Reevaluation of transit time distributions, mean transit times and their relation to catchment topography

S. Seeger and M. Weiler

Albert-Ludwigs-University of Freiburg, Germany

Correspondence to: S. Seeger (stefan.seeger@hydrology.uni-freiburg.de)

Abstract

The transit time of water is a fundamental property of catchments, revealing information about the flow pathways, source of water and storage in a single integrated measure. While several studies have investigated the relationship between catchment topography and transit times, there
5 are few studies that expanded the analysis to a wide range of catchments properties and assessed the influence of the selected transfer function model. We used stable water isotopes from mostly baseflow samples with lumped convolution models of time invariant transfer functions to estimate the transit time distributions of 24 meso-scale catchments covering different geomorphic and geologic regions in Switzerland. The sparse network of 13 precipitation isotope
10 sampling sites required the development of a new spatial interpolation method for the monthly isotopic composition of precipitation. A point-energy-balance based snow model was adapted to account for the seasonal water isotope storage in snow dominated catchments. Transit time distributions were estimated with three established transfer functions (exponential, gamma distribution and two parallel linear reservoirs). While the exponential transfer function proved to
15 be less suitable to simulate the isotopic signal in most of the catchments, the gamma distribution and the two parallel linear reservoirs transfer function reached similarly good model fits to the fortnightly observed isotopic compositions in discharge. Albeit, in many catchments the transit time distributions implied by equally well fitted models differed markedly of each other and in extreme cases, the resulting mean transit time differed by orders of magnitude. A more
20 thorough comparison showed that equally suited models corresponded with agreeing values of cumulated transit time distributions only between three and six month. The short term (< 30 days) component of the transit time distributions did not play a role because of the limited temporal resolution of the available input data. The long-term component ($>$ three years) could hardly be assessed by means of stable water isotopes resulting in ambiguous mean
25 transit time and hence questioning the relevance of a mean transit time determined with stable isotopes. Finally we investigated the relation between mean transit time estimates based on the three different transfer function types as well as other transit time properties and a range of topographical catchment characteristics. Depending on the selected transfer model, we found

a weak correlation between transit time properties and the ratio between flow path length over the flow gradient, drainage density and the mean discharge. The catchment storage derived from mean transit times and mean discharge did not show any clear relation to any catchment properties, indicating that in many studies the mean annual discharge may bias the MTT
5 estimates.

1 Introduction

Stable water isotopes or other natural constituents, like chloride, in precipitation act as environmental tracers whose signals are altered by hydrological processes, storage and mixing inside a catchment. Measurements of those environmental tracers in discharge can be used
10 to infer transit (or travel) time distributions (TTDs) and mean transit times (MTTs) on the catchment scale (Kirchner et al., 2010). These inferred TTDs and MTTs might in turn enable a deeper understanding of hydrological processes which cannot be assessed by discharge measurements alone.

Transit time estimations based on lumped convolution modelling approaches have been
15 carried out in various studies, reviewed by McGuire and McDonnell (2006), and subsequent studies like Soulsby and Tetzlaff (2008), Tetzlaff et al. (2009b), Hrachowitz et al. (2010), Roa-García and Weiler (2010), Lyon et al. (2010), Soulsby et al. (2011), Heidbüchel et al. (2012), and Capell et al. (2012). Lumped convolution modelling approaches are based on the convolution of an input signal with a transfer function (TF) to obtain an appropriate output signal. McGuire
20 and McDonnell (2006) pointed out that this widespread modelling approach was originally developed for groundwater systems (Małoszewski and Zuber, 1982) and assumes a hydrological steady state system (Małoszewski et al., 1983) and a determinable representative input.

Botter et al. (2010), Botter et al. (2011) and Rinaldo et al. (2011) developed a mathematical framework for catchment based tracer studies and they reached the conclusion that the steady
25 state assumption generally cannot hold in dynamically responding catchments. Botter et al. (2010) found that the steady state assumption is particularly unsuited to capture a catchment's short term behaviour. Rinaldo et al. (2011) pointed out that the TTD conditional on a specific

3

input time usually is not the same as its counterpart, the TTD conditional on a certain exit time. The input TTD will be continuous, while the exit TTD will be as discrete as the respective input. Despite this clear discrepancy, lumped convolution modelling approaches assume an equivalence of both TDD. This also means that an optimized TF is assumed to be suited to reflect a catchment's TTD, which remains a crude approximation as long as the catchment is not a steady state system with continuous input.

Some more recent studies (Hrachowitz et al., 2010; Heidbüchel et al., 2012, 2013) have abandoned the time-invariant TF approach in favour of convolution models with time-variant TFs. This approach allows for temporal variability of the TTDs, but it also greatly increases the computational cost and includes further assumptions. Another way to allow for time-variant TTDs is to abdicate the convolution approach altogether and to apply a more explicit modelling approaches like van der Velde et al. (2010) or Hrachowitz et al. (2013).

While these more recent approaches seem to be more suited to reproduce a natural catchment's TTDs, they all come at an increased cost. In order to keep the computational cost of the optimisation manageable, Heidbüchel et al. (2012), who estimated time-variant TTDs for two catchments, limited the number of free parameters in their TFs to one. Whereas Hrachowitz et al. (2010) stated that the estimation of time-variant TTDs based on a two parameter gamma distribution TF took about 150 h for one catchment. In a flux tracking approach of Hrachowitz et al. (2013), the size of the multidimensional data matrix required for flux-tracking increases with the square of the time series length and tends to exhaust the commonly available memory capacity rather fast.

So even though the lumped convolution modelling approach with time-invariant TTDs has several shortcomings and is likely to be superseded by more sophisticated modelling approaches in the future, up to date the only practical alternatives to consider a greater amount of catchments without additional assumptions to reduce the number of parameters using commonly available computing resources are time-invariant TTDs. Neither the fitting of sine waves (Małoszewski et al., 1983) nor the computation of damping ratios (Tetzlaff et al., 2009a) are suitable to account for time-variant TTDs. For the time being the lumped convolution approach with time-invariant

4

TTDs will likely stay the method of choice for studies which have another focus than the advancement of transit time estimation methods (e.g. Mueller et al. (2013)).

Several studies were dedicated to the investigation of the potential relationship between catchment properties and mean transit times. McGuire et al. (2005) as well as Tetzlaff et al. (2009b) found a strong correlation between MTTs and the ratio of the median overland flow distance to median flow path gradient (L/G) for two nested catchment studies. Hrachowitz et al. (2009), on the other hand, found no significant correlation between MTTs and L/G. They identified the catchments' proportions of responsive soils and their drainage densities as best predictors of MTTs. Soulsby and Tetzlaff (2008) and Capell et al. (2012) also found good correlations between MTTs and the proportions of responsive soils. Probably due to the small sample size of four catchments, Mueller et al. (2013) found no significant correlation between MTTs and any topographic index, but the highest correlation coefficient of 0.62 was obtained for the drainage density, however, they did not test for a correlation to L/G. In a comparative study Tetzlaff et al. (2009a) used the damping ratio of standard deviations of $\delta^{18}O$ in precipitation and discharge as transit time proxy (TTP) instead of MTTs to investigate catchments of various geomorphic regions across the Northern Hemisphere and also found a strong correlation to L/G. Considering time variant TTDs of zero order catchments, Heidbüchel et al. (2013) found that the relation between MTTs and catchment characteristics is not constant over time and hypothesized that internal catchment states as well as external forcings can alter the dominating factors that influence TTDs.

The objective of this study was to determine TTDs of 24 catchments in Switzerland and to assess the relationship of MTT and other proxies to catchment's topographical indices, with the final aim of finding a topography driven regionalisation method. Another focus of this study was laid on a comparison of the MTT estimates from different TFs to assess the suitability of different TF types. Furthermore, the influence of seasonal snow storage in alpine catchments necessitated the development of a snow module, which accounts for the isotopic composition of snow storage and melt water. The sparse network of precipitation isotope sampling sites required the development of a new spatial interpolation method for the monthly isotopic composition of precipitation.

2 Data

2.1 Study area

This study focused on 24 catchments distributed across the Swiss Plateau and the Swiss Alps (see Fig. 1), selected based on the following criteria: least possible human influence, glaciers

5 covering less than 5 % of the catchment area, possibility for collecting isotope samples and data availability. The catchment area, mean elevation and average annual precipitation is listed for all catchments in Table 1. The mean catchment elevations are between 472 m and 2369 m a.s.l. and their areas range from 0.7 to 351 km². The dominating landcovers within these catchments are elevation dependent, with agricultural areas dominating at lower elevations (< 800 m),

10 grasslands, pastures and forests at mid altitudes (800–1400 m) and grasslands or sparsely vegetated areas at higher elevations > 1700 m. Minor fractions of the catchments *Schaechen* and *Dischmabach* (2 and 5 %, respectively) are glaciated and around 10 % of the catchments *Biber* and *Aabach* are covered with permanent wetlands or open water.

Mean annual catchment precipitation sums range from 1012 to 2600 mm. The seasonal

15 distribution of precipitation is fairly balanced, with 54 to 61 % of annual precipitation occurring during the summer half year. Primarily elevation dependent temperature differences cause a range of discharge regimes from pluvial for the colline and submontane catchments to nival for the more alpine catchments. Different underlying geologies, from crystalline and limestone in the Alps to flysch and molasse in the Swiss Plateau, in connection with varying topographical

20 conditions led to a variety of soils and further differences in discharge behavior among the catchments.

2.2 Discharge data and meteorological data

The Swiss Federal Office for the Environment (FOEN) provided the daily discharge data for most of the catchments. Discharge data for the catchments *Luempenenbach*, *Erlenbach* and

25 *Vogelbach* were obtained from the Swiss Federal Institute for Forest, Snow and Landscape

6

Research (WSL). Additional discharge data for the catchments *Roethebach* and *Emme* were provided by the Amt für Abwasser und Umwelt (AWA) of the Swiss Canton Berne.

The climate data, including average catchment precipitation, temperature, relative air humidity, wind speed and global radiation for 100 m elevation bands in each catchment based on interpolated site data from the national meteorological service of Switzerland (MeteoSwiss) were provided by the PREVAH working group (Viviroli et al., 2009a,b).

2.3 Discharge isotope data

All isotopic compositions in this study are expressed in the δ notation according to the VSMOW-standard. Water samples at the catchment outlets were taken fortnightly from mid 2010 to end 2012. The 100 mL samples were analyzed for stable water isotopes with a PICARRO cavity ringdown spectrometer at the Chair of Hydrology at the University of Freiburg, Germany. According to the manufacturer's specifications the measurement accuracy for δ^{18}O and δ^{2}H is 0.16 and 0.6‰, respectively. Additional discharge isotope data before 2010 for the catchment *Rietholzbach Mosnang* and its subcatchment *Oberer Rietholzbach* was received from the Institute for Atmospheric and Climate Science (IAC) of the Swiss Federal Institute of Technology (ETH), Zurich. Therefore, the available discharge isotope time series for those two catchments extend further into the past, though no discharge isotope samples for the subcatchment *Oberer Rietholzbach* have been taken after February 2010. Due to limited financial and logistic resources, the sampling frequency remained temporally sparse and samples mostly include baseflow conditions.

2.4 Precipitation isotope data

In our study region, the ratio of δ^{18}O to δ^{2}H in precipitation shows no seasonal variation. Therefore, we assume that the δ^{18}O and δ^{2}H data records convey virtually the same information. Lyon et al. (2009) presented a study in a different climatic setting, where this assumption would be untenable. As the data availability was better for δ^{18}O than for δ^{2}H, only δ^{18}O values were considered in our analyses.

The National Network for the Observation of Isotopes in the Water Cycle (NAQUA-ISOT) of the Federal Office for the Environment (FOEN) of Switzerland measures stable water isotopes (δ^{18}O and δ^2H) in the precipitation at monthly intervals at 13 sites. Supplemental data were taken from 5 sites of the Austrian Network of Isotopes in Precipitation (ANIP) and 5 sites of the Global Network of Isotopes in Precipitation (GNIP). Figure 1 shows the positions of these sites. The highest data availability is given for the period between July 1992 and October 2011, where at least for eleven sites monthly values were available.

3 Methods

3.1 Derivation of topographic indices

In order to derive topography based indices for the 24 catchments, a topographic terrain analysis based on a digital elevation model (DEM) with a resolution of 25 m was carried out with the free open source software SAGA-GIS (Conrad et al., 2013). In a first step, the SAGA module "Channel Network" was used to derive the channel network for each catchment. The required drainage area initiation threshold was adapted manually for each catchment to achieve the best agreement between the computed channel networks and the channel networks observed in maps and areal imagery, in our case from Google Maps WMS (Web Map Service) layers.

The SAGA module "Overland Flow Distance to Channel network" was used to calculate the flow path lengths L as well as their respective horizontal and vertical components (L_{h} and L_{v}) for the 24 catchments. Furthermore, the flow gradient G was computed as the ratio $L_{\mathrm{v}}/L_{\mathrm{h}}$. These values were aggregated for each catchment by computing each catchment's median values. Eventually, the ratio L/G was computed for each of the study catchments. Additionally, the topographic wetness indices (TWI) were computed with the module "Topographic Wetness Index" (Böhner and Selige, 2006) and again aggregated by computing their median values for each catchment. Drainage densities (DD) were computed as the ratio of channel length to catchment area.

8

3.2 Spatial interpolation of precipitation isotope data

The isotopic composition of precipitation was not directly measured within the catchments. Instead, the following procedure to interpolate the available site data was applied:

I Based on the δ^{18}O values of the three measurement sites *Meiringen*, *Guttannen* and *Grimsel*, which lie along an elevation transect in the Bernese Alps between 632 and 1950 m a.s.l. (see the bold red line in the map in Fig. 1), average elevation gradients $\overline{g}$ for each month were computed. It was assumed that these gradients are representative for the whole study area.

II Monthly and average monthly δ^{18}O values corrected to the sea level elevation (i_{s} and $\overline{i_{\mathrm{s}}}$) were computed for every measurement site s as follows:

$$i_{\mathrm{s}} = I_{\mathrm{s}} + h_{\mathrm{s}} \cdot \overline{g} \tag{1}$$
$$\overline{i_{\mathrm{s}}} = \overline{I_{\mathrm{s}}} + h_{\mathrm{s}} \cdot \overline{g} \tag{2}$$

where I_{s} is the isotopic composition for measurement site s with the site elevation h_{s} for a certain month and year, while $\overline{I_{\mathrm{s}}}$ is the same, but averaged over all years for each month.

III The average monthly elevation corrected δ^{18}O values $\overline{i_{\mathrm{s}}}$ for all measurement sites were spatially interpolated using kriging (Delhomme, 1978), implemented in the *gstat*-package (Pebesma, 2004) for R. This resulted in continuous maps of average monthly sea level δ^{18}O values for every point p within the study region for each month of the year.

IV To derive the δ^{18}O value for a certain location p at a specific year and month, I_{p}, the following equations were used:

$$d_{\mathrm{s}^*} = i_{\mathrm{s}^*} - \overline{i_{\mathrm{s}^*}} \tag{3}$$
$$i_{\mathrm{p}} = \overline{i_{\mathrm{p}}} - d_{\mathrm{s}^*} \tag{4}$$
$$I_{\mathrm{p}} = i_{\mathrm{p}} - h_{\mathrm{p}} \cdot \overline{g} \tag{5}$$

9

First, the measurement site closest to the location p was chosen, denoted as s*. In Eq. (3), the deviation d_{s*} for a specific month's $\delta^{18}O$ value to its according average monthly value was computed for the measurement site s*. By subtracting this deviation from the average monthly sea level $\delta^{18}O$ value at the location p, obtained from the interpolation in step III, the specific month's sea level $\delta^{18}O$ value at point p was estimated in Eq. (4). Finally h_p, the elevation of the point of interest, was taken into account to obtain the actual $\delta^{18}O$ value I_p at the location p in Eq. (5). Since most measurement sites have data gaps during the investigation period, s* for the same p can refer to different sites for different time steps.

3.3 Transit time proxy

To compare the results obtained by the lumped convolution approach using time-invariant TFs with a more simplistic approach, we also adopted the transit time proxy approach described by Tetzlaff et al. (2009a). Instead of using the inverse transit time proxy used by Tetzlaff et al., we used its reciprocal value, the transit time proxy (TTP) which is computed according to eq. 6:

$$P = \frac{\sigma_{C_p}}{\sigma_{C_Q}} \tag{6}$$

The TTP, denoted as P in eq. 6, is computed as the ratio of the standard deviations of $\delta^{18}O$ values in precipitation (σ_{C_p}) and discharge (σ_{C_Q}). The TTP reflects the precipitation input signal's damping in the discharge and proved to be proportional to MTT estimates (Tetzlaff et al., 2009a). Instead of long time series of climatic input data and stream discharge measurements, this approach only requires time series of the isotopic compositions of precipitation and stream water.

3.4 Model framework

The model framework in this study is based on the TRANSEP-framework (Weiler et al., 2003), without the distinction of event and pre-event water and extended by a snow module

to encounter the specific conditions in alpine catchments. Figure 2 provides an overview on the model structure and the data flow.

3.4.1 Distributed snow modelling

Since several of the selected catchments are notably influenced by snow accumulation and snow
5 melt processes, the implementation of a snow model was crucial. Due to a lack of suitable snow data for the calibration of a simple parametrised snow model and the availability of the appropriate climatic input data, a point-energy-balance based approach was chosen. This study uses a modified implementation of *ESCIMO* (*E*nergy balance *S*now *C*over *I*ntegrated *MO*del by Strasser and Marke, 2010), based on *ESCIMO.spread* and requires hourly input values for air
10 temperature, precipitation amount, wind speed, relative humidity as well as for incoming short- and longwave radiation. To account for the available input data, the following modifications were made:

- change of time step length from hourly to daily (significant snowfall rate of $0.5\,\mathrm{mm\,h^{-1}}$ to reset the albedo to its maximum value was adapted to $2\,\mathrm{mm\,d^{-1}}$)

15 - calculation of incoming longwave radiation with available input data and an empirical relationship given in Sicart and Hock (2010)

Like the original *ESCIMO*, this modified version predicts melt water amounts and sublimation for one point. Under the simplifying assumptions of complete mixing in the snow pack and negligible influence of fractionation processes, further minor modifications like the computation
20 of weighted averages of snow pack and new snow enabled the prediction of average isotopic compositions of the snow pack and hence the melt water. Due to the distinct elevation dependence of snow accumulation and melt processes, it was decided to run the snow module individually for each 100 m elevation band in each catchment. Melt water amounts (including precipitation not retained in the snow pack), sublimation from the snow pack and the isotopic
25 composition of the melt water for all elevation bands of a catchment were then aggregated to calculate the total catchment wide liquid input for the next modelling steps.

3.4.2 Lumped discharge and isotope modelling

Discharge and its isotopic compositions were simulated with two similar lumped convolution models. Both of these models require effective precipitation as their input. The effective precipitation was computed with a rainfall-loss module. While the proposed modelling
5 framework is not bound to any particular method for computing the effective precipitation, we used the approach described by Jakeman and Hornberger (1993), which computes effective precipitation based on a storage index that underlies a decay rate depending on temperature. For further details see Jakeman and Hornberger (1993) or Weiler et al. (2003).

Discharge Q for a certain time step t is described by a convolution of the hydraulic transfer
10 function $h(\tau)$ with all preceding effective precipitation values p_{eff} (Weiler et al., 2003):

$$Q(t) = \int\limits_0^t h(\tau) p_{\text{eff}}(t - \tau) \mathrm{d}\tau \tag{7}$$

The tracer concentration in discharge $C(t)$ is computed in a similar way. Instead of the effective precipitation, the mass weighted isotopic composition of the precipitation, $C_P(t)$, is convoluted
15 by the tracer transfer function, $g(\tau)$ (Stewart and McDonnell, 1991; Weiler et al., 2003; Hrachowitz et al., 2010):

$$C(t) = \frac{\int_0^t g(\tau) p_{\text{eff}}(t - \tau) C_P(t - \tau) \mathrm{d}t}{\int_0^t g(\tau) p_{\text{eff}}(t - \tau) \mathrm{d}\tau} \tag{8}$$

In this time-invariant modelling approach, the optimized tracer TF can be considered to
20 represent the respective catchments TTD.

3.4.3 Transfer functions

Table 2 shows all TFs used in this study: the widely used exponential model (EM), described by Małoszewski and Zuber (1982); the more flexible gamma distribution model (GM), described

12

by Kirchner et al. (2000) and the two parallel linear reservoir (TPLR) model (Weiler et al., 2003). Both, the GM as well as the TPLR, have special cases in which they are equal to the EM.

The discharge convolution module was mainly needed as an auxiliary mean to constrain the parameters of the rainfall-loss module. As initial testing revealed, the TPLR was clearly
5 outperforming the GM and the EM as hydraulic TF and was therefore a priori selected as the sole hydraulic TF $h(\tau)$ of this study.

Regardless whether previous transit time studies mentioned different tracer TFs or not, for catchment comparisons most of them focused on one of them: McGuire et al. (2005) and Mueller et al. (2013) chose the EM; Hrachowitz et al. (2010), Soulsby et al. (2011), Birkel
10 et al. (2012) and Heidbüchel et al. (2012) chose the GM while Roa-García and Weiler (2010) selected the TPLR. An exception is a nested catchment study by Capell et al. (2012), who fitted GM as well as TPLR to eight catchments and considered both model types throughout the analysis of the results. In this study we refrained from an a priori selection of the tracer TF type and chose to optimise our models for each of the three TF types.

15 ## 3.5 Model optimisation and uncertainty

Due to the large amount of optimisations (three models with seven to nine parameters for 24 catchments) Monte Carlo sampling was deemed impracticable for this study. Instead, a multi objective optimisation approach using the NSGA-II algorithm after Deb et al. (2002), implemented in the R-package *mco* by Trautmann et al. (2013) was selected.

20 Three objective functions were applied to evaluate the model: $KGE'(Q)$ and $KGE'(\log(Q))$ were selected to compare the simulated discharge values against the observed values and $KGE^-(C)$ was used to compare the simulated isotopic composition of the discharge against the $\delta^{18}O$ values observed in the discharge. KGE' is the modified Kling-Gupta Efficiency after Gupta et al. (2009) and Kling et al. (2012), which consists of a combination of the correlation
25 coefficient, the ratio of standard deviations and the ratio of mean values. For the evaluation of simulated isotope concentrations, possible biases caused by the spatial interpolation of sparse input data had to be ignored. Therefore a reduced variant of the KGE', called KGE^-, that only takes into account the correlation coefficient and the ratio of standard deviations was applied.

When dealing with multiple separate objective functions, there is no clear best solution, as the improvement of one objective function value can impair the value of another objective function. All combinations of objective function values where this is the case are Pareto optimal. The multi-objective NSGA-II optimisation algorithm (Deb et al., 2002) was run with a population size $N = 1500$ over the curse of $I = 20$ generations, which lead to a total number of $N \times I = 30000$ model runs for each of the 24 catchments and the three TF types. The NSGA-II algorithm returns N parameter sets, but usually only a subset of them are Pareto optimal. In case the first run of the algorithm did not generate at least 300 Pareto optimal parameter sets, the found solutions were remembered and the algorithm was repeated as often as needed to reach at least 300 Pareto optimal parameter sets for each catchment.

Not all of the Pareto optimal parameter sets lead to sensible solutions, as at a certain point minimal improvements in respect to the value of one objective function lead to substantial deterioration of the values of the other objective functions. Similarly to combining three single objective functions into one for the Kling-Gupta Efficiency (Gupta et al., 2009), we used D_0, the euclidean distance to the ideal point (in our case zero), to evaluate the overall goodness of a parameter set:

$$D_0 = \sqrt{(1 - KGE'(Q))^2 + (1 - KGE'(\log(Q)))^2 + (1 - KGE^-(C))^2} \tag{9}$$

In Eq. (9) Q is the discharge amount, C is the isotopic composition in the discharge. KGE' and its reduced variant KGE^- were defined earlier.

The results of the iterative meta-heuristic NSGA-II algorithm are not suited to be used within the established Generalized Likelihood Uncertainty Estimation (Beven and Binley, 1992) method, which would require large numbers of parameter sets obtained by random sampling over the whole parameter value ranges. Therefore another approach to estimate model uncertainty was utilised. All parameter sets with a D_0 smaller than the 10% quantile of all parameter sets' D_0 were considered acceptable. Parameter- and prediction uncertainties were then given by the ranges encompassed by all acceptable parameter sets and their respective simulation results. Most comparisons and analysis presented in this study refer to the median values of all acceptable solutions.

14

To compare the characteristics of TTDs resulting from the three TF types across all catchments, we started by identifying the best TF type for each catchment, i.e. the TF type with the highest median $E_r(C)$ value amongst all acceptable solutions. This set of the best models served as a reference against which the three model types were compared. We compared the models under two aspects: time after which a certain cumulated TTD value is reached and the cumulated TTD value after a certain time. Coefficients of determination as well as the mean ratio of the reference values and the respective values of a specific model were computed.

4 Results

4.1 Spatial interpolation of isotopes in precipitation

Monthly elevation gradients of $\delta^{18}O$, averaged over the time period from mid 1992 to the end of 2011, computed along the three *NAQUA-ISOT* sites *Meiringen*, *Gutannen* and *Grimsel* reached values between $-0.10\,‰$ per $100\,m$ for January and $-0.25\,‰$ per $100\,m$ for September, with an overall mean value of $-0.21\,‰$ per $100\,m$. This is in good agreement with the values reported for the same region by Siegenthaler and Oeschger (1980) and Mueller et al. (2013). The interpolated average monthly $\delta^{18}O$ values at sea level shown in Fig. 3 reveal a seasonal pattern, where $\delta^{18}O$ values at sea level from May to September are higher and far more homogeneous than from October to April. Biggest differences occur from December to March, where $\delta^{18}O$ values at sea level clearly decline in a south-eastern direction. A qualitative validation of the interpolation based predictions can be found in Appendix A.

4.2 Model optimisation and parameter identifiability

For some catchments, the intended number of 300 Pareto optimal solutions was exceeded after the first run and it could easily be increased to 1000, for other catchments the required number of 300 Pareto optimal solutions demanded several repetitions of the optimization algorithm. Consequently, the number of acceptable solutions and the quality of the obtained Pareto fronts varied between the catchments and the TF types, so that the final analyses were based on 30–

100 (10% of 300–1000) parameter sets for each catchment and TF type. The parameters of the rainfall loss module after Jakeman and Hornberger (1993) could hardly be identified – in many cases two of the three parameters spanned over wide ranges of the whole possible value ranges. For the TPLR hydraulic transfer model, τ_f and ϕ could be identified quite well, while the values for τ_s often covered large parts of the possible value range. Unsurprisingly, the EM with only one parameter showed the best parameter identifiability amongst all tracer TF types. Even when the parameters of the rainfall-loss models proved to be unidentifiable, in most cases τ_m of the EM could be constrained to rather narrow ranges. Only for the catchments *Aabach* and *Mentue* τ_m varied by orders of magnitude. The two parameters of the GM generally proved to be identifiable, even though in some cases they had a notable range. As expectable, parameter identifiability for the three parameter TPLR was the lowest. Similarly to the TPLR hydraulic TF, the TPLR tracer TF's τ_f and ϕ tended to be more identifiable than its τ_s.

4.3 Rainfall-discharge model

Independent from the three different tracer TF types, the rainfall-runoff component of the model performed equally satisfactorily for most of the studied catchments, reaching KGE' and $\text{KGE}'_{\log}$ values between 0.7 and 0.9 for most of them (see Fig. 4). Notable exceptions were *Riale di Calneggia*, whose KGE' value of 0.6 was still acceptable but below the values of the other catchments, *Erlenbach* and *Vogelbach* with $\text{KGE}'_{\log}$ values around 0.5 and *Oberer Rietholzbach* with KGE' values below 0.3 and $\text{KGE}'_{\log}$ values around 0.6. Not only the values of the discharge based objective functions, but also the optimised parameter values for the rainfall-runoff component of the model turned out to have the same values, no matter which tracer TF type was part of the multi-objective optimisation. Unsurprisingly, the application of the snow module proved to be essential for the good performance of the rainfall-runoff model for catchments at higher elevations.

16

4.4 Isotopic composition model

4.4.1 Performance

Objective function values for the prediction of isotopic compositions in discharge for the three different TF types are listed in the lower part of Fig. 4, while the left column of Fig. 5 shows
5 the simulated and observed $\delta^{18}O$ values for five catchments, which were selected to represent the range of all observed catchment behaviours. For the four catchments *Guerbe*, *Sitter* (see third row of Fig. 5), *Riale di Calneggia* and *Schaechen*, all models performed similarly well. Comparison of simulated and observed $\delta^{18}O$ values in discharge as well as the objective function values suggest a less satisfactory performance of the EM for the other catchments.
10 Beyond that, it is not possible to name an overall superior TF type: The three parameter TPLR often reached the highest objective function values, but for some catchments the two parameter GM reached higher values. For many catchments the GM and TPLR performed very similarly, even though the simulated $\delta^{18}O$ values in discharge were not the same for the two model types, as the GM tended to produce more short term variability than the TPLR.

15 #### 4.4.2 Prediction bias

Regardless of the applied TF types, all predicted $\delta^{18}O$ time series in discharge were biased in one or the other direction (for some examples see the result of the bias calculation shown in the middle column of Fig. 5). A negative prediction bias means that the predicted $\delta^{18}O$ values in discharge were lower than the respective observed values. These biases were not taken into
20 account for the computation of the respective objective function values. For most catchments, the bias for all three TF types varied within a range of 0.5‰ $\delta^{18}O$. Larger differences between different TF types' bias values were observed for the catchments at higher elevation, with a maximum bias for the catchment *Dischmabach*, where the bias of the TPLR was around $-0.2‰\ \delta^{18}O$, while the biases of the other TF types were distinctly higher and reached 2‰
25 $\delta^{18}O$. An elevation dependent grouping was observed: the 16 catchments at mean elevations up to 1300 m.a.s.l. showed negative biases around $-0.7‰$ (ranging from -0.1 to $-1.3‰$),

17

while seven catchments with higher mean elevations showed more positive biases between -0.2 and $2\permil$. The transition between those two groups is not gradually but abrupt. Being the only catchment south of the Alps, *Riale di Calneggia* with a mean elevation of nearly $2000\,\mathrm{m.a.sl.}$ showed a distinctively negative bias around $-2\permil$ for all three TF types.

4.4.3 Intercomparison of transfer functions

Despite the quite similar performances of the simulations based on TPLR and GM TFs, clear differences of the TTD shapes were observed (Fig. 6). The differences concerning TTD shapes and tailings were reflected by differences in MTT estimates for the different TF types (see Fig. 7).

The MTTs for all TFs agreed only for two catchments: *Schaechen* (MTT of 1.2 years) and *Sitter* (MTTs between 0.7 and 0.9 years). For the other catchments, MTT estimates of different TF types occasionally varied by orders of magnitude (see Table 5). One example is the catchment *Langeten* (see top of Fig. 5): while the EM results in a MTT of 2.3 years, GM and TPLR result in MTTs of 29 and 67.2 years, respectively.

Spearman rank correlation coefficients (ρ) and Pearson correlation coefficients (r) and their respective p values were computed to assess the relationships between the MTTs estimated with the three different TF types as well as the TTP (Fig. 8). Correlations between the EM and TPLR proved to be the lowest ($r = 0.49$ and $\rho = 0.53$), but still significant (both p values < 0.05). For all other combinations the correlations were highly significant with correlation coefficients between 0.6 and 0.8 and p values below 0.005. The TTP correlated with all TF types' MTT estimates and reached highly significant (p values < 0.005) rank correlation coefficients between 0.61 (for TPLR based MTT estimates) and 0.87 (for EM based MTTs estimates). Generally, the Pearson correlation coefficients, which assumes a linear relationship, were smaller than the Spearman rank correlation coefficients.

The comparison of the cumulated TTDs of the three model types (examples for five selected catchments in the right column of Fig. 5) showed that the differences between the model types were greatest towards the longer transit times (>2 years). For some catchments there were also notable differences between different model types towards the shortest transit times (<1 month).

Instead of discussing the cumulated TTD curves for all 24 catchments of each of the three TF types individually, Fig. 9 shows the coefficients of determination and the mean cumulated TTD value ratios between a specific model type and the respective best model for each catchment (as described at the end of Sect. 3.5). Figure 9 shows that for the GM and TPLR the coefficient

5 of determination as well as the mean value ratios reached values close to one around a time of three months. This means that after an elapsed time of around three months these two TF types led to very similar cumulated TTD values. For longer and shorter times, the coefficients of determination declined and the mean value ratios started to diverge from one, which means that the cumulated TTDs were generally less similar and further apart from the respective best

10 cumulated TTD.

4.4.4 Relation between topographic indices and mean transit times

Without discussing all topographic indices (see Table 3) in detail, it seems noteworthy to point out that L, L/G and DD were significantly ($p < 0.05$) correlated with each other. Furthermore, G significantly correlated with TWI and elevation, whereas L/G also significantly correlated

15 with elevation. The higher the catchments, the bigger were the gradients G, the smaller the ratios L/G and the smaller the TWI values. The catchments *Aach*, *Aabach* and (to a lesser degree) *Mentue* showed the lowest values for G and consequently extremely high values for L/G. To remove a distortion of the results caused by a leverage effect, correlations between transit time metrics and topographic indices were computed for all and for all but those three catchments.

20 Even though there were some significant correlations to topographic catchment characteristics when all catchments were considered (table not shown), the picture got much clearer when those three flattest catchments were omitted (Tab. 4).

We observed a high agreement between the cumulated TTD fractions of the first three months (from now on CF3M) for GM and TPLR (see Fig. 9). On the other hand, the TTDs tailings and

25 MTTs varied notably between different models and proved to be less identifiable. Therefore we decided to include CF3M as an apparently more consistent transit time metric than MTT into this analysis. For all model types the transit time metrics CF3M (first section of Tab. 4) as well as MTT (second section of Tab. 4) showed significant ($p < 0.05$) Pearson correlations and

19

Spearman rank correlations to the ratio L/G. Furthermore, there were significant correlations between CF3M values of all TF types and the drainage densities DD and significant correlations to MTT for some TF types.

However, the strongest correlations were found between the transit time metrics (CF3M and MTTs including TTPs) and the mean discharge of the catchments ($\overline{Q}$), which is also correlated to other topographic properties. $\overline{Q}$ primarily is a consequence of external climatic forcings, namely precipitation input and potential evapotranspiration. In order to neutralize the influence of this dominant external forcing, we multiplied MTT estimates with $\overline{Q}$-values to estimate catchment storage volumes. After this step, the correlations to topographic characteristics generally decreased (third section of Tab. 4). The only remaining significant rank correlations to topographic indices were those between EM and GM based MTT estimates and L/G and between EM based MTT estimates and DD.

5 Discussion

5.1 Modelling framework and optimisation procedure

In order to estimate the effective precipitation amounts, the discharge amounts were considered during the multi-objective optimisation procedure. The relatively simple TPLR discharge convolution module managed to predict annual discharge reasonably well for most catchments. As it turned out, the optimised parameters for the rainfall loss module and the discharge convolution module did not depend on the chosen isotopic TF. This suggests, that both of them could have been calibrated before and independently from the isotopic convolution module and only once for all TFs, as done by Weiler et al. (2003) – an approach that reduces the complexity of the optimisations and therefore frees computational resources. Considering the low parameter identifiability of the three parameter rainfall loss module after Jakeman and Hornberger (1993), the use of another rainfall loss module might be advisable.

20

5.2 Applicability of the precipitation isotope interpolation method

Despite the availability of precipitation isotope concentration data being suboptimal (insufficient precipitation isotope data, directly measured only within a few of the study catchments and a sparse measurement network in a region with distinct topography), the

5 interpolation method described in Sect. 3.2 proved to work fairly well. Assuming the observed negative prediction biases for the lower catchments and the positive biases for the higher situated catchments can, for the most part, be explained by systematic errors of other model components (see next section), the interpolation method can be considered suitable for this application (see also Appendix A). More sophisticated interpolation procedures, taking other influence

10 factors such as air temperatures, precipitation amounts, windward-leeward effects and dominant weather situations into account, are conceivable, but to the authors' knowledge up to the present there is no such interpolation method for the given temporal and spatial scales available and its development clearly exceeded the scope of this study.

5.3 Prediction bias of streamwater stable isotopes

15 The convolution model could adequately reproduce the seasonal variations of the isotope concentrations in streamwater, however all predictions exhibited a bias. For most of the catchments, the biases were independent from the applied TF, indicating that the systematic bias was not caused by the choice of TFs. Upon closer inspection, three possible reasons for this bias have to be considered:

20   First, there could be a bias in the precipitation isotopes, caused by incorrect assumptions made during the interpolation of the sparse measurement site data. The resulting biases could be positive or negative and are more likely to occur in regions where the surrounding measurement sites are further apart and the catchment elevations exceed the elevations of the measurement sites.

25   Another error source for the input isotope concentration of alpine catchments could be assumptions made for the snow module. Particularly the assumption of isotopical homogeneous melt from the snow pack without significant enrichment is debatable as Taylor et al. (2001) as

well as Unnikrishna et al. (2002) observed a range of melt water $\delta^{18}O$ values of up to 3‰ around the snow pack's mean isotopic composition. Furthermore, Taylor et al. (2001) measured an overall $\delta^{18}O$ enrichment of around 0.3‰ for the entire melt water amount. While this could explain deviations during the ablation period, it is not sufficient to explain the observed overall bias values of around 1‰ for the alpine catchments, unless the enrichment effect observed in the two aforementioned studies, both of them conducted in the Californian Sierra Nevada, is more pronounced for our study region.

The third possible cause of the prediction biases is inherent to the model, more precisely its rainfall-loss module. Since there is no representation of a soil storage, where winter- and summer precipitation can mix to a certain extent, the simulated evapotranspiration, occurring predominantly during summer, consists almost entirely of the isotopically heavier summer precipitation. On the other hand, nearly all of the isotopically lighter winter precipitation is routed to discharge. While it is likely, that the largest part of the yearly evapotranspiration stems from summer precipitation and that a larger fraction of winter precipitation contributes to discharge, it can be assumed that the missing model representation of a mixing soil storage necessarily leads to a prediction bias towards lighter discharge isotope concentrations. This kind of bias might be prevalent at the non-alpine catchments, where all predictions have a slightly negative bias between 0 and -1‰ $\delta^{18}O$, while no such bias can be recognised when the interpolated precipitation isotope concentrations are compared to the validation site data (see Appendix A) in the same region.

5.4 Temporal scope of the modelling approach

As all simulated values can only be compared to the observed values, the coarse temporal resolution of the isotopic input data (fortnightly data in streamflow and monthly bulk sampled precipitation isotope data) is not suited to evaluate the short-term components of the TTDs. At the same time, the increased dampening of the seasonal variation of the $\delta^{18}O$ signal in precipitation after a few years inevitably leads to a point, where the measurement uncertainties and faster components of the TTD wholly conceal the part of the signal which is caused by the long tailing of the TTD, which in turn also excludes long tailings of a TTD from an objective

22

evaluation using stable water isotopes. However, the ratio between high frequency variations and complete dampening during different times of the years seems to define the fraction between the fast and slow part of the TPLR. Hence, the proportion of relative young water (< 2 years) and much older water (>10 years) can be estimated with a good certainty for most watersheds using the TPLR transfer function.

The inter-model comparison in Fig. 9 suggests that, at least for the available fortnightly stream sample data in combination with the monthly bulk sampled precipitation isotope data, the model optimisation is most sensitive on an intermediate time scale between one month and a year. During these time scales, the estimated cumulated discharge fractions of the more flexible TPLR and GM are very similar. When we compared the TFs to arbitrarily normalised variants of themselves (forcing the cumulated TTD to reach unity after 20 years), it turned out that the latter lead to the exact same simulated isotopic compositions in discharge, even when their TTDs' tailings were distinctly compressed and had notably lower MTTs than their not normalised variants.

This might help to explain the low identifiability of the TPLR model's parameter representing the MTT of the slow reservoir τ_s. The long term tailing of a TTD simply does not matter in respect to an objective function based on natural precipitation's $\delta^{18}O$ in discharge, no matter how long the input data time series is. To asses this part of a catchment's TTD, a tracer with an extended temporal scope, like ^3H, would be required. This was already emphasised by McDonnell et al. (2010), Stewart et al. (2010) and Stewart et al. (2012).

5.5 Meaningfulness of the mean transit time

As mentioned in the previous section, a TTD containing longer transit times cannot be properly assessed solely with a cyclical annually varying environmental tracer like ^{18}O or ^2H. Still, it is possible to fit an arbitrary TF with any kind of long-term tailing to the measured environmental tracer data. A wide range of sufficiently flexible transfer functions should be able to produce acceptable predictions of temporally sparse measurements of $\delta^{18}O$ in discharge. However, this is not enough to ensure an appropriate representation of a TTD's long-term tailing. The long-term tailing of a TTD strongly affects MTT estimates without having any discernible

23

impact on the predicted time series. Thus, reliable MTT estimates are not possible without the consideration of a tracer with extended temporal scope.

Even though the MTT estimates vary between the different model types (see Fig. 7), Fig. 8 indicates that the MTT estimates are not random, as there are significant, yet not very strong, correlations between the MTT estimates based on different TF types. It turns out that in respect to MTT estimates relying solely on stable water isotope data, TTP values seem to be just as good as more complex convolution models: both can be used for a general classification into catchments with short, intermediate and long MTTs, neither can provide sound absolute values for MTT.

Given a sufficiently high measurement frequency, stable water isotope data should be suited to characterise the short term and intermediate part of a catchment's TTD, but it certainly does not contain enough information to determine complete TTDs or actual MTTs of a catchment.

5.6 Relationship between MTT and topography

Despite the distinct differences between different MTT estimates based on different TF types, the results in Table 4 suggest a significant correlation between MTTs (and the TTP) and the ratio L/G for all TFs.

McGuire et al. (2005) also reported a strong correlation between MTTs estimated by the EM and L/G for the *Lookout Creek* catchment and six of its subcatchments in the H. J. Andrews Experimental Forest in the central western Cascades of Oregon, USA. Tetzlaff et al. (2009b) likewise found the strongest correlation between MTTs and L/G for three Scottish catchments and their subcatchments, while the study of Hrachowitz et al. (2009) did not find a significant correlation between MTTs estimated by the GM and L/G for 20 catchments in the Scottish Highlands. Though, according to the method description in Hrachowitz et al. (2009) the stream network for all of the 20 Scottish catchments was computed with a fixed stream initiation threshold. At least for our study area, in some cases a fixed stream initiation threshold area caused large discrepancies between the computed and the observed channel networks and consequently led to different values for L as well as G. Therefore it cannot be excluded that

24

Hrachowitz et al. (2009) found no significant correlation between MTTs and L/G because they worked with values for L and G which were derived with fixed stream initiation thresholds.

However, in this study most of the observed correlations were only significant as long as the external climatic forcing was not taken into account. The correlation between MTTs and mean annual discharge was higher than for any of the topographical indices. For two hypothetical catchments, which share identical properties regarding geology, topography, soils and vegetation, the catchment with the higher effective precipitation would undoubtedly have higher turnover rates and hence lower MTTs. Consequently, a catchment's MTT actually always will be determined by two components: external forcings (precipitation and potential evapotranspiration) and catchment internal properties. When the aim of a study is the assessment of the influence of catchment properties (like topography) on MTTs, it would appear that to be essential to take external forcings into account. Yet, many studies (e.g. McGuire et al., 2005; Tetzlaff et al., 2009a,b; Hrachowitz et al., 2009; Soulsby et al., 2011; Mueller et al., 2013) did not account for this and directly compared MTTs of catchments with differing mean annual precipitation or discharge amounts. This practice is likely to, at least partially, obscure the true influence of the (non-climatic) catchment properties.

Troch et al. (2013) found strong evidence for a general co-evolution of catchment properties and climatic influences. When climate as well as catchment properties determine MTTs, but at the same time there is a relation between climate and catchment properties which leads to collinearity between many of the catchment properties, it gets near to impossible to identify catchment properties that actually control MTTs independently from the climatic influences, unless there is a possibility to compare different catchments that underlie identical climatic forcings.

Together with the aforementioned issue, the uncertainties connected to the determination of MTTs (Which is the most appropriate TF? Is the time-invariant TF approach suited at all? How can the TTDs tailing properly assessed?), will lead to high degrees of uncertainty for any approach to regionalise MTTs.

On top of that, Heidbüchel et al. (2013) showed that depending on varying external forcings and internal catchment states MTTs can be highly variable. This means that the outcome of any

catchment TTD comparison study is likely to be influenced by time-variant climatic conditions prior to and during the time when the catchments were studied. The reliability of the results might be impaired, when the observational time series do not cover representative periods. Furthermore, Heidbüchel et al. (2013) showed that in some years topographical characteristics

5 might be a good predictor, while in other years, with different external forcings, other factors, like soil characteristics or underlying geology, have greater influence on the observed TTDs. Consequently, when TTDs are considered as time-invariant, it is possible to miss temporally critical relations.

6 Conclusions

10 In this study, we used three different TF types in a time-invariant lumped convolution modelling approach to estimate the TTDs and MTTs of 24 meso-scale catchments in Switzerland on the basis of $\delta^{18}O$ data. We showed, that different TF types could be used to reach similarly acceptable fits to fortnightly sampled $\delta^{18}O$ data in discharge. A comparison of the cumulated TTDs of those equally well performing models indicated that their cumulated values tended to

15 agree at an intermediate time scale between three months and one year, while they diverged on shorter and even more so on longer time scales. From a certain point on, differences in TTD tailings did not influence the predicted $\delta^{18}O$ values in discharge at all. Hence, to properly assess a catchment's TTD on all time scales, a higher sampling frequency of precipitation and discharge would be needed for more information on the catchment's short term behaviour and

20 a more persistent tracer is required to determine the catchment's actual long term behaviour.

The poorly identifiable tailings of the TTDs greatly influenced MTT estimates, which partially exhibited high uncertainties. For catchments with longer MTTs, different model types' MTT estimates could differ by orders of magnitude while the available data was not suited to determine the most appropriate model type. In many cases the EM proved to be less appropriate

25 than the more flexible GM and the TPLR. Given the fact, that the easily computable TTP values showed significant rank correlations to MTT estimates of all TF types, they might serve as

a coequal replacement for them, as long as the latter are as underdetermined as in this study and only relative differences among the catchments are the focus.

The results of this study suggest that seemingly good correlations between MTTs and the ratio of median flow path lengths over median flow path gradients L/G are mainly caused by the mean annual discharge, which considerably influences these topographic indices as well as the MTTs. In order to assess the actual influence of topographic indices on MTTs in catchment comparison studies, the dominant influence of climate should be considered and removed.

Appendix A

Validation of the interpolated precipitation isotope data

A1 Origin of the validation data

Within the frame project of this study, bulk precipitation samples have been taken to determine the isotopic composition of the precipitation at five sites in Central Switzerland. With lengths of not more than one year and limited spatial coverage, these time series were of little use as model input data. Three of those sites, *Benglen*, *Schallenberg* and *Aeschau* have been chosen to validate the interpolated precipitation isotope data.

Further isotope composition data was thankfully obtained from Mueller et al. (2013), who collected precipitation bulk samples for the summer half years of 2010 and 2011 for four small alpine catchments in the Ursern Valley in southern Central Switzerland. Data from the two sites *Bonegg* and *Laubgaedem* were included into the validation data to extend the their elevation range up to 1720 m.

The Institute for Atmospheric and Climate Science (IAC) of the Swiss Federal Institute of Technology Zurich maintains the field measurement site *Messtelle Büel* within the catchment *Rietholzbach* for which fortnightly bulk sample data for $\delta^{18}O$ from 1994 until the beginning of the year 2010 were available.

27

A2 Reasons for the qualitative validation

The method described in Sect. 3.2 was not only applied to obtain precipitation isotope compositions for the studied catchments, but also for all available validation sites. Unfortunately, the temporal resolutions of the monthly interpolation derived predictions and the sub-monthly observed $\delta^{18}O$ time series were not the same. To aggregate isotope composition data to a coarser time scale, mass weighted averaging would be required, but the respective precipitation amounts to the bulk sample isotope data were not available. Hence, a quantitative validation of the interpolation based predictions was not possible, instead a qualitative comparison was made.

A3 Comparison of predictions and validation data

Figure A1 shows the monthly predicted $\delta^{18}O$ values obtained by the interpolation procedure described in Sect. 3.2 plotted with the on-site measured validation data. All validation time series have been collected over shorter periods than one month and thus exhibit more variance and higher amplitudes than the monthly predictions. Nevertheless, a qualitative comparison of predicted and validation data indicates a reasonably good performance of the interpolation method.

References

Beven, K. J. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, DOI:10.1002/hyp.3360060305, 1992.

Birkel, C., Soulsby, C., Tetzlaff, D., Dunn, S. M., and Spezia, L.: High-frequency storm event isotope sampling reveals time-variant transit time distributions and influence of diurnal cycles, Hydrol. Process., 26, 308–316, DOI:10.1002/hyp.8210, 2012.

Böhner, J. and Selige, T.: Spatial prediction of soil attributes using terrain analysis and climate regionalisation, Göttinger Geographische Abhandlungen, 115, 13–28, available at: http://www.saga-gis.org/en/about/references.html, last access: 12 February 2014, 2006.

Botter, G., Bertuzzo, E., and Rinaldo, A.: Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, Water Resour. Res., 46(3), DOI:10.1029/2009WR008371, 2010.

Botter, G., Bertuzzo, E., and Rinaldo, A.: Catchment residence and travel time distributions: The master equation, Geophys, Res. Lett., 38(11), DOI:10.1029/2011GL047666, 2011.

Capell, R., Tetzlaff, D., Hartley, A. J., and Soulsby, C.: Linking metrics of hydrological function and transit times to landscape controls in a heterogeneous mesoscale catchment, Hydrol. Process., 26, 405–420, DOI:10.1002/hyp.8139, 2012.

Conrad, O., Wichmann, V., Olaya, V., and Ringeler, A.: SAGA GIS (version 2.1), http://www.saga-gis.org, last access: 6 May 2014, 2013.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation, 6, 182–197, DOI:10.1109/4235.996017, 2002.

Delhomme, J. P.: Kriging in the hydrosciences, Adv. Water Resour., 1, 251–266, DOI:10.1016/0309-1708(78)90039-8, 1978.

DeWalle, D. and Edwards, P.: Seasonal isotope hydrology of three Appalachian forest catchments, Hydrol. Process., 11, 1895–1906, 1997.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, J. Hydrol., 377, 80–91, DOI:10.1016/j.jhydrol.2009.08.003, 2009.

Heidbüchel, I., Troch, P. A., Lyon, S. W., and Weiler, M.: The master transit time distribution of variable flow systems, Water Resour. Res., 48, W06520, DOI:10.1029/2011WR011293, 2012.

Heidbüchel, I., Troch, P. A., and Lyon, S.W.: Separating physical and meteorological controls of variable transit times in zero-order catchments, Water Resour. Res. 49(11):7644–7657, DOI:10.1002/2012WR013149, 2013.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., Dawson, J. J. C., and Malcolm, I. A.: Regionalization of transit time estimates in montane catchments by integrating landscape controls, Water Resour. Res., 45, W05421, DOI:10.1029/2008WR007496, 2009.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., Malcolm, I. A., and Schoups, G.: Gamma distribution models for transit time estimation in catchments: physical interpretation of parameters and implications for time-variant transit time assessment, Water Resour. Res., 46, W10536, DOI:10.1029/2010WR009148, 2010.

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

Jakeman, A. J. and Hornberger, G.: How much complexity is warranted in a rainfall-runoff model?, Water Resour. Res., 29, 2637–2649, DOI:10.1029/93WR00877, 1993.

Kirchner, J. W., Feng, X., and Neal, C.: Fractal stream chemistry and its implications for contaminant transport in catchments, Nature, 403, 524–7, DOI:10.1038/35000537, 2000.

Kirchner, J. W, Tetzlaff, D., and Soulsby, C.: Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments, Hydrol. Process., 24, 1631–1645, DOI:10.1002/hyp.7676, 2010.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424-425, 264–277, DOI:10.1016/j.jhydrol.2012.01.011, 2012.

Lyon, S. W., Desilets, S. L. E., and Troch, P. A.: A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus δ 18 O, Hydrol. Process., 23, 2095–2101, DOI:10.1002/hyp.7326, 2009.

Lyon, S. W., Laudon, H., Seibert, J., Mörth, M., Tetzlaff, D., and Bishop, K. H.: Controls on snowmelt water mean transit times in northern boreal catchments, Hydrol. Process., 24, 1672–1684, DOI:10.1002/hyp.7577, 2010.

Małoszewski, P. and Zuber, A.: Determining the turnover time of groundwater systems with the aid of environmental tracers, J. Hydrol., 57, 207–231, 1982.

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

Maloszewski, P.: Program flow PC version 2-3-September, GSF-Institut für Hydrologie D-85764 Neuherberg b. München, Fed. Rep. Germany, 1996.

McDonnell, J. J., McGuire, K. J., Aggarwal, P. K., Beven, K. J., Biondi, D., Destouni, G., Dunn, S. M., James, A., Kirchner, J. W., Kraft, P., Lyon, S. W., Maloszewski, P., Newman, B., Pfister, L., Rinaldo, A., Rodhe, A., Sayama, T., Seibert, J., Solomon, K., Soulsby, C., Stewart, M. K., Tetzlaff, D., Tobin, C., Troch, P. A., Weiler, M., Western, A., Wörman, A., and Wrede, S.: How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis, Hydrol. Process., 24, 1745–1754, DOI:10.1002/hyp.7796, 2010.

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

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., and Seibert, J.: The role of topography on catchment-scale water residence time, Water Resour. Res., 41, W05002, DOI:10.1029/2004WR003657, 2005.

Mueller, M. H., Weingartner, R., and Alewell, C.: Importance of vegetation, topography and flow paths for water transit times of base flow in alpine headwater catchments, Hydrol. Earth Syst. Sci., 17, 1661–1679, DOI:10.5194/hess-17-1661-2013, 2013.

Pebesma, E. J.: Multivariable geostatistics in S: the gstat package, Comput. Geosci., 30, 683–691, DOI: 10.1016/j.cageo.2004.03.012, 2004.

Rinaldo, A., Beven, K. J., Bertuzzo, E., Nicotina, L., Davies, J., Fiori, A., Russo, D., and Botter, G.:Catchment travel time distributions and water flow in soils, Water Resour. Res., 47, DOI:10.1029/2011WR010478, 2011.

Roa-García, M. C. and Weiler, M.: Integrated response and transit time distributions of watersheds by combining hydrograph separation and long-term transit time modeling, Hydrol. Earth Syst. Sci., 14, 1537–1549, DOI:10.5194/hess-14-1537-2010, 2010.

Sicart, J. and Hock, R.: Sky longwave radiation on tropical Andean glaciers: parameterization and sensitivity to atmospheric variables, J. Glaciol., 56, 854–860, DOI:10.3189/002214310794457182, 2010.

Siegenthaler, U. and Oeschger, H.: Correlation of ^{18}O in precipitation with temperature and altitude, Nature, 285, 317, DOI:10.1038/285314a0, 1980.

Soulsby, C. and Tetzlaff, D.: Towards simple approaches for mean residence time estimation in ungauged basins using tracers and soil distributions, J. Hydrol., 363, 60–74, DOI:10.1016/j.jhydrol.2008.10.001, 2008.

31

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

Stewart, M. K. and McDonnell, J. J.: Modeling base flow soil water residence times from deuterium concentrations, Water Resour. Res., 27, 2681–2693, DOI:10.1029/91WR01569, 1991.

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

Stewart, M. K., Morgenstern, U., McDonnell, J. J., and Pfister, L.: The "hidden streamflow" challenge in catchment hydrology: a call to action for stream water transit time analysis, Hydrol. Process., 26, 2061–2066, DOI:10.1002/hyp.9262, 2012.

Strasser, U. and Marke, T.: *ESCIMO.spread* – a spreadsheet-based point snow surface energy balance model to calculate hourly snow water equivalent and melt rates for historical and changing climate conditions, Geosci. Model Dev., 3, 643–652, DOI:10.5194/gmd-3-643-2010, 2010.

Taylor, S., Feng, X., Kirchner, J. W., Osterhuber, R., Klaue, B., and Renshaw, C. E.: Isotopic evolution of a seasonal snowpack and its melt, Water Resour. Res., 37, 759–769, DOI:10.1029/2000WR900341, 2001.

Tetzlaff, D., Seibert, J., McGuire, K. J., Laudon, H., Burns, D. A., Dunn, S. M., and Soulsby, C.: How does landscape structure influence catchment transit time across different geomorphic provinces?, Hydrol. Process., 23, 945–953, DOI:10.1002/hyp.7240, 2009a.

Tetzlaff, D., Seibert, J., and Soulsby, C.: Inter-catchment comparison to assess the influence of topography and soils on catchment transit times in a geomorphic province; the Cairngorm mountains, Scotland, Hydrol. Process., 1886, 1874–1886, DOI:10.1002/hyp.7318, 2009b.

Trautmann, H., Steuer, D., and Mersmann O.: mco: Multi criteria optimization algorithms and related functions, http://cran.r-project.org/package=mco, 2013.

Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T., and Sawicz, T.: Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution, Hydrol. Earth Syst. Sci., 17(6):2209–2217, DOI:10.5194/hess-17-2209-2013, 2013.

Unnikrishna, P., McDonnell, J. J., and Kendall, C.: Isotope variations in a Sierra Nevada snowpack and their relation to meltwater, J. Hydrol., 260, 38–57, DOI:10.1016/S0022-1694(01)00596-0, 2002.

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

Viviroli, D., Zappa, M., Gurtz, J., and Weingartner, R.: An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools, Environ. Modell. Softw., 24, 1209–1222, DOI:10.1016/j.envsoft.2009.04.001, 2009a.

Viviroli, D., Zappa, M., Schwanbeck, J., Gurtz, J., and Weingartner, R.: Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part I: Modelling framework and calibration results, J. Hydrol., 377, 191–207, DOI:10.1016/j.jhydrol.2009.08.023, 2009b.

Weiler, M., McGlynn, B. L., McGuire, K. J., McDonnell. J. J.: How does rainfall become runoff? A combined tracer and runoff transfer function approach, Water Resour. Res., 39, 1315, DOI:10.1029/2003WR002331, 2003.

Table 1. Areas, elevations, and mean annual precipitation sums of the 24 studied catchments.

catchment name	gauging station	catchment ID	area [km²]	mean elev. [m]	min elev. [m]	max elev. [m]	prcp [mm a⁻¹]
Dischmabach	Davos	DIS	43.2	2369	1663	3139	1391
Ova da Cluozza	Zernez	OVA	26.9	2364	1519	3160	1053
Riale di Calneggia	Cavergno	RIA	23.9	1986	881	2908	2104
Allenbach	Adelboden	ALL	28.8	1852	1293	2742	1651
Schaechen	Buerglen	SCH	107.9	1719	487	3260	1687
Sitter	Appenzell	SIT	88.2	1301	768	2500	1870
Biber	Biberbrugg	BIB	31.6	999	827	1495	1639
Alp	Einsiedeln	ALP	46.5	1154	845	1894	2112
Luempenenbach	–	ALP_L	0.9	1336	1092	1508	2615
Erlenbach	–	ALP_E	0.7	1359	1117	1650	2168
Vogelbach	–	ALP_V	1.6	1335	1038	1540	2161
Sense	Thoerishaus	SEN	351.2	1068	554	2184	1270
Ilfis	Langnau	ILF	187.9	1037	681	2087	1450
Emme	Eggiwil	EMM	127	1285	743	2216	1559
Roethebach	Eggiwil	ROE	54.1	991	731	1542	1099
Guerbe	Burgistein	GUE	55.4	1037	556	2152	1241
Mentue	Yvonand	MEN	105.0	679	447	926	1060
Langeten	Huttwil	LAN	60.3	760	598	1100	1195
Aach	Salmsach	AAC	50.0	472	408	560	1095
Ergolz	Liestal	ERG	261.2	584	305	1165	1012
Aabach	Moenchaltorf	AAB	55.6	635	519	1092	1081
Murg	Waengi	MUR	76.8	648	467	1036	1281
Rietholzbach	Mosnang	RIE	3.2	794	671	938	1555
Oberer Rietholzbach	–	RIE_O	0.9	815	748	938	1670

34

Table 2. Overview of transfer functions with specification of the parameters and analytical mean transit time (MTT).

Transfer function	Parameters	analytical MTT
Linear reservoir (EM) $g(\tau) = \frac{1}{\tau_\mathrm{m}} \exp\left(\frac{-\tau}{\tau_\mathrm{m}}\right)$	τ_m mean transit time	τ_m
Gamma Distribution (GM) $g(\tau) = \frac{\tau^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp(-\tau/\beta)$	α shape parameter β scale parameter	$\alpha\beta$
Two parallel linear reservoirs (TPLR) $h(\tau) = g(\tau) = \frac{\phi}{\tau_\mathrm{f}} \exp\left(-\frac{\tau}{\tau_\mathrm{f}}\right) + \frac{1-\phi}{\tau_\mathrm{s}} \exp\left(-\frac{\tau}{\tau_\mathrm{s}}\right)$	ϕ fraction of fast reservoir τ_f MTT of fast reservoir τ_s MTT of slow reservoir	$\phi\tau_\mathrm{f} + (1-\phi)\tau_\mathrm{s}$

Table 3. Results of the topographic analysis. L is the flowpath length, G the flow gradient, DD the drainage density and TWI the topographic wetness index.

catchment ID	L [m]	G [m m^{-1}]	L/G [m]	DD [km km^{-2}]	TWI [–]
DIS	647	0.33	1961	1.14	9.52
OVA	616	0.46	1339	0.91	8.87
RIA	647	0.46	1407	1.17	9.10
ALL	423	0.31	1365	1.53	9.27
SCH	646	0.38	1700	1.08	9.38
SIT	329	0.27	1219	2.14	9.48
BIB	207	0.16	1294	2.80	9.96
ALP	196	0.21	933	3.72	9.69
ALP_L	155	0.17	912	4.52	9.61
ALP_E	169	0.20	845	4.75	9.67
ALP_V	193	0.28	689	3.30	9.22
SEN	227	0.20	1135	2.63	9.76
ILF	157	0.30	523	3.54	9.00
EMM	286	0.27	1059	2.18	9.43
ROE	210	0.18	1167	2.34	9.67
GUE	258	0.19	1358	3.06	9.88
MEN	364	0.08	4550	1.32	10.83
LAN	308	0.11	2800	1.40	9.85
AAC	481	0.02	24 050	1.17	11.67
ERG	421	0.15	2807	1.05	9.99
AAB	407	0.04	10 175	1.49	10.92
MUR	219	0.10	2190	2.29	10.07
RIE	194	0.18	1078	2.59	9.51
RIE_O	254	0.15	1693	1.86	9.46

Table 4. Pearson correlation coefficients (r) and Spearman rank correlation coefficients (ρ) between catchment characteristics and different transit time metrics of different models. Significant correlations (p-value < 0.05) are printed in boldface. The shown correlations were computed for 21 of the 24 catchments (the three flattest catchments *Mentue* (MEN), *AAbach* (AAB) and *Aach* (AAC) were omitted).

model	area r	ρ	elevation r	ρ	L r	ρ	G r	ρ	L/G r	ρ	DD r	ρ	TWI r	ρ	$\overline{Q}$ r	ρ
							cumulated fraction of TTD after three months									
EM	-0.32	-0.34	0.12	**0.5**	-0.29	-0.28	0.06	0.38	**-0.53**	**-0.63**	**0.61**	**0.48**	-0.03	-0.19	**0.66**	**0.69**
GM	-0.26	-0.27	0.07	0.34	-0.35	-0.43	0	0.23	**-0.55**	**-0.71**	**0.68**	**0.62**	0.05	-0.07	**0.65**	**0.59**
TPLR	-0.24	-0.26	0.13	0.36	-0.28	-0.4	0.09	0.24	**-0.59**	**-0.7**	**0.63**	**0.6**	-0.02	-0.08	**0.66**	**0.61**
							MTT									
EM	0.25	0.35	-0.3	**-0.5**	0.11	0.29	-0.3	-0.38	**0.63**	**0.63**	**-0.5**	**-0.48**	0.18	0.19	**-0.68**	**-0.69**
GM	0.23	0.33	-0.42	**-0.51**	-0.01	0.12	**-0.46**	-0.42	**0.77**	**0.48**	-0.39	-0.38	0.39	0.22	**-0.63**	**-0.73**
TPLR	0.07	0.06	0.12	-0.14	0.21	0.23	-0.13	-0.22	**0.52**	**0.46**	-0.35	**-0.44**	0.13	0.13	-0.26	**-0.46**
TTP	0.06	0.29	-0.25	-0.38	0.1	0.34	-0.31	-0.24	**0.7**	**0.64**	-0.44	**-0.52**	0.22	0.14	-0.54	**-0.66**
							MTT $\times \overline{Q}$ = catchment storage volume									
EM	-0.08	0.02	0.16	0	0.33	0.31	0.07	0.03	0.34	**0.45**	-0.44	**-0.44**	-0.19	-0.25	-0.1	-0.13
GM	0.07	0.18	-0.23	-0.37	0.05	0.16	-0.36	-0.34	**0.67**	0.49	-0.38	-0.41	0.24	0.08	**-0.48**	**-0.5**
TPLR	-0.21	-0.14	0.33	0.04	0.1	0.07	0.01	-0.12	0.09	0.22	0	-0.25	-0.08	-0.04	0.29	-0.2
TTP	-0.3	-0.23	0.21	0.12	0.3	0.14	0.03	0.06	0.42	0.3	-0.36	-0.28	-0.14	-0.31	0.04	0.15+

Table 5. Mininum, median and maximum MTT in years for each catchment and transfer function type and the respective transit time proxy (TTP) values.

| catchment	GM min	GM med	GM max	TPLR min	TPLR med	TPLR max	EM min	EM med	EM max	TTP
DIS	9.16	10.5	12	84.5	96.7	102	1.42	1.51	1.54	10.16
OVA	3.32	4.81	8.46	32.1	67.7	94.9	0.87	0.91	0.94	6.13
RIA	0.43	0.6	0.75	0.65	1.79	72.7	0.45	0.57	0.68	3.77
ALL	4.33	6.25	8.65	22.3	70.5	108	1.53	1.58	1.65	8.90
SCH	1.11	1.21	1.25	1.11	1.2	1.27	1.11	1.2	1.26	6.76
SIT	0.69	0.92	1.7	0.79	0.9	68	0.61	0.68	0.74	4.08
BIB	0.52	0.76	1.23	0.78	8.49	71.6	0.3	0.36	0.4	3.97
ALP	0.35	0.46	0.57	0.58	0.8	35	0.24	0.28	0.32	3.49
ALP L	1.7	1.91	2.24	18.5	65.7	80.8	0.52	0.55	0.59	3.22
ALP E	0.14	0.21	0.35	0.52	0.57	1.51	0.15	0.17	0.18	2.52
ALP V	0.66	1.02	1.48	1.47	47.3	78.8	0.18	0.27	0.35	4.02
SEN	2.36	5.82	17.6	3.13	39	84	1.25	1.34	1.43	5.31
ILF	5.31	8.88	20.7	3.62	12.1	88.6	1.49	1.55	1.75	8.52
EMM	1.14	1.66	2.77	3.18	12.3	71.5	0.39	0.42	0.48	3.38
ROE	6.5	12.2	25.6	5.31	59.5	108	0.29	1.77	2.45	9.77
GUE	1.04	1.33	2.52	1.43	5.55	75.4	0.84	1.04	1.15	5.02
MEN	12.9	18.2	20.7	23.7	69.3	105	0.02	0.86	1.84	5.01
LAN	20.4	29	31.8	21.6	67.2	122	2.17	2.31	2.36	23.29
AAC	1.05	1.6	2.96	1.72	19.6	98.7	0.76	1.03	1.1	4.35
ERG	7.42	18.8	29	18.2	81.4	114	1.45	1.63	1.86	8.68
AAB	11.7	18.5	21.6	4.28	17.2	105	0.01	0.09	2.36	5.71
MUR	12.4	20.5	26.4	8.79	74	112	1.71	1.79	1.82	10.75
RIE	2.45	6.44	14.3	4.05	15.8	68.9	0.79	1.34	2.02	7.28
RIE O	2.02	8.44	26.4	3.1	21.9	100	0.91	2.02	4.04	8.71
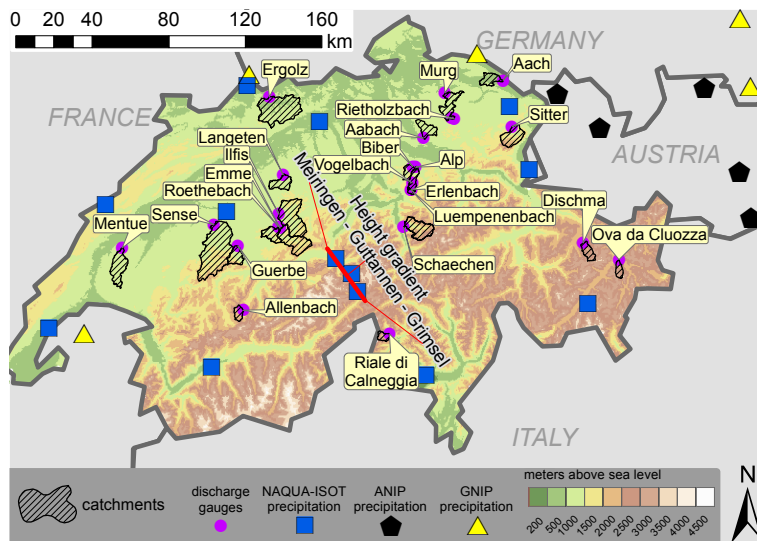
Fig. 1. Map of the study area with elevation and catchment borders. The not shown catchment *Oberer Rietholzbach* is a subcatchment of the *Rietholzbach*-catchment. The symbols indicate positions of isotope measurement sites of various sources.
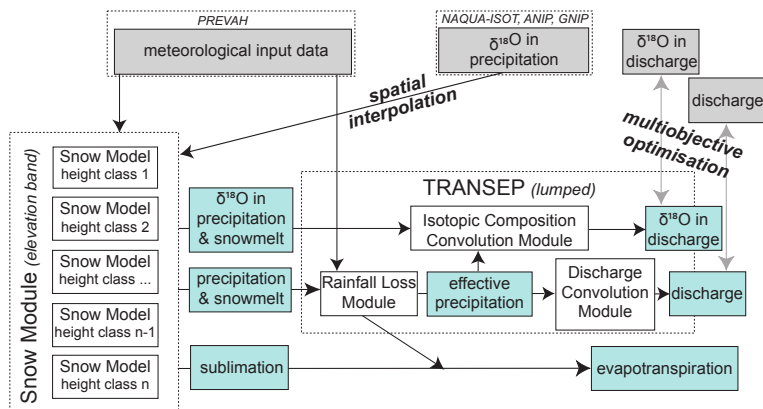
Fig. 2. Overview scheme of the model modules. Grey boxes represent input data, blue boxes represent data computed by model modules (white boxes).
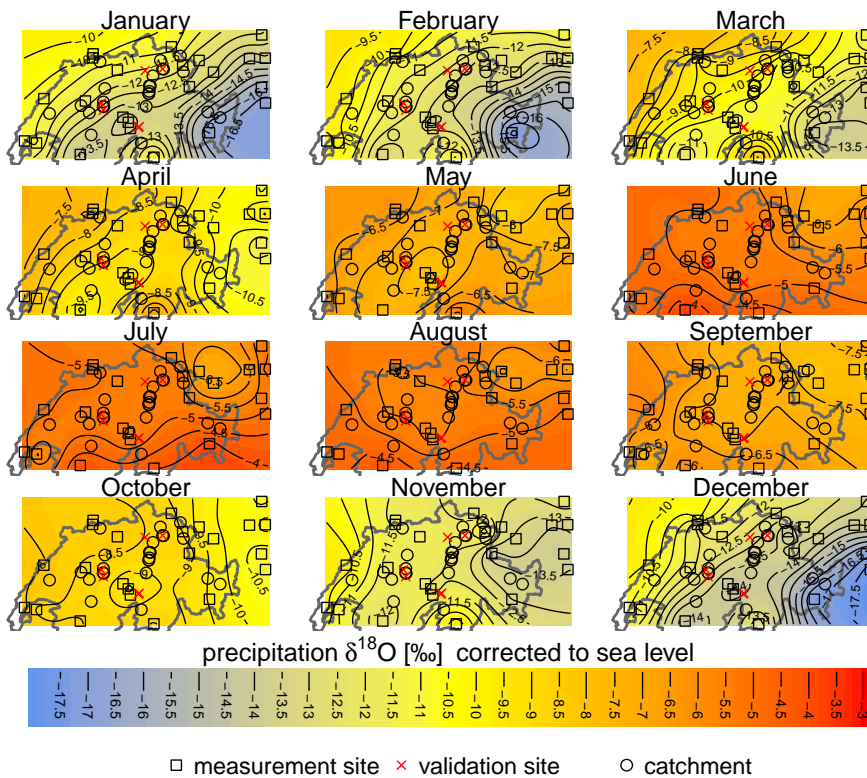
precipitation δ^{18}O [‰] corrected to sea level

□ measurement site   × validation site   ○ catchment

Fig. 3. Monthly maps of interpolated sea level precipitation δ^{18}O values.
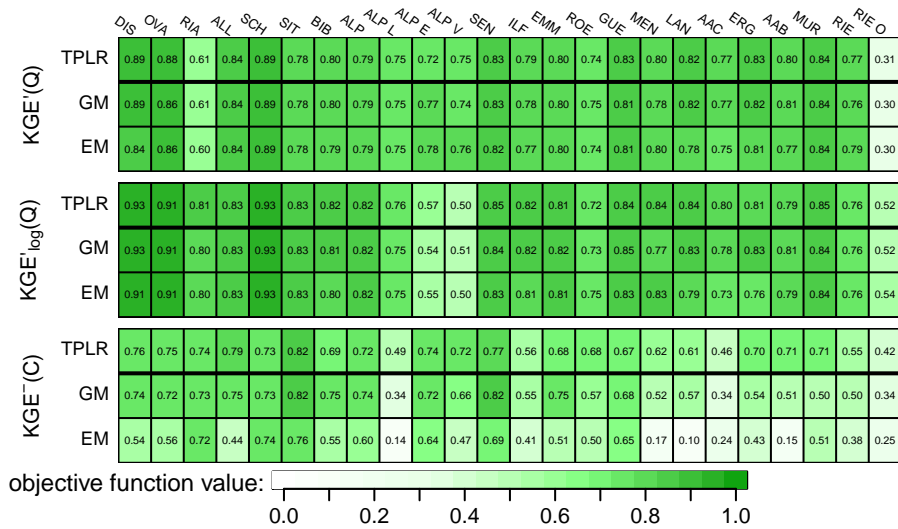
Fig. 4. Values of the three objective functions for all catchments for the three different transfer function types.
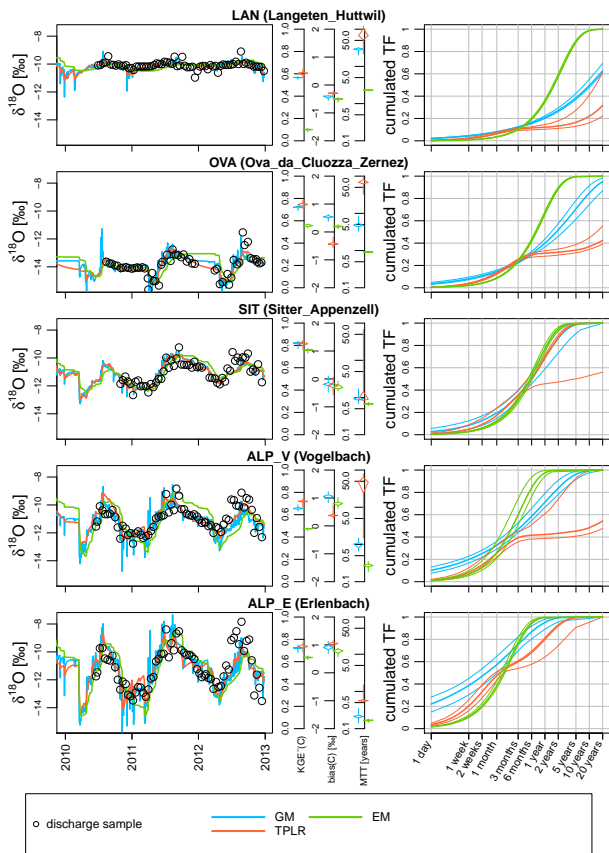
Fig. 5. Optimisation results for selected catchments. Left: observed and predicted isotope concentrations in discharge. Right: cumulated TTDs (thicker lines represent the median values of all accepted solutions, thinner lines indicate their range). Centre: objective function values for isotopic composition predictions, biases of the predictions and MTTs implied by the optimised TFs; lines indicate the full value range, diamonds indicate the 25, 50 and 75 percentiles of the accepted solutions.
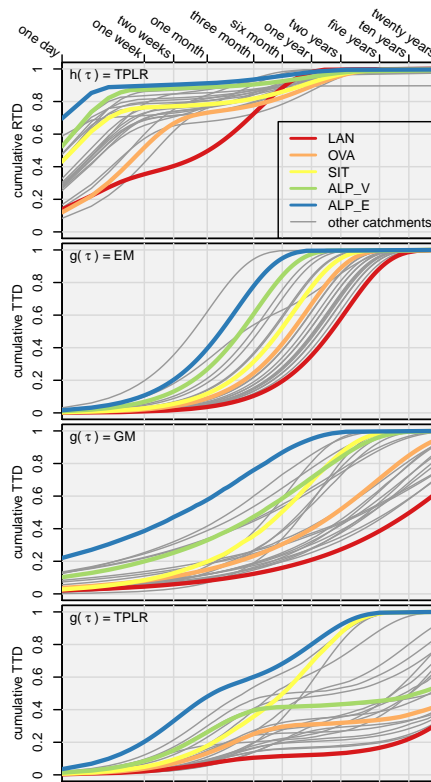
Fig. 6. Top: cumulated response time distributions (RTD) from the rainfall-discharge module; lower three subfigures: cumulated TTDs obtained by the three TF types. The shown curves are the median values of all (i.e. 30–100) accepted solutions.
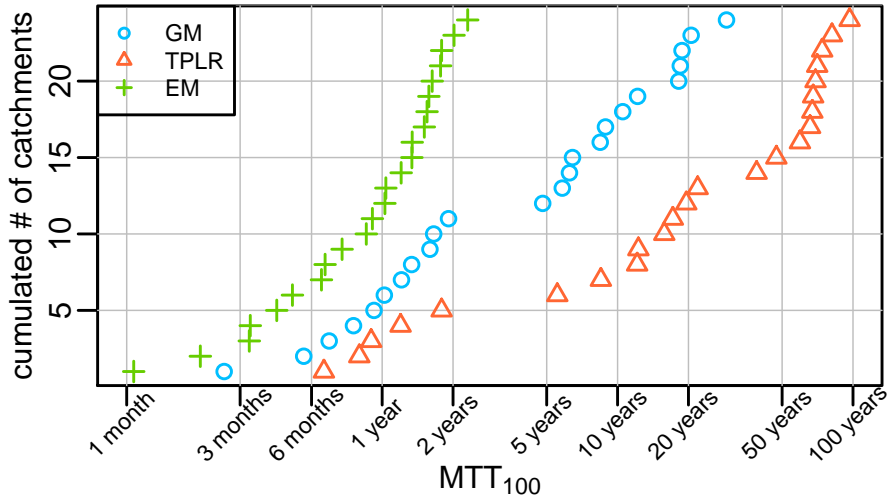
Fig. 7. Cumulative distribution of catchment MTT estimates based on the three TF types.
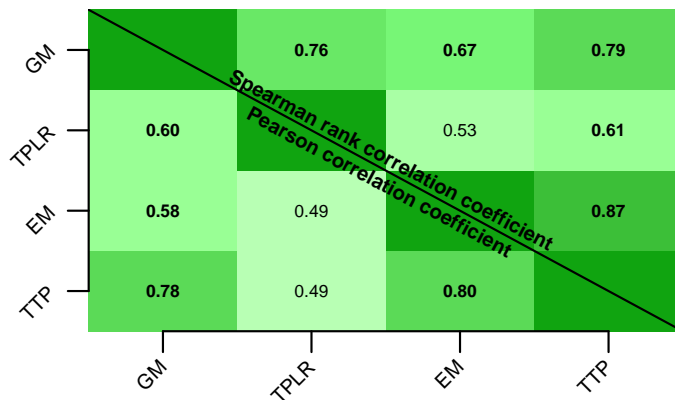
Fig. 8. Combined matrix of Pearson correlation coefficients (lower left) and Spearman rank correlation coefficients (upper right) for MTTs of all catchments derived by the three different TF types and the TTP. All correlations were significant (p values < 0.05), correlation coefficients with p values < 0.005 are printed in boldface.
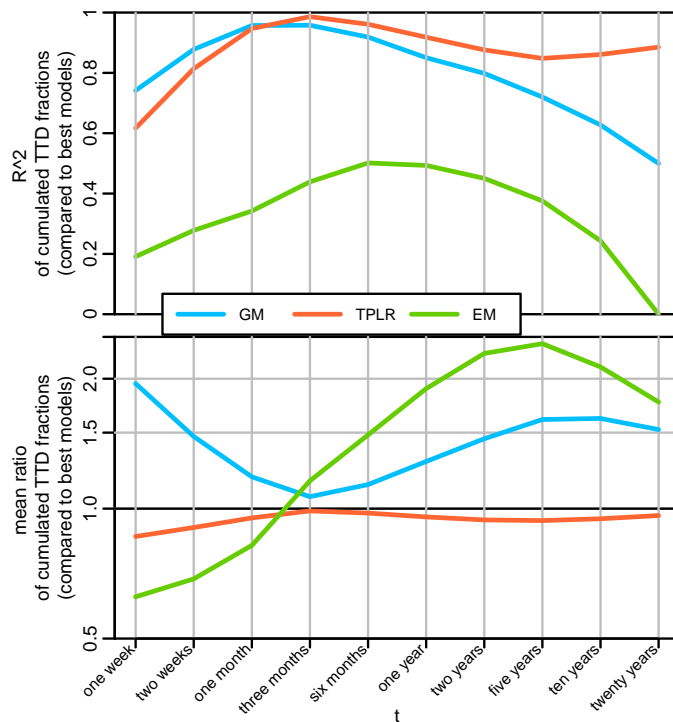
Fig. 9. Comparison of cumulated discharge fractions after certain elapsed times. Top: correlation coefficients between TTDs of specific TFs and a selection of the best TTDs for each catchment. Bottom: mean value ratios between specific and selected best TFs.
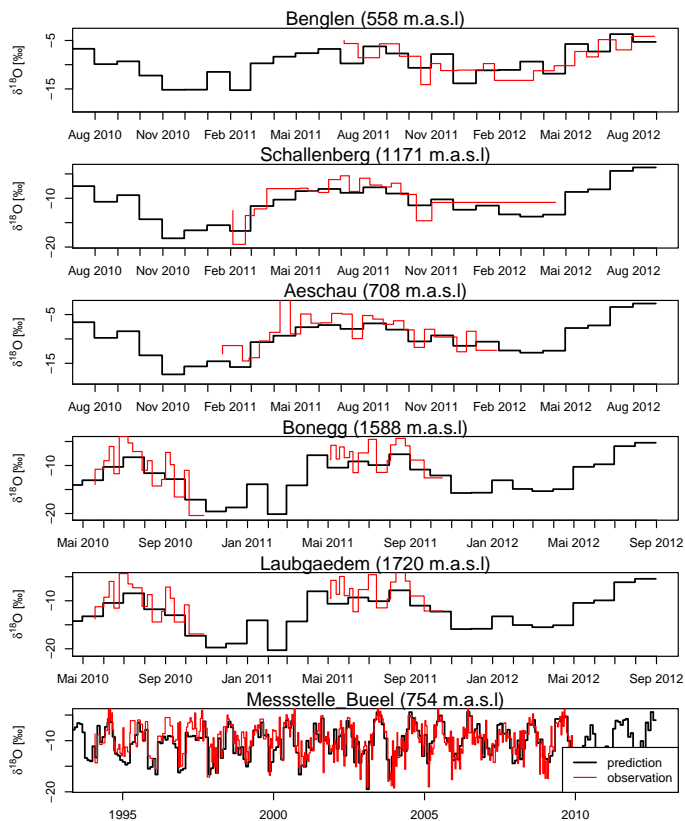
Fig. A1. Comparison between measured $\delta^{18}O$ values (red lines) in precipitation and values obtained by the spatial interpolation method (black lines.)