Testing the truncation of travel times with StorAge Selection functions using deuterium and tritium as tracers

Nicolas B. Rodriguez[1,2], Laurent Pfister[1], Erwin Zehe[2], and Julian Klaus[1]

[1]Catchment and Eco-Hydrology Research Group, Environmental Research and Innovation Department, Luxembourg Institute of Science and Technology, Belvaux, Luxembourg
[2]Institute of Water Resources and River Basin Management, Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence: Nicolas Rodriguez (nicolas.rodriguez@list.lu)

Abstract. Catchment travel time distributions (TTDs) are an efficient concept to summarize the time-varying 3-dimensional transport of water and solutes to an outlet in a single function of water age and to estimate catchment storage by leveraging information contained in tracer data (e.g. ^{2}H and ^{3}H). It is argued that the increasing use of the stable isotopes of O and H compared to tritium as tracers has truncated our vision of streamflow TTDs, meaning that the long tails associated with old water are neglected. However, the reasons for the truncation of the TTD tails are still obscured by methodological and data limitations. In this study, we went beyond these limitations and tested the hypothesis that streamflow TTDs calculated using only deuterium (^{2}H) or only tritium (^{3}H) are different. We similarly tested if the mobile catchment storage (derived from the TTDs) associated with each tracer is different. For this we additionally constrained a model successfully developed to simulate high-frequency stream deuterium measurements with about 30 stream tritium measurements over the same period (2015–2017). We used data from the forested headwater Weierbach catchment (42 ha) in Luxembourg. The streamflow TTDs were estimated in unsteady conditions by using both tracers coherently within a framework based on StorAge Selection (SAS) functions. We found equal TTDs and equal mobile storage between the ^{2}H- and ^{3}H-derived estimates. The truncation hypothesis was thus rejected. The small differences we found could be explained by the calculation uncertainties and by a limited sampling frequency for tritium. Using both stable and radioactive isotopes of H as tracers reduced the age and storage uncertainties. Although tritium and stable isotopes had redundant information about younger water, using both tracers better exploited the more specific information about longer ages that ^{3}H inherently contains, and it could be even better in the next decades. The two tracers thus had overall different information contents. Tritium was however slightly more informative and cost-effective than stable isotopes for travel time analysis. We thus reiterate the call of Stewart et al. (2012) to measure tritium in the streams for travel time analysis, and emphasize the need for high-frequency tritium sampling in future studies to match the resolution in stable isotopes.

1 Introduction

Sustainable water resource management is based upon a sound understanding of how much water is stored in catchments, and how it is released to the streams. Isotopic tracers such as deuterium (^{2}H), oxygen 18 (^{18}O), and tritium (^{3}H) have become
25   the cornerstone of several approaches to tackle these two critical questions (Kendall and McDonnell, 1998). For instance, hydrograph separation using stable isotopes of O and H (Buttle, 1994; Klaus and McDonnell, 2013) has unfolded the difference between catchments hydraulic response (i.e. streamflow) and chemical response (e.g. solutes) (Kirchner, 2003) related to the different concepts of water celerity and water velocity (McDonnell and Beven, 2014). Isotopic tracers have also been the backbone to unravel water flow paths in soils (Sprenger et al., 2016), and to distinguish soil water going back to the atmosphere
30   and flowing to the streams (Brooks et al., 2010; McDonnell, 2014; McCutcheon et al., 2017; Berry et al., 2018; Dubbert et al., 2019).

    The travel time distribution (TTD) is nevertheless the concept relying the most on isotopic tracers (McGuire and McDonnell, 2006). TTDs provide a concise summary of water flow paths to an outlet by leveraging the information on storage and release contained in tracer input-output relationships. TTDs are essential to link water quantity to water quality (Hrachowitz et al.,
35   2016), for example by allowing calculations of stream solute dynamics from a hydrological model (Rinaldo and Marani, 1987; Maher, 2011; Benettin et al., 2015a, 2017a). TTDs are commonly calculated from isotopic tracers in many sub-disciplines of hydrology and thus have the potential to link the individual studies focused on the various compartments of the critical zone (e.g. groundwater and surface water) (Sprenger et al., 2019). ^{3}H has been used as an environmental tracer since the late 1950s (Begemann and Libby, 1957; Eriksson, 1958; Dinçer et al., 1970; Hubert, 1971; Martinec, 1975) and it gained particular
40   momentum in the eighties with its use in diverse TTD models (Małoszewski and Zuber, 1982; Stewart et al., 2010). It is argued that ^{3}H contains more information on the age of water than stable isotopes due to its radioactive decay (Stewart et al., 2012). For example, low tritium content generally indicates old water in which most of the ^{3}H from nuclear tests has decayed. Despite its potential, ^{3}H is used only rarely in travel time studies nowadays (Stewart et al., 2010), most likely because high precision analyses are laborious (Morgenstern and Taylor, 2009) and rather expensive. In contrast, the use of stable isotopes in travel
45   time studies has soared in the last three decades (Kendall and McDonnell, 1998; McGuire and McDonnell, 2006; Fenicia et al., 2010; Heidbuechel et al., 2012; Klaus et al., 2015a; Benettin et al., 2015a; Pfister et al., 2017; Rodriguez et al., 2018). This is notably due to the fast and low-cost analyses provided by recent advances in laser spectroscopy (e.g. Lis et al., 2008; Gupta et al., 2009; Keim et al., 2014) and the associated technological progress in sampling techniques of various water sources (Berman et al., 2009; Koehler and Wassenaar, 2011; Herbstritt et al., 2012; Munksgaard et al., 2011; Pangle et al., 2013;
50   Herbstritt et al., 2019). According to Stewart et al. (2012) and Stewart and Morgenstern (2016), the limited use of ^{3}H may have cause a biased or "truncated" vision of stream TTDs, in which the the older ages associated with long TTD tails remain mostly undetected by stable isotopes. Longer mean travel times (MTT) were inferred from ^{3}H than from stable isotopes in several studies employing both tracers (Stewart et al., 2010). Longer MTTs may have profound consequences for catchment storage, usually estimated from TTDs as $S = Q \times MTT$ (with Q the flux through the catchment), assuming steady-state flow

55 conditions (i.e. $S(t) = \overline{S(t)} = S$, $Q(t) = \overline{Q(t)} = Q$, $MTT(t) = \overline{MTT(t)} = MTT$) (McGuire and McDonnell, 2006; Soulsby et al., 2009; Birkel et al., 2015; Pfister et al., 2017). Under this assumption, a truncated TTD would result in an underestimated MTT thus an underestimated catchment storage. A different perspective on catchment storage and on its relation with travel times may however be adopted by calculating storage from unsteady TTDs.

A water molecule that reached an outlet has only one age, defined as the time it took to get there. This age is not affected
60 by the isotopes (^{2}H, ^{3}H, and ^{18}O) carried by the molecule and used as tracers, because they do not influence its flow path or its advective velocity or its self diffusion in water (Devell, 1962). The use of different methods of travel time analysis for stable isotopes of O and H and for ^{3}H (e.g. amplitudes of seasonal variations vs. radioactive decay) was hence first pointed out as a main reason for the discrepancies in MTT (Stewart et al., 2012). Further research is thus needed for developing mathematical frameworks that coherently incorporate stable isotopes of O and H and ^{3}H in travel time calculations. Moreover,
65 several limiting assumptions were used in previous studies employing ^{3}H to derive MTTs, which are in themselves insufficient statistics to describe various aspects (e.g. shape, modes, percentiles) of the TTDs. For example, the steady-state assumption has been used in almost all ^{3}H travel time studies (McGuire and McDonnell, 2006; Stewart et al., 2010; Cartwright and Morgenstern, 2016; Duvert et al., 2016; Gallart et al., 2016). Yet, time variance is a fundamental characteristic of TTDs (Botter et al., 2011; Rinaldo et al., 2015), and it has been acknowledged in simulations of stream ^{3}H only very recently (Visser et al.,
70 2019). Recharge models are also often employed to account only indirectly for the impact of evapotranspiration fluxes (ET) on the catchment inputs in ^{3}H (Stewart et al., 2007) and for the TTD of ET. In contrast for stable isotopes, explicit considerations of ET and of the influence of its TTD on the streamflow TTD are becoming common (van der Velde et al., 2015; Visser et al., 2019). Finally, more guidance on the calibration of the TTD models against ^{3}H measurements is needed (see e.g. Gallart et al., 2016). Especially, uncertainties of ^{3}H-inferred age estimates may have been overlooked, while these could explain the
75 differences with the stable isotope-inferred age estimates.

Besides methodological problems, the reasons for the age differences (hence apparent storage or mixing) are still not well understood, because little is known about the true age information content of ^{3}H compared to stable isotopes. First, ^{3}H sampling in catchments typically differs from stable isotope sampling in terms of frequency and flow conditions. Stable isotope records in precipitation and in the streams have lately shown increasing resolution, covering a wide range of flow conditions (McGuire
80 et al., 2005; Benettin et al., 2015a; Birkel et al., 2015; Pfister et al., 2017; von Freyberg et al., 2017; Visser et al., 2019; Rodriguez and Klaus, in review). Tritium records in precipitation and streams are on the other hand usually at a monthly resolution in many places around the globe (IAEA and WMO, 2019; IAEA, 2019; Halder et al., 2015). Only a handful of travel time studies employing ^{3}H report more than a dozen stream samples for a given site and for different conditions than baseflow (e.g. Małoszewski et al., 1983; Visser et al., 2019). This general focus on baseflow ^{3}H sampling introduces by definition a bias
85 towards older water. Second, the natural variability of ^{3}H compared to that of stable isotopes has rarely been documented. ^{3}H in precipitation has returned to the pre-bomb levels, and like stable isotopes it shows a clear yearly seasonality (e.g. Stamoulis et al., 2005; Bajjali, 2012). However, ambiguous age estimates may still be obtained with ^{3}H in the northern hemisphere

because the current precipitation has similar ^{3}H concentrations than water recharged in the 1980s (Stewart et al., 2012). Higher sampling frequencies of precipitation ^{3}H are almost nonexistent. Rank and Papesch (2005) revealed a short term variability of precipitation ^{3}H likely due to different air masses, observed also during complex meteorological conditions such as hurricanes (Östlund, 2013). ^{3}H in streams also show some yearly seasonality (Różański et al., 2001; Rank et al., 2018), but short term dynamics are not well understood because high frequency data sets are limited. Dinçer et al. (1970) showed that short-term stream tritium variations can be caused by the melting of the snowpack from the current and the previous winters. In addition, the seasonally higher values of precipitation ^{3}H in Spring could explain some of the ^{3}H peaks observed in the large rivers (Rank et al., 2018). More studies employing both ^{3}H and stable isotopes and comparing their age information content are therefore crucial to understand travel times in catchments from a multi-tracer perspective.

In this study, we go beyond previous work and test the hypothesis that stream TTDs and the associated catchment storage are different (considering their uncertainties) when inferred from stable isotopes or from ^{3}H measurements used in a coherent mathematical framework for both tracers. For this, we use high frequency isotopic tracer data from an experimental headwater catchment in Luxembourg. Here we focus on the stable isotope of H (deuterium ^{2}H) for which we have more precise measurements. A transport model based on TTDs was recently developed and successfully applied to simulate a two-year high frequency (sub-daily) record of ^{2}H in the stream (Rodriguez and Klaus, in review). Here, we additionally constrain the same model within the same mathematical framework against nearly 30 stream samples of ^{3}H collected during highly varying flow conditions over the same period as for ^{2}H. We do not assume steady-state and do not rely on a recharge model by employing StorAge Selection functions to account for the type and the variability of the TTDs of Q and ET that affect the water age balance in the catchment. The tracer input-output relationships and the ^{3}H radioactive decay are accounted for in the method, which reduces ^{3}H age ambiguities. We provide guidance on how to jointly calibrate the model to both tracers and on how to derive likely ranges of storage estimates and travel time measures other than the MTT. This work addresses the following related research questions:

– Are travel times and storage inferred from a common transport model for ^{2}H and ^{3}H in disagreement?

– Are the water age information contents of ^{2}H and ^{3}H similar?

2 Methods

2.1 Study site description

This study is carried out in the Weierbach catchment, which has been the focus of an increasing number of investigations in the last few years about streamflow generation (Glaser et al., 2016, 2019b; Scaini et al., 2017, 2018; Carrer et al., 2019; Rodriguez and Klaus, in review), biogeochemistry (Moragues-Quiroga et al., 2017; Schwab et al., 2018), and pedology and geology (Juilleret et al., 2011; Gourdol et al., 2018).

The Weierbach is a forested headwater catchment of 42 ha located in northwestern Luxembourg (Fig. 1). The vegetation consists mostly of deciduous hardwood trees (European beech and Oak), and conifers (*Picea abies* and *Pseudotsuga menziesii*). Short vegetation covers a riparian area that is up to 3 m wide and that surrounds most of the stream. The catchment morphology is a deep V-shaped valley in a gently sloping plateau. The geology is essentially Devonian slate of the Ardennes massif, phyllades, and quartzite (Juilleret et al., 2011). Pleistocene Periglacial Slope Deposits (PPSD) cover the bedrock and are oriented parallel to the slope (Juilleret et al., 2011). The upper part of the PPSD (~ 0–50 cm) has higher drainable porosity than the lower part of the PPSD (~ 50–140 cm) (Gourdol et al., 2018; Martínez-Carreras et al., 2016). Fractured and weathered bedrock lies from ~ 140 cm depth to ~ 5 m depth on average. Below ~ 5 m depth lies the fresh bedrock that can be considered impervious. The climate is temperate and semi oceanic. The flow regime is governed by the interplay of seasonality between precipitation and evapotranspiration. Precipitation is fairly uniformly distributed over the year, and averages 953 mm/yr over 2006–2014 (Pfister et al., 2017). The runoff coefficient over the same period is 50 %. Streamflow (Q) is double-peaked during wetter periods (Martínez-Carreras et al., 2016), and single peaked during drier periods occurring normally in summer when evapotranspiration (ET) is high.

Based on previous modeling (e.g. Fenicia et al., 2014; Glaser et al., 2019b) and experimental studies (e.g. Martínez-Carreras et al., 2016; Juilleret et al., 2016; Scaini et al., 2017; Glaser et al., 2018), Rodriguez and Klaus (in review) proposed a perceptual model of streamflow generation in the Weierbach. In this model, the first and flashy peaks of double-peaked hydrographs are generated by precipitation falling directly into the stream, by saturation excess flow from the near-stream soils, and by infiltration excess overland flow in the riparian area. The second peaks are generated by delayed lateral subsurface flow. The lateral fluxes are assumed higher at the PPSD/bedrock interface due to the hydraulic conductivity contrasts (Glaser et al., 2016, 2019b). Lateral subsurface flows are thus accelerated when groundwater rises after a rapid vertical infiltration through the soils (Rodriguez and Klaus, in review). The model based on travel times presented in this study was developed in a step-wise manner based on this hypothesis of streamflow generation. The model's ability to simulate stream δ^2H dynamics helped to further confirm that these flow processes are active in the Weierbach (Rodriguez and Klaus, in review). Water flow paths and streamflow generation processes in this catchment are however not completely resolved. Other studies carried out in the Colpach catchment (containing the Weierbach) highlighted the potential role of lateral preferential flow through macropores in the highly heterogeneous soils for the generation first peaks of the hydrographs (Angermann et al., 2017; Loritz et al., 2017), contrary to the understanding from various studies in the Weierbach (Glaser et al., 2016, 2019a).

2.2 Hydrometric and tracer data

In this study we use precipitation (J, in mm/h), ET (mm/h), Q (mm/h), and δ^2H (‰) and ^{3}H (Tritium Units, T.U.) measurements in precipitation ($C_{P,2}$ and $C_{P,3}$ respectively) and in the stream ($C_{Q,2}$ and $C_{Q,3}$ respectively). Here a subscript 2 indicates deuterium (^{2}H) and a subscript 3 indicates tritium (^{3}H). The analysis in this study focuses on the period October 2015–October 2017 during which most samples were collected at higher frequencies than in the past (Fig. 2). Details on the hydrometric data collection (J, ET, Q), and on the ^{2}H sample collection and analysis are given in Rodriguez and Klaus (in review).
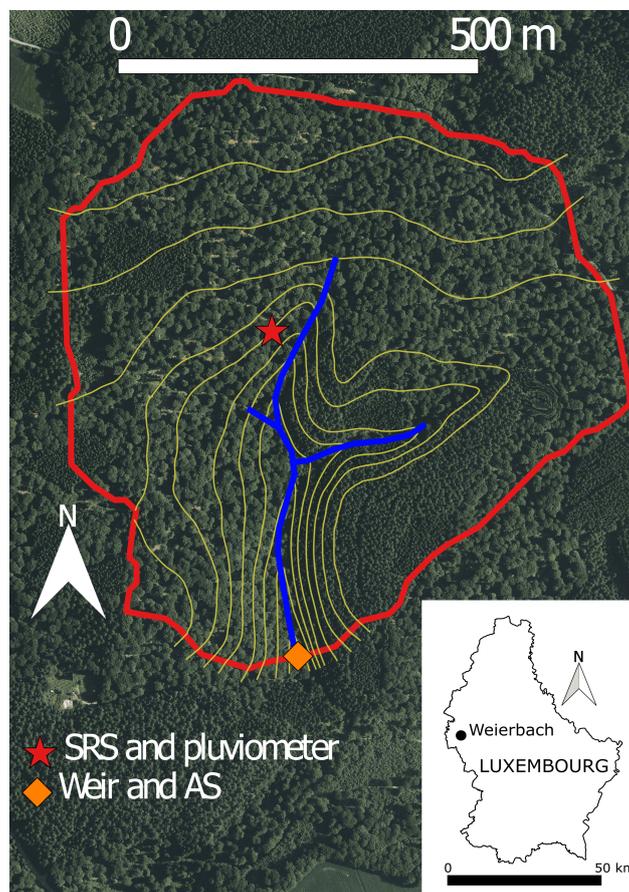
Figure 1. Map of the Weierbach catchment and its location in Luxembourg. The weir is located at coordinates (5°47'44" E, 49°49'38" N). SRS is the sequential rainfall sampler. AS is the stream autosampler. The elevation lines go by increments of 5 m from 460 m.a.s.l. downstream close to the weir location to 510 m.a.s.l. at the northern catchment divide.

The 1088 stream samples analyzed for ^{2}H were collected manually or automatically with an autosampler (AS, Fig. 1), resulting in samples every 15 hours on average over October 2015–October 2017. The 525 precipitation samples analyzed for ^{2}H were collected approximately every 2.5 mm rain increment (i.e. on average every 23 hours) with a sequential rainfall sampler (SRS) and in addition as bulk samples on a bi-weekly basis. The samples were analyzed at the Luxembourg Institute

155   of Science and Technology (LIST) using an LGR Isotope Water Analyzer, yielding for ^{2}H an analytical accuracy of 0.5 ‰ (equal to the LGR standard accuracy), and a precision maintained <0.5 ‰ (quantified as one standard deviation of the measured samples and standards).

The 24 stream samples analyzed for ^{3}H were selected from manual bi-weekly sampling campaigns to cover various flow ranges. The samples were analyzed by the GNS Science Water Dating Laboratory (Lower Hutt, New Zealand), which provides

160   high precision tritium measurements using electrolytic enrichment and liquid scintillation counting (Morgenstern and Taylor,

2009). The precision of the stream samples varies from roughly 0.07 T.U. to roughly 0.3 T.U., but is usually around 0.1 T.U. Monthly values of ^{3}H in precipitation were obtained for the Trier station (60 km from the Weierbach) until 2016 from the WISER database of the International Atomic Energy Agency (IAEA) (IAEA and WMO, 2019; Stumpp et al., 2014). The 2017 values were obtained from the Radiologie group of Bundesanstalt für Gewässerkunde (Schmidt et al., submitted). ^{3}H in

165 precipitation before 1978 was calculated by regression with data from Vienna, Austria (Stewart et al., 2017).

For both ^{2}H and ^{3}H, the time series of tracer in precipitation was interpolated between two consecutive samples (e.g. A and B) as being equal to the value of the next sample (i.e. B). Since no measurements of J, Q, ET, and $C_{P,2}$ are available before 2010, we looped back their values of the period October 2010–October 2015 periodically before 2010 as a best estimate of their past values. We aggregated the input data (J, ET, Q, $C_{P,2}$, $C_{P,3}$) to a resolution $\Delta t = 4$ hours, which is small enough to

170 capture the variability of flows and tracers in the input and simulate the variability of the flows and tracers in the output.
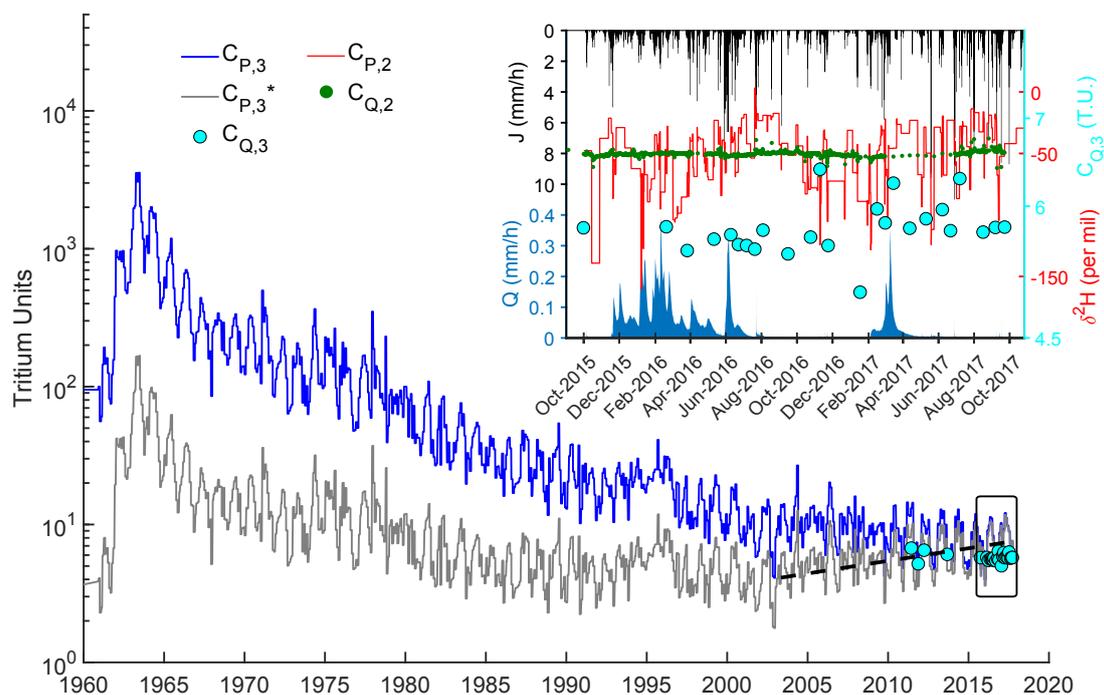
Figure 2. Data used in this study: ^{3}H in precipitation ($C_{P,3}$), the corresponding tritium activities accounting for radioactive decay until 2017 ($C^*_{P,3}$), δ^2H in precipitation $C_{P,2}$ (inset), precipitation J (inset), streamflow Q (inset), ^{3}H measurements in the stream ($C_{Q,3}$ both plots), and δ^2H in the stream ($C_{Q,2}$, inset). The period contained in the inset is represented as a rectangle in the bigger plot. The dashed line visually represents the increasing trend in $C^*_{P,3}$ that emerges as the effect of bomb peak tritium disappears (i.e. $C_{P,3}(t-T)$ stops decreasing around 2000 so $C^*_{P,3}(T,t) = C_{P,3}(t-T)e^{-\alpha T}$ starts decreasing with increasing T).

2.3 Mathematical framework

Mathematically, the streamflow TTD is related to the stream tracer concentrations $C_Q(t)$ according to the following Eq. (1):

$$C_Q(t) = \int\limits_{T=0}^{+\infty} C_P^*(T,t)\overleftarrow{p_Q}(T,t)\,dT \tag{1}$$

175   where T is water age, t is time of observation, $C_Q(t)$ is the stream tracer concentration, $\overleftarrow{p_Q}$ (probability distribution function, p.d.f.) is the stream backward TTD (Benettin et al., 2015b), and $C_P^*(T,t)$ is the tracer concentration of the water parcel reaching the outlet at time t with age T (this parcel was in the inflow at time $t-T$, its travel time is thus T). This equation is always verified for the exact (usually unknown) TTD, because it simply expresses the fact that the stream concentration is the volume-weighted arithmetic mean of the concentrations of the water parcels with different travel times at the outlet. $C_P^*(T,t)$ depends on T and t as separate variables if the tracer concentration of a water parcel in the catchment changes between injection time

180   $t-T$ and observation time t. For solutes like silicon and sodium, the concentration can increase with age (Benettin et al., 2015a). For ^{3}H, radioactive decay with a constant $\alpha = 0.0563\ \text{yr}^{-1}$ implies $C_{P,3}^*(T,t) = C_{P,3}(t-T)\,e^{-\alpha T}$, where $C_{P,3}(t-T)$ is the concentration in precipitation measured at $t-T$. For ^{2}H, $C_{P,2}^*(T,t) = C_{P,2}(t-T)$. Thus, the streamflow TTD simultaneously verifies Eq. (2) and (3):

$$C_{Q,2}(t) = \int\limits_{T=0}^{+\infty} C_{P,2}^*(T,t)\overleftarrow{p_Q}(T,t)\,dT = \int\limits_{T=0}^{+\infty} C_{P,2}(t-T)\overleftarrow{p_Q}(T,t)\,dT \tag{2}$$

185   $$C_{Q,3}(t) = \int\limits_{T=0}^{+\infty} C_{P,3}^*(T,t)\overleftarrow{p_Q}(T,t)\,dT = \int\limits_{T=0}^{+\infty} C_{P,3}(t-T)\,e^{-\alpha T}\,\overleftarrow{p_Q}(T,t)\,dT \tag{3}$$

Practically, when measurements of ^{2}H and ^{3}H are used to inversely deduce the TTD by using Eq. 2 and 3, different TTDs may be found. These different TTDs may be called $\overleftarrow{p_{Q,2}}$ and $\overleftarrow{p_{Q,3}}$ for instance, referring to ^{2}H and ^{3}H, respectively. To avoid introducing more variables and to avoid confusion, we do not use the names $\overleftarrow{p_{Q,2}}$ and $\overleftarrow{p_{Q,3}}$ and we instead refer to the TTDs "constrained" by a given tracer, using a common symbol $\overleftarrow{p_Q}$. We do this also to stress that the exact (true) TTD

190   must simultaneously verify both Eq. (2) and (3), and that two different TTDs $\overleftarrow{p_{Q,2}}$ and $\overleftarrow{p_{Q,3}}$ cannot physically exist. This is a fundamental difference from previous work that assumed two different TTDs, using for example Eq. (3) for ^{3}H and another method for ^{2}H (the sine-wave approach) (e.g. Małoszewski et al., 1983). The framework in this study also uses the fact that the same functional form of streamflow TTD needs to simultaneously explain both tracers to be valid, unlike previous work that used different TTD models for different tracers (Stewart and Thomas, 2008).

195   ## 2.4 Transport model based on TTDs

Most of the previous travel time studies using tritium assumed steady-state and an analytical shape for the stream TTD, and fitted the parameters of the analytical function using the framework described in 2.3. In this study, the TTDs are unsteady (i.e.

time-varying) and cannot be analytically described. Still, they can be calculated by numerically solving the "Master Equation" (Botter et al., 2011). This method has been applied in several recent studies (e.g. van der Velde et al., 2015; Harman, 2015;

200   Benettin et al., 2017b), and is described in more details by Benettin and Bertuzzo (2018). The numerical method used to solve this equation in this study is described by Rodriguez and Klaus (in review).

Essentially, the Master Equation is a water balance equation where storage and fluxes are labeled with age categories. The Master Equation is thus a partial differential equation expressing the fact that the amount of water in storage having age T changes with time because of new water introduced by precipitation $J(t)$, because of water aging, and because of losses to

205   catchment outflows $ET(t)$ and $Q(t)$. Solving the Master Equation requires knowledge (or an assumption about the shape) of the StorAge Selection (SAS) functions Ω_Q and Ω_{ET} of outflows Q and ET, which conceptually represent how likely water ages in storage are to be present in the outflows at a given time. Solving the Master Equation yields the distribution of ages in storage at every moment, that can be represented in a cumulative form with age-ranked storage S_T, defined as the amount of water in storage (e.g. 10 mm) younger than T (e.g. 1 year) at time t. $T \rightarrow S_T$ is just a mathematical change of variable,

210   and it has no meaning respective to the location or depth of a certain water parcel with age T in the catchment. By definition $\lim_{T \rightarrow +\infty} S_T = S(t)$, where $S(t)$ is catchment storage. Ω_Q and Ω_{ET} are functions of S_T and cumulative distributions functions (c.d.f.) for numerical convenience. SAS functions are closely linked to TTDs, such that one can be found from the other using the following expression (here for Q, but valid for other outflows):

$$\overleftarrow{p_Q}(T,t) = \frac{\partial}{\partial T}(\Omega_Q(S_T,t)) \tag{4}$$

215   The partial derivative with respect to age T ensures the transition from c.d.f. to p.d.f. Assuming a parameterized form for Ω_Q and Ω_{ET} and calibrating their parameters using the framework defined in 2.3 yields time-varying TTDs constrained by the tracers in the outflows.

In this study, we assumed that Ω_{ET} is a function of only S_T and it is gamma distributed with a mean parameter μ_{ET} (mm) and a scale parameter θ_{ET} (mm). Rodriguez and Klaus (in review) showed that in the Weierbach, a weighted sum of

220   three components in the streamflow SAS function is more consistent the superposition of streamflow generation processes (i.e. saturation excess flow, saturation overland flow, lateral subsurface flow, see Sect. 2.1) than a single component. This means that Ω_Q is written as a weighted sum of three c.d.f.s (see appendix A1) (Rodriguez and Klaus, in review):

$$\Omega_Q(S_T,t) = \lambda_1(t)\,\Omega_1(S_T) + \lambda_2(t)\,\Omega_2(S_T) + \lambda_3(t)\,\Omega_3(S_T) \tag{5}$$

$\lambda_1(t)$, $\lambda_2(t)$, and $\lambda_3(t)$ are time-varying weights summing to 1. Essentially, $\lambda_1(t)$ is the smallest weight and it is parame-

225   terized to increase sharply during flashy streamflow events, using parameters λ_1^*, f_0, S_{th} (mm), and ΔS_{th} (mm). $\lambda_2(t) = \lambda_2$ is calibrated, and $\lambda_3(t)$ just deduced by difference. Ω_1 is a cumulative uniform distribution over S_T in $[0, S_u]$ (with S_u a parameter in mm). Ω_1 represents the young water contributions associated with short flow paths during flashy streamflow events. Ω_1 corresponds to processes in the near stream area: saturation excess flow, saturation overland flow, and rain on the stream

(Rodriguez and Klaus, in review). Ω_2 and Ω_3 are gamma-distributed with mean parameters μ_1 and μ_2 (mm), and scale parameters θ_1 and θ_2 (mm) respectively. Ω_2 and Ω_3 represent older water that is always contributing to the stream. This older water consists of groundwater stored in the weathered bedrock that flows laterally in the subsurface. Note that we used the same functional form of $\Omega_Q(S_T, t)$ for ^{2}H and ^{3}H to keep the functional form of the TTDs consistent between the tracers.

2.5 Model initialization and numerical details

Numerically solving the Master Equation requires an estimation of catchment mobile storage $S(t)$. In this context, $S(t)$ represents the sum of "dynamic" (or "active") storage and "inactive" (or "passive") storage (Fenicia et al., 2010; Birkel et al., 2011; Soulsby et al., 2011; Hrachowitz et al., 2013). In this study the model is initialized in October 1915 with storage $S(t=0) = S_{ref}$. This initial value is chosen large enough to sustain Q and ET during drier periods and to store water that is sufficiently old to satisfy Eq. (1). $S(t)$ is then simply deduced from the water balance as $S(t) = S_{ref} + \int_{x=0}^{t}(J(x) - Q(x) - ET(x))\,dx$. The initial age distribution in storage is exponential with a mean of 1.7 years, the estimated Mean Residence Time (MRT) by Pfister et al. (2017). The model is then run with time steps $\Delta t = 4$ hours and age resolution $\Delta T = 8$ hours. This way the computational cost is balanced with the resolution of the simulations in δ^2H. The period October 1915–October 2015 serves as a long spin-up period to completely remove the impact of the initial conditions. This means that S_{ref} and the initial age distribution in storage do not influence the results over October 2015–October 2017. $ET(t)$ is taken equal to potential evapotranspiration $PET(t)$ except that it tends non-linearly towards 0 (using a constant smoothing parameter n) when storage $S(t)$ decreases below $S_{ref} - S_{root}$ (mm), where S_{root} is a parameter accounting for the water amount accessible by ET (appendix A2).

2.6 Model calibration

The parameters of the SAS functions and the other model parameters were calibrated using a Monte Carlo technique. In total, 12 parameters were calibrated (Table 1). The initial ranges were selected based on parameter feasible values (e.g. f_0 between 0 and 1 by definition), on previous estimations (e.g. S_{th}), on hydrological data (e.g. S_u and ΔS_{th} deduced from average precipitation depths), and on initial tests on the parameter ranges (e.g. μ and θ). These ranges allow a wide range of shapes of SAS functions while minimizing numerical errors (occurring for example for $S_T > S(t)$).

Unlike our previous modeling work in this catchment (Rodriguez and Klaus, in review), we fixed the initial storage in the model S_{ref} to 2000 mm. We did this to reduce the degrees of freedom when sampling the parameter space in order to limit the impact of numerical errors on the calibration. These errors are due to numerical truncation of $\Omega_Q(S_T, t)$ when a considerable part (e.g. a few percent) of its tail extends above $S(t)$. This occurs when parameters μ_2, μ_3, θ_2, and θ_3 are too large compared to S_{ref} when the latter is also randomly sampled. Choosing a constant large value for S_{ref} thus guarantees the absence of truncation errors. S_{ref} has little influence on the storage deduced from travel times since the ages sampled by streamflow are governed only by μ_2, μ_3, θ_2, and θ_3. These parameters are independent of S_{ref} as long at it allows sufficiently old water to reside in storage, which is ensured by its large value and by the long spin-up period we used (100 years).

Table 1. Model parameters

Symbol	Type	Unit	Initial range	Description[a]
S_{th}	Calibrated	mm	$[20, 200]$	Storage threshold relative to S_{min} separating "dry" and "wet" periods
ΔS_{th}	Calibrated	mm	$[0.1, 20]$	Threshold in short term storage changes identifying "first" peaks in hydrographs
S_u	Calibrated	mm	$[1, 50]$	Range of the uniformly distributed Ω_1
f_0	Calibrated	–	$[0, 1]$	Young water coefficient for the dry periods
λ_1^*	Calibrated	–	$[0, 1]^b$	Maximum value of the weight $\lambda_1(t)$
λ_2	Calibrated	–	$[0, 1]$	Constant[c] value of the weight $\lambda_2(t)$
μ_2	Calibrated	mm	$[0, 1600]$	Mean parameter of the gamma distributed Ω_2
θ_2	Calibrated	mm	$[0, 100]$	Scale parameter of the gamma distributed Ω_2
μ_3	Calibrated	mm	$[0, 1600]$	Mean parameter of the gamma distributed Ω_3
θ_3	Calibrated	mm	$[0, 100]$	Scale parameter of the gamma distributed Ω_3
μ_{ET}	Calibrated	mm	$[0, 1600]$	Mean parameter of the gamma distributed Ω_{ET}
θ_{ET}	Calibrated	mm	$[0, 100]$	Scale parameter of the gamma distributed Ω_{ET}
S_{root}	Constant	mm	150	Water amount accessible by ET
m	Constant	–	1000	Smoothing parameter for the calculation of $\lambda_1(t)$
n	Constant	–	20	Smoothing parameter for the calculation of $ET(t)$ from $PET(t)$
Δt^*	Constant	hours	8	Width of the moving time window used to calculate short term storage variations $\overline{\Delta S(t)}$

[a] Details about the equations involving these parameters are given in appendix A1 and in Rodriguez and Klaus (in review)
[b] λ_1^* is in fact sampled between 0 and $1 - \lambda_2 \leq 1$ to ensure that $\sum_{n=1}^{3} \lambda_k(t) = 1$
[c] $\lambda_1(t)$ varies, λ_2 is constant, and $\lambda_3(t)$ varies and it is deduced using $\lambda_3(t) = 1 - \lambda_2 - \lambda_1(t)$

The first step of the Monte Carlo procedure we employed consists in randomly sampling parameters from the uniform prior distributions with ranges defined in Table 1. 12,096 sets of the 12 calibrated parameters were sampled as a Latin Hypercube (LHS, Helton and Davis, 2003). The model was then run over October 1915–October 2017, and its performance was evaluated over October 2015–October 2017. We evaluated model performance in a multi-objective manner, by using separate objective functions for ^{2}H and ^{3}H. For deuterium, we used the Nash-Sutcliffe Efficiency (NSE):

$$E_2 = 1 - \frac{\sum_{k=1}^{N_2}(C_{Q,2}(t_k) - \delta^2 H(t_k))^2}{\sum_{k=1}^{N_2}(\delta^2 H(t_k) - \overline{\delta^2 H})^2} \tag{6}$$

where $N_2 = 1,016$ is the number of deuterium observations in the stream. For tritium, we used the Mean Absolute Error:

$$E_3 = \sum_{j=1}^{N_3} |C_{Q,3}(t_j) - {}^3H(t_j)| \tag{7}$$

where $N_3 = 24$ is the number of tritium observations in the stream. We used the MAE for tritium because it is common to report errors in T.U., and because of the limited variance of stream ^{3}H (due to the low number of samples and the low

variability) making the NSE less appropriate (Gallart et al., 2016). The behavioral parameter sets that are used for uncertainty calculations and further analysis were selected based on threshold values L_2 and L_3 for the performance measures E_2 and E_3 respectively. Parameter sets were considered behavioral for deuterium simulations if $E_2 > L_2 = 0$, and behavioral for tritium simulations if $E_3 < L_3 = 0.5$ T.U. We subsequently refer to these parameter sets and corresponding simulations as

275 "constrained by deuterium", "constrained by tritium", and as "constrained by both" when both performance criteria were used. We chose these constraints to get reasonable model fits to the data, to obtain a comparable number of behavioral parameter sets for ^{2}H and ^{3}H, and to maximize the amount of information gained about the parameters when adding a constraint on the model performance for a tracer. This information gain was assessed with the Kullback-Leibler Divergence D_{KL} between the posterior parameter distributions inferred from various combinations of constraints L_2 and L_3 (Sect. 2.7).

280 ## 2.7 Information contents of ^{2}H and ^{3}H

Loritz et al. (2018) and Loritz et al. (2019) recently used information theory to detect hydrological similarity between hill-slopes of the Colpach catchment and to compare topographic indexes in the Attert catchment in Luxembourg. Thiesen et al. (2019) used information theory to build an efficient predictor of rainfall-runoff events. Here we leverage information theory to evaluate our model parameter uncertainty. For this we calculated the expected information content of the posterior parameter

285 distributions constrained by deuterium or tritium using the Shannon entropy $\mathcal{H}$:

$$\mathcal{H}(X|^iH) = -\sum_{k=1}^{n_I} f(I_k) \log_2 f(I_k) \tag{8}$$

In this equation, the parameter X (e.g. μ_1) takes values (e.g. 125 mm) falling in intervals I_k (e.g. $[100, 150]$ mm) that do not intersect each other and which union $\cup_{k=1}^{n_I} I_k$ equals I_X, the total interval of values on which X is defined (e.g. $[50, 500]$ mm). The definitions of the n_I intervals I_k for each parameter depend on the binning of the parameter values, given in Table

290 2. The posterior probability distribution f defines the probability of the parameter X to be in a certain state (i.e. to take a value falling in an interval I_k), when constrained by the criterion $E_2 > L_2 (i = 2)$ or $E_3 < L_3 (i = 3)$. When using the logarithm of base 2, $\mathcal{H}$ is expressed in bits of information contained in the posterior distribution f. The uniform distribution over I_X has the maximum possible entropy. Lower values of $\mathcal{H}$ thus indicate that the posterior distribution is not flat, hence less uncertain than the uniform prior distribution. In general, lower values of $\mathcal{H}$ indicate less uncertain parameters.

295 We also use the Kullback-Leibler Divergence D_{KL} to evaluate the gain of information when ^{3}H is used in addition to ^{2}H to constrain model predictions or vice versa:

$$D_{KL}(X|(^2H \cap {}^3H), X|^iH) = \sum_{k=1}^{n_I} f(I_k) \log_2 \frac{f(I_k)}{g(I_k)} \tag{9}$$

where f is the posterior distribution constrained by $E_2 > L_2$ and $E_3 < L_3$, and g is the posterior distribution constrained only by $E_2 > L_2 (i = 2)$ or only by $E_3 < L_3 (i = 3)$. D_{KL} is expressed in bits of information gained when the knowledge about

300   a parameter posterior distribution is updated by adding another tracer. D_{KL} can also be used to evaluate the gain of information from prior to posterior parameter distributions (by using $g =$ prior and $f =$ posterior). Calculating D_{KL} also requires binning the parameter values to define the intervals I_k and calculate the distributions f and g. The binning for each parameter (Table 2) was chosen such that the resulting histograms visually reveal the underlying structure of the parameter values, while avoiding uneven features and irregularities (e.g. very spiky histograms).

305   3   Results

3.1   Calibration results

148 parameter sets were behavioral for deuterium simulations, with E_2 ranging from $L_2 = 0$ to 0.24. 181 parameter sets were behavioral for tritium simulations, with E_3 ranging from 0.24 T.U. to $L_3 = 0.5$ T.U. Additionally, 16 parameter sets were behavioral for both tritium and deuterium simulations, with E_2 ranging from $L_2 = 0$ to 0.19 and E_3 ranging from 0.36 T.U. to 310 $L_3 = 0.5$ T.U. These solutions show that a reasonable agreement between the model fit to ^{2}H and the model fit to ^{3}H can be found.

     The behavioral posterior parameter distributions constrained by deuterium or tritium or by both generally had similar ranges than their prior distributions, except notably for μ_2, θ_2, μ_3, and θ_3 (Table 2). To assess the reduction of parameter uncertainty, we calculated and compared the entropy of the prior and of the posterior distributions (Table 2). A visual inspection of the 315   posterior distributions was also made, and we show here only the parameters μ_2, θ_2, μ_3, and θ_3 (Fig. 3) that directly control the range of older water ages in streamflow, since they act mostly on the right-hand tail of the gamma components in Ω_Q. These parameters thus also have a direct influence on the catchment storage inferred via age-ranked storage S_T.

     Essentially, the results (Table 2 and Fig. 3) reveal that the parameter ranges decreased by adding information on ^{2}H or ^{3}H or both. This effect is particularly noticeable for f_0 and λ_1^*, which saw their upper boundary decrease, and for μ_2 and μ_3, 320   which saw their lower boundary increase considerably. These results also show that the posterior distributions depart from the uniform prior distributions when considering ^{2}H alone or ^{3}H alone (i.e. $\mathcal{H}(X|^iH) < \mathcal{H}(X)$ in Table 2). This effect is not very pronounced for most parameters, but clearly visible for λ_1^*, for μ_2 and μ_3 (e.g. uneven distributions of points in Fig. 3), and for μ_{ET}. The posterior distributions become considerably narrower when considering both tracers, since $\mathcal{H}(X|(^2H \cap ^3H))$ is much lower than $\mathcal{H}(X)$, which is visually represented by the distribution of points tending to cluster towards a corner in Fig. 325   3. Generally, more was learned about the likely parameter values by adding a constraint on ^{2}H simulations after constraining ^{3}H simulations than the opposite (i.e. generally $D_{KL}(X|(^2H \cap ^3H), X|^3H) \geq D_{KL}(X|(^2H \cap ^3H), X|^2H)$). Noticeable exceptions to this are the parameters μ_2, θ_2, and θ_3, which are more related to the older ages in streamflow and to catchment storage than the other parameters.

     Simulations of stream δ^2H captured both the slow and the fast dynamics of the observations when constrained by $E_2 > 0$ 330   (blue bands and blue curve Fig. 4). This is not the case for δ^2H simulations constrained only by $E_3 < 0.5$ T.U. (red bands).
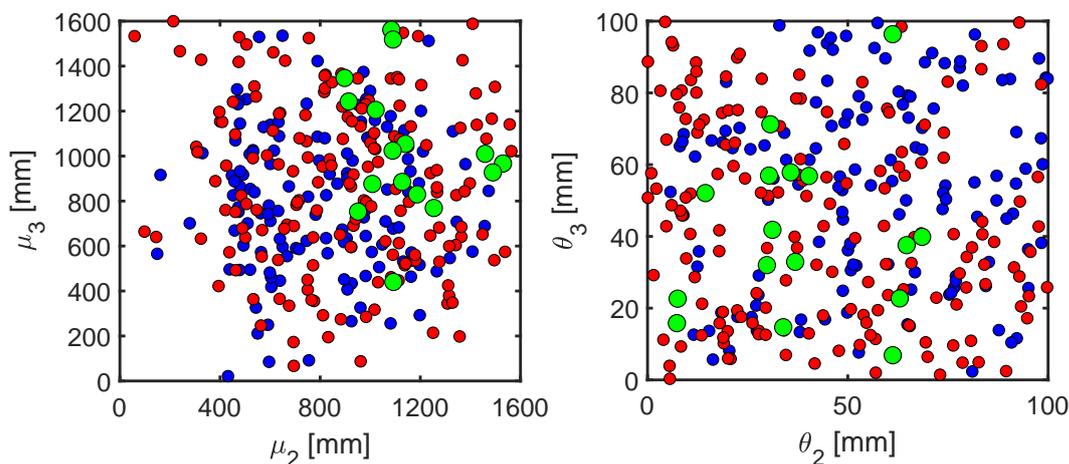
Figure 3. Distributions of SAS function mean (μ, left panel) and scale (θ, right panel) behavioral parameters directly controlling the selection of older ages by streamflow, constrained by deuterium (148 blue dots), or tritium (181 red dots), or both (16 green dots).

This shows that ^{3}H contains some information that is not in common with ^{2}H about the transport processes to the stream. Yet, simulations constrained by both criteria (green bands) have a smaller variability than those constrained only by $E_2 > 0$, suggesting that ^{3}H nevertheless contains some information that is common with ^{2}H.

Simulations of stream ^{3}H generally matched the observations better in 2017 than before 2017 (red bands and red curve
335 in Fig. 5). Some simulations (red bands) nevertheless matched the observations before 2017 relatively well. Similar to δ^2H simulations, both the slow and the fast simulation dynamics seemed necessary to reproduce the variability in ^{3}H observations (especially in 2017), although more stream samples would be needed to confirm that the model is accurate between the current measurement points. The higher stream ^{3}H values in 2017 that are better reproduced by the model correspond to an extended dry period during which streamflow responses are mostly flashy and short-lasting hydrographs. The associated ^{3}H values are
340 closer to precipitation ^{3}H, mostly around 10 T.U. The stream reaction to those higher values suggest a considerable influence of recent rainfall events on the stream, that steady-state TTD models relying only on tritium decay would probably struggle to simulate. This also suggests a stronger influence of old water in 2016 than in 2017 (see Sect. 4.4). Simulations constrained by deuterium (blue bands) tended to overestimate stream ^{3}H. Simulations constrained by both criteria (green bands) worked well in 2017, but they overestimated stream ^{3}H before 2017. Similar to δ^2H simulations, this suggests that ^{2}H and ^{3}H have common
345 but also distinct information contents about transport processes to the stream.
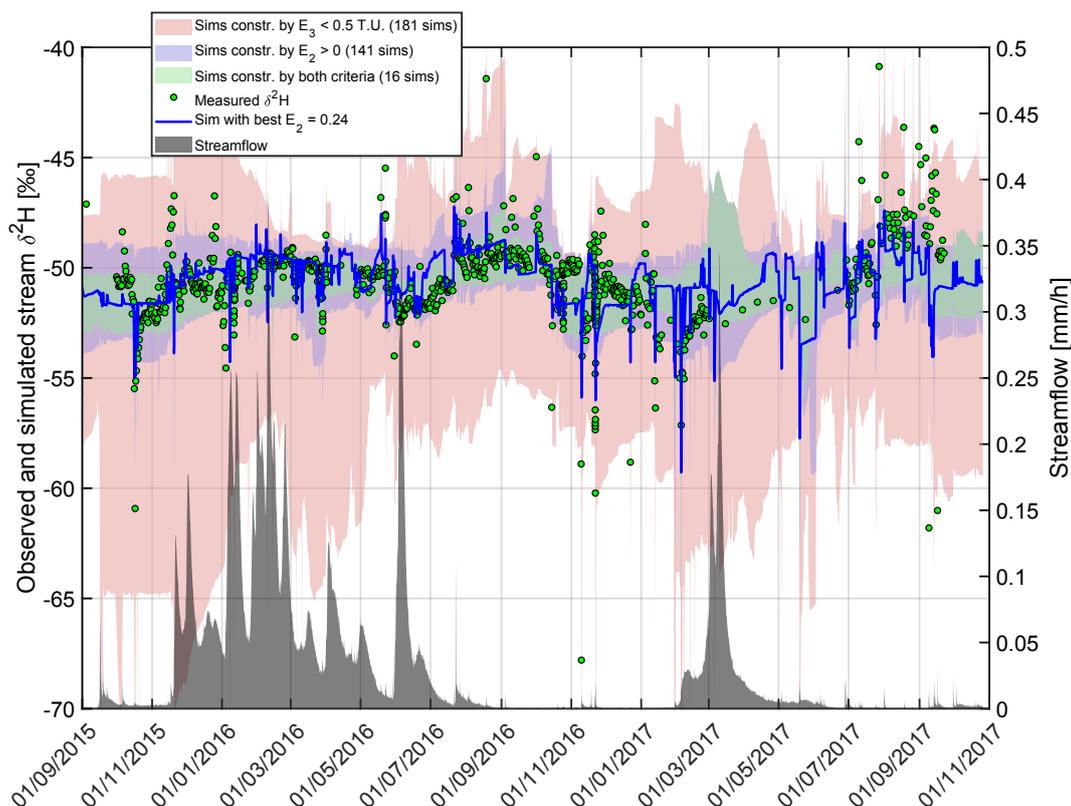
Figure 4. Simulations in deuterium. E_2 is the Nash-Sutcliffe efficiency in deuterium, and E_3 is the Mean Absolute Error in tritium units.

3.2 Storage and travel time results

For each behavioral parameter set, we calculated $\overleftarrow{P_Q}(T)$, the average stream TTD weighted by streamflow (over 2015–2017) in cumulative form (Fig. 6). Visually, there are no striking differences between $\overleftarrow{P_Q}(T)$ constrained by deuterium or by tritium, except a slightly wider spread for simulations constrained by tritium. The $\overleftarrow{P_Q}(T)$ constrained by both tracers clearly differ. The

350 associated curves (Fig. 6c) show a much narrower spread. They are also slightly shifted towards higher ages. We calculated various statistics of the distributions $\overleftarrow{P_Q}(T)$ constrained by the different performance criteria to compare them quantitatively (Table 3). This shows that the $\overleftarrow{P_Q}(T)$ constrained only by tritium systematically correspond to higher ages (and lower young water fractions) than those constrained only by deuterium. However, these age differences are small and could be explained by the uncertainties, which are larger for the younger age fractions, and systematically higher for tritium than for deuterium. The

355 $\overleftarrow{P_Q}(T)$ constrained by both tracers systematically correspond to the highest ages (and the lowest young water fractions). The corresponding uncertainties are much lower than when using individual tracers.
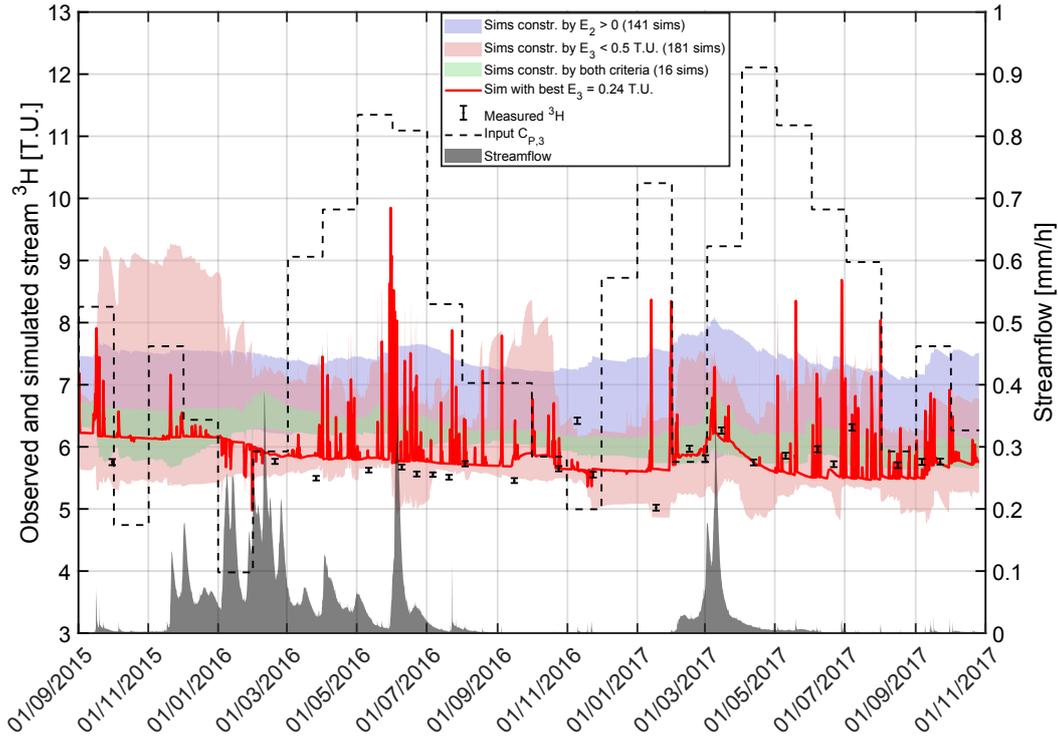
Figure 5. Simulations of stream concentrations in tritium compared to observations and to the variability in precipitation.

We defined Ω_{tail} as the weighted sum of the two gamma components in Ω_Q:

$$\Omega_{tail}(S_T) = \frac{1}{\lambda_2 + \lambda_3^*} \left(\lambda_2 \, \Omega_2(S_T) + \lambda_3^* \, \Omega_3(S_T) \right) \tag{10}$$

where $\lambda_3^* = 1 - \lambda_2 - \lambda_1^*$. Ω_{tail} thus represents the right-hand tail of the SAS function Ω_Q, allowing us to study the asymptotic

360   behavior of the function in detail. In particular, this asymptotic behavior is time-invariant when plotted against S_T, because Ω_2
and Ω_3 are functions of S_T only. The behavioral parameter sets were thus directly used to calculate the curves $(S_T, \Omega_{tail}(S_T))$.
These curves show similar differences for ^2H and ^3H than the curves $(T, \overleftarrow{P_Q}(T))$ (Fig. 7): a slightly wider spread is observed
for Ω_{tail} constrained by tritium than deuterium (Fig. 7b), and the Ω_{tail} constrained by both tracers tend to converge to a narrow
envelope of curves slightly shifted towards higher storage values (Fig. 7c).

365   To quantitatively study the implications of different Ω_{tail} for storage estimations, we computed statistics of a storage measure
derived from these curves (Table 4). The 95$^{\text{th}}$ percentile of Ω_{tail}, called S_{95P} (black crosses in Fig. 7) allows for estimating
total mobile storage $S(t)$ from Ω_{tail}. In average, the Ω_{tail} constrained by tritium or by both tracers yielded higher mobile
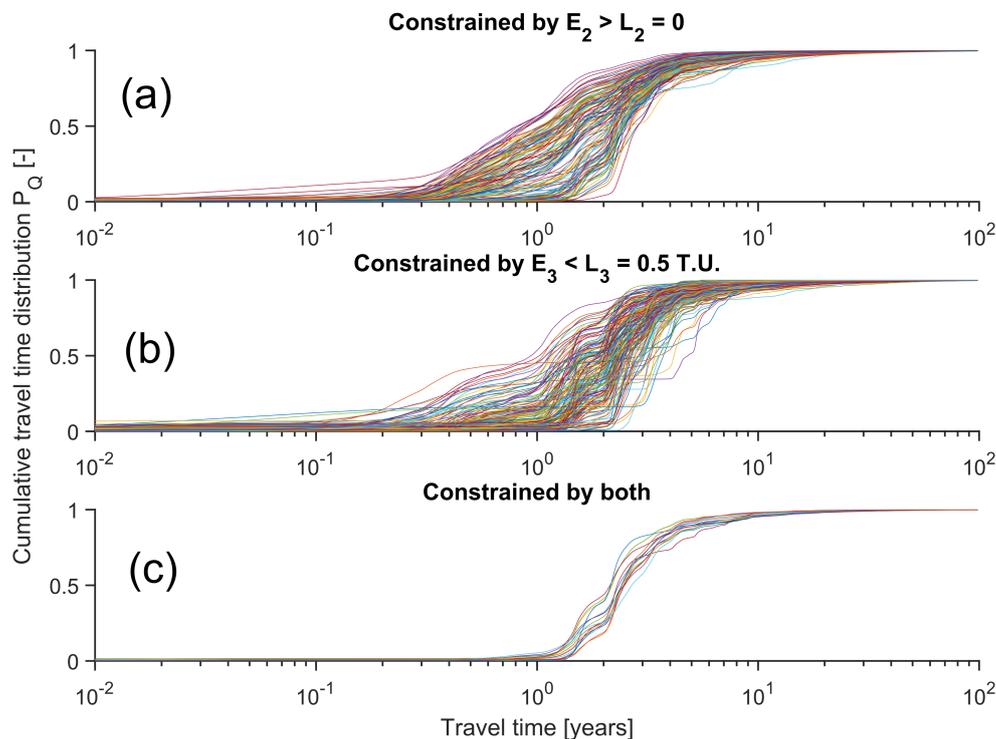
Figure 6. Flow weighted (2015–2017) cumulative stream TTDs for the behavioral parameter sets constrained by [2]H (a), by [3]H (b), and by both (c).

storage $S(t)$ and smaller spread in $S(t)$ (Fig. 7 and Table 4). Overall, the mobile storage $S(t)$ values estimated from the tracers are mutually consistent when considering the uncertainties.

370 **4 Discussion**

4.1 Reconciliation of water ages from stable and radioactive isotopes of H

Our work shows that streamflow TTDs and the related catchment mobile storage $S(t)$ can still be estimated in unsteady conditions by using "ranked" SAS functions $\Omega(S_T, t)$ (Harman, 2015). Similar to Visser et al. (2019), we propose to coherently use the measurements of stream ^{2}H and ^{3}H to calibrate the parameters of the SAS functions, here defined in the age-ranked

375 domain $S_T \in [0, +\infty[$ instead of the cumulative residence time domain $P_S \in [0, 1]$. The calibrated tail of the streamflow SAS function Ω_Q (called here Ω_{tail}) could thus be used to approximate mobile storage $S(t)$ instead of defining the value a priori. The SAS functions also allowed us to estimate the unsteady TTDs defined in the age domain T and their statistics (mean, median, etc.). Differences between the various statistics of the TTDs were smaller than the uncertainties of the calculations when comparing the results obtained with ^{2}H alone and with ^{3}H alone. Similarly, the derived storage estimates were consistent
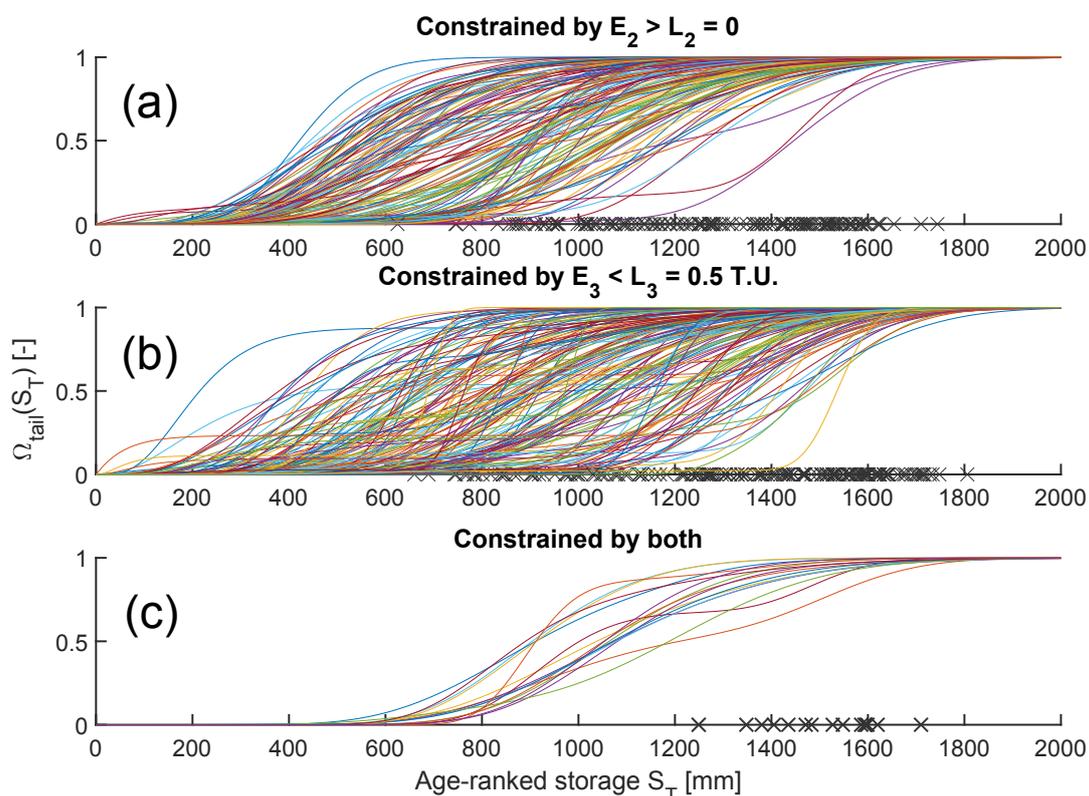
Figure 7. Cumulative right-hand tail Ω_{tail} of streamflow SAS functions for the behavioral parameter sets constrained by ^{2}H (a), by ^{3}H (b), and by both (c). Ω_{tail} is defined as the weighted sum of the two gamma components in Ω_Q. The black crosses indicate S_{95P} for each curve, i.e. the 95th percentile of Ω_{tail}

380   between ^{2}H and ^{3}H. The hypothesis of truncation of the TTD tails when using stable isotopes is therefore rejected at the present time based on the data from the Weierbach catchment. Moreover, stable isotopes do not seem to underestimate the related catchment storage compared to tritium.

These findings were made possible for a number of reasons. First, we treated ^{2}H and ^{3}H equally by calculating TTDs using a coherent mathematical framework for both tracers (i.e. same method and same functional form of TTD). Even though we
385   clearly distinguished tritium from deuterium by accounting for the relationship between water ages and tritium activities $C_{Q,3}$ (term $\exp(-\alpha\,T)$ in Eq. (3)), we did not use it directly to calculate T to avoid biases due to mixing of various ages at the outlet (Bethke and Johnson, 2008) and to avoid the age ambiguity caused by tritium from nuclear tests (Stewart et al., 2012). Also, we did not use multiple control volumes having different TTDs determined by tracer measurements in their input and output (Małoszewski et al., 1983; Uhlenbrook et al., 2002; Stewart et al., 2007; Stewart and Thomas, 2008). This way, we

390   avoided adding large uncertainties related to difficulties in characterizing end members and gathering representative samples (Delsman et al., 2013). Second, we explicitly accounted for unsteady conditions, which has been done in only one other study using tritium (Visser et al., 2019). This allowed us to estimate realistic average TTDs corresponding to the catchment inflows, outflows, and internal flows that are highly time variant. Third, our tritium stream sampling was not focused solely on baseflow hence not biased towards old water. Fourth, we considered the entire TTDs by using various percentiles and statistics, and

395   not only the MTT which is highly influenced by the improbable extreme values of T. This means that even if there is water older than e.g. 1,000 years in streamflow, it can be neglected if it represents less than e.g. 0.000001 % of the volume. Finally, we explicitly accounted for parameter uncertainty. This is important because absolute values without an uncertainty estimate cannot be reliably interpreted.

4.2   Yet, tritium seems to reveal older water!

400   Even though the uncertainties are sufficient to account for the differences between [2]H- and ^{3}H-derived age and storage measures, it is worth noticing that ^{3}H systematically gave higher estimates (tables 3 and 4). The hypothesis of different transport velocities between water molecules containing deuterium and water molecules containing tritium can be rejected, because their self diffusion in water are equal (Devell, 1962), and their advective velocities are the same. On the one hand, the slightly higher ages derived from tritium seem related to the apparent absence of responses of stream ^{3}H to the high precipitation ^{3}H in 2016,

405   indicative of the dominance of old water compared to 2017. On the other hand, tritium simulations included many small peaks corresponding to flashy streamflow responses associated with young water (Fig. 5). Only some of those simulated peaks could be confirmed by the presence of stream measurements at those times, especially in 2016. More stream ^{3}H samples during these flashy events would probably support even further these simulations of young water in streamflow and shift the TTDs constrained by tritium towards younger water. We thus interpret the observed small age differences rather as the consequence

410   of a limited tritium sampling resolution (bi-weekly) that may still be biased towards hydrological recessions during which the youngest water fractions are absent by definition. Tritium and stable isotopes of O and H sampled synchronously at high resolution would thus pave the way for further research on stream water ages from a multi-tracer perspective.

It is also interesting that the age and storage measures estimated from a joint use of ^{2}H and ^{3}H are the highest (tables 3 and 4). In the end, tritium may have helped revealing the presence of old water in streamflow. However, it did so only when combined

415   with deuterium. It is commonly assumed that ^{3}H is more informative about old water because of radioactive decay that relates lower tritium activities to increasing water ages (Stewart et al., 2010). However, as shown by Stewart et al. (2012) and in Fig. 2, current tritium values of the water recharged in 1980–2000 are similar to the tritium values of the water recharged today. Thus, the younger water disrupts the relationship between water age and tritium values. Adding supplementary information about the younger water in the calibration with the high frequency ^{2}H measurements may have partly helped "filtering" the currently

420   complex relationship between water ages and tritium values, leveraging the potential of tritium for revealing the tail of the TTDs. The fact that water ages in the Weierbach are limited to about 5 years (Table 3) could be another reason for the limited

information of ^{3}H about older water. ^{3}H decays by only about 25 % in 5 years, meaning that all the tritium activities of the water in the Weierbach have varied by at most $\sim$2 T.U. since water entered the catchment. This is much lower than the 10 T.U. amplitude of tritium variations in precipitation. Thus in catchments with limited residence times, radioactive decay may only

425 give information that is redundant with the natural variability of the tracer in precipitation. In a few decades, water recharged in 1980–2000 may have completely left the catchments or may be a negligible part of storage, such that the $\log(^3H)$ of stored water may increase linearly with water age (see the recent increasing trend in $C_{P,3}^*$ in Fig. 2). Thus in a few decades, tritium could be even more informative about old water contributions because there may be no age ambiguity anymore. Furthermore, the oscillations of tritium in precipitation over long time scales (> 10 years) recently detected and related to cycles of solar

430 magnetic activity (Palcsu et al., 2018) may give stream tritium concentrations even more age-specific meaning. Therefore it is important to re-iterate the call of Stewart et al. (2012) to start sampling tritium in streams now and for the next decades to use it in travel time analyses.

4.3   Age information contents of stable and radioactive isotopes

The fact that we found equal travel time and storage measures when using ^{2}H alone or ^{3}H alone does not mean that it is not

435 worth sampling both. Our results show that more information was learned about storage and travel times (all the $D_{KL} > 0$) by using both tracers together, which resulted in lower uncertainties (lowest entropy $\mathcal{H}(X\,|\,(^2H \cap {}^3H))$ in Table 2, narrower groups of curves in Fig. 6 and 7, lower standard deviations in tables 3 and 4). This was possible because the composite SAS functions (Eq. (5)) allowed us to independently constrain different parts of the same streamflow TTD with one tracer or the other, reducing the potential trade-offs between the shapes suggested by one tracer or the other. In addition, the streamflow

440 TTD was constrained using only stream samples. On the contrary, Stewart et al. (Table I, 2010) showed three studies where multiple TTDs corresponding to different end members (e.g. surface runoff, groundwater) are constrained by tracers sampled in the associated outlets. Although reasonable fits were shown for the samples from the different end members, the fit of the combined TTD for the stream samples was not systematically checked (Uhlenbrook et al., 2002; Stewart et al., 2007; Stewart and Thomas, 2008).

445   For future studies it is worth mentioning the amount of information gained per isotopic sample or per euro invested in sample analysis. This amount of information can be calculated for a given tracer by summing for all parameters the Kullback-Leibler divergences D_{KL} (see Sect. 2.7) between the prior and the posterior parameter distributions. With deuterium, we learned 13.55 bits of information with 1385 samples, representing about 9.79×10^{-3} bits per sample or about 9.79×10^{-4} bits per euro. With tritium, we learned 14.85 bits of information with only 24 samples. We thus have a much higher relative information content

450 of 0.619 bits per sample. However tritium analyses are more expensive, so the information content is only 1.44×10^{-3} bits per euro. It should be noted that for tritium the precipitation samples were not included in this cost as they were analyzed by the IAEA. Thus tritium was overall more informative than deuterium about water ages, and it was also more cost-effective. One reason for this is that tritium considerably informed us about the travel times in ET because it constrained the posterior of μ_{ET} well (Table 2) that controls directly the ages in ET. This also highlights the importance of considering explicitly ET in

455 streamflow travel time calculations (van der Velde et al., 2015; Visser et al., 2019) However, ^{2}H resulted in lower uncertainties for nearly all other parameters (e.g. lower Shannon entropy $\mathcal{H}(X|^2H)$, Table 2). This is most likely due to the much higher sampling frequency for deuterium that allows for constraining the simulations better than with bi-weekly tritium measurements (see the simulation envelopes Fig. 4 and 5). From our experience in the Weierbach catchment, we estimate that for ^{2}H, a weekly sampling to cover the damped variations of δ^2H (i.e. about 100 samples over 2015–2017) complemented with an event-based

460 high-frequency sampling (every 15 hours) of the flashy responses (i.e. about 300 samples over 2015–2017) could have given us as much information as the complete time series. This suggests that a more strategic sampling of ^{2}H could outperform ^{3}H in terms of cost-efficiency. The amount of information learned from the isotopic data probably scales non-linearly and probably reaches a plateau as the number of observation points grows. In the future, it would be useful to further use information theory (e.g. entropy conditional on sample size) to know how many measurements are enough and when to sample isotopes for

465 maximum information gain on water ages. This would imply artificially re-sampling a higher-frequency isotopic time series using various strategies (e.g. Pool et al., 2017; Etter et al., 2018) and re-calibrating the model many times, which would come with an exorbitant computational price.

  In the end, stable and radioactive isotopes of H have different information contents. For example, they lead to different Shannon entropy $\mathcal{H}$ for the posteriors. Also, the Kullback-Leibler divergence D_{KL} was never 0, indicating that adding one

470 tracer after the other still allowed us to learn something about parameter values. Finally, the parameters corresponding to the best simulations in ^{2}H did not correspond to those for ^{3}H and vice versa. Our results suggest that ^{3}H is more informative about old water thanks to its radioactive decay. Yet, stable and radioactive isotopes have information in common about young water. For example, both ^{2}H and ^{3}H stream samples showed reactions to precipitation ^{2}H and ^{3}H values during flashy streamflow events, revealing the role of young water during these events. This was previously unobserved for tritium because of the

475 sampling focused on periods outside events (Stewart et al., 2010). The theoretical span of 0–4 years pointed out in Stewart et al. (2010) should however not be taken as the only range of ages where ^{18}O, ^{2}H, and ^{3}H have redundant information. As clearly written in Stewart et al. (2010), this limit corresponds to a steady-state exponential TTD only, while other TTD shapes (or unsteady TTDs) could yield much higher limits. More importantly, this limit can be lowered by the seasonality of the input function (see Stewart et al., 2010, p. 1647).

480 **4.4 Limitations and way forward**

The storage value derived from unsteady travel times constrained by tracer data (Table 4, ∼1200–1700 mm) is noticeably larger than the maximum storage ($\simeq$ 250 mm) estimated from point measurements of porosity and water content (Martínez-Carreras et al., 2016), from water balance analyses (Pfister et al., 2017) combined with recession techniques (Carrer et al., 2019), and from the values used in a distributed hydrological model ($\leq$ 700 mm, Glaser et al., 2016). Our storage value is more

485 consistent with the ∼1600 mm derived from depth to bedrock and porosity data used for the Colpach catchment (containing the Weierbach) that was modeled with CATFLOW (Loritz et al., 2017). Large differences between hydrometrically-derived and tracer-derived storage estimates are not uncommon (Soulsby et al., 2009; Fenicia et al., 2010; Birkel et al., 2011) and in

fact highlight the ability of tracers to reveal the existence of stored water that is not directly involved in streamflow generation (Dralle et al., 2018; Carrer et al., 2019). This "hydraulically disconnected" storage is nevertheless important to explain the

490 long residence times in catchments. More research is thus needed for improving the conceptualization of storage and unifying storage terminology and the various estimates obtained from tracers or other techniques. The storage value we found is not in complete contradiction with the previous estimates if we consider their uncertainties. Hydrological measurements (J, Q, and especially ET) are highly uncertain (Waichler et al., 2005; Graham et al., 2010; Buttafuoco et al., 2010; McMillan et al., 2012; McMahon et al., 2013) and their errors are accumulated in long term water balance calculations. An explicit consideration

495 of those uncertainties in the future could reconcile the different storage estimates. Furthermore, it is worth remembering that simplifying storage from a complex spatially-distributed quantity to a simple compact 1D water column neglects the importance of subsurface heterogeneity, surface topography, and bedrock topography for the storage and release of water. As a result, upscaling local point measurements of storage capacity that are not representative of the whole subsurface is very likely to under or overestimate the true storage capacity of the whole catchment. This is even more true if the new techniques

500 used to scan the subsurface over larger areas such as Electrical Resistivity Tomography (ERT) are themselves associated with uncertainties, requiring adaptations (Gourdol et al., 2018) and site-specific independent knowledge (Parsekian et al., 2015).

Our conclusions rest on the assumption that the model captures the water ages in the Weierbach accurately, which was validated by the acceptable performance of the simulations. Still, the performance in δ^2H or in ^{3}H could be improved in the future by testing other models of composite SAS functions. The best NSE for deuterium simulations (called E_2) was 0.24,

505 which is lower than the values reported in a number of studies using SAS functions (van der Velde et al., 2015; Harman, 2015; Benettin et al., 2017b). It should be pointed out again the NSE may not be the most appropriate objective function to characterize performance against the δ^2H time series from the Weierbach (see Rodriguez and Klaus, in review). Future work could look for more appropriate objective functions for δ^2H, especially with respect to the information gained from model calibration. The best MAE for tritium simulations (called E_3) was 0.24. This is slightly higher than values of RMSE (close

510 to 0.10) reported in a number of studies using tritium (Stewart et al., 2007; Stewart and Thomas, 2008; Duvert et al., 2016). However these studies had only a few stream samples, while Gusyev et al. (2013) report for instance a RMSE of 1.62 T.U. for 15 stream samples. Stream δ^2H seems to suggest larger fraction of young water than the simulations (c.f. underestimation of many flashy events in Fig. 4). Stream ^{3}H data seems to suggest larger fractions of old water than the simulations (c.f. overestimation of tritium activities over March–September 2016 in Fig. 5). A model passing through all observation points

515 may thus show larger differences between the TTDs constrained by deuterium and the TTDs constrained by tritium. However, there are not enough ^{3}H stream samples compared to ^{2}H, so a comparison of the TTDs from this hypothetical ideal model could be misleading.

The simulations in deuterium were better for decreasing δ^2H than for increasing δ^2H (better simulations of the flashy events in δ^2H pointing downwards, Fig. 4). This is probably because the increases in δ^2H generally correspond to drier periods,

520 during which $C_{Q,2}$ starts reacting stronger to $C_{P,2}$ indicating that young water fractions (controlled by $\lambda_1(t)$ in the model)

are higher than expected. During drier periods, we expect an increase in the non-linearity of the processes delivering young water to the stream. For example, the decreasing extent of the stream network and of saturated areas observed in the Weierbach during drier conditions (Antonelli et al., in reviewa, i) is caused by decreasing groundwater levels (Glaser et al., 2019a) and it could reduce the amounts of young water reaching the stream (c.f. van Meerveld et al., 2019). However, streamflow is lower

525 during drier conditions, so the fractions of young water can still increase because of a less pronounced dilution of the young water in streamflow compared to wet periods. On the other hand, preferential flow observed in the soils of the Weierbach catchment and in the direct vicinity (Jackisch et al., 2017; Angermann et al., 2017; Scaini et al., 2017, 2018) may become more relevant during drier conditions and could increase the amount of young water contributing to streamflow, especially because precipitation intensities can be much higher in summer (due to thunderstorms) than in winter. The parameterization

530 of the streamflow SAS functions via $\lambda_1(t)$ (Eq. (A5)) includes—to some extent—the effect of wet vs. dry conditions and the role of precipitation intensity, but it seems not to fully capture how these factors influence young water fractions in the stream. Testing other parameterizations of $\lambda_1(t)$ or including other information such as soil moisture or groundwater levels in the current parameterization of $\lambda_1(t)$ may improve the simulations. Finally, the uncertainty of precipitation $\delta^2 H$ could be higher during drier periods, because precipitation amounts can be too small (e.g. < 1 mm) over several weeks or because the

535 precipitation intensities can be too high (e.g. > 5 mm/h) to be captured efficiently by the sequential rainfall sampler. This may lead to inaccuracies in the input data thus to the inability of the model to simulate the corresponding flashy events in stream $\delta^2 H$. The representation of precipitation $\delta^2 H$ could thus be improved in the future by using more recent sampling techniques (e.g. Michelsen et al., 2019).

The simulations overestimated ^3H in the stream in 2015–2016 compared to 2017 (Fig. 5). In 2017 the simulations were better

540 because the model used more of the young water (< 7 days old, using Ω_1) to simulate the variability and the higher values of stream ^3H than in 2016. The lower ^3H in 2015–2016 could be caused by an increased age in the older water components in 2015–2016 compared to 2017, due to changes in the importance of different subsurface flow paths in the Weierbach caused by a wetter period. The old water components Ω_2 and Ω_3 (Eq. (5)) represent subsurface flows likely occurring in the lower soils and following bedrock topography (Glaser et al., 2016; Rodriguez and Klaus, in review) and in weathered bedrock fractures

545 (Scaini et al., 2018). We used functions of S_T only for these components, meaning that the ranges of ages they select do not change considerably with time (because the distribution of S_T is rather stable). Including explicitly a dependence on time for Ω_2 and Ω_3 could help to better represent e.g. the fracture flows or deep groundwater flows in the catchment and improve ^3H simulations in 2015–2016. Eventually, the monthly resolution of ^3H in precipitation is coarser than the biweekly sampling in the stream, which can hinder accurate simulations. An increase in sampling resolution of tritium in the stream to better

550 constrain the TTDs in the future will need to be followed by a considerable increase of sampling resolution in precipitation (Rank and Papesch, 2005).

Although we found much lower deviations for the age and storage measures constrained by deuterium and tritium together (tables 3 and 4), it has to be acknowledged that this is also because there are only few accepted solutions (16), while there about

10 times more when using ^{2}H alone or ^{3}H alone. Yet, the associated curves (Fig. 6 and 7) fall close to each other, so the lower

555 deviations have to be due also to lower uncertainties. A lower number of accepted solutions is in the end inevitable as it is an inherent consequence of using several performance measures independently as opposed to using a combined objective function (e.g. Hrachowitz et al., 2013; Rodriguez et al., 2018). Less strict threshold criteria for behavioral solutions could increase the number of accepted solutions but they would accept less accurate simulations, which could lead to misleading conclusions. More stream ^{3}H measurements would on the other hand allow the use of more advanced objective functions, which could lead

560 to more accepted solutions. Eventually, the input data measured over 2010–2017 and used to spin up the model from 1960 to 2010 (J, ET, Q, and $C_{P,2}$) could be unrepresentative of the real hydrometeorological and isotopic conditions of 1960–2015 due for instance to nonstationarity or climate change. These changing conditions could affect the modeled residence times in storage and thus the estimated streamflow travel times (Wilusz et al., 2017). Different methods to spin up the model could be tested in the future (Hrachowitz et al., 2011), especially to assess the effect the effect of changing hydrometeorological and

565 isotopic conditions on the estimation of travel times. For this, isotope tracer records that span several decades like the ones that can be reconstructed from pearl mussels shells (Pfister et al., 2018, 2019) represent a crucial asset.

5 Conclusions

Stable isotopes of O and H and tritium are indispensable tracers to infer the streamflow TTD and derive storage estimates in catchments. Our study addressed an emerging concern about the possible deficiency of stable isotopes to infer the whole

570 streamflow TTD. We went beyond previous data and methodological limitations and thus we did not find that stable isotopes are blind to old water fractions compared to tritium in our experimental catchment in Luxembourg. However, we found that stable isotopes and tritium do have different information contents on water ages. In fact, inferring the streamflow TTD from a joint use of both tracers better exploits their respective age information contents, which results in lower uncertainties. Even if ^{3}H appeared to be slightly more cost-effective and informative than ^{2}H, a smart sampling of the stable isotopes could outperform

575 tritium. Future work could compare streamflow TTD and storage from the two tracers in larger catchments where older water is expected, to give tritium more time to decay and better leverage its ability to point the presence of very old water out. We therefore recommend to: (1) keep sampling tritium in as many places as possible, as emphasized by Stewart et al. (2012); but also (2) to sample tritium at the highest frequency possible and synchronously with stable isotopes if possible. This is particularly important for the isotopic measurements in precipitation that drive all model simulations, regardless of functional

580 forms of TTD and their parameter values. Overall this work shows that more tracer data is naturally better to gather more information about the catchments functions of storage and release.

Data availability. The tritium input data until 2016 used in this study can be obtained from the WISER database portal of the International Atomic Energy Agency (values for 2017 will be accessible there too in the future, please ask Axel Schmidt from Bundesanstalt

für Gewässerkunde in the meantime). The rest of the data used in this study is the property of the Luxembourg Institute of Science and
585 Technology (LIST) and can be obtained by request to the corresponding author after approval by LIST.

Appendix A: Model equations

A1   Parameterization of the SAS functions

In this section we provide further details on the equations used in the model. The composite streamflow SAS function Ω_Q used
in this study is:

590 $$\Omega_Q(S_T,t) = \lambda_1(t)\,\Omega_1(S_T) + \lambda_2(t)\,\Omega_3(S_T) + \lambda_3(t)\,\Omega_1(S_T) \tag{A1}$$

$\Omega_1(S_T)$ is a cumulative uniform distribution for S_T in $[0,S_u]$, where S_u (mm) is a calibrated parameter representing the
amount of stored young water potentially contributing to flashy streamflow responses. Thus:

$$\Omega_1(S_T) = \begin{cases} \frac{S_T}{S(t)}, & S_T \in [0,S_u] \\ 1, & S_T > S_u \end{cases} \tag{A2}$$

$\Omega_2(S_T)$ and $\Omega_3(S_T)$ are direct functions of S_T and are gamma-distributed:

595 $$\Omega_2(S_T) = \frac{1}{\Gamma(\frac{\mu_2}{\theta_2})}\,\gamma(\frac{\mu_2}{\theta_2},\frac{S_T}{\theta_2}) \tag{A3}$$

$$\Omega_3(S_T) = \frac{1}{\Gamma(\frac{\mu_3}{\theta_3})}\,\gamma(\frac{\mu_3}{\theta_3},\frac{S_T}{\theta_3}) \tag{A4}$$

where Γ is the gamma function, γ is the lower incomplete gamma function, μ_2 and μ_3 (mm) are mean parameters (calibrated),
and θ_2 and θ_3 (mm) are scale parameters (calibrated).

$\lambda_1(t)$, $\lambda_2(t)$, and $\lambda_3(t)$ sum to 1. These are simply time-varying weights giving each component (i.e. c.d.f. Ω) a dynamic
600 contribution to streamflow generation. In particular, $\lambda_1(t)$ is made highly time-variant to represent the flashy hydrographs
that have an on-off type of response to precipitation. $\lambda_2(t)$ is considered constant and calibrated to keep the parameterization
parsimonious. $\lambda_3(t) = 1 - \lambda_2 - \lambda_1(t)$ is deduced by difference for parsimony as well. Since $\Omega_1(S_T)$ represents young water
contributions and previous studies in the Weierbach showed that event water contributions depend on the catchment wetness
and on precipitation intensity (Wrede et al., 2015; Martínez-Carreras et al., 2015), $\lambda_1(t)$ was parameterized using storage $S(t)$
605 and a proxy storage variations $\overline{\Delta S(t)}$ (see Rodriguez and Klaus (in review) for more details):

$$\lambda_1(t) = \lambda_1^*\,[f(t) + (1 - f(t))\,g(t)] \tag{A5}$$

where $\lambda_1^* \in [0, 1]$ (no units) is a calibrated parameter representing the maximum value of $\lambda_1(t)$, and $f(t) \in [0, 1]$ and $g(t) \in [0, 1]$ are given by:

$$f(t) = f_0 \left(1 - \tanh \left[\left(\frac{S(t)}{S_{min} + S_{th}} \right)^m \right] \right) \tag{A6}$$

610 $$g(t) = 1 - \exp \left(-\frac{\overline{\Delta S(t)}}{\Delta S_{th}} \right) \tag{A7}$$

$f_0 \in [0, 1]$ (no units) is a calibrated parameter guaranteeing a minimum for $\lambda_1(t)$ during dry periods, $S_{min} = \min(S(t))$, and S_{th} (mm, calibrated parameter) is a storage threshold relative the to minimum storage S_{min} separating wet ($S(t) > S_{min} + S_{th}$) from dry periods ($S(t) < S_{min} + S_{th}$). $m = 1000$ is a fixed parameter used to smooth the function f with respect to $S(t)$. $\overline{\Delta S(t)}$ is a proxy of storage variations calculated as a moving average of storage variations over a time window $\Delta t^* = 2\,\Delta t$:

615 $$\overline{\Delta S(t)} = \max \left(\frac{1}{3} \sum_{j=0}^{2} \Delta S(t - j\Delta t), 0 \right) \tag{A8}$$

with $\Delta S(t) = \Delta t \left(J(t) - Q(t) - ET(t) \right)$. $\overline{\Delta S(t)}$ essentially increases during precipitation events and decreases when $Q(t)$ or $ET(t)$ are high. ΔS_{th} is a threshold in $\overline{\Delta S(t)}$ above which $g(t)$ tends to 1, allowing $\lambda_1(t)$ to increase and decrease sharply during flashy streamflow events.

A2 Actual evapotranspiration

620 Actual evapotranspiration $ET(t)$ is calculated from potential evapotranspiration $PET(t)$ using the formula:

$$ET(t) = PET(t) \tanh \left[\left(\frac{S(t)}{S_{root}} \right)^n \right] \tag{A9}$$

where $S_{root} = S_{ref} - 150$ is a fixed parameter (mm) representing the storage threshold $S(t) = S_{root}$ below which $ET(t)$ starts decreasing from $PET(t)$ towards 0. This decrease is smoothed by the fixed coefficient $n = 20$. S_{root} accounts for the water available for evaporation and plant transpiration until the capillary forces offer too much resistance. This formula thus
625 represents the decrease in water losses to the atmosphere under water limited conditions.

Author contributions. LP and JK designed the project and obtained the funding for this study. NR carried out the field and lab work. NR and JK did the modelling part. NR, LP, EZ, and JK jointly structured the manuscript, and NR wrote the manuscript with contributions from JK, EZ, and LP

Competing interests. The authors declare no competing interests

630   *Acknowledgements.* We thank Uwe Morgenstern from GNS Science and Axel Schmidt from Bundesanstalt für Gewässerkunde (BfG) for providing access to the 2017 precipitation tritium data. We thank Laurent Gourdol for his help with the preparation of the tritium input data, and for useful discussions about estimating TTDs with tritium measurements. We thank Uwe Ehret for providing Matlab scripts to compute information theory measures ($\mathcal{H}$ and D_{KL}). Nicolas Rodriguez and Laurent Pfister (FNR/CORE/C14/SR/8353440/STORE-AGE), and Julian Klaus (FNR/CORE/C17/SR/11702136/EFFECT) acknowledge funding for this study from the Luxembourg National

635   Research Fund (FNR). This study also contributes to and benefited from the "Catchments as Organized Systems" (CAOS FOR 1598, IN-TER/DFG/14/9476192/CAOS2) research unit funded by DFG, FNR, FWF. We thank Jérôme Juilleret for his help to collect tritium samples in the field. The authors acknowledge support by the state of Baden-Württemberg through bwHPC. We thank Barbara Glaser for her help with the input data (PET).

References

640 Angermann, L., Jackisch, C., Allroggen, N., Sprenger, M., Zehe, E., Tronicke, J., Weiler, M., and Blume, T.: Form and function in hillslope hydrology: characterization of subsurface flow based on response observations, Hydrology and Earth System Sciences, 21, 3727–3748, https://doi.org/10.5194/hess-21-3727-2017, 2017.

Antonelli, M., Glaser, B., R., T., Klaus, J., and Pfister, L.: Saturated areas through the lens: 1. Spatio-temporal variability of surface saturation documented through Thermal Infrared imagery, Hydrological Processes, in reviewa.

645 Antonelli, M., Glaser, B., R., T., Klaus, J., and Pfister, L.: Saturated areas through the lens: 2. Spatio-temporal variability of streamflow generation and its relationship with surface saturation, Hydrological Processes, in reviewb.

Bajjali, W.: Spatial variability of environmental isotope and chemical content of precipitation in Jordan and evidence of slight change in climate, Applied Water Science, 2, 271–283, https://doi.org/10.1007/s13201-012-0046-1, 2012.

Begemann, F. and Libby, W.: Continental water balance, ground water inventory and storage times, surface ocean mixing rates and
650 world-wide water circulation patterns from cosmic-ray and bomb tritium, Geochimica et Cosmochimica Acta, 12, 277 – 296, https://doi.org/10.1016/0016-7037(57)90040-6, 1957.

Benettin, P. and Bertuzzo, E.: *tran*-SAS v1.0: a numerical model to compute catchment-scale hydrologic transport using StorAge Selection functions, Geoscientific Model Development, 11, 1627–1639, https://doi.org/10.5194/gmd-11-1627-2018, https://www.geosci-model-dev.net/11/1627/2018/, 2018.

655 Benettin, P., Bailey, S. W., Campbell, J. L., Green, M. B., Rinaldo, A., Likens, G. E., McGuire, K. J., and Botter, G.: Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH, USA, Water Resources Research, 51, 9256–9272, https://doi.org/10.1002/2015WR017552, http://doi.wiley.com/10.1002/2015WR017552, 2015a.

Benettin, P., Rinaldo, A., and Botter, G.: Tracking residence times in hydrological systems: forward and backward formulations, Hydrological Processes, https://doi.org/10.1002/hyp.15034, 2015b.

660 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, Hydrological Processes, 31, 2982–2986, https://doi.org/10.1002/hyp.11243, http://doi.wiley.com/10.1002/hyp.11243, 2017a.

Benettin, P., Soulsby, C., Birkel, C., Tetzlaff, D., Botter, G., and Rinaldo, A.: Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments, Water Resources Research, 53, 1864–1878, https://doi.org/10.1002/2016WR020117,
665 http://doi.wiley.com/10.1002/2016WR020117, 2017b.

Berman, E. S. F., Gupta, M., Gabrielli, C., Garland, T., and McDonnell, J. J.: High-frequency field-deployable isotope analyzer for hydrological applications, Water Resources Research, 45, https://doi.org/10.1029/2009WR008265, 2009.

Berry, Z. C., Evaristo, J., Moore, G., Poca, M., Steppe, K., Verrot, L., Asbjornsen, H., Borma, L. S., Bretfeld, M., Hervé-Fernández, P., Seyfried, M., Schwendenmann, L., Sinacore, K., De Wispelaere, L., and McDonnell, J.: The two water worlds hypothesis: Addressing
670 multiple working hypotheses and proposing a way forward, Ecohydrology, 11, e1843, https://doi.org/10.1002/eco.1843, e1843 ECO-16-0180.R2, 2018.

Bethke, C. M. and Johnson, T. M.: Groundwater Age and Groundwater Age Dating, Annual Review of Earth and Planetary Sciences, 36, 121–152, https://doi.org/10.1146/annurev.earth.36.031207.124210, 2008.

Birkel, C., Soulsby, C., and Tetzlaff, D.: Modelling catchment-scale water storage dynamics: reconciling dynamic storage with tracer-inferred
675 passive storage, Hydrological Processes, 25, 3924–3936, https://doi.org/10.1002/hyp.8201, http://doi.wiley.com/10.1002/hyp.8201, 2011.

Birkel, C., Soulsby, C., and Tetzlaff, D.: Conceptual modelling to assess how the interplay of hydrological connectivity, catchment storage and tracer dynamics controls nonstationary water age estimates, Hydrological Processes, https://doi.org/10.1002/hyp.10414, http://dx.doi.org/10.1002/hyp.10414, 2015.

Botter, G., Bertuzzo, E., and Rinaldo, A.: Catchment residence and travel time distributions: The master equation, Geophysical Research
680   Letters, 38, https://doi.org/10.1029/2011GL047666, 2011.

Brooks, J. R., Barnard, H. R., Coulombe, R., and McDonnell, J. J.: Ecohydrologic separation of water between trees and streams in a Mediterranean climate, Nature Geoscience, 3, 100–104, https://doi.org/10.1038/ngeo722, 2010.

Buttafuoco, G., Caloiero, T., and Coscarelli, R.: Spatial uncertainty assessment in modelling reference evapotranspiration at regional scale, Hydrology and Earth System Sciences, 14, 2319–2327, https://doi.org/10.5194/hess-14-2319-2010, 2010.

685   Buttle, J.: Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins, Progress in Physical Geography: Earth and Environment, 18, 16–41, https://doi.org/10.1177/030913339401800102, 1994.

Carrer, G. E., Klaus, J., and Pfister, L.: Assessing the Catchment Storage Function Through a Dual-Storage Concept, Water Resources Research, 55, 476–494, https://doi.org/10.1029/2018WR022856, 2019.

Cartwright, I. and Morgenstern, U.: Contrasting transit times of water from peatlands and eucalypt forests in the Australian Alps determined
690   by tritium: implications for vulnerability and the source of water in upland catchments, Hydrology and Earth System Sciences, 20, 4757–4773, https://doi.org/10.5194/hess-20-4757-2016, 2016.

Delsman, J. R., Essink, G. H. P. O., Beven, K. J., and Stuyfzand, P. J.: Uncertainty estimation of end-member mixing using generalized likelihood uncertainty estimation (GLUE), applied in a lowland catchment, Water Resources Research, 49, 4792–4806, https://doi.org/10.1002/wrcr.20341, 2013.

695   Devell, L.: Measurements of the Self-diffusion of Water in Pure Water, H2O-D2O Mixtures and Solutions of Electrolytes, Acta Chemica Scandinavica, 16, 2177 – 2188, https://doi.org/10.3891/acta.chem.scand.16-2177, 1962.

Dinçer, T., Payne, B. R., Florkowski, T., Martinec, J., and Tongiorgi, E.: Snowmelt runoff from measurements of tritium and oxygen-18, Water Resources Research, 6, 110–124, https://doi.org/10.1029/WR006i001p00110, 1970.

Dralle, D. N., Hahm, W. J., Rempe, D. M., Karst, N. J., Thompson, S. E., and Dietrich, W. E.: Quantification of the seasonal hillslope water
700   storage that does not drive streamflow, Hydrological Processes, 32, 1978–1992, https://doi.org/10.1002/hyp.11627, 2018.

Dubbert, M., Caldeira, M. C., Dubbert, D., and Werner, C.: A pool-weighted perspective on the two-water-worlds hypothesis, New Phytologist, 0, https://doi.org/10.1111/nph.15670, 2019.

Duvert, C., Stewart, M. K., Cendón, D. I., and Raiber, M.: Time series of tritium, stable isotopes and chloride reveal short-term variations in groundwater contribution to a stream, Hydrology and Earth System Sciences, 20, 257–277, https://doi.org/10.5194/hess-20-257-2016,
705   2016.

Eriksson, E.: The Possible Use of Tritium' for Estimating Groundwater Storage, Tellus, 10, 472–478, https://doi.org/10.1111/j.2153-3490.1958.tb02035.x, 1958.

Etter, S., Strobl, B., Seibert, J., and van Meerveld, H. J. I.: Value of uncertain streamflow observations for hydrological modelling, Hydrology and Earth System Sciences, 22, 5243–5257, https://doi.org/10.5194/hess-22-5243-2018, 2018.

710   Fenicia, F., Wrede, S., Kavetski, D., Pfister, L., Hoffmann, L., Savenije, H. H. G., and McDonnell, J. J.: Assessing the impact of mixing assumptions on the estimation of streamwater mean residence time, Hydrological Processes, 24, 1730–1741, https://doi.org/10.1002/hyp.7595, 2010.

Fenicia, F., Kavetski, D., Savenije, H. H. G., Clark, M. P., Schoups, G., Pfister, L., and Freer, J.: Catchment properties, function, and conceptual model representation: is there a correspondence?, Hydrological Processes, 28, 2451–2467, https://doi.org/10.1002/hyp.9726, 2014.

Gallart, F., Roig-Planasdemunt, M., Stewart, M. K., Llorens, P., Morgenstern, U., Stichler, W., Pfister, L., and Latron, J.: A GLUE-based uncertainty assessment framework for tritium-inferred transit time estimations under baseflow conditions, Hydrological Processes, 30, 4741–4760, https://doi.org/10.1002/hyp.10991, 2016.

Glaser, B., Klaus, J., Frei, S., Frentress, J., Pfister, L., and Hopp, L.: On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum, Water Resources Research, 52, 8317–8342, https://doi.org/10.1002/2015WR018414, 2016.

Glaser, B., Antonelli, M., Chini, M., Pfister, L., and Klaus, J.: Technical note: Mapping surface-saturation dynamics with thermal infrared imagery, Hydrology and Earth System Sciences, 22, 5987–6003, https://doi.org/10.5194/hess-22-5987-2018, 2018.

Glaser, B., Antonelli, M., Hopp, L., and Klaus, J.: Intra-catchment variability of surface saturation – insights from longterm observations and simulations, Hydrology and Earth System Sciences Discussions, 2019, 1–22, https://doi.org/10.5194/hess-2019-203, 2019a.

Glaser, B., Jackisch, C., Hopp, L., and Klaus, J.: How meaningful are plot-scale observations and simulations of preferential flow for catchment models?, Vadose Zone Journal, https://doi.org/10.2136/vzj2018.08.0146, 2019b.

Gourdol, L., Clément, R., Juilleret, J., Pfister, L., and Hissler, C.: Large-scale ERT surveys for investigating shallow regolith properties and architecture, Hydrology and Earth System Sciences Discussions, 2018, 1–39, https://doi.org/10.5194/hess-2018-519, 2018.

Graham, C. B., van Verseveld, W., Barnard, H. R., and McDonnell, J. J.: Estimating the deep seepage component of the hillslope and catchment water balance within a measurement uncertainty framework, Hydrological Processes, 24, 3631–3647, https://doi.org/10.1002/hyp.7788, 2010.

Gupta, P., Noone, D., Galewsky, J., Sweeney, C., and Vaughn, B. H.: Demonstration of high-precision continuous measurements of water vapor isotopologues in laboratory and remote field deployments using wavelength-scanned cavity ring-down spectroscopy (WS-CRDS) technology, Rapid Communications in Mass Spectrometry, 23, 2534–2542, https://doi.org/10.1002/rcm.4100, 2009.

Gusyev, M. A., Toews, M., Morgenstern, U., Stewart, M., White, P., Daughney, C., and Hadfield, J.: Calibration of a transient transport model to tritium data in streams and simulation of groundwater ages in the western Lake Taupo catchment, New Zealand, Hydrology and Earth System Sciences, 17, 1217–1227, https://doi.org/10.5194/hess-17-1217-2013, 2013.

Halder, J., Terzer, S., Wassenaar, L. I., Araguás-Araguás, L. J., and Aggarwal, P. K.: The Global Network of Isotopes in Rivers (GNIR): integration of water isotopes in watershed observation and riverine research, Hydrology and Earth System Sciences, 19, 3419–3431, https://doi.org/10.5194/hess-19-3419-2015, 2015.

Harman, C. J.: Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, Water Resources Research, 51, 1–30, https://doi.org/10.1002/2014WR015707, http://dx.doi.org/10.1002/2014WR015707, 2015.

Heidbuechel, I., Troch, P. A., Lyon, S. W., and Weiler, M.: The master transit time distribution of variable flow systems, Water Resources Research, 48, https://doi.org/10.1029/2011WR011293, 2012.

Helton, J. and Davis, F.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliability Engineering & System Safety, 81, 23 – 69, https://doi.org/10.1016/S0951-8320(03)00058-9, 2003.

Herbstritt, B., Gralher, B., and Weiler, M.: Continuous in situ measurements of stable isotopes in liquid water, Water Resources Research, 48, https://doi.org/10.1029/2011WR011369, 2012.

Herbstritt, B., Gralher, B., and Weiler, M.: Continuous, near-real-time observations of water stable isotope ratios during rainfall and through-fall events, Hydrology and Earth System Sciences, 23, 3007–3019, https://doi.org/10.5194/hess-23-3007-2019, 2019.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., and Malcolm, I. A.: Sensitivity of mean transit time estimates to model conditioning and data availability, Hydrological Processes, 25, 980–990, https://doi.org/10.1002/hyp.7922, 2011.

Hrachowitz, M., Benettin, P., Breukelen, B. M. V., Fovet, O., Howden, N. J. K., Ruiz, L., Velde, Y. V. D., and Wade, A. J.: Transit times — the link between hydrology and water quality at the catchment scale, Wiley Interdisciplinary Reviews: Water, https://doi.org/10.1002/wat2.1155, 2016.

Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., and Soulsby, C.: What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, Hydrology and Earth System Sciences, 17, 533–564, https://doi.org/10.5194/hess-17-533-2013, 2013.

Hubert, P.: Etude par le tritium de la dynamique des eaux du Lac Léman. Apport du tritium à la limnologie physique, Ph.D. thesis, Hydrologie. Université de Paris, 1971.

IAEA: Global Network of Isotopes in Rivers. The GNIR Database, https://nucleus.iaea.org/wiser, 2019.

IAEA and WMO: Global Network of Isotopes in Precipitation. The GNIP Database, https://nucleus.iaea.org/wiser, 2019.

Jackisch, C., Angermann, L., Allroggen, N., Sprenger, M., Blume, T., Tronicke, J., and Zehe, E.: Form and function in hillslope hy-drology: in situ imaging and characterization of flow-relevant structures, Hydrology and Earth System Sciences, 21, 3749–3775, https://doi.org/10.5194/hess-21-3749-2017, https://www.hydrol-earth-syst-sci.net/21/3749/2017/, 2017.

Juilleret, J., Iffly, J., Pfister, L., and Hissler, C.: Remarkable Pleistocene periglacial slope deposits in Luxembourg (Oesling): pedological implication and geosite potential, Bulletin de la Société des naturalistes luxembourgeois, 112, 125–130, 2011.

Juilleret, J., Dondeyne, S., Vancampenhout, K., Deckers, J., and Hissler, C.: Mind the gap: A classification system for integrating the subsolum into soil surveys, Geoderma, 264, 332 – 339, https://doi.org/10.1016/j.geoderma.2015.08.031, soil mapping, classification, and modelling: history and future directions, 2016.

Keim, R. F., Kendall, C., and Jefferson, A.: The Expanding Utility of Laser Spectroscopy, Eos, Transactions American Geophysical Union, 95, 144–144, https://doi.org/10.1002/2014EO170007, 2014.

Kendall, C. and McDonnell, J. J.: Isotope tracers in catchment hydrology, Elsevier, Amsterdam, https://doi.org/10.1016/B978-0-444-81546-0.50001-X, 1998.

Kirchner, J. W.: A double paradox in catchment hydrology and geochemistry, Hydrological Processes, 17, 871–874, https://doi.org/10.1002/hyp.5108, http://doi.wiley.com/10.1002/hyp.5108, 2003.

Kirchner, J. W.: Aggregation in environmental systems - Part 1: Seasonal tracer cycles quantify young water fractions, but not mean transit times, in spatially heterogeneous catchments, Hydrology and Earth System Sciences, 20, 279–297, https://doi.org/10.5194/hess-20-279-2016, 2016.

Klaus, J. and McDonnell, J.: Hydrograph separation using stable isotopes: Review and evaluation, Journal of Hydrology, 505, 47 – 64, https://doi.org/10.1016/j.jhydrol.2013.09.006, 2013.

Klaus, J., Chun, K. P., McGuire, K. J., and McDonnell, J. J.: Temporal dynamics of catchment transit times from stable isotope data, Water Resources Research, 51, 4208–4223, https://doi.org/10.1002/2014WR016247, 2015a.

Koehler, G. and Wassenaar, L. I.: Realtime Stable Isotope Monitoring of Natural Waters by Parallel-Flow Laser Spectroscopy, Analytical Chemistry, 83, 913–919, https://doi.org/10.1021/ac102584q, pMID: 21214188, 2011.

Lis, G., Wassenaar, L. I., and Hendry, M. J.: High-Precision Laser Spectroscopy D/H and 18O/16O Measurements of Microliter Natural Water Samples, Analytical Chemistry, 80, 287–293, https://doi.org/10.1021/ac701716q, pMID: 18031060, 2008.

Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J., and Zehe, E.: Picturing and modeling catchments by representative hillslopes, Hydrology and Earth System Sciences, 21, 1225–1249, https://doi.org/10.5194/hess-21-1225-2017, 2017.

Loritz, R., Gupta, H., Jackisch, C., Westhoff, M., Kleidon, A., Ehret, U., and Zehe, E.: On the dynamic nature of hydrological similarity, Hydrology and Earth System Sciences, 22, 3663–3684, https://doi.org/10.5194/hess-22-3663-2018, 2018.

Loritz, R., Kleidon, A., Jackisch, C., Westhoff, M., Ehret, U., Gupta, H., and Zehe, E.: A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation, Hydrology and Earth System Sciences, 23, 3807–3821, https://doi.org/10.5194/hess-23-3807-2019, 2019.

Maher, K.: The role of fluid residence time and topographic scales in determining chemical fluxes from landscapes, Earth and Planetary Science Letters, 312, 48–58, https://doi.org/10.1016/j.epsl.2011.09.040, http://linkinghub.elsevier.com/retrieve/pii/S0012821X11005607, 2011.

Małoszewski, P. and Zuber, A.: Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability, Journal of Hydrology, 57, 207 – 231, https://doi.org/10.1016/0022-1694(82)90147-0, 1982.

Małoszewski, P., Rauert, W., Stichler, W., and Herrmann, A.: Application of flow models in an alpine catchment area using tritium and deuterium data, Journal of Hydrology, 66, 319 – 330, https://doi.org/10.1016/0022-1694(83)90193-2, 1983.

Martinec, J.: Subsurface flow from snowmelt traced by tritium, Water Resources Research, 11, 496–498, https://doi.org/10.1029/WR011i003p00496, 1975.

Martínez-Carreras, N., Wetzel, C. E., Frentress, J., Ector, L., McDonnell, J. J., Hoffmann, L., and Pfister, L.: Hydrological connectivity inferred from diatom transport through the riparian-stream system, Hydrology and Earth System Sciences, 19, 3133–3151, https://doi.org/10.5194/hess-19-3133-2015, 2015.

Martínez-Carreras, N., Hissler, C., Gourdol, L., Klaus, J., Juilleret, J., Iffly, J. F., and Pfister, L.: Storage controls on the generation of double peak hydrographs in a forested headwater catchment, Journal of Hydrology, 543, 255 – 269, https://doi.org/10.1016/j.jhydrol.2016.10.004, 2016.

McCutcheon, R. J., McNamara, J. P., Kohn, M. J., and Evans, S. L.: An evaluation of the ecohydrological separation hypothesis in a semiarid catchment, Hydrological Processes, 31, 783–799, https://doi.org/10.1002/hyp.11052, 2017.

McDonnell, J. J.: The two water worlds hypothesis: ecohydrological separation of water between streams and trees?, Wiley Interdisciplinary Reviews: Water, 1, 323–329, https://doi.org/10.1002/wat2.1027, 2014.

McDonnell, J. J. and Beven, K. J.: Debates on Water Resources: The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph, Water Resources Research, 50, 5342–5350, https://doi.org/10.1002/2013WR015141, 2014.

McGuire, K. J. and McDonnell, J. J.: A review and evaluation of catchment transit time modeling, Journal of Hydrology, 330, 543–563, https://doi.org/10.1016/j.jhydrol.2006.04.020, 2006.

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., and Seibert, J.: The role of topography on catchment-scale water residence time, Water Resources Research, 41, https://doi.org/10.1029/2004WR003657, 2005.

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R.: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, Hydrology and Earth System Sciences, 17, 1331–1363, https://doi.org/10.5194/hess-17-1331-2013, 2013.

McMillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality, Hydrological Processes, 26, 4078–4111, https://doi.org/10.1002/hyp.9384, 2012.

Michelsen, N., Laube, G., Friesen, J., Weise, S. M., Bait Said, A. B. A., and Müller, T.: Technical note: A microcontroller-based automatic rain sampler for stable isotope studies, Hydrology and Earth System Sciences, 23, 2637–2645, https://doi.org/10.5194/hess-23-2637-2019, 2019.

Moragues-Quiroga, C., Juilleret, J., Gourdol, L., Pelt, E., Perrone, T., Aubert, A., Morvan, G., Chabaux, F., Legout, A., Stille, P., and Hissler, C.: Genesis and evolution of regoliths: Evidence from trace and major elements and Sr-Nd-Pb-U isotopes, CATENA, 149, 185 – 198, https://doi.org/10.1016/j.catena.2016.09.015, 2017.

Morgenstern, U. and Taylor, C. B.: Ultra low-level tritium measurement using electrolytic enrichment and LSC, Isotopes in Environmental and Health Studies, 45, 96–117, https://doi.org/10.1080/10256010902931194, pMID: 20183224, 2009.

Munksgaard, N. C., Wurster, C. M., and Bird, M. I.: Continuous analysis of $\delta18O$ and δD values of water by diffusion sampling cavity ring-down spectrometry: a novel sampling device for unattended field monitoring of precipitation, ground and surface waters, Rapid Communications in Mass Spectrometry, 25, 3706–3712, https://doi.org/10.1002/rcm.5282, 2011.

Palcsu, L., Morgenstern, U., Sültenfuss, J., Koltai, G., László, E., Temovski, M., Major, Z., Nagy, J. T., Papp, L., Varlam, C., Faurescu, I., Túri, M., Rinyu, L., Czuppon, G., Bottyán, E., and Jull, A. J. T.: Modulation of Cosmogenic Tritium in Meteoric Precipitation by the 11-year Cycle of Solar Magnetic Field Activity, Scientific Reports, 8, 12 813, https://doi.org/10.1038/s41598-018-31208-9, 2018.

Pangle, L. A., Klaus, J., Berman, E. S. F., Gupta, M., and McDonnell, J. J.: A new multisource and high-frequency approach to measuring $\delta2H$ and $\delta18O$ in hydrological field studies, Water Resources Research, 49, 7797–7803, https://doi.org/10.1002/2013WR013743, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013WR013743, 2013.

Parsekian, A. D., Singha, K., Minsley, B. J., Holbrook, W. S., and Slater, L.: Multiscale geophysical imaging of the critical zone, Reviews of Geophysics, 53, 1–26, https://doi.org/10.1002/2014RG000465, 2015.

Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., and McDonnell, J. J.: Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments, Hydrological Processes, 31, 1828–1845, https://doi.org/10.1002/hyp.11134, 2017.

Pfister, L., Thielen, F., Deloule, E., Valle, N., Lentzen, E., Grave, C., Beisel, J.-N., and McDonnell, J. J.: Freshwater pearl mussels as a stream water stable isotope recorder, Ecohydrology, 11, e2007, https://doi.org/10.1002/eco.2007, 2018.

Pfister, L., Grave, C., Beisel, J.-N., and McDonnell, J. J.: A global assessment of freshwater mollusk shell oxygen isotope signatures and their relation to precipitation and stream water, Scientific Reports, 9, 4312, https://doi.org/10.1038/s41598-019-40369-0, 2019.

Pool, S., Viviroli, D., and Seibert, J.: Prediction of hydrographs and flow-duration curves in almost ungauged catchments: Which runoff measurements are most informative for model calibration?, Journal of Hydrology, 554, 613 – 622, https://doi.org/10.1016/j.jhydrol.2017.09.037, 2017.

Rank, D. and Papesch, W.: Isotopic composition of precipitation in Austria in relation to air circulation patterns and climate, chap. 2, pp. 19–35, International Atomic Energy Agency (IAEA), 2005.

Rank, D., Wyhlidal, S., Schott, K., Weigand, S., and Oblin, A.: Temporal and spatial distribution of isotopes in river water in Central Europe: 50 years experience with the Austrian network of isotopes in rivers, Isotopes in Environmental and Health Studies, 54, 115–136, https://doi.org/10.1080/10256016.2017.1383906, pMID: 29082751, 2018.

Różański, K., Froehlich, K., and Mook, W. G.: Environmental Isotopes in the Hydrological Cycle, Principles and Applications. VOLUME III: Surface water, IAEA and UNESCO, 2001.

Rinaldo, A. and Marani, A.: Basin scale-model of solute transport, Water Resources Research, 23, 2107–2118, https://doi.org/10.1029/WR023i011p02107, 1987.

Rinaldo, A., Benettin, P., Harman, C. J., Hrachowitz, M., McGuire, K. J., van der Velde, Y., Bertuzzo, E., and Botter, G.: Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes, Water Resources Research, 51, 4840–4847, https://doi.org/10.1002/2015WR017273, http://dx.doi.org/10.1002/2015WR017273, 2015.

Rodriguez, N. B. and Klaus, J.: Catchment travel times from composite StorAge Selection functions: superposition of streamflow generation processes, Water Resources Research, in review.

Rodriguez, N. B., McGuire, K. J., and Klaus, J.: Time-Varying Storage-Water Age Relationships in a Catchment With a Mediterranean Climate, Water Resources Research, 54, 3988–4008, https://doi.org/10.1029/2017wr021964, 2018.

Scaini, A., Audebert, M., Hissler, C., Fenicia, F., Gourdol, L., Pfister, L., and Beven, K. J.: Velocity and celerity dynamics at plot scale inferred from artificial tracing experiments and time-lapse ERT, Journal of Hydrology, 546, 28 – 43, https://doi.org/10.1016/j.jhydrol.2016.12.035, 2017.

Scaini, A., Hissler, C., Fenicia, F., Juilleret, J., Iffly, J. F., Pfister, L., and Beven, K.: Hillslope response to sprinkling and natural rainfall using velocity and celerity estimates in a slate-bedrock catchment, Journal of Hydrology, 558, 366 – 379, https://doi.org/10.1016/j.jhydrol.2017.12.011, 2018.

Schmidt, A., Frank, G., Stichler, W., Duester, L., Steinkopff, T., and Stumpp, C.: Tritium in precipitation and surface waters in Germany, Hydrological Processes, submitted.

Schwab, M. P., Klaus, J., Pfister, L., and Weiler, M.: Diel fluctuations of viscosity-driven riparian inflow affect streamflow DOC concentration, Biogeosciences, 15, 2177–2188, https://doi.org/10.5194/bg-15-2177-2018, 2018.

Soulsby, C., Piegat, K., Seibert, J., and Tetzlaff, D.: Catchment-scale estimates of flow path partitioning and water storage based on transit time and runoff modelling, Hydrological Processes, 25, 3960–3976, https://doi.org/10.1002/hyp.8324, 2011.

Soulsby, C., Tetzlaff, D., and Hrachowitz, M.: Tracers and transit times: windows for viewing catchment scale storage?, Hydrological Processes, 23, 3503–3507, https://doi.org/10.1002/hyp.7501, 2009.

Sprenger, M., Leistert, H., Gimbel, K., and Weiler, M.: Illuminating hydrological processes at the soil-vegetation-atmosphere interface with water stable isotopes, Reviews of Geophysics, 54, 674–704, https://doi.org/10.1002/2015RG000515, 2016.

Sprenger, M., Stumpp, C., Weiler, M., Aeschbach, W., Allen, S. T., Benettin, P., Dubbert, M., Hartmann, A., Hrachowitz, M., Kirchner, J. W., McDonnell, J. J., Orlowski, N., Penna, D., Pfahl, S., Rinderer, M., Rodriguez, N., Schmidt, M., and Werner, C.: The demographics of water: A review of water ages in the critical zone, Reviews of Geophysics, 0, https://doi.org/10.1029/2018RG000633, 2019.

Stamoulis, K., Ioannides, K., Kassomenos, P., and Vlachogianni, A.: Tritium Concentration in Rainwater Samples in Northwestern Greece, Fusion Science and Technology, 48, 512–515, https://doi.org/10.13182/FST05-A978, 2005.

Stewart, M. K. and Morgenstern, U.: Importance of tritium-based transit times in hydrological systems, Wiley Interdisciplinary Reviews: Water, 3, 145–154, https://doi.org/10.1002/wat2.1134, 2016.

Stewart, M. K. and Thomas, J. T.: A conceptual model of flow to the Waikoropupu Springs, NW Nelson, New Zealand, based on hydrometric and tracer (^{18}O, Cl,^{3}H and CFC) evidence, Hydrology and Earth System Sciences, 12, 1–19, https://doi.org/10.5194/hess-12-1-2008, 2008.

Stewart, M. K., Mehlhorn, J., and Elliott, S.: Hydrometric and natural tracer (oxygen-18, silica, tritium and sulphur hexafluoride) evidence for a dominant groundwater contribution to Pukemanga Stream, New Zealand, Hydrological Processes, 21, 3340–3356, https://doi.org/10.1002/hyp.6557, 2007.

Stewart, M. K., Morgenstern, U., and McDonnell, J. J.: Truncation of stream residence time: how the use of stable isotopes has skewed our concept of streamwater age and origin, Hydrological Processes, 24, 1646–1659, https://doi.org/10.1002/hyp.7576, 2010.

Stewart, M. K., Morgenstern, U., McDonnell, J. J., and Pfister, L.: The 'hidden streamflow' challenge in catchment hydrology: a call to action for stream water transit time analysis, Hydrological Processes, 26, 2061–2066, https://doi.org/10.1002/hyp.9262, 2012.

Stewart, M. K., Morgenstern, U., Gusyev, M. A., and Małoszewski, P.: Aggregation effects on tritium-based mean transit times and young water fractions in spatially heterogeneous catchments and groundwater systems, Hydrology and Earth System Sciences, 21, 4615–4627, https://doi.org/10.5194/hess-21-4615-2017, 2017.

Östlund, G. H.: Hurricane Tritium I: Preliminary Results on Hilda 1964 and Betsy 1965, pp. 58–60, American Geophysical Union (AGU), https://doi.org/10.1029/GM011p0058, 2013.

Stumpp, C., Klaus, J., and Stichler, W.: Analysis of long-term stable isotopic composition in German precipitation, Journal of Hydrology, 517, 351 – 361, https://doi.org/10.1016/j.jhydrol.2014.05.034, 2014.

Thiesen, S., Darscheid, P., and Ehret, U.: Identifying rainfall-runoff events in discharge time series: a data-driven method based on information theory, Hydrology and Earth System Sciences, 23, 1015–1034, https://doi.org/10.5194/hess-23-1015-2019, 2019.

Uhlenbrook, S., Frey, M., Leibundgut, C., and Maloszewski, P.: Hydrograph separations in a mesoscale mountainous basin at event and seasonal timescales, Water Resources Research, 38, 31–1–31–14, https://doi.org/10.1029/2001WR000938, 2002.

van der Velde, Y., Heidbüchel, I., Lyon, S. W., Nyberg, L., Rodhe, A., Bishop, K., and Troch, P. A.: Consequences of mixing assumptions for time-variable travel time distributions, Hydrological Processes, 29, 3460–3474, https://doi.org/10.1002/hyp.10372, 2015.

van Meerveld, H. J. I., Kirchner, J. W., Vis, M. J. P., Assendelft, R. S., and Seibert, J.: Expansion and contraction of the flowing stream network changes hillslope flowpath lengths and the shape of the travel time distribution, Hydrology and Earth System Sciences Discussions, 2019, 1–18, https://doi.org/10.5194/hess-2019-218, https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-218/, 2019.

Visser, A., Thaw, M., Deinhart, A., Bibby, R., Safeeq, M., Conklin, M., Esser, B., and Van der Velde, Y.: Cosmogenic Isotopes Unravel the Hydrochronology and Water Storage Dynamics of the Southern Sierra Critical Zone, Water Resources Research, 0, https://doi.org/10.1029/2018WR023665, 2019.

von Freyberg, J., Studer, B., and Kirchner, J. W.: A lab in the field: high-frequency analysis of water quality and stable isotopes in stream water and precipitation, Hydrology and Earth System Sciences, 21, 1721–1739, https://doi.org/10.5194/hess-21-1721-2017, 2017.

Waichler, S. R., Wemple, B. C., and Wigmosta, M. S.: Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest, Hydrological Processes, 19, 3177–3199, https://doi.org/10.1002/hyp.5841, 2005.

Wilusz, D. C., Harman, C. J., and Ball, W. P.: Sensitivity of Catchment Transit Times to Rainfall Variability Under Present and Future Climates, Water Resources Research, https://doi.org/10.1002/2017WR020894, http://dx.doi.org/10.1002/2017WR020894, in press, 2017.

Wrede, S., Fenicia, F., Martínez-Carreras, N., Juilleret, J., Hissler, C., Krein, A., Savenije, H. H. G., Uhlenbrook, S., Kavetski, D., and Pfister, L.: Towards more systematic perceptual model development: a case study using 3 Luxembourgish catchments, Hydrological Processes, 29, 2731–2750, https://doi.org/10.1002/hyp.10393, 2015.

Table 2. Parameter ranges and information measures before and after calibration to isotopic data

Parameter	S_{th}	ΔS_{th}	S_u	f_0	λ_1^*	λ_2	μ_2	θ_2	μ_3	θ_3	μ_{ET}	θ_{ET}
Unit	mm	mm	mm	–	–	–	mm	mm	mm	mm	mm	mm
Range for $E_2 > L_2$	[21, 196]	[0.14, 20]	[1.4, 50]	[0, 1]	[0, 0.76]	[0, 1]	[149, 1530]	[6, 100]	[21, 1561]	[2, 100]	[51, 926]	[1, 99]
Range for $E_3 < L_3$	[20, 200]	[0.1, 20]	[1, 50]	[0, 1]	[0, 0.92]	[0, 1]	[58, 1564]	[0, 100]	[67, 1600]	[0, 100]	[0, 959]	[1, 100]
Range for $E_2 > L_2$ and $E_3 < L_3$	[25, 177]	[0.17, 20]	[11, 49]	[0, 0.82]	[0, 0.76]	[0.1, 1]	[897, 1530]	[7, 69]	[440, 1561]	[7, 96]	[51, 120]	[3, 98]
Binning[a]	[20:20:200]	[0:2:20]	[0:5:50]	[0:0.1:1]	[0:0.05:1]	[0:0.1:1]	[0:100:1600]	[0:10:100]	[0:100:1600]	[0:10:100]	[0:50:1600]	[0:10:100]
$\mathcal{H}(X)$[b]	10.53	11.03	11.03	11.03	12.32	11.03	13.29	11.03	13.29	11.03	16.61	11.03
$\mathcal{H}(X\,|\,^2H)$	10.36	10.86	10.93	10.7	9.87	10.63	11.43	10.8	11.93	10.73	9.2	10.76
$\mathcal{H}(X\,|\,^3H)$	10.3	10.96	11	10.96	10.73	10.76	12.39	10.8	12.72	10.8	5.08	10.7
$\mathcal{H}(X\,|\,(^2H \cap {}^3H))$	8.37	9.8	8.4	9.67	8.8	9.67	8.24	7.24	9.67	9.14	2.33	8.24
$D_{KL}(X\,|\,(^2H \cap {}^3H), X\,|\,^2H)$	1.3	1.2	2.49	1	1.2	0.9	4.09	3.42	2.49	1.59	4.29	2.39
$D_{KL}(X\,|\,(^2H \cap {}^3H), X\,|\,^3H)$	1.59	1.2	2.66	1.4	1.1	1.13	3.59	3.02	2.79	1.23	6.78	2.49

[a]Binning is indicated as $[a : b : c]$, where a is the left edge of the first bin, b is the bin width, and c is the right edge of the last bin.

For instance, $[7 : 2 : 11]$ indicates data sorted with the two bins $[7, 9]$ and $[9, 11]$

[b]$\mathcal{H}$ and D_{KL} are expressed in bits.

Table 3. Statistics of $\overleftarrow{P_Q}(T)$ constrained by deuterium or tritium

Age statistics	^{2}H ($E_2 > 0$) [mean ± std]	^{3}H ($E_3 < 0.5$ T.U.) [mean ± std]	^{2}H and ^{3}H [mean ± std]
10th percentile [years]	0.78 ± 0.49	1.10 ± 0.57	1.44 ± 0.11
25th percentile [years]	1.16 ± 0.56	1.54 ± 0.59	1.85 ± 0.22
Median age [years]	1.77 ± 0.55	2.19 ± 0.64	2.38 ± 0.15
75th percentile [years]	2.78 ± 0.61	3.07 ± 0.74	3.26 ± 0.39
90th percentile [years]	4.64 ± 1.27	4.79 ± 1.41	5.19 ± 0.86
Mean age [years]	2.90 ± 0.54	3.12 ± 0.59	3.45 ± 0.28
F_{yw}[a] [%]	1.5 ± 1.6	1.8 ± 2.3	0.61 ± 0.53
F(T < 6 months) [%]	10 ± 8.6	6.3 ± 8.2	0.75 ± 0.58
F(T < 1 year) [%]	24 ± 17	11 ± 12	2.1 ± 1.5
F(T < 3 years) [%]	77 ± 8.5	71 ± 16	70 ± 6.6

The mean and standard deviations are calculated from all retained behavioral solutions for a given criterion. [a] Fraction of "young water" (Kirchner, 2016), younger than 0.2 years

Table 4. Storage estimate S_{95P} constrained by deuterium or tritium

Statistics of S_{95P}	^{2}H ($E_2 > 0$)	^{3}H ($E_3 < 0.5$ T.U.)	^{2}H and ^{3}H
Mean ± st. dev. [mm]	1275 ± 245	1335 ± 279	1488 ± 135
Median ± st. dev. [mm]	1281 ± 245	1392 ± 279	1505 ± 135
Min [mm]	625	660	1249
Max [mm]	1744	1806	1710

S_{95P} is calculated as the 95th percentile of Ω_{tail} (eq. 10)