Effects of reduced stream water isotope data availability on modelled isotope time series

Fatemeh Babaei[1,2], Meenakshi Arora[2], Heye Bogena[3], Julian Klaus[1] and Andrew W. Western[2]

1 University of Bonn, Department of Geography, Hydrology, Bonn, Germany

2 University of Melbourne, Department of Infrastructure Engineering, Hydrology and Water Resources, Melbourne, Australia. Email: fbabaei@student.unimelb.edu.au

3 Forschungszentrum Jülich GmbH, Agrosphere Institute (IBG-3), 52425 Jülich, Germany

Key points

- Weekly, fortnightly and monthly stream water isotope datasets were used to test the effect of data availability on the selection of representative modelled isotope time series.
- The simulation selected from fortnightly data was close to weekly, while the monthly isotope data showed weaker agreement with the weekly benchmark.
- Reduced isotope records can therefore alter which simulation is selected and how well it reproduces broader stream water behaviour.

Abstract

Stable water isotopes are widely used to evaluate how well hydrological models reproduce observed catchment transit time and flow path behaviour, but the available isotope record is often limited in temporal density. Here, we examined how different temporal resolutions of isotope time series affect modelled stream isotope simulations in the Weierbach catchment, Luxembourg. Weekly, fortnightly, and monthly subsets were derived from the available stream δ^2H record, and the best matching simulation was identified for each subset from a previously saved ensemble. The fortnightly- and monthly-based simulations were then compared against the weekly isotope observations, which served as a common benchmark. The weekly-based simulation achieved an NSE of 0.47, the fortnightly-based simulation was 0.46, and the monthly-based simulation was lower with an NSE of 0.19. The time-series comparison showed that the monthly-based simulation departed more strongly from the observed dynamics, whereas the fortnightly-based simulation had less deviation from the weekly benchmark. These results indicate that reduced isotope data availability can affect which simulation in a Monte Carlo set is identified as most representative and how well it reproduces broader stream water behaviour. Moderate reduction (fortnightly data) in isotope data availability may still support a useful model representation, but stronger reduction (monthly data) can noticeably weaken consistency with observed isotope dynamics. The results are relevant for water quality and catchment management applications where isotope time series are used to understand catchment residence time and guide hydrological model evaluation under limited data availability.

Keywords

Stable water isotopes, stream isotope time series, model simulation, transport dynamics, observation frequency

Introduction

Stable water isotopes are widely used to compare observed and simulated stream water behaviour because isotope time series integrate information on how catchments store, mix, and release water. This is important not only for hydrological interpretation, but also for broader water-quality and water-resources questions that depend on realistic simulation of stream water composition. In practice, however, isotope records are often limited in temporal density, and previous work has shown that the amount of isotope data available can influence model performance and the realism of the resulting simulations (Stevenson et al., 2021). Here, we compare modelled stream isotope time series obtained from weekly, fortnightly, and monthly observation subsets in order to assess how reduced isotope data availability affects the identified simulation and its agreement with a denser weekly benchmark.

Methods

The study uses isotope data from the Weierbach catchment in Luxembourg (Figure 1), a small forested headwater catchment that has been extensively monitored and used in previous studies on streamflow generation, isotope dynamics, and catchment transport modelling. The broader dataset includes hydrometric variables and water isotope observations. Previous work has shown that the high-frequency stream δ^2H record captures a wide range of hydrological conditions in the catchment (Hissler et al., 2021). Rainfall, discharge, and evapotranspiration data formed part of the existing modelling workflow and provide the hydrometric context for the stream isotope simulations (Figure 2).

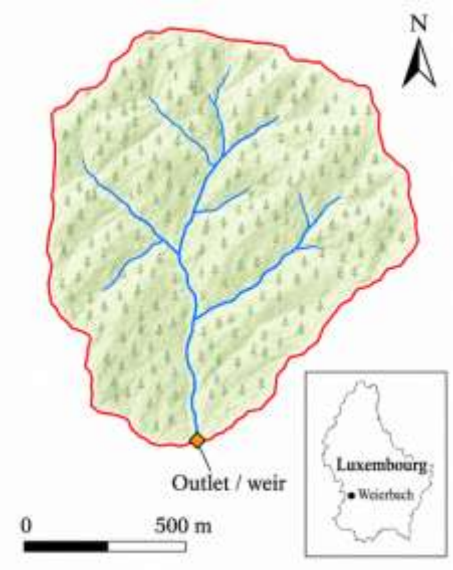

Figure 1 Map of Weierbach catchment and its location in Luxembourg

In this study, we used stream δ^2H simulation time series from a Monte Carlo ensemble of StorAge Selection (SAS) (Benettin et al., 2022) model runs for the Weierbach catchment (Rodriguez & Klaus, 2019). In SAS models, the distribution of ages in the catchment is tracked through a mass balance for each age and SAS functions determine the selection of water from that age distribution that contributes to outflows either to the stream or atmosphere. In this case three SAS functions, one for each of three runoff components, are weighted together. Model parameters include the parameters control the storage threshold, event-water contribution, component weights, and the shape and scale of the gamma-based age-selection SAS functions (Table 1); full equations and parameter definitions are given by Rodriguez and Klaus (2019). The Python implementation followed the modelling framework and parameterization described by Rodriguez and Klaus (2019). The ensemble was generated from a range of alternative SAS parameter sets, and each run produced a complete simulated stream isotope time series. The saved ensemble simulations were re-evaluated under reduced observation subsets. Three reduced stream isotope time series were defined from the available stream observations: weekly, fortnightly, and monthly. For each scheme, one observed sample closest to the target interval was retained, resulting in 96 weekly, 52 fortnightly, and 24 monthly observations. These reduced subsets were used to represent progressively lower isotope data availability under practical monitoring conditions.

Table 1 Brief description and units of the SAS parameters

Parameter	Unit	Description
Sth	mm	Storage threshold relative to minimum storage, separating dry and wet conditions
dSth	mm	Storage-variation threshold controlling activation of flashy/event responses
Su	mm	Range of the uniformly distributed young-water SAS component
f0	–	Young-water coefficient during dry conditions
lambda1	–	Maximum value of the time-varying young-water component weight
lambda2	–	Weight of the second runoff SAS component
mu2	mm	Mean parameter of the gamma-distributed second runoff component
theta2	mm	Scale parameter of the gamma-distributed second runoff component
mu3	mm	Mean parameter of the gamma-distributed third runoff component
theta3	mm	Scale parameter of the gamma-distributed third runoff component
muET	mm	Mean parameter of the gamma-distributed evapotranspiration SAS function
thetaET	mm	Scale parameter of the gamma-distributed evapotranspiration SAS function

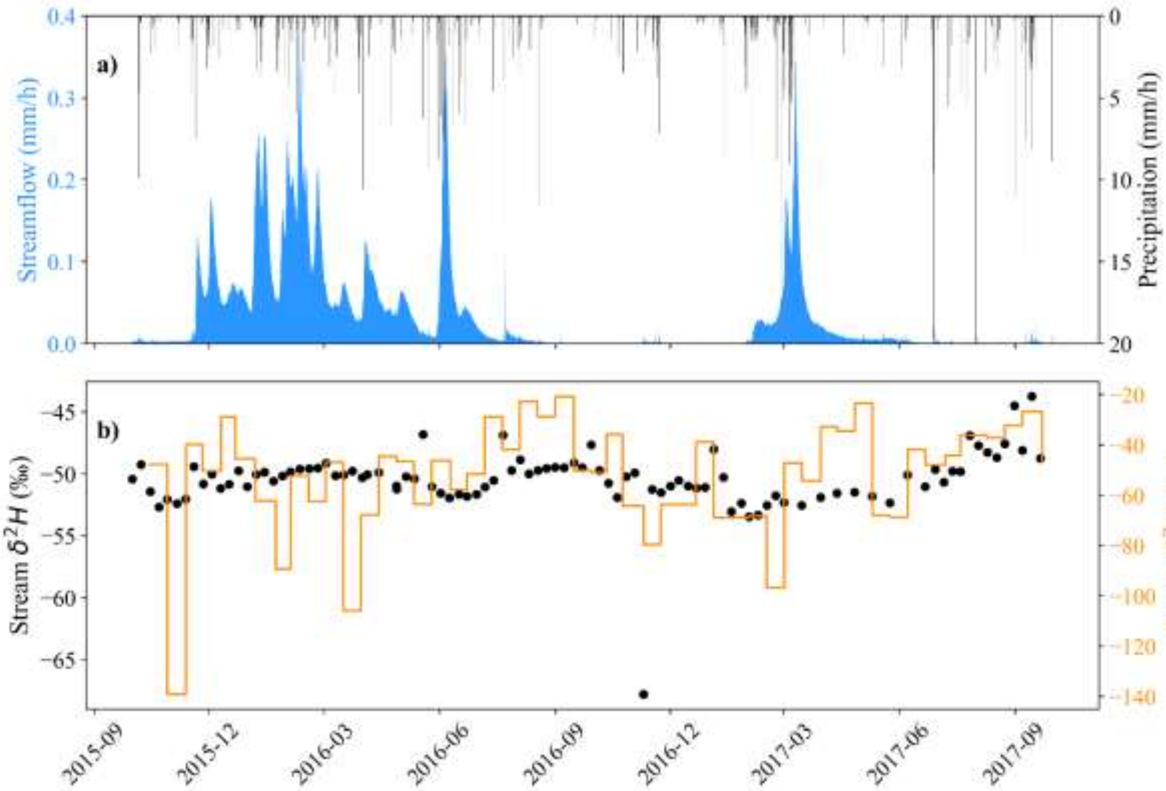

Figure 2 Hydrometric and isotope input data for the main SAS model run period: a) streamflow and precipitation input data, and b) weekly stream δ^2H observations together with precipitation δ^2H input

The SAS model was run at a 4-hourly time step. Monte Carlo simulations were used to explore the parameter space, producing 12,096 total ensemble simulations. The same prior Monte Carlo parameter set was used for the weekly, fortnightly, and monthly analyses to avoid differences caused by parameter resampling. For each observation subset, all saved ensemble simulations were screened against the corresponding observation set using the Nash–Sutcliffe efficiency (NSE), and the best matching simulation for that subset was identified. Root mean square error (RMSE) was also calculated to quantify the magnitude of isotope mismatches between observed and simulated stream δ2H values. The NSE of each model selected using the weekly, fortnightly, and monthly data sets was then re-calculated against the weekly observations, which were used as a common benchmark. The simulation identified from the weekly data set was treated as the reference case because weekly observations represented the densest sampling considered in this study.

Results

The comparison showed that reduced isotope data availability affected the identified stream isotope simulation. Relative to the weekly benchmark, the fortnightly data resulted in a model with similar performance (NSE = 0.46) to that based on the weekly isotope observations (NSE = 0.47). In contrast, using the monthly data resulted in a model with an NSE of only 0.19.

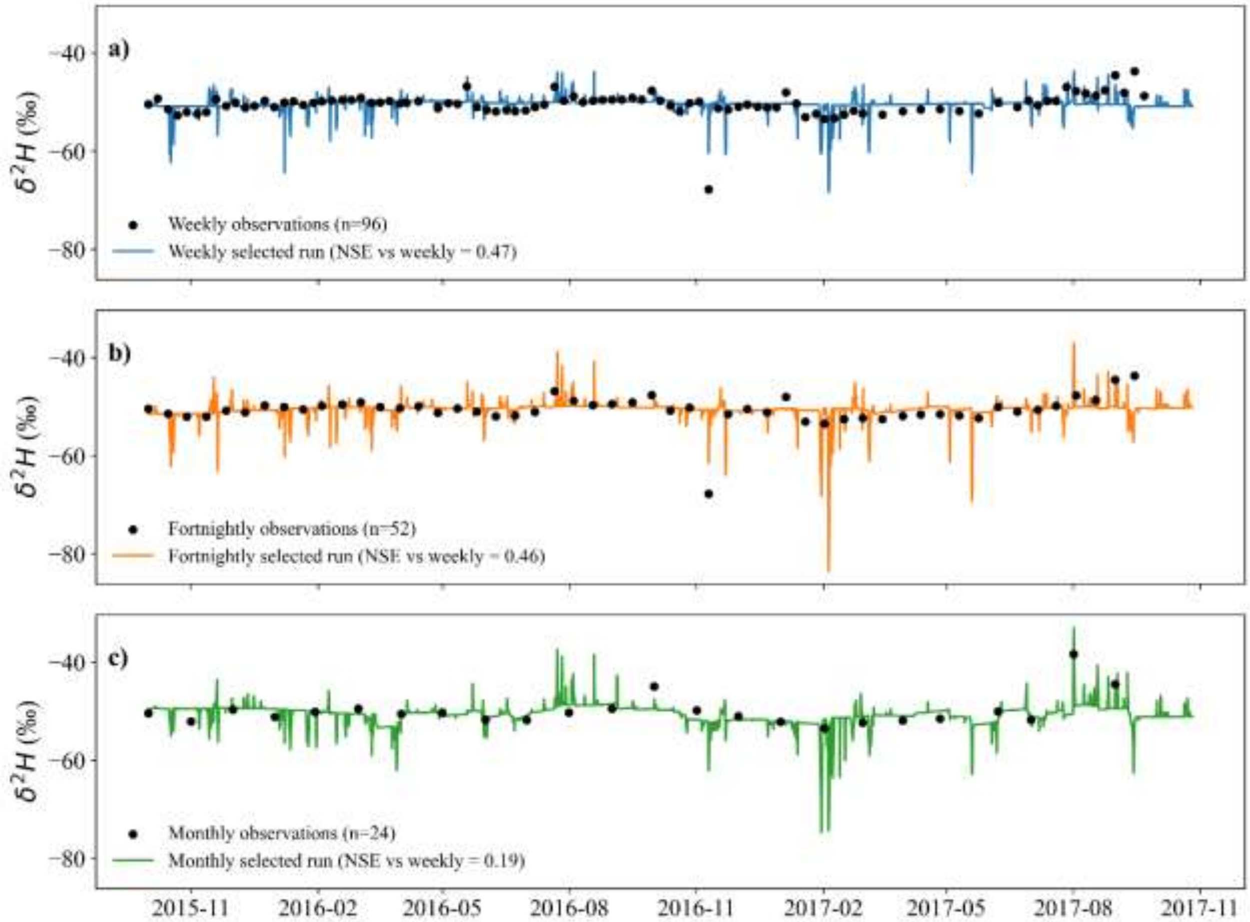

Figure 3 Simulated stream isotope time series for the weekly, fortnightly, and monthly observation subsets

The model selection based on fortnightly data remained broadly similar to that selected using the weekly data and reproduced the main seasonal pattern in stream δ^2H, including the relatively stable isotope values through March 2016 and the more enriched values toward the end of the record. In contrast, the model selected using the monthly data departed more strongly from the observed dynamics and produced larger mismatches, particularly during short term isotope fluctuations, where the monthly simulation missed some of the higher and lower δ^2H values visible in the weekly observations. Thus, reducing the isotope time series from weekly to fortnightly caused only limited deterioration in the identified simulation, whereas reduction to monthly led to a much clearer decline in agreement with the weekly benchmark (Figure 3).

Table 2 Selection and weekly benchmark performance metrics for the simulations selected using weekly, fortnightly, and monthly stream isotope observations

Selected simulation	Number of observations	NSE selection	RMSE (‰) selection	Weekly benchmark NSE	Weekly benchmark RMSE (‰)
Weekly	96	0.47	1.77	0.47	1.77
Fortnightly	52	0.55	2.03	0.46	1.79
Monthly	24	0.67	1.81	0.19	2.19

The results indicate that the effect of reduced isotope data availability cannot be assessed only from the fit to the subset used for model selection. Table 2 further shows that performance against the selection subset alone can be misleading. The monthly data had the highest selection NSE and a relatively low selection RMSE, but its RMSE increased to 2.19‰ against the weekly benchmark, indicating larger isotope mismatches under the denser observation record. In contrast, the fortnightly data showed only a small change in RMSE between the selection subset and weekly benchmark, suggesting that it preserved the broader isotope dynamics more consistently. The absolute deviation comparison also showed larger departures of the monthly data from the weekly benchmark than the fortnightly (Figure 4). This suggests that reduced isotope records can alter which simulation is identified as most representative and weaken its agreement with broader stream water behaviour.

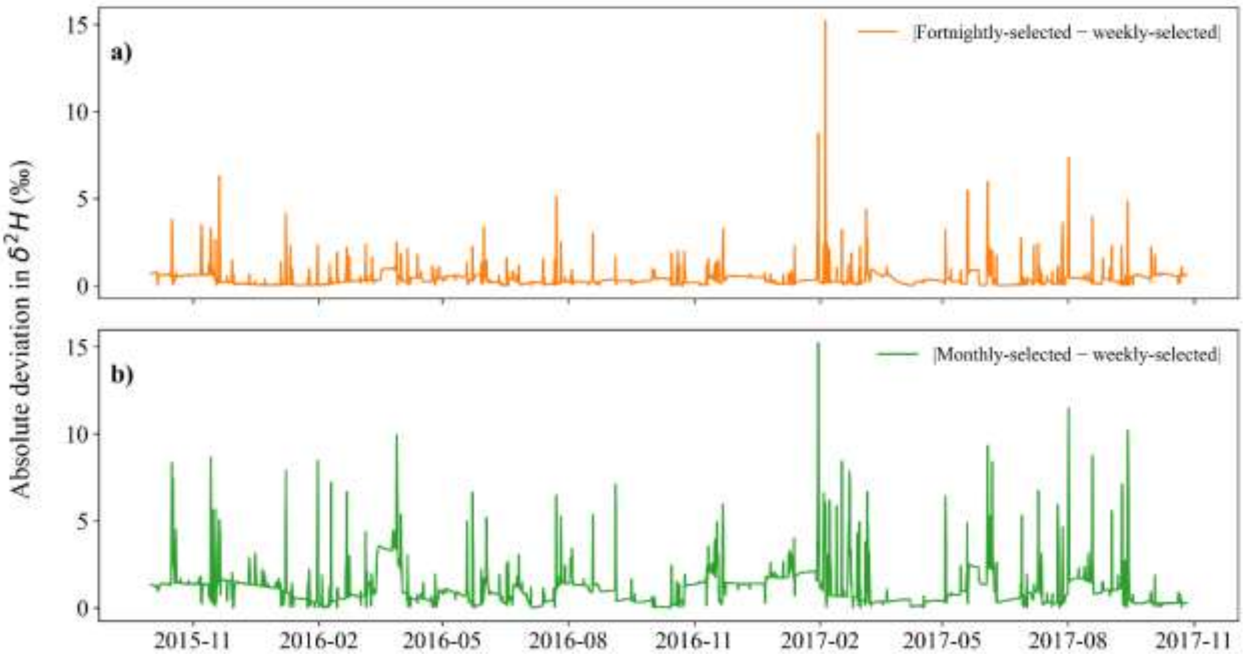

Figure 4 Absolute deviations of a) fortnightly and b) monthly data simulations from the weekly benchmark simulation

To examine whether reduced sampling frequency also changed the identified parameter values, the parameters of the weekly, fortnightly, and monthly-selected simulations were normalized to their prior ranges and compared (Figure 5). The monthly simulation differed more strongly from the weekly reference for several parameters, particularly Sth, Su, lambda1, lambda2, mu3, theta3, and thetaET. This indicates that monthly sampling did not only reduce benchmark performance, but also selected a different region of the Monte Carlo parameter space. The fortnightly simulation also differed from weekly simulation for some parameters, but these differences had a smaller effect on the weekly-benchmark isotope performance.

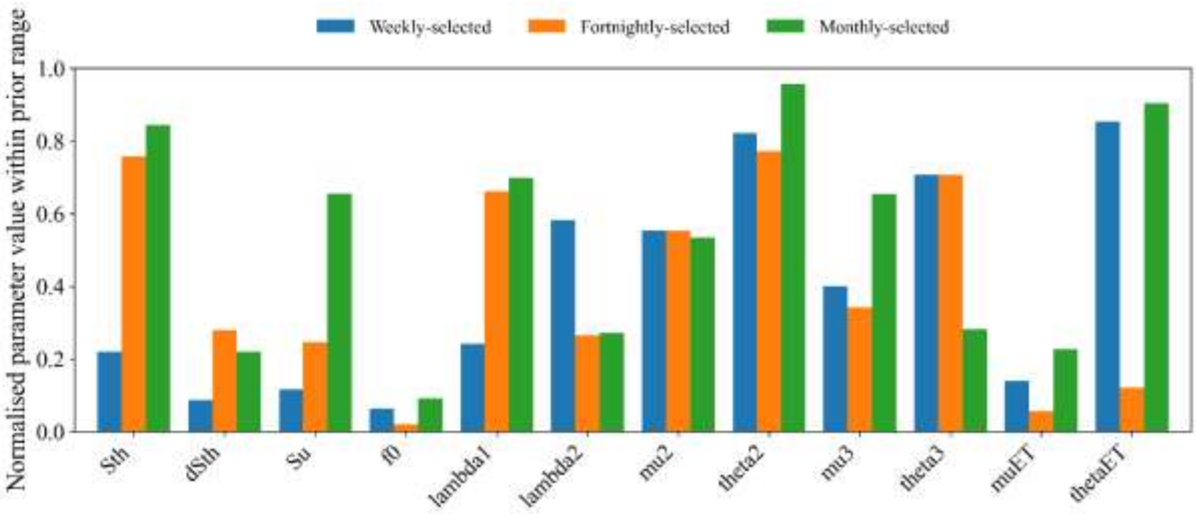

Figure 5 Normalised SAS parameter values of the simulations selected using weekly, fortnightly, and monthly stream isotope observations. Parameter values were normalised to their prior ranges to allow comparison across parameters with different units and magnitudes

Discussion

The comparison suggests that for this catchment and study period, the practical consequences of reduced isotope data availability are not linear. In the present case, moving from weekly to fortnightly observations did not strongly alter the identified simulation, whereas the shift to monthly data changed the simulated isotope behaviour more noticeably. This indicates that some reduction in temporal density may still preserve much of the structure needed to identify a plausible stream isotope simulation, but beyond a certain point the reduced record no longer captures enough variability to guide model selection reliably.

The results therefore highlight a sampling-design trade-off between data density, sampling effort, and simulation accuracy. In this case, fortnightly sampling produced nearly the same benchmark performance as weekly sampling, while requiring substantially fewer observations. Monthly sampling, however, led to a much larger loss of agreement with the weekly benchmark. The parameter comparison supports this interpretation, as the monthly-selected simulation was associated with a more distinct parameter combination than the fortnightly-selected simulation, suggesting that sparse sampling can affect both model performance and parameter-set identification.

For applications concerned with stream water composition, solute transport, or broader water quality interpretation, this trade-off is an important distinction because the selected simulation is then based on a less complete representation of observed dynamics. These findings suggest that moderate reductions in sampling frequency may be effective, but that performance loss can increase more rapidly once the sampling interval becomes too coarse. However, this result is based on one catchment and one study period, and further tests across different sites, periods, and hydrological conditions would be needed before making general sampling recommendations.

Conclusions

This study shows that reducing the available stream isotope time series can change the simulation identified as most representative of observed stream water behaviour. Using weekly observations as the reference case, model selection based on fortnightly data remained close to the weekly benchmark, whereas the monthly-based simulation showed a marked reduction in performance. The results suggest that moderate reduction in isotope data availability may still support a useful model representation, but stronger reduction can noticeably weaken consistency with observed isotope dynamics. For this case study, the results show that reduced isotope records can still identify useful simulations, but only while enough temporal information is retained to preserve consistency with the denser benchmark, highlighting the importance of data availability when using isotope time series to guide hydrological model evaluation. These results also have implications for sampling design, although analysis of a broader range of catchments would be required to produce general results.

References

Benettin, P., Rodriguez, N. B., Sprenger, M., Kim, M., Klaus, J., Harman, C. J., Van Der Velde, Y., Hrachowitz, M., Botter, G., McGuire, K. J., Kirchner, J. W., Rinaldo, A., & McDonnell, J. J. (2022). Transit Time Estimation in Catchments: Recent Developments and Future Directions. *Water Resources Research*, *58*(11), e2022WR033096. https://doi.org/10.1029/2022WR033096

Hissler, C., Martínez-Carreras, N., Barnich, F., Gourdol, L., Iffly, J. F., Juilleret, J., Klaus, J., & Pfister, L. (2021). The Weierbach experimental catchment in Luxembourg: A decade of critical zone monitoring in a temperate forest - from hydrological investigations to ecohydrological perspectives. *Hydrological Processes*, *35*(5), e14140. https://doi.org/10.1002/hyp.14140

Rodriguez, N. B., & Klaus, J. (2019). Catchment Travel Times From Composite StorAge Selection Functions Representing the Superposition of Streamflow Generation Processes. *Water Resources Research*, *55*(11), 9292–9314. https://doi.org/10.1029/2019WR024973

Stevenson, J. L., Birkel, C., Neill, A. J., Tetzlaff, D., & Soulsby, C. (2021). Effects of streamflow isotope sampling strategies on the calibration of a tracer-aided rainfall-runoff model. *Hydrological Processes*, *35*(6), e14223. https://doi.org/10.1002/hyp.14223