtering approach takes into account that concentrations are expected to vary more when the amount of precipitation is small, that very high solute concentrations are unlikely, and that dry deposition is most likely to affect samples collected after extended periods without rain (see Fig. S5). Samples falling into any of these three categories (soft criteria), were flagged and their chloride concentrations were compared to the concentration of the sample collected immediately prior and/or afterward.

Only if the concentration of the sample in question was much higher than the neighboring samples did we consider it an outlier (hard criterion).

This filtering led to an exclusion of 6.6 % of the chloride samples from the 7-hourly data set. Removing these points reduced the chloride input-output flux ratio from 1.06 to 0.75. A mass balance of a passive tracer smaller than unity is not ideal. However, in our analysis it was more important that the variability of the input signal represented the one expected for a conservative tracer, than that its mass balance was correct. After the dry deposition filtering, the new water fractions calculated from chloride and stable isotopes were broadly similar across a range of precipitation thresholds (Fig. 8 in the main text), over different sampling intervals (Fig. 9 in the main text), and with/without discharge-weighting (Table 1 in the main text).

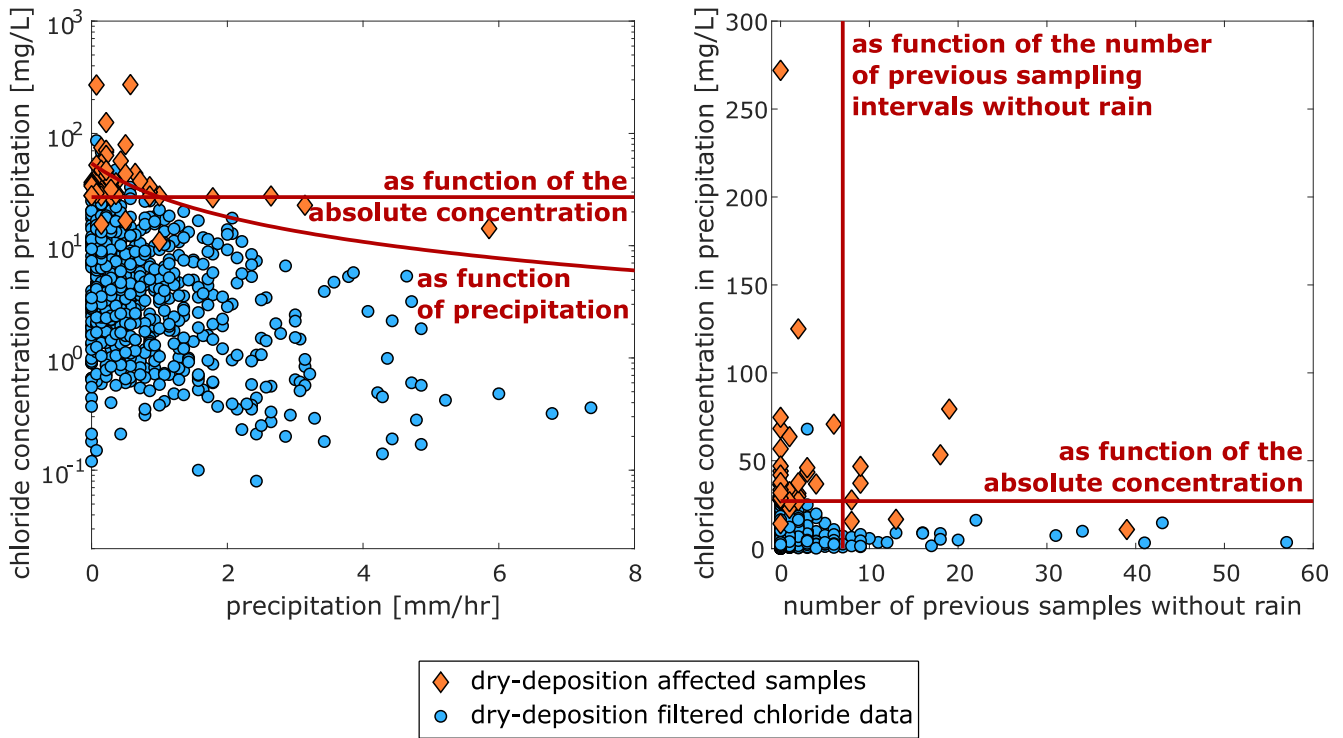


Figure S5: Chloride concentrations in precipitation were tested for effects of dry deposition. We assumed that dry deposition would lead to exceptionally high absolute concentrations, lead to very high concentrations at low precipitation rates, or play an important role after longer periods without precipitation. These acted as soft criteria (shown as red lines in the above figures), and flagged samples were removed if their concentrations were higher than the previous or following samples by more than twice the median concentration of all samples. Samples considered to be affected by dry deposition and removed from the data set are shown in orange (diamonds), and the final data set of 7-hourly chloride concentrations in precipitation is shown in blue (circles).

Comparison of transit time distributions from spectral fitting and ensemble hydrograph separation

We estimated transit time distributions from the power spectra of the tracer time series. The gamma model, when multiplied by the precipitation power spectrum, fitted the streamwater tracer power spectrum closely (Fig. 13a,b in the main text, and Fig. S6a,b and S7a,b). The fitted gamma parameters yielded transit time distributions that correspond closely to those estimated by ensemble hydrograph separation, both for 7-hourly (Fig. 13c,d in the main text) and weekly data (Fig. S6c,d and S7c,d).

Nevertheless, mean transit times calculated by multiplying the scale factors θ by the shape factors k of the fitted gamma distributions differed substantially, with longer mean transit times estimated from oxygen-18 than from deuterium (Table S1).

Table S1: Shape factors k (\pm standard errors) and scale factors θ (\pm standard errors) of the fitted gamma distributions, and resulting mean transit times (\pm standard errors) derived from unweighted and volume-weighted spectra of oxygen-18 and deuterium at Upper Hafren (7-hourly sampling), Lower Hafren (weekly sampling), and Tanllwyth (weekly sampling).

			Scale factor θ (yr)	Shape factor k (-)	Mean transit time (yr)	Ratio ^a
Upper Hafren (7-hourly sampling):						
Oxygen-18	—	unweighted	80.540 ± 16.718	0.276 ± 0.005	22.209 ± 4.627	35.124
Deuterium	—	unweighted	1.353 ± 0.128	0.467 ± 0.006	0.632 ± 0.060	
Oxygen-18	—	volume-weighted	0.323 ± 0.025	0.434 ± 0.005	0.140 ± 0.011	1.708
Deuterium	—	volume-weighted	0.159 ± 0.012	0.517 ± 0.007	0.082 ± 0.006	
Lower Hafren (weekly sampling):						
Oxygen-18	—	unweighted	9.022 ± 2.557	0.448 ± 0.020	4.041 ± 1.159	1.896
Deuterium	—	unweighted	4.094 ± 0.821	0.521 ± 0.019	2.131 ± 0.434	
Oxygen-18	—	volume-weighted	6.331 ± 1.840	0.398 ± 0.019	2.519 ± 0.742	2.102
Deuterium	—	volume-weighted	2.647 ± 0.526	0.453 ± 0.017	1.199 ± 0.242	
Tanllwyth (weekly sampling):						
Oxygen-18	—	unweighted	1.731 ± 0.300	0.552 ± 0.020	0.955 ± 0.169	1.516
Deuterium	—	unweighted	0.994 ± 0.134	0.634 ± 0.020	0.630 ± 0.087	
Oxygen-18	—	volume-weighted	2.161 ± 0.462	0.463 ± 0.020	1.000 ± 0.218	2.270
Deuterium	—	volume-weighted	0.816 ± 0.117	0.540 ± 0.019	0.440 ± 0.065	

^a ratio between mean transit times determined from oxygen-18 and deuterium.

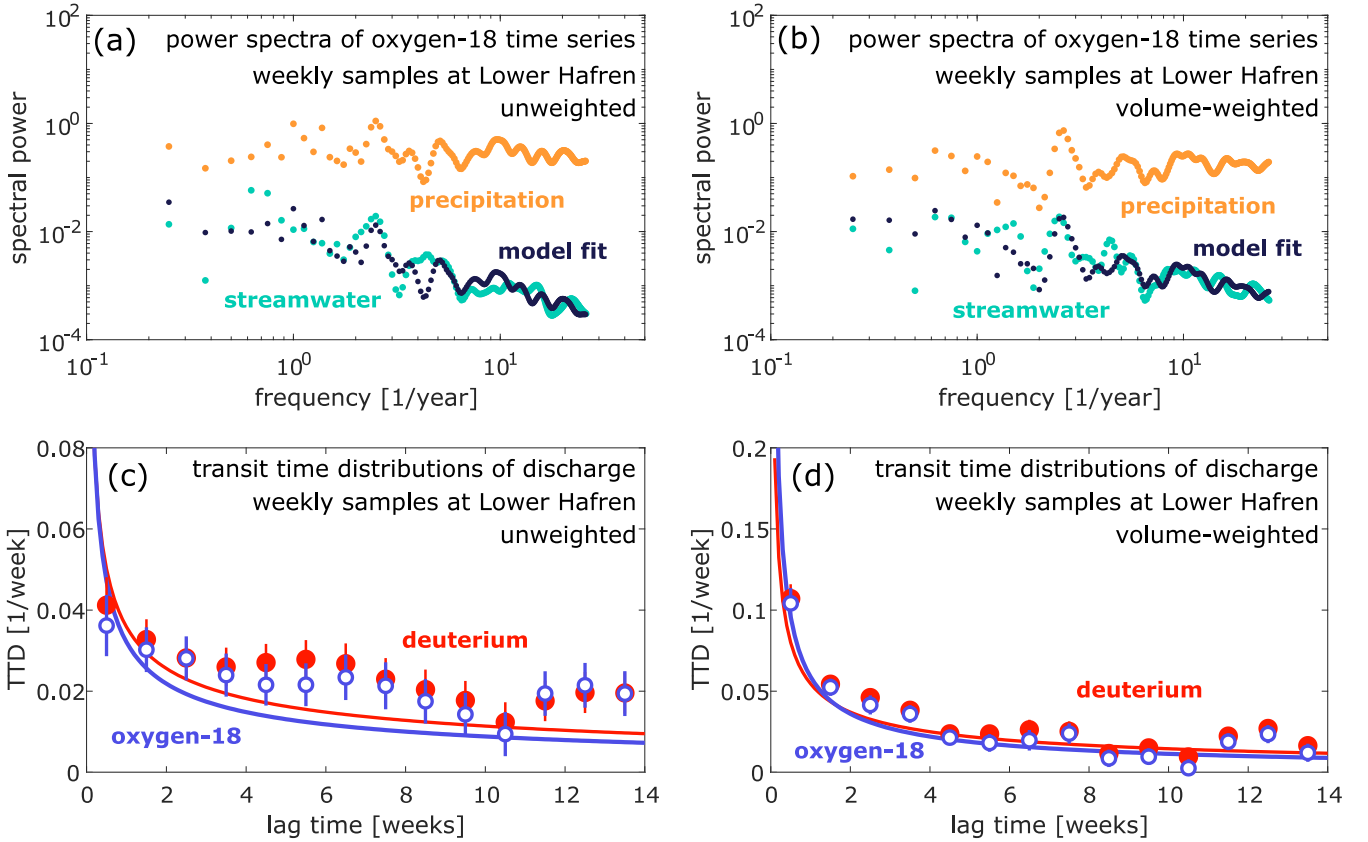


Figure S6: Fits of convolution models to unweighted (a) and volume-weighted (b) power spectra of oxygen-18 in precipitation and streamwater at Lower Hafren, and gamma distributions estimated by spectral fitting for both deuterium (red lines) and oxygen-18 (blue lines), compared to ensemble hydrograph separation estimates (dots, with standard errors) of unweighted (c) and volume-weighted (d) transit time distributions.

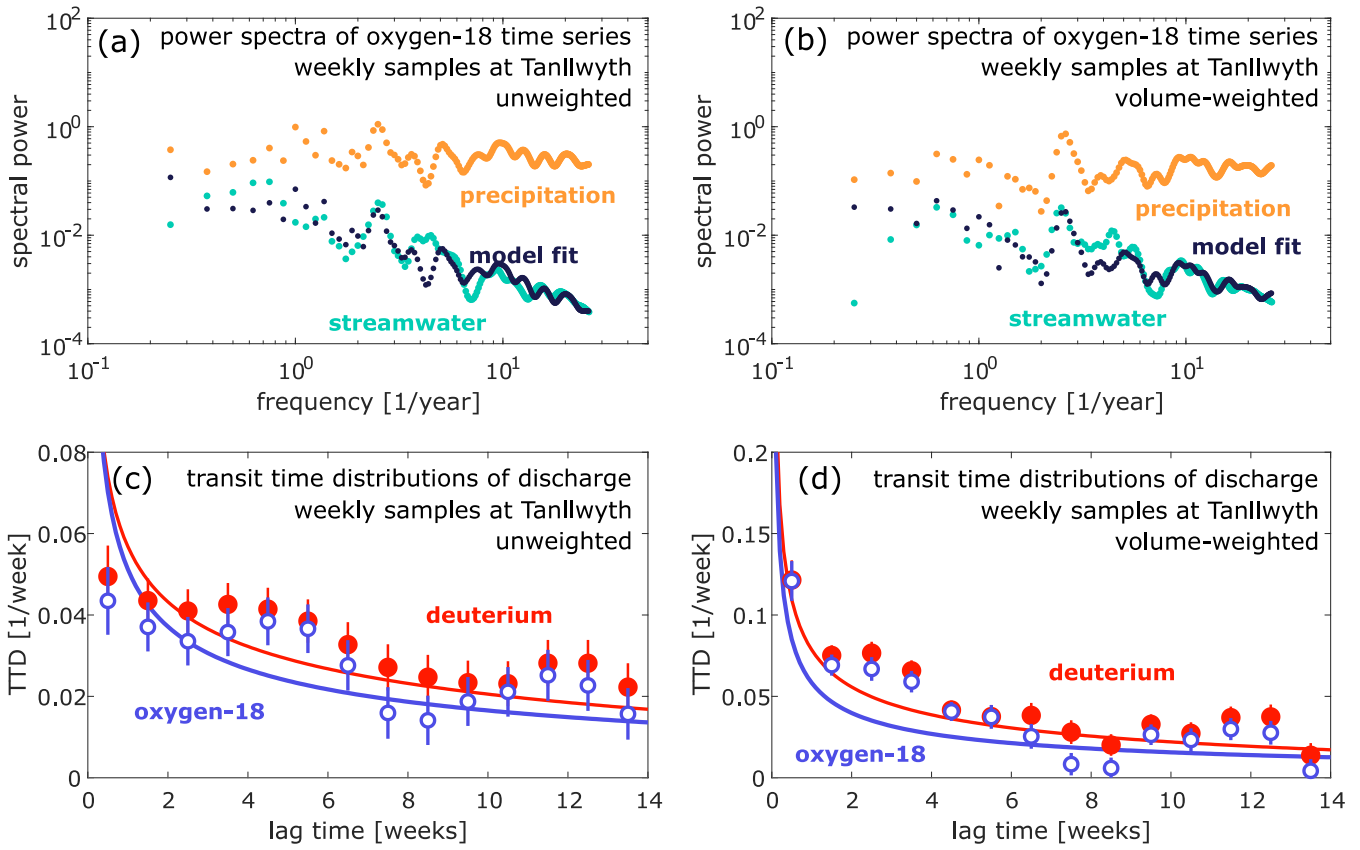


Figure S7: Fits of convolution models to unweighted (a) and volume-weighted (b) power spectra of oxygen-18 in precipitation and streamwater at Tanllwyth, and gamma distributions estimated by spectral fitting for both deuterium (red lines) and oxygen-18 (blue lines), compared to ensemble hydrograph separation estimates (dots, with standard errors) of unweighted (c) and volume-weighted (d) transit time distributions.

References

Kirchner, J. W.: Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: theory and benchmark tests, *Hydrol. Earth Syst. Sci.*, 23, 303-349, <https://doi.org/10.5194/hess-23-303-2019>, 2019.