AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR020117

Key Points:
- Hydrologic transport is described through StorAge Selection (SAS) functions
- High-resolution isotope data allow constraining different model approaches
- Time-variant transport processes are addressed with a single power-law-shaped SAS function

Correspondence to:
P. Benettin,
paolo.benettin@epfl.ch

Citation:
Benettin, P., C. Soulsby, C. Birkel, D. Tetzlaff, G. Botter, and A. Rinaldo (2017), Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments, *Water Resour. Res.*, *53*, 1864–1878, doi:10.1002/2016WR020117.

Received 11 NOV 2016
Accepted 7 FEB 2017
Accepted article online 11 FEB 2017
Published online 3 MAR 2017

© 2017. American Geophysical Union. All Rights Reserved.

Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments

Paolo Benettin[1], Chris Soulsby[2], Christian Birkel[2,3], Doerthe Tetzlaff[2], Gianluca Botter[4], and Andrea Rinaldo[1,4]

[1]Laboratory of Ecohydrology ENAC/IIE/ECHO, École Polytechinque Fédérale de Lausanne, Lausanne, Switzerland, [2]Northern Rivers Institute, School of Geosciences, University of Aberdeen, Aberdeen, UK, [3]Department of Geography, University of Costa Rica, San José, Costa Rica, USA, [4]Dipartimento ICEA, Università degli Studi di Padova, Padua, Italy

Abstract We use high-resolution tracer data from an experimental site to test theoretical approaches that integrate catchment-scale flow and transport processes in a unified framework centered on selective age sampling by streamflow and evapotranspiration fluxes. Transport processes operating at the catchment scale are reflected in the evolving residence time distribution of the catchment water storage and in the age selection operated by out-fluxes. Such processes are described here through StorAge Selection (SAS) functions parameterized as power laws of the normalized rank storage. Such functions are computed through appropriate solution of the master equation defining formally the evolution of residence and travel times. By representing the way in which catchment storage generates outflows composed by water of different ages, the main mechanism regulating the tracer composition of runoff is clearly identified and detailed comparison with empirical data sets are possible. Properly calibrated numerical tools provide simulations that convincingly reproduce complex measured signals of daily deuterium content in stream waters during wet and dry periods. Results for the catchment under consideration are consistent with other recent studies indicating a tendency for natural catchments to preferentially release younger available water. The study shows that power law SAS functions prove a powerful tool to explain catchment-scale transport processes that also has potential in less intensively monitored sites.

1. Introduction

The problem of how water and solutes are concurrently transported along the various (and varying) hydrological pathways of a catchment has long been the subject of inquiry [e.g., *Rinaldo and Marani*, 1987; *Rinaldo et al*., 1989; *Maloszewski et al*., 1992]. Early studies were based on the use of catchment travel times because a central concept was the instantaneous unit hydrograph (IUH), seen [*Rodriguez-Iturbe and Valdes*, 1979; *Gupta et al*., 1980] as the travel time distribution to the catchment outlet of a unit runoff-producing rainfall pulse. Such a distribution was, in turn, argued to be the distribution of contact times between fixed and mobile phases—a main driver of biogeochemical mass exchange governing transport phenomena at catchment scales. This was a partially misleading approach, because the characterization of the time spent by a parcel of water within a watershed was incorrect. In fact, although the age dynamics of the water in storage within a catchment system was correctly assumed to affect the chemical composition of hydrologic fluxes (including solute concentrations in runoff measured in the field), there was little awareness that much (if not most) of runoff had been stored in the catchment for time periods much longer than event waters [e.g., *Stewart and McDonnell*, 1991; *Kirchner*, 2003; *Bishop et al*., 2004; *McGuire et al*., 2007]. Thus, while water fluxes (and thus IUHs) are controlled by celerities reflecting how flow paths are activated, solute transport needs to track the actual velocities of molecules along hydrologic pathways thus resulting in radically different travel times [e.g., *Botter et al*., 2010; *Beven*, 2012; *McDonnell and Beven*, 2014].

Any consistent catchment transport framework must thus formally characterize the age dynamics in hydrologic systems in order to be able to incorporate both hydrograph and tracer information. As experimental evidence has mounted on the links between the tracer composition of waters in storage and in output fluxes with the age of streamflows [e.g., *Weiler et al*., 2003; *Hrachowitz et al*., 2013; *Soulsby et al*., 2009; *Tetzlaff et al*., 2015], common ground has been identified on the need to properly characterize the age dynamics in

hydrologic systems to reproduce both hydrograph and tracer information [e.g., *Rinaldo et al*., 2015]. Hence, it is not surprising that a variety of environmental and biological tracers [e.g., *Dahlke et al*., 2015; *Klaus et al*., 2015; *Beyer et al*., 2016] are used to track the movement of water, thus placing hydrologic transport at the interface between hydrology and biogeochemistry [*Hrachowitz et al*., 2016]. Recently, the characterization of hydrologic transport has also impacted plant physiology studies, as it has been experimentally shown that water transpired by the plants may be sampled from a pool that is different from that generating stream water in some environments [*McDonnell*, 2014; *Evaristo et al*., 2015].

Since early formulations, travel time distributions have been extensively explored [see *McGuire and McDonnell*, 2006] and a new generation of theoretical transport models has been introduced in recent years [*Botter et al*., 2011; *van der Velde et al*., 2012; *Harman*, 2015], which focus on catchment-scale nonstationary transport. This new approach brought to light a number of important aspects, which include (1) acknowledging the strong nonstationary character of travel time distributions and their dependence on hydrologic variability [*Botter et al*., 2010]; (2) distinguishing between the age distribution of the water stored within a catchment and the age distribution of the fluxes like discharge and evapotranspiration which remove water from the catchment storage [*Botter et al*., 2011; *van der Velde et al*., 2012; *Harman*, 2015]; (3) introducing parallels between forward-in-time and backward-in-time descriptions of water travel times [*Benettin et al*., 2015a; *Calabrese and Porporato*, 2015]; (4) and extending the framework to the age distribution of fluxes other than discharge, like evapotranspiration [*Botter*, 2012; *Soulsby et al*., 2016a].

Following the development of this new theoretical framework, applications have started to appear that model and understand nonstationary hydrologic transport in real settings. However, most of the applications are based on a specific assumption of local "complete mixing" (or "random sampling," as illustrated in section 2), and only few studies [*van der Velde et al*., 2015; *Harman*, 2015; *Queloz et al*., 2015; *Kim et al*., 2016] have actually explored the full potential of these new methods. More extensive implementations that do not need a priori assumptions are thus needed, to realize the full potential of the method to different settings, including locations with sparsely gauged basins. For this reason, studies conducted at highly monitored research locations can be of guidance for the applications of the method in different geographical areas and hydrologic conditions.

Here we utilize data from an experimental site in the Scottish Highlands, the Bruntland Burn, where daily isotope measurements were made in precipitation and streamflow over a 3 year period. Our overall objective was to test recent catchment-scale formulations of transport by nonstationary travel time distributions against this unique data set. More specifically we aimed to (1) use a StorAge Selection Approach (SAS) to simulate the isotope dynamics of streamflow; (2) demonstrate how the size of catchment-scale storage influences the time-varying age of water fluxes in streamflow; and (3) assess how effectively the approach constrains the travel time distribution of the catchment.

The paper is organized as follows: Section 2 summarizes the methods used to simulate hydrologic transport and to reproduce the tracer signals of a catchment. Section 3 introduces the Bruntland Burn catchment and the available data and details the model used to simulate its isotopic signature. Sections 4 and 5 present and discuss the model results and their interpretation. Finally, the main conclusions are summarized in section 6.

2. Methods: The Storage Selection Approach

The StorAge Selection (SAS) approach denotes the recent catchment-scale formulation of transport by nonstationary travel time distributions. While its comprehensive description can be found in the commentary by *Rinaldo et al*. [2015], some essential elements are summarized in the following.

The main task of the SAS approach is to keep track of the volumes of precipitation which have entered the catchment at different times in the past, which remain in storage and have not been released yet. To accomplish this, the "age-ranked storage" $S_T(T,t)$ [*van der Velde et al*., 2012; *Harman*, 2015] is introduced, which describes at any time t the cumulative volumes of water in storage as ranked by their age T. The units of S_T are those of a volume (hence typically mm or m^3) and, for example, $S_T = 50$ mm refers to the youngest 50 mm of storage, $S_T = 200$ mm to the youngest 200 mm and so on. Each value of S_T can be used to tag a specific precipitation event that occurred in the past: $S_T = 200$ mm identifies, for example, the precipitation

event that occurred just before the most recent 200 mm of storage entered the catchment. The variable S_T is naturally bounded by the total catchment storage $S_{tot}(t)$ so it can be useful to introduce the normalized rank storage $P_S = S_T/S_{tot}$, which is confined in $[0, 1]$ and is dimensionless. In this case, the rank storage is expressed in a fractional form, so for example $P_S = 0.2$ refers to the youngest 20% of stored water. The rank storage changes in time because new precipitation (J) bring new volumes of water into the catchment, and because other volumes are released to discharge (Q) and evapotranspiration (ET). The essence of the approach is to assign each outflow a SAS function $\omega(S_T)$ (or equivalently its integral $\Omega(S_T)$) that expresses how much each different volume in storage contributes to that outflow. The use of these functions represents a practical advance as they effectively collapse complex 3-D transport dynamics into a single curve. The governing equations can be written in a number of equivalent forms which have been proposed in the literature [e.g., *Botter et al.*, 2011; *van der Velde et al.*, 2012; *Harman*, 2015; *Benettin et al.*, 2015b]. Here we employ the cumulative version, which is suitable for numerical solutions

$$\frac{\partial S_T(T,t)}{\partial t} + \frac{\partial S_T(T,t)}{\partial T} = J(t) - Q(t)\,\Omega_Q(S_T(T,t),t) - ET(t)\,\Omega_{ET}(S_T(T,t),t). \tag{1}$$

The SAS functions can have any shape, but they must comply with the constrain

$$\Omega_Q(S_T \to S_{tot}(t),t) = \Omega_{ET}(S_T \to S_{tot}(t),t) = 1, \tag{2}$$

which is verified when probability density functions are used to parameterize $\omega(S_T,t)$. The discharge age distributions $p_Q(T,t)$ can be obtained from equation (1) as

$$p_Q(T,t) = \omega(S_T,t)\,\frac{\partial S_T}{\partial T}. \tag{3}$$

Equation (3) shows how stream age distributions result from combining the available storage volumes (defined by S_T at any time) with the catchment tendency to release water from different components of the storage (expressed by the SAS functions). The same selection process can be used to quantify the solutes or tracers that are released to the stream from the catchment storage. As each stored volume $S_T(T,t)$ can be associated with its solute (or tracer) concentration $C_S(S_T,t)$, the stream concentration $C_Q(t)$ is obtained by integrating the contribution of all the discharged volumes

$$C_Q(t) = \int_0^{S_{tot}} C_S(S_t,t)\,\omega_Q(S_T,t)\,dS_T. \tag{4}$$

In the particular case of a conservative environmental tracer such as chloride or stable water isotopes, the stored concentration is equal to that of precipitation at the time it entered the catchment, but in general C_S needs to take into account possible biogeochemical processes, like, e.g., decays and evapoconcentration [*Bertuzzo et al.*, 2013] or mass transfer processes [*Benettin et al.*, 2015b]. For the particular case of stable water isotopes, simple models can be introduced to take into account the possible occurrence of evaporative fractionation [*Gat*, 1996] (see Appendix B).

The description of transport as a selection process facilitates considerable flexibility for the balance equation (1). One can for example divide the total ET into evaporation and transpiration contributions, or identify additional outflows, like, e.g., interception. Each different outflow must then be assigned a SAS function, which can be used to assess the impact of the given flux on the storage balance. The SAS approach is also potentially well-suited to investigate possible ecohydrological separation between water used by vegetation and streamflow [see *McDonnell*, 2014], as it is intrinsically based on how the different outflows sample the storage pool. However, it should be noted that it is difficult to gather appropriate field data to constrain a catchment-scale model, because a time series of representative sap flow from the whole catchment (and hence diverse assemblages of vegetation cover) would be required. Still, the method can be used to check whether the removal of water by ET can have an indirect effect on C_Q [*Queloz et al.*, 2015].

From an operational point of view, there are two main options for using the SAS function approach (Figure 1). The first option requires a partitioning of the watershed into conceptual compartments representing different runoff sources, where water volumes of different ages are assumed to be sampled uniformly. As the uniform sampling is also known as Random Sampling (RS), this approach is here termed multi-RS, to recall that the RS scheme is applied separately to each compartment. This approach has been widely used in the literature to

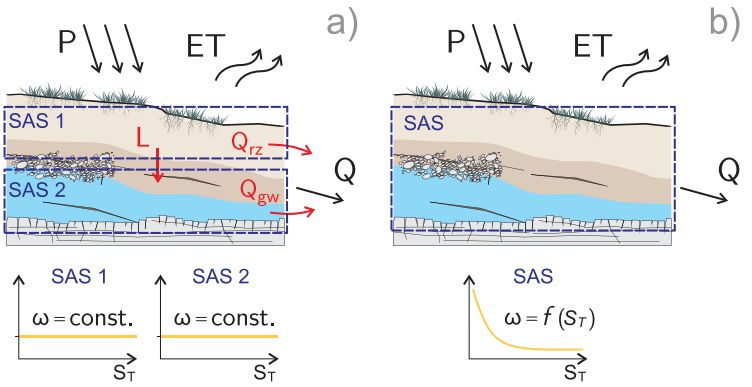

Figure 1. Examples of two main approaches to hydrologic transport using the StorAge Selection (SAS) functions: (a) multi RS approach, where the SAS is uniform in each partitioning of the catchment, but a full hydrologic model is needed to estimate the internal fluxes (red arrows in the plot); (b) direct SAS approach, where the SAS function is nonuniform but all the hydrologic fluxes can be measured. In this study, only the direct SAS approach is employed.

simulate the transport of stable water isotopes [e.g., *Birkel et al.*, 2015], conservative solutes of atmospheric or agricultural origins [*Hrachowitz et al.*, 2013; *Benettin et al.*, 2013; *Hrachowitz et al.*, 2015], and more recently also nonconservative geogenic solutes [*Benettin et al.*, 2015b; *Tunaley et al.*, 2016]. One of the main advantages is that the subdivision into functional units (e.g., soil and groundwater compartments, Figure 1a) is flexible and the local RS assumption allows easy mathematical manipulations of equations (1–4). Although this method is easy to implement, it suffers from a rigid compartmentalization of the system (which is obviously continuous) and from the need for a coupled hydrologic model to estimate the internal fluxes from and to the different stores. The second option considers the catchment as a whole and requires just one catchment-scale SAS function for each outflow (Figure 1b). This option is computationally more expensive, but has a major advantage; the hydrologic fluxes involved in the computation are catchment-scale fluxes (basically, *P*, *ET*, and *Q*) that can typically be measured or estimated without the need of a full hydrologic model. The challenge with this "direct-SAS" approach is then to assess whether one single storage selection function, coupled to the measured in- and out-fluxes, can capture the main dynamics of hydrologic transport at catchment scale. The only parameters that need to be calibrated are those that define the shape of the SAS functions (and the total storage in some cases); hence, their number is typically limited to 3–5. In this study, we make use of the "direct-SAS" approach and the details of the implementation are given in section 3.2.

3. Case Study: The Bruntland Burn Catchment

3.1. Data and Catchment Description

The Bruntland Burn catchment (Figure 2) is a 3.2 km^2 subbasin of the Girnock Burn and both sites have been the focus of long-term tracer-based hydrological investigations [*Birkel et al.*, 2011]. The dominant processes generating streamflow have been identified through hydrometric observations coupled with frequent sampling for stable water isotopes in precipitation, soil water, groundwater, and streamflow since 2011 (see *Tetzlaff et al.* [2014] for full details). Briefly, the catchment, which receives around 1000 mm of precipitation a year (<5% of which is snow), is underlain by low-permeability granite and metamorphic rocks. Land cover is mostly characterized by heather (*Calluna* and *Erica* spp. are the dominant shrubs) moorland with around 10% forest cover of Scots Pine (*Pinus sylvestris*). The topography reflects glaciation, with a wide valley bottom (at ~250 m), fringed by steeper slopes which reach up to 560 m (Figure 2). The lower slopes of the catchment (around 70% of the area) are covered by glacial drift deposits (mainly a sandy-silt loam with abundant clasts). Hydrogeophysical surveys have shown that this averages depths of ~5 m on the steeper slopes and ~25 m in the valley bottom and is mostly saturated, acting as a significant store of groundwater which maintains base flows [*Soulsby et al.*, 2016b]. The catchment soil cover is characterized by peats (histosols) 0.5–4 m which cover around 20% of the area mostly occupying the flat valley bottoms area in riparian areas fringing the stream channel network [*Tetzlaff et al.*, 2014]. These are saturated for much of the year and generate saturation overland flow as the dominant mechanism of storm runoff response. The steeper slopes are mostly covered by shallow (<0.7 m) podzols (spodosols) which tend to be mostly freely draining and recharge deeper groundwater.

Water samples were collected on a daily basis for isotope analysis; precipitation samples were cumulative over a 24 h period, while stream water samples were instantaneously collected at 9 A.M. each day. The

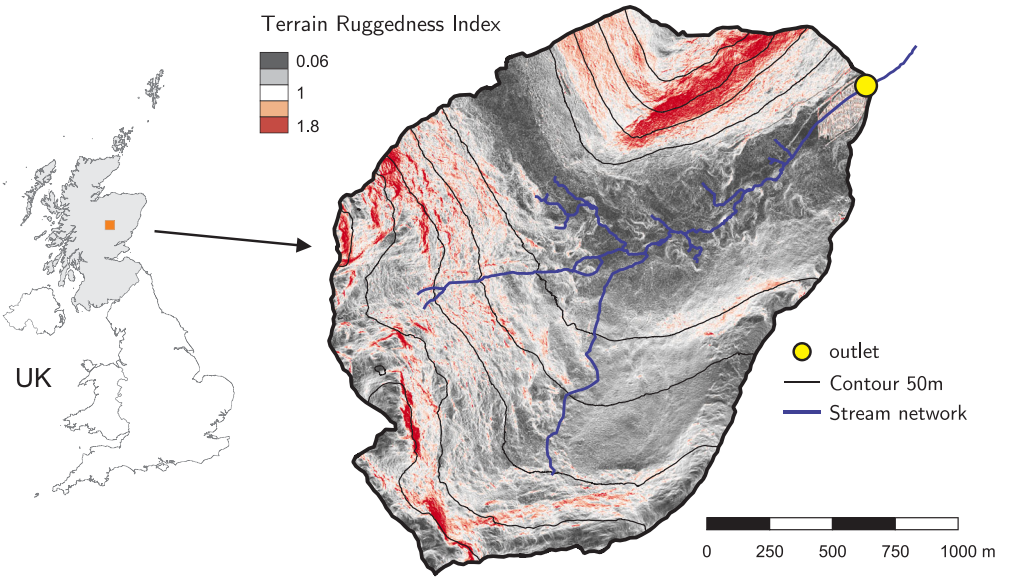

Figure 2. Map of the Brundtland Burn (BB) catchment, with a detailed Terrain Ruggedness Index derived from a 1 m^2 LiDAR digital elevation model with monitoring at the catchment outlet (yellow circle). The ruggedness of the BB catchment clearly differentiates the hillslopes (white to red colors) from the relatively flat and wide valley bottom area (gray areas). However, the high-resolution elevation model exhibits important microtopographical features (white color) in the valley bottom close to the stream network that might affect water flow paths and transport.

sampling period encompassed 3 years from 1 June 2011 to 31 May 2014; during this period the catchment experienced the coldest winter for 50 years (2012/2013), the driest summer for 10 years (in 2013), and the wettest winter (December 2013/January 2014) for 20 years, so it was subject to marked variability in storage [*Soulsby et al*., 2015]. Samples were returned to the laboratory and analyzed using a Los Gatos DLT-100 laser spectroscope for deuterium (δ^2H) with a precision of 0.4‰ after calibration with Vienna Standard Mean Ocean Water (VSMOW) standards. Precipitation samples are more likely to be more depleted in δ^2H in winter and enriched in summer, but day-to-day variation can be marked.

Figure 3 shows the measured deuterium data sets. As fluctuations in precipitation are roughly 5 times larger than those in streamflow, the *y* axes of the figure were rescaled around the long-term means, to allow a direct visual comparison between the two time series. The resulting plot shows that the two signals have an almost synchronous periodicity. Moreover, during individual storm events (see gray vertical bars in Figure 3), streamflow content can have a local positive/negative peak which is related to local positive/negative fluctuations in precipitation. Streamflow deuterium content is also characterized by high-frequency fluctuations that are not directly correlated with precipitation content and may look like measurement noise, but are actually much larger than the typical analytical error. This probably reflects patches of fractionated waters that are heterogeneously displaced from the peatlands (especially during smaller events), making the high-frequency fluctuations a local dynamic that is difficult to link with the average catchment states [*Sprenger et al*., 2017].

As typically observed in catchment tracer studies [see *Kirchner*, 2003], the streamflow response in terms of isotopic content is strongly damped with respect to precipitation, despite the stream responsiveness to storm events. Previous work has shown that this conceptually reflects the mixing of varying fluxes of younger, less damped overland flow from the peats soils with more stable fluxes of older more damped groundwater [*Tetzlaff et al*., 2014]. In addition, the mean flow-weighted isotopic content of precipitation is −59.8‰, while that of discharge is −57.4‰; this is indicative of the effect of evaporative fractionation from areas of surface water on the saturated peatland during warmer periods in the spring and summer [*Lessels et al*., 2016].

3.2. Model Implementation and Calibration

The implementation of the direct-SAS approach requires knowledge of the catchment-scale fluxes and storage. We took *P*, *Q*, and the estimated potential evapotranspiration *PET* from *Soulsby et al*. [2015]. Relative

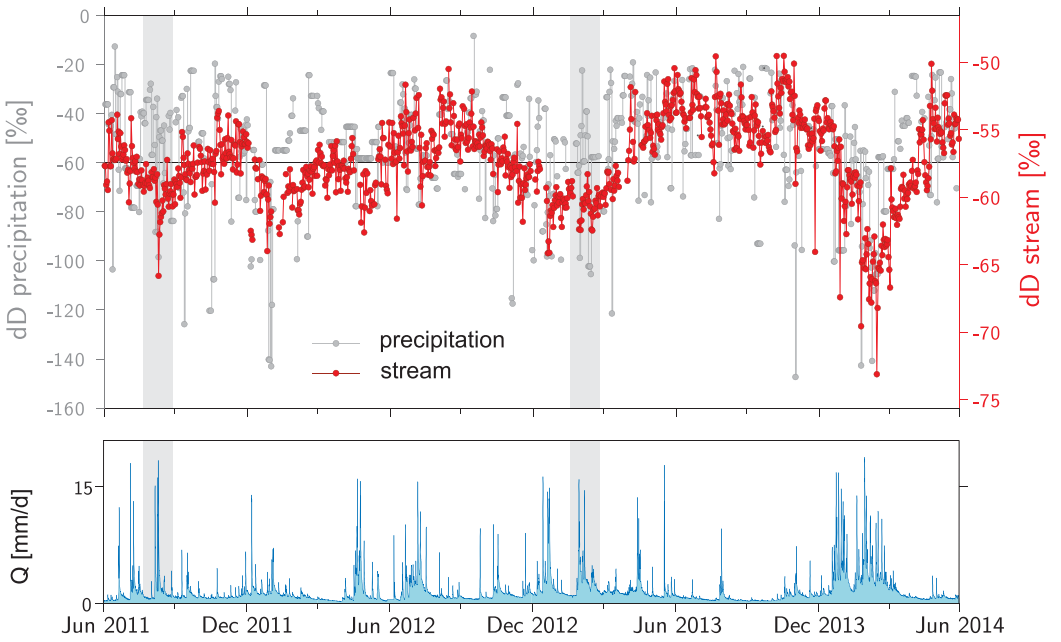

Figure 3. Measured data from 1 June 2011 to 1 June 2014. (top) Daily measured deuterium content of precipitation (gray color, left axis) and streamflow (red color, right axis). The axes have different scalings (roughly 1:5) around the long-term means (black continuous line), to allow visual comparison between the two data sets. Two events are also singled-out in the plot, to highlight the isotopic behavior during individual storm events. (bottom) Measured hourly discharge.

storage variations were then computed solving the water balance $dS(t)/dt = P(t) - \ell_f(t)\,PET(t) - Q(t)$, where $\ell_f(t)$ is a limiting factor for potential evapotranspiration obtained as min $\{1, S(t)/(2\,std(S))\}$. Storage variations computed with this simplified methodology show no trend in the observed 3 year period.

The model aims to conceptualize and simulate how each precipitation input is incorporated within the catchment storage and later released through discharge and evapotranspiration. As the isotopic content of precipitation (C_J) is known, the model can be used to predict the isotopic content of discharge (C_Q). The comparison between measured and modeled C_Q over the 3 year available record is then used to calibrate the main transport parameters.

The main task of the model is to update the ranked storage S_T (together with its isotopic composition $C_S(S_T, t)$) and evaluate the SAS functions at each time step. The SAS functions used in this study are defined over the normalized rank storage $P_S = S_T/S_{tot}$, because it allows a power law shape [*Queloz et al.*, 2015] of the kind

$$\omega(S_T/S_{tot}) = \omega(P_S) = k\,(P_S)^{k-1}, \tag{5}$$

where the only parameter involved, k, indicates the "preference" for releasing water from younger/older storage components. The preferential release of younger waters corresponds to values $k < 1$ (where the lower the value of k, the higher the preference), while the affinity for older waters corresponds to $k > 1$. In the case where the released particles are sampled almost uniformly from all the ranked storage, then $k \approx 1$, which is equivalent to the case of a well-mixed system. The SAS function as defined by equation (5) does not vary in time, but it can be made time-variant to simulate the evolving transport mechanisms during the different phases of the hydrologic response [see *van der Velde et al.*, 2015; *Harman*, 2015; *Kim et al.*, 2016]. A simple way to achieve this is to define a measure of catchment wetness (e.g., some normalized storage variations) and let the parameter k vary linearly between two end-members k_1 and k_2. A calibration procedure can then determine whether the time variance of the SAS function is strong and whether it justifies the required added parameter.

In this study, two different model configurations are tested. The first (Model 1) is a simpler model with fixed SAS functions, where the parameter k_Q is fixed to a value k_{Q0}. The second model (Model 2) attempts to evaluate the possible time variance of the SAS functions by letting the parameter k_Q vary between two end-

Table 1. Summary of the Calibration Parameters

Parameter	Symbol	Low. Bound	Upper Bound
SAS param. for Q	k_Q	0.2	3
SAS param. for ET	k_{ET}	0.2	3
Average total storage (mm)	$\bar{S}_{tot}$	500	4000
Fractionation factor	α	0.95	1.00

members (k_{Q1} and k_{Q2}), based on the storage variations of the system. In this case, a wetness state is defined as $ws(t)=(S(t)-S_{min})/(S_{max}-S_{min})$ and the expression for the SAS function exponent becomes $k_Q(t)=k_{Q1}+(1-ws(t))(k_{Q2}-k_{Q1})$. The potential effect of evaporative fractionation is taken into account using the formula from Appendix B, which requires a further parameter, α. In case calibration identifies a value $\alpha < 1$, this indicates that fractionation has a significant impact. Finally, the model needs knowledge of the total storage, but only storage variations can be assessed from purely hydrologic measurements. Hence, a parameter indicating the initial catchment storage S_0 is required. Model 1 hence needs four parameters while Model 2 needs five.

The models are run at hourly time steps using a Euler-Forward scheme based on the operational routine outlined in Appendix A. The required initial conditions (i.e., the initial storage distribution S_T and its isotopic content) are obtained by running a spin-up period of 8 years, based on the repetition of the hydrologic data from the first year of measurements (where the storage returns to its initial value). The isotopic content of the initial S_T was set to $-58‰$. The model is calibrated using the Markov Chain Monte Carlo algorithm $DREAM_{ZS}$ [*Vrugt et al.*, 2009; *ter Braak and Vrugt*, 2008], using the uniform prior parameter ranges summarized in Table 1. The calibration was run from 1 June 2011 to 1 June 2014 (1095 days). The results from the MCMC calibration are formally evaluated using the Deviance Information Criterion (DIC) [*Spiegelhalter et al.*, 2002]. The DIC is based on the computation of the deviance, which is a statistic expressing the distribution of model residuals. DIC discounts the model performance based on the additional degrees of freedom related to the number of calibrated parameters. A lower DIC value, regardless of its actual value, indicates the best model. Details on the DIC computations are reported in Appendix C.

4. Results

The results of the calibration are shown in Figure 4, where the posterior parameter distributions of the two model configurations are reported together with some metrics of their performance. The only difference between the two configurations is in the choice of the SAS function for Q: endowed with just one

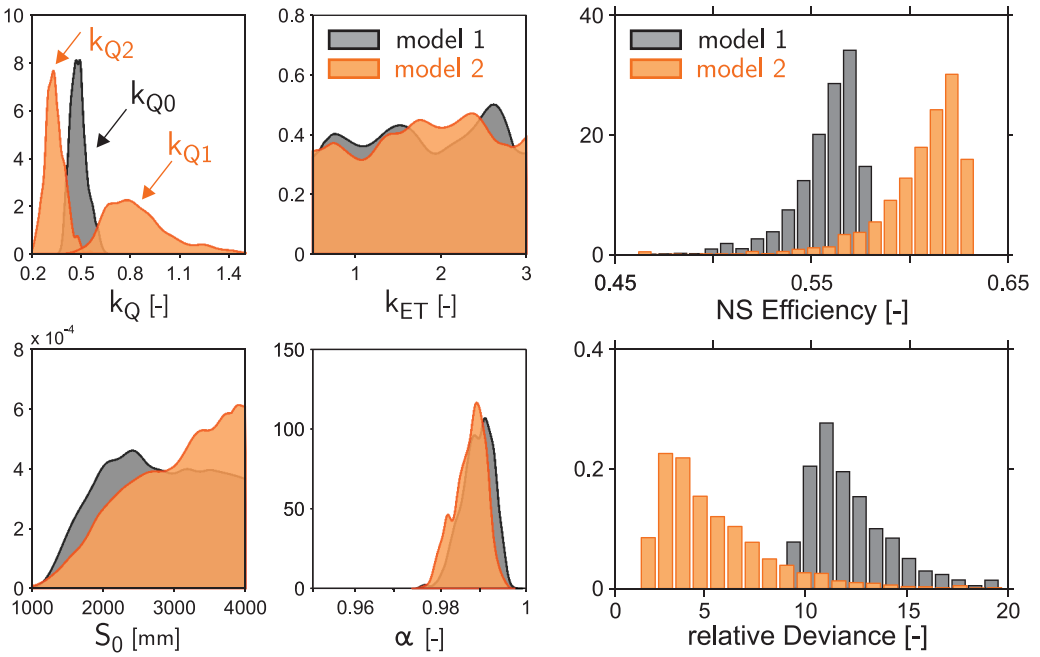

Figure 4. Results from the two calibrated models: posterior parameter distributions (left) and model performances (right). Model 2 differs from Model 1 because the ω_Q function is time-variant and needs two parameters (k_{Q1} and k_{Q2}) instead of one parameter (k_{Q0}).

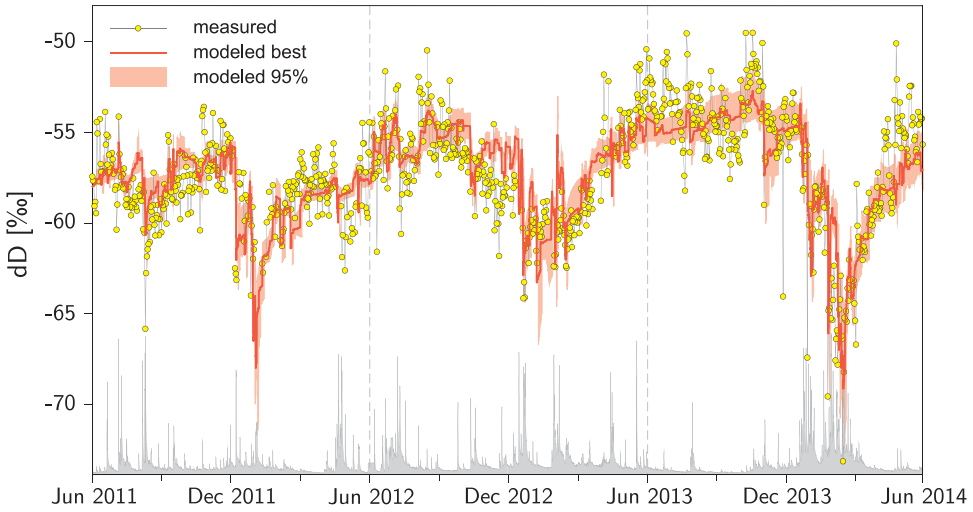

Figure 5. Posterior simulations of the selected Model 2. Hourly simulated deuterium content (red band indicating the 95% of the posterior simulation and line series indicating the best performance) compared to daily measurements (yellow dots). The lower part of the plot qualitatively shows the discharge time series, to aid the identification of wet and dry periods.

parameter k_{Q0} in Model 1, while for Model 2 it is allowed to change with wetness, ranging from a maximum value (k_{Q1}) during dry periods to a minimum value (k_{Q2}) during wet periods.

For both configurations, the SAS function parameter for *ET* (k_{ET}) is not identified. This is likely due to the more limited role of evapotranspiration in the low energy Scottish climate and wet soils, which makes it difficult to assess the affinity of *ET* fluxes for the younger/older stored waters. The fractionation parameter α, instead, is clearly identified and the optimal values are lower than 1 for both models, indicating a mild but significant isotopic enrichment of stream water due to evaporative fractionation. The initial storage parameter S_0 is only partially identified, because any value larger than $S_0 \approx 2300$ mm provides equally good results. This issue is rather common in transport studies [see *Benettin et al.*, 2015c; *Birkel et al.*, 2011; *Hrachowitz et al.*, 2015] and is further discussed in section 5. Finally, the parameter k_Q is well identified in both configurations, and Model 2 shows a clear advantage in using a time variant SAS function, otherwise the two endmembers k_{Q1} and k_{Q2} would be close to the stationary k_{Q0}. The variation of the SAS function with catchment wetness plays a significant role in the simulation performance. Although both configurations result in a satisfactory simulation of the measured signal (Nash-Sutcliffe (NS) efficiencies typically higher than 0.55), Model 2 clearly shows better performances. The model deviance, indeed, shows that Model 2 has a higher accuracy, and even accounting for its higher complexity, Model 2 still has a lower relative DIC (11.5) compared to Model 1 (14.7), meaning that the extra parameter is justified. Consequently, the results from Model 2 were selected and used for the following analysis.

To illustrate the model results in terms of stream isotopic content and age, a sample of 500 combinations extracted from the posterior parameter distributions was used to rerun Model 2. The best combination in terms of deuterium transport ($k_{Q1}=0.36$, $k_{Q2}=0.80$, $k_{ET}=0.90$, $S_0=2400$, $\alpha=0.991$) is used as reference.

The simulated stream deuterium content is shown against daily measured values in Figure 5. The model reproduces the general periodicity of the measured signal as well as much of the main event-based variations (e.g., December 2011 and December 2013). However, the simulation is not able to reproduce the highest-frequency variations of the signal which are not related to particular wetness/climatic conditions (notably during summer 2013), confirming the idea that these fluctuations are due to local and heterogeneous dynamics which cannot be addressed by an integrated model.

The simple method used to account for evaporative fractionation proves generally effective in increasing the mean (flow-weighted) isotopic content of discharge $\overline{\delta D_Q}$. Indeed, the best performance has $\overline{\delta D_Q}=-58.2$ ‰ obtained with $\alpha=0.991$, while the same model with $\alpha=1$ (i.e., no fractionation) results in $\overline{\delta D_Q}=-59.5$‰. The increase in more enriched discharge samples is most pronounced during (and immediately after) dry spells (like in the summer of 2013) and is needed to match the observed behavior during such periods.

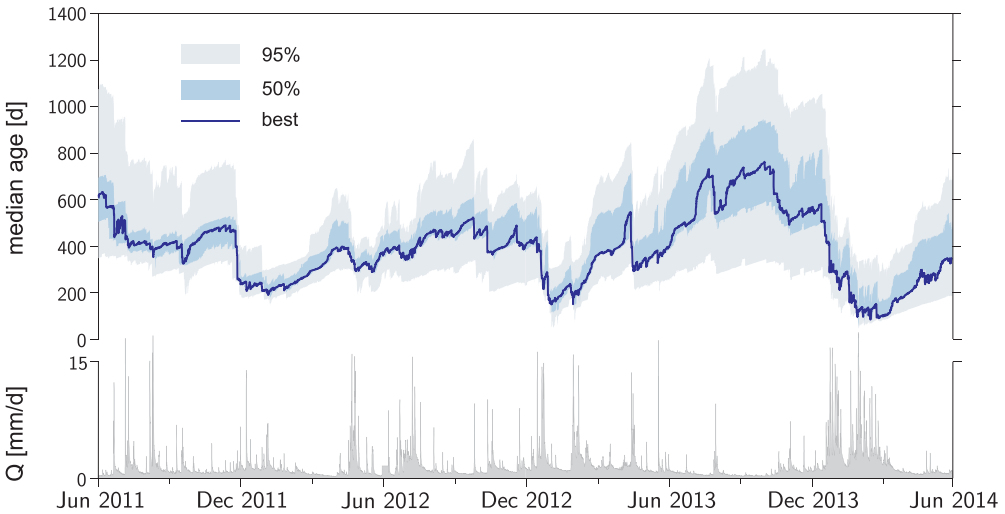

Figure 6. Simulations of the median age of streamflow using the posterior parameter distributions of Model 2. The colored bands indicate 50 and 95% of the posterior simulation, while the line series indicates the simulation which best performs in terms of deuterium transport. The discharge time series is provided at the bottom to visually assess the impact of hydrologic conditions on the median age.

We use the median age (i.e., the age which is not exceeded by 50% of the water particles) as a representative metric of streamflow age. The median is less impacted than other statistics by the poor identifiability of the older water components as it does not need to specify how old that older component is. As the model computes the age distributions at any time step, a complete time series of the median age is available, which shows the variability of the metric in response to the changing hydrologic state of the system. Figure 6 shows the evolution of the median age of stream water exiting the system for the posterior parameters sample. The plot reveals that the parameters have a large impact on the median, which is characterized by a high degree of uncertainty. Nevertheless, the dynamics are similar for all parameter combinations, with pronounced decreases in the median age during and immediately after storm events, and large increases during drier periods. The uncertainty is also larger during drier periods, where the influence of the total storage (and hence its uncertainty) is higher.

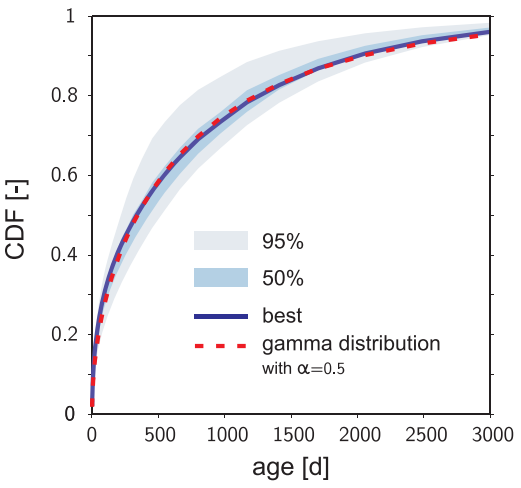

Figure 7. Marginal age distribution of streamflow from the posterior parameter sample. The colored bands indicate 50 and 95% of the posterior simulation, while the line series indicates the simulation which best performs in terms of deuterium transport. A gamma distribution with shape parameter α=0.5 and mean value 750 days is superimposed to the simulation (red dashed line).

The age distributions computed at any time step can be averaged to obtain a long-term age distribution [*Heidbuechel et al.*, 2012], also known as the marginal distribution. The flow-weighted marginal distribution from the posterior parameter sample is shown in Figure 7. The distribution is very smooth because it is the ensemble average of more than 26,000 different curves (i.e., the number of hourly time steps in the simulated 3 years). The marginal distribution represents the long-term average behavior of the system; thus, it is a useful metric to quantify some characteristic properties of a catchment. In this case, the marginal distribution predicts that, on average, only 10% of the flow is younger than 10 days, while 40−60% is older than 1 year. The model also estimates that about 5−15% is older than 5 years, but these estimates of the oldest streamflow components are more sensitive to the uncertainty of the parameters. The smooth shape of the marginal distribution can be easily approximated by a gamma distribution (Figure 7). Indeed, a gamma with shape parameter α=0.5 and mean

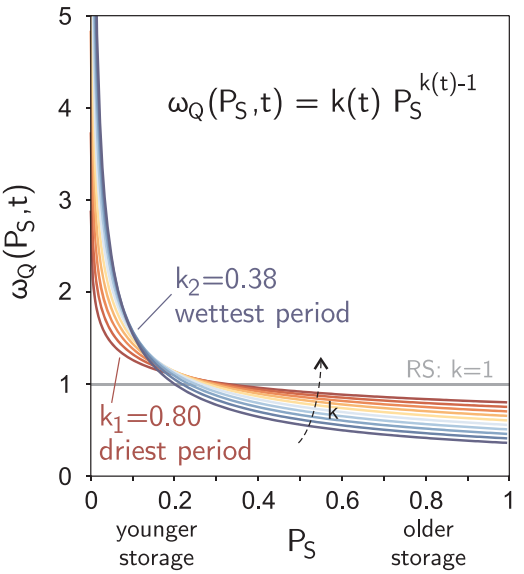

Figure 8. Variation of the SAS function ω_Q with catchment wetness. The affinity for the younger storage volumes is more pronounced during wet periods. The parameters k_1 and k_2 are those of the best model simulation.

value 750 days almost overlaps with the marginal distribution obtained from the best model parameters. This result is similar to that found by *Benettin et al.* [2015c] for the Plynlimon Upper Hafren catchment, thus reinforcing the idea that stationary travel time distributions (like the gamma, often used in the literature) can be seen as the ensemble average of a set of time-variant age distributions. Hence, this result helps bridge the recent advances on water age theory and traditional estimates of stationary catchment travel time distributions [*McGuire and McDonnell*, 2006].

5. Discussion

Our modeling exercise confirms a general preference for catchments to release the younger available water, as opposed to other systems like lysimeters where the dominant vertical percolation creates a preference for the older available components [*Queloz et al.*, 2015; *Kim et al.*, 2016]. This behavior is suitably captured by power-law-shaped SAS functions with exponents smaller than one. In the optimally calibrated model, the exponent k_Q (Figure 8) varies from 0.38 to 0.80 from the wettest period (high storage) to the driest (low storage), with an average value during the 3 years of simulation of about 0.5. Hence, the tendency of the catchment for releasing young water to streamflow appears stronger when the storage is higher (and vice versa). The same pattern was noted by *Harman* [2015] and termed "inverse storage effect," to stress that contrary to classically defined well-mixed systems, a precipitation pulse that lands on the catchment during high-storage conditions will make a relatively larger contribution to event streamflow than during low-storage conditions, as rapid hydrological flow paths are most strongly activated at such times. This analysis was made possible by the availability of accurate hydrological and isotopic data sets at the Bruntland Burn.

The average value k_Q=0.5 for the ω_Q parameter is equal to the optimal value k_{Q0} of the tested model configuration with stationary SAS function (Model 1). This reinforces the idea that, on average, a fixed value k_Q =0.5 is representative of the average catchment state and that perturbations of that value arise when the system is forced by dry/wet conditions. To characterize the deviations from the average catchment state, we have used here the variations in catchment storage, but different metrics could be introduced to describe it, like discharge or possibly the dynamic expansion of the river network and its riparian areas [*Godsey and Kirchner*, 2014]. Furthermore, one could use a hydrological model to derive metrics of catchment state that match, for example, the hysteretic patterns between water age and catchment wetness which have been observed in other modeling application studies [*Hrachowitz et al.*, 2015; *Benettin et al.*, 2015c].

The fact that the parameter k_Q proves meaningful and not just a pure calibration artifact is a significant finding that may help guide the development of transport models at locations with scarce tracer or water quality monitoring. As a first-order approximation, one could use a stationary ω_Q function with $k_Q \approx 0.5$, which for entire watersheds proves to be more realistic than a well-mixed assumption ($k_Q = 1$). Then values in the range $[0.3, 0.7]$ can be individually tested without the need for a full calibration.

Given the flashy nature of the hydrograph in this upland catchment, it is interesting that only around 10% of the annual flow appears to have an age less than 10 days. While this initially appears surprising, it is consistent with the high water storage in the riparian peats which is largely displaced by incoming precipitation leading to low (<20%) "new water" fractions in overland flow and storm runoff and a damping of the isotope signal [*Tetzlaff et al.*, 2014].

Our work attempted to estimate the uncertainty related to water age, which is not often done in transport studies. The results revealed a rather large uncertainty on the median water age (Figure 6), but the main

age dynamics (e.g., the sharp decrease in water age after a storm event) are consistent across the different parameter combinations. Moreover, the estimates give an order of magnitude of the (flow-weighted) median water age, which in this case is 330 days. This result is generally consistent with the 420 days found independently by *Soulsby et al*. [2015] through an alternative modeling approach based on empirical work in the catchment.

Further comparisons can be made with this conceptually based tracer-aided models at the Bruntland Burn catchment. In the study by *Soulsby et al*. [2015], a conceptual rainfall-runoff model is used which links three main storage zones representing groundwater in the drift, the steeper hillslopes and saturated peat soils in the riparian area [*Birkel et al*., 2015]. Water and tracer fluxes are distributed between and routed through these stores in a nonstationary way according to an algorithm that tracks the nonlinear expansion and contraction of saturation in the valley bottom riparian areas corresponding to the peat soils [*Soulsby et al*., 2016a]. Flux tracking subroutines in the model show that in dry periods older water ($>$3 years old) from the groundwater store sustains base flows, and storm runoff is usually generated from younger waters ($<$0.3 years) in overland flow from the peat. In the wettest periods, the saturation area expands and receives drainage from the hillslope store, further increasing the influence of younger waters on the stream when storage is high, also reproducing an "inverse-storage" effect. The convergence of the results of the direct SAS approach with that of the conceptual models is reassuring and gives additional confidence in our new estimates. Moreover, it highlights the utility of the SAS approach in being able to reproduce the deuterium time series and derive the time-variant travel time distribution from an elegant, parsimonious model (only four or five parameters) which can capture the contribution of older and younger waters sources with no a priori assumptions about partitioning the catchment into different hydrological response units that is needed in spatially distributed models [*van Huijgevoort et al*., 2016].

The main source of uncertainty in the modeling is represented by the storage parameter, which controls the timescale of the transport processes, without exerting a major influence on the local age dynamics. The storage uncertainty is indeed a structural problem in hydrology because the older water components are poorly represented in streamflow, and they usually have a rather uniform tracer signal or chemical signature that makes it difficult to quantify their actual age. The conceptual model reported by *Soulsby et al*. [2015] was subject to the same problem, although likely similar storage magnitudes to those calibrated by the SAS approach ($>$2000 mm) were found. For this reason, alternative and independent methods can be used to help constrain the parameter values. One such example is the young water fraction proposed by *Kirchner* [2016], which quantifies the fraction of water younger than 2–3 months and could be directly compared to the model estimates. Possible improvements can also be achieved through the use of other tracers whose input varies on a much larger timescale than stable isotopes, allowing a better quantification of the older storage volumes and ages [e.g., *Stewart et al*., 2012; *Kolbe et al*., 2016]. Additionally, the physical characteristics and volume of older water stores can be estimated through geophysical techniques. In the case of the Bruntland Burn, electrical resistivity tomography (ERT) surveys of the soil and drift cover and its water content along representative transects provided a basis for catchment-scale extrapolation of total storage. These methods constrained the catchment storage between 1600 and 3000 mm [*Soulsby et al*., 2015], which is again consistent with the findings of the SAS model.

6. Conclusions

The main findings of this work can be summarized as follows:

1. The availability of unique, high-resolution, and long-term data sets at the Bruntland Burn experimental catchment allows testing and constraining different model approaches that help unravel the problem of water and solute transport through a catchment.
2. The implementation of a transport model through the "direct-SAS" approach (i.e., without a priori separation of the watershed into assumed functional units) proved feasible, computationally efficient and it was able to reproduce to a high goodness of fit the observed deuterium signature at the Bruntland Burn. Results indicate that a fixed ω_Q function already provides satisfactory results, but the use of a time-variant ω_Q which follows the storage variations of the system was able to provide significantly improved performances.

3. The age distributions resulting from the calibrated model show a pronounced time variance of streamflow age, both on a seasonal basis and during the different phases of the hydrologic response. Still, the long-term streamflow age distribution (marginal distribution), which can be used to subsume some long-term properties of the catchment, is rather regular and can be compared to smooth probability density functions like the gamma distribution.
4. The isotope simulations, calibrated storage, and time variant-age distributions are consistent with those reported by more complex conceptual models which make a priori assumptions about the catchment structure and hydrological function and whose specific goals is the simultaneous simulation of streamflow and tracer dynamics. The resulting selection of ages from stores of younger and older water is also consistent with empirical hydrometric and geophysical evidence from the catchment.
5. The observed isotopic enrichment could be effectively reproduced by modeling isotopic fractionation similarly to an evapoconcentration process.

Overall, our work demonstrates the suitability of the method to address catchment-scale transport problems and highlights its potential for application in a range of contexts. Having a flexible tool grounded on a solid theoretical basis is a concrete step forward for our understanding of transport processes and addressing important environmental challenges. Among these challenges, we foresee a central role for SAS function approaches in understanding the storage selection mechanisms of the evapotranspiration fluxes, in the application to large-scale transport of complex yet harmful chemicals like nitrates, and in the proper estimation of solute mass loads exiting a catchment.

Appendix A: Steps to Implement the Water Age Balance

The numerical implementation of equation (1) is based on updating at each time the rank storage S_T (together with its isotopic composition $C_S(S_T)$) and evaluating the SAS functions $\omega(S_T)$. The method requires the system fluxes (J, Q, ET) and the isotopic composition of precipitation C_J. Moreover, the initial conditions S_{T0} and $C_S(S_{T0})$ must be provided. At each time step i of length Δt, the following steps can be performed:

1. In case new precipitation inputs enter the system, update the rank storage distribution by introducing the new volume $S_T(1)=J(i)\,\Delta t$ with its solute concentration $C_S(1)=C_J(i)$, and shifting all other elements by one position in the rank.
2. Evaluate the cumulative SAS functions $\Omega_Q(S_T)$ and $\Omega_{ET}(S_T)$.
3. Update the rank storage to account for particles removed by $Q(i)$ and $ET(i)$: $\Delta S_T/\Delta t=-Q(i)\,\Omega_Q-ET(i)\,\Omega_{ET}$.

Along with the fundamental routine, further operations can be implemented:

1. Keep track of the age T of each element of the rank storage.
2. Compute the age distribution $p_Q=\omega_Q\,\Delta S_T/\Delta T$ and $p_{ET}=\omega_{ET}\,\Delta S_T/\Delta T$.
3. Update the solute concentration of the ranked storage to account for processes f that may cause a solute to be nonconservative (e.g., evaporative fractionation or decay): $C_S=f(S_T,T)$.
4. Compute solute concentrations in the outputs $C_Q(i)=\omega_Q\cdot C_S$ and $C_{ET}(i)=\omega_{ET}\cdot C_S$.

Note that the variable S_T has infinite support, but numerically one needs to work with a finite number of elements. It is hence convenient to aggregate the oldest volumes after a certain threshold S_{th} (e.g., 95% of S_{tot}) and consider them as a whole. From an age point of view, this means that the ages beyond that threshold are sampled uniformly. The value of S_T should be high enough to avoid numerical inaccuracies, but at the same time as low as possible to avoid considering a large amount of negligible storage contributions.

Appendix B: Accounting for Evaporative Fractionation

A simple way to account for evaporative fractionation is to assume that the overall ET flux is depleted in isotopic content by a factor α with respect to the water storage. This depletion causes the remaining water to get isotopically enriched with time.

Similarly to the case of evapoconcentration, a pulse i of water that enters the catchment at time t_i, and is subject to ET during its permanence inside the catchment, undergoes an increase in mass concentration that can be quantified as [*Queloz et al.*, 2015, Appendix A]

$$\frac{dC_i(T)}{dT} = (1-\alpha)\,\frac{ET(t_i+T)}{S(t_i+T)}\,\omega_{ET}(T)\,C_i(T). \tag{B1}$$

The terms which express the impact of *ET* on the water balance can be grouped together to give a more compact form

$$\frac{dC_i(T)}{dT} = (1-\alpha)\,f_{ET}(T)\,C_i(T), \tag{B2}$$

where the term $f_{ET}(T)$ indeed expresses how much of the precipitation input *i* is evapotranspired in time. Equation (B2) can be directly implemented to update the mass concentration of each volume stored within the catchment. However, the isotopic content is usually expressed not as a mass concentration, but as an abundance with respect to a standard in per mil units: $\delta = (R_{sample} - R_{stand})/R_{stand} \times 1000$, where R is the ratio of the heavy isotope to the lighter one [*Craig*, 1961]. An equivalent form of equation (B2) for the δ notation is then derived. Assuming that the amount of the abundant isotope in the sample is approximately equal to the one in the standard (which is a typical assumption) leads to a linear relationship between mass concentration and isotopic content in delta notation

$$\delta = 1000\left(\frac{C}{C_{stand}} - 1\right), \tag{B3}$$

where C_{stand} is the mass concentration of the isotope in the standard. By substitution of equation (B3) into equation (B2), one finally obtains an expression for the isotopic enrichment of the water pulse in δ notation

$$\frac{d\delta_i(T)}{dT} = (1-\alpha)\,f_{ET}(T)\,[\delta_i(T) + 1000]. \tag{B4}$$

Appendix C: Deviance Information Criterion (DIC)

The Deviance Information Criterion (DIC) [*Spiegelhalter et al*., 2002] is computed as

$$DIC = \bar{D} - p_D, \tag{C1}$$

where $\bar{D}$ is the posterior mean deviance and represents the "adequacy" of the model, while p_D expresses the degrees of freedom of the model and represents its complexity. The deviance is computed as

$$D(\theta) = -2\log\,\{p(y|\theta)\} + C, \tag{C2}$$

where θ is a set of parameter values, y represents the observed data, $p(y|\theta)$ is the model likelihood as computed from the MCMC calibration, and C is a constant that depends solely on the measured data so it is irrelevant to model selection. To estimate p_D, we use the expression suggested by *Gelman et al*. [2004]

$$p_D = p_V = \frac{1}{2}\,var(D), \tag{C3}$$

where *var*(*D*) is the variance of the posterior deviance distribution. The posterior deviance distribution and its statistics are directly computed from the results of the MCMC calibration. Other examples of the use of DIC in travel time model selection can be found in *Massoudieh et al*. [2013].

Acknowledgments
We are grateful to the European Research Council (ERC) VeWa project (GA335910) and NERC/JIP SIWA project (NE/MO19896/1) for funding. A.R. acknowledges the financial support from the ENAC school at EPFL. C.B. acknowledges support from the University of Costa Rica (project 217-B4-239 and the Isotope Network for Tropical Ecosystem Studies (ISONet)). Data to support this study are provided by the Northern Rivers Institute, University of Aberdeen and are available by the authors. The authors wish to thank Ype van der Velde, Arash Massoudieh, Jean-Raynald de Dreuzy, and an anonymous referee for the useful review comments.

References

Benettin, P., Y. van der Velde, S. E. A. T. M. van der Zee, A. Rinaldo, and G. Botter (2013), Chloride circulation in a lowland catchment and the formulation of transport by travel time distributions, *Water Resour. Res.*, *49*, 4619–4632, doi:10.1002/wrcr.20309.

Benettin, P., A. Rinaldo, and G. Botter (2015a), Tracking residence times in hydrological systems: Forward and backward formulations, *Hydrol. Processes*, *29*, 5203–5213, doi:10.1002/hyp.15034.

Benettin, P., S. W. Bailey, J. L. Campbell, M. B. Green, A. Rinaldo, G. E. Likens, K. J. McGuire, and G. Botter (2015b), Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH, USA, *Water Resour. Res.*, *51*, 9256–9272, doi:10.1002/2015WR017552.

Benettin, P., J. W. Kirchner, A. Rinaldo, and G. Botter (2015c), Modeling chloride transport using travel time distributions at Plynlimon, Wales, *Water Resour. Res.*, *51*, 3259–3276, doi:10.1002/2014WR016600.

Bertuzzo, E., M. Thomet, G. Botter, and A. Rinaldo (2013), Catchment-scale herbicides transport: Theory and application, *Adv. Water Resour.*, *52*, 232–242, doi:10.1016/j.advwatres.2012.11.007.

Beven, K. J. (2012), *Rainfall-Runoff Modelling: The Primer*, John Wiley and Sons, Ltd, Chichester, U. K., doi:10.1002/9781119951001.

Beyer, M., B. Jackson, C. Daughney, U. Morgenstern, and K. Norton (2016), Use of hydrochemistry as a standalone and complementary groundwater age tracer, *J. Hydrol.*, *543*(Part A), 127–144, doi:10.1016/j.jhydrol.2016.05.062.
Birkel, C., C. Soulsby, and D. Tetzlaff (2011), Modelling catchment-scale water storage dynamics: Reconciling dynamic storage with tracer-inferred passive storage, *Hydrol. Processes*, *25*(25), 3924–3936, doi:10.1002/hyp.8201.
Birkel, C., C. Soulsby, and D. Tetzlaff (2015), Conceptual modelling to assess how the interplay of hydrological connectivity, catchment storage and tracer dynamics controls nonstationary water age estimates, *Hydrol. Processes*, *29*, 2956–2969, doi:10.1002/hyp.10414.
Bishop, K., J. Seibert, S. Khler, and H. Laudon (2004), Resolving the double paradox of rapidly mobilized old water with highly variable responses in runoff chemistry, *Hydrol. Processes*, *18*(1), 185–189, doi:10.1002/hyp.5209.
Botter, G. (2012), Catchment mixing processes and travel time distributions, *Water Resour. Res.*, *48*, W05545, doi:10.1029/2011WR011160.
Botter, G., E. Bertuzzo, and A. Rinaldo (2010), Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, *Water Resour. Res.*, *46*, W03514, doi:10.1029/2009WR008371.
Botter, G., E. Bertuzzo, and A. Rinaldo (2011), Catchment residence and travel time distributions: The master equation, *Geophys. Res. Lett.*, *38*, L11403, doi:10.1029/2011GL047666.
Calabrese, S., and A. Porporato (2015), Linking age, survival, and transit time distributions, *Water Resour. Res.*, *51*, 8316–8330, doi:10.1002/2015WR017785.
Craig, H. (1961), Isotopic variations in meteoric waters, *Science*, *133*(3465), 1702–1703, doi:10.1126/science.133.3465.1702.
Dahlke, H. E., A. G. Williamson, C. Georgakakos, S. Leung, A. N. Sharma, S. W. Lyon, and M. T. Walter (2015), Using concurrent DNA tracer injections to infer glacial flow pathways, *Hydrol. Processes*, *29*(25), 5257–5274, doi:10.1002/hyp.10679, hYP-15-0084.R1.
Evaristo, J., S. Jasechko, and J. J. McDonnell (2015), Global separation of plant transpiration from groundwater and streamflow, *Nature*, *525*(7567), 91–94, doi:10.1038/nature14983.
Gat, J. R. (1996), Oxygen and hydrogen isotopes in the hydrologic cycle, *Annu. Rev. Earth Planet. Sci.*, *24*(1), 225–262, doi:10.1146/annurev.earth.24.1.225.
Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004), *Bayesian Data Analysis*, 2nd ed., Chapman and Hall, New York.
Godsey, S., and J. Kirchner (2014), Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order, *Hydrol. Processes*, *28*(23), 5791–5803, doi:10.1002/hyp.10310.
Gupta, V. K., E. Waymire, and C. T. Wang (1980), A representation of an instantaneous unit hydrograph from geomorphology, *Water Resour. Res.*, *16*(5), 855–862, doi:10.1029/WR016i005p00855.
Harman, C. J. (2015), Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, *Water Resour. Res.*, *51*, 1–30, doi:10.1002/2014WR015707.
Heidbuechel, I., P. A. Troch, S. W. Lyon, and M. Weiler (2012), The master transit time distribution of variable flow systems, *Water Resour. Res.*, *48*, W06520, doi:10.1029/2011WR011293.
Hrachowitz, M., H. Savenije, T. A. Bogaard, D. Tetzlaff, and C. Soulsby (2013), What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, *Hydrol. Earth Syst. Sci.*, *17*(2), 533–564, doi:10.5194/hess-17-533-2013.
Hrachowitz, M., O. Fovet, L. Ruiz, and H. H. G. Savenije (2015), Transit time distributions, legacy contamination and variability in biogeochemical $1/f$ scaling: how are hydrological response dynamics linked to water quality at the catchment scale?, *Hydrol. Processes*, *29*, 5241–5256, doi:10.1002/hyp.10546.
Hrachowitz, M., P. Benettin, B. M. V. Breukelen, O. Fovet, N. J. K. Howden, L. Ruiz, Y. V. D. Velde, and A. J. Wade (2016), Transit times the link between hydrology and water quality at the catchment scale, *WIREs Water*, *3*, 629–657, doi:10.1002/wat2.1155.
Kim, M., L. A. Pangle, C. Cardoso, M. Lora, T. H. M. Volkmann, Y. Wang, C. J. Harman, and P. A. Troch (2016), Transit time distributions and storage selection functions in a sloping soil lysimeter with time-varying flow paths: Direct observation of internal and external transport variability, *Water Resour. Res.*, *52*, 7105–7129, doi:10.1002/2016WR018620.
Kirchner, J. W. (2003), A double paradox in catchment hydrology and geochemistry, *Hydrol. Processes*, *17*(4), 871–874, doi:10.1002/hyp.5108.
Kirchner, J. W. (2016), Aggregation in environmental systems - Part 1: Seasonal tracer cycles quantify young water fractions, but not mean transit times, in spatially heterogeneous catchments, *Hydrol. Earth Syst. Sci.*, *20*(1), 279–297, doi:10.5194/hess-20-279-2016.
Klaus, J., C. E. Wetzel, N. Martnez-Carreras, L. Ector, and L. Pfister (2015), A tracer to bridge the scales: On the value of diatoms for tracing fast flow path connectivity from headwaters to meso-scale catchments, *Hydrol. Processes*, *29*(25), 5275–5289, doi:10.1002/hyp.10628.
Kolbe, T., J. Marais, Z. Thomas, B. W. Abbott, J.-R. de Dreuzy, P. Rousseau-Gueutin, L. Aquilina, T. Labasque, and G. Pinay (2016), Coupling 3D groundwater modeling with cfc-based age dating to classify local groundwater circulation in an unconfined crystalline aquifer, *J. Hydrol.*, *543*(Part A), 31–46, doi:10.1016/j.jhydrol.2016.05.020.
Lessels, J. S., D. Tetzlaff, C. Birkel, J. Dick, and C. Soulsby (2016), Water sources and mixing in riparian wetlands revealed by tracers and geospatial analysis, *Water Resour. Res.*, *52*, 456–470, doi:10.1002/2015WR017519.
Maloszewski, P., W. Rauert, P. Trimborn, A. Herrmann, and R. Rau (1992), Isotope hydrological study of mean transit times in an alpine basin (Wimbachtal, Germany), *J. Hydrol.*, *140*(1), 343–360, doi:10.1016/0022-1694(92)90247-S.
Massoudieh, A., N. Lu, X. Liang, T. H. Nguyen, and T. R. Ginn (2013), Bayesian process-identification in bacteria transport in porous media, *J. Contam. Hydrol.*, *153*, 78–91, doi:10.1016/j.jconhyd.2013.08.004.
McDonnell, J. J. (2014), The two water worlds hypothesis: Ecohydrological separation of water between streams and trees?, *WIREs Water*, *1*, 323–329, doi:10.1002/wat2.1027.
McDonnell, J. J., and K. J. Beven (2014), 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 Resour. Res.*, *50*, 5342–5350, doi:10.1002/2013WR015141.
McGuire, K. J., and J. J. McDonnell (2006), A review and evaluation of catchment transit time modeling, *J. Hydrol.*, *330*(3–4), 543–563, doi:10.1016/j.jhydrol.2006.04.020.
McGuire, K., M. Weiler, and J. McDonnell (2007), Integrating tracer experiments with modeling to assess runoff processes and water transit times, *Adv. Water Resour.*, *30*(4), 824–837, doi:10.1016/j.advwatres.2006.07.004.
Queloz, P., L. Carraro, P. Benettin, G. Botter, A. Rinaldo, and E. Bertuzzo (2015), Transport of fluorobenzoate tracers in a vegetated hydrologic control volume: 2. Theoretical inferences and modeling, *Water Resour. Res.*, *51*, 2793–2806, doi:10.1002/2014WR016508.
Rinaldo, A., and A. Marani (1987), Basin scale-model of solute transport, *Water Resour. Res.*, *23*(11), 2107–2118, doi:10.1029/WR023i011p02107.
Rinaldo, A., A. Marani, and A. Bellin (1989), On mass response functions, *Water Resour. Res.*, *25*(7), 1603–1617.
Rinaldo, A., P. Benettin, C. J. Harman, M. Hrachowitz, K. J. McGuire, Y. van der Velde, E. Bertuzzo, and G. Botter (2015), Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes, *Water Resour. Res.*, *51*, 4840–4847, doi:10.1002/2015WR017273.

Rodriguez-Iturbe, I., and J. B. Valdes (1979), The geomorphologic structure of hydrologic response, *Water Resour. Res.*, *15*(6), 1409–1420, doi:10.1029/WR015i006p01409.
Soulsby, C., D. Tetzlaff, and M. Hrachowitz (2009), Tracers and transit times: Windows for viewing catchment scale storage?, *Hydrol. Processes*, *23*(24), supplement, 3503–3507, doi:10.1002/hyp.7501.
Soulsby, C., C. Birkel, J. Geris, J. Dick, C. Tunaley, and D. Tetzlaff (2015), Stream water age distributions controlled by storage dynamics and nonlinear hydrologic connectivity: Modeling with high-resolution isotope data, *Water Resour. Res.*, *51*, 7759–7776, doi:10.1002/2015WR017888.
Soulsby, C., C. Birkel, and D. Tetzlaff (2016a), Characterizing the age distribution of catchment evaporative losses, *Hydrol. Processes*, *30*(8), 1308–1312, doi:10.1002/hyp.10751.
Soulsby, C., J. Bradford, J. Dick, J. P. McNamara, J. Geris, J. Lessels, M. Blumstock, and D. Tetzlaff (2016b), Using geophysical surveys to test tracer-based storage estimates in headwater catchments, *Hydrol. Processes*, *30*, 4434–4445, doi:10.1002/hyp.10889.
Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002), Bayesian measures of model complexity and fit, *J. R. Stat. Soc., B*, *64*(4), 583–639, doi:10.1111/1467-9868.00353.
Sprenger, M., D. Tetzlaff, C. Tunaley, J. Dick, and C. Soulsby (2017), Evaporation fractionation in a peatland drainage network affects stream water isotope composition, *Water Resour. Res.*, doi:10.1002/2016WR019258, in press.
Stewart, M. K., and J. J. McDonnell (1991), Modeling base flow soil water residence times from deuterium concentrations, *Water Resour. Res.*, *27*(10), 2681–2693, doi:10.1029/91WR01569.
Stewart, M. K., U. Morgenstern, J. J. McDonnell, and L. Pfister (2012), The hidden streamflow challenge in catchment hydrology: A call to action for stream water transit time analysis, *Hydrol. Processes*, *26*(13), 2061–2066, doi:10.1002/hyp.9262.
ter Braak, C. J. F., and J. A. Vrugt (2008), Differential evolution Markov chain with snooker updater and fewer chains, *Stat. Comput.*, *18*(4), 435–446, doi:10.1007/s11222-008-9104-9.
Tetzlaff, D., C. Birkel, J. Dick, J. Geris, and C. Soulsby (2014), Storage dynamics in hydropedological units control hillslope connectivity, runoff generation, and the evolution of catchment transit time distributions, *Water Resour. Res.*, *50*, 969–985, doi:10.1002/2013WR014147.
Tetzlaff, D., J. Buttle, S. K. Carey, K. McGuire, H. Laudon, and C. Soulsby (2015), Tracer-based assessment of flow paths, storage and runoff generation in northern catchments: A review, *Hydrol. Processes*, *29*(16), 3475–3490, doi:10.1002/hyp.10412.
Tunaley, C., D. Tetzlaff, J. Lessels, and C. Soulsby (2016), Linking high-frequency doc dynamics to the age of connected water sources, *Water Resour. Res.*, *52*, 5232–5247, doi:10.1002/2015WR018419.
van der Velde, Y., P. J. J. F. Torfs, S. E. A. T. M. van der Zee, and R. Uijlenhoet (2012), Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, *Water Resour. Res.*, *48*, W06536, doi:10.1029/2011WR011310.
van der Velde, Y., I. Heidbchel, S. W. Lyon, L. Nyberg, A. Rodhe, K. Bishop, and P. A. Troch (2015), Consequences of mixing assumptions for time-variable travel time distributions, *Hydrol. Processes*, *29*(16), 3460–3474, doi:10.1002/hyp.10372.
van Huijgevoort, M. H. J., D. Tetzlaff, E. H. Sutanudjaja, and C. Soulsby (2016), Using high resolution tracer data to constrain water storage, flux and age estimates in a spatially distributed rainfall-runoff model, *Hydrol. Processes*, *30*, 4761–4778, doi:10.1002/hyp.10902.
Vrugt, J., C. T. Braak, C. Diks, B. Robinson, J. Hyman, and D. Higdon (2009), Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlinear Sci. Numer. Simul.*, *10*(3), 271–288, doi:10.1515/IJNSNS.2009.10.3.273.
Weiler, M., B. L. McGlynn, K. J. McGuire, and J. J. McDonnell (2003), How does rainfall become runoff? A combined tracer and runoff transfer function approach, *Water Resour. Res.*, *39*(11), 1315, doi:10.1029/2003WR002331.