Exploring tracer information in a small stream to reduce the uncertainty and enhance the process interpretation of transient storage models

Enrico Bonanno[1,2], Günter Blöschl[2], Julian Klaus[3]

5   [1] Catchment and Eco-Hydrology Group, Luxembourg Institute of Science and Technology, Belvaux, Luxembourg.

[2] Institute of Hydraulic and Water Resources Engineering, Vienna University of Technology, Vienna, Austria.

[3] Institute of Geography, University of Bonn, Bonn, Germany.

10   *Correspondence to:* Enrico Bonanno (bonanno@hydro.tuwien.ac.at)

Abstract. The transport of solutes in river networks is controlled by the interplay of processes such as in-stream solute transport and the exchange of water between the stream channel and dead zones, in-stream sediments, and the hyporheic zone. Transient storage models (TSMs) are a powerful tool for testing hypotheses related to solute

15   transport in streams. However, TSM parameters are often non-identifiable leading to an unclear understanding of the processes controlling solute transport in streams. In this study, we increased parameter identifiability in a set of tracer breakthrough experiments by combining global identifiability analysis and dynamic identifiability analysis. We compared our results to inverse modelling approaches (OTIS-P) and the commonly used random sampling approach for TSMs (OTIS-MCAT). Compared to OTIS-P, our results informed about sensitivity and

20   identifiability of TSM parameters on the entire feasible parameter space. Our results clearly improved parameter identifiability compared to OTIS-MCAT that often indicated non-identifiability of TSM parameters. Non-identifiable results led to wrong solute retention times in the storage zone and the exchange flow with the storage zone, with a difference respectively up to four and two orders of magnitude compared to results with identifiable TSM parameters. The severe differences in the transport metrics between results obtained from our proposed

25   approach and results from OTIS-MCAT model also resulted in contrasting interpretation of the hydrologic processes controlling solute transport at the study site. Thus, our outcomes point to the risks of interpreting TSM results when even one of the TSM parameters is non-identifiable. Our results showed that there is clear potential for increasing parameter identifiability in random sampling approaches for TSMs and for advancing our understanding of hydrological processes controlling in-stream solute transport.

30

35

1 Introduction

40

Modelling of stream water movement is pivotal for understanding how nutrients, solutes and pollutants are transported downstream and ultimately affect stream water quality along river networks (Krause et al., 2011; Rathfelder, 2016; Smith, 2005). Experimental studies of water flow and solute transport in the stream channel commonly rely on the observation of tracer breakthrough curves (BTC, i.e., the measurement of the concentration

45   over time of a tracer released in an upstream section). Such a BTC reflects stream discharge (Beven et al., 1979; Butterworth et al., 2000) and longitudinal tracer advection and dispersion (Gooseff et al., 2008). A milestone in the study of solute transport was that in-stream solutes and water are exchanged with slowly-moving channel waters, the dead zones (Hays, 1966), and with the saturated area between the stream channel and the adjacent groundwater (i.e., the hyporheic zone, Triska et al., 1989; White, 1993). This hydrologic exchange results in a

50   skewed non-Fickian BTC with a pronounced tail, which makes the advection-dispersion equation (ADE) unable to correctly describe the observed tracer transport in stream channels (Bencala & Walters, 1983; Castro & Hornberger, 1991). Despite the large amount of studies, the numerous contradictory outcomes (Ward & Packman, 2019), coupled with uncertainties of the model parameters (Ward et al., 2017), raise the question about how informative such modelling results are (Knapp & Kelleher, 2020).

55   Considerable potential in reducing uncertainty of the processes controlling solute transport in streams lies in modelling the tail of the BTC, since it contains information on the transient storage inside the stream channels (Bencala et al., 2011). For simulating the retentive effect of dead zones on solute transport, Hays (1966) modelled the tail of the BTC by introducing a second differential equation in addition to the ADE. Following a similar approach, (Bencala & Walters (1983) described the solute transport in streams as a pure advection-dispersion

60   transport, coupled with a hydrologic exchange term between the stream channel and a single, homogeneously mixed volume that delays the solute movement downstream (Transient Storage Model - TSM). The estimation of TSM parameters often rely on the use of inverse modelling approaches via nonlinear regression algorithm that can return precise estimation of TSM parameters with a narrow 95% confidence interval (OTIS-P; Runkel, 1998). While this approach found extensive application in past decades, it does not allow a comprehensive assessment

65   about the identifiability of the TSM parameters (Ward et al., 2017; Knapp & Kelleher, 2020). This is because inverse modelling approach do not provide information on performances and parameter identifiability on the entire feasible parameter space (Ward et al., 2017). Also, the parameter sets obtained after inverse modelling approach do not necessarily indicate meaningful results, as non-identifiable parameters can provide good BTC fitting despite being uncertain and non-identifiable (Kelleher et al., 2019). These considerations, coupled with the lack

70   of knowledge of the modeller over the distribution of parameters and their performance after OTIS-P results, led to a progressive increase of studies addressing identifiability in TSMs via random sampling of parameters and global identifiability analysis (OTIS-MCAT model; Ward et al., 2017; Knapp & Kelleher, 2020; Kelleher et al., 2019).

Random sampling approaches provide information on parameter identifiability and accuracy on the feasible

75   parameter space, however they rarely show identifiability for all the TSM parameters (Knapp & Kelleher, 2020). Kelleher et al. (2013) found that the parameters associated with the transient storage process are not identifiable for a large variety of the stream reaches and experiments they investigated. Other studies have shown that TSM parameters are often not-independent and poorly identifiable (Camacho & González, 2008; Kelleher et al., 2019;

Hydrology and
Earth System
Sciences
Open Access
Discussions
EGU

Knapp & Kelleher, 2020; Wagener et al., 2002; Ward et al., 2017; Wlostowski et al., 2013). Despite these findings

80   and the crucial need for parameter identifiability, only few studies have explored the reliability of TSM results obtained via inverse modelling approach, and model interpretation is often based on a single set of parameters without testing their robustness (Knapp & Kelleher, 2020).

Addressing the identifiability of TSM parameters is a pressing issue, since we are still unable to link specific physical processes with the parameters derived from BTC studies (Ward & Packman, 2019). This problem is

85   commonly related to the over-simplistic approach of TSM, which is unable to distinguish between the effects of eddies, pools, and the hyporheic zone (Gooseff et al., 2008; Zaramella et al., 2006). To overcome this limitation, the TSM has been modified to include multiple storage zones (Choi et al., 2000), sorption kinetics for reactive tracers (Gooseff et al., 2005; Kelleher et al., 2019), and changes of residence time distributions in the storage zone (Haggerty et al., 2002). While these changes increased the quality of the model fit, they also came at the cost of

90   increased dimensionality with a further reduction of parameter identifiability and certainty, leaving the open question what physical processes exactly are associated with the transient storage modelling (Kelleher et al., 2019; Knapp & Kelleher, 2020).

The observed strong non-identifiability for TSM parameters in random-sampling studies may have three causes. First, there is no common strategy for selecting parameter ranges and the number of parameter sets in TSM

95   simulations. To obtain reliable results, Ward et al. (2017) indicated that modelling studies need to apply TSM on a large number of parameter sets (between 10,000 and 100,000) over a parameter range spanning at least two orders of magnitude. While for some studies, the non-identifiability of parameters might be explained by the low number of parameter sets (less than 10,000) and the relatively narrow selected parameter range (Camacho & González, 2008; Wagener et al., 2002; Wlostowski et al., 2013), non-identifiability was also found when a rather

100   large number of parameter sets and wide range were used (Kelleher et al., 2013; Kelleher et al., 2019; Ward et al., 2017). This is bringing up the question *if* and *when* TSM parameters are actually meaningful (Knapp & Kelleher, 2020).

A second cause related to uncertain results in the random-sampling approach for TSM parameters relates to the selected parameters chosen for TSM simulations. The parameters describing the advection-dispersion process

105   (stream velocity, cross-sectional area of the stream channel, and the longitudinal dispersion) are known to be the best identifiable in the TSM (Ward et al., 2017) and once they are kept constant in the random sampling of the parameters, they can drive strong changes in the parameters describing the transient storage process (Knapp & Kelleher, 2020). This dependency indicates that the use of a constant value for an identifiable parameter may result in a mis-estimation of the other TSM parameters. While it is not recommended to keep constant a rather

110   identifiable parameter in random-sampling approaches for TSMs, no study investigated the role that a variable velocity has on identifiability of TSM parameters. This leads to the question on how meaningful, sensitive, and uncertain the transient storage parameters are when stream velocity varies or is kept constant.

A third cause for non-identifiable TSM parameters relates to the selected approach for addressing parameter identifiability. The identifiability analysis used in most studies is based on the Generalized Likelihood Uncertainty

115   Estimation that assesses parameter certainty across the entire observed BTC (GLUE, Beven & Binley, 1992; Camacho & González, 2008; Kelleher et al., 2013; Kelleher et al., 2019; Ward et al., 2017). However, such global identifiability analysis is unable to find informative sections of the simulated BTC with respect to a certain parameter and to unequivocally link a given parameter to a specific process (Wagener et al., 2003; Wagener et

120 al., 2002; Wagener & Kollat, 2007). This information is particularly important for BTC modelling, since advection-dispersion parameters are physically responsible for the bulk solute transport in the stream and they are therefore expected to act on the rising limb and peak of the BTC (Gooseff et al., 2008). Contrary, the parameters describing the exchange between the stream channel and the transient storage zone are responsible for delaying solute transport compared to the advective-dispersive transport, acting on the falling limb and tail of the BTC (Runkel, 2002). By investigating parameter sensitivity and identifiability across the entire BTC, global

125 identifiability analysis is unable to capture an increase in parameter identifiability towards the tail of the BTC. However, studies addressing the identifiability of TSM parameters over time in different sections of the BTC reported an increase of identifiability for transient storage parameters on the tail of the BTC (Wagener et al., 2002; Scott et al., 2003; Wlostowski et al., 2013; Kelleher et al., 2013). We hypothesise that this information is key in designing a successive parameter sampling in a constrained parameter space – ultimately reducing the uncertainty

130 affecting parameters describing solute retention in streams.

A robust assessment of transient storage parameters would not only improve the model fit of tracer transport and decrease parameter uncertainty, but it might also lead to stronger interpretation on the physical processes controlling solute transport in streams. TSM parameters are often used to calculate metrics on the solute exchange between the stream channel and the transient storage zone and the residence time of solutes in the coupled system

135 (Thackston and Schnelle, 1970; Hart et al., 1999; Morrice et al., 1997; Runkel, 2002). These metrics are pivotal to address the potential for nutrient cycling, microbial activity, and the development of hot-spots in river ecosystems (Mulholland et al., 1997; Smith, 2005; Triska et al., 1989; Krause et al., 2017). However, no study so far indicated and evaluated *if* and *how much* the interpretation of hydrologic processes changes when TSM parameters are identifiable and when they are not, due to the enunciated challenges in TSMs.

140 To address these challenges, we have organised this contribution around three questions related to the key challenges of parameter identifiability in transient storage modelling:

1) How does the identifiability and the information content of model parameters associated with transient storage processes change by using fixed and varying velocity in the random-sampling of TSM parameters?

145 2) Does the identifiability analysis on specific sections of the BTC reduce the parameter uncertainty in random-sampling of TSM?

With the outcomes of these questions we will address:

3) How does the residence time of solute in the transient storage zone and hydrologic process interpretation vary when TSM parameters are identifiable and when they are not?

150

2 Study site and methods

2.1 Study site and data

The studied stream reach (49°49'38"N, 5°47'44"E) is located in western Luxembourg, downstream of the Weierbach experimental catchment (Hissler et al., 2021). The stream channel is unvegetated with a slope of ≃6

155 % and consists of deposited colluvium material and fragmented schists (up to 50 cm depth) with local outcrops of fractured slate bedrock in the streambed. The flow regime is governed by the interplay of seasonality between precipitation and evapotranspiration (Rodriguez et al., 2021; Rodriguez & Klaus, 2019) with a persistent discharge

between autumn and spring, and little to no discharge during summer months (discharge arithmetic mean equal
to 6.5 l/s, median of 1.7 l/s, St.Dev. of 11.52 l/s between Aug 2018 and Feb 2020; Bonanno et al., 2021). To test

160   our objectives, we have carried out three tracer experiments with an instantaneous tracer injection at three different
flow (Q) conditions: 6th December 2018, Q = 2.52 l/s (E1); 23rd January 2019, Q = 9.05 l/s (E2); 28th January
2019, Q = 22.79 l/s (E3). For each experiment, we prepared an NaCl solution using 2 l of stream water and 100 g
of reagent-grade NaCl. We injected the solution in a turbulent pool at the beginning of the stream reach to assure
complete mixing in the stream water. Electric conductivity (EC) was measured via a portable conductivity meter

165   (WTW) 55 m downstream of the injection point and converted into Cl[-] concentration via an EC-Cl[-] regression
line (R^2 = 0.9999).

2.2 Advection-Dispersion equation and Transient Storage Model formulation

The one-dimensional Fickian-type advection and dispersion equation describes the joint effect of flow velocity
and turbulent diffusion on solute transport (Beltaos & Day, 1978; Taylor, 1921, 1954). The differential form of

170   ADE reads:

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} + \frac{1}{A} \frac{\partial}{\partial x} \left(AD \frac{\partial C}{\partial x} \right) \qquad\qquad\qquad \text{Eq.1}$$

Where t is time [T], x is the distance from the injection point along the stream reach [L], A [L^2] is the cross-
sectional area of flow, v [L/T] is the average flow velocity, D [L^2/T] is the longitudinal dispersion coefficient, and
C is the concentration of the observed tracer above background levels [M/L^3]. The solution of the differential form

175   of ADE for an instantaneous solute injection at x = 0 [L] reads:

$$C(t) = \frac{M}{A(4\pi Dt)^{1/2}} exp \left[-\frac{(L-vt)^2}{4Dt} \right] \qquad\qquad\qquad \text{Eq. 2}$$

Where M is the injected solute mass [M], t is time [T], and L is length of the investigated reach [L].

The TSM describes the solute transport in streams by combining the advection-dispersion process in the stream
channel through a hydrologic exchange with an external storage zone. The model equations read (Bencala &

180   Walters, 1983):

$$\begin{cases} \frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} + \frac{1}{A} \frac{\partial}{\partial x} \left(AD \frac{\partial C}{\partial x} \right) + \frac{q_L}{A}(C_L - C) + \alpha(C_{TS} - C) \\ \frac{\partial C_{TS}}{\partial t} = -\alpha \frac{A}{A_{TS}}(C_S - C) \end{cases} \qquad \text{Eq.3}$$

where the hydrologic exchange with the transient-storage zone is driven by the exchange coefficient α [1/T] and
the area of the transient storage zone, A_{TS} [L^2]. Here, we will refer to A, v, and D as "advection-dispersion
parameters" and to A_{TS} and α as "transient storage parameters"; the five parameters are referred to as "TSM

185   parameters". The solute concentration in the main channel and the transient storage zone are C and C_S [M/L^3],
respectively. The performances of both ADE and TSM results are evaluated using the Root Mean Squared Error
objective function (RMSE), which is the most commonly used objective function in solute-transport studies (Ward
et al., 2018; Wlostowski et al., 2017; Zaramella et al., 2016).

2.3 Iterative modelling approach to obtain TSM parameters

190   Several sampling approaches were previously used to estimate parameter uncertainty in TSMs, such as Monte
Carlo sampling (Wagener et al., 2002; Wagner & Harvey, 1997; Ward et al., 2013), Latin hypercube sampling

(LHS, Kelleher et al., 2019), and Monte Carlo coupled with a behavioural threshold (Kelleher et al., 2013; Ward et al., 2017). Here, we use LHS to sample from the selected parameter space, due to LHS higher efficiency compared to the classic Monte Carlo approach (Yin et al., 2011).

195 We simulated our tracer experiments with the ADE by sampling advection-dispersion parameters via LHS to avoid initial assumptions that could impact the parameter estimates (Figure 1). The RMSE value of the best-performing ADE parameter set was indicated as RMSE$_{ADE}$. Similar to the Monte Carlo approach coupled with behavioural threshold (Kelleher et al., 2013; Ward et al., 2017), we simulated the three tracer experiments with the TSM through a step-wise approach with *n* TSM iterations (*n* is number of iterations, Figure 1). To obtain

200 reliable TSM results, Ward et al. (2017) suggested a minimum amount of parameter sets between 10,000 and 100,000. Thus, in each TSM iteration we simulated 115,000 parameter sets. Results of each TSM iteration include RMSE values for the 115,000 parameter sets, and results of global identifiability analysis of TSM parameters. Global identifiability analysis was conducted through parameter vs RMSE plots (Wagener et al., 2003), parameter distribution plots (Ward et al., 2017), regional sensitivity analysis (Kelleher et al., 2019; Wagener & Kollat, 2007),

205 and parameter distribution plots (Wagener et al., 2002; Ward et al., 2017). Globally identifiable parameters satisfy the following criteria: univocal peak of performance in parameter vs RMSE plots and in parameter distribution plots (Ward et al., 2017) and CDF corresponding to the best 0.1 % of the results deviating from the 1:1 line and from parameter CDF corresponding to the best 10 % of the results (Kelleher et al., 2019). To evaluate the degree of identifiability of a certain parameter, we also evaluated the two-sample Kolmogorov-Smirnov (K-S) test which

210 calculates the maximum distance *K* and the corresponding *p*-value between two cumulative distribution functions, $F(P_{0.1})$ and $F(P_{10})$, by:

$$[K, p] = max|F(P_{0.1}) - F(P_{10})| \hspace{2cm} \text{Eq. 4}$$

Where $F(P_{0.1})$ and $F(P_{10})$ are the cumulative distribution function of a parameter *P* respectively for the best 0.1% and the best 10% of the results. Following the approach of Ouyang et al. (2014), we grouped parameter

215 identifiability in four categories: highly identifiable ($K > 0.25$, $p \leq 0.05$), moderately identifiable ($0.1 \leq K \leq 0.25$, $p \leq 0.05$), poorly identifiable ($K < 0.1$, $p \leq 0.05$), and non-identifiable ($p > 0.05$).

100 best performing parameter sets for each iteration were analysed with the DYNamic Identifiability Analysis (DYNIA, Wagener et al., 2002) to address the role of TSM parameters on the BTC and the change of information content. The dynamic identifiability analysis identifies regions of the observed data that are sensitive (or not) to

220 the investigated model parameter, and it can be used to test model structure, to design specific experiments, and to relate the model parameters to a specific simulated model response (Wagener et al., 2004). The dynamic identifiability analysis yields the distribution of the likelihood as function of the parameter values and the information content of the parameters over time. The information content is expressed as one minus the width of the 90 % confidence interval over the entire parameter range (Wagener et al., 2002). A wide 90 % confidence

225 interval indicates that various parameter values are associated to equally good performances resulting in a low information content. Conversely, narrow 90% confidence intervals and corresponding high information content values suggests that the best performing parameters are contained in a relatively narrow range compared to the feasible range. A detailed description on how to read the plots used to address the global identifiability analysis and the description of the dynamic identifiability analysis algorithm are reported in Appendix A.

230 The first TSM iteration investigated model performance over parameter ranges defined from the ADE results and literature values (Table 1). The following TSM iterations sampled via LHS over parameter ranges defined by the

Hydrology and
Earth System
Sciences
Discussions

Open Access
EGU

results of the previous TSM iteration. Namely, if the global identifiability analysis from the previous TSM iteration indicated that the investigated parameter is identifiable, the best 1 % of the results were used to define its parameter space in the successive TSM iteration. When the identifiability criteria were not met, the parameter

235   space investigated in the successive TSM iteration was increased or, for the case of A_{TS} and α, it was constrained based on the dynamic identifiability analysis result (information content above 0.66 on the BTC tail). This condition was chosen by the evidence that transient storage parameters A_{TS} and α are often non-identifiable via global identifiability analysis (Camacho & González, 2008; Ward et al., 2013; Ward et al., 2017; Kelleher et al., 2019), but they are identifiable on the BTC tail (Kelleher et al., 2013; Wagener et al., 2002; Wlostowski et al.,

240   2013).

While the first TSM iteration was conducted to investigate the identifiability of all the possible combinations in the feasible parameter space reported in literature and from results of ADE (Table 1), the successive iterations excluded pairs of v and A whose product was outside the value of the discharge evaluated via dilution gauging ±10 %. This condition was chosen to respect results from Schmadel et al. (2010), who reported that the discharge

245   error from dilution gauging method is $\simeq$8 %. The approach reported in Figure 1 for TSM was used also in the case where v was assumed fixed and equal to v_{peak}. The modelling was finalized once every TSM parameter indicated global identifiability via the enunciated criteria and the Kolmogorov-Smirnov test resulted in $K > 0.1$ and $p \leq 0.05$ for each TSM parameter.

We compared our results with both inverse modelling results (OTIS-P), and the most-common random sampling

250   approach in TSMs (OTIS-MCAT). OTIS-P is an inverse-modelling scheme that minimise residual sum of squares between the modelled BTC and the observed BTC. OTIS-P model estimates the best-fitting TSM parameter values and their 95 % confidence interval. As indicated in Runkel (1998), we carried out multiple OTIS-P iterations and interrupted them when parameter values calibrated via OTIS-P changed less than 0.1 % between subsequent runs. OTIS-MCAT solves the TSM for the selected number of parameter sets and addresses their identifiability with a

255   global identifiability analysis (Ward et al., 2017). Compared to our approach, OTIS-MCAT considers Monte Carlo parameter sampling instead of LHS, velocity equal to v_{peak} and it does not foresee iterative parameter sampling from results of dynamic identifiability analysis. Thus, we here indicate as "OTIS-MCAT results" the results we obtained after the first TSM iteration when v was assumed fixed and equal to v_{peak}.

2.4 Hydrologic interpretation of TSM results

260   The TSM parameter sets obtained after OTIS-P, OTIS-MCAT, and the proposed iterative TSM approach were used to compute some hydrologic metrics to interpret solute transport in streams. We here computed the average distance a molecule travels in the stream channel before entering the transient storage zone (L_s [L], Mulholland et al., 1997):

$$L_s = \frac{v}{\alpha} \qquad \qquad \text{Eq.5}$$

265   The average time spent by a molecule in the transient storage zone (T_{sto} [T]) is evaluated as (Thackston and Schnelle, 1970):

$$T_{sto} = \frac{A_{TS}}{\alpha A} \qquad \qquad \text{Eq.6}$$

We computed the average water flux through the storage zone per unit length of stream channel to interpret the magnitude of flux between the stream channel and the transient storage zone (q_s [L^2/T], Harvey et al., 1996):

270
$$q_s = \alpha A \qquad\qquad\qquad\qquad\qquad \text{Eq.7}$$

However, L_s, T_{sto}, and q_s metrics are not able to encompass both the role of advective transport and of the transient storage. Thus, we also evaluated F_{MED} [-] that accounts for the median travel time due to advection-dispersion and transient storage and for the travel time only due to advection-dispersion (Runkel, 2002):

$$F_{MED} \cong \left(1 - e^{\left(-L\frac{\alpha}{v}\right)}\right)\frac{A_{TS}}{A_{TS}+A} \qquad\qquad\qquad\qquad \text{Eq.8}$$

275   Increasing values of F_{MED} have to be interpreted as increasing relative importance of the storage zone in the solute transport downstream (Gooseff et al., 2013; Runkel, 2002).

3. Results

3.1 Identifiability of TSM parameters and comparison with OTIS-P and OTIS-MCAT results

The identifiability of TSM parameters was studied for the three tracer experiments injections (indicated as E1,
280   E2, E3, cfr. paragraph 2.1) for two distinct cases: stream velocity considered as a variable parameter, and velocity considered equal to v_{peak} ($v = v_{peak}$). Global Identifiability analysis results are here reported via only parameter values plotted against the corresponding RMSE values.

3.1.1 Transient storage modelling with stream velocity as varying model parameter

After the first TSM iteration the global identifiability analysis indicated that v, D, and α parameters are sensitive
285   with a unique and identifiable performance peak (K of K-S test always > 0.22 and p < 0.05 for each tracer experiment). However, A and A_{TS} appeared non- or poorly-identifiable for the three investigated BTCs (Figure 2, green dots, p-value of the K-S test for A_{TS} > 0.05 for each tracer experiment). The dynamic identifiability analysis provided clearer insights into the effect of the TSM parameters on the BTC and their identifiability ranges compared to the global identifiability analysis. v and α were confirmed to be the most identifiable and informative
290   parameters in the rising limb, the peak and the tail of the BTC (information content > 0.66; Figure 3b, h). A and D were mostly identifiable and informative during the rising limb and the tail of the BTC (information content > 0.50; Figure 3c-f). A_{TS} was uncertain and non-informative in most sections of the BTC (information content < 0.33; Figure 3i, j). However, the identifiability of A_{TS} increased in the tail of the BTC, where the information content was above 0.66 for A_{TS} below 5.356 m^2 for E1 (Figure 3i, j), and for A_{TS} below 5.4315 m^2 and 4.6404 m^2
295   respectively for the BTCs of E2 and E3 (results not shown).

The global identifiability of TSM parameters increased through the iterative model approach and when A_{TS} or α were poorly or non-identifiable (p-value of the K-S test for A_{TS} > 0.05) TSM performances approached at best $RMSE_{ADE}$ (Figure 2, green, yellow and blue dots). After four (for E1 and E2) or five (for E3) TSM iterations, the parameter values plotted against the corresponding RMSE values showed univocal increase of performance
300   toward unique values for v, A, D, α, and A_{TS} (Figure 2, orange dots), and the RMSE of the best performing parameter sets decreased below $RMSE_{ADE}$ (Figure 2, black horizontal line). Also, the CDF corresponding to the best 0.1 % of the results deviated both from the 1:1 line and from parameter CDF corresponding to the best 10 % of the results (results not shown). These conditions, coupled with the K of K-S test always larger than 0.1 (average

K for all the TSM parameters equal to 0.36, and p-value < 0.05) indicated parameter identifiability and the

305   finalization of the iterative TSM approach.

The dynamic identifiability analysis for the last TSM iteration showed that the advection-dispersion parameters were important in controlling the rising limb and the tail of the BTC (results reported only for E1, Figure 4a-f), while α was particularly important for controlling the tail (Figure 4g, h) and A_{TS} for controlling the rising limb and the tail of the BTC (Figure4i, j).

310   **3.1.2 Transient storage modelling with stream velocity set equal to v_{peak}**

The OTIS-MCAT results produced low p-values for each TSM parameter after the K-S test ($p < 0.05$, $K > 0.12$) indicating parameter identifiability. However, global identifiability analysis showed that the distribution of TSM parameters did not converge towards univocal and optimal parameter values suggesting that identifiability of TSM parameters was rather uncertain with TSM outcomes performing worse than the ADE (Figure 5, green dots). The

315   global identifiability of TSM parameters increased considerably through the iterative model approach. After the third TSM iteration, the best performing parameter sets approached unique parameter values (Figure 5, blue dots) and the CDF corresponding to the best 0.1% of the results deviated from 1:1 line and from the CDF of the best 10% of the results (results not shown). These conditions, together with K of K-S test always > 0.25 and p-value < 0.05 for each TSM parameter and tracer experiment, showed a clear increase of identifiability compared to the

320   initial OTIS-MCAT results. The increase of parameter identifiability was followed by a sharp increase of model performance, with the best performing parameter sets at the end of the iterative approach having RMSE values below $RMSE_{ADE}$ for all the investigated BTCs (Figure 5, blue dots and black line). Dynamic identifiability analysis for the last TSM iteration indicated that A and D control respectively the falling limb and the rising limb of the BTC (Figure 6a-d, results of E1). α controlled both on the rising limb, falling limb and tail of the BTC (Figure 6e-

325   f) and A_{TS} controlled both the falling limb and the tail of the BTC (Figure 6g, h).

Results from OTIS-P showed parameter identifiability with narrow 95% confidence range for the A_{TS} and A, while D and α parameters were estimated with higher uncertainty due to larger 95% confidence range (Figure 2, 4). The parameter sets obtained via OTIS-P (Figure 2, 4, red vertical dashed line) were close to approach the best fitting results at the end of the used iterative approach, regardless the fact the velocity was considered as a variable

330   parameter (Figure 2), or was it considered equal to v_{peak} (Figure 5, Table 2).

3.2 Variation of transport metrics with increasing identifiability of TSM parameters

The investigated transport metrics showed high uncertainty as long TSM parameters were poorly or non-identifiable (Figure 2, 5, green and yellow dots). This was particularly evident after the first and second TSM iterations, when the 100 best performing parameter sets showed T_{sto} values spanning over nine orders of magnitude

335   (Figure 7d-f), while both L_s and q_s spanned over three orders of magnitude (Figure 7a-c, g-i). When TSM parameters were poorly identifiable, the values of the metrics obtained when stream velocity was considered as a variable parameter (Figure 7, blue boxplots, first TSM iteration) and when stream velocity was considered equal to v_{peak} (OTIS-MCAT, Figure 7, orange boxplots, first TSM iteration) showed relevant differences. When v was considered variable together with the others TSM parameters, the best performing parameter set after the first

340   TSM iteration showed a non-negligible role of transient storage in solute transport for the investigated discharge conditions. This was indicated by the high values of L_s (from ~2 km for E1 to ~69 m for E3), by the rather low

exchange flux q_s (from 0.06 l/s for E1 to 8.8 l/s for E3), and by the long solute residence time in the storage zone T_{sto} (ranging from ~ 140 days for E1 to ~ 15 hrs for E3). Conversely, very different values were obtained for the transport metrics when v was fixed equal to v_{peak}. The results from OTIS-MCAT showed a rather fast exchange

345 flux of the active stream with the transient storage zone (q_s ranged from ~23 l/s for E1 to ~121 l/s for E3), a rather similar L_s for the three tracer experiments (~10 m), and T_{sto} increased with increasing discharge (from ~12 sec for E1 to ~3 sec for E3).

However, once the TSM parameters were identifiable all the transport metrics converged toward constrained values and consistent with OTIS-P results (Figure 7). This was achieved whether stream velocity was kept fixed

350 or was variable in the modelling procedure. Results of the last TSM iteration showed that the investigated transport metrics have low dispersion around the median, and that the median almost coincides with the result of the best performing parameter set (Figure 7, red dots). When all TSM parameters were identifiable for the three tracer experiments, the transport metrics showed increasing q_s (from ~2.7 l/s for E1 to ~23 l/s for E3), increasing L_S (from ~50 m for E1 to ~100 m for E3), and decreasing T_{sto} (from ~150 s for E1 to ~33 s for E3) with increasing

355 discharge conditions (from E1 to E3). F_{med} did not change widely between the TSM iterations since the median of the best performing 100 parameter sets varied always between 0.04 and 0.2 (Figure 7j-l). However, together with the other investigated transport metrics, the dispersion of F_{med} values around the median decreased with increasing identifiability of TSM parameters.

4. Discussion

360 ### 4.1 The importance of the identifiability of TSM parameter for correct interpretation of hydrological processes

Our results demonstrated that poor or non-identifiability of TSM parameters can result in a wrong hydrological interpretation of the processes controlling solute transport in streams. Our results showed that with increasing discharge conditions L_s and q_s increased, T_{sto} decreased, and F_{med} was rather stable for simulations where the TSM

365 parameters were identifiable (cfr paragraph 3.2). The low uncertainty and the values of the investigated transport metrics suggested that the transient storage at the experimental site was most probably controlled by in-stream dead-zones (Boano et al., 2014). The observed link of L_s, q_s, and T_{sto} values with discharge (Figure 7) also suggested that the transient storage at our site became less important in controlling solute transport with increasing discharge conditions. This could be explained as with increasing discharge conditions the wet stream areas and

370 the water depth increased more than the wetted perimeter that embeds dead-zoned and streambed heterogeneities (Gooseff et a., 2008). This would have caused a progressive increase of piston-flow transport and a reduced role of low-flow areas for in-stream solute retention with increasing discharge condition.

However, if we would have based the process interpretation on simulations before we reached identifiability of TSM parameters, the conclusions would have been different. The values for the transport metrics obtained when

375 v was considered variable together with the others TSM parameters, together with published results about solute residence time in the hyporheic zone and in in-stream channel (Boano et al., 2014; Gooseff et al., 2005) could have been interpreted in a way that the transient storage was controlled by in-stream dead-zones during high-discharge events and by a low rate hyporheic exchange that lasted several weeks at low flow conditions (Figure 7, blue boxplots and first TSM iteration). Conversely, results from OTIS-MCAT might have been interpreted in

380 a way that transient storage of the studied stream channel was controlled by dead-zones at the lowest flow

conditions and by in-stream turbulences that caused solute retention in the transient storage zone to last ~3 seconds during high-flow events (Figure 7, orange boxplots and first TSM iteration).

Compared to the inverse modelling approach (OTIS-P), we were able to interpret modelling results without the uncertainty typically associated with inverse modelling outcomes that do not assess the identifiability of the

385   parameters and of the associated hydrological metrics across their feasible spectrum of values (Ward et al., 2017; Kelleher et al., 2019). Furthermore, by reaching identifiability of TSM parameters, we also obtained a low uncertainty of the transport metrics and a more robust process interpretation compared to the standard random sampling approach, which showed poor identifiability of TSM parameters (Figure 2, 5, green dots). The results presented in this work offer new insights into the role that identifiable TSM parameters have on the interpretation

390   of solute transport processes in streams. The limitations of relying on poorly identifiable parameters in TSMs are generally well known (Wagner & Harvey, 1997; Ward et al., 2017; Kelleher et al., 2019), however the majority of TSM studies did not address parameter uncertainty and/or draw conclusions based on non-identifiable transient storage parameters (see the commentary of Knapp & Kelleher, 2020). Identifiability of TSM parameters is nowadays crucial, since the amount of contradicting interpretation of TSM results hamper our ability to link

395   specific physical processes and hydrologic conditions to an increase or decrease of transient storage parameters (Ward & Packman, 2019). Contradicting hydrologic interpretation and modelling outcomes can be driven by different hydrologic and morphological setting of the investigated stream reach (Gooseff et al., 2005, Kelleher et al., 2013). Our study highlights how the poor or non-identifiability of TSM parameters can result in a wrong estimation of solute transport metrics in streams and can play a major role on the hydrologic interpretation of

400   modelling results.

4.2 Challenges associated to parameter identifiability in TSMs

We showed that non-identifiability of α and A_{TS} in TSM can result from the assumption $v = v_{peak}$, the selected number of parameter sets, and the parameter space used for the random sampling. Our results indicated that v interacts with α and A_{TS} in transient storage models. This was particularly evident when v was variable together

405   with the other TSM parameters, and the non-identifiability of A_{TS} was coupled with identifiability of v and α (Figure 2, green and yellow dots). On the contrary, A_{TS} was found to be identifiable and α to be non-identifiable when v was fixed equal to v_{peak} (Figure 5, yellow dots). It is known that a separate evaluation of the advection-dispersion parameters from the transient storage parameters can result in misevaluation of transient storage parameters due to the high parameter interaction (Knapp & Kelleher, 2020). However, no study so far investigated

410   the role of the assumption v equals v_{peak} on the non-identifiability of α or A_{TS} in random sampling approach for TSMs. Despite the observed interaction between v, α and A_{TS}, our study also showed that when all TSM parameters are identifiable the best performing parameter sets showed similar values no matter if the stream velocity was fixed equal to v_{peak}, or was considered a variable parameter (Table 2).

Our results showed that non-identifiability of transient storage parameters might indicate inaccurate TSM results.

415   This was evident from TSM iterations showing non-identifiability of α and A_{TS}, with the best model performances approaching the $RMSE_{ADE}$ (Figure 2, 5, black line). This outcome indicated that non-identifiability of α or A_{TS} is linked to an underestimation of transient storage process with the optimal modelled BTCs mimicking the ADE. Similar to our results, many authors showed non-identifiability of TSM parameters in random-sampling approach. Previous research found identifiable A_{TS} coupled with non-identifiable α (Camacho & González, 2008; Kelleher

420   et al., 2013; Wagener et al., 2002; Wlostowski et al., 2013), while other TSM applications found α to be
identifiable coupled with non-identifiability for A_{TS} (Kelleher et al., 2019), or α or A_{TS} to be both non-sensitive
and non-identifiable (Camacho & González, 2008; Ward et al., 2013; Ward et al., 2017). Random-sampling
approach are generally considered more informative than inverse-modelling approach (Ward et al., 2017; Knapp
and Kelleher, 2020), however our results indicate that model outcomes showing non-identifiability of transient

425   storage parameters should be used with particular caution for model interpretation due to the rather different
parameter estimation when TSM parameters were identifiable and non-identifiable (Figure 2, 5).

Identifiability of TSM is commonly studied via random sampling approaches using between 800 and 100,000
parameter sets sampled from a parameter space spanning several orders of magnitude (Table 1). Our study
demonstrated that it is unlikely to reach parameter identifiability via random-sampling approach using less than

430   100,000 parameter sets when investigating a rather large parameter space of TSM parameters (Table 1). Our
results showed identifiability only after the third TSM iteration, between 230,000 and 345,000 parameter sets and
by narrowing the investigated parameter range twice (Figure 2, 5, blue dots). While the range and the order of
magnitude of advection-dispersion parameter can be estimating by using the ADE, the ranges where α and A_{TS} are
identifiable are never known a-priori and random sampling approaches need to target a parameter space large

435   enough to capture the distribution of transient storage parameters on their entire feasible space. Thus, the peak of
performance for the transient storage parameters can be so narrow that it can be missed by the random sampling
approach or by only a low number of selections. Similar conclusions have been obtained by Ward et al. (2017),
who found by using the OTIS-MCAT model via 100'000 parameter sets that the TSM parameters were identifiable
only for one of the three investigated BTCs.

440   Other studies coupled random sampling approaches with behavioural thresholds to derive the most sensitive range
of TSM parameters and reduce parameter uncertainty, yet this was done to constrain only the range of A (Kelleher
et al., 2013; Ward et al., 2017). Here, we demonstrated the importance of coupling the behavioural threshold
approach with the global and dynamic identifiability analyses in successive TSM iterations (Kelleher et al., 2013;
Ward et al., 2017). The high information content (> 0.66, eg. Figure 3j) of α and A_{TS} on the tail of the BTC

445   provided valuable information to constrain the parameter space in successive TSM iterations. This approach
eventually allowed us to identify TSM parameters.

4.3 How TSM parameters control the rising limb, the peak, and the falling limb of the BTC

The relative high information content of transient storage parameters on the rising limb and the peak of the BTC,
coupled with the high information content of advection-dispersion parameters on the falling limb and tail of the

450   BTC (Figure 4, 6 8a, c, e) showed that all TSM parameters control solute arrival-time and solute retention in
stream channels. This outcome is in contradiction with the common interpretation of TSM parameters, where it is
assumed that the advection-dispersion parameters control the solute arrival time, while transient-storage
parameters are assumed to control the tail of the BTC (Bencala, 1983; Bencala et al., 2011; Bencala & Walters,
1983; Smith, 2005; Runkel, 2002). Following this common interpretation of the role of TSM parameters on the

455   BTC, some authors decomposed the BTC in an advective part and a transient storage part (Ward et al., 2019;
Wlostowski et al., 2017). This decomposition allows to quantify the role of advection-dispersion and transient
storage embedded in the BTC. However, it also implicitly assumes a negligeable role of advection-dispersion

parameters on the BTC tail, and of transient-storage parameters on the rising limb and peak of the BTC which is in contradiction with our findings (Figures 4, 6, 8).

460   Our work demonstrated that the assumption v equals v_{peak} might not be representative of the advection role on solute transport in streams. The assumption of stream velocity equalling v_{peak} implies that v_{peak} can encompass the effect of advection on the entire BTC or at least in the rising-limb and peak of the BTC. However, when v was used as calibration parameter, our results showed that v is one of the least meaningful parameters for simulating the peak of the BTC at low discharge (Figure 4a, b; 8a), while higher information content for v is obtained at

465   higher discharge rates for values larger than v_{peak} at the peak of the BTC (Figure 8c, e, dynamic identifiability plots not shown). Our results also highlighted how assuming v equals v_{peak} caused a stronger influence of α and weaker influence of A_{TS} on the BTC compared to the case when v is variable. Indeed, the dynamic identifiability analysis for the case v equals v_{peak} underestimated the role of A and A_{TS} on the rising limb and peak of the BTC and overestimated the role of D and α on the rising limb of the BTC compared to the case when v was considered

470   variable together with the other TSM parameters (Figures 4, 6).

Several studies have addressed the identifiability of the TSM parameters for different sections of the BTC and showed both similar and contrasting outcomes to our findings (Figure 8g-l, Wagner & Harvey, 1997; Wagener et al., 2002; Scott et al., 2003; Wlostowski et al., 2013; Kelleher et al., 2013). The sensitivity of TSM parameters in different sections of the BTC might be driven by different physical settings or discharge condition of the study

475   sites. For example, the sensitivity of the TSM to α and A_{TS} is expected to increase for dispersive streams and alluvial stream channels, compared to mountain reaches with low or null hydrologic exchange with the hyporheic zone (Gooseff et al., 2005; Kelleher et al., 2013). However, our analysis also suggests that the different results on the importance of TSM parameters for certain sections of the BTC reported in Figure 8 could be driven by the selected random sampling approach and the non-identifiability of TSM parameters.

480   Results from Wagner & Harvey (1997) and Scott et al. (2003) are in partial agreement with our results for the case v equals v_{peak} (Figure 8b, d, f, g, h) suggesting a non-negligeable influence of D on the rising and the falling limb of the BTC and a significant role of α and A_{TS} on controlling the peak, the falling limb, and the tail of the BTC. However, our results also highlight a non-relevant role of A on the rising limb and the peak of the BTC in our experiments (Figure 8b, f). Our results also show the role of α on controlling the rising limb of the BTC

485   (Figure 8b, d, f), a result that was not found by other studies (Scott et al., 2003; Wagner & Harvey, 1997). This difference might be driven by different hydrologic conditions and physical settings of the study sites, by the methodologies used for accounting parameters sensitivity, by the parameter sampling procedure, or the strategy used to obtain the best-fitting parameter set.

Consistent with our results, the dynamic identifiability analysis of TSM parameters by Wagener et al. (2002) and

490   Wlostowski et al. (2013) indicated high identifiability of A on the rising limb of the BTC, while α and A_{TS} controlled the falling limb and the tail of the BTC (Figure 8h, j). However, they also report lower information content for D, A_{TS} and α on the BTC compared to our study. This difference may arise from the parameter range and number of simulations chosen by the authors that could affect the TSM results. Plots of the parameter values against the corresponding objective function in Wagener et al. (2002) and the regional sensitivity analysis in

495   Wlostowski et al. (2013) do not indicate parameter identifiability for A_{TS}, D and α. These results together with our identifiability plots when TSM parameters were poorly identifiable (Figures 2, 5, green and yellow plots) suggest that the range and the number of the parameter sets chosen by the authors could have been unsufficient to obtain

global sensitivity and identifiability of D, A_{TS} and α parameters. Similar to Wagener et al. (2002) and Wlostowski et al. (2013) results, our dynamic identifiability analysis showed no influence of A_{TS} on the majority of the BTC when it was non-identifiable (Figure 3).

Eventually, results from Kelleher et al., (2013) emphasize the roles that A and D have on large sections of the BTC (Figure 8k, l). While this is consistent with our findings (Figure 8a, c, e), Kelleher et al., (2013) also indicate that transient storage parameters have a rather weak influence only on the tail of the BTC (Figure 8k, l). Different sensitivity of TSM to transient storage parameters could be driven by the different approach used for evaluating the sensitivity (i.e. Sobol' sensitivity analysis). However, our results suggest that the number of parameter sets (42,000) selected by Kelleher et al. (2013) might be too small to obtain identifiability of TSM parameters compared to the rather large parameter range chosen for their Monte Carlo sampling (Table 1). Results by Kelleher et al., (2013) are very similar to our TSM iterations for cases where α was non-identifiable (v equals to v_{peak}, Figure 5 yellow dots, dynamic identifiability plots not shown). We also demonstrated that our results after the first and second TSM iterations are not useful for interpreting transient storage process, because of the non-identifiability of the TSM parameters and the low model performances ($RMSE \geq RMSE_{ADE}$ (Figure 5a-p, green and yellow dots).

5 Conclusion

There is a clear need in stream hydrology to better identify TSM parameters for simulating solute transport in streams. Here we address parameter identifiability by combining global identifiability analysis with dynamic identifiability analysis in an iterative approach to reduce parameter uncertainty in TSMs. Our results show that v interacts with the transient storage parameters. Namely, when v was variable together with the other TSM parameters, we found non-identifiable A_{TS} coupled with identifiable α. On the contrary, when v was assumed equal to v_{peak}, A_{TS} was found identifiable and α non-identifiable. We proved that non-identifiability of transient storage parameters can result in modelled BTC mimicking the ADE. Non-identifiable TSM simulations also severely misevaluated the solute retention time in the storage zone and the exchange flow with the storage zone, with a difference respectively up to four and two orders of magnitude compared to the results when the TSM parameters were identifiable.

We here validated our initial hypothesis that the BTC tail contains critical information on transient storage parameters, since we clearly reduced parameter uncertainty compared to the standard random-sampling approach coupled with the global identifiability analysis. Results obtained via inverse modelling approach are generally considered less informative about identifiability of TSM parameters compared to random-sampling approaches. However, random sampling approach rarely achieved identifiability in TSMs. As a complement to the existing body of literature, our work shows that the non-identifiability of α or A_{TS} occurring in prior TSM studies might be related to a lack of modelled transient storage exchange due to the narrow peak of performances that can be easily missed by the rather small number of simulations compared to the investigated parameter range. When all the TSM parameters were identifiable, the best performing parameter sets and the evaluated transport metrics converged toward unique values, regardless the fact the velocity was considered as a variable parameter or equal to v_{peak}. This allowed us a robust assessment of the hydrological processes governing the solute transport in the investigated site.

Our work highlights how both parameter evaluation and process interpretation in TSMs should be used with particular caution even if one parameter between α or A_{TS} is found uncertain and non-identifiable. Stream

hydrologists are currently unable to obtain univocal physical process interpretation from modelling results due to contradictory interpretation of TSM parameters and lack of parameter identifiability in the published studies. Our work casts new lights on the opportunity to increase parameter identifiability and achieve stronger hydrologic

540 interpretation of the processes governing solutes transport in streams.

6 Acknowledgements

7 References

Bahr, J. M., and J. Rubin (1987), Direct comparison of kinetic and local equilibrium formulations for solute transport affected by surface reactions, Water Resour. Res. 23(3), 438–452, doi:198710.1029/WR023i

550   003p00438.

Beltaos, S., & Day, T. J. (1978). Field Study of Longitudinal Dispersion. *Canadian Journal of Civil Engineering*, *5*(4), 572–585. https://doi.org/10.1139/l78-062

Bencala, K. E. (1983). Simulation of solute transport in a mountain pool-and-riffle stream with a kinetic mass transfer model for sorption. *Water Resources Research*, *19*(3), 732–738.

555 Bencala, K. E., & Walters, R. A. (1983). Simulation of solute transport in a mountain pool-and-riffle stream: A transient storage model. *Water Resources Research*, *19*(3), 718–724. https://doi.org/10.1029/WR019i003p00718

Bencala, K. E., Gooseff, M. N., & Kimball, B. A. (2011). Rethinking hyporheic flow and transient storage to advance understanding of stream-catchment connections. *Water Resources Research*, *47*(3), 1–9.

560   https://doi.org/10.1029/2010WR010066

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298. https://doi.org/10.1002/hyp.3360060305

Beven, K., Gilman, K., & Newson, M. (1979). Flow and flow routing in upland channel networks. *Hydrological Sciences Bulletin*, *24*(3), 303–325. https://doi.org/10.1080/02626667909491869

565 Bonanno, E., Blöschl, G., & Klaus, J. (2021). Flow directions of stream-groundwater exchange in a headwater catchment during the hydrologic year. *Hydrological Processes*, *35*(8), 1–18. https://doi.org/10.1002/hyp.14310

F. Boano, J. W. Harvey, A. Marion, A. I. Packman, R. Revelli, L. Ridolfi, A. Wörman, Hyporheic flow and transport processes. Rev. Geophys. 52, 603–679 (2014).

570 Butterworth, J. A., Hewitt, E. J., & McCartney, M. P. (2000). Discharge Measurement Using Portable Dilution Gauging Flowmeters. *Water and Environment Journal*, *14*(6), 436–441. https://doi.org/10.1111/j.1747-6593.2000.tb00291.x

Camacho, L. A., & González, R. A. (2008). Calibration and predictive ability analysis of longitudinal solute transport models in mountain streams. *Environmental Fluid Mechanics*, *8*(5–6), 597–604. https://doi.org/10.1007/s10652-008-9109-0

Castro, N. M., & Hornberger, G. M. (1991). Surface-subsurface water interactions in an alluviated mountain stream channel. *Water Resources Research*, *27*(7), 1613–1621. https://doi.org/10.1029/91WR00764

Choi, J., Harvey, J. W., & Conklin, M. H. (2000). Characterizing multiple timescales and storage zone interaction that affect solute fate and transport in stream. *Water Resour. Res.*, *36(6)*(6), 1511–1518.

Gooseff, M. N., Wondzell, S. M., Haggerty, R., & Anderson, J. (2003). Comparing transient storage modeling and residence time distribution (RTD) analysis in geomorphically varied reaches in the Lookout Creek basin, Oregon, USA. *Advances in Water Resources*, *26*(9), 925–937. https://doi.org/10.1016/S0309-1708(03)00105-2

Gooseff, M. N., LaNier, J., Haggerty, R., & Kokkeler, K. (2005). Determining in-channel (dead zone) transient storage by comparing solute transport in a bedrock channel-alluvial channel sequence, Oregon. *Water Resources Research*, *41*(6), 1–7. https://doi.org/10.1029/2004WR003513

Gooseff, M. N., Bencala, K. E., Wondzell, S. M., Service, U. F., Northwest, P., & Sciences, O. F. (2008). Solute transport along stream and River Networks. In S. P. Rice, A. G. Roy, & B. L. Rhoads (Eds.), *River Confluences, Tributaries and the Fluvial Network*. John Wiley & Sons, Ltd. Retrieved from https://onlinelibrary.wiley.com/doi/book/10.1002/9780470760383

Gooseff, M. N., Briggs, M. A., Bencala, K. E., McGlynn, B. L., & Scott, D. T. (2013). Do transient storage parameters directly scale in longer, combined stream reaches? Reach length dependence of transient storage interpretations. *Journal of Hydrology*, *483*, 16–25. https://doi.org/10.1016/j.jhydrol.2012.12.046

Haggerty, R., Wondzell, S. M., & Johnson, M. A. (2002). Power-law residence time distribution in the hyporheic zone of a 2nd-order mountain stream. *Geophysical Research Letters*, *29*(13), 18-1-18-4. https://doi.org/10.1029/2002GL014743

Hart, D. R., Mulholland, P. J., Marzolf, E. R., DeAngelis, D. L., & Hendricks, S. P. (1999). Relationships between hydraulic parameters in a small stream under varying flow and seasonal conditions. *Hydrological Processes*, *13*(10), 1497–1510. https://doi.org/10.1002/(SICI)1099-1085(199907)13:10<1497::AID-HYP825>3.0.CO;2-1

Harvey, J W, Wagner, B. J., & Bencala, K. E. (1996). Evaluating the Reliability of the Stream Tracer Approach to Characterize Stream Subsurface Water Exchange. *Water Resources Research*.

Harvey, Judson W, & Wagner, B. J. (2000). Quantifying hydrologic interactions between streams and their subsurface hyporheic zones. In: *Streams and Ground Waters*. (ed. by J. A. Jones & P. J. Mulholland), 3-44: Academic Press, San Diego, USA. https://doi.org/10.1016/B978-012389845-6/50002-8

Hays, J. R. (1966). Mass transport mechanisms in open channel flow. *PhD Dissertation*.

Hissler, C., Martínez-Carreras, N., Barnich, F., Gourdol, L., Iffly, J. F., Juilleret, J., et al. (2021). The Weierbach experimental catchment in Luxembourg: A decade of critical zone monitoring in a temperate forest - from hydrological investigations to ecohydrological perspectives. *Hydrological Processes*, *35*(5), 1–7. https://doi.org/10.1002/hyp.14140

Kelleher, C., Wagener, T., McGlynn, B., Ward, A. S., Gooseff, M. N., & Payn, R. A. (2013). Identifiability of transient storage model parameters along a mountain stream. *Water Resources Research*, *49*(9), 5290–5306. https://doi.org/10.1002/wrcr.20413

615 Kelleher, Christa, Ward, A., Knapp, J. L. A., Blaen, P. J., Kurz, M. J., Drummond, J. D., et al. (2019). Exploring Tracer Information and Model Framework Trade-Offs to Improve Estimation of Stream Transient Storage Processes. *Water Resources Research*, *55*(4), 3481–3501. https://doi.org/10.1029/2018WR023585

Knapp, J. L. A., & Cirpka, O. A. (2017). Determination of hyporheic travel time distributions and other
620 parameters from concurrent conservative and reactive tracer tests by local-in-global optimization. *Water Resources Research*, *53*(6), 4984–5001. https://doi.org/10.1002/2017WR020734

Knapp, J. L. A., & Kelleher, C. (2020). A Perspective on the Future of Transient Storage Modeling: Let's Stop Chasing Our Tails. *Water Resources Research*, *56*(3), 1–7. https://doi.org/10.1029/2019WR026257

625 Krause, S., Hannah, D. M., Fleckenstein, J. H., Heppell, C. M., Kaeser, D., Pickup, R., et al. (2011). Interdisciplinary perspectives on processes in the hyporheic zone. *Ecohydrology*, *4*(26 November 2010), 481–499. https://doi.org/10.1002/eco

Krause, S., Lewandowski, J., Grimm, N. B., Hannah, D. M., Pinay, G., McDonald, K., et al. (2017). Ecohydrological interfaces as hot spots of ecosystem processes: Ecohydrological interfaces
630 as hot spots. *Water Resources Research*, 53, 6359– 6376. https://doi.org/10.1002/2016WR019516

Lemke, D., Liao, Z., Wöhling, T., Osenbrück, K., & Cirpka, O. A. (2013). Concurrent conservative and reactive tracer tests in a stream undergoing hyporheic exchange. *Water Resources Research*, *49*(5), 3024–3037. https://doi.org/10.1002/wrcr.20277

Morrice, J. A., Valett, H. M., Dahm, C. N., & Campana, M. E. (1997). Alluvial characteristics,
635 groundwater–surface water exchange and hydrological retention in headwater streams. *Hydrological Processes*, *11*(3), 253–267. https://doi.org/10.1002/(SICI)1099-1085(19970315)11:3<253::AID-HYP439>3.0.CO;2-J

Mulholland, P. J., Marzolf, E. R., Webster, J. R., Hart, D. R., & Hendricks, S. P. (1997). Evidence that hyporheic zones increase heterotrophic metabolism and phosphorus uptake in forest streams.
640 *Limnology and Oceanography*, *42*(3), 443–451. https://doi.org/10.4319/lo.1997.42.3.0443

Rathfelder, K. (2016). *Modelling Tools for Estimating Effects of Groundwater Pumping on Surface Waters.* Province of B.C.: Ministry of Environment, Water Science Series.

Rodriguez, N., Pfister, L., Zehe, E., & Klaus, J. (2021). Testing the truncation of travel times with StorAge Selection functions using deuterium and tritium as tracers. *Hydrology and Earth System Sciences*
645 *Discussions*, (October), 1–37. https://doi.org/10.5194/hess-2019-501

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

Runkel, R. L. (1998). *One-dimensional transport with inflow and storage (OTIS): a solute transport model*
650 *for streams and rivers. U.S. Geol. Surv. Water Resour. Invest. Rep.; 98–4018; Denver.* University of

Michigan Library. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/wrcr.20277#wrcr20277-bib-0038

Runkel, R. L. (2002). A new metric for determining the importance of transient storage. *Journal of the North American Benthological Society*, *21*(4), 529–543. https://doi.org/10.2307/1468428

Schmid, B. H. (2003). Temporal moments routing in streams and rivers with transient storage. *Advances in Water Resources*, *26*(9), 1021–1027. https://doi.org/10.1016/S0309-1708(03)00086-1

Schmid, B. H. (2008). Can Longitudinal Solute Transport Parameters Be Transferred to Different Flow Rates? *Journal of Hydrologic Engineering*, *13*(6), 505–509. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:6(505)

Scott, D. T., Gooseff, M. N., Bencala, K. E., & Runkel, R. L. (2003). Automated calibration of a stream solute transport model: Implications for interpretation of biogeochemical parameters. *Journal of the North American Benthological Society*, *22*(4), 492–510. https://doi.org/10.2307/1468348

Smith, J. W. N. (2005). *Groundwater – surface water interactions in the hyporheic zone* (2005th ed.). Bristol: Environment Agency. Retrieved from www.environment-agency.gov.uk

Taylor, G. (1921). Diffusion by continuous movements. *Proceedings of the London Mathematics Society*, *2*(20), 196–212.

Taylor, G. (1954). The dispersion of matter in turbulent flow through a pipe. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *223*(1155), 446–468. https://doi.org/10.1098/rspa.1954.0130

Triska, F. J., Kennedy, V. C., Avanzino, R. J., Zellweger, G. W., & Bencala, K. E. (1989). Retention and Transport of Nutrients in a Third-Order Stream: Hyporheic Processes. *Ecological Society of America*, *70*(6), 1893–1905. Retrieved from http://www.jstor.org/stable/1938120

Wagener, T., Lees, M. J., & Wheater, H. S. (2002). A toolkit for the development and application of parsimonious hydrological models. *Water Resources Publications*, *2*, 34.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., & Gupta, H. V. (2003). Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological Processes*, *17*(2), 455–476. https://doi.org/10.1002/hyp.1135

Wagener, Thorsten, & Kollat, J. (2007). Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox. *Environmental Modelling and Software*, *22*(7), 1021–1033. https://doi.org/10.1016/j.envsoft.2006.08.017

Wagener T, Camacho LA, Wheater HS. (2002). Dynamic identifiability analysis of the transient storage model for solute transport in rivers. Journal of Hydroinformatics 4: 199–211.

Wagner, B. J., & Harvey, J. W. (1997). Experimental design for estimating parameters of rate-limited mass transfer: Analysis of stream tracer studies. *Water Resources Research*, *33*(7), 1731–1741. https://doi.org/10.1029/97WR01067

Ward, A. S., & Packman, A. I. (2019). Advancing our predictive understanding of river corridor exchange. *WIREs Water*, *6*(1), e1327. https://doi.org/10.1002/wat2.1327

Ward, A. S., Payn, R. A., Gooseff, M. N., McGlynn, B. L., Bencala, K. E., Kelleher, C. A., et al. (2013). Variations in surface water-ground water interactions along a headwater mountain stream:

690   Comparisons between transient storage and water balance analyses. *Water Resources Research*, *49*(6), 3359–3374. https://doi.org/10.1002/wrcr.20148

Ward, A. S., Kelleher, C. A., Mason, S. J. K., Wagener, T., McIntyre, N., McGlynn, B., et al. (2017). A software tool to assess uncertainty in transient-storage model parameters using Monte Carlo simulations. *Freshwater Science*, *36*(1), 195–217. https://doi.org/10.1086/690444

695   Ward, A. S., Schmadel, N. M., & Wondzell, S. M. (2018). Time-Variable Transit Time Distributions in the Hyporheic Zone of a Headwater Mountain Stream. *Water Resources Research*, *54*(3), 2017–2036. https://doi.org/10.1002/2017WR021502

Ward, A. S., Wondzell, S. M., Schmadel, N. M., Herzog, S., Zarnetske, J. P., Baranov, V., et al. (2019). Spatial and temporal variation in river corridor exchange across a 5th-order mountain stream network.
700   *Hydrology and Earth System Sciences*, *23*(12), 5199–5225. https://doi.org/10.5194/hess-23-5199-2019

White, D. S. (1993). Perspectives on Defining and Delineating Hyporheic Zones. *Journal of the North American Benthological Society*, *12*(1), 61–69. https://doi.org/10.2307/1467686

Wlostowski, A. N., Gooseff, M. N., & Wagener, T. (2013). Influence of constant rate versus slug injection
705   experiment type on parameter identifiability in a 1-D transient storage model for stream solute transport. *Water Resources Research*, *49*(2), 1184–1188. https://doi.org/10.1002/wrcr.20103

Wlostowski, A. N., Gooseff, M. N., Bowden, W. B., & Wollheim, W. M. (2017). Stream tracer breakthrough curve decomposition into mass fractions: A simple framework to analyze and compare conservative solute transport processes. *Limnology and Oceanography: Methods*, *15*(2), 140–153.
710   https://doi.org/10.1002/lom3.10148

Yin, J., Lu, W., Xin, X., & Zhang, L. (2011). Application of Monte Carlo sampling and Latin Hypercube sampling methods in pumping schedule design during establishing surrogate model. *ISWREP 2011 - Proceedings of 2011 International Symposium on Water Resource and Environmental Protection*, *1*, 212–215. https://doi.org/10.1109/ISWREP.2011.5892983

715   Zaramella, M., Marion, A., & Packman, A. I. (2006). Applicability of the Transient Storage Model to the hyporheic exchange of metals. *Journal of Contaminant Hydrology*, *84*(1–2), 21–35. https://doi.org/10.1016/j.jconhyd.2005.12.002

Zaramella, M., Marion, A., Lewandowski, J., & Nützmann, G. (2016). Assessment of transient storage exchange and advection-dispersion mechanisms from concentration signatures along breakthrough
720   curves. *Journal of Hydrology*, *538*(May), 794–801. https://doi.org/10.1016/j.jhydrol.2016.05.004

725

730 **Table 1. Summary of publications that address identifiability of TSM parameters with random sampling approach. We reported the used number of parameter sets and the parameter ranges, while in parenthesis it is reported the method used for the parameter sampling. "Double step" indicates that the sampling procedure was divided in two steps. In the first step, _A_ varied across a broad range and in the second step, it was varied across a narrowed range to cover the most sensitive-range of the parameter domain. Each of the two steps has a number of simulations equal to**
735 **half of the total number indicated in the table.**

Authors	Number of parameter sets	Range of TSM parameters	
Wagner & Harvey, 1997	800 (Monte Carlo)	A (m²)	0.02-0.6
		D (m²/s)	0.025-0.8
		A_{TS} (m²)	0.01-2
		α (1/s)	0.000005-0.001
Wagener et al., 2002	1,000 (Monte Carlo)	A (m2)	0.3-1.05
		D (m²/s)	0.1-0.225
		A_{TS} (m²)	0.1-0.5
		α (1/s)	0.00035-0.0025
Wlostowski et al., 2013	2,000 (Monte Carlo)	A (m²)	0.5-1.0
		D (m²/s)	0.5-1.5
		A_{TS} (m²)	0.05-0.5
		α (1/s)	10-4-10-3
Kelleher et al., 2013	42,000 (Double step Monte Carlo)	A (m2)	0.001-1.0 (In the second step, min and max A values chosen from the best 1,000 results of the first step)
		D (m²/s)	0.001-1.0
		A_{TS} (m²)	0.001-0.01
		α (1/s)	10-5-10-3
Ward et al., 2013	100,000 (Monte Carlo)	A (m²)	+-50% A_{peak}
		D (m2/s)	0.0001-5
		ATS (m2)	0.01-10
		α (1/s)	10-8-10-1
Ward et al., 2017	100,000 (Double step Monte Carlo)	A (m²)	0.1-1 (0.3-0.5 in the second step)
		D (m²/s)	0.01-10
		A_{TS} (m²)	0.01-1
		α (1/s)	10-5-10-1
Kelleher et al., 2019	27,000 (LHS)	A (m2)	1.0 - 3.0
		D (m2/s)	0.001 - 10
		A_{TS} (m²)	0.01 - 1
		α (1/s)	10-6 - 10-2
This manuscript	Second ADE – 35,000 (LHS)	v (m/s)	v_{peak} · 0.8 - velocity of the first recorded increase of concentration in the BTC
		A (m²)	+-20% A_{peak}
		D (m²/s)	0.0001 - D_{best} · 1.2
This manuscript	First TSM iteration – 115,000 (LHS)	v (m/s)	+-50% v_{ADE}
		A (m²)	+-50% A_{ADE}
		D (m²/s)	0.0001 - D_{ADE} · 2
		A_{TS} (m2)	0.00001 - 20
		α (1/s)	0.00001 - 0.1

740 **Table 2: Summary of the TSM results. OTIS-MCAT results refer to the case v = vpeak without any successive modification of the parameter space via dynamic identifiability analysis results. "Iterative TSM" indicate the best parameter sets obtained after the iterative TSM approach presented in Figure 1 and applied for the cases v = variable and v = vpeak. The best TSM results are indicated with bold font.**

		v	A	D	α	A_{TS}	$RMSE$
E1	ADE	0.0681	0.0395	0.0965	/	/	1.9423
	OTIS-P	0.0739	0.0364	0.0637	0.0006	0.0074	**0.6159**
	OTIS-MCAT	0.0739	0.0351	0.1339	0.0119	0.0051	2.7421
	Iterative TSM $v = variable$	0.0728	0.0369	0.0522	0.0013	0.0073	0.7229
	$v = v_{peak}$	0.0739	0.0359	0.0534	0.0013	0.0077	0.7681
E2	ADE	0.1746	0.054	0.1599	/	/	0.9982
	OTIS-P	0.1774	0.0509	0.1151	0.0016	0.0077	0.4152
	OTIS-MCAT	0.1774	0.0604	0.1271	0.0137	0.0033	1.4429
	Iterative TSM $v = variable$	0.1790	0.0523	0.1131	0.0018	0.0067	**0.3377**
	$v = v_{peak}$	0.1774	0.0528	0.1154	0.0015	0.0065	0.3696
E3	ADE	0.262	0.0874	0.2525	/	/	0.9894
	OTIS-P	0.275	0.081	0.1404	0.005	0.0144	**0.2544**
	OTIS-MCAT	0.275	0.0849	0.2441	0.0259	0.0073	1.2612
	Iterative TSM $v = variable$	0.2861	0.0818	0.1286	0.0064	0.0145	0.2697
	$v = v_{peak}$	0.275	0.083	0.1603	0.0037	0.0123	0.3109

745

750

755

760

765
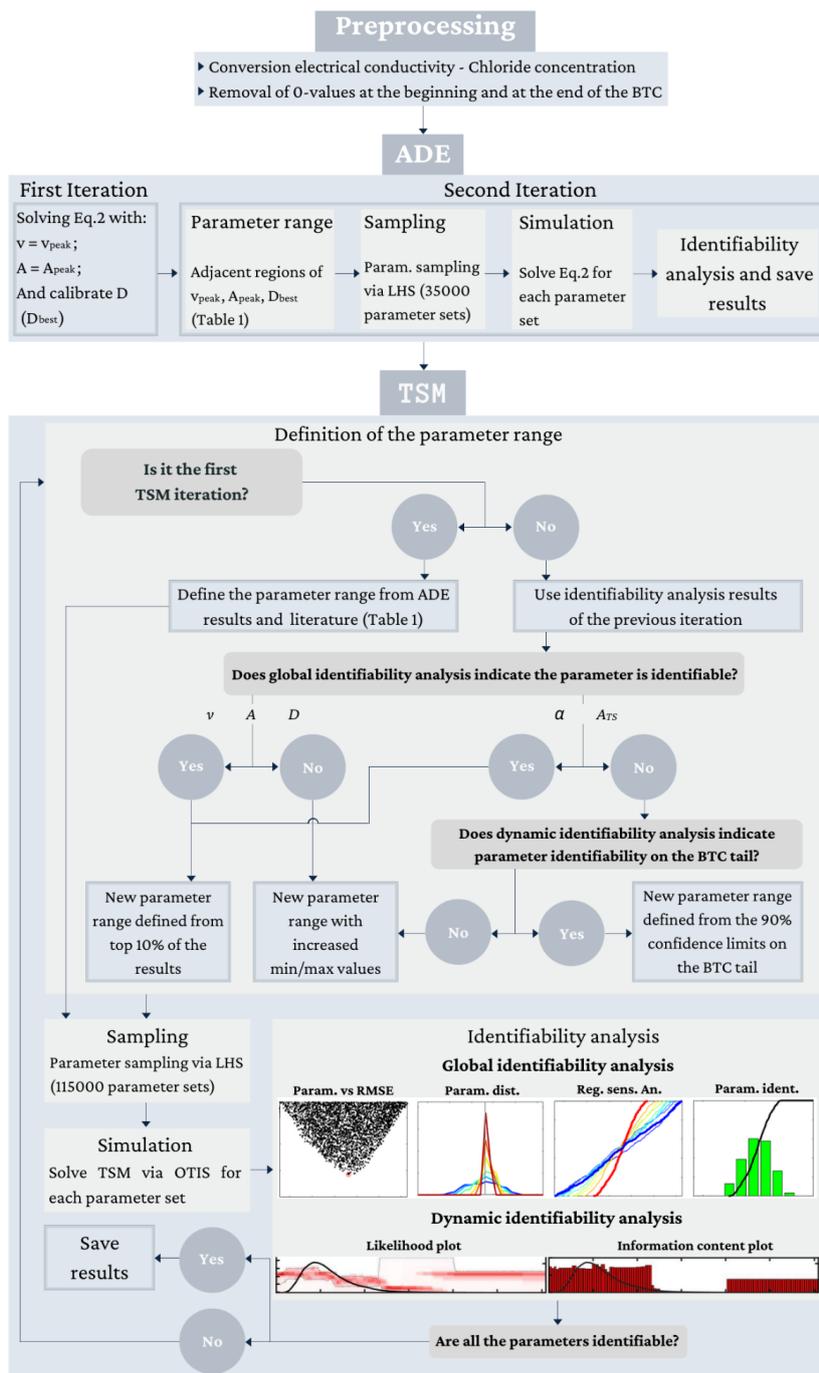
Figure 1: Conceptual modelling workflow. The parameters have the following unit of measurements: velocity v [m/s], cross-sectional area A [m2], longitudinal dispersion coefficient D [m2/s], exchange coefficient α [1/s], area of the transient storage zone ATS [m2].
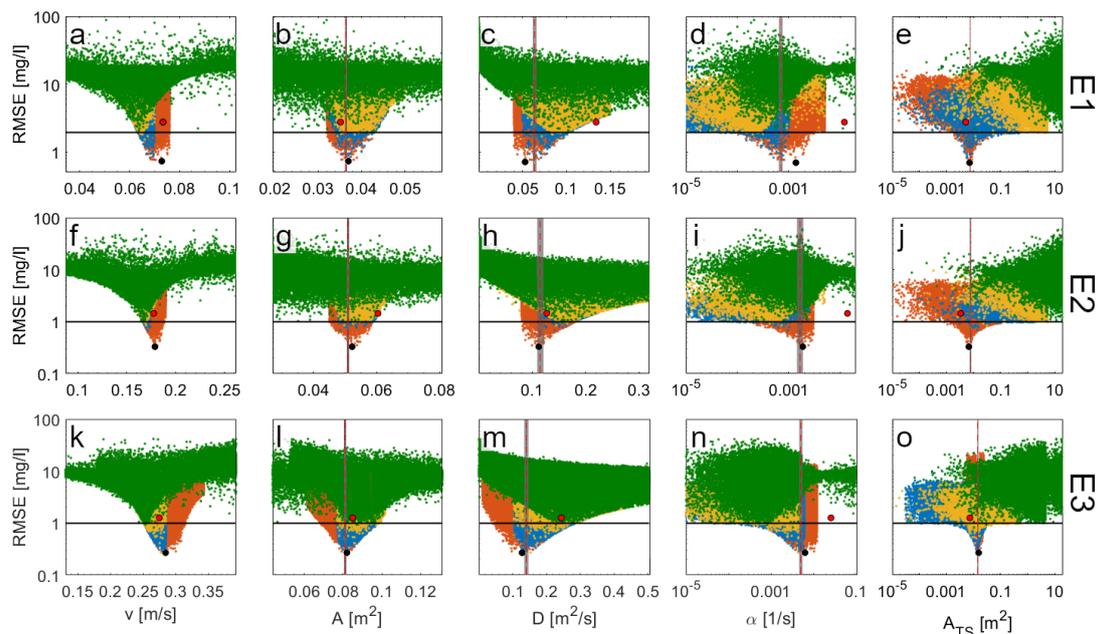
Figure 2. Parameter values plotted against the corresponding RMSE values for the TSM results conducted for the tracer injections (a-e) E1, (f-j) E2, and (k-o) E3. (a-j) Green, yellow, blue and orange dots indicate results respectively for the first, second, third, and fourth TSM iterations. (k-l) Green dots indicate results for the first and second TSM iterations, while yellow, blue and orange dots indicate results respectively for the, third, fourth, and fifth TSM iterations. Each TSM iteration was conducted via 115,000 parameter sets. The red dots indicate OTIS-MCAT results (best parameter set after the first TSM iteration for v equals vpeak) while the black dots indicate the best performing parameter value after the used iterative TSM approach. The horizontal black line indicates the $RMSE_{ADE}$ (Table 2). Vertical dashed red line indicates OTIS-P results, while the 95% confidence range for OTIS-P results are indicated via vertical grey areas.
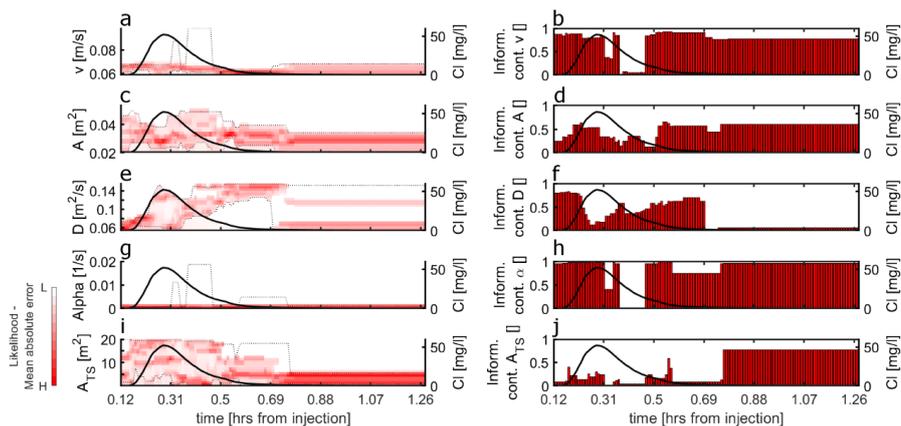
Figure 3. Dynamic identifiability analysis of TSM parameters for the first TSM iteration (E1, v considered as a varying model parameter). (a), (c), (e), (g), (i) likelihood distribution as function of parameter values at each time step. Black line indicates the observed BTC, dashed black lines indicate the 90% confidence limits. (b), (d), (f), (h), (j) indicate parameter information content (red bars) at each time step. Black line indicates the observed BTC.
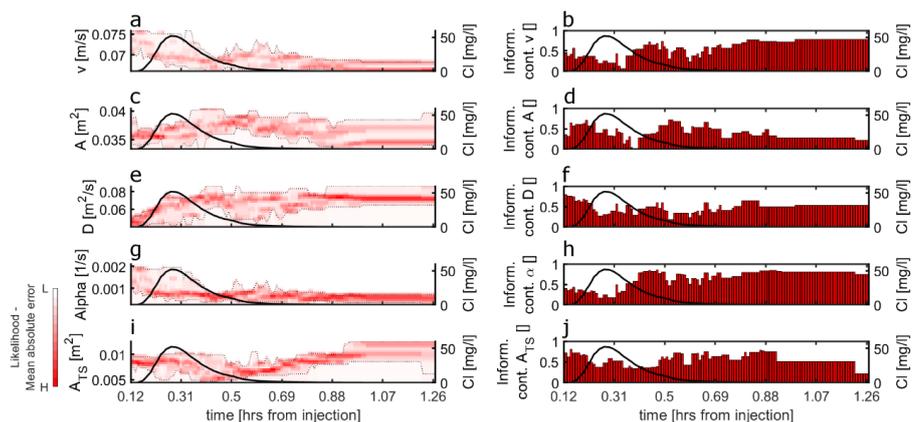
Figure 4. Same as Figure 3, but reporting dynamic identifiability results for the fourth and last TSM iteration (E1; *v* considered as a varying model parameter).

Hydrology and
Earth System
Sciences
Discussions
Open Access
EGU
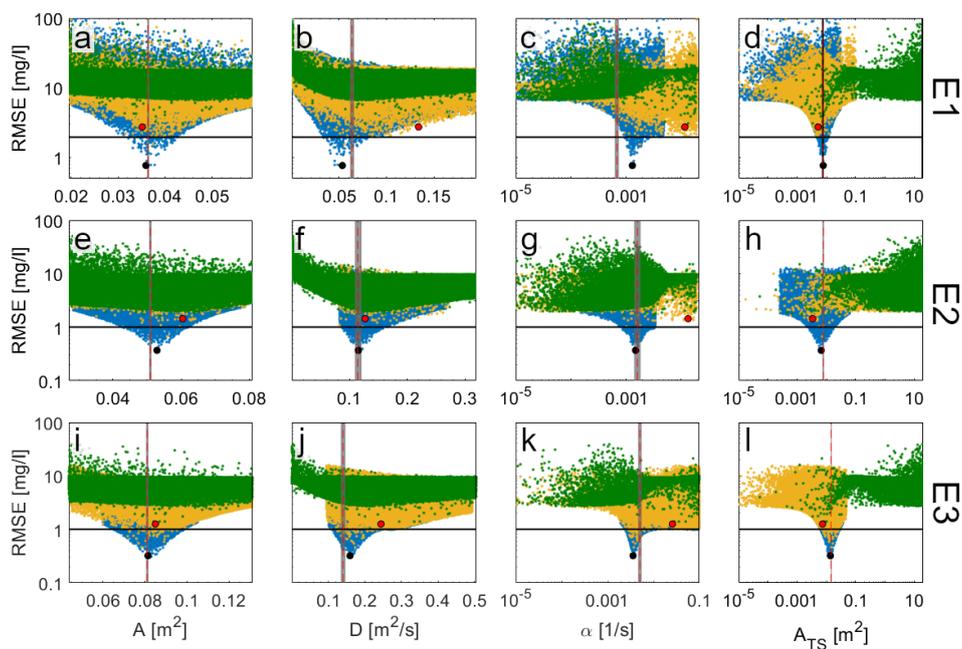
Figure 5. Same as Figure 2, but reporting TSM results when velocity was considered equal to v_{peak}.

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU
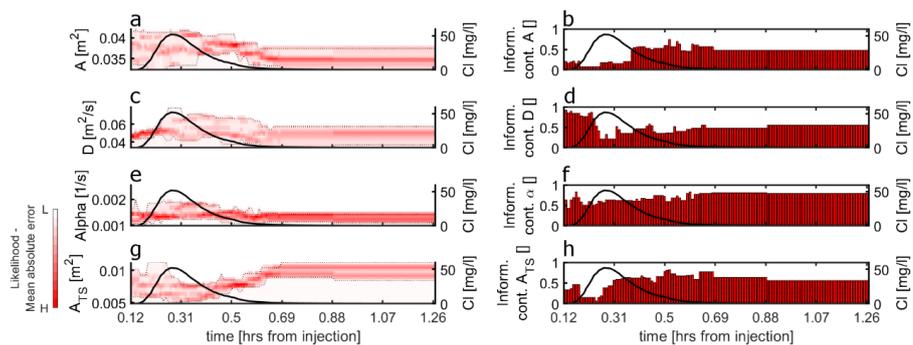
Figure 6. Same as Figure 3, but reporting dynamic identifiability results for the third TSM iteration (E1, $v = v_{peak}$).

850

855

860
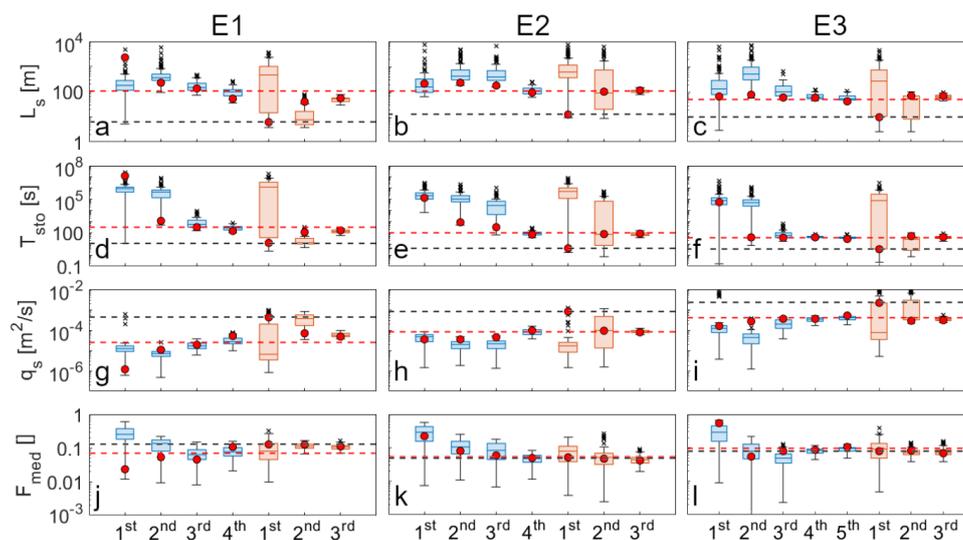
Figure 7. Boxplot of the investigated transport metrics for the best 100 parameter sets for the three simulated experiments. (a-c) L_s, (d-f) T_{sto}, (g,i) q_s, (j-l) F_{med}. Results are reported for (a, d, g, j, m) E1, (b, e, h, k, n) E2, and (c, f, i, l ,o) E3. On the x axis we indicated the *n-th* TSM iteration. Blue and red boxplots indicate results when velocity was respectively a varying model parameter and when it was kept fixed and equal to v_{peak}. Red dots indicate the transport metric values obtained via the parameter sets with lower RMSE. The red and the black horizontal dashed lines indicate respectively the transport metric obtained using the OTIS-P results and OTIS-MCAT results.
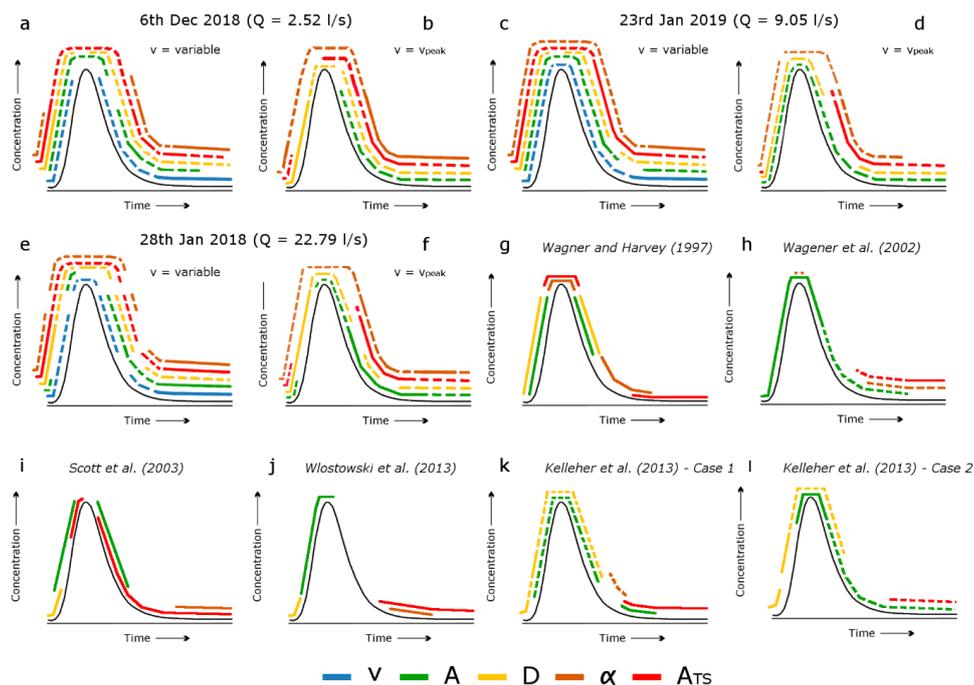
880 **Figure 8. Qualitative plots of the TSM parameter influence on different sections of the BTC. (a) and (b) qualitative parameter information content on the BTC for E1, (c, d) E2, and (e, f) E3. (g) Wagner and Harvey, 1997; (h) Wagener et al., (2002); (i) Scott et al., (2003); (j) Wlostowski et al., 2013; (k) Kelleher et al., (2013) for the case of a dispersive mountain stream (Case 1) and (l) Kelleher et al., (2013) for the case of a small low-flow mountain stream (Case 2). In plots (a-f) solid lines indicate an information content above 0.66 while dashed lines indicate an information content between 0.33 and 0.66. Plot (g) has been modified from Figure 7 in Wagener et al., (2002) in order to fit our 0.66 and 0.33 threshold classification in term of information content. Plots (k) and (l) indicate by solid and dashed lines if the parameters influence the model output by itself or through interactions (cfr. Section 6.1 Kelleher et al., 2013). Plots (g) and (i) describe the parameter influence evaluated via sensitivity evaluation (cfr. p. 1733, Wagner and Harvey, 1997) and dimensionless scaled sensitivities (cfr. Table 1 Scott et al., 2002), therefore the parameter influence is here described only using solid line. Plot (j) describes the parameter influence after DYNIA analysis, however information content plots were not reported by the authors, therefore the solid lines indicate the areas for the best performing parameters as indicated in Figure 2 of Wlostowski et al. (2013).**

895

Appendix A - Parameter sensitivity and identifiability

The interpretation of the parameter space is based on the sensitivity and identifiability of the *i-th* parameter on the

900   chosen model (the TSM) via a selected objective function used to compare model results with the observation (the
BTC) (Kelleher et al., 2019; Wagener et al., 2003; Wagener & Kollat, 2007; Ward et al., 2017; Wlostowski et al.,
2013). A parameter is called sensitive whenever a variation in the parameter value causes variations in the TSM
performances (Kelleher et al., 2019). A parameter is identifiable whenever the best-fit value of that parameter is
constrained on a relative narrow range across the entire distribution of the possible parameter values (Ward et al.,

905   2017). To assess identifiability of TSM parameters, we used parameter vs likelihood plots, identifiability plots,
regional sensitivity analysis plots and parameter distribution plots.

Parameter vs likelihood plots visualize the distribution of the investigated values of a certain parameter plotted
against the corresponding values of the objective function (Wagener et al., 2003; Wagener & Kollat, 2007).
Identifiable parameters are described in parameter vs likelihood plots by a univocal increase of model

910   performances approaching a certain optimum-value of the parameter (Figure A1a). Non-identifiable parameters
are described in parameter vs likelihood plots by a not-univocal increase of performances of the model in certain
parameter range (Figure A1b). Parameter distribution plots show probability density function (PDF) divided by
behavioural sets (from top 20% to top 0.1% of the results for the selected objective function) (Ward et al., 2017).
Identifiable parameters are indicated by narrow range of the PDF relative to the smaller behavioural sets (top

915   0.1%, 0.5% and 1% of the results) compared to a wider range of the PDF relative to the larger behavioural sets
(top 5%, 10% and 20% of the results) (Figure A1c). Non-identifiable parameters are defined by equally wide PDF
for the different investigated behavioural sets (Figure A1d). Regional sensitivity analysis plots are obtained after
dividing the population of the parameter by behavioural sets (from top 10% of the results to top 1% of the results
with 1% step for the selected objective function, Ward et al., 2017; Kelleher et al., 2019). Each objective function

920   population so obtained was transformed into cumulative distribution functions (CDFs) for equal size bins of the
parameter range (Kelleher et al., 2019; Wagener & Kollat, 2007). Sensitive parameters are identified by CDF for
the top 1% of the results deviating from the CDF for the top 10% of the results (Figure A1e). If the CDFs lay on
the 1:1 line, then the objective function is uniformly distributed across the parameter range which indicates
parameter unsensitivity (Figure A1f). Identifiability plots display the CDF of the objective function across the

925   selected parameter range (Wagener et al., 2002; Ward et al., 2017). The slope of the CDF will be higher in the
parameter interval where the model is more sensitive to that parameter. The measure of the local gradient of the
cumulative distribution will be represented by the height of the bar plot in each equally-sized bin across the
parameter space. Higher bars and steeper gradients of the CDF line indicate greater model performances in that
parameter range and, therefore, parameter sensitivity and identifiability (Figure A1g). On the contrary, equal eight

930   of the bars and similar gradients of the CDF line indicate that the parameter is unsensitive and non-identifiable
(Figure A1h).

The plots used to address the global sensitivity analysis indicate parameter identifiability and sensitivity on the
entire observed BTC, however they are unable to address if the *i-th* parameter describes the process it is meant to
represent or if the role of the *i-th* parameter on the model is constant in time (Wagener & Kollat, 2007). To address

935   identifiability and sensitivity of the *i-th* parameter on the different sections of the BTC we applied DYNIA
algorithm which steps are reported in Figure A2 (Wagener et al., 2002).
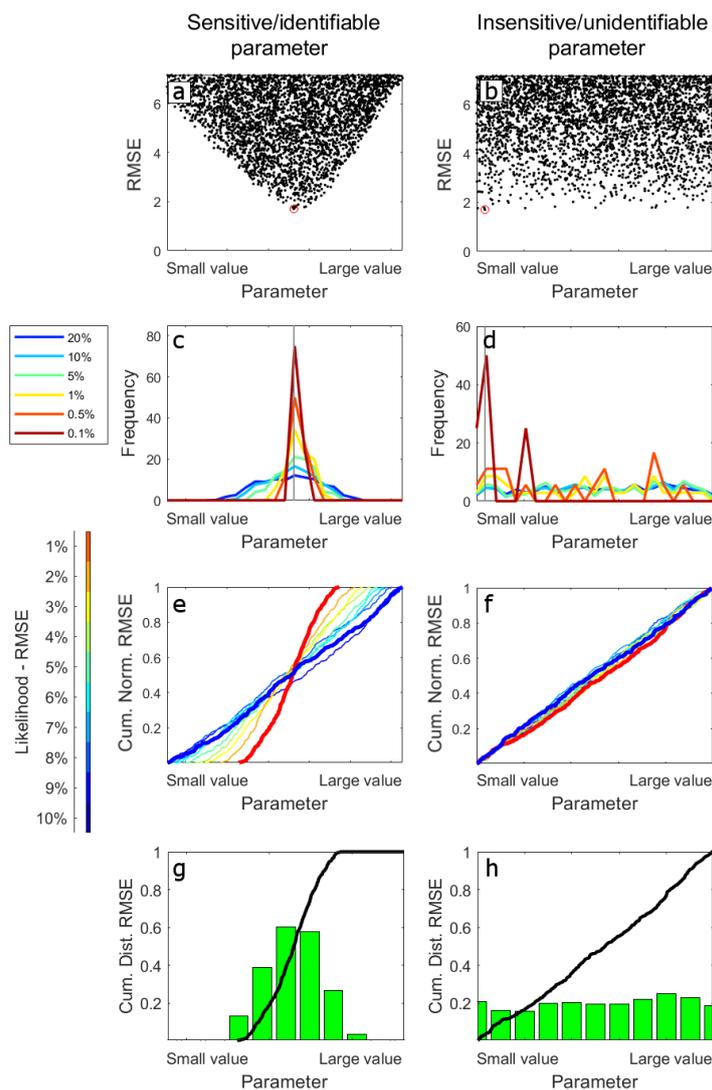
Figure A1: Examples of the four types of visualizations intended for parameter identifiability and sensitivity with the plots in the first column (a, c, e, and g) reporting an example of plots for sensitive and identifiable parameter and plots in the second column (b, d, f, and h) reporting an example of plots for insensitive and non-identifiable parameter. (a) and (b) parameter vs likelihood plots; (c) and (d) parameter distribution plots for the top 20, 10, 5, 1, and 0.1% of the results; (e) and (f) regional sensitivity analysis plots from the top 1% to the top 10% of the results; (g) and (h) identifiability plots for the top 1% of the model results.
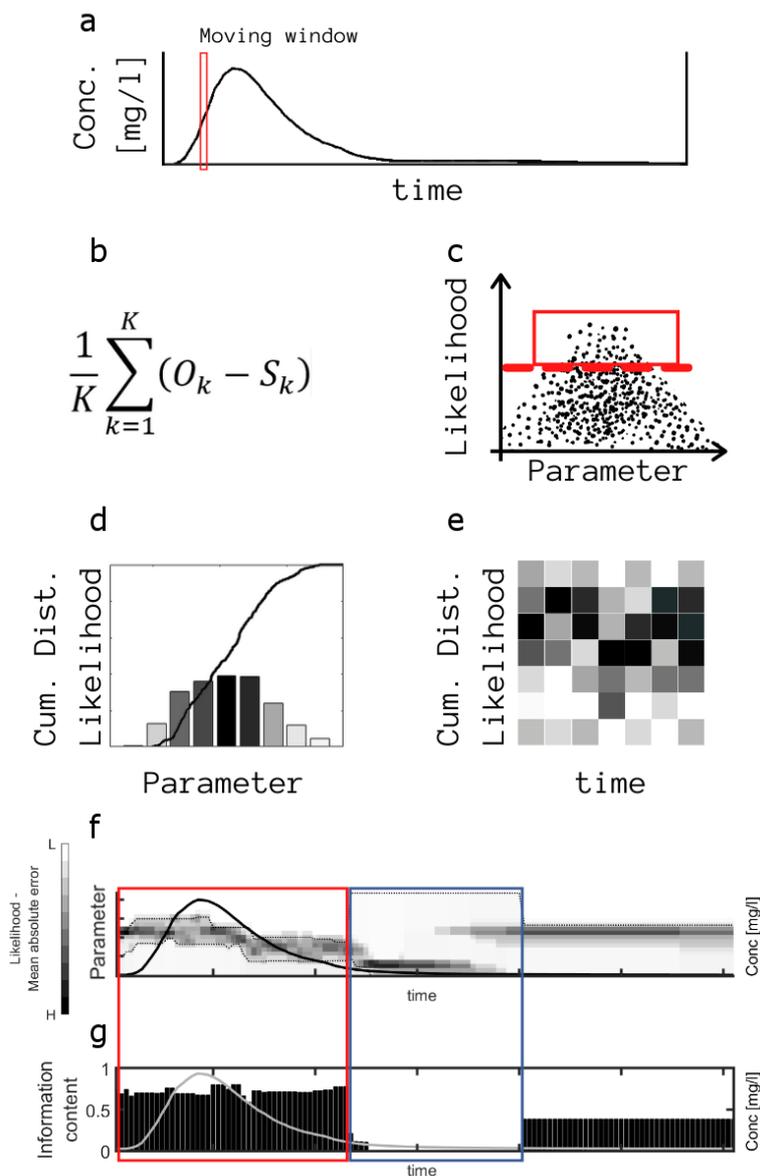
Figure A2. Dynamic identifiability analysis algorithm flowchart. (a) The BTC is subdivided in moving windows (size equal to three times the BTC timestep, Wagener et al., 2002); (b) In each moving window the likelihood (efficiency) of every TSM simulation is evaluated via mean absolute error (Wagener & Kollat, 2007); (c) an efficiency-threshold is chosen (e.g. top 10%); (d) for the chosen model results, the cumulative distribution function is built for each investigated parameter; (e) steps from (b) to (d) are repeated for each moving window and model likelihood for the investigated parameter is plotted over time (white: minimum likelihood; black: maximum likelihood). (f) cumulative

Hydrology and
Earth System
Sciences
Open Access
Discussions
EGU

distribution function of the parameter distribution is plot vs the observed BTC together with 90% confidence limits. Narrow limits indicate identifiable parameter while wide limits indicate unidentifiable parameter. (g) a second plot reports the metric of one minus the normalized distance between the 90% confidence limits. Small values of this metric indicate that the selected time window contain a narrow identifiability range for the investigated parameter and, 955   therefore, that it is informative on that part of the BTC (Wagener et al., 2002).