manuscript gmd-2017-305

tran-SAS v1.0: a numerical model to compute catchment-scale hydrologic transport using StorAge Selection functions

Dear Editor,

we have modified the paper according to the comments provided by the referees. As the numerical model did not need any major change, we have not created a new model release. Following the suggestion by Dr. Gross, the current release (tran-SAS v1.0) has been assigned a zenodo DOI (https://doi.org/10.5281/zenodo.1203600) and this has been included in the code availability section. Overall, we are grateful to the three referees and to Dr. Gross for their constructive comments which helped improve our manuscript.

Yours sincerely,

Paolo Benettin and Enrico Bertuzzo

RESPONSE TO REFEREES' COMMENTS

In the following, referees' comments are reported in italic. Our responses follow point-by-point

REVIEWER 1

> *This is a technical paper documenting a model code built on previous developments. This is entirely in keeping with the scope of the journal. A link to open source code is provided (github). I found this paper exceptionally clearly presented. It's quite easy for the reader to understand what the model does. Illustrative graphics and sensible notational shorthand help in that. The numerical accuracy test is useful. It seems the authors have taken care of numerical efficiency.*

We thank Referee 1 for her/his very positive assessment of the paper.

> *I only had 2 comments related to references to more complex applications:*
>
> *"On p10, l2-3 a catchment with legacy agricultural inputs is put forward as an example of a diluting system such as that simulated here with a synthetic dataset. I don't think that's correct because agricultural inputs are subject to reactive transport, not just dilution, which is not implemented in this code. Please come up with a better example."*
>
> *"On p15, l25-26 the authors claim that reactive transport can be easily implemented. I would question this general statement as especially agricultural solute transport can be quite complex as dissolution, precipitation and re-mobilisation as well as spatial variables (e.g. temperature) matter greatly. Please limit this statement to "simple" reactive transport."*

We agree with these useful comments on reactive transport. Indeed, the transport of agricultural solutes (particularly nitrate and phosphorus) is far more complex than a simple dilution and we do not want to convey the idea that a conservative transport model is appropriate in those situations. Agricultural inputs were mentioned just as an example of input that, due to regulation, can undergo rather drastic reductions. We now refer explicitly to the case of chloride (Page 10 Lines 1-2), that can have an agricultural origin (see e.g. van der Velde et al., 2010 and Martin et al., 2004) and a substantially simpler biogeochemical cycling.

REVIEWER 2

> *This study introduces a numerical transport modeling package, tran-SAS, which is aimed to simulate solute transport and water residence time at a catchment scale. In particular, the authors explore the computational stability as well as the numerical accuracy of the proposed model. The manuscript is well written, easy to follow, and quite interesting to the hydrologic transport community. A few minor and relatively major points are listed below.*

We thank Reviewer 2 for the positive comments.

1) Section 1, Line 21: Please revise the first sentence. The new transport model has improved the capabilities in terms of what? I understand your point, but it worth it to make it clear for a general audience. A suggestion is to add two or three sentences on, e.g., how the new transport model can be expressed in different ways depending on the ease of its application in a desired study (Botter et al., GRL2011 vs. van der Velde et al., WRR2012 vs. Harman, WRR2015 vs. Benettin et al., WRR2017). Or, for instance, how this new transport model is much less biased to spatial aggregation as opposed to the traditional approaches assigning the TTD a priori (Danesh-Yazdi et al., GRL2017).

We think that an exhaustive discussion on the differences between SAS function parametrization methods already exists in the literature (see e.g. Harman, 2015), so we have modified our introduction to accommodate this Reviewer's suggestion, but limited to the comment that the SAS approach is less biased to spatial aggregation (P2 L22).

2) Section 2, Line 7: Characterization of the SAS function is also part of the requirements (as emphasized in section 2, line 21) for solving the age distribution of the water storage. Please revise this sentence, accordingly.

We have now specified all the requirements for the application of the approach at the beginning of Section 4 (Application Example).

3) Equation 3: Isn't this conditional on no precipitation takes place at time t? What about those conditions when part of the input precipitation falls directly into the river, contributing to the streamflow? Or what about those conditions when a major portion of the input precipitation contributes rapidly to the streamflow?

This boundary condition is coupled to equation (1) where precipitation $J(t)$ appears explicitly. For any age $\varepsilon > 0$ (even very small), integration of eq (1) results in $S_T(\varepsilon, t) > 0$ (if precipitation occurred during ε).

4) Section 2, Lines 20-24: S0 and k have been written in different formats in the manuscript (i.e., at one place as bold and italic, and at another place as normal). Please make them consistent throughout the manuscript.

We decided to remove the bold format and only use italic in the formulas. Thanks for pointing this out.

5) Section 2.4, title: I know in their former papers, the authors have already emphasized on the distinction between the random sampling and the well-mixed conditions. As such, I am not sure why they equivalently put them together in this title.

Indeed, there are differences between a "well-mixed" and a "randomly-sampled" reservoir and we do not imply that the two coincide. In a catchment-scale formulation this distinction has no practical consequences and equation (7) holds in both cases. As readers are usually more familiar with the concept of "well-mixed", we preferred to keep the section title as is.

6) Section 3, Line 18: You already called ST (0, t0) = 0 a "boundary" condition in Eq. (3).

This is true, but equation (8) is now an ordinary differential equation in the variable T, so the condition T=0 is formally an initial condition.

7) Section 3, Line 22: I am not following this last sentence.

The entire sentence (P6 L17-23) has been reformulated and in particular:
"Water entered after t_0 gradually replaces the water entered before t_0 and for very large T the solution reaches (asymptotically) the total storage in the system, as no water that had entered before t_0 is still present in the system."

8) Page 7, Line 12: Not sure what does 1 in ΩQ[1, j − 1] imply? It is essentially ST (i, j − 1), so you meant i instead of 1?

The term e[j] represents an estimate of the event water so it refers to the first element in the rank storage. As the indexes i and j go from 0 to N, we have now indicated the first element of the rank storage as ST[i = 0, j-1].

9) Page 10, Line 12; Page 11, Lines 1-2: This is an important conclusion, but with a relatively weak reasoning. The difference between the curves in Figure 3b after year 2 is not really 2 significant. Author might want to provide another example that clearly demonstrates this conclusion.

We agree with this comment. The conclusion is more of a general consequence of the young storage preference, but it is not well visible in Figure 3b. We have specified this (P10 L13-14) in the revised text.

10) Does the tran-SAS package also include the Markov Chain Monte Carlo calibration scheme (with reference to Page 15, Line 8)? If yes, please add a few lines on how such a scheme is embedded within the package. If no, why not to include?

The MCMC package could not be included in our package for copyright reasons, but the structure of the model function is fully compatible with the DREAM_ZS [ter Braak and Vrugt, 2008, Vrugt et al., 2009] software for matlab, freely available at http://faculty.sites.uci.edu/jasper/software/.

11) The examples include two different ways of parameterizing the SAS functions, that is, using the power-law and the gamma functions. However, there is no discussion about which model provides a better solution to TTD and CQ. This is a missing, but quite important information for the users of this model and should be well addressed in the manuscript.

The examples actually included parametrization using power-law or beta functions. It is not possible to tell *apriori* which function provides a better solution because it depends on the specific application and on the desired degree of complexity of the model. For example, the beta function is a more general case than the power-law, so in principle it provides more accurate solutions, but it also makes use of more parameters and it involves longer computations.

12) Section 1, Line 22: "such as" instead of "like"? Also, at Page 14, Line 13.

13) Section 2, Line 24: "expressed in terms of" instead of "expressed as"?

14) Section 2: You might want to define CQ(t) as well to complete your definitions here.

15) Page 11, Line 11: Define the acronym for the random sampling, i.e., RS, earlier in the manuscript where it was first mentioned.

Done, we thank Reviewer 2 for these corrections.

REVIEWER 3

The authors present a very useful Matlab implementation of the StorAge Selection modeling framework.

We thank Referee 3 for her/his positive evaluation of the model.

The implementation is essentially a solute transport model, because the solute concentrations are one of the state variables. With the SAS framework it is possible to calculate solute concentrations "offline", by storing the travel time distribution of stream flow for select (sampled) times, and multiplying these with the tracer input history. This is efficient for a small number of samples and a large number of tracers. Perhaps not a common case.

From the manuscript, it is not clear if the model supports simultaneous calculation of multiple solutes. Perhaps I missed that. It would be useful.

The present implementation of the SAS framework is chiefly oriented to modeling the time-evolution of one solute in a hydrologic system. Extending the code to the case of multiple solutes is an easy task because the water carrier (and its transit time distributions) remains the same. Hence, one only needs to duplicate the equations that involve solute transport (or re-run the code with modified initial and boundary conditions). We preferred to keep this basic code simple and intuitive, and let the user adapt the model to more advanced transport problems.

I would like to see a stronger encouragement by the authors to test the parameter space for each new case. The example parameters are very hypothetical.

We agree with Referee 3 that the parameter space should be widely explored and this has been stressed in the revised version (P9 L23-24, P10 L17-22 and P11 L1-2).

The model description is accurate and easy to understand (for someone who has worked with a different implementation of a SAS model). I hope one of the other reviewers is a "SAS dummy" who can ask the questions that seem obvious to me.

We thank Referee 3 for this positive comment as, indeed, we put quite some effort to make the code description and implementation easy to understand.

I have a few comments specific to the text:

P5 L19, Eq 6: This implementation is equivalent to the fractional StorAge Selection (fSAS) implementation, right?

Mathematically, equation (6) becomes a fSAS after the variable transformation $S_T(T,t) \rightarrow f(T,t) = S_T(T,t)/S(t)$.

Section 3.3: I would like the authors to elaborate on the discussion of the "old pool". Transient tracers like tritium and chlorine-36 demand that the age distribution of the old pool is accurately represented. Or at least in the concentration in the "old pool" needs to be represented.

We agree with Referee 3. The problem of what is to be considered as “old” also depends on the considered tracer and its characteristic input timescales (see Benettin *et al*., 2017a for a discussion on this). In the case of tracers like tritium and chlorine-36, a much longer spin-up is advised to limit the impact of *apriori* assumptions on the initial old pool concentration. Also in this case a much longer timestep (e.g. weeks) could be used in the computations. We have expanded the discussion on this at P8 L17-19 and P9 L2-3.

> *P9 L6: "long term" = 4 years? P9 L13: S0=1000? mm? P9 L11: I understand the parameterization of the example is not intended to represent the hydrogeological conditions of the particular data set. Nevertheless, I find the random sample (kET=1) surprising, as I would expect the vegetation to have even the slightest preference for younger water. Perhaps the authors can warn the reader that these parameters should not be considered "valid" for any catchment and encourage the user of the tranSAS to vary all parameters of the example case drastically if applied to a specific setting to test the sensitivity.*

As mentioned in previous comments, we fully agree with Referee 3 on this point. These were just hypothetical parameters (although they are similar to parameters found in small catchments in wet climates, e.g. Benettin et al., 2017b) and should not be taken as representative of a general catchment behavior. This has been clarified in the revised manuscript (P9 L23-24, P10 L17-22 and P11 L1-2). We also believe that, thanks to the short computational times, the *tran*SAS code facilitates sensitivity analyses.

> *Figure 3d, please clarify that this is the stream flow TTD.*

> *P10 L1: "solutes with a yearly period".... like stable isotopes of water? (These aren't really solutes.)*

Done, thank you.

> *P10 L8: The range in median ages can vary much more. It all depends on the fictional parameters you enter into your model. It might be more relevant to compare the nonrandom-sampling cases with the random-sampling case. Or reiterate that any power with a k < 1 prefers younger water and will therefore have a younger TTDs (right? or is this not alwyas the case?)*

Age estimates are typically more sensitive to model parameters than solute concentration estimates. This has been specified within the expanded discussion on the sensitivity of model results (P10 L17-22 and P11 L1-2). The relationship between the age distribution and the value of parameter k is not straightforward as it also depends on which portion of the age distribution is considered.

> *P10 L10: This dilution example is interesting. Is it true that the stream solute concentration is the inverse of the TTD in the random sampling case (k=1)? It might be worth mentioning.*

We did not fully understand this comment.

> *The inverse problem, a step increase of a contaminant input relates more directly to the TTD. I do like this example because it is more optimistic about the potential to reduce environmental contamination. And it illustrates an important aspect of transient contaminant flow, that even with zero input, stream concentrations can increase due to the variable hydrology.*

We thank Referee 3 for this feedback.

> *P15 L6: "less than a second" for a 4 year time series? How much longer does the ode113 solution take?*

On an ordinary PC, the test-case implementation (4 years spin-up + 4 years run, power-law SAS functions with k=0.7, 24-hour timestep) runs in less than a second for the modified Euler Scheme and in about 30 seconds for the ode113 solution.

> *P15 L28: "chronology of the inputs is irrelevant" Not quite sure how to interpret this. The chronology of a constant input decaying tracer (e.g. tritium for the last 30 years) is irrelevant, in the sense that it doesn't matter "when" the precipitation entered the catchment, but it does matter "how long ago". I know what is menat, but it reads like this model is only relevant for tracers with input fluctuations, which isn't the case (as long as the tracer decays on relevant time scales).*

In our view, the impact of input "chronology" is twofold: it expresses the time-variability of the input and it also determines the residence time of the input in the system (traditionally seen as the interval between present time and entrance time). In this paragraph we wanted to warn the reader that sometimes solute concentration can be driven by factors that do not depend on when the input entered the system nor on how long it remained in the system. We have better specified this second point (P16 L22).

REFERENCES:

Benettin, P., Bailey, S. W., Rinaldo, A., Likens, G. E., McGuire, K. J., & Botter, G. (2017a). Young runoff fractions control streamwater age and solute concentration dynamics. Hydrological Processes, 31(16), 2982–2986. https://doi.org/10.1002/hyp.11243

Benettin, P., Soulsby, C., Birkel, C., Tetzlaff, D., Botter, G., & Rinaldo, A. (2017b). Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments. Water Resources Research, 53(3), 1864–1878. https://doi.org/10.1002/2016WR020117

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

Martin, C., Aquilina, L., Gascuel-Odoux, C., Molénat, J., Faucheux, M. and Ruiz, L. (2004), Seasonal and interannual variations of nitrate and chloride in stream waters related to spatial and temporal patterns of groundwater concentrations in agricultural catchments. Hydrol. Process., 18: 1237–1254. doi:10.1002/hyp.1395

ter Braak, C. J. F., & Vrugt, J. a. (2008). Differential Evolution Markov Chain with snooker updater and fewer chains. Statistics and Computing, 18(4), 435–446. https://doi.org/10.1007/s11222-008-9104-9

van der Velde, Y., de Rooij, G. H., Rozemeijer, J. C., van Geer, F. C., & Broers, H. P. (2010). Nitrate response of a lowland catchment: On the relation between stream concentration and travel time distribution dynamics. Water Resources Research, 46(11). https://doi.org/10.1029/2010WR009105

Vrugt, J. A., ter Braak, C. J. F., Diks, C., Robinson, B. A., Hyman, J. M., & Higdon, D. (2009). Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. International Journal of Nonlinear Sciences and Numerical Simulation, 10(3), 271–288. https://doi.org/10.1515/IJNSNS.2009.10.3.273

tran-SAS v1.0: a numerical model to compute catchment-scale hydrologic *tran*sport using StorAge Selection functions

Paolo Benettin[1] and Enrico Bertuzzo[2]

[1]Laboratory of Ecohydrology ENAC/IIE/ECHO, École Polytechinque Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
[2]Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, Venice, Italy

Correspondence to: Paolo Benettin (paolo.benettin@epfl.ch)

Abstract. This paper presents the '*tran*-SAS' package, which includes a set of codes to model solute transport and water residence times through a hydrological system. The model is based on a catchment-scale approach that aims at reproducing the integrated response of the system at one of its outlets. The codes are implemented in MATLAB and are meant to be easy to edit, so that users with minimal programming knowledge can adapt them to the desired application. The problem of large-scale solute transport has both theoretical and practical implications. On one side, the ability to represent the ensemble of water flow trajectories through a heterogeneous system helps unraveling streamflow generation processes and allows making inferences on plant-water interactions. On the other side, transport models are a practical tool that can be used to estimate the persistence of solutes in the environment. The core of the package is based on the implementation of an age Master Equation (ME), which is solved using general StorAge Selection (SAS) functions. The age ME is first converted into a set of ordinary differential equations, each addressing the transport of an individual precipitation input through the catchment, and then it is discretized using an explicit numerical scheme. Results show that the implementation is efficient and allows the model to run in short times. The numerical accuracy is critically evaluated and it is shown to be satisfactory in most cases of hydrologic interest. Additionally, a higher-order implementation is provided within the package to evaluate and, if necessary, to improve the numerical accuracy of the results. The codes can be used to model streamflow age and solute concentration, but a number of additional outputs can be obtained by editing the codes to further advance the ability to understand and model catchment transport processes.

1 Introduction

The field of hydrologic transport focuses on how water flows through a watershed and mobilizes solutes towards the catchment outlets. The proper representation of transport processes is important for a number of purposes such as understanding streamflow generation processes (Weiler et al., 2003; McGuire and McDonnell, 2010; McMillan et al., 2012), modeling the fate of nutrients and pollutants (Jackson et al., 2007; Hrachowitz et al., 2015), characterizing how watersheds respond to change (Kauffman et al., 2003; Oda et al., 2009; Danesh-Yazdi et al., 2016; Wilusz et al., 2017) and estimating solute mass export to stream (Destouni et al., 2010; Maher, 2011). The spatiotemporal evolution of a solute is typically expressed (Rinaldo and Marani, 1987; Hrachowitz et al., 2016) as a combination of displacements, due to the carrier motion, and biogeochemical reactions, due to the interactions with the surrounding environment.

Water trajectories within a catchment are usually considered from the time water enters as precipitation to the time it leaves as discharge or evapotranspiration. As watersheds are heterogeneous and subject to time-variant atmospheric forcing, water flowpaths have marked spatiotemporal variability. For this reason, a formulation of transport by travel time distributions (see Cvetkovic and Dagan, 1994; Botter et al., 2005) can be particularly convenient as it allows transforming complex 3D trajectories into a single variable: the travel time, i.e. the time elapsed from the entrance of a water particle to its exit.

While early catchment-scale approaches (see McGuire and McDonnell, 2006) focused on the identification of an appropriate shape for the travel time distributions (TTD), emphasis has recently been put on a new generation of catchment-scale transport models, where TTDs result from a mass balance equation rather than being assigned *a priori* (Botter et al., 2011). As a concequence, TTDs change through time, as observed experimentally (e.g. Queloz et al., 2015a; Kim et al., 2016) and as required for consistency with mass conservation. This approach has the advantage of being consistent with the observed hydrologic fluxes and follows from the formulation of an age Master Equation (ME) (Botter et al., 2011), describing the age-time evolution of each individual precipitation input after entering the catchment. The key ingredient of this new approach is the "StorAge Selection" (SAS) function, which describes how storage volumes of different ages contribute to discharge (and evapotranspiration) fluxes. The direct use of SAS functions has already provided insights on water age in headwater catchments (van der Velde et al., 2012, 2015; Harman, 2015; Benettin et al., 2017b; Wilusz et al., 2017), intensively managed landscapes (Danesh-Yazdi et al., 2016), lysimeter experiments (Queloz et al., 2015b; Kim et al., 2016), reach-scale hyporheic transport (Harman et al., 2016), and it has also been applied to non-hydrologic systems like bird migrations (Drever and Hrachowitz, 2017). In principle, applications can be extended to any system where the chronology of the inputs plays a role in the output composition.

The new theoretical formulation has improved capabilities, including being less biased to spatial aggregation (Danesh-Yazdi et al., 2017) as opposed to traditional methods like the lumped convolution approach (e.g. Maloszewski and Zuber, 1993), but the numerical implementation of the governing equations is more demanding~~than in traditional methods like the lumped convolution approach (e.g. Maloszewski and Zuber, 1993)~~. This can represent a barrier to the diffusion of the new models, preventing their widespread use in transport processes investigation. To make the use of the new theory more accessible, the *tran*-SAS package includes a basic numerical model that solves the age ME using arbitrary SAS functions. The model is developed to simulate the transport of tracers in watershed systems, but it can be extended to other hydrologic systems (e.g. water circulation in lakes and oceans). The numerical code is written in MATLAB and it is intended to be intuitive and easy to edit, hence minimal programming knowledge should be sufficient to adapt it to the desired application.

The specific objectives of this paper are: i) provide a numerical model that solves the Age Master Equation with any form of the SAS functions in a computationally efficient way, ii) show the potential of the model for simulating catchment-scale solute transport, and iii) assess the numerical accuracy of the model for different aggregation timesteps.

2 Model Description

The model implemented in *tran*-SAS solves the age ME by means of general SAS functions and uses the solution to compute the concentration of an ideal tracer (conservative and passive to vegetation uptake) in streamflow. The model is described here using hydrologic terminology and applications.

2.1 Definitions

The general theoretical framework relies on the works by Botter et al. (2011); van der Velde et al. (2012); Harman (2015); Benettin et al. (2015b)~~and requires knowledge of the input/output water fluxes to/from the catchment and the initial water storage. Moreover, to compute the evolution of solute concentration in the storage and in the out-fluxes, the input solute concentration must be known. We~~ . Here, we consider a typical hydrologic system with precipitation $J(t)$ as input and evapotranspiration $ET(t)$ and streamflow $Q(t)$ as outputs. The total system storage is obtained as $S(t) = S_0 + V(t)$ where S_0 is the initial storage in the system and $V(t)$ are the storage variations obtained from the hydrologic balance equation $dV/dt = J - ET - Q$. ~~Tracer concentration in precipitation is indicated as $C_J(t)$.~~

The system state variable is the age distribution of the water storage. Indeed, at any time t, the water storage is comprised of precipitation inputs that occurred in the past and that have not left the system yet. Each of these past inputs can be associated with an age T, representing the time elapsed since its entrance into the watershed. Hence, at any time t the storage is characterized by a distribution of ages $p_S(T,t)$. Similarly, discharge and evapotranspiration fluxes are characterized by age distributions $p_Q(T,t)$ and $p_{ET}(T,t)$, respectively. Each water parcel in storage can also be characterized by its solute concentration $C_S(T,t)$, which in case of an ideal tracer is equal to the concentration of precipitation upon entering the catchment $C_J(t-T)$. Tracer concentration in streamflow is indicated as C_Q. A useful, transformed version of the storage age distribution is the rank storage S_T which is defined as $S_T(T,t) = S(t)\int_0^T p_S(\tau,t)d\tau$ and represents the volume in storage younger than T at time t.

The key element of the formulation is the SAS function, which formalizes the functional relationship between the age distribution of the system storage and that of the outflows. Different forms have been proposed to express the SAS function directly as a function of age or as a derived distribution of the storage age distribution, (e.g. *absolute*, *fractional* or *ranked* SAS functions, see Harman (2015)). For numerical convenience, SAS functions are here expressed ~~as~~ in terms of cumulative probability distributions (CDF) of the rank storage, for both discharge ($\Omega_Q(S_T,t)$) and evapotranspiration ($\Omega_{ET}(S_T,t)$). Namely, $\Omega_Q(S_T,t)$ is, at any time t, the fraction of total discharge which is produced by $S_T(T,t)$. Hence, it is equal to the fraction of discharge younger than T. The corresponding probability density functions are indicated as $\omega_Q(S_T,t)$ and $\omega_{ET}(S_T,t)$. Main model variables are illustrated in Figure 1.

2.2 The Age Master Equation

The age ME (Botter et al., 2011) can be seen as a hydrologic balance applied to every parcel of water stored in the catchment. Two different equations can be formulated, that describe the forward-in-time or the backward-in-time process (Benettin et al.,

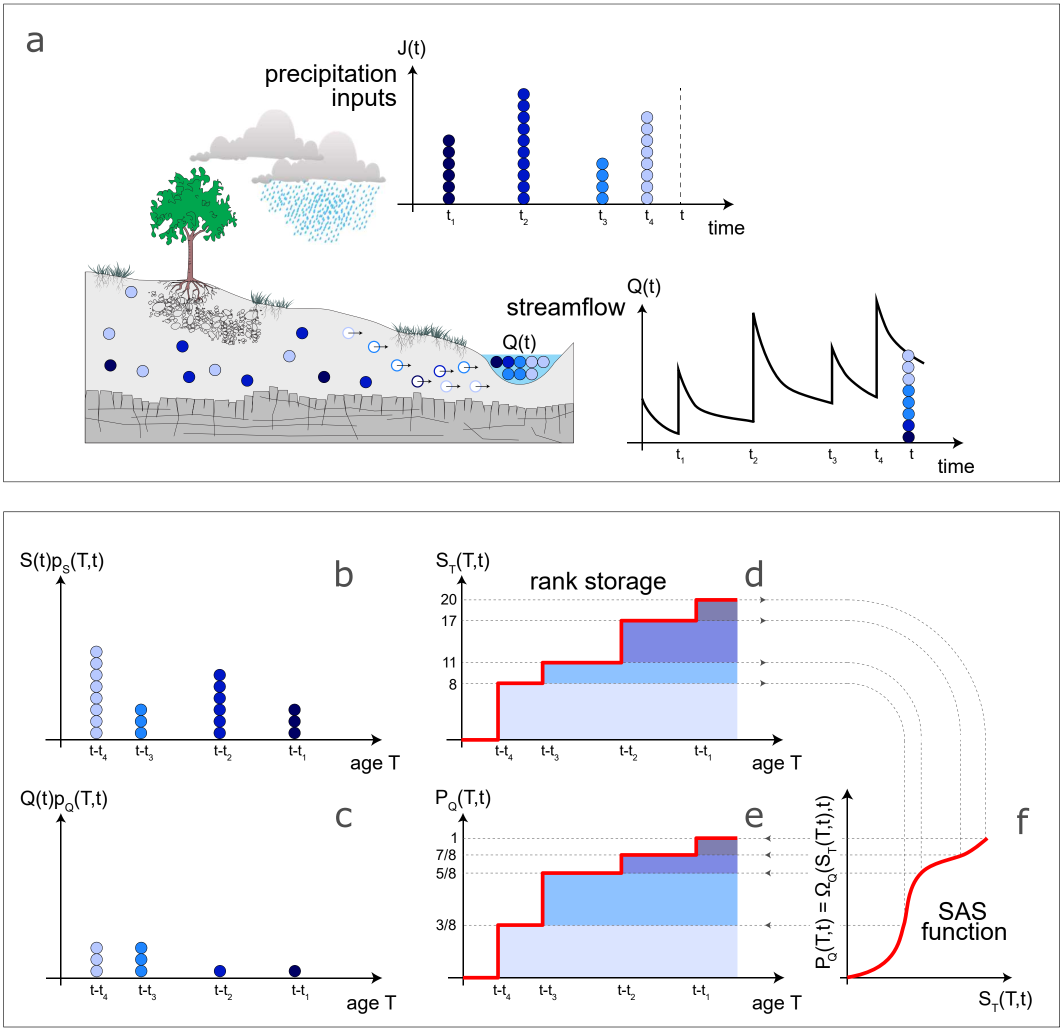

Figure 1. Conceptual illustration of the main variables of the theoretical formulation. Precipitation volumes are represented through coloured circles, with darker colours indicating the older precipitations with respect to current time t. Due to transport and mixing processes, precipitation volumes are retained in the catchment storage and released to streamflow (plot a). Both the catchment storage and its outfluxes are characterized by a distribution of ages (plots b and c). For example, the youngest water (age $t-t_4$, light blue colour) accounts for 8/20 of the storage and 3/8 of streamflow. By cumulating such distributions one gets the rank storage $S_T(T,t)$ and the cumulative discharge age distribution $P_Q(T,t)$ (plots d and e, red lines). The relationship between $S_T(T,t)$ and $P_Q(T,t)$ is quantified by the SAS function $\Omega_Q(S_T,t)$ (plot f).

2015b; Calabrese and Porporato, 2015; Rigon et al., 2016). Here, we focus on the backward form, as it is the most convenient to model solute concentration in streamflow. The backward form of the ME can be written in a number of equivalent forms that have been proposed in the literature (e.g. Botter et al., 2011; van der Velde et al., 2012; Harman, 2015). Here, we employ the cumulative version, which has a less intuitive physical interpretation but a better suitability to numerical implementation. The complete set of equations reads:

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

$$\text{Initial Condition: } S_T(T,t=0) = S_{T_0}\,, \tag{2}$$

$$\text{Boundary Condition: } S_T(T=0,t) = 0\,, \tag{3}$$

where the initial condition S_{T_0} indicates some initial distribution of the rank storage at time 0. Note that to ensure that p_S, p_Q and p_{ET} are distributions over the age domain $(0,+\infty)$, the SAS functions must verify the condition $\Omega_Q(S_T \to S(t),t) = \Omega_{ET}(S_T \to S(t),t) = 1$. This condition, however, is automatically verified as the SAS functions were defined as CDFs.

The solution of equation (1) gives the rank storage $S_T(T,t)$, from which the discharge age distributions $p_Q(T,t)$ can be obtained as:

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

where $P_Q(T,t)$ is the cumulative distribution of $p_Q(T,t)$ and $P_Q(T,t) = \Omega_Q(S_T,t)$ by definition. Stream solute concentration $C_Q(t)$ follows from:

$$C_Q(t) = \int_0^\infty C_S(T,t)p_Q(T,t)dT\,. \tag{5}$$

The same reasoning applies to the age distributions and concentration of the evapotranspiration flux.

2.3 The SAS functions

As explained in section 2.1, SAS functions are CDF's over the finite interval $[0,S(t)]$. A simple class of probability distributions that is suitable to serve as SAS function is the power-law distribution (Queloz et al., 2015b; Benettin et al., 2017b), which takes the form:

$$\Omega(S_T,t) = \left[\frac{S_T(T,t)}{S(t)}\right]^{\cancel{\mathbf{k}}\,k} = \left[\cancel{\frac{S_T(T,t)}{\mathbf{S_0}+V(t)}}\;\frac{S_T(T,t)}{S_0+V(t)}\right]^{\cancel{\mathbf{k}}\,k} \tag{6}$$

The parameter ~~$\mathbf{k} \in (0,+\infty)$~~ $k \in (0,+\infty)$ controls the affinity of the outflow for relatively younger/older water in storage. Specifically, $k<1$ [$k>1$] implies affinity for young [old] water, whereas the case $k=1$ represents "random sampling", i.e. outfluxes select water irrespective of its age. ~~$\mathbf{k}$~~ k can be conveniently made time-variant (e.g. dependent on the system wetness) to account for possible changes in the properties of the system (see van der Velde et al., 2015; Harman, 2015). Equation (6) also requires knowledge of the initial storage in the system ~~$\mathbf{S_0}$~~ S_0, which can be difficult to estimate experimentally and it is often treated as a calibration parameter. When using power-law SAS functions for both Q and ET, the system only requires 3 calibration parameters: k_Q, k_{ET} and S_0. Different classes of probability distributions can be used to have more flexibility in the SAS function shape, e.g. the beta (van der Velde et al., 2012; Drever and Hrachowitz, 2017) or the Gamma (Harman, 2015;

Wilusz et al., 2017) distributions. Such functions can be more difficult to implement numerically, but they are usually available in software libraries.

2.4 The special case of well-mixed/random-sampling

In case all the outflows remove the stored ages proportionally to their abundance, the outflow age distributions become a perfect sample (or *random* sample, RS) of the storage age distribution. The SAS functions in this case assume the linear form $\Omega_Q(S_T,t) = \Omega_{ET}(S_T,t) = S_T(T,t)/S(t)$ and equation (1) has analytical solution (Botter, 2012):

$$p_S(T,t) = p_Q(T,t) = \frac{J(t-T)}{S(t)} \exp\left[-\int_{t-T}^{t} \frac{Q(\tau)+ET(\tau)}{S(\tau)}\, d\tau\right] \tag{7}$$

Equation (7) can be seen as a generalization of the linear reservoir equation to fluctuating storage. Indeed, in the special case of a stationary system, where $J = Q + ET$ and the ratio J/S is a constant c, equation (7) takes the simple form $p_S(T) = c\,\exp(-cT)$.

3 Model Implementation

3.1 Problem Discretization

Equation (1) does not have exact solution, except for the particular case of randomly sampled storage (section 2.4), so in general a numerical implementation is required. Following the approach by Queloz et al. (2015b) and Harman (2015), the partial differential equation (1) is first converted into a set of ordinary differential equations using the method of characteristics. Indeed, along a characteristic line of the type $t = T + t_0$, equation (1) simplifies into an ordinary differential equation in the single variable T:

$$\frac{dS_T(T,T+t_0)}{dT} = J(T+t_0) - Q(T+t_0)\,\Omega_Q(S_T,T+t_0) - ET(T+t_0)\,\Omega_{ET}(S_T,T+t_0)\,, \tag{8}$$

with initial conditions $S_T(0,t_0) = 0$. In this context, reformulating the problem along characteristic lines means following the variable $S_T(T,T+t_0)$, i.e. the fraction of storage younger than the water input entered in t_0. This can be equally interpreted as the amount of water storage entered after time t_0. The solution $S_T(T,T+t_0)$ starts from the value 0, corresponding to the initial time t_0~~when the input enters~~. Then, as time (and age) grows, $S_T(T,T+t_0)$ increases when precipitation J introduces younger water into the system and decreases when out-fluxes Q and ET withdraw water younger than T. ~~The solution eventually~~ Water entered after t_0 gradually replaces the water entered before t_0 and for very large T the solution reaches (asymptotically) the total storage in the system, as ~~all the water is younger than the input~~ no water that had entered ~~in~~ before t_0 is still present in the system.

We discretize time and age using the same time steps $\Delta T = \Delta t = h$, resulting in $T_i = i\cdot h$ and $t_j = j\cdot h$, with $i,j \in \mathbb{N}$ and we use the convention that the discrete variables T_i and t_j refer to the beginning of the timestep. To simplify the notation,

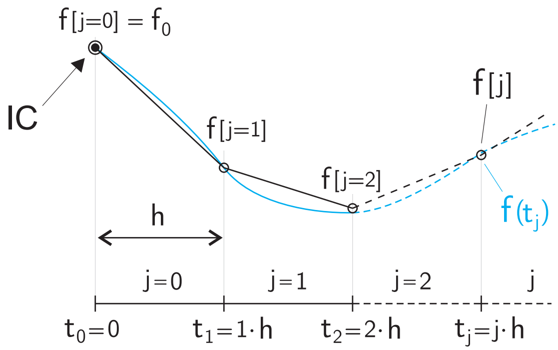

Figure 2. Illustration of the conventions used to discretize the time domain. Time steps have a fixed length h (e.g. 12 hours) and each time step j starts in $t_j = j \cdot h$. The numerical evaluation of a function f at time t_j is indicated as $f[j]$.

square brackets are used to indicate the numerical evaluation of a function and the indexes i and j are used for T_i and t_j respectively. For example, $f[i,j]$ indicates the numerical evaluation of function $f(T_i, t_j)$. The conventions used for the discretization are illustrated in Figure 2. For numerical convenience and because real-world data often represent an average over a certain time-interval, all fluxes (J, Q, ET) are considered as averages over the time step h (e.g., $J[j] = 1/h \int_{jh}^{(j+1)h} J(\tau)d\tau$). As a consequence, storage variations obtained from a hydrologic balance are linear during a timestep and each value refers to the beginning of the timestep.

To solve equation (8), we implement a forward Euler scheme. This explicit numerical scheme is intuitive and fast to solve, and its numerical accuracy is shown to be satisfactory for many hydrologic applications (see model verification, section 5.1). By terming $\Omega[i,j] = \Omega(S_T[i,j], t_j)$, the discretized problem becomes:

$$S_T[i+1, j+1] = S_T[i,j] + h \cdot (J[j] - Q[j]\,\Omega_Q[i,j] - ET[j]\,\Omega_{ET}[i,j]) \tag{9}$$

for $i, j \in [0, N]$, with N indicating the number of timesteps in the simulation, and boundary condition $S_T[0,j] = 0$. In a pure forward Euler scheme, this boundary condition implies that $\Omega[0,j] = \Omega(0, t_j) = 0$, meaning that no input can be part of an output during the same timestep. This can be a limitation for catchment applications, where "event" water is often not negligible and it can bear important information on catchment form and function. For this reason, in equation (9) we use a modified Ω^* defined as:

$$\Omega^*[i,j] = \Omega(S_T[i,j] + e[j], t_j) \tag{10}$$

where $e[j]$ is an estimate of the youngest water stored in the system at the end of time step j. Such an estimate is here obtained as ~~$e[j] = \max(0, J[j] - Q[j]\,\Omega_Q[1, j-1] - ET[j]\,\Omega_{ET}[1, j-1]$~~$e[j] = \max(0, J[j] - Q[j]\,\Omega_Q[0, j-1] - ET[j]\,\Omega_{ET}[0, j-1]$, i.e. it is a water balance for current precipitation input using the SAS functions evaluated at previous timestep. The classic Euler scheme is returned if $e[j] = 0$. This modification of the classic numerical scheme only affects the behavior of the youngest age in the system and it is a simple and efficient way to account for transport of event water. The accuracy of this numerical scheme is evaluated in Section 5.1.

3.2 Numerical routine

The model solves equation (9) by implementing an external for-loop on j (i.e. on the chronologic time) and an internal for-loop on i (i.e. on the ages). This means that during one timestep j all the characteristic curves (equation (9)) are updated by one timestep. The internal loop is implemented using vector operations. The vector length is indicated as n_j and it depends on the number of age classes (which is also the number of characteristic curves) that are included in the computations at time j (see section 3.3). At any time step, the two fundamental operations to solve the discretized ME are:

– compute $\Omega^*_Q[i,j]$ and $\Omega^*_{ET}[i,j]$ using equation (10);

– compute $S_T[i,j]$ using equation (9) for $i \in [1, n_j]$;

To compute the model output, further operations are required. In particular:

– update $C_S[i,j] = C_J[i-j]$, valid for conservative solutes entering through precipitation

– compute $p_Q[i,j] \cdot h = \Omega_Q[i,j] - \Omega_Q[i-1,j]$;

– compute $C_Q[j] = \sum_{i=1}^{n_j} C_S[i,j] \cdot p_Q[i,j] \cdot h$;

Starting from these basic routines, many additional operations can be implemented, to e.g. characterize the non-conservative behavior of solutes or to compute some age distribution statistics.

3.3 Additional numerical details

A first issue that the model needs to take into account is that age distributions are defined over an age domain $[0, +\infty)$, meaning that the rank storage is made of an infinite number of elements where the oldest elements typically represent infinitesimal stored volumes. To have a finite number of elements in the computations, an arbitrary old fraction of rank storage can be considered as a single undifferentiated volume of "older" water. This allows merging a high number of very little residual volumes into a single "old" pool. ~~Such~~ Note that the term "old" should be used carefully as its definition depends on the particular system under consideration and it may differ depending on the characteristic timescales of the solute used to infer water age (Benettin et al., 2017a). The old pool is here defined as the volume $S_T(T,t) > S_{th}$, where S_{th} is a numerical parameter that can be fixed for each different application. S_{th} also defines the age T_{th}, corresponding to $S(T = T_{th}, t) = S_{th}$, which indicates the oldest age that is computed individually. Numerically, the parameter S_{th} controls the number n_j of age classes (or equivalently rank storage volumes) that are taken into account in the computations. S_{th} should be chosen so that the number of elements used in the computations remains small but the numerical accuracy is not compromised. It can be convenient to define a non-dimensional threshold $f_{th} \in [0,1]$ such that $S_{th} = f_{th}\, S(t)$. In this case, a value $f_{th} = 0.9$ means that the old pool comprises the oldest 10% of the water storage. When $f_{th} = 1$ no old pool is taken into account. Once a storage element is merged to the old pool, its individual age and concentration properties cannot be retrieved, but the mean properties of the old pool like the mean solute concentration are preserved.

A second, connected problem regards the initial conditions of the system, i.e. the unknown storage age distribution and solute concentration to be used at the beginning of the calculations. In the absence of information, the initial storage can be considered as one single old pool, hence the initial number of age classes n_0 is equal to 1. Once computations start, new elements are introduced and accounted for in the balance, reducing the impact and the influence of the initial conditions. The old pool gets progressively smaller (and vector length n_j larger) until it reaches the stationary value defined by S_{th}. An initial spinup period can be used to initialize the ME balance and reduce the size of the initial old water pool. This is particularly indicated when modeling solutes with long turnover times like tritium. The influence of the initial conditions decreases with time, but given the long timescales that may characterize transport processes, it is likely never completely exhausted. This has little impact on the output concentration but it limits the maximum computable age to the time elapsed since the start of the simulation.

The computational time of a simulation can be reduced by not accounting for zero-precipitation inputs as they have no influence in the balance but increase the number of operations required at each time step. In such a case, however, the position of an element in the vector does not correspond with its age anymore and age has to be counted separately. To keep the model intuitive, we decided to not remove zero-precipitation inputs.

4 Application Example

Application of the approach requires knowledge of the input/output water fluxes to/from the catchment, the input solute concentration and the initial conditions on the water storage magnitude and concentration. Then, a SAS function must be specified for each outflow. The code comes with example virtual data that can be used to evaluate the model capabilities. Four years of hydrologic data were obtained from recorded precipitation and streamflow at the Mebre-Aval station near Lausanne (CH). Evapotranspiration was obtained from regional daily estimates around the Lausanne area and modified to match the long-term mass balance. On average, yearly precipitation is 1100 mm, discharge is 580 mm (53$\%$ of precipitation) and evapotranspiration is 520 mm. The storage variations, computed by solving the hydrologic balance, were normalized to the interval [0,1] to serve as a non-dimensional metric of catchment wetness (variable wi). Overall, the data are not meant to be representative of a particular location, but they constitute a realistic set of hydrologic variables to test the model.

The code was run on the example data using the 4 illustrative shapes for the discharge SAS function listed in Table 1. All simulations share the following settings: 12-h timestep, 4-year spinup period obtained by repeating the example data, storage threshold f_{th}=1 (i.e., no old-pool schematization), initial storage parameter S_0=1000, evapotranspiration SAS function selected as a power law with parameter k=1 (equivalent to a random sampling). The different shapes for the discharge SAS function were selected to test different functional forms (power law, power law time variant, beta distribution) and to illustrate the transition from the preferential release of younger water volumes (examples ω_1 and ω_2) to the random sampling case (ω_3) and the preferential release of older waters (ω_4). The time-variant power-law SAS (ω_1) was obtained by using equation (6) with a time-variant exponent $k(t) = k_{Q1} + [1 - wi(t)]\,(k_{Q2} - k_{Q1})$, with parameters k_{Q1} and k_{Q2} corresponding to the exponent k during the wettest ($wi = 1$) and driest ($wi = 0$) conditions. This parameter choice is used for illustration purposes and should not be taken as representative of a general catchment behavior.

Table 1. Description of the discharge SAS functions used in the application. All the functions were tested with the same initial total storage S_0=1000 mm.

name	type	parameters	value
ω_1	power law time variant	k_{Q1}	0.3
		k_{Q2}	0.9
ω_2	power law	k_Q	0.7
ω_3	random sampling	-	-
ω_4	beta	a	1.5
		b	0.8

Two different examples of solute transport were simulated in the test. In the first case, solute input concentration was generated by adding noise to a sinusoidal wave with annual cycle. This example can be representative of atmospheric ~~solutes~~ tracers with a yearly period (like stable water isotopes). In the second case, the initial storage was set to a concentration of 100 mg/l and any subsequent input was assigned a concentration of 0 mg/l, causing the system to dilute. This example can be representative of a diluting system, e.g. a catchment with ~~agricultural inputs~~ conservative agricultural inputs like chloride (Martin et al., 2004; van der Velde et al., 2010) that undergoes a step reduction. Results of both examples are shown in Figure 3.

Each discharge SAS function simulates different transport mechanisms and provides rather different outputs, both in terms of water age and streamflow concentration. In the first solute transport example (Figure 3a), discharge concentration gets progressively damped and shifted as the SAS function moves from younger-water preference to older-water preference. The travel time distributions extracted for February 15th, 2016 (Figure 3d) show that the median age of streamflow may vary by a factor of 3-8 simply based on the selection of the SAS function ~~.~~ (i.e. leaving the storage parameter unchanged). The affinity for younger water is rather typical in catchments, at least during wet conditions, while the release of older water is more representative of soil columns or aquifers. The second solute example (Figure 3b) evaluates the “memory” of a system, i.e. the time needed to adapt to a new condition. Again, the preferential release of older storage volumes and the implied lack of young water in streamflow makes the system response more damped. However, this also means that the old water gets depleted faster, hence in the long term ~~(e.g. after 2 years in Figure 3b) the trend is~~ the trend may be reversed and the residual legacy of the initial conditions ~~is~~ may be stronger in systems with a high affinity for younger water. This is visible in Figure (Figure 3b) right after year 2, although the effect is very mild in this case. The time-variant SAS function (ω_1) is particularly illustrative in this example, because it shows that streamflow concentration can increase in time (e.g. around year 1 in Figure 3b), even in the absence of new solute input, just as a consequence of the changing transport mechanisms.

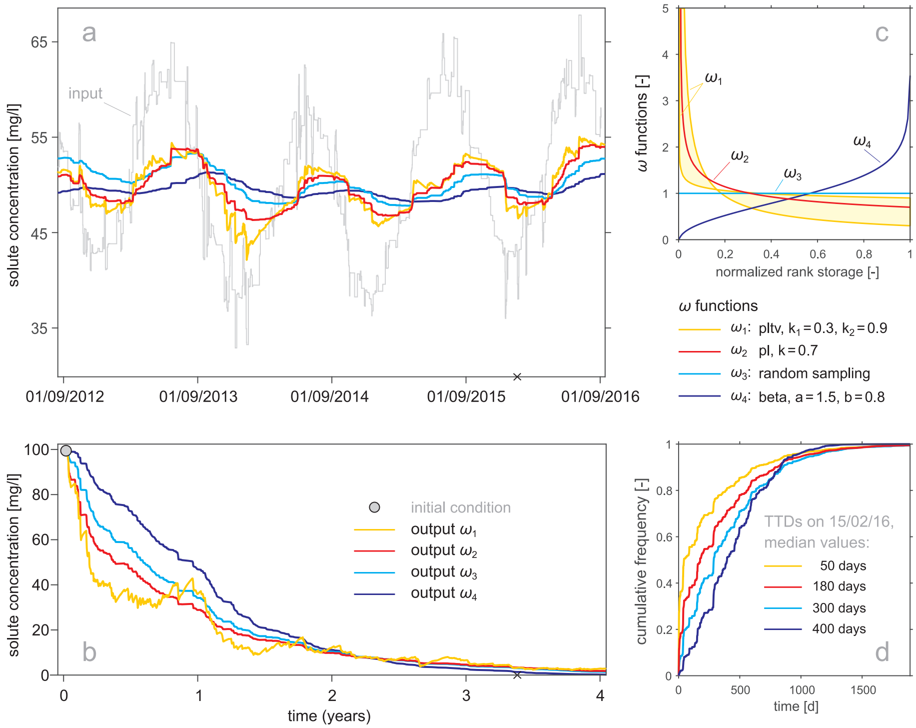

Figure 3. Example of results that can be obtained from the model. a) streamflow solute response in case of synusoidal tracer input; b) streamflow solute response in case of step-reduction of the tracer input; c) illustration of the different ω_Q used in the simulations and listed in Table 1 (as ω_1 is time-variant, its possible shapes are represented by a colored band); d) cumulative streamflow travel time distributions (TTDs) extracted on a specific day (15 February 2016, indicated with a cross in plots a) and b)). All simulations share the same settings and only differ in the choice of the ω_Q function.

Overall, these quick examples were used to illustrate the model capabilities ~~, but many other applications to solute transport can be designed and addressed through the model~~and to show that results may change significantly depending on the choice of the parameters. A sensitivity analysis is generally advised to identify the parameters that have the highest impact on model results. For example, previous catchment studies (e.g. Benettin et al., 2017b) highlighted the challenge in constraining the SAS function of ET flux when based on streamflow concentration measurements only. As a consequence, the hypothesis of random sampling for the ET flux is often as valid as the preference for the younger/older stored water, but it is more parsimonious. Different models outputs are affected by parameters in different ways, and water ages (for example the median age, Figure 3d) are typically more sensitive than solute concentration to parameters variations. The low computational times of the model aid the development of sensitivity analyses.

5 Discussion

5.1 Model verification

We evaluate here the numerical accuracy of the model in computing the solution of the age ME (i.e., the rank storage S_T) and streamflow concentration C_Q. The numerical model is first evaluated by comparing our modified Euler solution (equation 9) to a numerical implementation of the analytic solution (equation 1). This comparison is only possible for the case of random sampling (~~RS,~~ see section 2.4), as no analytic solution is usually available for other transport schemes. Then, the comparison is made for other shapes of the SAS function, approximating the 'true' solution with a higher-order implementation of equation (8). As in section 4, comparisons are made on the example dataset, using daily average fluxes and the sinusoidal tracer input concentration.

For the RS comparison, the analytic solution was obtained by implementing equation (7) at daily scale, considering that fluxes are piecewise constant while the storage is piecewise linear during the timestep. The numerical solution for the RS was obtained by setting both Ω_Q and Ω_{ET} as power laws with parameters $k_Q = k_{ET} = 1$. The numerical model was run for 8 different aggregation timesteps h: 1, 2, 3, 4, 6, 8, 12, 24 hours. For each run, the resulting streamflow concentration and one rank storage (corresponding to the end of day 2745) were used for comparison with the analytic solution. Models were run for 8 years using 4 years of spinup. To allow direct comparisons across different aggregation timesteps, streamflow concentrations were extracted at the end of each day, resulting in 8 different timeseries (one per h) of 2920 elements. The timeseries were then normalized by the mean and standard deviation of the analytic solution. A timeseries of model errors on streamflow concentration (err_{C_Q}) was finally obtained from the difference between the analytic and the numerical (normalized) solutions. The rank storage was evaluated on the entire age domain every 24 hours (again, to allow comparisons across different timesteps). To avoid comparisons between cumulative functions, the rank storage was used to compute the storage age pdf p_S (see Section 2.1). The errors on p_S were obtained from the difference between the analytic and the numerical solutions. In this case, the error timeseries (err_{p_S}) consists, for each of the 8 aggregation timesteps, of 2745 elements. For additional comparisons, the performance of our numerical implementation ("EF*") was compared to the classic implementation of the forward Euler scheme ("EF", i.e., equation (10) with $e[j] = 0$). Results are obtained for 4 different values of the initial storage S_0: 300, 500, 1000, 2000 mm. The standard deviations of err_{p_S} and err_{C_Q} are shown in Figure 4 as a function of the aggregation timestep. The EF and EF* implementations almost have the same error on p_S, indicating that accounting for the event water does not have a major impact on the overall solution of the age ME. However, as different ages do not contribute equally to streamflow, the event water can have a larger impact on streamflow concentration. This is evident in the performance on err_{C_Q}, where the modified EF* implementation is about one order of magnitude more accurate than the classic Euler scheme. The error is on average smaller than 10^{-2} the variance of the C_Q signal, which is lower than most measurement errors. The performance on err_{C_Q} also shows that the errors tend to grow with decreasing values of the mean storage, i.e. when the storage gets depleted (or filled) faster. The error of the EF* scheme shows a good stability. This is not surprising as the RS case resembles a linear reservoir (see Section 2.4) with a coefficient c approximately equal to the mean ratio between the fluxes and the storage $\langle J/S \rangle$ during a timestep. The stability condition for the Euler Forward scheme in the case of a linear reservoir requires that $c < 2/h$

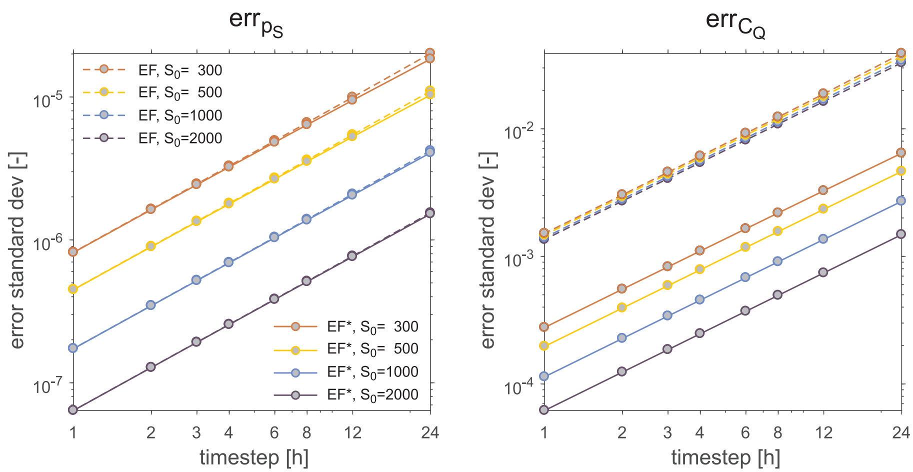

Figure 4. Numerical errors on the storage age distribution (left panel) and on streamflow concentration (right panel) as a function of the aggregation timestep. The error timeseries are summarized through their standard deviation. Each plot shows the performance of 2 different numerical schemes: classic Euler Forward (EF) and modified Euler Forward (EF*, which is the default model version). The EF* implementation shows significant improvements with respect to EF in the accuracy of streamflow concentration.

(no fast decay). In typical hydrologic applications, fluxes are usually much smaller than the storage, hence $\langle J/S \rangle \ll 1/h$ and the EF solution is stable.

Results show that the numerical implementation of the ME is satisfactory for the RS solution both in terms of accuracy and stability. However, solutions other than the RS case may be more challenging owing to the non-uniform age selection played by the outflows. For this reason, we tested power-law SAS functions (equation (6)) with different values of the exponent k: 0.2, 0.3, 0.5, 0.7, 1, 1.2, 1.5, 2, 3. The same exponent was used each time for both Ω_Q and Ω_{ET}. The model was run with a fixed initial storage $S_0 = 1000$, for the same timespan and aggregation timesteps as in the RS case, and the performance was again evaluated in terms of err_{p_S} and err_{C_Q}. Given the lack of analytical solutions, we approximated the true solution by using a higher-order implementation (built-in MATLAB solver 'ode113' (Shampine and Reichelt, 1997)) for equation (8). An example of C_Q timeseries obtained from the different values of k for $h = 24$ hours is reported in Figure 5. The C_Q timeseries are rather different, being progressively more lagged and damped for increasing values of k. Although the residual with respect to the higher-order solution can occasionally be up to 1.3 mg/l, it is on average very low compared to the signal, so in this case the accuracy of the model is satisfactory even for $h = 24$ hours. Note that for this dataset, the parameters of the SAS function (k =0.2 and S_0 =1000) imply that 30$\%$ of the input, on average, becomes output during the same day. The residuals are overall low and do not accumulate during the 8-year simulation, suggesting that even the 24-hour simulation is stable. The performance on C_Q was further evaluated in the same way as for the RS case: we normalized the concentration signals and obtained the error timeseries err_{C_Q} from the difference with the higher-order solution. Similarly, we computed the errors err_{p_S} with respect to the higher-order solution for simulation day 2745. The standard deviations of the errors are shown in Figure 6 for different values of k and aggregation timesteps. The errors on p_S grow for increasing preference of the SAS functions for

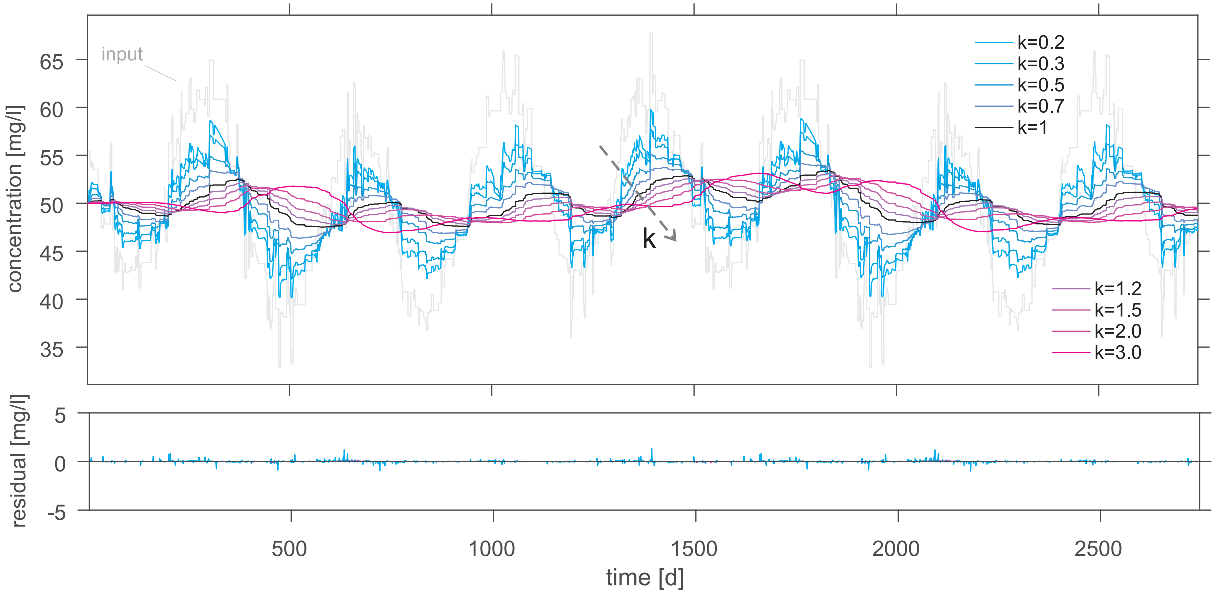

Figure 5. Solute concentration (C_Q) timeseries obtained from power-law SAS functions with parameter $S_0 = 1000$ and parameter $k \in [0.2, 3.0]$, using a 24-hour timestep (top panel). The timeseries are rather different, being progressively more lagged and damped for increasing values of k. The difference with the higher-order solution forms the residual timeseries (bottom panel, same scale as top panel). Residuals are overall limited and they do not cumulate during the 8-year simulation.

the younger stored volumes (lower values of k). This indicates that the young water preference is a more challenging numerical condition for the solution of the age ME. This behavior is to be mostly attributed to the errors on the youngest waters in storage. Although we use a modified version of the EF scheme to account for the presence of event water in the outflows (equation (10)), this approximation has some limitations. In particular, the youngest age in storage ($e[j]$) is quantified through the SAS function from previous timestep, so it may give rise to errors at the onset of intense storm events. The interpretation of the behavior of the error on C_Q (Figure 6b) is less straightforward as the errors on the solution p_S can be amplified in various ways by the different SAS functions. Errors appear not too dissimilar for k in the range 0.5-1.2 and they all are reduced by 1 order of magnitude moving from daily to hourly timesteps. The more “extreme" age selections (i.e. $k \leq 0.3$ and $k \geq 2$) tend to result in higher errors, although the error magnitude remains low (less than 10^{-2} the signal variance) and the solution is stable.

These examples suggest that the behavior of the system can be interpreted using a (non-linear) reservoir analogy. Each individual water parcel can be seen as a depleting reservoir that decreases in time owing to the particular outflow removal (equation 8). This removal is mediated by the SAS functions, so it can become large corresponding to high values of $\omega(S_T, t)$, potentially leading to an unstable fast-decay. The depletion pattern of the reservoir is rather complex as it is nonlinear and it changes at every timestep, but it suggests that very pronounced age selections should be considered carefully and checked for potential numerical instabilities. Note that for illustration purposes the effects of the two power-law SAS function parameters k and S_0 were presented separately (Figures 4 and 6), but they should be considered together as lower storage values may

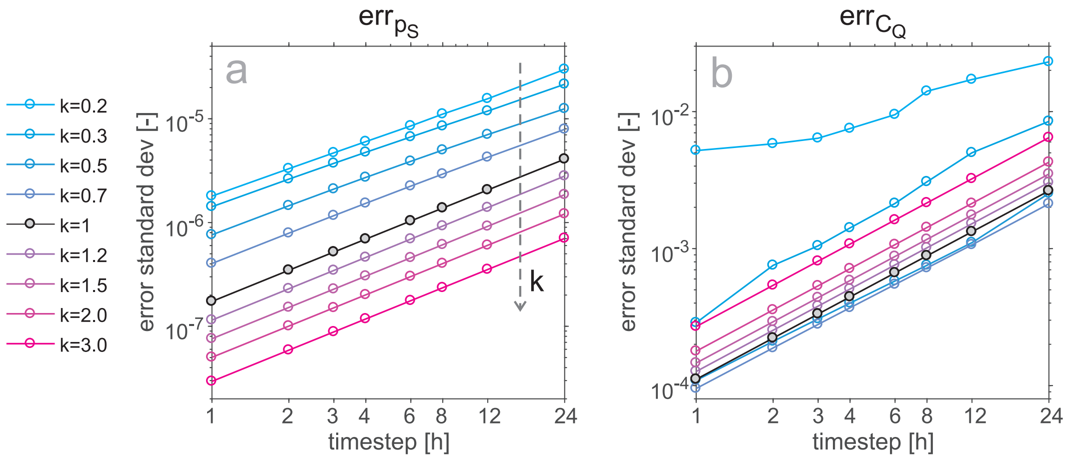

Figure 6. Numerical errors on the storage age distribution (a) and on streamflow concentration (b) as a function of the aggregation timestep. The error timeseries are summarized through their standard deviation. Each plot shows the model performance for several shapes of the SAS function, parameterized as a power-law distribution with parameter k (equation (6)). The color code is the same as in Figure 5. The random-sampling case (i.e. $k = 1$) is indicated in black and it is equivalent to the curves featuring $S_0 = 1000$ in Figure 4.

enhance the selection of younger/older waters and increase the numerical errors. The model was here tested for several shapes of the SAS functions on a realistic hydrochemical dataset. Although every dataset is different and it would be impossible to do a model verification valid for all applications, these results provide some first guidelines as to where the explicit numerical implementation may become critical.

5.2 Model applicability, limitations and perspectives

The model is based on a catchment-scale approach, so it only requires catchment-scale fluxes like precipitation, discharge and evapotranspiration. These fluxes can often be measured (or modeled in the case of ET) without the need for a full hydrologic model. Moreover, the 'pure' SAS function approach implies that, differently from previous approaches (e.g. Bertuzzo et al., 2013; Benettin et al., 2015a), the transport equations which are solved in the model are completely decoupled from the way fluxes were obtained. This notably reduces the number of involved parameters and it simplifies the applicability of the model to different datasets and contexts. Although more research is needed to classify the expected shapes of the SAS functions based on measurable catchment properties, one can quickly obtain first-order evaluations of solute transport by using SAS functions already tested in the literature (e.g. van der Velde et al., 2015; Harman, 2015; Queloz et al., 2015b; Benettin et al., 2017b; Wilusz et al., 2017) and a reasonable choice of the initial storage S_0.

The use of an explicit numerical scheme has the potential of greatly reducing the computational times. Short aggregation timesteps are generally recommended, especially when testing the affinity for younger storage volumes (e.g. equation (6) with parameter $k < 0.3$), but in case larger timesteps (e.g. $h =$24 h) prove satisfactory, the model can typically run in less than a second on a normal computer. The short computational times make the use of calibration techniques easier and the model structure is directly compatible with the DREAM (Vrugt et al., 2009; ter Braak and Vrugt, 2008) calibration packages. The model can be made faster by not considering the zero-precipitation times but, as explained in section 3.3, this improvement is currently not implemented to keep the model more intuitive.

The model is based on a catchment-scale formulation of transport processes, so it cannot provide spatial information unless the system is partitioned into a series of spatial compartments (e.g. Soulsby et al., 2015). Even in this case, one would need to know the fluxes to/from each compartment, hence losing one of the main advantages of the general SAS approach. The catchment-scale nature of the formulation also implies that SAS functions have a conceptual character and they cannot be determined directly from physical properties of the system. Their general shape, however, can be traced back to elementary advection-dispersion processes (Benettin et al., 2013) and the mechanistic basis for time-variable SAS functions has recently been highlighted (Pangle et al., 2017).

Although the numerical accuracy of the computations has to be evaluated for each different application, section 5.1 provides some first guidelines to cases where the numerical accuracy may not be satisfactory. Systems whose storage is quickly depleted by the fluxes are prone to inaccuracies and instabilities. This can happen, for instance, if the system storage is small compared to the fluxes and the SAS functions have a very strong preference for some storage portions. In such cases, higher order schemes may become desirable. The model package already provides a higher-order solution to equation (8) (obtained through the MATLAB built-in function ’ode113’), that can help evaluating the numerical accuracy of the results.

The codes implemented in the *tran*-SAS package can be used to simulate the transport of conservative solutes through a catchment. This represents a first step towards the modeling of large-scale solute transport. ~~Reactive~~ Simple reactive transport equations can be easily implemented in the main model routine (section 3.2) using effective formulations that integrate biogeochemical processes across the catchment heterogeneity (Rinaldo and Marani, 1987). Being based on a travel time formulation of transport, the model is obviously not suited to simulate the circulation of solutes for which the chronology of the inputs ~~is~~ and the age of water are irrelevant. For a number of cases of interest, however, both the time of entry into the catchment and the residence time of water within the catchment storage may play an important role in the transport process. Many such examples have been addressed in the literature using a catchment-scale approach, including the case of nitrate export from agricultural catchments (Botter et al., 2006; van der Velde et al., 2012), solutes influenced by evapoconcentration effects (Queloz et al., 2015b), pesticide transport (Bertuzzo et al., 2013; Lutz et al., 2017) and solutes produced by mineral weathering (Benettin et al., 2015a). The provided codes are designed to be easy to understand, so that they can be easily customized by the user and adapted to different contexts and applications. The next step is then to adapt the model to real-world problems, where solutes’ non-conservative behavior has to be taken into account.

6 Conclusions

The *tran*-SAS package includes a basic implementation of the age Master Equation (equation 1) using general SAS-functions. The codes can be used to simulate the transport of solutes through a catchment and to evaluate water residence times. The package is ready-to-go and it includes some example data that can be used to test the main model features. The codes are extensively commented so that they can be edited according to the user's needs. The model is based on a catchment-scale formulation of solute transport and it only relies on measurable data. Main model equations are implemented using an explicit Euler scheme that allows to reduce computational times. The numerical accuracy of the model was verified on the example data and was shown to be generally satisfactory even at larger (e.g. daily) computation timesteps. The most critical cases are those in which the stored water parcels are rapidly removed by the outflows. This situation can occur when the SAS function assumes very high values for some stored water volumes. In such cases, higher-order model implementations (provided within the package) should be used to check the numerical accuracy of the solution. The model allows to test different SAS functions and evaluate solute transport in the catchment storage and outflows. Applications can be oriented to different catchments and solutes, advancing our ability to understand and model catchment transport processes.

7 Code and Data availability

~~A maintained code package with~~ The current model release, including example data and documentation, is available at https://doi.org/10.5281/zenodo.1203600. A maintained GitHub project is available at the following GitHub repository: https://github.com/pbenettin/tran-SAS

Acknowledgements. The authors thank Andrea Rinaldo and Gianluca Botter for the useful discussions that inspired this work, and Damiano Pasetto for support in the numerical implementation of the model equations. PB thanks the ENAC school at EPFL for financial support.

References

Benettin, P., Rinaldo, A., and Botter, G.: Kinematics of age mixing in advection-dispersion models, Water Resources Research, 49, 8539–8551, doi:10.1002/2013WR014708, http://dx.doi.org/10.1002/2013WR014708, 2013.

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, doi: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, doi:10.1002/hyp.15034, 2015b.

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, doi:10.1002/hyp.11243, https://onlinelibrary.wiley.com/doi/abs/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, doi:10.1002/2016WR020117, http://dx.doi.org/10.1002/2016WR020117, ~~2017.~~ 2017b.

Bertuzzo, E., Thomet, M., Botter, G., and Rinaldo, A.: Catchment-scale herbicides transport: Theory and application, Advances in Water Resources, 52, 232 – 242, doi:10.1016/j.advwatres.2012.11.007, 2013.

Botter, G.: Catchment mixing processes and travel time distributions, Water Resources Research, 48, doi:10.1029/2011WR011160, 2012.

Botter, G., Settin, T., Marani, M., and Rinaldo, A.: A stochastic model of nitrate transport and cycling at basin scale, Water Resources Research, 42, doi:10.1029/2005WR004599, http://dx.doi.org/10.1029/2005WR004599, w04415, 2006.

Botter, G., Bertuzzo, E., Bellin, A., and Rinaldo, A.: On the Lagrangian formulations of reactive solute transport in the hydrologic response, Water Resources Research, 41, doi:10.1029/2004WR003544, 2005.

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

Calabrese, S. and Porporato, A.: Linking age, survival, and transit time distributions, Water Resources Research, 51, 8316–8330, doi:10.1002/2015WR017785, http://dx.doi.org/10.1002/2015WR017785, 2015.

Cvetkovic, V. and Dagan, G.: Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations, Journal of Fluid Mechanics, 265, 189–215, doi:10.1017/S0022112094000807, 1994.

Danesh-Yazdi, M., Foufoula-Georgiou, E., Karwan, D. L., and Botter, G.: Inferring changes in water cycle dynamics of intensively managed landscapes via the theory of time-variant travel time distributions, Water Resources Research, 52, 7593–7614, doi:10.1002/2016WR019091, http://dx.doi.org/10.1002/2016WR019091, 2016.

Danesh-Yazdi, M., Botter, G., and Foufoula-Georgiou, E.: Time-variant Lagrangian transport formulation reduces aggregation bias of water and solute mean travel time in heterogeneous catchments, Geophysical Research Letters, 44, 4880–4888, doi:10.1002/2017GL073827, http://dx.doi.org/10.1002/2017GL073827, 2017.

Destouni, G., Persson, K., Prieto, C., and Jarsj, J.: General Quantification of Catchment-Scale Nutrient and Pollutant Transport through the Subsurface to Surface and Coastal Waters, Environmental Science & Technology, 44, 2048–2055, doi:10.1021/es902338y, 2010.

Drever, M. C. and Hrachowitz, M.: Migration as flow: using hydrological concepts to estimate the residence time of migrating birds from the daily counts, Methods in Ecology and Evolution, pp. n/a–n/a, doi:10.1111/2041-210X.12727, http://dx.doi.org/10.1111/2041-210X.12727, 2017.

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, doi:10.1002/2014WR015707, http://dx.doi.org/10.1002/2014WR015707, 2015.

Harman, C. J., Ward, A. S., and Ball, A.: How does reach-scale stream-hyporheic transport vary with discharge? Insights from rSAS analysis of sequential tracer injections in a headwater mountain stream, Water Resources Research, 52, 7130–7150, doi:10.1002/2016WR018832, http://dx.doi.org/10.1002/2016WR018832, 2016.

Hrachowitz, M., Fovet, O., Ruiz, L., and Savenije, H. H. G.: Transit time distributions, legacy contamination and variability in biogeochemical $1/f^{\alpha}$ scaling: how are hydrological response dynamics linked to water quality at the catchment scale?, Hydrological Processes, doi:10.1002/hyp.10546, http://dx.doi.org/10.1002/hyp.10546, 2015.

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, doi:10.1002/wat2.1155, 2016.

Jackson, B., Wheater, H., Wade, A., Butterfield, D., Mathias, S., Ireson, A., Butler, A., McIntyre, N., and Whitehead, P.: Catchment-scale modelling of flow and nutrient transport in the Chalk unsaturated zone, Ecological Modelling, 209, 41 – 52, doi:http://dx.doi.org/10.1016/j.ecolmodel.2007.07.005, http://www.sciencedirect.com/science/article/pii/S0304380007003572, 2007.

Kauffman, S. J., Royer, D. L., Chang, S., and Berner, R. A.: Export of chloride after clear-cutting in the Hubbard Brook sandbox experiment, Biogeochemistry, 63, 23–33, doi:10.1023/A:1023335002926, https://doi.org/10.1023/A:1023335002926, 2003.

Kim, M., Pangle, L. A., Cardoso, C., Lora, M., Volkmann, T. H. M., Wang, Y., Harman, C. J., and Troch, P. A.: 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 Resources Research, doi:10.1002/2016WR018620, http://dx.doi.org/10.1002/2016WR018620, 2016.

Lutz, S. R., Velde, Y. V. D., Elsayed, O. F., Imfeld, G., Lefrancq, M., Payraudeau, S., and van Breukelen, B. M.: Pesticide fate on catchment scale: conceptual modelling of stream CSIA data, Hydrology and Earth System Sciences, 21, 5243–5261, doi:10.5194/hess-21-5243-2017, https://www.hydrol-earth-syst-sci.net/21/5243/2017/, 2017.

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, doi:10.1016/j.epsl.2011.09.040, http://linkinghub.elsevier.com/retrieve/pii/S0012821X11005607, 2011.

Maloszewski, P. and Zuber, A.: Principles and practice of calibration and validation of mathematical models for the interpretation of environmental tracer data in aquifers, Advances in Water Resources, 16, 173 – 190, doi:http://dx.doi.org/10.1016/0309-1708(93)90036-F, 1993.

Martin, C., Aquilina, L., Gascuel?Odoux, C., Molt, J., Faucheux, M., and Ruiz, L.: Seasonal and interannual variations of nitrate and chloride in stream waters related to spatial and temporal patterns of groundwater concentrations in agricultural catchments, Hydrological Processes, 18, 1237–1254, doi:10.1002/hyp.1395, 2004.

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

McGuire, K. J. and McDonnell, J. J.: Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlinearities, Water Resources Research, 46, n/a–n/a, doi:10.1029/2010WR009341, http://dx.doi.org/10.1029/2010WR009341, w10543, 2010.

McMillan, H., Tetzlaff, D., Clark, M., and Soulsby, C.: Do time-variable tracers aid the evaluation of hydrological model structure? A multimodel approach, Water Resources Research, 48, doi:10.1029/2011WR011688, http://doi.wiley.com/10.1029/2011WR011688, 2012.

Oda, T., Asano, Y., and Suzuki, M.: Transit time evaluation using a chloride concentration input step shift after forest cutting in a Japanese headwater catchment, Hydrological Processes, 23, 2705–2713, doi:10.1002/hyp.7361, http://dx.doi.org/10.1002/hyp.7361, 2009.

Pangle, L. A., Kim, M., Cardoso, C., Lora, M., Meira Neto, A. A., Volkmann, T. H. M., Wang, Y., Troch, P. A., and Harman, C. J.: The mechanistic basis for storage-dependent age distributions of water discharged from an experimental hillslope, Water Resources Research, 53, 2733–2754, doi:10.1002/2016WR019901, http://dx.doi.org/10.1002/2016WR019901, 2017.

Queloz, P., Bertuzzo, E., Carraro, L., Botter, G., Miglietta, F., Rao, P., and Rinaldo, A.: Transport of fluorobenzoate tracers in a vegetated hydrologic control volume: 1. Experimental results, Water Resources Research, 51, 2773–2792, doi:10.1002/2014WR016433, http://dx.doi.org/10.1002/2014WR016433, 2015a.

Queloz, P., Carraro, L., Benettin, P., Botter, G., Rinaldo, A., and Bertuzzo, E.: Transport of fluorobenzoate tracers in a vegetated hydrologic control volume: 2. Theoretical inferences and modeling, Water Resources Research, 51, 2793–2806, doi:10.1002/2014WR016508, http://dx.doi.org/10.1002/2014WR016508, 2015b.

Rigon, R., Bancheri, M., and Green, T. R.: Age-ranked hydrological budgets and a travel time description of catchment hydrology, Hydrology and Earth System Sciences, 20, 4929–4947, doi:10.5194/hess-20-4929-2016, http://www.hydrol-earth-syst-sci.net/20/4929/2016/, 2016.

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

Shampine, L. F. and Reichelt, M. W.: The MATLAB ODE Suite, SIAM Journal on Scientific Computing, 18, 1–22, doi:10.1137/S1064827594276424, https://doi.org/10.1137/S1064827594276424, 1997.

Soulsby, C., Birkel, C., Geris, J., and Tetzlaff, D.: Spatial aggregation of time-variant stream water ages in urbanizing catchments, Hydrological Processes, 29, 3038–3050, doi:10.1002/hyp.10500, http://dx.doi.org/10.1002/hyp.10500, 2015.

ter Braak, C. J. F. and Vrugt, J. A.: Differential Evolution Markov Chain with snooker updater and fewer chains, Statistics and Computing, 18, 435–446, doi:10.1007/s11222-008-9104-9, http://link.springer.com/10.1007/s11222-008-9104-9, 2008.

van der Velde, Y., Heidbchel, 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, doi:10.1002/hyp.10372, http://dx.doi.org/10.1002/hyp.10372, 2015.

van der Velde, Y., de Rooij, G. H., Rozemeijer, J. C., van Geer, F. C., and Broers, H. P.: Nitrate response of a lowland catchment: On the relation between stream concentration and travel time distribution dynamics, Water Resources Research, 46, doi:10.1029/2010WR009105, 2010.

van der Velde, Y., Torfs, P. J. J. F., van der Zee, S. E. A. T. M., and Uijlenhoet, R.: Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, Water Resources Research, 48, doi:10.1029/2011WR011310, w06536, 2012.

Vrugt, J., Braak, C. T., Diks, C., Robinson, B., Hyman, J., and Higdon, D.: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, International Journal of Nonlinear Sciences & Numerical Simulation, 10, 271–288, doi:10.1515/IJNSNS.2009.10.3.273, 2009.

Weiler, M., McGlynn, B. L., McGuire, K. J., and McDonnell, J. J.: How does rainfall become runoff? A combined tracer and runoff transfer function approach, Water Resources Research, 39, n/a–n/a, doi:10.1029/2003WR002331, http://dx.doi.org/10.1029/2003WR002331, 2003.

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, doi:10.1002/2017WR020894, http://dx.doi.org/10.1002/2017WR020894, in press, 2017.