



HydroMix v1.0: a new Bayesian mixing framework for attributing uncertain hydrological sources

Harsh Beria¹, Joshua R. Larsen², Anthony Michelin¹, Natalie C. Ceperley^{1,a}, and Bettina Schaeffli^{1,a}

¹Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

²School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

^anow at: Institute of Geography, University of Berne, Bern, Switzerland

Correspondence: Harsh Beria (harsh.beria@unil.ch)

Received: 15 March 2019 – Discussion started: 28 March 2019

Revised: 17 March 2020 – Accepted: 27 April 2020 – Published: 27 May 2020

Abstract. Tracers have been used for over half a century in hydrology to quantify water sources with the help of mixing models. In this paper, we build on classic Bayesian methods to quantify uncertainty in mixing ratios. Such methods infer the probability density function (PDF) of the mixing ratios by formulating PDFs for the source and target concentrations and inferring the underlying mixing ratios via Monte Carlo sampling. However, collected hydrological samples are rarely abundant enough to robustly fit a PDF to the source concentrations. Our approach, called HydroMix, solves the linear mixing problem in a Bayesian inference framework wherein the likelihood is formulated for the error between observed and modeled target variables, which corresponds to the parameter inference setup commonly used in hydrological models. To address small sample sizes, every combination of source samples is mixed with every target tracer concentration. Using a series of synthetic case studies, we evaluate the performance of HydroMix using a Markov chain Monte Carlo sampler. We then use HydroMix to show that snowmelt accounts for around 61 % of groundwater recharge in a Swiss Alpine catchment (Vallon de Nant), despite snowfall only accounting for 40 %–45 % of the annual precipitation. Using this example, we then demonstrate the flexibility of this approach to account for uncertainties in source characterization due to different hydrological processes. We also address an important bias in mixing models that arises when there is a large divergence between the number of collected source samples and their flux magnitudes. HydroMix can account for this bias by using composite likelihood functions that effectively weight the relative magnitude of source

fluxes. The primary application target of this framework is hydrology, but it is by no means limited to this field.

1 Introduction

Most water resources are a mixture of different water sources that have traveled via distinct flow paths in the landscape (e.g., streams, lakes, groundwater). A key challenge in hydrology is to infer source contributions to understand the flow paths to a given water body using a source attribution technique. A classic example is the two-component hydrograph separation model to quantify the proportion of groundwater and rainfall in streamflow, often referred to as “pre-event” water vs. “event” water (Burns et al., 2001; Klaus and McDonnell, 2013; Schmieder et al., 2016). Other examples include estimating the proportional contribution of rainfall and snowmelt to groundwater recharge (Beria et al., 2018; Jasechko et al., 2017; Jeelani et al., 2010), fog to the amount of throughfall (Scholl et al., 2011, 2002; Uehara and Kume, 2012), and soil moisture (at varying depths) and groundwater to vegetation water use (Ehleringer and Dawson, 1992; Evaristo et al., 2017; Rothfuss and Javaux, 2017).

The primary goal of such attribution in hydrology is to infer the contribution of different sources to a target water body; the tracer can be an observable compound like a dye, a conservative solute, or even a proxy for chemical composition such as electrical conductivity. The key requirement is that the concentration of the tracer is distinguishable between different sources. The stable isotope compositions of hydrogen and oxygen in water (subsequently referred to as

“stable isotopes of water”) are used as tracers in hydrology. Other commonly used tracers include electrical conductivity (Hoeg et al., 2000; Laudon and Slaymaker, 1997; Lopes et al., 2018; Pellerin et al., 2007; Weijs et al., 2013) and conservative geochemical solutes such as chloride and silica (Rice and Hornberger, 1998; Wels et al., 1991).

Classically, attribution analysis is done by assigning an average tracer concentration to each source, typically estimated from time or space averages of observed field data (Maule et al., 1994; Winograd et al., 1998), and then solving a series of linear equations. In order to express uncertainty in the attribution analysis, a tracer-based hydrograph separation approach was first proposed in the work of Genereux (1998) and has subsequently been used in many studies (Genereux et al., 2002; Koutsouris and Lyon, 2018; Zhu et al., 2019). Bayesian mixing approaches offer a useful alternative to classic hydrograph separation, as Bayesian approaches explicitly acknowledge the temporal variability of source tracer concentrations estimated from observed samples (Barbeta and Peñuelas, 2017; Blake et al., 2018). Rather than a single estimate of source contributions, Bayesian approaches yield full probability density functions (PDFs) of the fraction of different sources in the target mixture (Parnell et al., 2010; Stock et al., 2018), hereafter referred to as “mixing ratios”.

Bayesian mixing was first developed in ecology to estimate the proportion of different food sources to animal diets (Parnell et al., 2010; Stock et al., 2018). Hydrological applications of such models are still rare (Blake et al., 2018; Evaristo et al., 2016, 2017; Oerter et al., 2019). In a Bayesian mixing model, a statistical distribution is fitted to both the measured source tracer concentrations and to the measured tracer concentrations from the target (e.g., river, groundwater, vegetation). The distribution of the mixing ratios is then inferred via Bayesian inference. With recent advances in probabilistic programming languages like Stan (Carpenter et al., 2017), Bayesian inference has become a relatively simple task.

However, the key limitation with the above approach is that the source compositions are assumed to come from standard statistical distributions. Typically, the sources are assumed to be drawn from Gaussian distributions, which can be fully characterized by the mean and variance of the data available for each source (Stock et al., 2018). This limits both the potential applicability and the insights that can be gained from tracer information in hydrology because the sample mean and variance may not accurately reflect the statistical properties of the actual source composition, and the Gaussian approach represents an unnecessary simplification in cases in which a large amount of information on source composition is available.

An additional complication in hydrology comes from the fact that observed point-scale samples do not necessarily capture the tracer concentrations in the actual sources, which are distributed heterogeneously in space and whose contribution can be temporally variable depending on the state of

the catchment (Harman, 2015). For instance, if we were to characterize the contribution of snowmelt to groundwater, we would need to capture (1) the temporal evolution of the isotopic ratio of snowmelt, which strongly varies in space (Beria et al., 2018; Earman et al., 2006), and (2) the temporal evolution of the area actually covered by snow. This spatially and temporally distributed nature of the sources can be hard to account for in both analytical and Bayesian mixing approaches.

To overcome the limitations of source heterogeneity and the previously discussed restriction to Gaussian distributions, we present a new mixing approach for hydrological applications called HydroMix. This approach does not require a parametric description of observed source or target tracer concentrations. Instead, HydroMix formulates the linear mixing problem in a Bayesian inference framework similar to hydrological rainfall–runoff models (Kavetski et al., 2006a), wherein the mixing ratios of the different sources are treated as model parameters. Multiple model parameters can be inferred in such a setup, allowing for the parameterization of additional hydrologic processes that can modify source tracer concentrations (shown in Sect. 3.5). A more detailed account of the advantages and limitations of this new approach is given in Sect. 5.

In this paper, we first describe the theoretical details of HydroMix for a simple case study with two sources, one mixture and one tracer (Sect. 2). Section 3 presents synthetic and real-world case studies that demonstrate the accuracy, robustness, and flexibility of HydroMix. In the synthetic case study, we use a conceptual hydrologic model to simulate tracer concentrations. We also introduce a composite likelihood function that accounts for the magnitude of the different source fluxes. The real-world case study applies HydroMix in a high-elevation headwater catchment in Switzerland. The results of these applications are presented in Sect. 4 before summarizing the main outcomes, applicability, and limitations of HydroMix in Sect. 5.

2 Model description and implementation

A system with n sources mixing linearly in a target water body can be written as

$$\rho_1 S_1^k + \rho_2 S_2^k + \dots + \rho_n S_n^k = Y^k, \quad (1)$$

where Y^k is the concentration of the k th tracer in the target mixture, and S_i^k is the concentration of the k th tracer in source i ; ρ_i ($i = 1, \dots, n$) represents the fractions of all sources in the mixture, with $\sum_{i=1}^n \rho_i = 1$, corresponding to the aggregation of different sources in the mixture. In order to solve this system of linear equations, $n - 1$ different tracers are required.

Section 2.1 details the general modeling approach for a simplified system with two sources and one tracer. This is followed by a detailed discussion on the choice of the parameter inference approach used.

Linear mixing model with non-concomitant observed data

For a system with two sources that combine linearly to form a mixture, the mixing model can be formulated as

$$\rho S_1(t - \tau_1) + (1 - \rho) S_2(t - \tau_2) = Y(t), \tag{2}$$

where $S_1(t - \tau_1)$ is the tracer concentration in source 1 at time step $t - \tau_1$, $S_2(t - \tau_2)$ is the tracer concentration in source 2 at time step $t - \tau_2$, $Y(t)$ is the concentration of the mixture (i.e., the tracer concentration in the target) at time step t , ρ is the mixing ratio, and τ_i is the time delay between the time when source i enters the system and the time when it is observed in the mixture. As an example, for a case in which the two sources are snowmelt and rainfall and the mixture is groundwater, ρ represents the proportional groundwater recharged from snowmelt and τ represents the average time lag for rain or snowmelt to reach groundwater once they enter the soil. In other words, the time lag (τ) stands for any delay caused by tracer transport from the source to the output; we assume that the source components are conservative in nature.

The two parameters in this system, the mixing ratio (ρ) and the time delay (τ), can be inferred via classical Bayesian parameter inference, which is widely used in hydrology (Kavetski et al., 2006a, b; Schaeffli and Kavetski, 2017). This implies taking an observed time series of the target (e.g., the tracer concentration in groundwater) and building a vector of model residuals:

$$\varepsilon_t = \tilde{Y}_t - \hat{Y}_t, \tag{3}$$

where \tilde{Y}_t represents the observed mixture concentration and \hat{Y}_t represents the simulated mixture concentration. However, in real environmental systems like that of groundwater recharge from rainfall and snowmelt, there are four major difficulties that can prevent the inference of ρ and τ from the observed data.

- i. ρ and τ strongly vary in time depending on catchment conditions such as soil moisture (as previously discussed in the context of the “inverse storage effect”; Benettin et al., 2017; Harman, 2015).
- ii. Long time series of the tracer concentration in both the sources and mixture are rare.
- iii. The effect of seasonality in precipitation can make the inference of τ very difficult in the case that the goal is to understand intra-annual recharge dynamics.
- iv. The tracer concentrations in the different sources are generally measured at point scales, whereas the tracer concentration in the target integrates inputs over the entire source area.

Our practical solution to limitation (iv) is to assume that tracer concentrations in the two sources are functions of observable point processes:

$$S_i(t) = f_i(P_i(t)), \tag{4}$$

where the function f_i represents the transformation from the point to the catchment scale for source i . Limitation (iii) can be relaxed by assuming a long enough time step (e.g., long-term groundwater recharge dynamics), for which the observed samples are samples from the long-term ($\gg 1$ year) source and target compositions. This allows us to replace the time step t and $t + \tau$ with Δt and write Eq. (2) as

$$\rho S'_1(\Delta t) + (1 - \rho) S'_2(\Delta t) = Y'(\Delta t), \tag{5}$$

where the ' signifies the new time-integrated variables. Now, any observed point-scale tracer concentration p_i in a given source i or in the output (e.g., the isotopic ratio of snowmelt) can be assumed to represent a sample from a stationary process (from S'_1 , S'_2 , or Y'). This assumption is in fact implicitly underlying most of the existing hydrological mixing models in which point samples are used to characterize a spatial process and the time reference of the samples is discarded.

By utilizing all the available measurements $\{p_1^i\}_{i=1\dots n}$ and $\{p_2^j\}_{j=1\dots m}$ of the two sources in the above model, with n samples of source 1 and m samples of source 2, we can build $n \times m$ predictions and compare them with the q observed samples of the target as

$$\varepsilon_{ij}^k = \tilde{Y}_{\text{obs}}^k - \hat{Y}_{ij}, \tag{6}$$

where \tilde{Y}_{obs}^k is the k th observed target concentration out of a total number of q target concentrations.

Assuming that the residuals can be described with a Gaussian error model with a mean of zero and constant variance σ^2 ,

$$\varepsilon \sim N(0, \sigma^2), \tag{7}$$

we can compute the likelihood function of the residuals as the joint probability of all the residuals:

$$L_j(\tilde{Y}_{\text{obs}} | P_1, P_2, \theta) = \prod_{k=1}^q \prod_{j=1}^m \prod_{i=1}^n (2\pi\sigma^2)^{-0.5} \exp\left(-\frac{1}{2} \frac{(\tilde{Y}_{\text{obs}}^k - \hat{Y}_{ij})^2}{\sigma^2}\right), \tag{8}$$

where θ represents all the model parameters and P_i ($i = 1, 2$) is the observed point process (see Eq. 4). The above Gaussian error model could in principle be replaced with any other stochastic process. However, the Gaussian error model has been shown to be relatively robust in this kind of application (Lyon, 2013; Schaeffli and Kavetski, 2017).

In the case of linear mixing between two sources, the two model parameters considered at this stage are the mixing ratio ρ and the error variance σ^2 . The error variance can either be computed from the observed residuals or treated as a model parameter (Kuczera and Parent, 1998; Schaeffli et al., 2007). For the examples shown in this paper, the error variance is computed from the residuals.

In order to avoid numerical problems, we use the log-likelihood form of Eq. (8),

$$\log L_j(\tilde{Y}_{\text{obs}}|P_1, P_2, \theta) = \sum_{k=1}^q \sum_{j=1}^m \sum_{i=1}^n -0.5 \left[\log(2\pi\sigma^2) + \frac{(\tilde{Y}_{\text{obs}}^k - \hat{Y}_{ij})^2}{\sigma^2} \right], \quad (9)$$

for parameter inference in a Bayesian framework.

Following the general Bayes' equation, the posterior distribution of the model parameters can be written as

$$p(\theta|P_1, P_2, \tilde{Y}) = \frac{p(\tilde{Y}|\theta, P_1, P_2)p(\theta)}{p(\tilde{Y}|P_1, P_2)}, \quad (10)$$

where $p(\theta)$ is the prior distribution of the model parameters and $p(\tilde{Y}|\theta, P_1, P_2)$ is the likelihood function. The denominator of Eq. (10) can generally not be computed as that would require integration over the whole parameter space, which is computationally expensive, and that is why Eq. (10) is reduced to

$$p(\theta|P_1, P_2, \tilde{Y}) \propto p(\tilde{Y}|\theta, P_1, P_2)p(\theta). \quad (11)$$

Two methods are traditionally used in hydrology to sample from the posterior distribution from Eq. (11): Markov chain Monte Carlo (MCMC) sampling (Hastings, 1970; Metropolis and Ulam, 1949) and importance sampling (Glynn and Iglehart, 1989; Neal, 2001). In the case of MCMC sampling, a common approach is the Metropolis algorithm (Kuczera and Parent, 1998; Schaeffli et al., 2007; Vrugt et al., 2003). In importance sampling, the posterior distribution is obtained from weighted samples drawn from the so-called importance distribution. For typical multivariate hydrological problems, the only possible choices for the importance distribution are either uniform sampling over a hypercube or sampling from an over-dispersed multi-normal distribution (Kuczera and Parent, 1998). A stochastic process is defined as over-dispersed when the variance of the underlying distribution is greater than its mean (Inouye et al., 2017). The sampling distributions in such cases have large variance, allowing for sufficient sampling over the entire parameter range.

We implement an MCMC sampling algorithm using a Metropolis–Hastings (Hastings, 1970) criterion to infer the posterior distribution of the mixing ratio. For the synthetic case study (Sect. 3.1), we set up 10 parallel MCMC chains

to monitor convergence according to the classical Gelman–Rubin convergence criterion (Gelman and Rubin, 1992). Each chain is initiated by assigning a uniform prior distribution for the mixing ratio, and the mixing ratio varies between 0 and 1. For the subsequent case studies, we use importance sampling for the sake of simplicity. The prior distributions of additional model parameters (if applicable) are discussed in the corresponding case study section. Apart from the prior distribution of the model parameters, HydroMix requires tracer concentration of the different sources and of the mixture. The error model variance is not jointly inferred with other model parameters but calculated for each sample parameter set from the residuals according to Eq. (6).

3 Case studies

We provide a comprehensive overview of the performance of HydroMix based on a set of synthetic case studies (case studies in Sect. 3.1 and 3.2) and a real-world application to demonstrate the practical relevance for hydrologic applications (case studies in Sect. 3.4 and 3.5). The first case study demonstrates the ability of HydroMix to converge on the correct posterior distribution for synthetically generated data. The second case study uses a synthetic dataset of rain, snow, and groundwater isotopic ratios using a conceptual hydrologic model and compares the results of HydroMix to the actual mixing ratios assumed to generate the dataset. It then weights the source samples and evaluates the effect of weighting on the mixing ratio (case study in Sect. 3.3). In the last two case studies, HydroMix is applied to observed tracer data from an Alpine catchment in the Swiss Alps to infer source mixing ratios and an additional parameter (isotopic lapse rate).

3.1 Mixing using Gaussian distributions

In this example, source concentrations S_1 and S_2 are drawn from two Gaussian distributions with different means (μ_1, μ_2) and standard deviations (σ_1, σ_2) and combined to form the mixture Y with a constant mixing ratio ρ :

$$\rho S_1 + (1 - \rho)S_2 = Y. \quad (12)$$

Assuming the two distributions are independent, the resultant mixture is normally distributed with mean (μ_y) and variance (σ_y^2) defined as

$$\mu_y = \rho\mu_1 + (1 - \rho)\mu_2; \sigma_y^2 = \rho^2\sigma_1^2 + (1 - \rho)^2\sigma_2^2. \quad (13)$$

A given number of samples are drawn from the distributions of S_1 and S_2 and of the mixture Y . The posterior distribution of the mixing ratio, $p(\rho|\tilde{S}_1, \tilde{S}_2, \tilde{Y})$, is then inferred using HydroMix for (i) a case in which the two source distributions are clearly identifiable and (ii) a case in which the distributions have a large overlap. Different values of mixing ra-

tios are tested, with ratios varying from 0.05 to 0.95 in steps of 0.05.

The sensitivity of HydroMix to the number of samples drawn from S_1 , S_2 , and Y , along with the time to convergence, is assessed based on the sum of the absolute error between the estimated mixing ratio $\hat{\rho}$ and its true value ρ .

3.2 Mixing with a time series generated using a hydrologic model

In this case study, we build a conceptual hydrologic model wherein groundwater is assumed to be recharged directly from rainfall and snowmelt. Stable isotopes of deuterium ($\delta^2\text{H}$) are used to see how the isotopic ratio in groundwater evolves under different assumptions of rain and snow recharge efficiencies.

Synthetic time series are generated for precipitation, the isotopic ratio in precipitation, and air temperature at a daily time step. For generating the precipitation time series, the time between two successive precipitation events is assumed to be a Poisson process with the precipitation intensity following an exponential distribution (Botter et al., 2007; Rodriguez-Iturbe et al., 1999). Time series of air temperature and of isotopic ratios in precipitation are obtained by generating an uncorrelated Gaussian process with the mean following a sine function (to emulate a seasonal signal) and with constant variance (Allen et al., 2018; Parton and Logan, 1981). The separation of precipitation into rainfall (P_r) and snowfall (P_s) is done based on a temperature threshold approach (Harpold et al., 2017a), whereby the fraction of rainfall $f_r(t)$ at time step t is computed as a function of air temperature $T(t)$:

$$f_r(t) = \begin{cases} 0 & \text{if } T(t) < T_L \\ \frac{T(t)-T_L}{T_H-T_L} & \text{if } T_L \leq T(t) \leq T_H \\ 1 & \text{if } T(t) > T_H, \end{cases} \quad (14)$$

where T_L and T_H are the lower and upper threshold bounds. A double air temperature threshold approach has been shown to be more accurate than a single temperature threshold (Harder and Pomeroy, 2014; Harpold et al., 2017a, b). In this case study, T_L and T_H are set to -1 and $+1$ °C. The evolution of the snow water equivalent (SWE) in the snowpack (h_s) is computed as

$$\frac{dh_s(t)}{dt} = P_s(t) - M_s(t), \quad (15)$$

where M_s is the magnitude of snowmelt computed using a degree-day approach as proposed by Schaeffli et al. (2014):

$$M_s = \begin{cases} a_s(T(t) - T_m), & \text{if } T(t) > T_m \\ 0 & \text{otherwise} \end{cases}, \quad (16)$$

where a_s is the degree-day factor (set here to $2.5 \text{ mm } ^\circ\text{C}^{-1} \text{ d}^{-1}$) and T_m is the threshold temperature at which snow starts to melt (set to 0 °C). Enhanced heat

exchange processes happening during rain-on-snow events are not explicitly considered as this lies beyond the scope of this paper. The snowpack is assumed to be fully mixed, and the isotopic ratio of snowpack is computed as

$$\frac{d(h_s(t)C_s(t))}{dt} = C_p(t)P_s(t) - C_s(t)M_s(t), \quad (17)$$

where C_s is the isotopic ratio of snowpack and C_p is the isotopic ratio of precipitation. The amount of groundwater recharge (R) is the sum of groundwater recharged from rainfall and snowmelt:

$$R(t) = R_r P_r(t) + R_s M_s(t), \quad (18)$$

where R_r and R_s are the rainfall and snowmelt recharge efficiencies. Recharge efficiency is defined as the fraction of rainfall or snowmelt that reaches the groundwater and is assumed to be a constant value. The groundwater storage is assumed to be fully mixed, and the isotopic ratio of groundwater is computed as

$$\begin{aligned} \frac{d(G(t)C_g(t))}{dt} &= R_r C_p(t) P_r(t) + R_s C_s(t) M_s(t) \\ &\quad - C_g(t) Q(t), \end{aligned} \quad (19)$$

where C_g is the isotopic ratio in groundwater, G is the volume of groundwater, and Q is the amount of groundwater outflow to the stream defined as

$$Q(t) = k(G(t) - G_C), \quad (20)$$

where k is the recession coefficient and G_C is constant groundwater storage that does not interact with the stream (added here to avoid zero storage and thus very small outflow). This formulation follows the linear groundwater reservoir assumption used in numerous hydrological modeling frameworks (Beven, 2011). The volume of groundwater storage is computed as

$$\frac{dG(t)}{dt} = R(t) - Q(t). \quad (21)$$

The model is run for a period of 100 years, allowing the system to reach a long-term steady state. The parameters used to generate daily precipitation, air temperature, and precipitation isotopic ratios are shown in Table 4. The number of yearly precipitation events is set to 30. The snow accumulation and the degree-day snowmelt models are then used to compute the number of snowfall days and snowmelt events. The static volume of groundwater that does not interact directly with the stream, G_C , is set to 1000 mm.

Only the last 2 years of the model runs are used to obtain the time series of isotopic ratios in rainfall, snowmelt, and groundwater. These years are then used to estimate the mixing ratio of snowmelt in groundwater, which is the fraction of groundwater recharged from snowmelt. Rainfall and

snowmelt samples are the two sources and groundwater samples represent the mixture. For the HydroMix application, all the modeled rainfall and snowmelt samples generated using the hydrologic model are used, whereas for groundwater, only one isotopic ratio per month is used (randomly sampled). The mixing ratios inferred using HydroMix are compared to the actual recharge ratio obtained from the hydrologic model as

$$R_s^a = \frac{\sum_t R_s M_s(t)}{\sum_t R(t)}, \quad (22)$$

where R_s^a represents the proportion of groundwater recharge derived from snowmelt summed over all the time steps. The numerical implementation of the evolution of the isotopic ratio in snowpack and groundwater is given in the Appendix.

3.3 Weighting mixing ratios in the hydrologic model

In Sect. 3.2, rainfall and snowmelt samples are not weighted by the magnitude of their fluxes while computing the mixing ratios with HydroMix. As all rainfall and snowmelt samples are used, the weights are implicitly determined by the number of rainfall and snowmelt events instead of their magnitudes. This is a general problem in all mixing approaches and has not been adequately acknowledged in the literature. Ignoring the weights may lead to biased mixing estimates if the proportional contribution of one of the components (e.g., rainfall or snowmelt) is low but the number of samples obtained to represent that component is proportionally much higher (Varin et al., 2011). For example, in a given catchment, the amount of total snowfall may be a small proportion of the annual precipitation, but the number of days when snowmelt occurs may be comparable to the total number of rainfall days in a year. If this is not specified a priori, HydroMix may overestimate the proportion of groundwater being recharged from snowmelt. To account for this, we introduce a weighting factor in the likelihood function originally formulated in Eq. (8) to make a new composite likelihood (Varin et al., 2011):

$$L_j(\tilde{Y}_{\text{obs}} | P_1, P_2, \theta) = \prod_{k=1}^q \prod_{j=1}^m \prod_{i=1}^n \left[\left(2\pi\sigma^2 \right)^{-0.5} \exp \left(-\frac{1}{2} \frac{(\tilde{Y}_{\text{obs}}^k - \hat{Y}_{ij})^2}{\sigma^2} \right) \right]^{w_i w_j}, \quad (23)$$

where i and j correspond to snowmelt and rainfall samples, and the weights w_i and w_j reflect the proportion of snowmelt and rainfall contributing to groundwater recharge (Vasdekis et al., 2014); w_i is expressed as

$$w_i = \frac{R_i S_i}{\sum_{i=1}^n R_i S_i}, \quad (24)$$

where R_i is the snowmelt magnitude and S_i is the isotopic ratio of the i th snowmelt event. Rain weights (w_j) are also

expressed similarly to Eq. (24). The obtained mixing ratio estimates are then compared with the unweighted estimates (in Sect. 3.2) to see if weighting by magnitude makes a significant difference.

3.4 Real case study: snow ratio in groundwater in Vallon de Nant

The objective of this case study is to infer the proportional contributions of snow versus rainfall to the groundwater of an Alpine headwater catchment, Vallon de Nant (Switzerland), using stable water isotopes.

3.4.1 Catchment description

Vallon de Nant is a 13.4 km² catchment located in the Vaud Alps in the southwest of Switzerland (Fig. 1), with elevation ranging from 1253 m to 3051 m a.s.l. Steep slopes form a major part of the catchment with a mean catchment slope of around 36° (Thornton et al., 2018). At lower elevations, a dense forest dominated by *Picea abies* covers 14 % of the catchment area. At around 1500 m a.s.l., there is an active pasture area with scattered trees and an open forest dominated by *Larix decidua*. Additional species scattered throughout the catchment include *Pinus sp.*, *Alnus sp.*, and *Acer pseudoplatanus*. Alpine meadows cover most of the higher-elevation land surfaces. Despite the relatively low elevation, there is a small glacier on its southwestern tip, which covers around 4.4 % of the catchment area, below which an extended moraine occupies 10.1 % of the catchment area. A large part (28 % of catchment area) of the hillslopes are composed of steep rock walls. At lower to mid-elevations, talus slopes account for about 6 % of the catchment area.

Vallon de Nant has a typical Alpine climate, with around 1900 mm of annual precipitation and a mean air temperature of 1.8 °C (Michelon, 2017). For this paper, long-term climate statistics are computed using the MeteoSwiss gridded precipitation and air temperature dataset for 1961–2015 (Isotta et al., 2013; MeteoSwiss, 2016, 2017). Applying a simple temperature threshold (0 and 1 °C) to observed precipitation indicates that, on average, 40 %–45 % of the total precipitation falls as snow in the catchment. There is a small degree of seasonality in precipitation, with higher precipitation between June and August and lower precipitation in the months of September and October.

3.4.2 Data collection

Vallon de Nant has been extensively monitored since February 2016. Water samples are collected from streamflow, rain, snowpack, and groundwater at different elevations, which are then analyzed for the isotopic ratios in deuterium ($\delta^2\text{H}$) and oxygen-18 ($\delta^{18}\text{O}$). Vallon de Nant is remotely located with very limited winter access, frequently experiencing winter avalanches. Due to these logistical constraints, snowmelt lysimeters or passive capillary samplers could not be set up

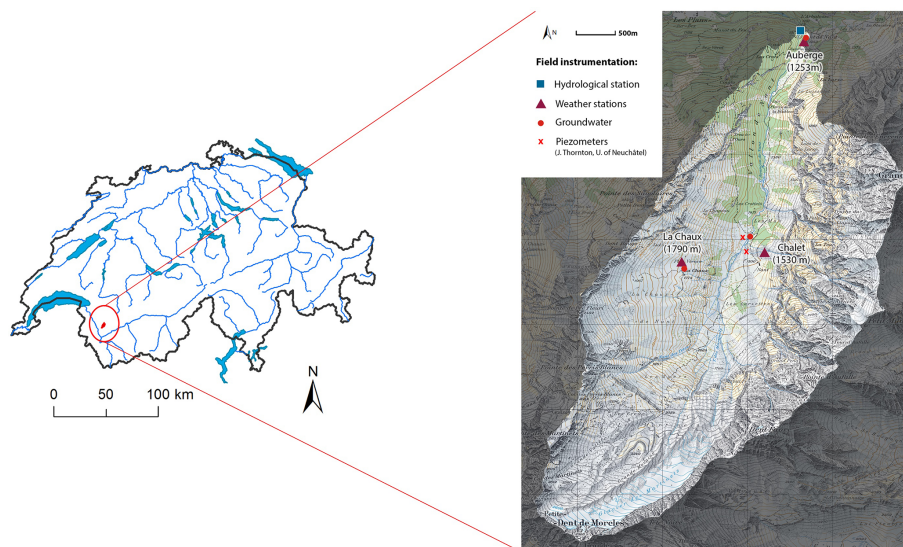


Figure 1. Map showing Vallon de Nant along with the locations of meteorologic and hydrologic observations as well as frequent sampling sites. Composite samples of precipitation were collected at the weather stations. Groundwater samples were collected at the groundwater monitoring points and the installed piezometers. The groundwater piezometers were installed by James Thornton from the University of Neuchâtel (Thornton et al., 2018).

Table 1. Summary of the isotopic data ($\delta^2\text{H}$ and $\delta^{18}\text{O}$) collected in Vallon de Nant between February 2016 and July 2017.

Sample name	Number of samples	Lowest elevation	Highest elevation
Rainfall	32	1253	1773
Top snowpack layer	80	1241	2455
Groundwater	22	1253	1779

to sample snowmelt water; accordingly, grab snowpack samples are used here as a proxy for snowmelt. A summary of the isotopic data is shown in Table 1.

3.4.3 Model implementation

HydroMix is used to estimate the proportion of snow recharging groundwater (subsequently referred to as the “snow recharge coefficient”). In order to obtain a PDF of the snow recharge coefficient, isotopic ratios in all the water samples from rain, snowpack, and groundwater are used. A uniform prior distribution is assigned to the snow recharge coefficient, which varies between 0 and 1, representing the entire range of possible values.

3.5 Introduction of an additional model parameter

In any mixing analysis, it may be useful or desirable for users to specify an additional model parameter that is able to modify the tracer concentrations based on their process understanding of the system. In the case of Alpine catchments with

large elevation gradients, stable isotopes in precipitation often exhibit a systematic trend with elevation, becoming more depleted in heavier isotopes with increasing elevation. This is also known as the “isotopic lapse rate” (Dansgaard, 1964; Friedman et al., 1964). In typical field campaigns, because of logistical challenges, precipitation samples are collected only at a few points in a catchment, with often fewer precipitation samples at high elevations. This leads to oversampling at lower elevations and undersampling at higher elevations, which can bias mixing estimates. This has been found to be especially relevant for hydrograph separation in forested catchments (Cayuela et al., 2019). To allow a process compensation for this, an additional lapse rate factor is introduced with which each observed point-scale sample (observed at a given elevation) is corrected to a reference elevation as follows:

$$\bar{r} = \frac{\sum_{j=1}^k [\alpha (e_j - e) + r] a_j}{\sum_{j=1}^k a_j}, \quad (25)$$

where r is the isotopic ratio in precipitation collected at elevation e , \bar{r} is the catchment-averaged isotopic ratio in precipitation, α is the isotopic lapse rate factor, e_j is the elevation of the j th elevation band, and a_j is the catchment area under the j th elevation band; the catchment is divided into k elevation bands. These bands are obtained by constructing a hypsometric curve of the catchment (Strahler, 1952).

The lapse rate factor is allowed to modify both rainfall and snowpack isotopic ratios to obtain a catchment-averaged isotopic ratio, which is then used in the mixing model. Using this formulation of an isotopic lapse rate makes the following implicit assumptions: (1) precipitation storms on aggre-

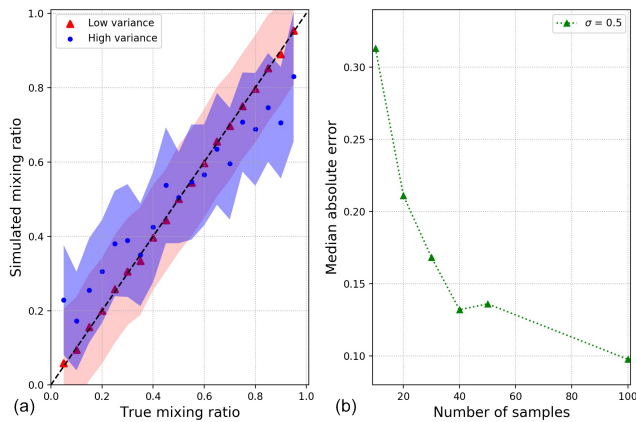


Figure 2. (a) Scatterplot showing the mixing ratio (ρ) values inferred using HydroMix for the low and high variance synthetic case in Table 3. The uncertainty band represents the inferred mixing ratio plus or minus the error standard deviation obtained from Eq. (13). The number of source and target samples is 100. (b) Performance of HydroMix in terms of the absolute error between the posterior mixing ratio mean and the true mean for the low variance dataset over all tested ratios plotted as a function of the number of samples drawn for the two sources.

gate move from the lower part of the catchment to the upper part of the catchment, thus creating a lapse rate effect, and (2) precipitation falls uniformly over the catchment. It is important to note that the isotopic lapse rate is different from the precipitation lapse rate; i.e., the rate of change of precipitation with elevation is different from the rate of change of the precipitation isotopic ratio with elevation.

It is important to note that the precipitation isotopic ratio is not only a function of elevation, but also depends on other factors such as the source of moisture origin, cloud condensation temperature, and secondary evaporation. Similarly, strong spatial variability exists in the isotopic ratio of snowmelt water, depending on catchment aspect, snow metamorphism, and wind distribution. This case study is a mere demonstration that HydroMix allows for the inference of additional parameters that can account for various physical processes that may modify isotopic ratios.

The prior distribution of the isotopic lapse rate is specified based on isotopic data collected across Switzerland under the Global Network of Isotopes in Precipitation (GNIP) program (IAEA/WMO, 2018). Using the monthly isotopic values collected between 1966 and 2014, average lapse rate values are obtained for both $\delta^2\text{H}$ and $\delta^{18}\text{O}$. These were $(-)$ 1.94‰ per 100 m for $\delta^2\text{H}$ and $(-)$ 0.27‰ per 100 m for $\delta^{18}\text{O}$ (Beria et al., 2018).

A uniform prior distribution is assigned to the isotopic lapse rate parameter, with the lower bound specified as 3 times the Swiss lapse rate for both $\delta^2\text{H}$ and $\delta^{18}\text{O}$. The observed isotopic lapse rate data from Switzerland suggest that average lapse rates are weakly negative; however, positive lapse rates can a priori not be excluded for the case study

catchment. Accordingly, we do not specify an upper lapse rate bound of zero but set it as 3 times the Swiss lapse rate (Table 2). In the case of Vallon de Nant, the elevation ranges from 1253 m to 3051 m a.s.l. For computing the Swiss lapse rate, the elevation range over which the monthly precipitation samples were collected was 300 m to 2000 m a.s.l. This difference in elevation ranges between Vallon de Nant and the GNIP network should be kept in mind during the interpretation of results.

4 Results

The results for the different case studies are discussed in the sections below.

4.1 Mixing with normal distributions

The mean and standard deviations used to generate the low and high variance source distributions for the synthetic case studies are summarized in Table 3. We randomly generated 100 samples from each of the two source distributions and from the target distribution and varied the mixing ratios between 0.05 and 0.95 in 0.05 increments. It should be noted that HydroMix permits using a different number of samples for the sources and for the mixture.

For the low variance case, the mixing ratio inferred with HydroMix with 1000 MCMC simulations closely reproduces the theoretical mean of the mixing ratios used to generate the synthetic data (Fig. 2a). However, for the high variance case, the inferred mixing ratios do not match the true underlying mixing ratios, especially for low and high mixing ratios. This is partly due to the poor identifiability of the sources (given that their distributions are highly overlapping) and partly due to the relatively small sample size of 100. The inferred mean should reproduce the theoretical mean with increasing sample size and we clearly see this for the low variance case in Fig. 2b, where the model performance markedly improves with an increasing number of samples. The performance is measured here in terms of the absolute error between the posterior mixing ratio mean and the true mean summed and averaged over all tested ratios from 0.05 to 0.95. We did not perform inferences for sample sizes larger than 100 as the computational requirement increases exponentially with increasing sample sizes.

The model converges fairly quickly for the low variance case after ~ 100 runs as shown in Fig. 3a. The obtained model residuals have zero mean and are approximately normally distributed as revealed by quantile–quantile plots (not shown), in line with the assumption of an unbiased normally distributed error model, as stated in Eq. (7).

Table 2. Prior distribution of the different model parameters as specified to HydroMix.

Variable	Prior distribution	Lower bound	Upper bound
Snow recharge coefficient	Uniform	0	1
Isotopic lapse rate in $\delta^2\text{H}$	Uniform	(−)5.82‰ per 100 m	(+)5.82‰ per 100 m
Isotopic lapse rate in $\delta^{18}\text{O}$	Uniform	(−)0.81‰ per 100 m	(+)0.81‰ per 100 m

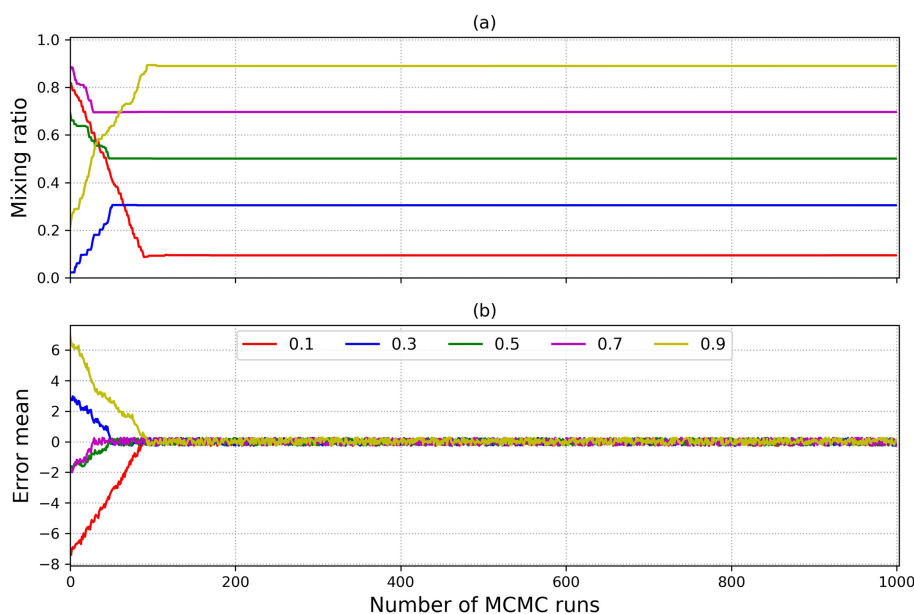


Figure 3. Diagnostic plots showing the convergence characteristics of MCMC chains for five different mixing ratios for the low variance dataset (shown in Table 3). Panels (a) and (b) show variations in the inferred mixing ratio and the error mean with increasing MCMC runs.

Table 3. Mean and variance of the two sources, S_1 and S_2 , drawn from normal distributions.

Dataset	μ_{S_1} (σ_{S_1})	μ_{S_2} (σ_{S_2})
Low variance	10 (0.5)	20 (0.5)
High variance	10 (5.0)	20 (5.0)

Table 4. Parameters used to generate time series of precipitation, air temperature, and isotopic ratios in precipitation. μ represents the mean, A is the amplitude, and φ is the time lag of the underlying sine function. For the precipitation process, μ is the mean intensity on days with precipitation. The resulting mean winter length (air temperature below 0 °C) is 119.5 d.

Variable	Parameter values
Precipitation	No. of events per year: 30, $\mu = 33.45 \text{ mm d}^{-1}$
Air temperature	$\mu = 4 \text{ }^\circ\text{C}$, $A = 8 \text{ }^\circ\text{C}$, $\varphi = -\pi/2$
Precipitation isotopic ratio	$\mu = (-80) \text{ }^\circ\text{‰}$, $A = 40 \text{ }^\circ\text{‰}$, $\varphi = -\pi/2$

4.2 Contribution of rain and snow to groundwater recharge using a hydrologic model

Figure 4 shows the variation in the isotopic ratio of groundwater over the entire 100-year period, showing that the system achieves a steady-state condition after ~ 15 years of simulation. The mixing ratio is estimated with HydroMix using (1) samples of the isotopic ratio in snowfall and (2) samples of the isotopic ratio in snowmelt. The two sample distributions differ, as shown in Fig. 5, where the variability of the isotopic ratio is lower in snowmelt when compared to snowfall. In the model at hand, this reduction is obtained because of mixing occurring within the snowpack, leading to homogenization and thus reducing the variability in the isotopic ratio of snowmelt. In field data, such a reduction in variability

is also generally observed (Beria et al., 2018) as a result of the homogenization as modeled here and from more complex snow physical processes, which lie beyond the scope of this study.

The mixing ratios inferred with HydroMix are very similar regardless of whether snowfall or snowmelt is used across the entire range of recharge efficiencies (Fig. 6). This provides confidence in the use of snowfall samples as a proxy for snowmelt when estimating mixing ratios. However, it is clear from Fig. 6 that an important bias emerges between the

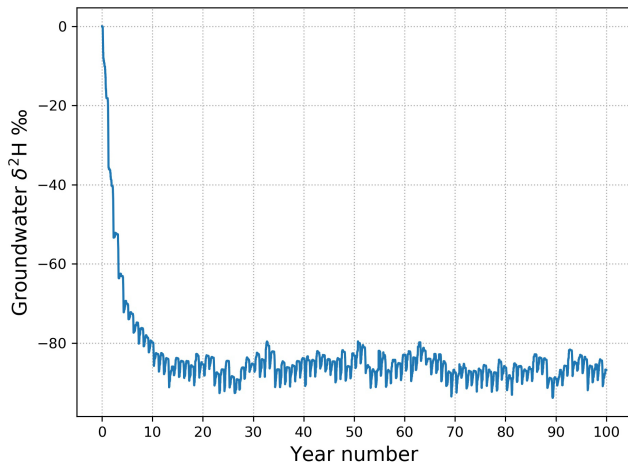


Figure 4. Evolution of the modeled isotopic ratio in groundwater over a 100-year period with $R_r = 0.3$ and $R_s = 0.6$.

estimated mixing ratio from HydroMix and the actual mixing ratio known from the hydrologic model, especially for low mixing ratios.

This bias can be expected to emerge when the source contributions are not weighted according to their fluxes, which to our knowledge has not been explicitly addressed in the hydrological literature. As already discussed in Sect. 3.3, the absence of sample weighting typically induces a bias when there is a large divergence between the number of samples taken over a certain period (e.g., 1 year) to characterize a source and the magnitude of source flux over that period (e.g., 40 snow and 10 rain samples taken to characterize the two sources, for which snow only accounts for a very small portion, e.g., 10 %, of the annual precipitation).

4.3 Effect of weights on estimates of mixing ratios using a hydrologic model

After taking into account the magnitude of rainfall and snowmelt events in the composite likelihood function of Eq. (23), it is clear that many of the unweighted biases can be removed (Fig. 7). The most significant improvement is seen at very low mixing ratios for which the divergence between the conceptual model and the mixing model estimates error is reduced by almost 50 %. In this study, we have used a relatively simple normalization-based weighting function (Eq. 25). Testing other weighting functions that have been proposed in the past (Vasdekis et al., 2014) is left for future research.

4.4 Inferring fraction of snow recharging groundwater in a small Alpine catchment along with an additional model parameter

Using the dataset from an Alpine catchment (Vallon de Nant, Switzerland), HydroMix estimates that 60 %–62 % of the

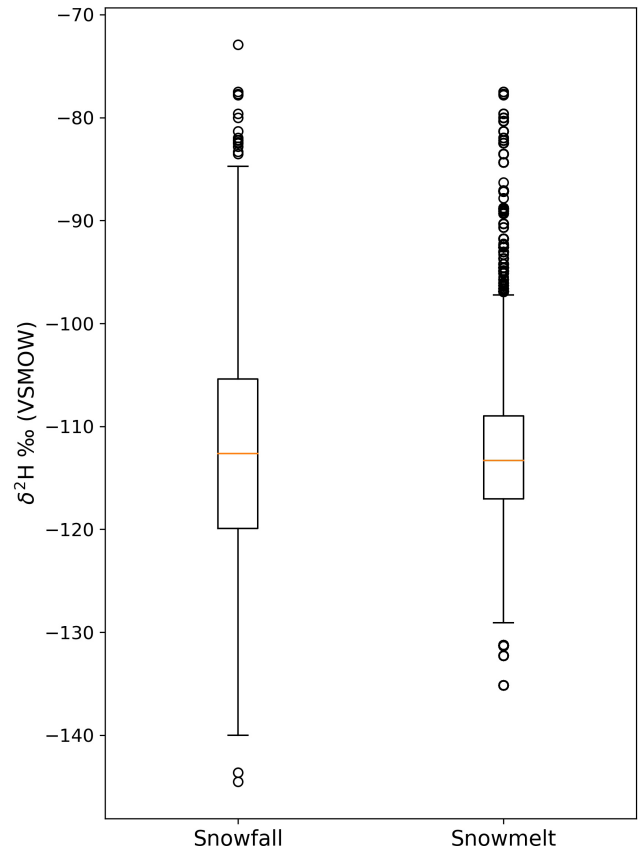


Figure 5. Box plot showing the variability in the isotopic ratio of snowfall and snowmelt as simulated by the hydrologic model. The box plot extends from the 25th to 75th percentile value, with the median value depicted by the orange line. The whiskers extend up to 1.5 times the interquartile range. The black circles are the outliers.

groundwater is recharged from snowmelt (using unweighted approach), with the full posterior distributions shown in Fig. 8a. This estimate is consistent for both of the isotopic tracers ($\delta^2\text{H}$ and $\delta^{18}\text{O}$), which are often used interchangeably in the hydrologic literature (Gat, 1996). Comparing this recharge estimate to the proportion of total precipitation that falls as snow (around 40 %–45 %; see Sect. 3.4.1) suggests that snowmelt is more effective at reaching the aquifer than an equivalent amount of rainfall falling at a different period of the year. Similar results have been obtained in a number of previous studies across the temperate and mountainous regions of the world (see Table 1 in the work of Beria et al., 2018, for a summary).

As can be seen from Fig. 8a, the estimated distribution of the snow ratio in groundwater is very narrow. This can be explained by the fact that we assume that the collected precipitation samples represent the variability actually occurring in the catchment. To overcome this limitation, we infer an additional parameter called the isotopic lapse rate that accounts for the spatial heterogeneity in terms of catchment elevation. As shown in Fig. 9, the posterior distributions

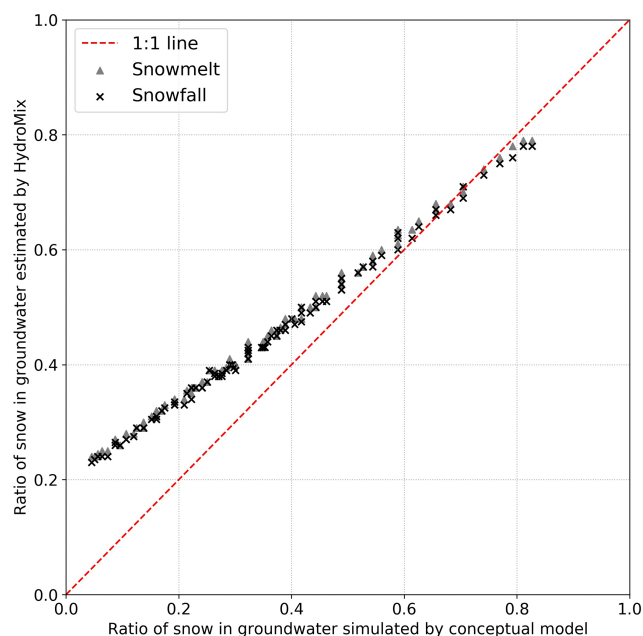


Figure 6. Ratios of snow in groundwater estimated with HydroMix plotted against ratios obtained from the hydrologic model for the last 2 years of simulation. Also shown are the separate results obtained by using samples of either snowmelt or snowfall. The full range of ratios is obtained by varying rainfall and snowmelt recharge efficiencies from 0.05 to 0.95. The numbers of rainfall, snowfall, and snowmelt days are 39, 24, and 107 in the last 2 years of simulation.

of the isotopic lapse rate (for both $\delta^2\text{H}$ and $\delta^{18}\text{O}$) largely overlap the spatially averaged isotopic lapse rate as estimated from precipitation isotopes across Switzerland. The overlap with the average Swiss isotope lapse rate suggests that our inferred lapse rates are reasonable, with the spread in the estimates likely reflecting the temporal variation in the catchment-specific isotope lapse rate that can develop from a wide range of moderating factors (e.g., air masses contributing precipitation without traversing the full elevation range of the catchment due to varying trajectories). The Swiss lapse rate is constructed as a long-term spatial average, whereas the inferred isotopic lapse rate in Vallon de Nant is constructed from the temporal variations in the isotopic ratios. These results demonstrate that it is relatively straightforward to jointly infer multiple parameters within the HydroMix modeling framework.

However, an important consequence of additional parameter inference without providing additional data or constraints is an increase in the degree of freedom, which can then increase the uncertainty in source contributions. This effect is seen in Fig. 8b, especially in contrast with the previous result in Fig. 8a, where the median mixing ratios of the posterior distributions remain similar (~ 0.6), but the spread increases drastically from 0.005 to 0.2.

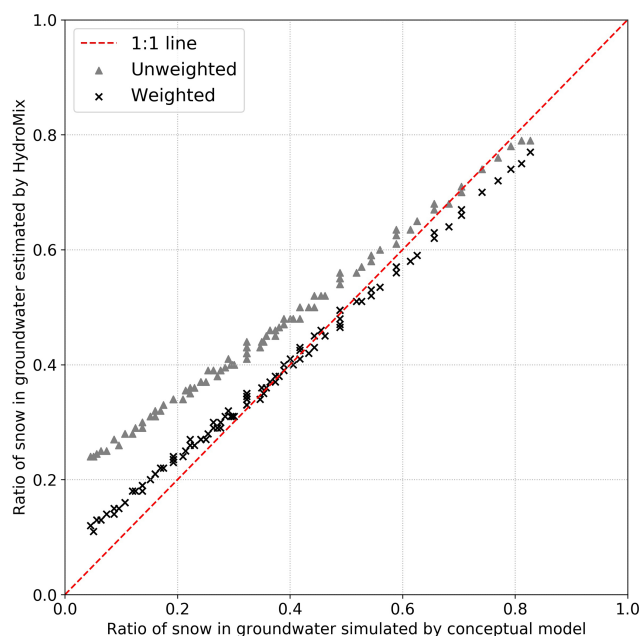


Figure 7. Ratios of snow in groundwater estimated using HydroMix plotted against ratios obtained from the hydrologic model for both weighted and unweighted mixing scenarios. The full range of ratios is obtained by varying rainfall and snowmelt recharge efficiencies from 0.05 to 0.95. The numbers of rainfall, snowfall, and snowmelt days are 39, 24, and 107 in the last 2 years of simulation.

5 Limitations and opportunities

As with all linear mixing models, the quality of the underlying data determines the accuracy and utility of the results. If the tracer compositions of the different sources are not sufficiently distinct, the uncertainty in the estimated mixing ratios will become very large. This means that if either the underlying data quality is poor or the source contribution dynamics are not well conceptualized, then the uncertainty in the mixing ratios will be too high to be useful.

In cases in which a large number of source samples are available, the computational requirements of HydroMix outweigh the benefit from using it. These are likely cases in which the statistical distribution of the source tracer composition is well understood, and therefore fitting a probability density curve to the source and target samples and then inferring the distribution of the mixing ratio using a probabilistic programming approach is more appropriate (Carpenter et al., 2017; Parnell et al., 2010; Stock et al., 2018). Also, HydroMix might not be an appropriate method in instances in which fitting statistical distributions to source and target compositions reflects a priori knowledge of the system.

A key difference between HydroMix and other Bayesian mixing approaches is that HydroMix parameterizes the error function, whereas other Bayesian approaches parameterize the statistical distribution of source and mixture compositions. Parameterizing source compositions requires large

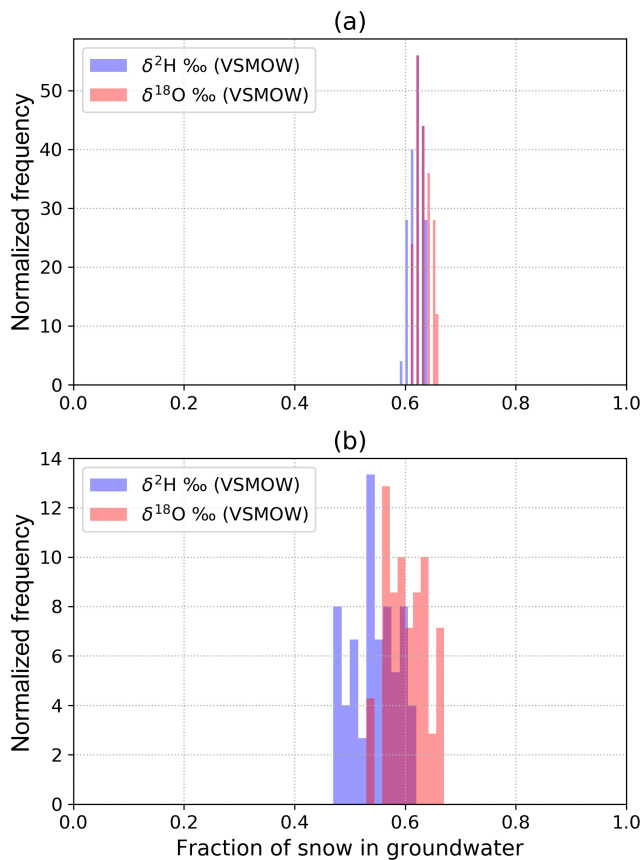


Figure 8. Histogram showing the fraction of snow recharging groundwater in Vallon de Nant using the isotopic ratios in $\delta^2\text{H}$ and $\delta^{18}\text{O}$ (a) without correcting for lapse rate and (b) after correcting for lapse rate.

sample sizes, which is seldom the case in tracer hydrology. Error parameterization offers a useful alternative and can also be verified against the posterior error distribution. In the case studies demonstrated in this paper, a normally distributed error model was found to be appropriate. However, error models other than Gaussian can be used by formulating the respective likelihood function.

HydroMix builds the model residuals by comparing all the observed source samples with all the observed samples of the target mixture, assuming that all available source and target samples are independent. Interestingly, the assumption of independence holds even if the source and target samples are taken at the same time, since the target samples result from water that has traveled for a certain amount of time in the catchment and hence is not related to the water entering the catchment. However, if a system has instantaneous mixing, then the source and target samples taken at the same moment in time will necessarily be strongly correlated. In such cases, the assumption of independent samples would not make sense and the method might give spurious results.

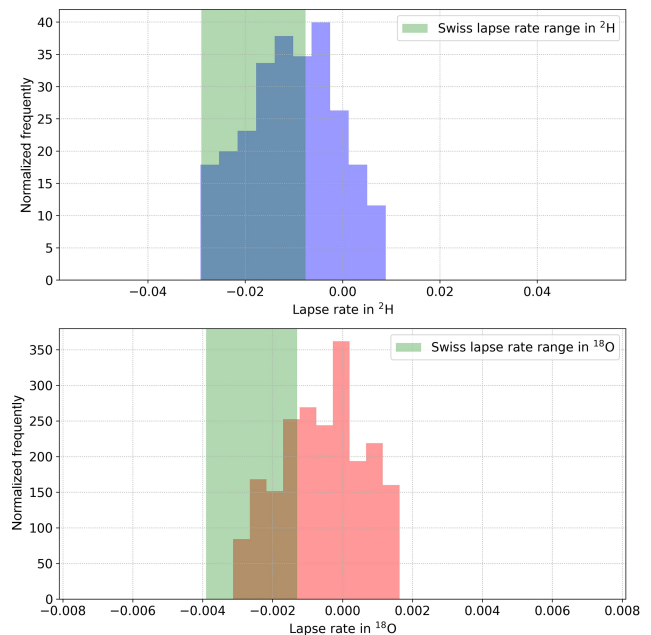


Figure 9. Histogram showing the posterior distribution of the isotope lapse rate parameter in $\delta^2\text{H}$ and $\delta^{18}\text{O}$. The green region shows the confidence bounds (significant at $\alpha = 0.01$) of the lapse rate computed over Switzerland by using inverse variance-weighted regression. The limits of the prior distribution of the isotopic lapse rates correspond to the limits of the x axis. The slope of the isotopic ratio when plotted against elevation for the Swiss-wide data is shown in Fig. 3 of Beria et al. (2018).

Finally, it is noteworthy that adding additional parameters to characterize the source tracer composition increases the degree of freedom of the model, which implies that adding such parameters leads to an increase in the uncertainty of the source contribution estimates unless new information, i.e., new observed data, is added to the model. This means that users who are interested in incorporating additional modification processes by adding parameters should ideally provide additional tracer data able to constrain this process, subject to tracer data being available.

For consistency and simplicity, the case studies and synthetic hydrological examples provided here focused on the contribution of rain and snow in recharging groundwater. However, it is important to emphasize that the opportunities to implement HydroMix extend to all cases in which mixing contributions are of interest and for which it is difficult to build extensive databases of source tracer compositions. Such examples include quantifying the amount of “pre-event” vs. “event” water in streamflow; pre-event water refers to groundwater and event water refers to rainfall or snowmelt. Another interesting use case might be to quantify the proportion of streamflow coming from the different source areas in a catchment to capture the spatial dynamics of streamflow. Other uses include quantifying the amount of fog contributing to throughfall, the proportion of glacial melt

vs. snowmelt flowing into a stream, the amount of vegetation water use from soil moisture at different depths vs. groundwater, the interaction between surface water and groundwater at the hyporheic zone (Leslie et al., 2017), and sediment fingerprinting to quantify the spatial origin of river sediments. In all of these cases, understanding source water contributions, both spatially and temporally, will improve the physical understanding of the system.

6 Conclusions

We develop a new Bayesian modeling framework for the application of tracers in mixing models. The primary application target of this framework is hydrology, but it is by no means limited to this field. HydroMix formulates the linear mixing problem in a Bayesian inference framework that infers the model parameters with a Metropolis–Hastings-based MCMC sampling algorithm based on differences between observed and modeled tracer concentrations in the target mixture using all possible combinations between all source and target concentration samples. This is especially useful in data-scarce environments where fitting probability distribution functions is not feasible. HydroMix also makes the inclusion of additional model parameters to account for source modification processes straightforward. Examples include known spatial or temporal tracer variations (e.g., isotopic lapse rates or evaporative enrichment).

An evaluation of HydroMix with data from different synthetic and field case studies leads to the following conclusions.

1. HydroMix gives reliable results for mixing applications with small sample sizes (< 20–30 samples). As expected, the variance in source tracer composition and the ensuing composition overlap determines the bias in the mixing ratio estimates. The bias in mixing ratio estimates increases with increasing variance in source tracer compositions. Mixing ratio estimates improve (in terms of lower error) with an increasing number of source samples.
2. As revealed by our synthetic case study with a conceptual hydrological model, at low source contributions (i.e., < 20 %), a strong divergence between the actual and estimated mixing ratios emerges. This arises if HydroMix assigns equal weights to all source samples by proportionally oversampling the less abundant source, which then leads to significant biases in mixing estimates. This problem is inherent to all mixing approaches and to our knowledge has not been adequately addressed in the literature.
3. The use of composite likelihoods to weight samples by their amounts can significantly reduce the bias in the mixing estimates. At low source proportions, the estimated mixing ratio improves by more than 50 % after

accounting for the amount of all the sources. We show this using a simple normalization-based weighting function. Future studies should explore the usage of different weighting functions that have been proposed in the past (Vasdekis et al., 2014).

4. A synthetic application of HydroMix to understand the amount of snowmelt-induced groundwater recharge revealed that using the snowfall isotopic ratio instead of the snowmelt isotopic ratio leads to similar mixing ratio estimates. This is particularly useful in high mountainous catchments, where sampling snowmelt is logistically difficult.
5. A real case application of HydroMix in a Swiss Alpine catchment (Vallon de Nant) showed a clear winter bias in groundwater recharge. About 60 %–62 % of the groundwater is recharged from snowmelt (unweighted mixing approach), while snowfall only accounts for 40 %–45 % of the total annual precipitation. This has also been previously suggested elsewhere in the European Alps (Cervi et al., 2015; Penna et al., 2014, 2017; Zappa et al., 2015).

To conclude, HydroMix provides a Bayesian approach to mixing model problems in hydrology that takes full advantage of small sample sizes. Future work will show the full potential of this approach in hydrology as well as other environmental modeling applications.

Appendix A

The equations below show the numerical implementation of the evolution of isotopic ratios in snowpack, isotopic ratios in groundwater, and groundwater storage at a daily time step.

$$C_s(t) = \frac{C_s(t-1)h_s(t-1) + C_p(t)P_s(t) - C_s(t-1)M_s(t)}{h_s(t-1) + P_s(t) - M_s(t)} \quad (\text{A1})$$

$$C_g(t) = \frac{C_g(t-1)G(t-1) + C_p(t)R_r P_r(t) + C_s(t)R_s M_s(t) - C_g(t-1)Q(t)}{G(t-1) + R_r P_r(t) + R_s M_s(t) - Q(t)} \quad (\text{A2})$$

$$G(t) = G(t-1) + R_r P_r(t) + R_s M_s(t) - k(G(t) - G_C) \quad (\text{A3})$$

Code and data availability. The model code is implemented in Python 2.7 and can be downloaded along with the dataset from Zenodo at <https://doi.org/10.5281/zenodo.3475429> (Beria et al., 2019). The most recent version of the model code is available on GitHub at https://github.com/harshberia93/HydroMix/tree/20191007_GMD (last access: 8 October 2019; harshberia93, 2019).

Author contributions. The paper was written by HB with contributions from all coauthors. HB and BS formulated the conceptual underpinnings of HydroMix. JRL helped in framing the statistical and hydrological tests to evaluate HydroMix. AM and NCC helped in compiling data used for model evaluation and provided critical feedback during model validation.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would also like to thank Lionel Benoit for his inputs on the formulation of the Bayesian mixing model. We thank the four anonymous reviewers and the editors for their constructive feedback that considerably improved the paper.

Financial support. This research has been supported by the Swiss National Science Foundation (grant no. PP00P2_157611).

Review statement. This paper was edited by Bethanna Jackson and reviewed by four anonymous referees.

References

- Allen, S. T., Kirchner, J. W., and Goldsmith, G. R.: Predicting spatial patterns in precipitation isotope ($\delta^2\text{H}$ and $\delta^{18}\text{O}$) seasonality using sinusoidal isoscapes, *Geophys. Res. Lett.*, 45, 4859–4868, <https://doi.org/10.1029/2018GL077458>, 2018.
- Barbeta, A. and Peñuelas, J.: Relative contribution of groundwater to plant transpiration estimated with stable isotopes, *Sci. Rep.*, 7, 10580, <https://doi.org/10.1038/s41598-017-09643-x>, 2017.
- Benettin, P., Bailey, S. W., Rinaldo, A., Likens, G. E., McGuire, K. J., and Botter, G.: Young runoff fractions control streamwater age and solute concentration dynamics, *Hydrol. Process.*, 31, 2982–2986, <https://doi.org/10.1002/hyp.11243>, 2017.
- Beria, H., Larsen, J. R., Ceperley, N. C., Michelon, A., Vennemann, T., and Schaefli, B.: Understanding snow hydrological processes through the lens of stable water isotopes, *Wiley Interdiscip. Rev. Water*, 5, e1311, <https://doi.org/10.1002/wat2.1311>, 2018.
- Beria, H., Larsen, J. R., Michelon, A., Ceperley, N. C., and Schaefli, B.: Data for the manuscript “HydroMix v1.0: a new Bayesian mixing framework for attributing uncertain hydrological sources” (Version 1.0), Zenodo, <https://doi.org/10.5281/zenodo.3475429>, 2019.
- Beven, K. J.: *Rainfall-runoff modelling: the primer*, Second Edi., John Wiley & Sons, 2011.
- Blake, W. H., Boeckx, P., Stock, B. C., Smith, H. G., Bodé, S., Upadhayay, H. R., Gaspar, L., Goddard, R., Lennard, A. T., Lizaga, I., Lobb, D. A., Owens, P. N., Peticrew, E. L., Kuzyk, Z. Z. A., Gari, B. D., Munishi, L., Mtei, K., Nebiyu, A., Mabit, L., Navas, A., and Semmens, B. X.: A deconvolutional Bayesian mixing model approach for river basin sediment source apportionment, *Sci. Rep.*, 8, 13073, <https://doi.org/10.1038/s41598-018-30905-9>, 2018.
- Botter, G., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response, *Water Resour. Res.*, 43, 2, <https://doi.org/10.1029/2006WR005043>, 2007.
- Burns, D. A., McDonnell, J. J., Hooper, R. P., Peters, N. E., Freer, J. E., Kendall, C., and Beven, K.: Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research Watershed (Georgia, USA), *Hydrol. Process.*, 15, 1903–1924, <https://doi.org/10.1002/hyp.246>, 2001.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A.: Stan?: A Probabilistic Programming Language, *J. Stat. Softw.*, 76, 1, <https://doi.org/10.18637/jss.v076.i01>, 2017.
- Cayuela, C., Latron, J., Geris, J., and Llorens, P.: Spatio-temporal variability of the isotopic input signal in a partly forested catchment: Implications for hydrograph separation, *Hydrol. Process.*, 33, 36–46, <https://doi.org/10.1002/hyp.13309>, 2019.
- Cervi, F., Corsini, A., Doveri, M., Mussi, M., Ronchetti, F., and Tazioli, A.: Characterizing the Recharge of Fractured Aquifers: A Case Study in a Flysch Rock Mass of the Northern Apennines (Italy), in: *Engineering Geology for Society and Territory*, vol. 3, edited by: Lollino, G., Arattano, M., Rinaldi, M., Giustolisi, O., Marechal, J.-C., and Grant, G. E., 563–567, Springer International Publishing, Cham., 2015.
- Dansgaard, W.: Stable isotopes in precipitation, *Tellus*, 16, 436–468, <https://doi.org/10.1111/j.2153-3490.1964.tb00181.x>, 1964.
- Earman, S., Campbell, A. R., Phillips, F. M., and Newman, B. D.: Isotopic exchange between snow and atmospheric water vapor: Estimation of the snowmelt component of groundwater recharge in the southwestern United States, *J. Geophys. Res.-Atmos.*, 111, D09302, <https://doi.org/10.1029/2005JD006470>, 2006.
- Ehleringer, J. R. and Dawson, T. E.: Water uptake by plants: perspectives from stable isotope composition, *Plant. Cell Environ.*, 15, 1073–1082, <https://doi.org/10.1111/j.1365-3040.1992.tb01657.x>, 1992.
- Evaristo, J., McDonnell, J. J., Scholl, M. A., Bruijnzeel, L. A., and Chun, K. P.: Insights into plant water uptake from xylem-water isotope measurements in two tropical catchments with contrasting moisture conditions, *Hydrol. Process.*, 30, 3210–3227, <https://doi.org/10.1002/hyp.10841>, 2016.
- Evaristo, J., McDonnell, J. J., and Clemens, J.: Plant source water apportionment using stable isotopes: A comparison of simple linear, two-compartment mixing model approaches, *Hydrol. Process.*, 31, 3750–3758, <https://doi.org/10.1002/hyp.11233>, 2017.
- Friedman, I., Redfield, A. C., Schoen, B., and Harris, J.: The variation of the deuterium content of natural waters in the hydrologic cycle, *Rev. Geophys.*, 2, 177–224, <https://doi.org/10.1029/RG002i001p00177>, 1964.

- Gat, J. R.: Oxygen and hydrogen isotopes in the hydrologic cycle, *Annu. Rev. Earth Planet. Sci.*, 24, 225–262, <https://doi.org/10.1146/annurev.earth.24.1.225>, 1996.
- Gelman, A. and Rubin, D. B.: Inference from Iterative Simulation Using Multiple Sequences, *Stat. Sci.*, 7, 457–472, <https://doi.org/10.1214/ss/1177011136>, 1992.
- Genereux, D.: Quantifying uncertainty in tracer-based hydrograph separations, *Water Resour. Res.*, 34, 915–919, <https://doi.org/10.1029/98WR00010>, 1998.
- Genereux, D. P., Wood, S. J., and Pringle, C. M.: Chemical tracing of interbasin groundwater transfer in the lowland rainforest of Costa Rica, *J. Hydrol.*, 258, 163–178, [https://doi.org/10.1016/S0022-1694\(01\)00568-6](https://doi.org/10.1016/S0022-1694(01)00568-6), 2002.
- Glynn, P. W. and Iglehart, D. L.: Importance Sampling for Stochastic Simulations, *Manage. Sci.*, 35, 1367–1392, <https://doi.org/10.1287/mnsc.35.11.1367>, 1989.
- Harder, P. and Pomeroy, J. W.: Hydrological model uncertainty due to precipitation-phase partitioning methods, *Hydrol. Process.*, 28, 4311–4327, <https://doi.org/10.1002/hyp.10214>, 2014.
- Harman, C. J.: Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, *Water Resour. Res.*, 51, 1–30, <https://doi.org/10.1002/2014WR015707>, 2015.
- Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., Schumer, R., and Steele, C. M.: Rain or snow: hydrologic processes, observations, prediction, and research needs, *Hydrol. Earth Syst. Sci.*, 21, 1–22, <https://doi.org/10.5194/hess-21-1-2017>, 2017a.
- Harpold, A. A., Rajagopal, S., Crews, J. B., Winchell, T., and Schumer, R.: Relative Humidity Has Uneven Effects on Shifts From Snow to Rain Over the Western U.S., *Geophys. Res. Lett.*, 44, 9742–9750, <https://doi.org/10.1002/2017GL075046>, 2017b.
- harshberia93: HydroMix, GitHub, available at: https://github.com/harshberia93/HydroMix/tree/20191007_GMD, last access: 8 October 2019.
- Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57, 97–109, <https://doi.org/10.1093/biomet/57.1.97>, 1970.
- Hoeg, S., Uhlenbrook, S., and Leibundgut, C.: Hydrograph separation in a mountainous catchment – combining hydrochemical and isotopic tracers, *Hydrol. Process.*, 14, 1199–1216, [https://doi.org/10.1002/\(SICI\)1099-1085\(200005\)14:7<1199::AID-HYP35>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-1085(200005)14:7<1199::AID-HYP35>3.0.CO;2-K), 2000.
- IAEA/WMO: Global Network of Isotopes in Precipitation. The GNIP Database, available at: http://www-naweb.iaea.org/naweb/ih/IHS_resources_gnip.html (last access: 31 January 2019), 2018.
- Inouye, D., Yang, E., Allen, G., and Ravikumar, P.: A Review of Multivariate Distributions for Count Data Derived from the Poisson Distribution, *Wiley Interdiscip. Rev. Comput. Stat.*, 9, e1398, <https://doi.org/10.1002/wics.1398>, 2017.
- Isotta, F. A., Frei, C., Weigluni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, *Int. J. Climatol.*, 34, 1657–1675, <https://doi.org/10.1002/joc.3794>, 2013.
- Jasechko, S., Wassenaar, L. I., and Mayer, B.: Isotopic evidence for widespread cold-season-biased groundwater recharge and young streamflow across central Canada, *Hydrol. Process.*, 31, 2196–2209, <https://doi.org/10.1002/hyp.11175>, 2017.
- Jeelani, G., Bhat, N. A., and Shivanna, K.: Use of $\delta^{18}\text{O}$ tracer to identify stream and spring origins of a mountainous catchment: A case study from Liddar watershed, Western Himalaya, India, *J. Hydrol.*, 393, 257–264, <https://doi.org/10.1016/j.jhydrol.2010.08.021>, 2010.
- Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, 42, 3, <https://doi.org/10.1029/2005WR004368>, 2006a.
- Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resour. Res.*, 42, 3, <https://doi.org/10.1029/2005WR004376>, 2006b.
- Klaus, J. and McDonnell, J. J.: Hydrograph separation using stable isotopes: Review and evaluation, *J. Hydrol.*, 505, 47–64, <https://doi.org/10.1016/j.jhydrol.2013.09.006>, 2013.
- Koutsouris, A. J. and Lyon, S. W.: Advancing understanding in data-limited conditions: estimating contributions to streamflow across Tanzania’s rapidly developing Kilombero Valley, *Hydrol. Sci. J.*, 63, 197–209, <https://doi.org/10.1080/02626667.2018.1426857>, 2018.
- Kuczera, G. and Parent, E.: Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm, *J. Hydrol.*, 211, 69–85, [https://doi.org/10.1016/S0022-1694\(98\)00198-X](https://doi.org/10.1016/S0022-1694(98)00198-X), 1998.
- Laudon, H. and Slaymaker, O.: Hydrograph separation using stable isotopes, silica and electrical conductivity: an alpine example, *J. Hydrol.*, 201, 82–101, [https://doi.org/10.1016/S0022-1694\(97\)00030-9](https://doi.org/10.1016/S0022-1694(97)00030-9), 1997.
- Leslie, D. L., Welch, K. A., and Lyons, W. B.: A temporal stable isotopic ($\delta^{18}\text{O}$, δD , d -excess) comparison in glacier meltwater streams, Taylor Valley, Antarctica, *Hydrol. Process.*, 31, 3069–3083, <https://doi.org/10.1002/hyp.11245>, 2017.
- Lopes Saraiva Okello, A. M., Uhlenbrook, S., Jewitt, G. P. W., Masih, I., Riddell, E. S., and Van der Zaag, P.: Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid mesoscale catchment, *Hydrol. Process.*, 32, 1334–1350, <https://doi.org/10.1002/hyp.11491>, 2018.
- Lyon, A.: Why are Normal Distributions Normal?, *Br. J. Philos. Sci.*, 65, 621–649, <https://doi.org/10.1093/bjps/axs046>, 2013.
- Maule, C. P., Chanasyk, D. S., and Muehlenbachs, K.: Isotopic determination of snow-water contribution to soil water and groundwater, *J. Hydrol.*, 155, 73–91, [https://doi.org/10.1016/0022-1694\(94\)90159-7](https://doi.org/10.1016/0022-1694(94)90159-7), 1994.
- MeteoSwiss: Documentation of MeteoSwiss Grid-Data Products: Daily Precipitation (final analysis): RhiresD, Zürich, available at: https://www.meteoswiss.admin.ch/content/dam/meteoswiss/fr/climat/le-climat-suisse-en-detail/doc/ProdDoc_RhiresD.pdf (last access: 31 January 2019), 2016.
- MeteoSwiss: Documentation of MeteoSwiss Grid-Data Products: Daily mean, minimum and maximum temperature, Zürich, available at: https://www.meteoswiss.admin.ch/content/dam/meteoswiss/de/service-und-publikationen/produkt/raeumliche-daten-temperatur/doc/ProdDoc_TabsD.pdf (last access: 31 January 2019), 2017.

- Metropolis, N. and Ulam, S.: The Monte Carlo Method, *J. Am. Stat. Assoc.*, 44, 335–341, <https://doi.org/10.1080/01621459.1949.10483310>, 1949.
- Michelon, A.: Weather dataset from Vallon de Nant, Switzerland, until July 2017, Zenodo, <https://doi.org/10.5281/ZENODO.1042473>, 2017.
- Neal, R. M.: Annealed importance sampling, *Stat. Comput.*, 11, 125–139, <https://doi.org/10.1023/A:1008923215028>, 2001.
- Oerter, E. J., Siebert, G., Bowling, D. R., and Bowen, G.: Soil water vapour isotopes identify missing water source for streamside trees, *Ecohydrology*, 21, e2083, <https://doi.org/10.1002/eco.2083>, 2019.
- Parnell, A. C., Inger, R., Bearhop, S., and Jackson, A. L.: Source partitioning using stable isotopes: coping with too much variation., *PLoS One*, 5, e9672, <https://doi.org/10.1371/journal.pone.0009672>, 2010.
- Parton, W. J. and Logan, J. A.: A model for diurnal variation in soil and air temperature, *Agric. Meteorol.*, 23, 205–216, [https://doi.org/10.1016/0002-1571\(81\)90105-9](https://doi.org/10.1016/0002-1571(81)90105-9), 1981.
- Pellerin, B. A., Wollheim, W. M., Feng, X., and Vörösmarty, C. J.: The application of electrical conductivity as a tracer for hydrograph separation in urban catchments, *Hydrol. Process.*, 22, 1810–1818, <https://doi.org/10.1002/hyp.6786>, 2007.
- Penna, D., Engel, M., Mao, L., Dell’Agnese, A., Bertoldi, G., and Comiti, F.: Tracer-based analysis of spatial and temporal variations of water sources in a glacierized catchment, *Hydrol. Earth Syst. Sci.*, 18, 5271–5288, <https://doi.org/10.5194/hess-18-5271-2014>, 2014.
- Penna, D., Zuecco, G., Crema, S., Trevisani, S., Cavalli, M., Pianezola, L., Marchi, L., and Borga, M.: Response time and water origin in a steep nested catchment in the Italian Dolomites, *Hydrol. Process.*, 31, 768–782, <https://doi.org/10.1002/hyp.11050>, 2017.
- Rice, K. C. and Hornberger, G. M.: Comparison of hydrochemical tracers to estimate source contributions to peak flow in a small, forested, headwater catchment, *Water Resour. Res.*, 34, 1755–1766, <https://doi.org/10.1029/98WR00917>, 1998.
- Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V., and Cox, D. R.: Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation, *P. Roy. Soc. London A*, 455, 3789–3805, 1999.
- Rothfuss, Y. and Javaux, M.: Reviews and syntheses: Isotopic approaches to quantify root water uptake: a review and comparison of methods, *Biogeosciences*, 14, 2199–2224, <https://doi.org/10.5194/bg-14-2199-2017>, 2017.
- Schaefli, B. and Kavetski, D.: Bayesian spectral likelihood for hydrological parameter inference, *Water Resour. Res.*, 53, 6857–6884, <https://doi.org/10.1002/2016WR019465>, 2017.
- Schaefli, B., Talamba, D. B., and Musy, A.: Quantifying hydrological modeling errors through a mixture of normal distributions, *J. Hydrol.*, 332, 303–315, <https://doi.org/10.1016/j.jhydrol.2006.07.005>, 2007.
- Schaefli, B., Nicótina, L., Imfeld, C., Da Ronco, P., Bertuzzo, E., and Rinaldo, A.: SEHR-ECHO v1.0: a Spatially Explicit Hydrologic Response model for ecohydrologic applications, *Geosci. Model Dev.*, 7, 2733–2746, <https://doi.org/10.5194/gmd-7-2733-2014>, 2014.
- Schmieder, J., Hanzer, F., Marke, T., Garvelmann, J., Warscher, M., Kunstmann, H., and Strasser, U.: The importance of snowmelt spatiotemporal variability for isotope-based hydrograph separation in a high-elevation catchment, *Hydrol. Earth Syst. Sci.*, 20, 5015–5033, <https://doi.org/10.5194/hess-20-5015-2016>, 2016.
- Scholl, M., Eugster, W., and Burkard, R.: Understanding the role of fog in forest hydrology: stable isotopes as tools for determining input and partitioning of cloud water in montane forests, *Hydrol. Process.*, 25, 353–366, <https://doi.org/10.1002/hyp.7762>, 2011.
- Scholl, M. A., Gingerich, S. B., and Tribble, G. W.: The influence of microclimates and fog on stable isotope signatures used in interpretation of regional hydrology: East Maui, Hawaii, *J. Hydrol.*, 264, 170–184, [https://doi.org/10.1016/S0022-1694\(02\)00073-2](https://doi.org/10.1016/S0022-1694(02)00073-2), 2002.
- Stock, B. C., Jackson, A. L., Ward, E. J., Parnell, A. C., Phillips, D. L., and Semmens, B. X.: Analyzing mixing systems using a new generation of Bayesian tracer mixing models, edited by: Nelson, D., *PeerJ*, 6, e5096, <https://doi.org/10.7717/peerj.5096>, 2018.
- Strahler, A. N.: Hypsometric (area-altitude) analysis of erosional topography, *GSA Bull.*, 63, 1117–1142, [https://doi.org/10.1130/0016-7606\(1952\)63\[1117:HAAOET\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1952)63[1117:HAAOET]2.0.CO;2), 1952.
- Thornton, J. M., Mariethoz, G., and Brunner, P.: A 3D geological model of a structurally complex Alpine region as a basis for interdisciplinary research, *Sci. Data*, 5, 180238, <https://doi.org/10.1038/sdata.2018.238>, 2018.
- Uehara, Y. and Kume, A.: Canopy Rainfall Interception and Fog Capture by *Pinus pumila* Regel at Mt. Tateyama in the Northern Japan Alps, Japan, Arctic, Antarctic, *Alp. Res.*, 44, 143–150, <https://doi.org/10.1657/1938-4246-44.1.143>, 2012.
- Varin, C., Reid, N., and Firth, D.: An overview of composite likelihood methods, *Stat. Sin.*, 21, 5–42, 2011.
- Vasdekis, V. G. S., Rizopoulos, D., and Moustaki, I.: Weighted pairwise likelihood estimation for a general class of random effects models, *Biostatistics*, 15, 677–689, <https://doi.org/10.1093/biostatistics/kxu018>, 2014.
- Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39, 8, <https://doi.org/10.1029/2002WR001642>, 2003.
- Weijis, S. V., Mutzner, R., and Parlange, M. B.: Could electrical conductivity replace water level in rating curves for alpine streams?, *Water Resour. Res.*, 49, 343–351, <https://doi.org/10.1029/2012WR012181>, 2013.
- Wels, C., Cornett, R. J., and Lazerte, B. D.: Hydrograph separation: A comparison of geochemical and isotopic tracers, *J. Hydrol.*, 122, 253–274, [https://doi.org/10.1016/0022-1694\(91\)90181-G](https://doi.org/10.1016/0022-1694(91)90181-G), 1991.
- Winograd, I. J., Riggs, A. C., and Coplen, T. B.: The relative contributions of summer and cool-season precipitation to groundwater recharge, Spring Mountains, Nevada, USA, *Hydrogeol. J.*, 6, 77–93, <https://doi.org/10.1007/s100400050135>, 1998.
- Zappa, M., Vitvar, T., Rücker, A., Melikadze, G., Bernhard, L., David, V., Jans-Singh, M., Zhukova, N., and Sanda, M.: A Trilateral program for estimating the link between snow resources and hydrological droughts, *Proc. Int. Assoc. Hydrol. Sci.*, 369, 25–30, <https://doi.org/10.5194/piahs-369-25-2015>, 2015.
- Zhu, X., Wu, T., Zhao, L., Yang, C., Zhang, H., Xie, C., Li, R., Wang, W., Hu, G., Ni, J., Du, Y., Yang, S., Zhang, Y., Hao, J., Yang, C., Qiao, Y., and Shi, J.: Exploring the contribution of precipitation to water within the active layer during the thawing pe-

riod in the permafrost regions of central Qinghai-Tibet Plateau by stable isotopic tracing, *Sci. Total Environ.*, 661, 630–644, <https://doi.org/10.1016/J.SCITOTENV.2019.01.064>, 2019.