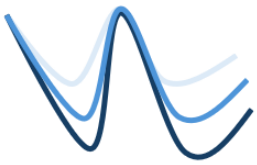
# Clab torino: A transdisciplinary environment to provide a challenge-based teaching model

WATSON

COST ACTION

# Storage and transit time estimation in hydrological systems with the «Water Isotope modelling for Transit time and Storage» (WITS) software



## Final report of working group 3 of the COST ACTION WATSON

COST
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the European Union

Julien Farlin, Alicia Correa, Sascha Müller, Stefanie Lutz, Michael Stockinger, Ladislav Holko, Stefano Ferris, Alessio Gentile, Martin Kralik

# Final report of working group 3 of COST Action CA19120: WATer isotopeS in the critical zONe from groundwater recharge to plant transpiration (WATSON)

https://watson-cost.eu/

Version 2.0

10.03.2025



This publication is based upon work from COST Action WATSON, CA19120, supported by COST (European Cooperation in Science and Technology).

# Table of Contents

# 1. Input-output modelling using environmental tracers

This document presents the theory and practice of input-output modelling approaches that can be used to gain information about storage volume, storage dynamics and tracer transit times in hydrological systems. It focuses specifically on the use of two environmental tracers that are part of the water molecule: the radioactive hydrogen isotope tritium ($^{3}H$), and its stable counterparts deuterium ($^{2}H$) and oxygen-18 ($^{18}O$). The aim of the document is to help the experimenter make informed decisions about modelling choices, depending on the available isotopic data and information about the system of interest, and the scientific or technical question to be answered. The limitations and advantages of the different choices are explained in a general way and illustrated by case studies. The report is accompanied by the R-based software WITS (Water Isotope modelling for Transit time and Storage) which implements most of the modelling options detailed in the text.

# 2. What is needed for lumped parameter modelling?

The minimum requirement to calibrate an input-output, or lumped parameter model is a measured input and a measured output tracer signal (see below for more details about tracer input length). The *input* is the concentration of tracer injected over time into the studied system. The *output* is the concentration of tracer measured at the outflow, or outlet, of the said system. The length of input and output necessary for calibration will depend on the tracer. Additional information concerning discharge at the system's outlet as well as rainfall and evapotranspiration can help refine the modelling (see paragraphs 3.8.2 and 3.8.3 respectively on variable flow modelling and on preparing the input function). Knowledge of the hydrogeological boundary conditions is useful to guide model choice and verify the plausibility of the obtained parameters.

# 3. Model concept

## 3.1 Basic considerations

All models presented in this report belong to the so-called **lumped-parameter model** family. The modelling problem is reduced to an input-output relationship described by a function called *transfer*, or *impulse response function*, whereby the hydrological system under study is considered to consist of one or more storage compartments supposed to be approximately homogeneous, each with uniform properties, boundary conditions and negligible spatial variability.

The *input* corresponds to the tracer injected into the hydrological system. This input may be tracer concentrations in precipitation or recharge water. The *output* is the tracer concentration at the outflow of the hydrological system under study (groundwater abstraction well, receiving stream, lysimeter, etc.). Any spatial characteristics of that pre-defined hydrological system (i.e., various flowpaths, hydrogeology, soil types, soil moisture distribution etc.) are summarized in the mathematical representation of the conversion from input to output by one or more *transfer functions*. The transfer functions represent how different flow paths, conceptualized as flow lines, transmit a tracer through the hydrological system (Figure 1). The functions summarize what is known or assumed with respect to storage characteristics and tracer release from the system under

investigation. Often, more than one function or combination of functions could a priori represent the feature of interest of the studied system. In such cases, each of them is a separate *model* of the system. Finally, a model is characterized by one or more parameters whose values must be estimated (or fitted or calibrated) using the tracer measurements.

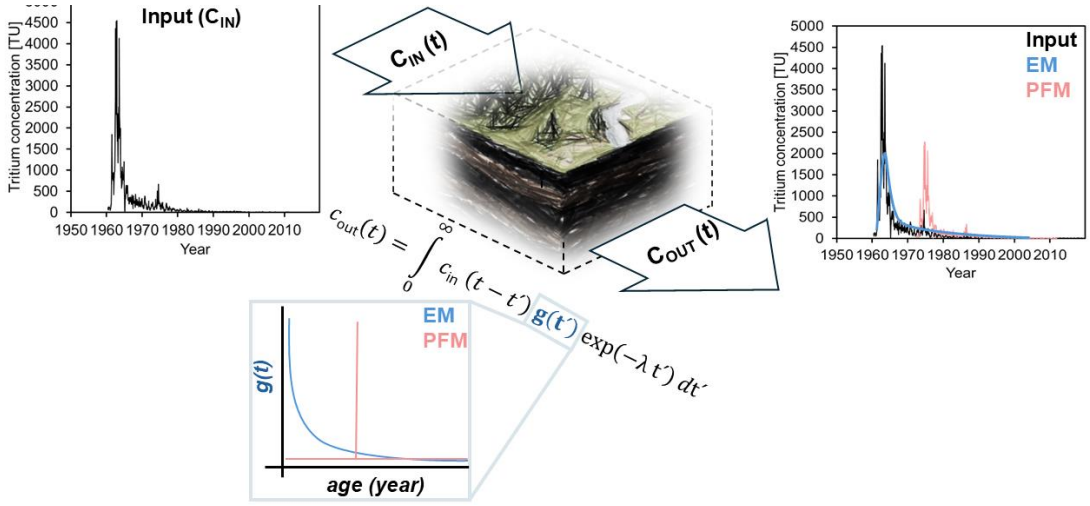



*Figure 1. Basic concept of tracer input-output modelling using lumped parameter models. The example shows a typical tritium input curve (left) and the output response corresponding to two different transfer functions, the exponential model (EM) and the piston-flow model (PFM). g(t) is the transfer function describing how each instantaneous tracer injection is released from the system over time (inspired by Suckow (2014)).*

When using isotopic tracers to calibrate lumped parameter models, a few basic assumptions must be considered:

a) Uniform spatial conditions for each considered compartment.
b) The tracer signal measured at the output integrates the total response all possible flow paths upstream of that point.
c) The isotopic composition of water leaving the system (output) is related to the isotopic composition of the input (usually precipitation) through the transit time distribution.
d) Flow conditions within the system can be either at steady state OR variable over time. Variable storage and variable fluxes must not be confused with the uniformity hypothesis (point a). Indeed, modelling variable flow can yield additional information about the system's storage dynamic.

The relationship between input, output and transfer function can be formalized by a convolution integral. WITS implements two different approaches: the classical lumped parameter models (Małoszewski & Zuber, 1982) and the storage selection functions (Botter, Bertuzzo, & Rinaldo, 2011).

The **classical lumped parameter models** have the following form (Małoszewski & Zuber, 1982):

$$c_{\mathrm{out}}(t) = \int_0^\infty c_{\mathrm{in}}\,(t-t')g(t')\exp(-\lambda\,t')\,dt' \qquad \textit{eq. 1}$$

here $c_{out}$ and $c_{in}$ are the time variable tracer concentrations in the output (for instance in a stream) and the input (for instance in precipitation), respectively, $t - t'$ is the time of tracer injection, $t'$ is the

transit time through the system, $g(t')$ is the transfer function and $\exp(-\lambda t')$ is a decay or degradation term modelling the dissipation of the tracer inside the system (by radioactive decay in the case of tritium). The transfer function or transit time distribution or impulse response function describes how an instantaneous tracer is released at the outlet over time. Theoretically, the integral from zero to infinity is necessary for all tracer injections to preserve mass balance (Małoszewski & Zuber, 2002). In practice, the convolution is carried out starting with the first measured tracer injection (for instance the 1950s to model the tritium bomb peak). Table 1 summarizes the most commonly used lumped parameter models transfer functions in isotope hydrology.

*Table 1. Commonly used lumped parameter model transfer functions in isotope hydrology*

| Model | Transfer function (TF) | Parameters | Mean transit time (MTT) |
|---|---|---|---|
| | *Convolution models* | | |
| Exponential (EM)[a] | $g(\tau) = (\frac{1}{\tau_m})^{\frac{-\tau}{\tau_m}}$ | $\tau_m$ = mean transit time | $\tau_m$ |
| Exponential-Piston (EPM)[b] | $g(\tau) = (f\tau_m)^{-1} \exp\ (-\frac{\tau}{f\tau_m}) + (\frac{1}{f}) - 1\ )$, g($\tau$)= 0 for $\tau < \tau_m(1-f)$ $f = \frac{1}{\eta}$ $\tau_{M-EM} = \frac{\tau_M}{\eta}$ $\tau_{M-PFM} = \frac{1-\eta}{\eta}\tau_M$ | $\tau_m$, $\eta = \frac{V}{V_{EM}}$ V= total Volume water in system $V_{EM}$ = Volume of water described by exponential transit time distribution | $\tau_m$ |
| Gamma (GM)[c] | $g(\tau) = (\frac{\tau^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)})^{\frac{-\tau}{\beta}}$ | α= shape parameter β = scale parameter | α*β |
| Dispersion Model (DM)[a] | $g(\tau) = \frac{1}{\tau\sqrt{\frac{4\pi P_D \tau}{\tau_m}}} exp\left[-\frac{(\frac{1-\tau}{\tau_m})^2}{\frac{4P_D\tau}{\tau_m}}\right]$ | $\tau_m$, $P_D$=dispersion parameter | $\tau_m$ |

[a] (Małoszewski & Zuber, 1982)
[b] Formulation from Stewart and Thomas (2008) from the original EPM by Małoszewski and Zuber (1982)
[c] (Amin & Campana, 1996)

The convolution describes how each tracer injection will be stored and released at later times, with the tracer concentration at the outlet being the sum of contributions from all previous injections depending on injection time.

The **storage selection functions** approach (Botter, Bertuzzo, & Rinaldo, 2011) is a reformulation of the classical lumped parameter models that replaces the tracer concentration vector $c_i(t)$ in the convolution by a resident tracer concentration vector which is ranked by residence time $C_s(t, T)$, with $T$ being the residence time through the compartment. Residence time is here defined as the time

spent by the tracer in the system since injection. Thus, $C_s(t,T)$ gives the concentration of a tracer (i.e. the tracer mass divided by the storage volume) at each time *t* with a residence time of *T* (*T* going from 0 to infinity) present in the storage compartment. This $C_s(t,T)$ vector is then convolved with a weighting function called the discharge age (or residence times) distribution that models how a tracer is released from the storage compartment as a function of its residence time following Benettin and Bertuzzo (2018)):

$$c_{\text{out}}(t) = \int_0^{\infty} C_S(T,t)\, p(T,t)\, \exp(-\lambda \mathrm{T})\; dT \qquad \textit{eq. 2}$$

with $p(\tau, t)$ being the discharge resident times distribution.

Just like the models of the classic lumped parameter model approach, the discharge age distribution describes how a tracer is released depending on its residence time and is composed of basic models that can be combined to represent different compartments and fluxes. In WITS, only one compartment with two fluxes (outflowing water flux and evapotranspiration flux) can be modelled using the storage selection functions.

Various functional shapes have been proposed empirically for the storage selection functions. They differ in the way in which the tracer is removed from each residence time "storage". The special case of a uniform selection across all residence times is equivalent to the exponential model of the classical lumped parameter approach (Benettin & Bertuzzo, 2018). Changing the shape of the selection function yields transit time distributions that are heavier or lighter tailed compared to the exponential model. Depending on the way short or long residence times are preferentially released, the selection function yields transit time distributions closer to other classical models such as the gamma model. WITS proposes three different models for the storage selection functions: fixed power law, time variant power law, and beta shape. The storage selection function for the power law model is defined as follows (Benettin & Bertuzzo, 2018):

$$\Omega(S_\tau, t) = \left[\frac{S_\tau(\tau, t)}{S_0 + V_0^{'}(t)}\right]^k \qquad \textit{eq. 3}$$

With $S_T(\tau, t)$ being the residence time ranked storage, i.e. the volume of tracer in storage with a residence time less than $\tau$ at time t, $S_0$ the initial storage and $V_0^{'}(t)$ the storage volume available to tracer at time t. $p(\tau, t)$ is then found from:

$$p(\tau, t) = \frac{\partial \Omega(S_\tau(\tau, t), t)}{\partial \tau} \qquad \textit{eq. 4}$$

Depending on the value of the k parameter, tracer mass in storage with longer or shorter residences times will be preferentially removed (Figure 2). For the power law model, k values smaller than one place more weight on the selection of shorter residence times, whereas k values above one give preference for the selection of longer residence times. A k value of one yields the exponential model (i.e. every residence time storage contributes equally to tracer flux).

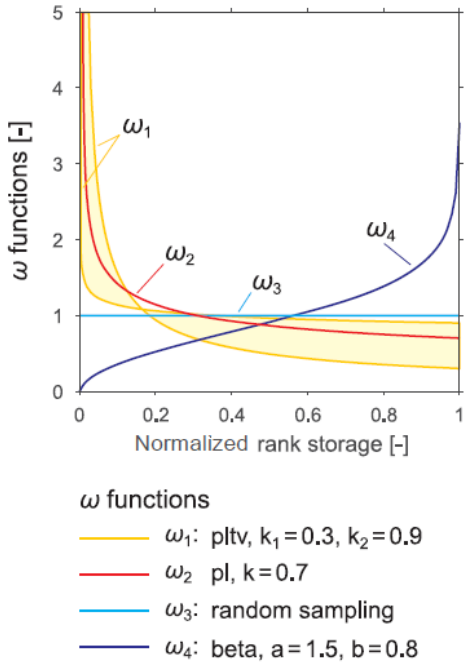



*Figure 2. Shapes of the three different storage selection functions implemented in WITS (Benettin & Bertuzzo, 2018). pl: power law, pltv: power law time variant, random sampling: equivalent to the exponential model (i.e., no preferential removal of transit time storage), beta: beta model. The storage selection functions ω are the probability density functions corresponding to the cumulative distribution functions Ω presented in eq. 3. and the age-ranked storage on the x-axis is presented as normalized rank storage ($P_S$=$S_T$/$S_{total}$). Note the effect of the k parameter on the preferential removal of the shorter transit times, and how the function converges to the constant value of random sampling as k approaches 1.*

## 3.2 Which useful information can be gained from lumped parameter models?

The calibrated lumped parameter model yields an **estimate of the storage volume** of water accessible to the tracer and its associated **mean transit time**. If the flow domain is split up in separate reservoirs the storage volume of each and its associated mean transit time is calculated. The *inverse problem* (or *model calibration* or *model fitting*) consists in finding the model parameters that best reproduces the observed output within a given error margin. A calibrated input-output model can be used to **predict** the output of another tracer if its input function is known (the *direct*, or *forward problem*). Typically, the direct problem is used in artificial tracing studies and water contamination problem to predict the evolution of an injected artificial tracer or contaminant concentration respectively (Małoszewski P. , 1993), allowing for possible differences in tracer behaviour (sorption, diffusion, and transformation processes. See for instance Farlin et al. (2022)). Concerning tracer differences, it is essential to keep in mind that water and tracer storage volumes (as well as mean transit time of water and of tracer) are *not necessarily equal*. The storage of mobile water is smaller than tracer storage if the ratio of microporosity to effective porosity is greater than one (see point 3.7 below for details). This fact is often forgotten in hydrological tracer studies, potentially yielding large overestimates of the amount of mobile water in storage (Zuber & Motyka, 1994).

## 3.3 Which systems and transport processes can be modelled?

Lumped parameter models can be used in any hydrological setting that can be conceptualized as a well-defined and bounded system. This includes surface water catchments, groundwatersheds, lysimeters, and column experiments. The system can be separated into different compartments that interact with each other and contribute with a tracer flux to the outlet. Within each compartment, advective and dispersive flow, sorption and desorption, tracer losses through radioactive decay or degradation as well as variable flow can be modelled. The compartments can be for instance:

- Two connected aquifers, one overlying the other, draining into a receiving stream (Haitjema, 1995).
- A double porous medium (aquifer, lysimeter, etc..) with mobile and immobile (stagnant) water zones (Stumpp, Stichler, & Małoszewski, 2009).
- A slowflow and a quickflow component, one sustaining baseflow, the other becoming activated when interflow and surface runoff occur, or in the case of karstic systems, when conduit flow is occurring (Małoszewski, Stichler, Zuber, & Rank, 2002)..

Water and tracer fluxes as well as storage of most hydrological systems are not at steady state, but variable in time. Both steady state and variable flow formulations are possible using lumped parameter models. It must be noted however that **modelling variable flow is not always desirable or indeed necessary**. In some cases, the steady state approximation will be sufficient to extract the isotopic information available, thus simplifying the fitting procedure and reducing model complexity.

## 3.4 Model calibration

Lumped parameter models predict the output ($C_{out}$) of a tracer as a function of its input ($C_{in}$) and of the transfer function $g(t')$ or the discharge resident times distribution $p(\tau, t)$. The parameters of the functions must be fitted using observations and are thus referred to as fitting or free parameters. Parameter fitting is performed by varying the parameter or parameters until a good match is obtained between observed and predicted output. In practice, the best fit or best fits are found by minimizing the residuals through the optimization of an objective function. Currently, the pre-defined objective functions in WITS are Nash-Sutcliff Efficiencies (NSE) and Mean Prediction Error (MPE).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(c_i - cm_i)^2}{\sum_{i=1}^{n}(c_i - \bar{c})^2} \qquad \textit{eq. 5}$$

$$MPE = \frac{\sqrt{\sum_{i=1}^{n}(c_i - cm_i)^2}}{n} \qquad \textit{eq. 6}$$

where n is the number of observations, c and cm the predicted and measured concentration respectively, and $\bar{c}$ the average of all measured values.

A good fit value does not necessarily mean that the chosen model is adequate. It is the responsibility of the modeler to select the models that best represents what is known of the modelled system and to be critical of model results. In doubtful cases, comparing the results of different possible models is advised, especially for increasing degrees of complexity (see the modelling guide for details). Additional compartments increase model flexibility, but at the cost of additional parameters which must be calibrated with the available data. Usually, each new compartment adds at least two

parameters to the model, one (or more) for the associated transfer function, and the another to set the fraction of flow in and out of the compartment. It is the modeler's responsibility to strike a balance between complexity and available constraint so that the information contained in the tracer and hydrological data is extracted as completely as possible without introducing undue non-uniqueness. A detailed overview of the models and fitting parameters included in WITS are shown in Table 1.

## 3.5 Minimum length of input time series

As any output measurement is usually the result of a superposition of separate tracer injections having occurred in the past, input measurements predating the first output measurement should be available. Hydrological systems store the history of tracer injection for a certain length of time which depends on storage and how each successive injection is released over time. Figure 3 shows the normalized tracer recovery ($C/C_0$) as a function of the normalized transit time ($t/\tau_m$) of an exponential model. 66 % of the tracer is recovered after one mean transit time, 95% after three mean transit times. This means that after three mean transit times 95 % of the total tracer volume in storage will have been replaced. Hence, this is the length of time necessary to characterize 95 % of the tracer storage in the system at any time. Thus, one can suggest as a rule of thumb that sampling of the input should begin at least three mean transit times before sampling at the outlet starts. As the mean transit time of tracer is usually unknown at the beginning of a study, the starting date of sampling at the input can only be estimated roughly and then verified a posteriori.

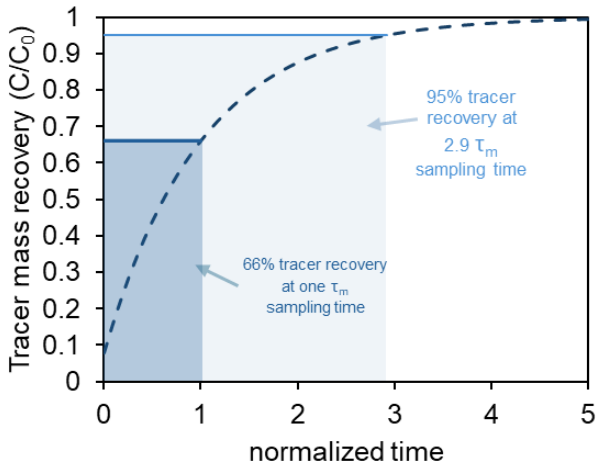



*Figure 3. Cumulative tracer mass recovery ($C/C_0$) as a function of normalized transit time ($t/\tau_m$) following an exponential model with a mean transit time of one year. In order to recover more than 95% of the tracer mass at the outlet of a hydrological system at least 3 times the mean transit time need to be sampled.*

If seasonal variations of the tracer input and output are used for modelling the length of observation is at least one year (and preferably longer as explained above). In the case of tritium, the local magnitude of the bomb peak must be known, as well as the subsequent decrease in atmospheric tritium activity. In the special case where the tritium peak has been completely flushed out of the hydrological system under consideration, the annual mean input, possibly weighted to take potential evapotranspiration losses back to the atmosphere during the summer months into account, is sufficient.

## 3.6 General strength and limitations of lumped parameter models

### 3.6.1 Limitations

Lumped parameter models are not spatially distributed. It is assumed that the whole flow domain under consideration is homogeneous in all its hydraulic characteristics and boundary conditions (same storage, recharge and tracer injection everywhere). Thus, spatial differences are assumed to be negligible at the scale of the whole system. If the experimenter has good reasons to suspect this not to be the case, lumped parameter modelling either (i) cannot help, (ii) provides a simple reference to compare more to complex approaches (i.e. physically based numerical models) against, (iii) should be applied to parts of the system where the homogeneity assumption can be upheld. In this latter case, the flow domain is split into different compartments, each homogeneous, in exchange with one another and each contributing to the outflow in various ways.

The modelled output and the validity of the calibration parameters is only as good as the input and the choice of the transfer function. Making sure the input has been prepared with the utmost care, considering all available information, is crucial for reliable results. Any information about tracer evolution over time observed in the output, but missing from the prepared input will not be modelled correctly and can lead to introducing modelling assumptions that are simple artefacts due to a faulty or doubtful input.

### 3.6.2 Strengths

Lumped parameter models are parsimonious, requiring as few as one fitting parameter, which is particularly relevant when the only information available is a tracer input and output. Setup is simple and if the tracer data is of good quality, calibration quick and unambiguous. Lumped parameter models are particularly easy to implement and can perform equally well or better than more complex approaches (Stumpp, Stichler, & Małoszewski, 2009) both in terms of reproducing observations, and in gaining insight into the hydrological system. They can furthermore be adapted to the hydrological situation in different ways, in particular by considering different compartments within the flow domain which transmit a tracer either in parallel or in series (see 3.8.1 below as well as Figure 6 and Figure 8).

## 3.7 Storage volume as the main parameter

No matter which approach is used (classical lumped parameter model or storage selection function), lumped parameter modelling yields primarily an estimate of the storage volume accessible to the tracer, either directly (in the case of the storage selection functions where it is an explicit fitting parameter) or indirectly through the estimated mean transit time. The relationship between the total storage volume $V_0$ and the mean transit time $\tau_M$ is (Małoszewski & Zuber, 1982):

$$\tau_M = \frac{V_0}{Q} \qquad \textit{eq. 7}$$

where Q is the mean flux entering or leaving the system. This equation relates three basic quantities: $\tau_M$ is an abstract number used for modelling purposes, the total storage volume $V_0$ is a physical quantity that cannot be measured directly and the flux Q is both physical and measurable. The relationship given in equation 7 is qualitatively always valid, even for variable flow (in which case, $\tau_M$

will vary with $Q(t)$ and $V_0(t)$), but the estimated value of the parameter $\tau_M$, and thus the calculated volume $V_0$ will depend on model choice.

The volume $V_0$ estimated from environmental tracers can be different from two volumes typically sought in hydrology: the total water storage and the dynamic storage ($V_D$). The relationship between the different volumes for a model (ground)watershed is illustrated on Figure 4. These differences are essential and *must be kept in mind* when modelling a tracer response and analysing the modelling results.

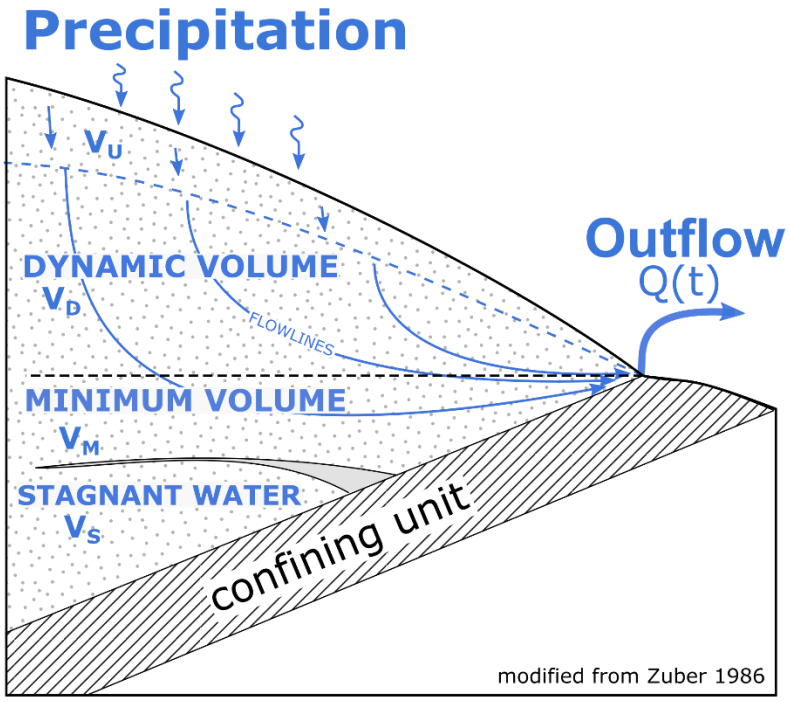



*Figure 4. A schematic view of the different volumes relevant for input-output modelling using environmental tracers (modified from Zuber, 1986). See text below for an explanation.*

The different volumes relevant for input-output modelling and their relationship are (Zuber, 1986):

- The dynamic volume $V_d$. It is the amount of water stored above the outlet of the system and which influences the outflow rate. The dynamic volume can be estimated using recession analysis. When this storage is exhausted, the water flow through and out of the system (and hence the tracer flux) stops. Water and tracer can still remain stored in any reservoir situated below the datum of the outlet until flow resumes.
- The minimum volume $V_m$. It is the volume still stored in the catchment when the outflow rate goes to zero.
- The stagnant volume $V_s$. It is the amount of stagnant water not accessible to tracer because it is isolated from all flow lines (i.e. there is no diffusive exchange with flowing water). Hence, this volume can neither be estimated from the outflow rate (as the dynamic volume) nor by tracer analysis (as the minimum volume). The stagnant volume is not to be confused with stagnant water zones situated within the flow domain where an exchange with flowing water by molecular diffusion is possible (see next point).

- The volume accessible to tracer $V_0$. It is the sum of the dynamic and minimum volumes ($V_0=V_d + V_m$). If stagnant micropores are present in the flow domain (for instance in the case of a fractured rock aquifer with a porous matrix), $V_0$ will represent the total volume of both hydraulically effective porosity $n_f$ and microporosity $n_p$ because for long term injections typical of environmental tracers (in the order of months or years) the tracers behaves as if transport were taking place through the entire pore volume (Zuber & Motyka, 1994).
- The volume of flowing water $V_w$ stored within the flow domain encompassed by the dynamic volume and the minimum volume will be the volume accessible to tracer $V_0$ corrected by the ratio of microporosity to effective porosity (Zuber, 1986). If this ratio is not known from independent measurements, the volume of water in storage cannot be calculated from tracer data alone, but tracer analysis will yield an upper bound equal to $V_0$. It must be understood that the difference between the volume of water stored $V_w$ and the volume accessible to tracer $V_0$ can be very large if the ratio of microporosity to effective porosity is greater than one.

$$V_0 = V_w * \left(1 + \frac{n_p}{n_f}\right) \qquad \textit{eq. 8}$$

  where $n_p$ is the porosity of the matrix and $n_f$ the effective porosity.
- The volume of water stored in the unsaturated zone $V_u$. This volume may not be negligible compared to the total volume accessible to tracer. In that case, a way to account explicitly for the corresponding the lag time between tracer injection and tracer arrival in the saturated zone could be added to the model.

If the experimenter has divided the flow domain into different reservoirs, the volume of each can be found from its corresponding mean transit time. For instance, for an exponential piston-flow model, calibration yields two mean transit times respectively for the exponential and the piston-flow component: $\tau_{EM}$ and $\tau_{PFM}$. Accordingly, the volume of each compartment is then given by:

$$V_{EM} = Q * \tau_{EM} \qquad \textit{eq. 9}$$

$$V_{PF} = Q * \tau_{PF} \qquad \textit{eq. 10}$$

Also, if the recharge area A and the mean porosity n are known or can be estimated, the equivalent water column height h is then:

$$h = V_0/A \qquad \textit{eq. 11}$$

Note

In WITS the estimation of storage volume is not defined as a model output (except in the case of the storage selection functions) and the user must calculate this variable with the above equations eq. 9 to eq. 11 if desired.

## 3.8 Modelling choices

### 3.8.1 Flow domain

Lumped parameter models are flexible in the hydrological elements that can be modelled. This flexibility is achieved by selecting or combining different basic models, each representing a part of the flow domain and then calibrate all the free parameters using the measured tracer input and output as well as ancillary information on the system. All basic models rest on the assumption that the tracer is transported along flow lines from the point of injection to the outlet (or from the entrance of a compartment to its exit if the model is composed of compartments in series). Flow lines are not simulated individually but lumped together in a model that is specific to the distribution of the transit time of all flow lines for that part of the system.

The basic models available in WITS are outlined below and an overview of their applicability is given in Figure 5. All these models can be implemented directly using the classical lumped parameter model approach. For the storage selection functions, some models can be implemented implicitly (the exponential and the gamma model) while others remain empirical without direct mechanistic relationship to hydraulic principles.

The models are:

- **Piston-flow model (PFM)**. If all flow lines have the same transit time and no exchange takes place between them (i.e. molecular diffusion and hydrodynamic dispersion are assumed negligible at the time and length scale of the problem), the piston flow model (PFM) approximately applies. This could be for instance the case for an unsaturated zone with approximately constant thickness or a quickflow contribution to streamflow where tracer transport can be considered to be nearly purely advective. In shallow unconfined aquifers with monitoring wells containing short screens PFM is applicable as well.
- **Exponential model (EM)**. When flow lines have different transit times due to different flow lengths and flow velocities and no exchange takes place between them (i.e. molecular diffusion and hydrodynamic dispersion are assumed negligible at the time and length scale of the problem), the exponential model (EM) applies. Hence, mixing occurs within the well screen or in the near vicinity, but not within the system. Aquifers discharging to streams or springs are settings where the EM can typically be applied.
- **Exponential Piston Flow model (EPM).** The EPM can be used in systems with two storage compartments *in series*: one storage is described by an EM, which is followed or preceded by a storage described by PFM. Typical examples are (i) a thick unsaturated zone (PFM) followed by an alluvial aquifer (EM) or (ii) a confined aquifer with a recharge area situated far enough from the discharge point so that the minimum flow time from recharge area to discharge is significantly greater than zero.

- **Gamma and dispersion models**. If flow lines have different transit times due to different flow lengths and flow velocities and are in exchange with each other due to hydrodynamic dispersion (i.e. a combination of mechanical dispersion and molecular diffusion), either the gamma model (GM) or the dispersion model (DM) can be used. Both the GM and the DM can yield a wide range of shapes of the transit time distribution. Both models may be used in a wide range of hydrogeological conditions where mixing due to hydrodynamic dispersion is supposed to play a role in tracer transport and storage.
- **Other models**. In case of the storage selection functions, the transit time distribution depends on the selection of SAS functions representing the different compartments (for instance the unsaturated zone and the groundwater storage) of the model. Thus, the transit time distributions obtained may not resemble any of the previously described models. These models offer more flexibility, but remain empirical, without physical basis guiding model choice.

The main defining parameter of all basic models is the mean transit time or alternatively, the storage volume (see Table 1, eq. 4, eq. 7). The additional parameter of the gamma model and the dispersion model describes the relative importance of dispersive fluxes compared to advective fluxes[1]. In the DM hydrodynamic dispersion is modelled using the dispersion parameter ($P_D$), which is a dimensionless constant giving the ratio of diffusive to advective flow rates (Table 1). The GM considers a degree of mixing going from 0 corresponding to no mixing between flow lines (i.e. the exponential model) to 1 corresponding to complete mixing *within the system* due to hydrodynamic dispersion (Amin & Campana, 1996). For the GM, two additional parameters of alpha (shape) and beta (scaling) are introduced, the product of which being the mean transit time (Table 1). All the aforementioned models have been derived from physical considerations (Lenda & Zuber, 1970) (Haitjema, 1995) (Amin & Campana, 1996) (Kirchner, Feng, & Neal, 2001). Other empirical forms have been used as transfer function with the classical lumped parameter model approach, but as they lack a physical basis that could guide their selection for modelling a particular hydrological compartment, they will not be presented here and are not implemented in WITS.

[1] The dispersion parameter is the inverse of the Peclet number, which is the ratio of the rate of diffusion to the rate of advection of a tracer through a medium.

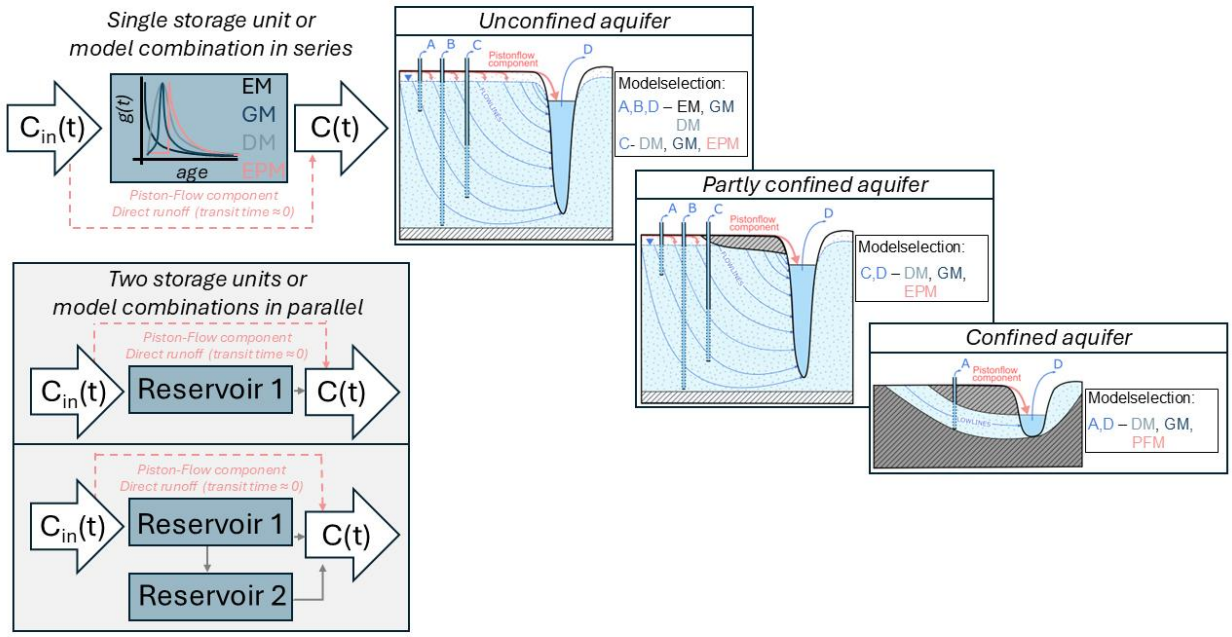



*Figure 5. An overview of lumped parameter model setups (left part of the figure and grey insets) as well as typical hydrological settings where the different basic models can be applied (after Małoszewski & Zuber, 1982). PFM: piston flow model, EM: exponential model, EPM: exponential piston flow model, DM: dispersion model. The gamma model (GM) is an alternative to the dispersion model. Note that the position of the screen in observation wells and the presence of impermeable layers close to the receiving surface water body is important to differentiate between an exponential model and an exponential piston flow model. The piston-flow component shown in red on the cross sections corresponds to an additional flux in parallel to the groundwater flux.*

These basic models can be combined to simulate more complex systems. The combination can be best represented as a series of compartments exchanging fluxes with the inflow and outflow, and possibly between themselves. Compartments can be arranged **in series** (i.e. one after the other in the direction of flow) or **in parallel** (i.e. the total flux is a combination of partial fluxes through each compartment). The most common combination in series is the so-called exponential piston flow model which consists of a piston-flow component and an exponential component (EPM, see examples below and Table 1), but any combination is feasible as long as the experimenter can justify the choices made from the available hydro(geo)logical information about the modelled system. Each compartment is assumed to be made up of flow lines arranged in a way that depends on the model chosen. The flow lines are either isolated from their neighbours (exponential and piston-flow models) or can exchange tracer via molecular diffusion or hydrodynamic dispersion (gamma and dispersion models). A compartment is characterized by a storage volume, a tracer release behaviour, and a fraction of total flux passing through it. Figure 6, Figure 7 and Figure 8 present examples of conceptualizations of hydrological systems making use of combined models, while the case studies will present some real-world applications.

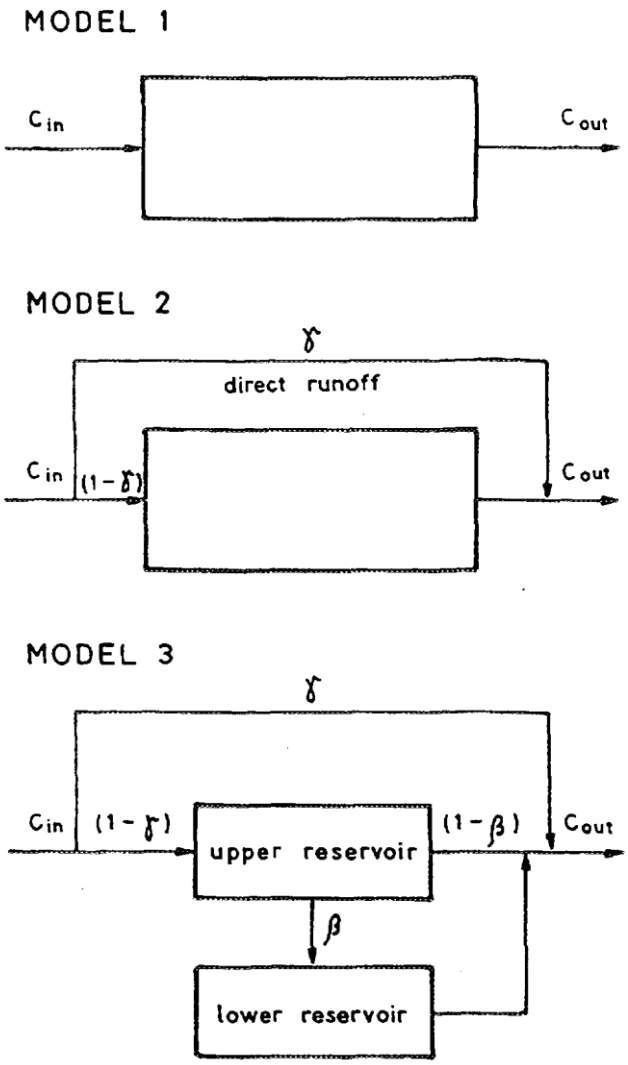



*Figure 6. Three different conceptualizations of a surface catchment (Małoszewski, Rauert, Stichler, & Herrmann, 1983) with increasing degree of complexity. Model 1 considers a single flow compartment. Model 2 distinguishes between two flow compartments: direct runoff and groundwater. Model 3 further divides the groundwater domain into two separate aquifers, with the upper aquifer recharging the lower aquifer, and both contributing separately discharge to the stream. All groundwater compartments were modelled using two alternative model: the exponential model and the dispersion model. Note that the flow into each reservoir for models 2 and 3 is quantified by an additional parameter (γ or β). In the original study, the γ and β parameters were estimated from hydrological data, independently of the isotope data. The isotope data was solely used to calibrate the parameters of each compartment and NOT the exchange between those.*

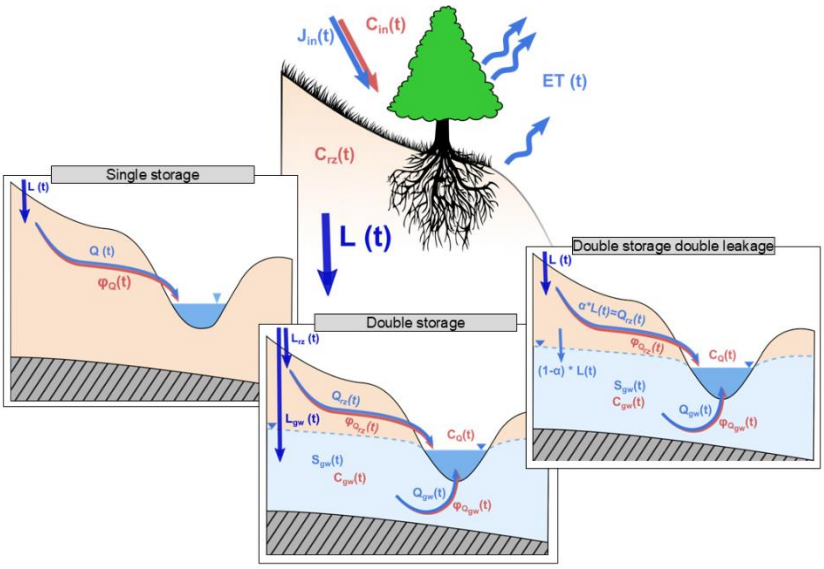



Figure 7. Different representations of a catchment considering evapotranspiration fluxes and with increasing number of storages and fitting parameters (after Benettin et al. (2013) ). Water fluxes are in blue, mass fluxes in red. Note that compared to the representations shown on Figure 6, the soil reservoir is considered as a separate compartment in order to model evapotranspiration dynamically using storage selection functions.

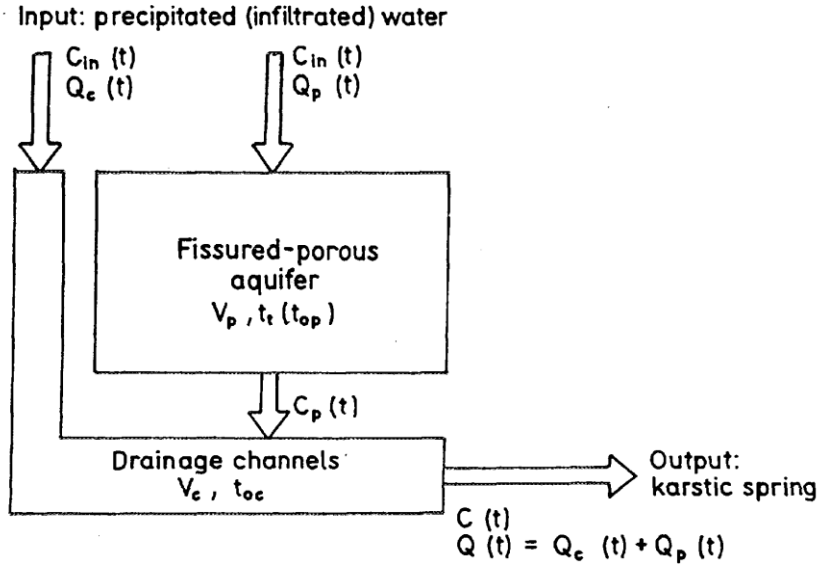



*Figure 8. Example of a concept model for a karstic aquifer (Małoszewski, Stichler, Zuber, & Rank, 2002). Tracer input and recharge water are either transported through drainage channels, or through the fissured-porous aquifer at flow rates $Q_p$ and $Q_c$ respectively. The fissured porous-aquifer discharges into the drainage channels at the flow rate $Q_p$. When $Q_c \neq 0$, direct flow from precipitated water also occurs. Flow rates Q and concentrations C were known, the parameters $V_p$, $V_c$ and $t_t$ and $t_{oc}$ were estimated from isotope data.*

Note

Input-output models can be setup in various ways. Models can be combined in series or in parallel. The difference between the two approaches is important to understand. The compartments of models combined in series share the same tracer flux, whereas the flux is split

between compartments for models combined in parallel. The exponential piston-flow model is a typical example of a model combined in series (Figure 5). A combination of a piston-flow compartment and an exponential (or dispersion) model in parallel (red dotted lines in Figure 5 and example shown in Figure 8) is not equivalent to an exponential piston-flow model because in the first case both compartments do not share the same flux.

The combination of models in series is implemented in WITS by allowing the user to add a piston-flow component to any other available model (in series approach). Calibrating models made up of compartments in parallel requires modelling each flux separately and can become more complex if the contribution from each compartment varies over time (see Małoszewski et al. (1983) and Małoszewski et al. (2002) for example applications). The only model of the kind programmed by default in WITS is a combination of two exponential models called the double exponential model (Morgenstern, et al., 2015) with a constant contribution from each compartment. This model has three fitting parameters: the relative proportion of total flux flowing through the two compartments and two mean transit times.

Because of the lumped parameter models' modular nature, the experimenter could be tempted to add more compartments, and hence more fitting parameters, than can be constrained unambiguously with the available information. While this problem is not inherent to lumped parameter models, but to any modelling attempt, it should be stressed that overfitting can and must be avoided when using lumped parameter models.

#### 3.8.2 Variable flow

Tracers are transported through hydrological systems following hydraulic head gradients, or for systems with stagnant water zones, concentration gradients. In most hydrological systems, storage will change over time, which will increase, decrease or even reverse gradients between inlet and outlet as well as between compartments. This will in turn change the water and tracer fluxes. Changing tracer flux caused by changes in storage can be modelled using lumped parameter models. In that case, the concentration term in equation eq. 1 is replaced by a mass flux term (tracer mass multiplied by volumetric flow rate). This flow term must be additionally estimated from available data.

If however the time scale of these variations as well as the magnitude of storage change is small compared to the time scale of tracer transport and total storage respectively, a steady state approximation will yield nearly the same results for lower computational costs, while also being less demanding in terms of data. Figure 9 presents an comparison of both steady state and variable flow fitting where the steady state formulation proved sufficient.

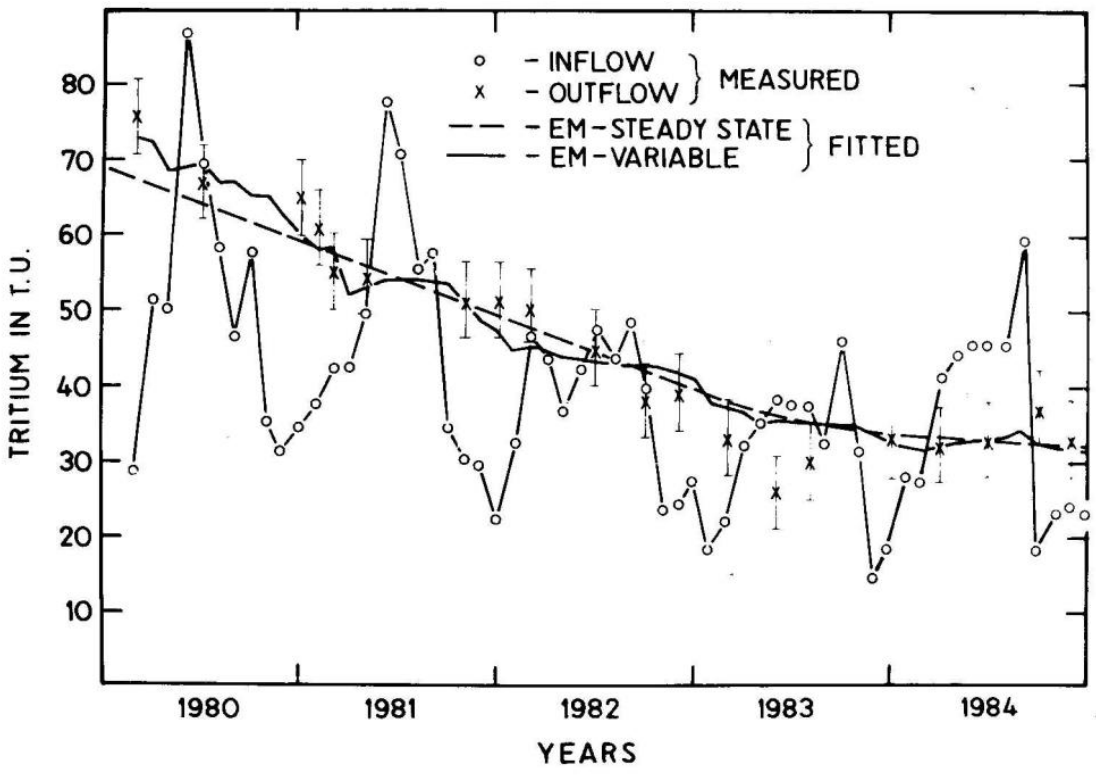



*Figure 9. Example of fitting comparison between steady state and variable flow fitting of tracer data using an exponential model (Zuber, Małoszewski, Stichler, & Herrmann, 1986). The variable flow formulation allows a better match of short term variations in tritium activity, but both approaches yielded very similar total storage volumes (0.85-1.04 x10$^6$ m$^3$ for the steady state model, 0.9 x10$^6$ m$^3$ for the variable flow model).*

The variable flow formulation for the classical lumped parameter models implemented in WITS is described in detail in Zuber (1986). Compared to the steady state formulation, the convolution integral is modified to account for variable flow $Q(t)$.

$$c(t)Q(t) = \int_{-\infty}^{t} c_{in}(\tau)\ Q_{in}(\tau)\ \ g_F(t,\tau)\exp[-\lambda(t-\tau)]\ d\tau \qquad \textit{eq. 12}$$

After rearrangement and some manipulations, one obtains the following convolution equation:

$$c(t) = \frac{1}{V_0^{'}(t)}\int_{-\infty}^{t} c_{in}(\tau)\ Q_{in}(\tau)\ \ C_I(z)\exp[-\lambda(t-\tau)]\ d\tau \qquad \textit{eq. 13}$$

Hence, the concentration at the outflow at time t depends on the storage volume $V_0^{'}(t)$, the time variable inflow concentration $C_{in}(\tau)$, variable inflow volume flux $Q_{in}(\tau)$, and the normalized transit time distribution $C_I(z)$, with z being the fraction of water having passed through the variable volume of the system in the time interval τ-t. For the exponential model, $C_I(z)$ is:

$$C_I(z) = \exp(-z) \qquad \textit{eq. 14}$$

with

$$z = \int_{\tau}^{t} \frac{Q(t')}{V_0^{'}(t)}\,dt' \qquad \textit{eq. 15}$$

$Q_{in}(\tau)$ is usually unknown, but can be estimated from the well-known recession equation

$$Q_{in}(t) = t_d \frac{dQ(t)}{dt} + Q(t) \qquad \textit{eq. 16}$$

where $t_d$ is the dynamic turnover time of the variable part of the system. The total volume of the system is then:

$$V_0'(t) = t_d Q(t) + V_M \qquad \textit{eq. 17}$$

In WITS, the exponential model and the dispersion model with or without a piston-flow component can be implemented with a variable flow formulation. The fitting parameters are those indicated in Table 1 (one for the exponential model, two for the dispersion model) together with the additional $t_d$ parameter.

Variable flow in the storage selection function approach is modelled in the same way but allowing for a second flux exiting each compartment (the evapotranspiration flux). eq. 18 presents the discharge age distribution equivalent to an exponential model but including two fluxes (rainfall and evapotranspiration) instead of only one. The discharge resident times distribution is further scaled by the corresponding flux in the case of a variable flow (eq.19) in the exact same way as the classical lumped parameter models (compare eq.14 and eq.18).

$$p(\tau, t) = \frac{J(\tau)}{V_0'(t)} \exp(-z) \qquad \textit{eq. 18}$$

where J is the rainfall flux.

z is similarly to eq.15 defining the fraction of water having passed through the variable volume of the system in the time interval τ-t but taking the sum of both fluxes leaving the compartment, namely Q (streamflow or discharge) and E (evapotranspiration).

$$z = \int_{\tau}^{t} \frac{Q(t') + E(t')}{V_0'(t)} dt' \qquad \textit{eq. 19}$$

with E being the evapotranspiration flux.

### 3.8.3 Evapotranspiration, snow storage and preparation of the input function

The input function is the concentration of tracer injected in the system over time. In most natural systems, injection takes place with precipitation proportionally to the injection flux (Małoszewski & Zuber, 1982). Nonetheless, the input function must usually be prepared from the raw measurements for modelling (Viville, Ladouche, & Bariac, 2006) (Stumpp, Stichler, & Małoszewski, 2009). If tracer is removed by evapotranspiration or stored temporarily in the snow cover, the tracer concentration in precipitation does not correspond to the tracer flux entering the system and should be modified accordingly. A simple example of snow storage modelling will be presented in the case studies section (paragraph 4.3). For modelling the effect of evapotranspiration on tracer storage, two possibilities are offered by the modelling software WITS:

- Calculation of the summer to winter recharge ratio. Long term mean evapotranspiration losses during the summer months can be estimated from the shift in the isotopic mean values of annual rainfall and outlet flux for either oxygen-18 or deuterium. From this shift, a summer to winter ratio is obtained and used to weight all summer months to reduce their

importance relatively to the winter months (Grabczak, Małoszewski, Rozanski, & Zuber, 1984). The weighting factor α applied to the summer months is found as follows:

$$\alpha = \left[\left(\overline{\delta P_w} - \delta G\right)\sum_i (P_i)_w\right] / \left[\left(\delta G - \overline{\delta P_s}\right)\sum_i (P_i)_s\right] \qquad \textit{eq. 20}$$

Where $\overline{\delta P}$ is the long-term mean weighted isotopic content in precipitation, δG is the mean isotopic content in groundwater, P is the precipitation amount, and s and w stand for summer and winter months respectively. This method is static and does not take interannual variations in summer evapotranspiration into account.

- Dynamic modelling of evapotranspiration losses. If evapotranspiration can be estimated from meteorological data, tracer losses to the atmosphere can be modelled as an additional flux leaving the system. This dynamic approach taking variable evapotranspiration fluxes into account is specific to the storage selection functions and cannot be used in the classical lumped parameter model approach.

The dynamic, time variable modelling of the evapotranspiration tracer flux is an advantage of the storage selection functions over the classical formulation of the lumped parameter models. If evapotranspiration can be neglected or simply modelled by a constant summer to winter ratio, classical lumped parameter models and storage selection functions are equivalent both for steady state and variable flow formulations, only differing in the underlying transfer function.

# 4. Case studies

## 4. 1 Long tritium time series in an alpine spring: the Weiβbach spring, Austria

Julien Farlin[1] and Martin Kralik[2]

[1] Luxembourg Water Agency, Luxembourg

[2] University of Vienna, Austria

The different springs discharging into the Weißbach stream in the Gamperdona valley close to the village of Nenzing in the Austrian Alps (WGS84 9.69°E, 47.15N, 978m, Figure 10) have been mapped and investigated for the first time by KRASSER (1952). More recent investigations were performed by SUTTERLÜTTI (1990), LEDITZKY et al. (1990) and KRALIK (2015). $\delta^{18}O$,$^{2}H$ and tritium have been sampled monthly by the department of the provincial government of Vorarlberg (Austria) in the northernmost of these springs since 1993. Due to the difficult accessibility of this Alpine spring most samples were taken by a tube diversion over the Meng stream to its northern bank. The discharge of all the Weißbach springs is estimated to be more than 400 litres per second, with a recharge area of 13.7 $km^2$ and a main altitude of 1730m. The spring sampled monthly discharged six litres per second only in 1987 (SUTTERLÜTTI 1990, LEDITZKY et al. 1990). The spring has its source at the base of a mighty Upper Triassic Hauptdolomite formation. KRALIK (2015) noticed the quick changes in tritium content between 7.8 and 12 TU in the year 2014 and concluded that the contribution of bomb-tritium in spring waters varies seasonally or in response to rainfall.

The tritium input was prepared by combining monthly tritium measurements from the Bregenz station (ANIP 2014) for the period from 1973 to the present and the Vienna station between 1961 and 1972. The slope of the regression line of Vienna monthly values against Bregenz being nearly equal to 1, the Vienna measurements were used without further correction. Results for various models are summarized in Table 2. Generally, the mean transit times estimated are greater than twenty-one years. For comparison, a tritium/helium-3 ($^3$H/3He) sampling of the Weißbach spring in August 2014 yielded a mean piston-flow transit time of 19 years (KRALIK, 2015). An exponential model was fitted first to serve as reference, being the simplest possible model corresponding to the hydrogeological setting (flow lines with different transit times and purely advective transport). Adding a piston-flow component simulating the existence of a minimal flow distance between recharge and discharge areas improved the fit significantly particularly for the measurements before 2005 (Figure 11).

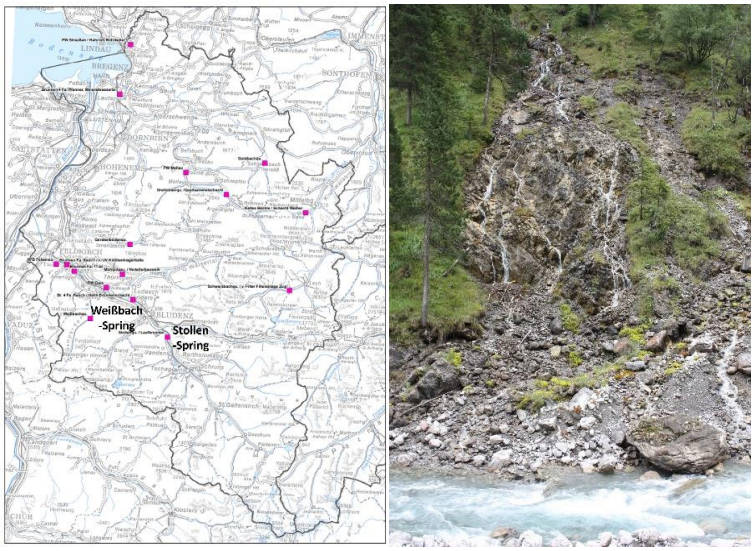



*Figure 10. Geographical situation and photograph of the alpine Weißbach springs in the provincial state of Vorarlberg on the western border of Austria (KRALIK 2015)*

The summer to winter ratio (eq. 20) could not be estimated using either $\delta^{18}O$ of $^2H$ measurements as the mean groundwater value of the Weißbach spring is the result of two effects combined: a potentially reduced tracer injection in the summer months and the altitude effect. The alpha parameter thus becomes an additional fitting parameter. As shown on Figure 12 using the exponential piston-flow model, the summer to winter ratio has to be greater than 0.5, and is most probably closer to 1 (i.e. no or negligible evapotranspiration losses during the summer months). This is a reasonable result for an alpine catchment where summer precipitation could generate immediate runoff and recharge without significant evapotranspiration losses.

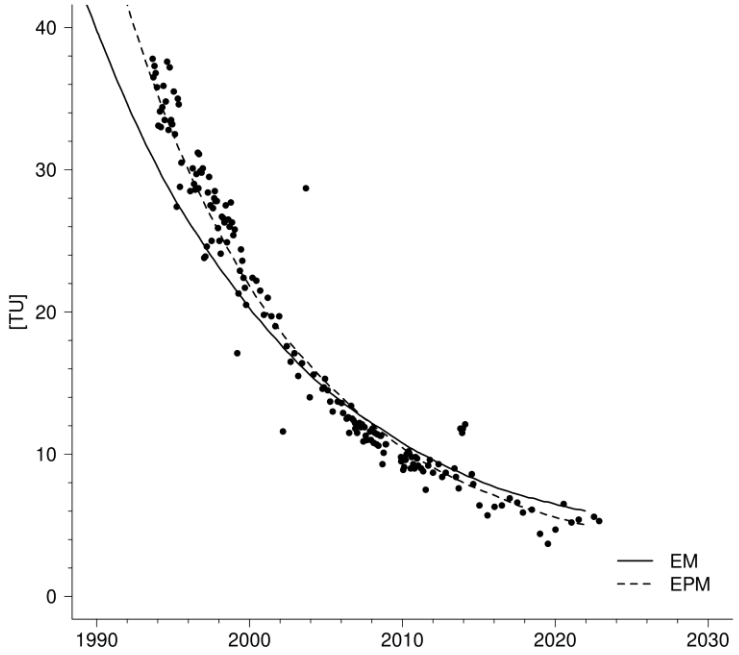



*Figure 11. Best fit comparison of an exponential model (EM) and an exponential piston-flow model (EPM)*

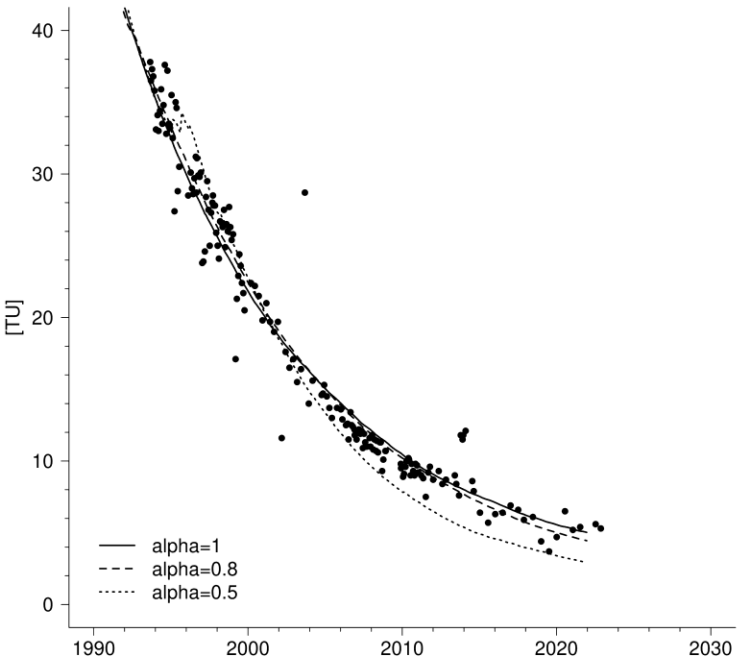



*Figure 12. Best fit of the EPM with different summer to winter infiltration ratios.*

Using a more complex model was also explored by fitting a gamma, dispersion and double exponential model to the output data (Figure 13) under the assumption that tracer transport undergoes some degree of hydrodynamic dispersion (gamma and dispersion models) or that total tracer flux is the result of two separate reservoirs with different volumes and mean transit times (double exponential model).

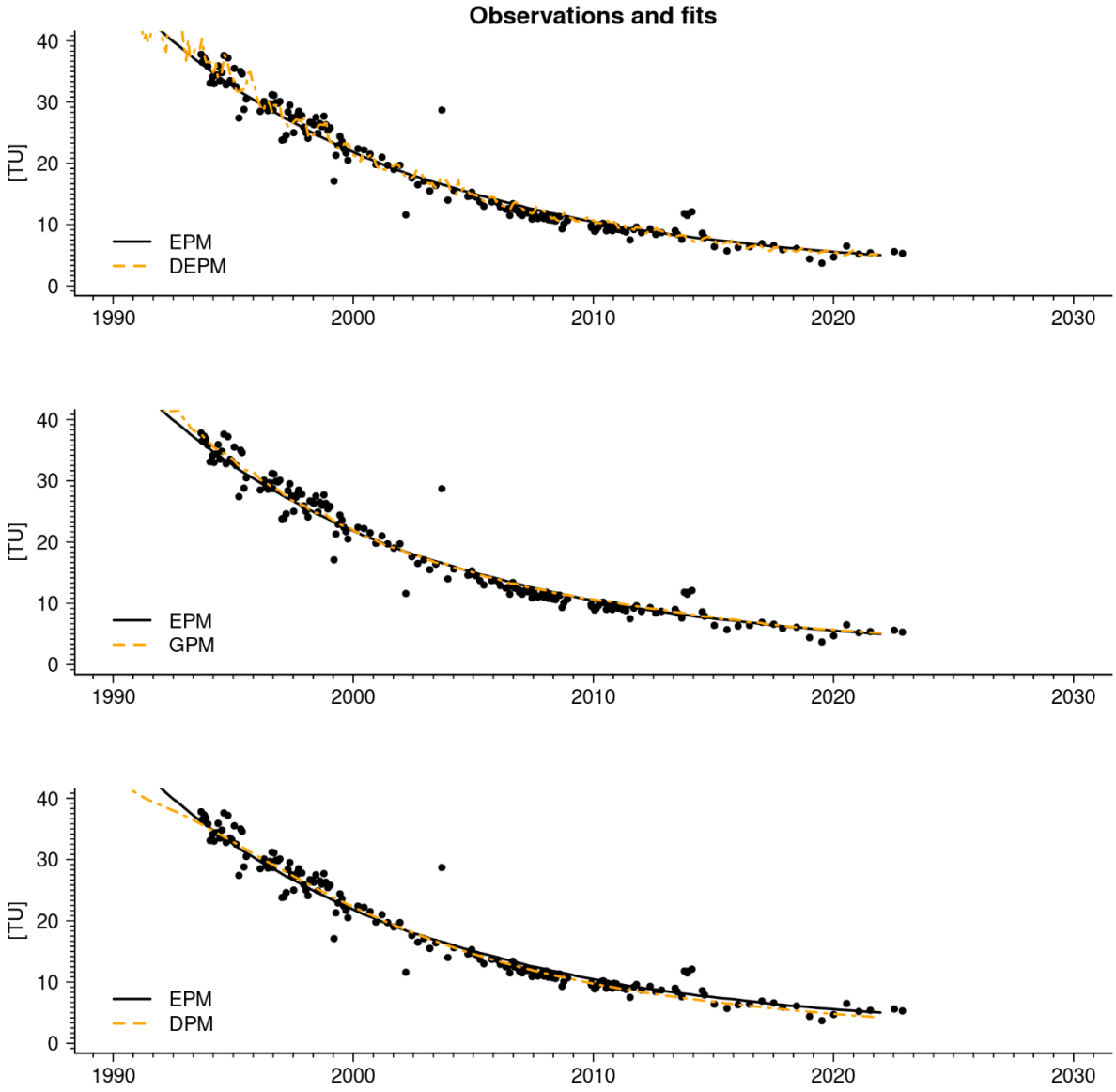



*Figure 13. Best fit comparison between an exponential piston-flow model (EPM) and either a double exponential piston-flow model (DEPM), gamma piston-flow model (GPM) or dispersion piston-flow model (DPM).*

The gamma and dispersion models simulate the additional effect of dispersion on tracer transport, i.e. tracer exchange between flow lines through mechanical dispersion and molecular diffusion,

whereas the double exponential model simulates the contribution of two reservoirs. In this case one reservoir would be a large groundwater storage and another reservoir is defined as a quickflow component of much smaller volume. The dispersion model improved the fit slightly by decreasing the overestimation of the exponential piston-flow model between 2005 and 2010, thus reducing the overall bias of the residuals from 0.27 TU to -0.02 TU. The gamma model yielded a worse fitting bias (0.39 instead of 0.27 TU) and did not reduce the range of the residuals. The long-term trend of the double exponential model followed as expected the predicted output of the exponential piston-flow model but was the only model that allowed a better match of the seasonal variations in the tritium output (Figure 14), thus reducing the range of the residuals, yet with an overall bias similar to the exponential model. Many remaining outliers evidently point towards additional runoff generation processes not captured by a simple annual cycle superimposed upon the slow decrease of the bomb peak tritium (Figure 15). Short-term variations in discharge following rainfall events as suggested by KRALIK (2015) may be one of those processes responsible for data outliers. In the absence of discharge measurements, this hypothesis could not be tested using the lumped parameter model's variable flow formulation. Considering the best two model (EPM and DPM) and a discharge of six litres per second, the total volume of water in storage is between 5.4 x $10^6$ $m^3$ and 11.1 x $10^6$ $m^3$. Without any information on fracture and matrix porosity, the volume of mobile water cannot be estimated (11.1 x $10^6$ $m^3$ being the upper estimate).

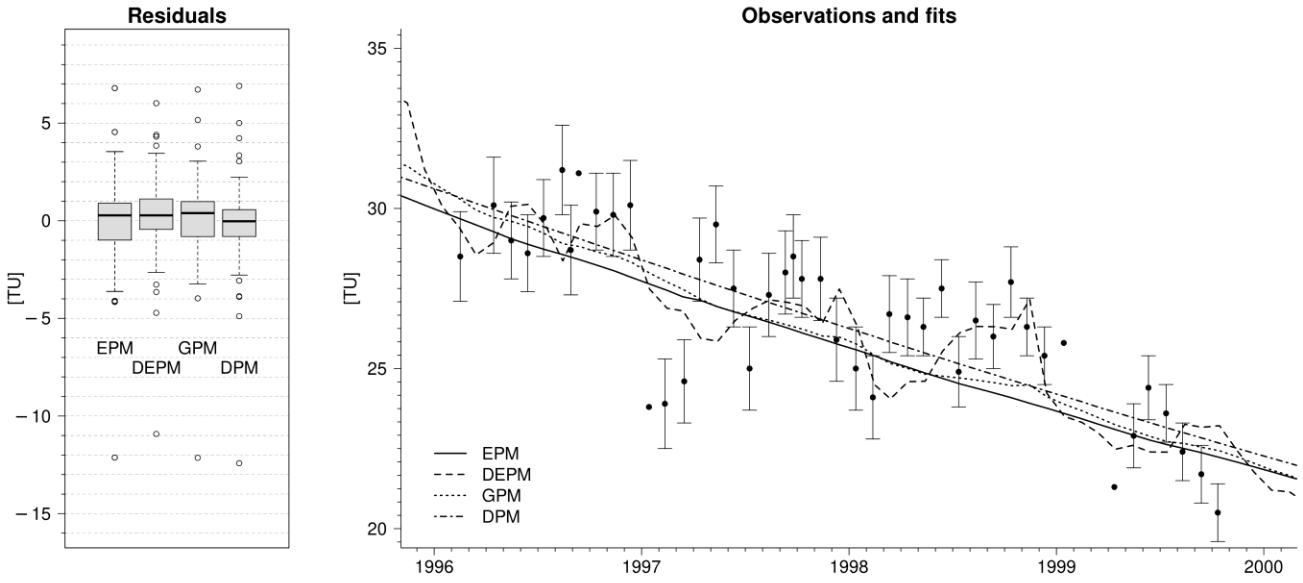



*Figure 14. Best fit comparison for an exponential piston-flow model (EPM), double exponential piston-flow model (DEPM), gamma piston-flow model (GPM) and dispersion piston-flow model (DPM) for the years 1996 to 2000. Left: Boxplots of the residuals for the entire observation period (1993-2022). Right: Time series. The error bars show the analytical uncertainties of the tritium measurements.*

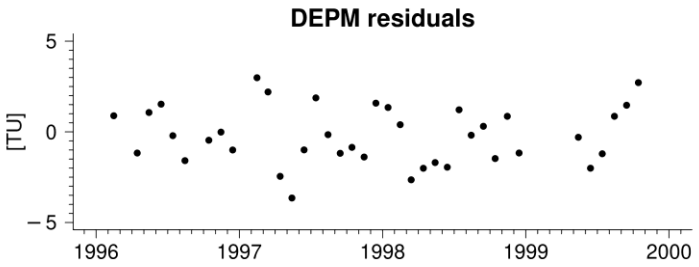



*Figure 15. Residuals of the DEPM model for the fit shown on Figure 14.*

*Table 2. Modelling results for the tritium data of the Weißbach spring. T: mean transit time without piston-flow. $T_{PF}$: mean transit time of the piston-flow component. $T_{EM}$: mean transit time of the quickflow reservoir of the double exponential model. The summer to winter ratio was set equal to one (i.e. all summer tritium contributes to the output).*

| Model | T [y] | $T_{PF}$ [y] | $T_{EM}$ [weeks] | $P_D/\alpha/f_{EM}$ [-] | volume in storage [$10^6$.$m^3$] | median error [TU] |
|---|---|---|---|---|---|---|
| EM | 32.22 | 0 | - | - | 6.2 | 0.3049 |
| EPM | 21.42 | 7.14 | - | - | 5.4 | 0.1743 |
| DEPM | 21.42 | 14.28 | 2 | 0.27 | 6.7 | 0.1682 |
| GPM | 32.86 | 14.28 | - | 0.55 | 8.9 | 0.1714 |
| DPM | 44.28 | 14.28 | - | 1.13 | 11.1 | 0.1684 |

**Bibliography**

ANIP – Austrian Network of Isotopes in Precipitation (2013): Österreichisches Netzwerk für Isotopen ($^{18}O$, $^{2}H$, $^{3}H$) in Niederschlägen und Oberflächengewässern. BMLFUW, Ämter d. Landesregierungen, ARC und Umweltbundesamt, Wien. **http://www.umweltbundesamt.at**.

KRALIK, M. (2015): Bestimmung der Mittleren Verweilzeiten von ausgewählten Brunnen- und Quellwässern in Vorarlberg mittels Isotopen- und Spurengasanalysen. 70 S., Unpubl. Ber. Amt der Vorarlberger Lrg. Bregenz.

KRASSER, L. (1952): Geologisches Gutachten über die Weissbachquellen im Gamperdonna. Unpubl. Gutachten, 17 S., Amt der Vorarlberger Landesregierung.

LEDITZKY, H.P., RAMSPACHER, P. & SUTTERLÜTTI, R. (1990): Hydrogeologische Untersuchungen zur Abklärung des Einzugsgebietes der Weißbachquellen (Gamperdonnatal. Vorarlberg). Steirische Beiträge zur Hydrogeologie, 41, 6-34.

SUTTERLÜTTI, R. (1990): Die Hydrogeologie des Gamperdonnatales unter besonderer Berücksichtigung der Weissbachquellen (Rätikon, Vorarlberg). Unpubl. Dissertation, 80 S., Technische Universität Graz, Graz.

## 4. 2 Using both tritium and oxygen-18: The Lange Bramke catchment, Germany

Alessio Gentile[1], Stefano Ferraris[1] and Julien Farlin[2]

[1]*Interuniversity Department of Regional and Urban Studies and Planning (DIST), Università degli Studi and Politecnico di Torino, 10125, Torino, Italy*

[2]*Luxembourg Water Agency, Luxembourg*

### Study site description

The Lange Bramke catchment (Figure 16), located on the windward slope of the Harz Mountains, Germany, covers an area of 0.76 $km^2$ with an elevation range of 543–700 m a.s.l. and is predominantly forested with Norwegian spruce (Maloszewski et al. 1999). The region, characterized by a nival hydrologic regime, receives an average annual precipitation of 1300 mm, with a runoff of 700 mm (Maloszewski et al. 1999).


*Figure 16 Geographical situation of the Lange Bramke catchment*

Herrmann and Schumann (2008) investigated the streamflow generation of this catchment and conceptualized three key successive hydrological processes:

I. The saturation of shallower soil layers is followed by rapid drainage through the macropores of the dominant podsolic brown soil (Maloszewski et al. 1999).
II. The drainage water produces an increase in subsurface pressure head (as indicated by the rise in piezometric level) inducing a mass displacement which is divided in *a)* vertical seepage to the unsaturated zone and *b)* lateral groundwater flow in the saturated zone. The 3.5 m thick unsaturated zone, consisting of residual weathering and allochthonic Pleistocene-age solifluidal materials (Maloszewski et al. 1999), remains unsaturated along the slopes of the catchment (Hauhs, 1989). This zone is extensively fragmented, thus

facilitating efficient seepage through the boundary between the unsaturated layer and the fractured-rock groundwater system.

III. The groundwater exfiltrates into the stream channels. The fractured-rock groundwater system, consisting of folded and intensely fractured sandstones, quartzites, and some slates of Early Devonian age (Maloszewski et al. 1999), has been confirmed as the primary source for water flows, while direct event water contribution was found to be negligible (Herrmann and Schumann, 2008).

In the Lange Bramke catchment, the groundwater recharge is quantitatively the most significant and continuous process throughout the year (Herrmann and Schumann, 2008). Indeed, the hydrogeological conditions seem to promote draining and vertical seepage (Maloszewski et al. 1999). The groundwater recharge process is also clear from the hydrometeorological data (Figure 17). In fact, it can be observed that from 1960 to 2007, precipitation consistently exceeded the sum of runoff and estimated evapotranspiration (Figure 17b). Over this period, the average annual precipitation, discharge and evapotranspiration were 1263 mm, 630 mm (both consistent with data reported in Maloszewski et al. 1999 and Lischeid, 2000) and 483 mm, respectively. The positive water balance suggests either *(i)* an average annual groundwater recharge of about 150 mm that could supply a deeper aquifer bypassing the catchment's outlet, or *(ii)* that the catchment-wide evapotranspiration calculated from point measurements is overestimated. Given the approximate nature of real evapotranspiration estimation, the latter possibility should not be discarded.

Herrmann and Schumann (2008) studied the runoff formation based on hydrological and hydrogeological measurements and the use of tracers. In this regard, both a historical time series (1980-2006) of monthly/multi-monthly tritium data in precipitation/discharge and a three-years-long (1983-1986) monthly/multi-daily time series of stable isotopes ($^{18}O$) in precipitation/discharge are available for this catchment.

The main objective of this work is to compare two modeling approaches, the classical lumped parameter model and the storage selection function widely described in the scientific literature (e.g., Zuber et al. 1986, Benettin and Bertuzzo, 2018) applying WITS to the tracers collected at the study site with the following aims:

1) investigate the effect of different *a)* flow regimes (steady and variable) and b*)* input weighting on the goodness of fit of an exponential lumped parameter model over discharge tritium data from 1980 to 2006 in the Lange Bramke catchment. We use the same "δ" metric used by Zuber et al. (1986) to evaluate the goodness of fit. We will further compare the results with the previous results by Zuber et al. (1986) which are based on data for the period 1980-1985, but also assessing the effect of applying the same modeling methodology to a time series that includes the 5-year period investigated by Zuber et al. (1986) but is approximately five times longer.

2) investigate the advantage of using the storage selection function approach, which includes the effect of evapotranspiration fluxes on streamwater age, by assuming a good mixing exponential model like the one used for the tritium data. We will investigate how the Mean Prediction Error (MPE), used as a best-fit metric, changes with variations in the $k_{ET}$ parameter, which governs the preference of the evapotranspirative flux to sample from the younger or

older portion of the storage. Finally, we will assess the effect of $k_{ET}$, and therefore of evapotranspiration, on the runoff Transit Time Distribution (TTD). A summary of the meaning of the parameters used in the storage selection function modeling approach will be presented below.

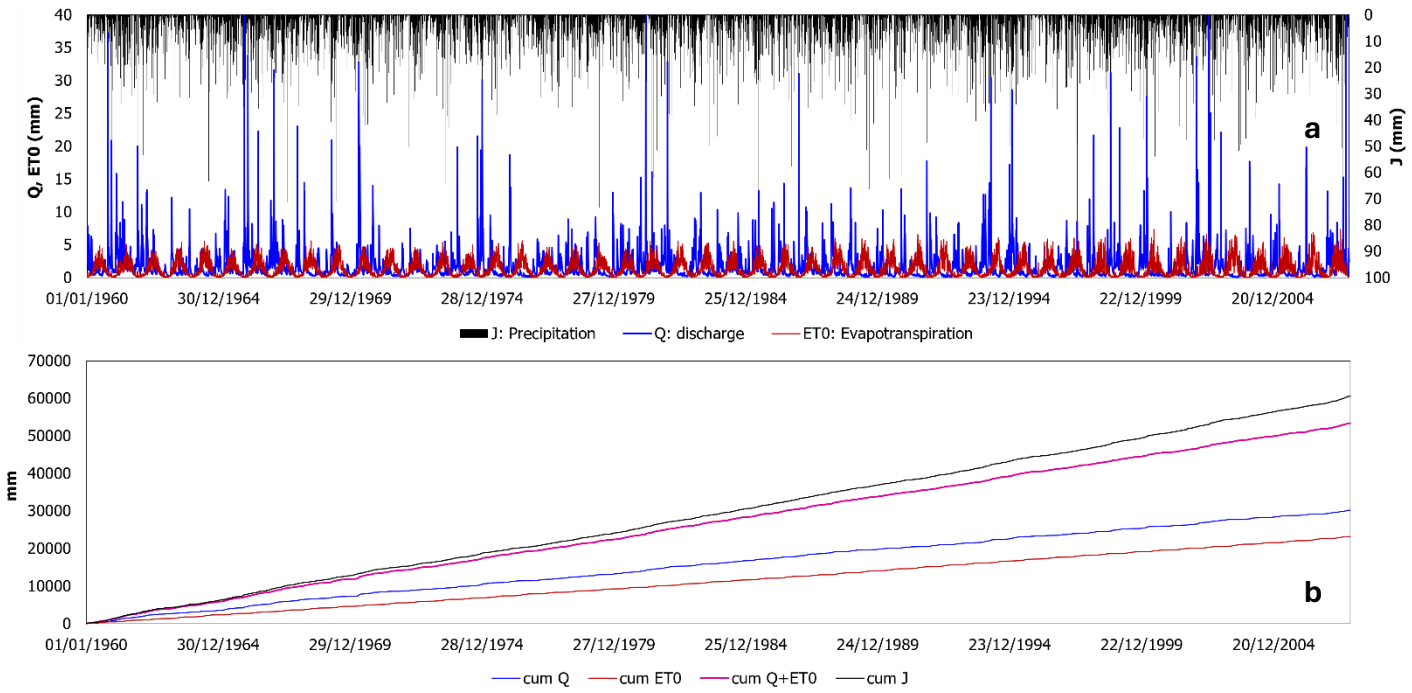



*Figure 17* ***a)*** *Measured inflow and outflow at the daily scale* ***b)*** *Cumulative inflow and outflow.*

## Classical Lumped Parameter Models with tritium data

The LPM approach was previously applied by Zuber et al. (1986) to tritium data measured in the Lange Bramke catchment. The authors used tritium data both in precipitation and discharge, collected from 1980 to 1985, and applied an exponential lumped parameter model under both steady (weighted and unweighted) and variable flow conditions.

The unweighted input steady-state conditions represent the case in which the ratio of summer to winter infiltration coefficient is equal to 1, while the weighted input steady-state conditions represent the case in which this ratio is comprised between 0 and 1. This ratio is represented by the *a*-parameter (Table 3). See the work of Zuber et al. (1986) for further methodological details. In this study as in Zuber et al. (1986), we consider two different conditions: $a = 1$ and $a = 0.5$. At weighting of "1" the isotopic signal in precipitation is directly used in the calculation. When on the contrary the precipitation isotope signal is believed to not fully contribute to the stream isotope signal during the summer month due to evaporation losses, weighting may be applied to reduce the injected tracer mass used for calculation. Here, a=0.5 is applied for the months from April to September.

To simulate variable flow conditions, it is necessary to estimate the dynamic turnover time ($t_d$), assumed to be constant in groundwater systems (Zuber et al. 1986). It can be obtained from hydrological measurements as the ratio of the dynamic storage and discharge (Zuber et al. 1986).

From 20-years observation Zuber et al. (1986) estimated a $t_d$ weighted mean value of 8.7 days and we have used the same value for the variable-flow simulation.

In summary, we simulated the discharge tritium data following the approach of Zuber et al. (1986) but using the output tritium data from 1980 to 2006. We then calculated the same δ metric used by the authors to assess the goodness of fit over the entire time series (Figure 18a, Table 3), as well as for the period 1980-1985 (Figure 18b, Table 3) previously investigated by Zuber et al. (1986). This approach allows us to determine whether the availability of additional output data after 1985 has improved the fit for the 1980-1985 period.

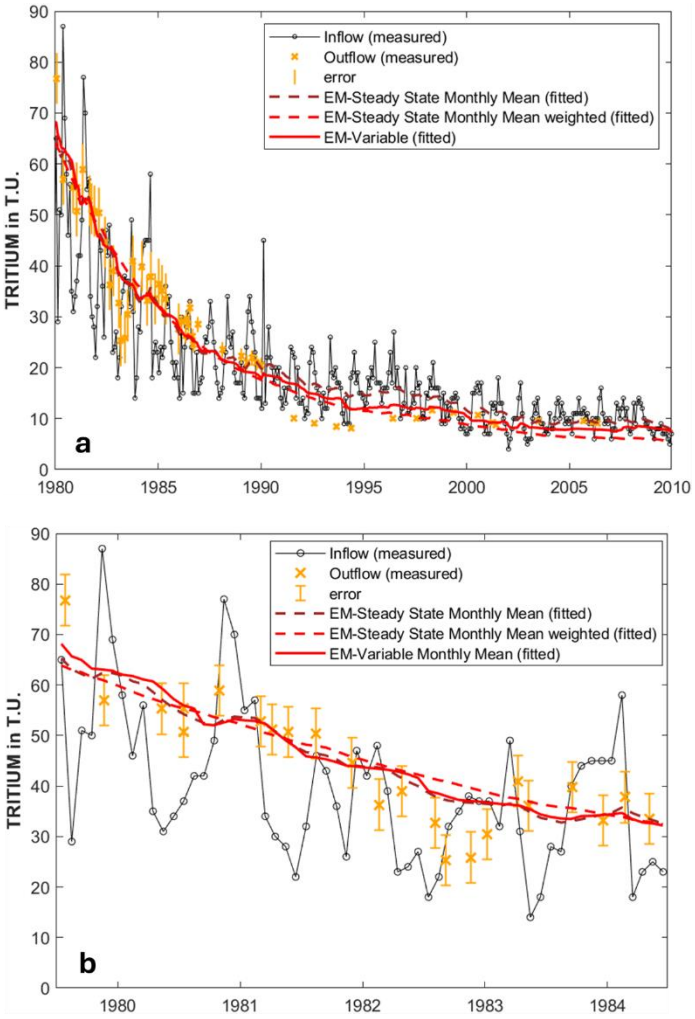



*Figure 18 Tritium data in the ouflow from the Lange Bramke catchment and fitted curves by considering **a)** all the available data in the period 1980-2006 and **b)** available data in the period 1980-1985 previously investigated by Zuber et al. (1986).*

***Table 3*** Summary of parameters used in the simulation and the accuracy of fitting (δ metric). MTT: mean transit time.

| Flow | *a* (-) | $t_d$ (d) | δ (T.U.) 1980-2006 | δ (T.U.) 1980-1985 | δ (T.U.) Zuber et al. (1986) | MTT (y) 1980-2006 | MTT (y) Zuber et al. (1986) |
|---|---|---|---|---|---|---|---|
| Steady | 1 | - | 0.63 | 1.21 | 1 | 1.55 | 1.8 |
| Steady (weighted) | 0.5 | - | 0.64 | 1.39 | 0.99 | 5.05 | 2.2 |
| Variable | - | 8.7 | 0.59 | 1.23 | 0.85 | 2.59 | - |

The results reported in Table 3 show that, by using a long-term (1980-2006) time series of tritium data, the weighted and unweighted steady-state flow conditions return a similar δ metric, but the mean transit time is longer than roughly 3.5 years when a weighting factor is applied. This result differs from what was found by Zuber et al. (1986) in which the mean transit time in steady conditions (weighted and unweighted) were similar by differing of 0.4 years (Zuber et al. 1986). This could be due to the inclusion of very different meteorological conditions in the period from 1980 to 2006 that could not be included in the much shorter period of five years between 1980 and 1985. Under variable flow condition the δ metric is better, but similar, to that obtained under steady-state conditions and this is consistent with results of Zuber et al. (1986). The auhors explained this result by considering that the variable storage of this catchment is only a small fraction of the total storage. By fitting the exponential lumped parameter model to the long-term (1980-2006) time series of tritium data resulted in a slighlty worse δ metric for the period investigated by Zuber et al. (1986) due to extremely low tritium values between 1990 and 1995.

## Storage selection functions with stable isotopes

We applied the storage selection function modelling approach to the Lange Bramke $^{18}O$ data collected in the period from 1983 to 1986. In this modelling approach three parameters can be calibrated: $k_Q$, $k_{ET}$ (the subscript *Q* refers to discharge, while the subscript *ET* refers to evapotranspiration) and $S_0$. If $0<k<1$ there is a preference of *Q* or *ET* fluxes for younger ages sampling from the storage. Conversely, if $k>1$ there is a preference of *Q* or *ET* fluxes for older ages sampling from the storage. The case of $k=1$ represents a random age sampling of *Q* or *ET* fluxes from the storage (i.e., no age preference) equivalent to the exponential model fitted above using tritium data. The parameter $S_0$ represents the initial total storage. Zuber et al. (1986) estimated a total storage of 1189 mm for the Lange Bramke catchment. Accordingly, we calibrate this parameter by considering a $S_0$ range from 500 to 1500 mm in order to include the best fit obtained from tritium measurements.

We investigated the sensitivity of $k_{ET}$, by considering $k_Q=1$ which corresponds to an exponential model (Benettin and Bertuzzo, 2018) and only varying the value of $k_{ET}$ during calibration. With this approach it is possible to understand the advantage of using the storage selection function approach which includes the effect of evapotranspiration fluxes. Thus, we perform 16 simulations by considering $k_{ET}$ values ranging from 0.3 to 5. The simulations consider a time-variant storage selection function in which we have estimated the wetness index from a normalization of the storage variations of the system (Benettin and Bertuzzo 2018). The initial isotope content of the storage input must be set when using the storage selection function modelling. This value was set equal to the

average isotopic composition ($\delta^{18}O$) of streamwater during low-flow conditions that was equal to -9.12 ‰.

The obtained results are reported in Figure 19. It can be observed that the Mean Prediction Error (MPE) is highly sensitive for $k_{ET}$ values below 1 and MPE decreases as $k_{ET}$ decreases (Figure 19b). For $k_{ET}$ values greater than 1, the MPE becomes insensitive. As previously mentioned, $k_{ET}$ values between 0 and 1 indicate a sampling preference of *ET* flux for younger ages of the storage, which in a catchment such as the Lange Bramke would physically correspond to the upper soil layers (where it is realistically possible for plant roots to uptake water and for evaporation to take place). Conversely, $k_{ET}$ values greater than 1 indicate a sampling preference of *ET* flux for older ages of the storage, which physically corresponds to the deeper layers of the catchment (out of the soil zone into the underlying aquifer), beyond the reach of plant roots. Therefore, $k_{ET}$ values greater than 1 should be considered highly unlikely.

It is also evident that the TTD on a mid-summer day (10-07-1985), when *ET* is most influential, is highly sensitive to low $k_{ET}$ values (recent water). As $k_{ET}$ increases, the frequency of shorter residence times also increases, as the *ET* flux mainly samples older water.

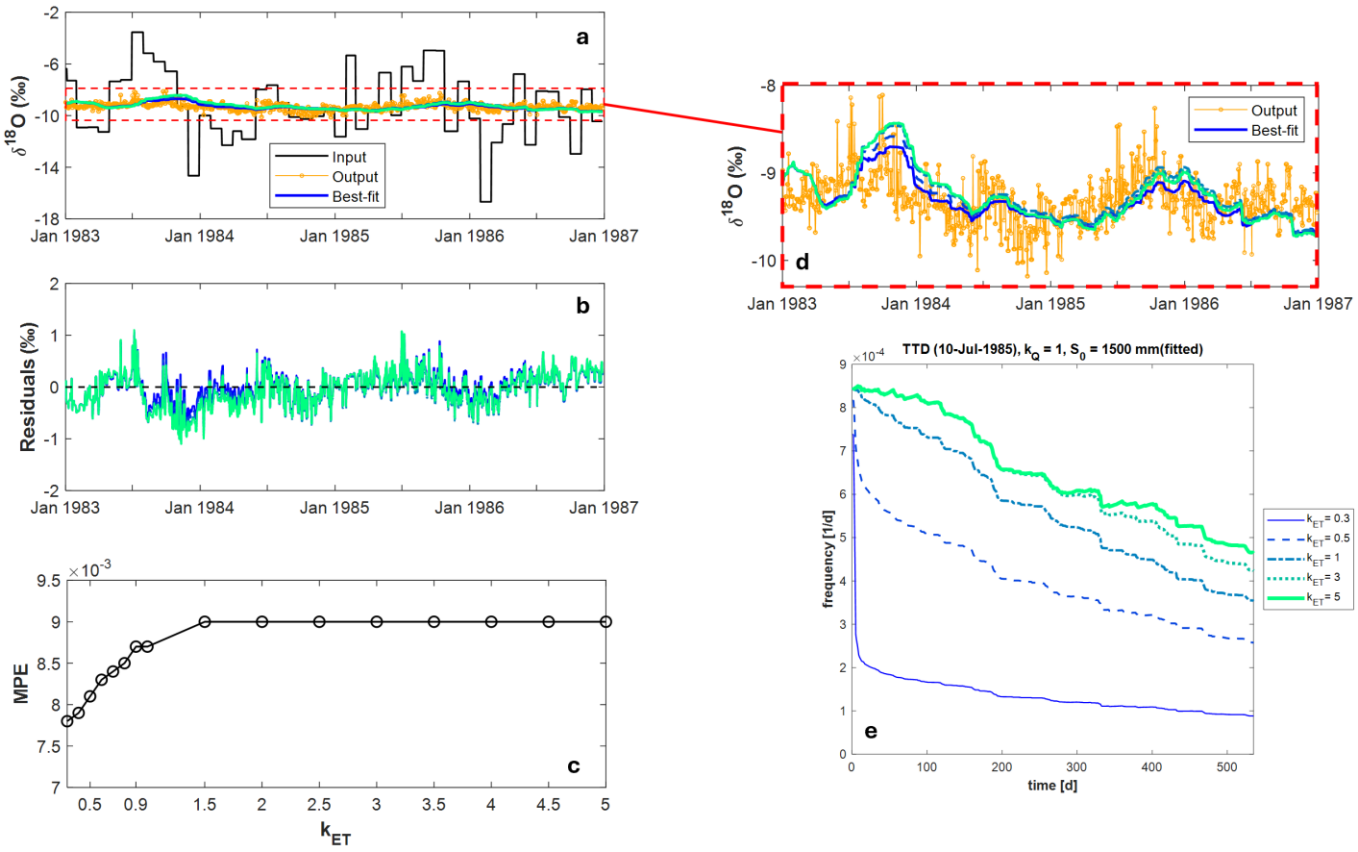



*Figure 19* ***a)*** *Observed isotopic composition in precipitation and streamflow with indication of modelled isotopic composition in the streamflow by using the SAS (with varying $k_{ET}$)* ***b)*** *Residuals (observed minus simulated isotopic composition) with varying $k_{ET}$* ***c)*** *Mean Prediction Error (MPE) with varying $k_{ET}$* ***d)*** *Focus on observed vs modelled (with varying $k_{ET}$) isotopic composition in streamflow* ***e)*** *Transit Time Distribution with varying $k_{ET}$ during a mid-summer day.*

## References

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

Hauhs, M.: Lange Bramke: An Ecosystem Study of a Forested Catchment, in: Acidic Precipitation: Case Studies, edited by: Adriano, D. C. and Havas, M., Springer, New York, NY, 275–305, https://doi.org/10.1007/978-1-4612-3616-0_9, 1989.

Herrmann, A. and Schumann, S.: Runoff formation in a small mountainous basin dominated by a fractured rock aquifer: results from the tracer-based integrated catchment approach (ICA), 12th Biennial International Conference of the Euromediterranean Network of Experimental and Representative Basins (ERB), Kraków, Poland, 103–109, 2008.

Lischeid, G.: Neither macropores nor interflow: generation of spiky doscharge peaks in small forested catchments and its implications on stream water chemistry, Runoff Generation and Implications for RIver Basin Modelling, Institut fur Hydrologie der Universitat Freiburg, 46–53, 2000.

Maloszewski, P., Herrmann, A., and Zuber, A.: Interpretation of tracer tests performed in fractured rock of the Lange Bramke basin, Germany, Hydrogeology Journal, 7, 209–218, https://doi.org/10.1007/s100400050193, 1999.

Zuber, A., Maloszewski, P., Stichler, W., and Herrman, A.: Tracer Relations in Variable Flow, 5th International Symposium on Underground Water Tracing, Institute of Geology and Mineral Exploration, Athens, 355–360, 1986.

## 4.3 Long-term data series of monthly stable isotopes measurements in precipitation and runoff: the Jalovecký Creek, Western Tatra Mountains, Slovakia

Ladislav Holko[1] and Julien Farlin[2]

[1]*Institute of Hydrology, Slovak Academy of Sciences, Bratislava, Slovakia*

[2]*Luxembourg Water Agency, Luxembourg*

The study site is a small mountain catchment (area 22.2 km$^2$, mean altitude 1500 m, mean slope 30°C, attitude range 820-2178 m a.s.l.) in northern Slovakia (Figure 20). Crystalline rocks (schists, gneiss, migmatite) and granodiorites are found in 48% and 21% of catchment area respectively, while Mesozoic rocks (mainly limestone and dolomite occupy about 7% of the catchment. Slope debris and glaciofluvial sediments cover 24%. The bedrock is weathered. The catchment is generally south oriented. Cambisol and podzols are the main soil types. The soils are shallow (soil depth up to 1 m) and contain a large percentage of skeletons (20%-60%). Forest (mainly spruce), dwarf pine and alpine meadows (including bare rocks) cover 44%, 31% and 25% of the catchment. Mean annual catchment precipitation and runoff is about 1500 mm and 1034 mm respectively. Mean annual air temperature at catchment mean altitude is about 3°C. The catchment represents a snowmelt dominated hydrological system. Snow cover is present there approximately from December until the beginning of May.

Monthly composite precipitation sampling and grab runoff sampling for stable isotopes started in the catchment in November 1990 and November 1991, respectively. Precipitation data is collected at the GNIP station in the downstream part of the Jalovecký Creek catchment (altitude 570 m a.s.l) which is outside of the catchment described above (Site 1 in Figure 20). Streamflow samples are collected at catchment outlet at 820 m a.s.l. (Site 2 in Figure 20).

Previous estimations of the mean transit time (MTT) by means of lumped models (exponential and dispersion) based on shorter data series (1991-1993 and 1991-2002) provided mean transit times of approximately 31 and 10-13 months respectively. Mean transit times of 9-19 months (27 months for the period 1991-1993) were obtained from the sine curve approach while a mean transit time of 12 months was calculated for the average weighted monthly isotope signatures in precipitation and runoff.

Figure 21 presenting available data shows slightly different mean $\delta^{18}O$ in runoff (output data series) in the periods 1991-1993 (around -11.1‰), 1998-2001 (-11.4‰), 2005-2014 (-11.7‰) and 2015-2022 (-11.2‰). Therefore, different lengths of the output data series were tested for parameter estimation using WITS.

The young water fractions estimated from the ratio of $\delta^{18}O$ amplitudes in runoff to that in precipitation were 0.05 (1991-1993), 0.09 (1998-2001), 0.08 (2006-2014) and 0.10 (2015-2022). Measurement of stable isotopes in precipitation at catchment mean altitude (1500 m) for the limited period (all together almost eight hydrological years) showed that the amplitude of $\delta^{18}O$ in precipitation was smaller than that at 570 m a.s.l. Thus, the young water fractions for the Jalovecký Creek catchment

are probably a few per cents greater than those calculated from the $\delta^{18}O$ amplitudes at the GNIP station.

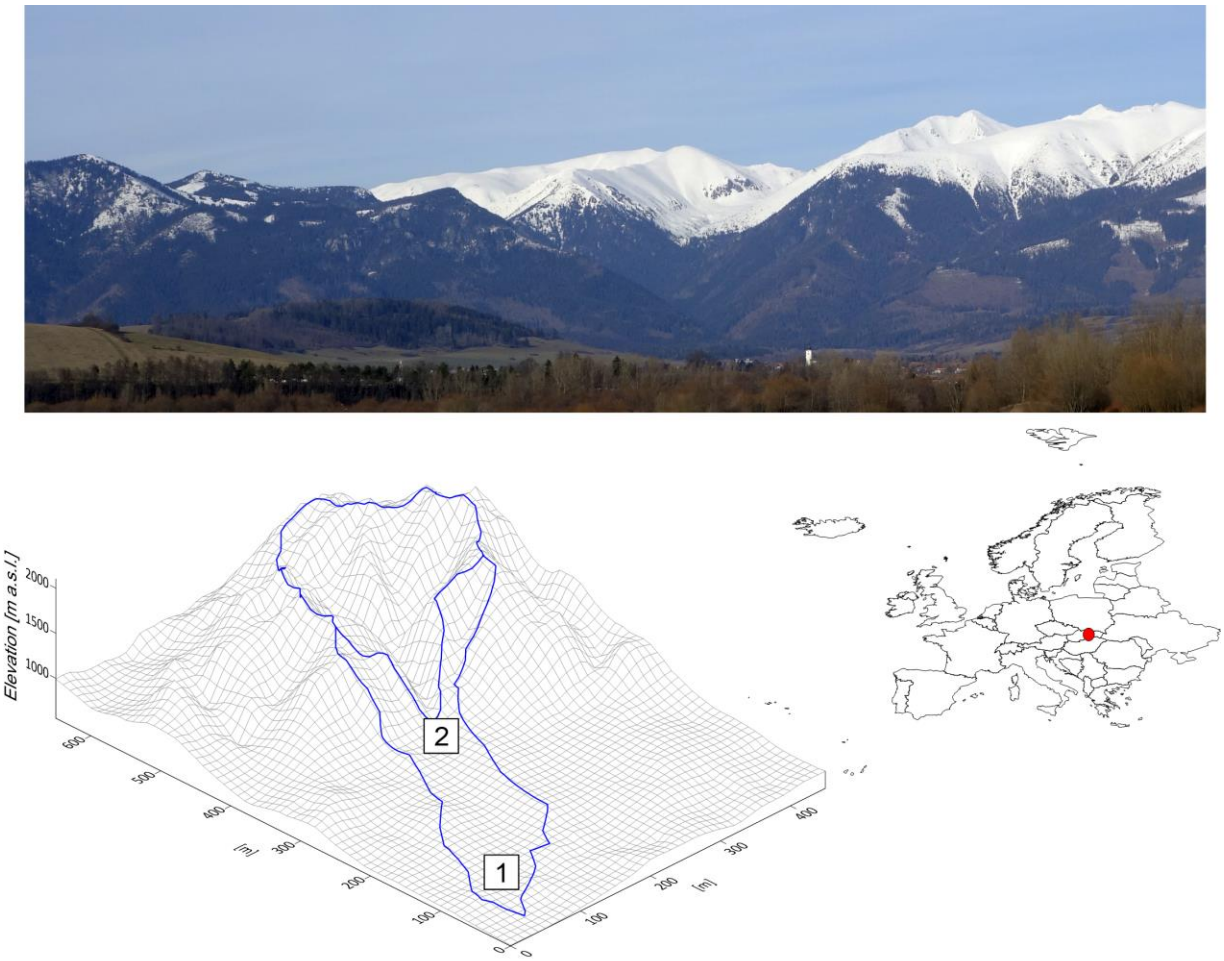



*Figure 20 The Jalovecký Creek catchment; isotopic composition of precipitation and runoff is measured at sites 1 and 2, respectively.*

The input data series for the mean transit time modelling using the lumped parameter models covered the entire period November 1991-October 2022. The inputs were not weighted or corrected. The exponential model provided smaller mean prediction errors (MPE) and shorter mean transit times (mostly 2-3 years) than the gamma and dispersion models (5-9 years). The exponential-piston flow model yielded long mean transit times of 7 years.

Calibration against monthly means yielded smaller MPEs, but greater ranges and mean values of residuals than calibration against the annual averages. While modelling with the exponential model based on the shorter output data 2006-2014, i.e. before the increase in measures streamflow $\delta^{18}O$ provided similar MTTs and MPEs as for periods 1991-2022 and 2006-2022, it was not possible to obtain reasonable results for the output period 2015-2022 with any input data series (neither covering the entire period 1991-2022, nor shortened periods 2013-2022 or 2015-2022). The best result considering the reasonability of calculated mean transit times, MPE and residuals was obtained from the exponential model calibrated against annual values.

Results from two simulations using an exponential model calibrated against measured streamflow isotopic composition from the periods 2006-2022 (best fit 1: MTT 2 years, MPE 0.02‰) and 2006-2014 (best fit 2: MTT 3 years, MPE 0.03‰) are shown on Figure 22.

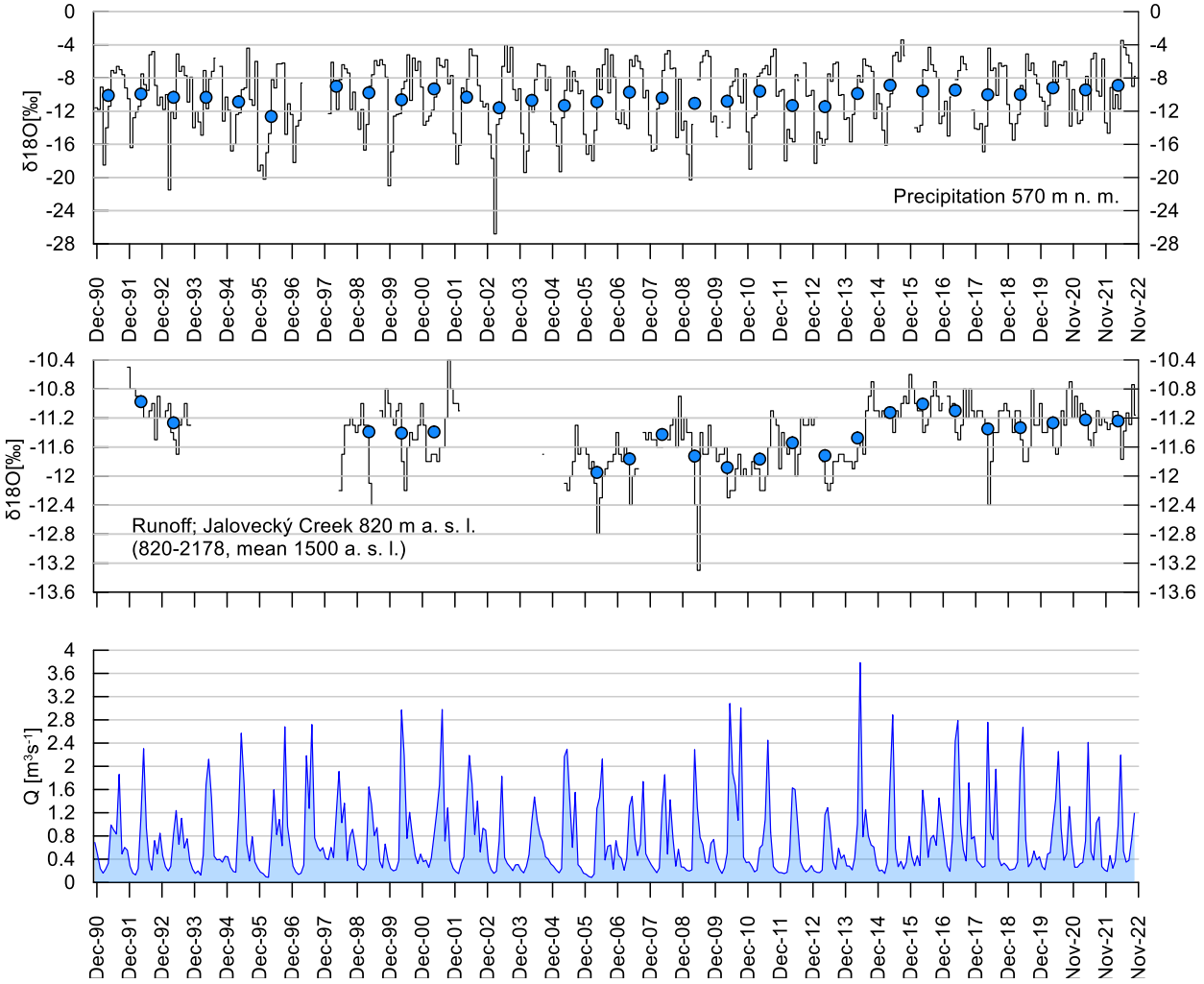



*Figure 21 Monthly $\delta^{18}O$ in precipitation (composite precipitation samples) and runoff (grab samples) and monthly catchment discharge in hydrological years 1991-2022 (November to October); the blue dots represent average annual values.*

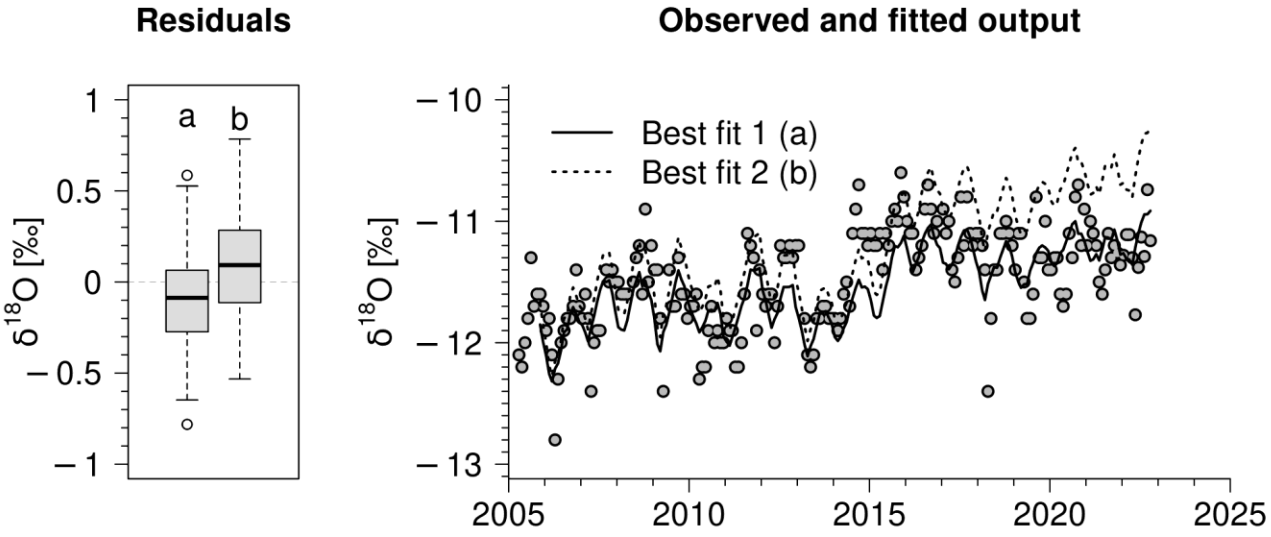



*Figure 22. Modelling results. Left panel – boxplots of residuals for Best fit 1 and Best fit 2 (the whiskers represent percentiles 10 and 90, the average is represented by the median; right panel - comparison of measured output data in hydrological years 2006-2022 with the results obtained from the exponential models fitted against input data from the period 2006-2022 (Best fit 1) and from the period 2006-2014 (Best fit 2).*

The calibration results shown on Figure 22 were obtained using the measured isotopic content in precipitation. Since however precipitation occurs between December and May in the form of snow, it is reasonable to assume that during that time, streamflow is not sustained by recent precipitation, but by percolation from deeper layers, and thus, that the input during the months were the catchment is snow covered deviates from raw precipitation. To explore this hypothesis, the input function used for a second modelling attempt was prepared as follows for the period 2004-2022:

- a value of either -11‰, -11.5‰ or -12‰ was assigned to the months December to March to simulate a steady groundwater contribution to streamflow without precipitation input. The value changed from year to year to correspond to the mean measured isotopic signal in streamflow for the previous four months (August to November).
- The snow melt pulse was then simulated by setting the input value to -16‰, -15‰ and -14‰ for April, May and June respectively. These values corresponding to the snow melt contribution were estimated from lysimeter measurements (Holko, unpublished data). In 2010 and 2011, snow melt started earlier due to a snow poor winter, and the snow melt contribution was accordingly shifted by one month (March to May).
- The measured isotopic values in rainfall were kept for the months July to November.

The predicted outputs for the raw isotopic rainfall measurements and the modified input function are shown on Figure 23. The annual dip in $\delta^{18}O$ values in the spring can be explained either by a snow melt release using the modified input function or can be fitted as the response to the annual cycle of the isotopic values in precipitation. Some features of the observed output are better matched with the modified input (for instance the tracer decrease and subsequent recovery in spring 2016 and spring 2018) although the overall bias is slightly larger (left panel on Figure 22). The obtained mean

transit times are however very similar (2.6 years using the raw rainfall measurements and 2.3 years using the modified input). Given the large influence of snow accumulation and snow melt on the hydrological response of the catchment (Figure 21) we conclude that:

- Snow accumulation and snow melt release modelled by modifying the input function using information about the period of snow accumulation and the probable isotopic value of snow melt (here from snow lysimeter measurements) is a priori realistic and matches output observations as well as the model using the raw precipitation.
- In cases of snow accumulation and subsequent prolonged snow melt pulse the tracer dip and the following recovery could be mistaken for seasonal tracer cycles and fitted erroneously using the isotopic composition of precipitation.

The matter of the preparation of the input function from precipitation data (rainfall and snow) and ancillary information in snow dominated catchments deserves more research and should be kept in mind when modelling the isotopic response of such systems.

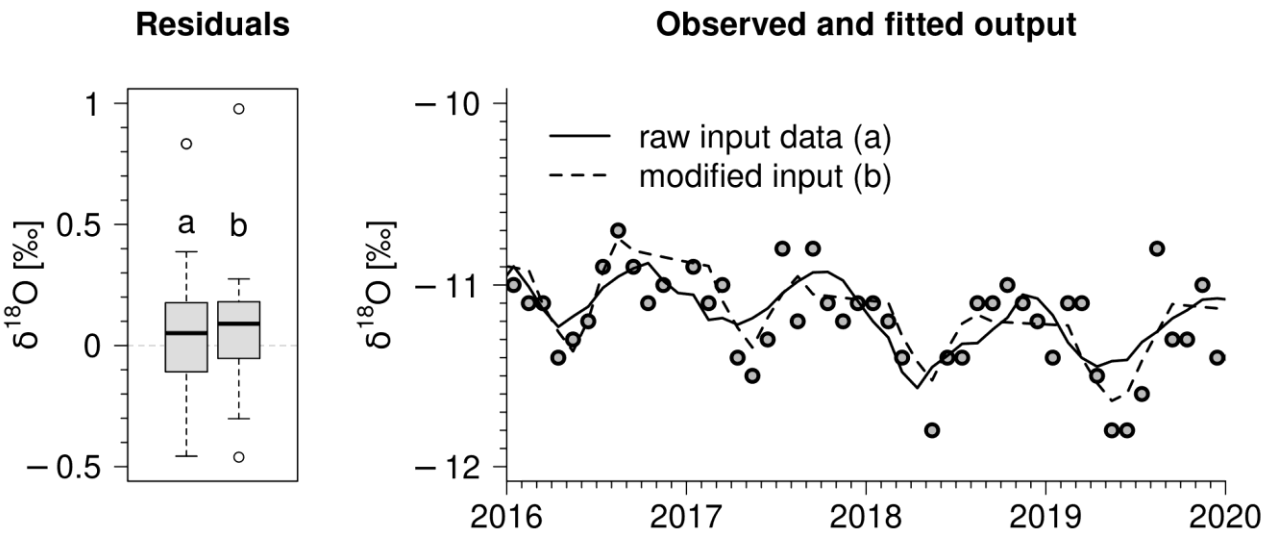



*Figure 23. Comparison between the best fit of an exponential model using the raw isotopic precipitation data as input and a modified input function simulating snow storage and the subsequent snow melt. Left: Residuals (a) raw data, (b) modified input. Right: Observed versus modelled output. The mean transit time was respectively 2.6 and 2.3 years.*

# 5. Modelling guide

The WITS (Water Isotope modelling for Transit time and Storage) program implemented in R is an interactive modelling platform which can be used to estimate transit time distributions in hydrological systems from isotopic tracer data either with the classical lumped parameter models or the storage selection functions. Tracer transit time distributions can be used to assess flow dynamics and calculate contaminant transport times.

The WITS framework is an alternative to other numerical tools (FlowPC, Tran-SAS, Lumpy, TracerLPM, Boxmodel, etc..) developed over the years to fit input-output models using environmental tracers. The novelty of WITS lies in the possibility of using different approaches in the **same** environment and thus more easily compare them, namely (i) the classical lumped parameter models in the steady state approximation, (ii) the classical lumped parameter models in transient mode and (iii) the storage selection functions in either steady state or transient mode. The authors of WITS hope this will provide students, scientists and consultants with an easier access to the estimation of storage and release dynamics in hydrological systems from environmental tracer data.

This report documents version 1.0 of WITS. The latest version of WITS can be downloaded from Github (https://github.com/srlutz/wits) and from the WATSON website (https://watson-cost.eu/storage-and-transit-time-estimation-in-hydrological-systems/).

## 5.1 Conceptual model

Before beginning with the calibration of the transfer function, the experimenter should prepare a conceptual sketch comprising all compartments that should be modelled, each represented by a box and arrows indicating the fluxes in and out of it and summarizing the number and meaning of the parameters characterizing the compartment (see examples on figures 5 to 8). A model that does not correspond to the flow domain can theoretically yield a good fit, but this fit is meaningless if it is hydrologically indefensible. The processes that need to be modelled depend on what the sought information is AND on the available data. Modelling for its own sake or with insufficient data is to be avoided. If only total storage is of interest for instance, a one compartment model may be sufficient. If quickflow response time is to be estimated additionally to the base flow contribution, a piston flow component **in parallel** is often necessary (Figure 8).

## 5.2 Model choice

The simplest basic model available to describe the transit time distribution of many hydrological systems (in particular groundwatersheds and surface water catchments) is the **exponential model**. This model is the exact solution of a flow balance equation for semi-confined aquifers of any shape, and for which the storage divided by the recharge rate is constant (Haitjema 1995). The exponential model describes purely advective flow (no exchange between flow lines due to diffusion or dispersion, no sorption and desorption exchanges with stagnant water in micropores). As such, it can serve as reference for modelling as the simplest possible model against which the performance and predictions of more complex models can be tested. If the fit does not improve with increasing model complexity, it is a strong argument to consider the exponential model as the most parcimonious model reproducing observations. The exponential model can be implemented both with the lumped parameter models and the storage selection function approaches.

## 5.3 Choosing between the classical lumped parameter models and the storage selection functions

Input-output models offer many different options for modelling. One of the first choices that must be made by the experimenter is whether to use the classical lumped parameter models or the storage selection functions. This chapter reviews the differences and similarities between both approaches in order to help WITS users make informed choices.

### 5.3.1 Underlying transit time distributions

As suggested in 5.2, using a simple model as starting point (in most cases the exponential model) can provide useful insight into the need of increasing model complexity. The exponential model is available in both approaches. The classical lumped parameter models offer a range of physically-based models for the transit time distribution besides the exponential model (see Table 1 and chapter 3.8.1). Each model corresponds to different geometries of the flow domain, or model different transport processes (pure advection, diffusion, hydrodynamic dispersion). The storage selection functions offer more flexibility in the choice of the underlying transit time distributions. These models however have yet to be derived from first principles. This means that their applicability to different hydrological situations is for the present unknown and that their selection as well as their parameterization remain empirical, without a clear relationship between parameter value and physical processes that could guide modelling and help validate or reject certain values as physically improbable. It is the experimenter's responsibility to check the physical plausibility of the predicted results in terms of preferential selection of transit times. For instance, if evapotranspiration is modelled using the storage selection functions with a power law, "older" storage should not be preferentially selected without a good reason, as deeper soil horizons in most cases contribute less to evapotranspiration than shallower ones, unless for instance particular plants with deep roots present in the study area can remove water from deeper layers.

### 5.3.2 Variable flow

Variable flow can be modelled both using the classical lumped parameter models and the storage selection functions as discussed in 3.8.2 using the same equation replacing normal time by cumulative flow exiting the system.

### 5.3.3 Variable storage

Variable storage is related to variable flow. As discussed in 3.8.2, the storage state variable drives the water and tracer fluxes, but it also controls flow path activation and deactivation. Changes in flow paths fluxes, and thus changes in the total transit time distribution at the system's outlet, can be modelled either by using a combination of models in parallel (see paragraph 3.8.2 and for instance Małoszewski et al. (2002) for an example application) or by switching between functions depending on storage (Benettin & Bertuzzo, 2018). The second approach is implemented in WITS for the storage selection functions using the power law time variant described in Benettin and Bertuzzo (2018).

### 5.3.4 Modelling evapotranspiration

As presented in 3.8.3, tracer losses via evapotranspiration can be modelled either statically with the classical lumped parameter models, or dynamically with the storage selection functions. The former

requires an estimation of the summer to winter ratio obtained either from stable isotope measurements or from hydrological data. In some cases (see the Vorarlberg case study), estimating independently even this single additional parameter is impossible with the data available, in which case it becomes an additional fitting parameter. This usually means that no unique solution can be obtained from tracer data.

Modelling evapotranspiration losses with the storage selection functions requires (i) an evapotranspiration time series and (ii) the calibration of the corresponding storage selection function. With the notable exception of weighable lysimeter studies (Stumpp, Stichler, & Małoszewski, 2009) evapotranspiration will always have to be *estimated* from meteorological data. It should be kept in mind that this estimation introduces hidden assumptions and potential biases in the modelling. Furthermore, the challenge of estimating the parameter(s) of the storage selection function for evapotranspiration should not be underestimated and has been discussed in this report (case study of the Lange Bramke) as well as by Sprenger et al. (2022). Trying to model a process with insufficient data for robust calibration should be avoided.

### 5.3.5 Numerical implementation

The convolution of the classical lumped parameter models is implemented numerically in an explicit way by discretizing the transfer function. On the other hand, the partial derivatives of the rank storage $S_T$ have no analytical solutions which means that a numerical approximation must be used for computation (Benettin & Bertuzzo, 2018). Both implementations have been benchmarked against simple models and yield the same result within an accuracy of less than 0.1%. The difference in numerical implementation however is reflected in calculation time. Calculations are faster with classical lumped parameter models compared to the storage selection functions.

## 5.4 Calibration

The aim of model fitting is to estimate the parameters of the transfer function that yields the best prediction of the measured output. If the fit is poor, modelling can be improved in two ways:

- By adapting the model structure (adding new compartments, changing the transfer function (section 3.8.1) or modelling variable flow (section 3.8.2))
- By modifying the input function (section 3.8.3)

These operations must be guided by a sound understanding of the system being modelled and based on physical effects that could explain a bad fit. For instance, the original input might not take properly into account snow melt storage (see the Jalovecky case study) or the effect of evapotranspiration. Manipulating the input function or the transfer function solely for the sake of improving the fit without a clear physically based justification should not be attempted.

Calibration should not be done to reproduce observations blindly, but to extract tracer information through modelling until the question at the root of the study has been answered or the residuals show no remaining structure. Any structure present in the residuals indicate remaining information that was not extracted by the model used. If for instance the amplitude of seasonal variations is underpredicted by the calibrated model, the time series of the residuals will display seasonal variations of the remaining amplitude. In that case, it is advised to modify the conceptual model for instance by adding another compartment. This should be done only *if a physical explanation can be*

*devised for the remaining structure of the residuals*. The new model will not be overfitted, because it is constrained by the remaining information contained in the residuals. On the other hand, the information remaining in the residuals may not be relevant for the question to be answered by modelling, in which case, it needs not be modelled. When for instance modelling the decrease of the bomb tritium over time to estimate groundwater storage, the seasonal variation in tritium content or tritium peaks of short duration in streamflow in response to heavy rainfall can be neglected, unless the response time to single rainfall events is also of interest. Using a simple exponential model as starting point and reference might provide useful insight before increasing model complexity.

Parameter estimation is performed by minimizing an objective function. Two objective functions are implemented in WITS: Nash-Sutcliffe (NSE, eq. 5) and Mean Prediction Error (MPE, eq. 6 (see 3.4)). For relatively well-defined problems and a good quality tracer dataset (i.e. with a high signal to noise ratio), both objective functions should yield equivalent results.

Note

MPE is in the unit of the tracer used for calibration, thus allowing a direct comparison between the mean calibration error and the analytical measurement error.

## 5.5 Transit time distributions

WITS allows the user to export the **forward** transit time distribution for a chosen simulation date. "Forward" in this context means that the distribution indicates the relative release of an instantaneous tracer injection with increasing transit time. In steady state cases, the transit time distribution does not vary in time, and any time will yield the same distribution. In variable flow cases, the transit time distribution of an instantaneous injection will depend in each time step on the cumulative tracer flux since injection, and may thus yield shapes that differ significantly from the known models at steady state (see Zuber (1986) and Benettin and Bertuzzo (2018) for examples).

## 5.6 Choice of the parameter range

WITS allows the user to define *a priori* for each fitting parameter the range of values in which the best solution will be sought. The user must be aware of two potential problems:

1 If the range is very broad (for instance, a mean transit time between 1 and 100 years), the numerical discretization of the parameter vector might be too coarse to find the absolute best solution. It is thus recommended to explore the parameter space sufficiently finely either by setting the discretisation parameter to a higher value than the default of 10 steps (being aware that computation time might increase substantially) or by dividing the parameter vector into smaller subranges (i.e., MTT=1-10; 11-20;21-30 etc.) and perform multiple runs, one for each subrange.

2 Setting the parameter space arbitrarily to a given range based on a priori expectations could lead to the best solution not being considered if it is outside of the chosen range. It is thus recommended to explore broadly a large parameter space, and then in a second step narrow down the search to the subspaces that are both physically realistic and yielding good fits to the measurements. A clear indication that the best fit might be outside of the chosen parameter range is when the estimated best fit value is equal to the upper or lower boundary of the range.

## 5.7 Plausibility of the obtained parameters

The data used for modelling may be faulty. The model chosen inadequate. Hence, the plausibility of the parameters obtained from fitting should always be assessed critically:

- Long mean transit times means that the tracer storage volume of the catchment is large. If the inflow or outflow rate is known or can be estimated, the storage volume can be calculated using eq. 7. Dividing this volume by the recharge area yield an average equivalent water column height that should be checked against available hydro(geo)logical knowledge (eq. 11). For instance, in an unconfined aquifer a calculated water column of 100 metres will not be plausible if the total thickness is less than this value.
- A good fit to tracer data may sometimes be achieved by using a large piston-flow component to shift the time response of the output, especially for recent tritium data only capturing the tail of the bomb peak. Such a shift might not be plausible if the recharge area is close to the outlet or if the unsaturated zone is too shallow to explain such a large response lag. Both effects can be checked by dividing the piston-flow time by the horizontal or vertical distance between recharge area and outflow location and critically assessing the plausibility of the tracer velocities obtained.

# References


Amin, I., & Campana, M. (1996). A general lumped parameter model for the interpretation of tracer data and transit time calculation in hydrologic systems. *Journal of Hydrology*, 179, 1-21.

Benettin, P., & Bertuzzo, E. (2018). tran-SAS v1.0: a numerical model to compute catchment-scale hydrologic transport using StorAge Selection functions. *Geoscientific Model Development*, 11, 1627-1639.

Benettin, P., van der Velde, Y., van der Zee, S., Rinaldo, A., & Botter, G. (2013). Chloride circulation in a lowland catchment and the formulation of transport by travel time distributions. *Water Resources Research*, 49, 4619-4632.

Botter, G., Bertuzzo, E., & Rinaldo, A. (2011). Catchment residence and travel time distributions: the master equation. *Geophysical Research Letters*, 38 (11).

Farlin, J., Gallé, T., Bayerle, M., & Pittois, D. (2022). On using lumped-parameter models in protection zones to predict aquifer response to reduced agrochemical input. *Journal of Hydrology*, 613.

Grabczak, J., Małoszewski, P., Rozanski, K., & Zuber, A. (1984). Estimation of the tritium input function with the aid of stable isotopes. *Catena*, 11, 105-114.

Haitjema, H. (1995). On the residence time distribution in idealized groundwatersheds. *Journal of Hydrology*, 172 (1-4), 127-146.

Jurgens, B., & Böhlke, J. E. (2012). *TracerLPM (Version 1): An Excel workboot for interpreting groundwater age distributions from environmental tracer data*. USGS 4-F3, 60 pp.

Kirchner, J., Feng, X., & Neal, C. (2001). Catchment-scale advection and dispersion as a mechanism for fractal scaling in strean tracer concentrations. *Journal of Hydrology*, 254, 82-101.

Lenda, A., & Zuber, A. (1970). Tracer dispersion in groundwater experiments. *Isotope Hydrology IAEA*, 619-642.

Małoszewski, P. (1993). Mathematical modelling of tracers experiments in fissured aquifers. *Freiburger Schriften zur Hydrologie*, 114 pages.

Małoszewski, P., & Zuber, A. (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. *Journal of Hydrology*, 57, 207-231.

Małoszewski, P., & Zuber, A. (2002). *Manual on lumped parameter models used for the interpretation of environmental tracer data in groundwaters*. IAEA technical report, 50 pages.

Małoszewski, P., Rauert, W., Stichler, W., & Herrmann, A. (1983). Application of flow models in an alpine catchment area using tritium and deuterium data. *Journal of Hydrology*, 66, 319-330.

Małoszewski, P., Stichler, W., Zuber, A., & Rank, D. (2002). Identifying the flow systems in a karstic-fissured-porous aquifer, the Schneealpe, Austria, by modelling of environmental 18O and 3H isotopes. *Journal of Hydrology*, 256, 48-59.

Morgenstern, U., Daughney, C., Leonard, G., Gordon, D., Donath, F., & Reeves, R. (2015). Using groundwater age and hydrochemistry to understand the sources and dynamices of nutrient contamination through the catchment into Lake Rotorua, New Zealand. *Hydrology and Earth System Sciences*, 19, 803-822.

Sprenger, M., Llorens, P., Gallart, F., Benettin, P., Allen, S., & Latron, J. (2022). Precipitation fate and transport in a Mediterranean catchment through models calibrated on plant an stream water isotope date. *Hydrology and Earth System Sciences*, 26, 4093-4107.

Stewart, M., & Thomas, J. (2008). A conceptual model of flow to the Waikoropupu Springs, NW Nelson, New Zealand, based on hydrometric and tracer evidence. *Hydrology and Earth System Sciences*, 12 (1), 1-19.

Stumpp, C., Stichler, W., & Małoszewski, P. (2009). Application of the environmental isotope delta 18 O to study water flow in unsaturated soils planted with different crops: Case study of a weightable lysimeter from the research field in Neuherberg, Germany. *Journal of Hydrology*, 368, 68-78.

Suckow, A. (2014). The age of groundwater- definitions, models, and why we do not need this term. *Applied Geochemistry*, 50, 222-230.

Viville, D., Ladouche, B., & Bariac, T. (2006). Isotope hydrological study of mean transit time in the granitic Strengbach catchment (Vosges massif, France): application of the FlowPC model with modified input function. *Hydrological Processes*, 20 (8), 1737-1751.

Zuber, A. (1986). On the interpretation of tracer data in variable flow systems. *Journal of Hydrology*, 86, 45-57.

Zuber, A., & Motyka, J. (1994). Matrix porosity as the most important parameter for fissured rocks for solute transport at large scales. *Journal of Hydrology*, 158, 19-46.

Zuber, A., Małoszewski, P., Stichler, W., & Herrmann, A. (1986). *Tracer relations in variable flow.* Proceedings of the 5th intermational symposium on underground water tracing. Institute of geology and mineral exploration, Athens. 355-360.