Understanding uncertainties when inferring mean transit times trough tracer based lumped parameter models in Andean tropical montane cloud forest catchments

E. Timbe[1,2], D. Windhorst[2], P. Crespo[1,3], H.-G. Frede[2], J. Feyen[1], L. Breuer[2]

[1]{Departamento de Recursos Hídricos y Ciencias Ambientales, Universidad de Cuenca, Cuenca, Ecuador}

[2]{Institute for Landscape Ecology and Resources Management (ILR), Research Centre for Bio Systems, Land Use and Nutrition (IFZ), Justus-Liebig-Universität Gießen, Germany}

[3]{Facultad de Ciencias Agropecuarias, Universidad de Cuenca, Cuenca, Ecuador}

Correspondence to: Edison Timbe (edison_timbe@yahoo.com)

Abstract

Weekly samples from surface waters, springs, soil water and rainfall were collected in a 76.9 km^2 mountain rain forest catchment and its tributaries in southern Ecuador. Time series of the stable water isotopes $\delta^{18}O$ and δ^2H were used to calculate mean transit times (MTTs) and the transit time distribution functions (TTDs) solving the convolution method for seven lumped parameter models. For each model setup, the Generalized Likelihood Uncertainty Estimation (GLUE) methodology was applied to find the best predictions, behavioral solutions and parameter identifiability. For the study basin, TTDs based on model types such as the Linear-Piston Flow for soil waters and the Exponential-Piston Flow for surface waters and springs performed better than more versatile equations such as the Gamma and the Two Parallel Linear Reservoirs. Notwithstanding both approaches yielded a better goodness of fit for most sites, but with considerable larger uncertainty shown by GLUE. Among the tested models, corresponding results were obtained for soil waters with short MTTs (ranging from 2 to 12 weeks). For waters with longer MTTs differences were found, suggesting that for those cases the MTT should be based at least on an intercomparison of several models. Under dominant baseflow conditions long MTTs for stream water $\geq$ 2 yr were detected, a phenomenon also observed for shallow springs. Short MTTs for water in the top soil layer indicate a rapid exchange of surface waters with deeper soil horizons. Differences in travel times between soils suggest that there is evidence of a land use effect on flow generation.

1 Introduction

The mean transit time (MTT) of waters provides a valuable primary description of the hydrological (Fenicia et al., 2010) and biochemical systems (Wolock et al., 1997) of a catchment and its sensitivity to anthropogenic factors (Landon et al., 2000; Turner et al., 2006; Tetzlaff et al., 2007; Darracq et al., 2010). Whereas the MTT describes the average time it takes for any given water parcel to leave the catchment, the transit time distribution function (TTD) describes the retention behavior of all those water parcels as a frequency function over time (McGuire and McDonnell, 2006). Together with the physical characteristics of the catchment, the MTT and TTD (for the particular case of soil water, MTT should be more properly understood as Mean Residence Time, and TTD as Residence Time Distribution function) allow inferring the recharge of aquifers (Rose et al., 1996), the bulk water velocities through its compartments (Rinaldo et al., 2011), and the interpretation of the water chemistry (Maher, 2011); all of which supports the design of prevention, control, remediation and restoration techniques. Additionally, MTT and TTD data are useful to reduce the uncertainty of results and improve input parameter identifiability for either hydrologic modeling studies (Weiler et al., 2003; Vache and McDonnell, 2006; McGuire et al., 2007; Capell et al., 2012) or solute movement analyses through soil and aquifers using mixing models (Iorgulescu et al., 2007; Barthold et al., 2010).

The stable water isotopes $\delta^{18}O$ and $\delta^{2}H$ are commonly used as environmental tracers for a preliminary assessment of the transport of water in watersheds with transit times less than 5 yr (Soulsby et al., 2000; Rodgers et al., 2005; Viville et al., 2006; Soulsby et al., 2009). For longer MTTs of up to 200 yr (Stewart et al., 2010), tritium radioisotopes are used to analyze the storage and flow behavior in surface water and shallow groundwater systems (Kendall and McDonnell, 1998), while, for example, carbon isotopes are employed for analyzing the dynamics of deep groundwater with ages of hundreds to thousands of years (Leibundgut et al., 2009).

Since Barnes and Bonell (1996), researchers in tracer hydrology use quasi distributed and conceptual models to encompass the non-linearity of the processes related to the transit states of the soil moisture dynamics (Botter et al., 2010; Fenicia et al., 2010). However, the use of such modeling approaches is only advisable after basic inferences about the underlying mixing processes and the way water is routed through the system have been drawn. Insights that can be provided by applying simpler lumped TTD functions as the models proposed by Maloszewski and Zuber (1982, 1993), which are based on quasi-linearity and steady state conditions. These models include the exponential (EM), piston (PM), or linear (LM) models, in which the MTT of the tracer is the only unknown variable, and also combinations of models such as the exponential-piston flow (EPM) and the linear-piston flow (LPM) models. Among the two-parameter lumped models, the dispersion model (DM), that considers simplifications of the general advection-dispersion equation, has been applied in environmental tracer studies (Maloszewski et al., 2006; Viville et al., 2006; Kabeya et al., 2006). Since almost one and a half decades ago, other lumped models are being exploited such as the two parameter Gamma model (GM) proposed by Kirchner et al. (2000), which is a more general and flexible version of the exponential model; and the Two Parallel Linear Reservoirs model (TPLR), a three-parameter function that combines two parallel reservoirs, each one represented by a single exponential distribution (Weiler et al., 2003). The use of these models for estimating the MTT in the compartments of a catchment has become a standard practice for the preliminary assessment of the catchment functioning. The advantage of the latter functions relies on that they allow the representation of different mixing processes in different system components, such as soil and groundwater. In contrast, simpler models assume instantaneous and complete mixing over the entire model domain (Hrachowitz et al., 2013). Regarding to lumped parameter models, McGuire and McDonnell (2006) presented in their study a compilation of the most frequently used models for deriving MTTs. Under the condition that a particular model ought to be concordant with the physical characteristics of the aquifer system, this condition hinders the applicability of lumped parameter models to poor gauged catchments with scarce or no information on the physical characteristics of the system. For these cases the authors believe that it is better to use an ensemble of models in order to be certain that the results or the inferences point in the same direction, or if not, to have a better idea of the uncertainties.

Particular for tropical zones the knowledge of hydrological functioning is still limited and investigation of system descriptors such as TTD and MTT are keys to improve our understanding of catchment responses (Murphy and Bowman, 2012; Brehm et al., 2008). This is especially the case for tropical mountain rainforest systems. In this study we focus on the San Francisco river basin, a mesoscale headwater catchment of the Amazon in Ecuador. Notwithstanding the recent characterization of the climate (Bendix et al., 2006), soils (Wilcke et al., 2002), water chemistry (Buecker et al., 2011) and hydrology (Plesca et al., 2012) of the basin, we are still lacking a perceptual model that explains the observations of chemical, hydrometric and isotopic variables and related processes (Crespo et al., 2012).

To enhance the understanding of the hydrological functioning of the San Francisco basin, this study focuses on the (i) estimation of the MTT in the different compartments of the catchment; (ii) characterization of the dominant TTD functions; and (iii) evaluation of the performance and uncertainty of the models used to derive the MTTs and TTDs. Translated into hypotheses the study reported in this paper aimed to test if

1) the diversity of the sampling sites allows evaluating the spatial variability in catchment hydrology, identifying the dominant processes, and screening the performance of the TTD models;

2) the multi-model approach and the identifiability of their parameters enable identification of the respective TTDs and MTTs.

The hypotheses are based in the following assumptions:

1) the used tracers are conservative, there are no stagnant flows in the system, and the tracer mean transit time τ represents the MTT of water (e.g. McGuire and McDonnell, 2006);
2) stationary conditions are dominant in the basin and lumped equations based on linear or quasi-linear behaviors are applicable (Heidbüchel et al., 2012);
3) from insights derived of related studies (Soulsby et al., 2010; McGuire and McDonnell, 2006; Rodgers et al., 2005), considering the drainage areas, the steepness of the topography and the shallow depth of the soil layers, the transit times of the sampling sites are less than 5 yr, making it possible to use δ^2H and δ^{18}O as tracers.

2 MATERIALS AND METHODS

2.1 Study area

The San Francisco tropical mountain cloud forest catchment (Fig. 1, Table 1), 76.9 km^2 in size, is located in the foothills of the Andean cordillera in South Ecuador, between Loja and Zamora, and drains into the Amazonian river system. Hourly meteorological data recorded at the Estación Científica San Francisco (ECSF, 1,957 m a.s.l.), El Tiro (2,825 m a.s.l.), Antenas (3,150 m a.s.l.) and TS1 (2,660 m a.s.l.) climate stations are available from the DFG funded Research Unit FOR816 (www.tropicalmountainforest.org). Monthly averages of the main meteorological parameters for the period 1998-2012 allow a description of their spatial and interannual variation. Mean annual temperature ranges from 15°C in the lower part of the study area (1,957 m a.s.l.) to 10°C on the ridge (3,150 m a.s.l.), with an altitude gradient of -0.57°C per 100 m, without marked monthly variability. The wind velocities of the prevailing south-easterlies reach average maximum daily values of 10 m s^{-1} between June and September, while wind velocities in the middle and lower catchment areas are fairly constant, equal to 1 m s^{-1}. The humid regime of the catchment is comparatively constant with the relative humidity varying between 84.5% in the lower parts to 95.5% at the ridges. Among all meteorological parameters, precipitation shows the largest spatial variability, with an average gradient of 220 mm per 100 m (Bendix et al., 2008b). However, this gradient is not constant throughout the catchment and shows substantial spatial variability (Breuer et al., 2013). Recent estimation of horizontal rainfall revealed its significance, contributing 5 to 35% of measured tipping bucket rainfall, respectively to the lower and ridge areas of the catchment

(Rollenbeck et al., 2011). Rainfall is marked by low rainfall intensities, generally less than 10 mm h^{-1} and high spatial variability. Annual rainfall is uni-modal distributed with a peak in the period April-June. Using the Thiessen method and considering horizontal rainfall, the precipitation depth amounted 2,321 mm in the period August 2010-July 2011, and 2,505 mm in the period August 2011-July 2012. A more detailed description of the weather and climate of the study area is given in Bendix et al. (2008a).

In line with findings of Crespo et al. (2012) in the same area, baseflow accounts for 85% of the total runoff (Table 1), notwithstanding the rapid and marked response of flows to extreme rainfall events. In just a few hours peak discharges are several times higher than baseflows (Fig. 2a), carrying considerable amounts of sediment and accompanied by drastic changes in some of the cross sections.

Major soil types are Histosols associated with Stagnasols, Cambisols and Regosols, while Umbrisols and Leptosols are present to a lesser degree (Liess et al., 2009). The geology is reasonable similar throughout the study area, consisting of sedimentary and metamorphic Paleozoic rocks of the Chiguinda unit with contacts to the Zamora batholith (Beck et al., 2008). The topography is characterized by steep valleys with an average slope of 63%, situated in the altitudinal range of 1,725 to 3,150 m a.s.l. (Table 1). Protected by the Podocarpus National Park, the southern part of the catchment is covered by pristine primary forest and sub-páramo. In the northern part, particular during the last two decades, land is being converted to grassland. Presently 68% of the catchment is covered by forest, 20% is sub-páramo, 6.5% is used as pasture and 3% is degraded grassland covered with shrubs (Goettlicher et al., 2009; Plesca et al., 2012). Landslides are present in the catchment, especially along the paved road between the cities Loja and Zamora.

2.2 Catchment composition and discharge measurements

The San Francisco catchment was subdivided into seven sub-catchments with areas ranging between 0.7 and 34.9 km^2, characterized by different land uses varying from pristine forest and sub-páramo to pasture areas (Fig. 1 and Table 1). In order to define baseflow conditions, each sub-catchment was equipped with a water level sensor (mini-diver, Schlumberger Water Services, Delft, NL). Reference discharge measurement, using the salt dilution method, where made frequently during the time of sampling. However, due to the high variability of the river bed for the sites QP, QZ and QR, only continuous records for sub-catchments FH, QN, QM, QC, and for the main outlet PL were considered as reliable to calculate stage-discharge curves and the hydrographs, as shown in Fig. 2a for PL. For the remaining sites, discharge measured at the moment of sampling was used.

2.3 Isotope sampling and analyses

Weekly water samples for isotope analysis were collected manually in the main river (Fig. 2b), its tributaries, creeks and springs in the period August 2010 to mid-August 2012 and later for soil water starting in September/November 2010 (Table 2), using 2 mL amber glass bottles. Soil water sampling was performed along two altitudinal transects covered by forest and pasture (Table 2), in 6 sites (Fig. 1) and 3 depths (0.10, 0.25 and 0.40 m) using wick-samplers. Wick-samplers were designed and installed as described by Mertens et al. (2007). Woven and braided 3/8 fiberglass wicks (Amatex Co. Norristown, PA, US) were unraveled over a length of 0.75 m and spread over a 0.30 m × 0.30 m × 0.01 m square plastic plate. The plate enveloped with fiberglass was covered with fine soil particles of the parent material and then set in contact with the undisturbed soil, respectively at the bottom of the organic horizon (0.10 m below surface), a transition horizon (0.25 m below surface) and a lower mineral horizon (0.40 m below surface). The low constant tension in the wick-samplers guarantees sampling of the mobile phase of soil water, avoiding isotope fractionation (Landon et al., 1999).

Along with the weekly sampling, event based rainfall samples for isotope analyses were collected manually in 1 L bottles using a Ø 25 cm funnel at 1900 m a.s.l. (Fig. 1). After every event, the sample bottles were covered with a lid and stored for analysis within a week in 2 mL amber glass bottles. Only sample volumes > 2 mL were suitable for permanent storage and measurements. Events with a sample volume below 2 mL were discarded. The end of a single rainfall event was marked by a time span of 30 min without rainfall, whereby a total of 946 samples were collected with an average duration of 3.2 h (varying from 0.25 to 19 h with up to 11 events per day). Since the solving of the convolution equation needs a continuous time step of input data (Maloszewski and Zuber, 1982), the time resolution of the input series was set to 7 days (Fig. 2c). In this sense, weekly mean isotopic signatures for smaller rainfall events during longer dry periods (only 5 among 104 weeks had no rainfall event > 2 mL sampling volume) were interpolated using antecedent and precedent measurements.

The final isotope signature used for the models represents:

- for rainfall water, the weighted mean of all events during each week (Sundays to Saturdays) using the rainfall data recorded at the nearby meteorological station (400 m to ECSF),
- for soil water samples, the weekly average isotope signal for each soil depth, and
- for stream, creek and spring water samples, an instantaneous isotopic concentration in time. These samples were not flux-weighted. For stream waters, only isotope samples from designated baseflow conditions were later considered (see Section 2.5).

The stable isotopes signatures of $\delta^{18}O$ and $\delta^{2}H$ are reported in per mil relative to the Vienna Standard Mean Ocean Water (VSMOW) (Craig, 1961). The water isotopic analyzes were performed using a compact wavelength-scanned cavity ring down spectroscopy based isotope analyzer (WS-CRDS) with a precision of 0.1 per mil for $\delta^{18}O$ and 0.5 for $\delta^{2}H$ (Picarro L1102-i, CA, US).

2.4 Isotopic gradient of rainfall

Throughout the catchment, the recorded rainfall time series from meteorological stations are correlated (r^2 was at least 0.6, based on weekly precipitation data). As the models in question are only driven by the isotope signal and not by the actual amount of incoming precipitation on site, a flux weighting based on a single station within the catchment (ECSF) was sufficient. Given the large altitudinal gradient in the San Francisco basin, it is to be expected that the input isotopic signal of rainfall for every sub-catchment varies according to its elevation (Dansgaard, 1964). In this regard, Windhorst et al. (2013) estimated this variation for the main transect of the catchment: -0.22‰ $\delta^{18}O$, -1.12‰ $\delta^{2}H$ and 0.6‰ deuterium excess per 100 m elevation gain. Applying this altitude gradient to the flux weighted isotope signal under the assumption that the incoming rainfall signal is the sole source of water, thereby excluding any unlikely source of water from outside the topographic catchment boundaries with a different isotope signal, it was possible to derive the recharge elevation and localized input signal in each sub-catchment. The derived recharge elevations were used to crosscheck that they are inside the topographic boundaries of every sub-catchment and comparable to their mean elevations.

The justification to adopt only the mentioned gradient to extrapolate the isotope signals, was based in previous studies on spatial and temporal variation of stable isotopes of rainfall in the same catchment, which revealed that, only the altitude effect is significant and that in this factor there is no influence of temperature, relative humidity and precipitation amount or intensity (Windhorst et al, 2013).

Since no marked fractionation was observed for all analyzed waters it is highly probable that similar estimations of MTT are derived using either $\delta^{18}O$ or $\delta^{2}H$ (Fig. 3). Therefore, in this study $\delta^{18}O$ was selected for further analysis.

2.5 Mean Transit Time estimation and Transit Time Distribution

Mean transit times were calculated based on stationary conditions. In the case of stream water this condition was fulfilled by considering only baseflow conditions (Heidbüchel et al., 2012), which were dominant in the catchment during the 2 yr observation period (Figs. 2a and 2b depict this characteristic for the main outlet), accounting for 85% of total runoff volume. Baseflow separations for streamflow were obtained through parameter fitting to the slope of the recessions in the observed hourly flows using the Water Engineering Time Series PROcessing tool (WETSPRO), developed by Willems (2009). To account for samples taken at baseflow conditions in sites where hydrometric records were not available, the specific discharges of the closer catchments with similar characteristics in terms of land use, size, and observed hydrologic behavior were used. In this sense, QZ, QR and QP were considered similar to QN, QM and QC (Table 1). In contrast, all spring and creek water samples were included in the analysis since their isotopic signatures were less influenced by particular rain events (as inferred from the smooth shape of the observed isotope signal) in the San Francisco catchment. In regard to soil water, we considered all samples, since each sample represents a volume weighted weekly average signature (isotopic signatures of particular high rainfall events are smoothed at a weekly time span).

For the calculation of MTTs, the authors used the lumped parameter approach. In this, the aquifer system is treated as an integral unit and the flow pattern is assumed to be constant as outlined in Maloszewski and Zuber (1982) for the special case of constant tracer concentration in time-invariant systems. In this case the transport of a tracer through a catchment is expressed mathematically by the convolution integral. The tracer output $C_{out}(t)$ and input $C_{in}(t)$ are related as function of time:

$$C_{out}(t) = \int_{-\infty}^{t} C_{in}(t') \exp[-\lambda(t-t')]\, g(t-t')dt' \qquad (1)$$

In the convolution integral, the stream outflow composition C_{out} at a time t (time of exit) consists of a tracer C_{in} that falls uniformly on the catchment in a previous time step t' (time of entry), C_{in} becomes lagged according to its transit time distribution $g(t-t')$; the factor $exp[-\lambda(t-t')]$ is used to correct for decay if a radioactive tracer is used (λ = tracer's radioactive decay constant). For stable tracers (λ = 0), and considering that the time span $t-t'$ is the tracer's transit time τ, Eq. (1) can be simplified and re-expressed as:

$$C_{out}(t) = \int_0^\infty C_{in}(t-\tau)g(\tau)d\tau \quad (2)$$

where the weighting function $g(\tau)$ or tracer's transit time distribution (TTD), describes the normalized distribution function of the tracer injected instantaneously over an entire area (McGuire and McDonnell, 2006). As it is hard to obtain this function by experimental means, the most common way to apply this lumped approach is to adopt a theoretical distribution function that better fits to the studied system. In general meaning, any type of a weighting function is understood as a model. In accordance, seven lumped parameter models to infer the MTTs for diverse water storages (stream, springs, creeks and soil water) were applied in this study. Results were evaluated on the basis of the best matches to a predefined objective function, their magnitude of uncertainty and the number of observations in the range of behavioral solutions. The equations for each of the lumped parameter models used are shown in Table 3. EM and LM reflect simpler transitions where the tracer's mean transit time τ is the only unknown variable. More flexible models consider a mixture of two different types of distribution. EPM includes piston and exponential flows, while the LPM accounts for piston and linear flows. In both cases the equations are integrated by the parameter η indicating the percentage contribution of each flow type distribution. The DM, derived from the general equation of advection-dispersion, is also one of the common models used in hydrologic systems (Maloszewski et al., 2006). In this model the fitting parameter D_p is related to the transport process of the tracer (Kabeya et al., 2006). In the GM, the product of the two shape parameters α and β equals τ. This method was successfully applied by Dunn et al. (2010) and Hrachowitz et al. (2010). The TPLR model (Weiler et al., 2003) is based on the parallel combination of two single exponential reservoirs (despite of its name TPLR follows exponential and not linear assumption), representing fast τ_f and slow flows τ_s, respectively. The flow partition between the two reservoirs is denoted by the parameter ϕ.

2.6 Convolution equation resolution

Due to the similarities between the seasonal isotopic fluctuations of the sampled effluents and rainfall signal, a constant interannual recharge of the aquifers was assumed. For each sampling site, the 2 yr isotopic data series were used as input for the models. To get stable results between two consecutive periods, these input isotope time series were repeated 20 times in a loop; an approach similar to the methodology presented by Munoz-Villers and McDonnell (2012) resulting in an artificial time series of 40 yr. It is common practice to extend the time series artificially by duplicating it (Hrachowitz et al., 2010 and 2011). This does not change the results; it rather gives the model more room to find stable results. Data of the last loop were considered for statistical treatment and analysis. The repetition of the input isotopic signal implies that the interannual variation is negligible; an acceptable assumption for the San Francisco catchment considering the high degree of similarity between the same months along the analyzed 2 yr period (Fig. 4). Comparable monthly isotopic seasonality of rainfall has been described by Goller et al. (2005) for the same study area and for nearby regions with similar climatic conditions, e.g., Amaluza GNIP station (http://www.iaea.org/water).

Modelled output results are available for the weekly time span chosen for the input function (an average signal of rainfall was distributed for every week at Wednesdays 12:00). These results were interpolated in order to perform statistical comparisons with instantaneous observed data. For soil waters, direct comparisons were performed between predictions and observed data.

2.7 Evaluation of model performance

The search for acceptable model parameters for each site was conducted through statistical comparisons of 10,000 simulations based on the Monte-Carlo method, considering a uniform random distribution of the variables involved in each model. For each site and model its performance was calculated using the Nash-Sutcliffe Efficiency (NSE). Quantification of errors and deviations from the observed data were respectively calculated by the root mean square error (RMSE) and the bias. MatLab version 7 was used for data handling and solving the convolution equation.

When looking for the optimum parameter range, we first set a wide range (maybe even unrealistic) to be sure to cover all possible solutions (Table 3). By checking the plots of these preliminary results we were able to identify the convergence of model solutions (we used NSE as the objective function for all model parameters), thereby making it possible, for a second simulation, to narrow down the parameter range for each variable. Once the variation ranges were identified and bounded, according to the largest solution peak for every site and for every variable, all the solutions 5% below the top NSE efficiency were selected. For these behavioral efficiencies, weighted quantiles between 0.05 and 0.95 (90% prediction limits) were calculated in order to refine limits of behavioral solutions for every variable. Using these limits, a final simulation for each site and model was performed (at this stage the 10,000 simulations were allowed to vary only for the corresponding final solution ranges). Results are shown in Tables 4 and 5, as well as in Annexes 1 and 2.

The before mentioned approach is based on the Generalized Likelihood Uncertainty Estimation (GLUE, Beven and Freer, 2001). The GLUE approach considers that several likely solutions are valid as long as efficiency of a particular simulation is above a pre-set, but subjective threshold. In this sense, considering the large number of sites and models used, no specific lower limit was set to discriminate predictions, but (as explained earlier) a range that depended on the top efficiency for each case. Only for the analysis of results and for intercomparison between predictions, we considered that a prediction was poor for NSE < 0.45.

The following three criteria were used to select the best solutions of MTTs and TTDs from the final model runs: 1) NSE; 2) magnitude of the uncertainty of the prediction, expressed as a percent of the predicted MTT value; and 3) percentage of observations covered by the range of behavioral solutions defined according to the second criteria.

3 Results

3.1 Soil water

Of all predictions the best matches of the models, with respect to the NSE objective function, ranged between 0.64 and 0.91 (Fig. 5a). When only the best goodness of fit was considered, the GM and the EPM models performed best in most of the sampled sites (13 from 18), followed by the DM, LM and LPM models (Fig. 5b). Only these models were considered for further mutual comparison. Even when the derived MTT values were similar among the models that best fitted the objective function (Fig. 6a, Table 4 and Annex 1), the LPM model performed best taking into consideration additional selection criteria, as shown in Figs. 6b and 6c. Fig. 7 depicts, for the LPM model applied to site C2, the uncertainty and the range of behavioral solutions for the two model parameters.

Considering results from the LPM model (Table 4), differences between observed and predicted values described by the RMSE are up to 1.72‰ and the larger absolute bias accounts for 0.181‰ (Table 4). Bearing in mind the ranges of behavioral solution, MTT results were between 2.3 to 6.3 weeks for pastures soils and between 3.7 to 9.2 weeks for forested soils, while parameterizations for η (ratio of the total volume to the volume in which linear flow applies) ranged from 0.84 to 2.23 and from 0.76 to 1.61 respectively.

Regarding to the shapes of the distribution functions, Fig. 8 shows the best matching results for two representative and comparable sampling sites (C2 for pastures and E2 for forest) for each lumped model (results for LM model are not included since best matching results for LPM were achieved with $\eta \approx 1$, see Table 4). These probability (PDF) and cumulative density functions (CDF) depict how water is routed through the system. In this sense, pasture sites generally show a faster and higher response of the tracer peak when compared to forest sites. The CDF (Figs. 8b and d) of all models are quite similar for the major part of the flows, even including the linear function LPM that averages the shape of the peaks described by the other models. Models based on exponential functions (EPM, DM, or GM in Figs. 8b and d) predict a small portion of the flow with an exponentially delayed tail, which is larger for forested sites than for pastures. Best distribution function results (based on highest NSEs) for all sampled sites, according to the type of land cover, are shown in Figs. 9a and b for the LPM and GM models applied to pasture sites, and in Figs. 9c and d for forest sites. Considering the range of possible or behavioral solutions (e.g., shaded area represents range of solutions for C2 site in Figs. 9a and b, and for E2 in Figs. 9c and d), distributions functions for each type of model and land cover are very similar between each sampled site.

3.2 River and tributaries

Considering all sites and models the criteria NSE > 0.45 was exceeded in 41 of the 63 predictions (9 sites per 7 models, Fig. 5a). Among the analyzed sites the TPLR model yielded the best matches for PL, SF, FH, QZ, QN, QM and QC, while the EPM model for the QR and QP sites (Fig. 5b). The GM model reached closest efficiencies when compared to the best match for every site. Consequently only the TPLR, EPM and GM models were further considered. Differences between MTT predictions for all sites are depicted in Fig. 10a and results from retained models in Table 5 and Annex 2. Although MTT results according to the best NSEs were reached using the TPLR model, compared to the GM or the EPM, these predictions also showed the largest uncertainties (Fig. 10b) and at the same time depicted the lowest number of observations inside the predicted range of behavioral solutions (Fig. 10c). Considering these additional selection criteria, EPM performed better. For stream water at the main outlet, Figs. 11-13 show the parameter uncertainties and behavioral solutions for the TPLR, GM and EPM models, respectively.

Considering results from the EPM model (Table 5, Fig. 10a), the fitting efficiencies reached a maximum NSE of 0.56 for the main stream, and NSEs between 0.48 and 0.58 for the main tributaries (Fig. 5a). The predicted MTT at catchment outlet was 2.0 yr with a η parameter of 1.84 (a similar value was estimated for the main river at the SF sampling site, MTT = 2.0 yr and η = 1.85) and varied from 2.0 (QM, η = 1.85) to 3.9 yr (QC, η = 1.97) for the main tributaries. Uncertainties of MTT predictions between sites were similar with a maximum range between 14.1% and 20.4% of the predicted MTT, as derived for the FH and QM sites (Table 5). Similarly, η ranged from 1.61 (QZ) to 2.21 (QP), the average value of η = 1.85 implies a 54% of volume portion of exponential flow and a 46% volume of piston flow; the uncertainty for the η parameter was 25% on average.

Figures 14a and 14b show the shape of the TTD for the main river outlet (PL), corresponding to the highest NSEs for EPM, GM and TPLR models. The curve for EPM shows a delayed peak that is not accounted in the GM or TPLR models (Fig. 14a), which in turn are very similar between them (at least after a short initial time since GM tends to infinity for times closes to cero). Besides, the latter models show a more delayed flow tail when compared to EPM, which show in general a faster transit time (Fig. 14b). Differences between stream water TTDs from the main sub-catchments considering EPM and GM models are shown in Figs. 15a and b. For comparison of the degree of similarities between sites, these plots include the range of behavioral solutions for the main outlet (PL), thereby being clear that apart from QC or QP, the remaining sites have similar (EPM or GM) transit time distribution functions.

3.3 Springs and creeks

Of 35 predictions (7 models for 5 sites) the criterion NSE > 0.45 was fulfilled in 20 cases. Sites with reduced isotope signal (small σ) yielded lower efficiencies (Fig. 5a, Table 5 and

Annex 2). Apart from TP and QRS, in the remaining sites the criterion NSE > 0.45 was reached at least by 5 models. TP, PLS and SFS sites were best described by using a TPLR model (Fig. 5b). In this regard, GM and EPM were the second and third best models. Figure 10a shows the MTT results predicted by the three models, while detailed information is given in Table 5 and Annex 2. As for stream waters, the EPM model performed best when looking at the uncertainties and the number of observed data inside the range of behavioral solutions (Figs. 10b and c).

Considering EPM, MTTs of 4.5 yr (NSE = 0.49, η = 1.74) for TP and 2.1 yr (NSE = 0.65, η = 1.84) for Q3 were estimated; while for springs, 2.0 yr (NSE = 0.69, η = 1.85) for PLS and 3.3 yr (NSE = 0.47, η = 1.42) for SFS. Results for the QRS site showed poor reliability due to the reduced amplitude of $\delta^{18}O$ in the observed data (Table 5), the lowest among the observed sites (σ = 0.17). Estimations of MTTs for this site was larger than 5 yr, and therefore beyond the level of applicability of the method for natural isotopic tracers.

Figures 14c and d show the TTD results of EPM, GM and TPLR models, for a representative site with long MTT (creek TP). This site show a distinctive more delayed time to the peak (for EPM model) and longer duration of flow tails compared to stream water (Figs. 14a and b). In Figs. 15c and d, the TTDs for all spring and creek sampled sites are shown for the EPM and GM models. In these figures, it is noticeable that the sites Q3 and PLS show the same patterns described previously for most of the stream waters (Figs. 14a and b), while some differences related to more delayed flow responses can be accounted for SFS, TP or QRS sites (Figs. 15c and d), which are more similar to QP and QC stream waters.

4 Discussion

For each soil water site, similar MTT results of a few weeks to months were obtained regardless of the lumped parameter model used (Fig. 6a, Table 4 and Annex 1). Although the LPM model did not yield predictions with the highest efficiencies (Fig. 5a), provided smaller ranges of uncertainty (Fig. 6b) and a larger number of observations inside them (Fig. 6c), advantages that could not be inferred by using only the best matches to NSE, for which GM and EPM models performed better than others (Fig. 5b). Using a LPM model, suitable to describe a partially confined aquifer with increasing thickness (Maloszewski and Zuber, 1982), we found MTTs varying from 2.3 to 6.3 weeks for pastures sites and from 3.7 to 9.2 weeks for forested soils. If we consider that only the top soil horizon was sampled (maximum sampled depth was 0.4 meters), these results are comparable to values between 7.5 and 31 weeks found in 2.0 meter soil columns of typical Bavarian soil using the DM model (Maloszewski et al., 2006). When analyzing the distribution function for soil waters, similarities between model results are evident (Figs. 8 and 9). Considering the range of possible solutions of each site (shaded areas in Figs. 9a-d), it is noticeable that the major part of the flow's transit can be described similarly by all models, even using the simpler function (LPM). For these sites, when considering exponential models (EPM, GM or DP), a small portion of the flow is depicted as having a delayed tail; however, compared to the magnitude of the total volume, an LPM distribution could still be considered as a reliable method to estimate MTTs.

Considering the LPM results for MTTs of soil water from pastures (4.3 weeks on average) and forest sites (5.9 weeks on average) as independent data sets, a two tailed *p-value* of 0.0075 for a Student's t-test was calculated, meaning that the difference between the two groups was statistically significant, although physical characteristics, like length, slope and altitude and meteorological conditions of the respective hill slopes were more or less similar. Land use effects, affecting soil hydraulic properties controlling the infiltration and flow of water, were detected in previous studies within the research area (Huwe et al., 2008). Confirming findings in other tropical catchments were published by Zimmermann et al. (2006) and by Roa-Garcia and Weiler (2010), who stated that under grazing the hydraulic conductivity decreased, overland and near surface flows increased, the storage capacity of the soil matrix declined, with feedbacks on the MTT of soil water. Similar insights were found by Tetzlaff et al. (2007) comparing two small catchments in Central Scotland Highlands of different land use.

For larger MTTs (> 1 yr), as derived for sampled surface waters and shallow springs, there were differences when predicted results among models were compared (Fig. 10a, Table 5 and Annex 2), especially for sites with strong damped signals of measured $\delta^{18}O$ (e.g. QRS and TP sites). When considering uncertainties, the EPM model performed significantly better when compared to the TPLR or GM models (Figs. 10b and c), although the latter two performed best for most of the sampled surface waters according to the NSE objective function (Figs. 5a and b).

When analyzing results from different models, dotty plots of model parameter uncertainty are very useful to display not only the magnitude of uncertainty but also its tendency. Similarly, the uncertainty bands of behavioral solutions can help to account for the sensitivity of the parameter uncertainty on $\delta^{18}O$ modeled results. For example, when predicted results for the PL site are compared, larger parameter uncertainty and skewness are notorious for TPLR than for EPM or GM models (Figs. 11a-c for TPLR; 12a-c for GM; 13a and b for EPM). At the same time EPM shows the highest sensitivity in modeled results (Figs. 11d, 12d, 13c). In order to contrast the signature of the effluent with younger waters such as rainfall, Figs. 11e, 12e, or 13d show the damped observed (and predicted) $\delta^{18}O$ signatures at the main outlet; a characteristic present in all analyzed surface waters. Considering the efficiencies reached by the predictions, we should keep in mind that ranges of behavioral solutions derived from a fixed 5% of the top NSE are generally smaller than a predefined lower limit for all waters, e.g., a predefined lower efficiency limit of 0.30 and 0.45 were used by Speed et al. (2010) and Capell et al., (2012), respectively.

For stream waters, as for springs and creeks, the main differences between EPM and GM (or TPLR) results consisted first in a delayed response of the tracer signal in the outlet, modeled by a parameter η > 1 (Table 5), while for GM or TPLR the response of the flow occurred instantaneously after the spread of the tracer along the catchment (Figs. 14 and 15, Annex 2); and secondly by a comparatively smaller exponential flow tails, which also means that in general the flow transport is faster considering EPM than GM or TPLR models. For these cases, regardless of the degree of efficiencies or uncertainties, the decision on which TTD is more reliable would depend on the conceptual knowledge of the functioning of the catchment. For the San Francisco catchment this can be gained through additional field experiments in selected sites or sub-catchments using either higher resolution samples from the effluents in order to analyze non steady conditions (Botter et al., 2011) or considering different mixing assumptions (Hrachowitz et al., 2013). Another approach could be to analyze longer time series of stable isotopes, or even to include radioactive isotopes as tritium, which would help to crosscheck results, as it has been claimed that, in some cases, the inferences of the processes using solely stables isotopes, underestimate the delayed part of the flow (Stewart et al., 2010).

Regardless of the used model, efficiencies of MTT for stream waters were lower than for soil waters. This was somehow expected, since the dampening effect on a catchment to sub-catchment scale generates a smoother signal filtering/averaging the heterogeneity observed at a single point along a precise transect. Since for most of the cases MTTs for soil waters showed an increasing trend according to increasing soil depth, longer MTTs corresponding to deeper soil layers are to be expected. Soil water below 0.4 m was not monitored within this study, given the shallow soil depth and the increasing fraction of rock material with depth, preventing the use of wick samplers.

The similarities and differences between models for sites with MTTs > 1 yr, as for stream and spring waters, gave insights about the importance to account for a proper TTD, defined according to the conceptual knowledge of the catchment's functioning, before calculating MTT. In this regard, the use of a multi-model approach and uncertainty analysis is believed essential as to be able of defining which functions describes in a better way the parameter identifiability and bounds of behavioral solutions. By considering best matches to NSE for stream waters, best predictions were obtained with the TPLR, EPM and GM models; being more flexible versions of a pure exponential distribution function (i.e. EM model), which help to account for non-linearities of the system. The same distribution functions were identified as good predictors of observed data in a related study by Weiler et al. (2003). When comparing the TPLR to EPM or GM models, the latter two take the non-linearity of the flow without splitting it in two reservoirs with different exponential behaviors into account, therefore yielding more identifiable results. However, findings by Weiler et al. (2003) suggest that the TPLR distribution function could achieve better predictions for runoff events generated by mixed fast and slow flows. In related studies using multiple models, the EPM model yielded the best predictions for surface and spring waters (Viville et al., 2006). Considering this model, in the San Francisco catchment, the average η = 1.85 value for surface waters (similar values were found for creeks: η = 1.79 and springs: η = 1.64) implies that a significant portion of old water (46%) is released previous to the new one (54%). The η value in this study is larger than the η value found in studies for stream water in temperate small headwaters catchments (η = 1.09, Kabeya et al., 2006; η = 1.28, McGuire et al., 2002; η = 1.37, Asano et al., 2002), and close to results published by Katsuyama et al. (2009) for two riparian groundwater systems (η = 1.6 and 1.7).

Regarding to the Gamma model, it was also identified as an applicable distribution function in headwater montane catchments with dominant baseflow in temperate climate (Hrachowitz et al., 2009a, 2010; Dunn et al., 2010). For our study area, a characteristic shape parameter $\alpha < 1$ (e.g. Fig. 12b and Annex 2) was found in all stream and spring sites meaning that an initial peak or a significant part of the flow was quickly transported to the river. Similar results were found recently for mountain catchments of comparable size in Scotland by Kirchner et al.

(2010), who also stated the importance for accounting the best distribution shape, which is usually assumed as purely exponential (α = 1). MTTs derived without the use of observed data, using a purely exponential model, frequently led to an overestimation of α and consequently an underestimation of MTTs. The higher flexibility of the GM model permits to account for the non-linearity in the behavior of a catchment system (Hrachowitz et al., 2010).

5 Conclusions

The research revealed that looking for the best TTD and its derived MTT is not only matter of accounting for the best fit to a predefined objective function, instead, it is recommended to (1) include in the analysis several potential TTD models, (2) assess the uncertainty range of predictions and (3) account for the parameter identifiability. Although the uncertainty range increases for MTTs larger than 1 to 2 yr, using simpler models that still yield acceptable fits to an objective function can help to reduce the uncertainty associated to the predictions. In this sense, using the best predictions from models like LPM for soil waters and EPM for surface and spring waters yielded a more reliable range of MTT inferences through lowering the uncertainty associated in the predictions of certain models. Sites that showed substantial differences in predictions between models (e.g. QRS or TP) were related to a strong reduction of the isotopic signal yielding larger uncertainties and extended MTT predictions getting close to the limitations of the used method. It is recommended to interpret these results with care, even to not consider them until longer time series of isotopic data are available.

The diversity of sampling sites and uncertainty analysis, based on the best fits to the objective function NSE and the identifiability of the parameters of the convolution equations of 7 conceptual models, allowed to define with adequate accuracy the ranges of variation of the mean transit times and the proper distributions functions for the main hydrological compartments of the San Francisco catchment. Pure exponential distributions (i.e. EM) provided the poorest predictions in all sites, suggesting non-linearities of the processes, as produced by preferential or bypass flow. On the other hand, models such as EPM or GM which have a better performance in terms of considering the non-linearity, in most cases yielded better fits to the observed data and at the same time better identifiability of its variables (τ, η or α).

For baseflow conditions, which are annually dominant in the catchment area, stream water at the main outlet (PL) and five tributaries (FH, QZ, QN, QR, QM) yielded similar MTT estimations, ranging from 1.8 to 2.5 yr, including uncertainty ranges; while the MTT estimation for two tributaries (QP and QC) were between 3.5 to 4.4 yr. Despite the similar contribution areas, 2 small creeks described contrasting transit times, TP between 4.2 and 5.1 yr, and Q3 between 1.9 and 2.2 yr. Springs showed a longer variation range, from 2.0 yr for PLS to larger than 5 yr for QRS. Considering the predominance of the stream water characteristics of the larger sub-catchments and the higher variability of smaller tributaries (creeks and springs), there is a clear indication that the heterogeneity of the small scale aquifers is averaged in large areas. In this sense, an in depth analysis on individual functioning or intercomparison between analyzed sites, which was beyond the scope of this paper, should be performed in selected areas using longer time series.

Two transects based on land cover characteristics showed differences in MTTs. Pastures have shorter ranges (2.3-6.3 weeks) than forested (3.7-9.2 weeks) areas. Considering the characteristics of the sampling sites (Table 1), results suggest a possible regulatory effect of land use on water movement. Although the representativeness of the sampled sites is low in comparison to the total catchment area, findings point out the potential of environmental tracer methods for estimating the effects of changes in vegetation, a task usually difficult to accomplish by conventional hydrometric methods.

Acknowledgments

The authors are very grateful for the support provided by Karina Feijo during the field sampling campaign which most of the times was conducted in harsh climatic conditions. Thanks are due to the German students spending throughout the research short-stays at the San Francisco Research Station helping with the realization of the aims of the project and more importantly for providing a friendly working environment. In this regard we like to acknowledge especially the dedication of Caroline Fries, Thomas Waltz and Dorothee Hucke. Furthermore, special thanks are due to Irene Cardenas for her unconditional support with the vast amount of lab analyses. Thanks are also due to Thorsten Peters of the University of Erlangen for providing meteorological data and the logistic support offered by Felix Matt and Jorg Zeilinger, and the administrative and technical staff of the San Francisco Research Station. The authors recognize that this research would not have been possible without the financial support of the German Research Foundation (DFG, BR2238/4-2) and the Secretaria Nacional de Ciencia, Tecnología e Innovación (SENESCYT). Last but not least, the authors thank the useful remarks provided by Markus Hrachowitz and from another unknown reviewer.

References

Asano, Y., Uchida, T. and Ohte, N.: Residence times and flow paths of water in steep unchannelled catchments, Tanakami, Japan, J. Hydrol., 261, 173–192, doi:10.1016/S0022-1694(02)00005-7, 2002.

Barnes, C. J. and Bonell, M.: Application of unit hydrograph techniques to solute transport in catchments, Hydrol. Process., 10, 793–802, doi:10.1002/(SICI)1099-1085(199606)10:6<793::AID-HYP372>3.3.CO;2-B, 1996.

Barthold, F. K., Wu, J., Vache, K. B., Schneider, K., Frede, H.-G. and Breuer, L.: Identification of geographic runoff sources in a data sparse region: hydrological processes and the limitations of tracer-based approaches, Hydrol. Process., 24, 2313–2327, doi:10.1002/hyp.7678, 2010.

Beck, E., Makeschin, F., Haubrich, F., Richter, M., Bendix, J. and Valerezo, C.: The ecosystem (Reserva Biológica San Francisco), in: Gradients in a Tropical Mountain Ecosystem of Ecuador, edited by: Beck, E., Bendix, J., Kottke, I., Makeschin, F., and Mosandl, R., Springer, Berlin, 1–13, 2008.

Bendix, J., Homeier, J., Ortiz, E. C., Emck, P., Breckle, S.-W., Richter, M. and Beck, E.: Seasonality of weather and tree phenology in a tropical evergreen mountain rain forest, Int. J. Biometeorol., 50, 370–384, doi:10.1007/s00484-006-0029-8, 2006.

Bendix, J., Rollenbeck, R., Fabian, P., Emck, P., Richter, M. and Beck, E.: Climate Variability, in: Gradients in a Tropical Mountain Ecosystem of Ecuador, edited by: Beck, E., Bendix, J., Kottke, I., Makeschin, F., and Mosandl, R., Springer, Berlin, 281–290, 2008a.

Bendix, J., Rollenbeck, R., Richter, M., Fabian, P. and Emck, P.: Climate, in: Gradients in a Tropical Mountain Ecosystem of Ecuador, edited by: Beck, E., Bendix, J., Kottke, I., Makeschin, F., and Mosandl, R., Springer, Berlin, 63–73, 2008b.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, doi:10.1016/S0022-1694(01)00421-8, 2001.

Boiten, W.: Hydrometry, Taylor & Francis, The Netherlands, 2000.

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, W03514, 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, L11403, doi:10.1029/2011GL047666, 2011.

Brehm, G., Homeier, J., Fiedler, K., Kottke, I., Illig, J., Nöske, N. M., Werner, F. A. and Breckle, S. W.: Mountain rain forests in southern Ecuador as a hotspot of biodiversity – limited knowledge and diverging patterns, in: Gradients in a Tropical Mountain Ecosystem of Ecuador, edited by: Beck, E., Bendix, J., Kottke, I., Makeschin, F., and Mosandl, R., Springer, Berlin, 15–23, 2008.

Breuer, L., Windhorst, D., Fries, A. and Wilcke, W.: Supporting, regulating, and provisioning hydrological services, in ecosystem services, biodiversity and environmental change, in: A Tropical Mountain Ecosystem of South Ecuador, edited by: Bendix, J., Beck, E., Bräuning, A., Makeschin, F., Mosandl, R., Scheu, S., and Wilcke, W., Springer, Berlin, 107–116, 2013.

Buecker, A., Crespo, P., Frede, H.-G. and Breuer, L.: Solute behaviour and export rates in neotropical montane catchments under different land-uses, J. Trop. Ecol., 27, 305–317, doi:10.1017/S0266467410000787, 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.

Craig, H.: Standard for reporting concentrations of deuterium and oxygen-18 in natural waters, Science, 133, 1833, doi:10.1126/science.133.3467.1833, 1961.

Crespo, P., Buecker, A., Feyen, J., Vache, K. B., Frede, H.-G. and Breuer, L.: Preliminary evaluation of the runoff processes in a remote montane cloud forest basin using mixing model analysis and mean transit time, Hydrol. Process., 26, 3896–3910, doi:10.1002/hyp.8382, 2012.

Dansgaard, W.: Stable isotopes in precipitation, Tellus, 16, 436–468, doi:10.1111/j.2153-3490.1964.tb00181.x, 1964.

Darracq, A., Destouni, G., Persson, K., Prieto, C. and Jarsjo, J.: Scale and model resolution effects on the distributions of advective solute travel times in catchments, Hydrol. Process., 24, 1697–1710, doi:10.1002/hyp.7588, 2010.

Dunn, S. M., Birkel, C., Tetzlaff, D. and Soulsby, C.: Transit time distributions of a conceptual model: their characteristics and sensitivities, Hydrol. Process., 24, 1719–1729, doi:10.1002/hyp.7560, 2010.

Fenicia, F., Wrede, S., Kavetski, D., Pfister, L., Hoffmann, L., Savenije, H. H. G. and McDonnell, J. J.: Assessing the impact of mixing assumptions on the estimation of streamwater mean residence time, Hydrol. Process., 24, 1730–1741, doi:10.1002/hyp.7595, 2010.

Goettlicher, D., Obregon, A., Homeier, J., Rollenbeck, R., Nauss, T. and Bendix, J.: Land-cover classification in the Andes of southern Ecuador using Landsat ETM plus data as a basis for SVAT modelling, Int. J. Remote Sens., 30, 1867–1886, doi:10.1080/01431160802541531, 2009.

Goller, R., Wilcke, W., Leng, M. J., Tobschall, H. J., Wagner, K., Valarezo, C. and Zech, W.: Tracing water paths through small catchments under a tropical montane rain forest in south Ecuador by an oxygen isotope approach, J. Hydrol., 308, 67–80, doi:10.1016/j.jhydrol.2004.10.022, 2005.

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.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., Dawson, J. J. C., Dunn, S. M. and Malcolm, I. A.: Using long-term data sets to understand transit times in contrasting headwater catchments, J. Hydrol., 367, 237–248, doi:10.1016/j.jhydrol.2009.01.001, 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., Soulsby, C., Tetzlaff, D. and Malcolm, I. A.: Sensitivity of mean transit time estimates to model conditioning and data availabilityle, Hydrol. Process., 25, 980–990, doi:10.1002/hyp.7922, 2011.

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.

Huwe, B., Zimmermann, B., Zeilinger, J., Quizhpe, M. and Elsenbeer, H.: Gradients and patterns of soil physical parameters at local, field and catchment scales, in: Gradients in a Tropical Mountain Ecosystem of Ecuador, edited by: Beck, E., Bendix, J., Kottke, I., Makeschin, F., and Mosandl, R., Springer, Berlin, 375–386, 2008.

Iorgulescu, I., Beven, K. J. and Musy, A.: Flow, mixing, and displacement in using a data-based hydrochemical model to predict conservative tracer data, Water Resour. Res., 43, W03401, doi:10.1029/2005WR004019, 2007.

Kabeya, N., Katsuyama, M., Kawasaki, M., Ohte, N. and Sugimoto, A.: Estimation of mean residence times of subsurface waters using seasonal variation in deuterium excess in a small headwater catchment in Japan, Hydrol. Process., 21, 308–322, doi:10.1002/hyp.6231, 2006.

Katsuyama, M., Kabeya, N. and Ohte, N.: Elucidation of the relationship between geographic and time sources of stream water using a tracer approach in a headwater catchment, Water Resour. Res., 45, W06414, doi:10.1029/2008WR007458, 2009.

Kendall, C. and McDonnell, J. J.: Isotope Tracers in Catchment Hydrology, Elsevier, Amsterdam, The Netherlands, 1998.

Kirchner, J. W., Feng, X. H. and Neal, C.: Fractal stream chemistry and its implications for contaminant transport in catchments, Nature, 403, 524–527, 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.

Landon, M. K., Delin, G. N., Komor, S. C. and Regan, C. P.: Comparison of the stable-isotopic composition of soil water collected from suction lysimeters, wick samplers, and cores in a sandy unsaturated zone, J. Hydrol., 224, 45–54, doi:10.1016/S0022-1694(99)00120-1, 1999.

Landon, M. K., Delin, G. N., Komor, S. C. and Regan, C. P.: Relation of pathways and transit times of recharge water to nitrate concentrations using stable isotopes, Ground Water, 38, 381–395, doi:10.1111/j.1745-6584.2000.tb00224.x, 2000.

Leibundgut, C., Maloszewski, P. and Külls, C.: Environmental tracers, in: Tracers in Hydrology, John Wiley & Sons, Ltd., Chichester, UK, 13–56, doi: 10.1002/9780470747148.ch3, 2009.

Liess, M., Glaser, B. and Huwe, B.: Digital soil mapping in southern Ecuador, Erdkunde, 63, 309–319, doi:10.3112/erdkunde.2009.04.02, 2009.

Maher, K.: The role of fluid residence time and topographic scales in determining chemical fluxes from landscapes, Earth Planet. Sci. Lett., 312, 48–58, doi:10.1016/j.epsl.2011.09.040, 2011.

Maloszewski, P. and Zuber, A.: Determining the turnover time of groundwater systems with the aid of environmental tracers, 1. models and their applicability, J. Hydrol., 57, 207–231, 1982.

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

Maloszewski, P., Maciejewski, S., Stumpp, C., Stichler, W., Trimborn, P. and Klotz, D.: Modelling of water flow through typical Bavarian soils: 2. environmental deuterium transport, Hydrol. Sci. J.-J. Sci. Hydrol., 51, 298–313, doi:10.1623/hysj.51.2.298, 2006.

McDonnell, J. J., McGuire, K., Aggarwal, P., Beven, K. J., Biondi, D., Destouni, G., Dunn, S., James, A., Kirchner, J., Kraft, P., Lyon, S., Maloszewski, P., Newman, B., Pfister, L., Rinaldo, A., Rodhe, A., Sayama, T., Seibert, J., Solomon, K., Soulsby, C., Stewart, M., Tetzlaff, D., Tobin, C., Troch, P., Weiler, M., Western, A., Worman, 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., DeWalle, D. R. and Gburek, W. J.: Evaluation of mean residence time in subsurface waters using oxygen-18 fluctuations during drought conditions in the mid-Appalachians, J. Hydrol., 261, 132–149, doi:10.1016/S0022-1694(02)00006-9, 2002.

McGuire, K. J., Weiler, M. and McDonnell, J. J.: Integrating tracer experiments with modeling to assess runoff processes and water transit times, Adv. Water Resour., 30, 824–837, doi:10.1016/j.advwatres.2006.07.004, 2007.

Mertens, J., Diels, J., Feyen, J. and Vanderborght, J.: Numerical analysis of passive capillary wick samplers prior to field installation, Soil Sci. Soc. Am. J., 71, 35–42, doi:10.2136/sssaj2006.0106, 2007.

Munoz-Villers, L. E. and McDonnell, J. J.: Runoff generation in a steep, tropical montane cloud forest catchment on permeable volcanic substrate, Water Resour. Res., 48, W09528, doi:10.1029/2011WR011316, 2012.

Murphy, B. P. and Bowman, D. M. J. S.: What controls the distribution of tropical forest and savanna?, Ecol. Lett., 15, 748–758, doi:10.1111/j.1461-0248.2012.01771.x, 2012.

Plesca, I., Timbe, E., Exbrayat, J.-F., Windhorst, D., Kraft, P., Crespo, P., Vache, K. B., Frede, H.-G. and Breuer, L.: Model intercomparison to explore catchment functioning: results from a remote montane tropical rainforest, Ecol. Model., 239, 3–13, doi:10.1016/j.ecolmodel.2011.05.005, 2012.

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, W07537, doi:10.1029/2011WR010478, 2011.

Roa-Garcia, 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.

Rodgers, P., Soulsby, C. and Waldron, S.: Stable isotope tracers as diagnostic tools in upscaling flow path understanding and residence time estimates in a mountainous mesoscale catchment, Hydrol. Process., 19, 2291–2307, doi:10.1002/hyp.5677, 2005.

Rollenbeck, R., Bendix, J. and Fabian, P.: Spatial and temporal dynamics of atmospheric water inputs in tropical mountain forests of South Ecuador, Hydrol. Process., 25, 344–352, doi:10.1002/hyp.7799, 2011.

Rose, T. P., Davisson, M. L. and Criss, R. E.: Isotope hydrology of voluminous cold springs in fractured rock from an active volcanic region, northeastern California, J. Hydrol., 179, 207–236, doi:10.1016/0022-1694(95)02832-3, 1996.

Soulsby, C., Malcolm, R., Helliwell, R., Ferrier, R. C. and Jenkins, A.: Isotope hydrology of the Allt a' Mharcaidh catchment, Cairngorms, Scotland: implications for hydrological pathways and residence times, Hydrol. Process., 14, 747–762, doi:10.1002/(SICI)1099-1085(200003)14:4<747::AID-HYP970>3.0.CO;2-0, 2000.

Soulsby, C., Tetzlaff, D. and Hrachowitz, M.: Tracers and transit times: windows for viewing catchment scale storage?, Hydrol. Process., 23, 3503–3507, doi:10.1002/hyp.7501, 2009.

Speed, M., Tetzlaff, D., Soulsby, C., Hrachowitz, M. and Waldron, S.: Isotopic and geochemical tracers reveal similarities in transit times in contrasting mesoscale catchments, Hydrol. Process., 24, 1211–1224, doi:10.1002/hyp.7593, 2010.

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.

Tetzlaff, D., Malcolm, I. A. and Soulsby, C.: Influence of forestry, environmental change and climatic variability on the hydrology, hydrochemistry and residence times of upland catchments, J. Hydrol., 346, 93–111, doi:10.1016/j.jhydrol.2007.08.016, 2007.

Turner, J., Albrechtsen, H. J., Bonell, M., Duguet, J. P., Harris, B., Meckenstock, R., McGuire, K., Moussa, R., Peters, N., Richnow, H. H., Sherwood-Lollar, B., Uhlenbrook, S. and van Lanen, H.: Future trends in transport and fate of diffuse contaminants in catchments, with special emphasis on stable isotope applications, Hydrol. Process., 20, 205–213, doi:10.1002/hyp.6074, 2006.

Vache, K. B. and McDonnell, J. J.: A process-based rejectionist framework for evaluating catchment runoff model structure, Water Resour. Res., 42, W02409, doi:10.1029/2005WR004247, 2006.

Viville, D., Ladouche, B. and Bariac, T.: Isotope hydrological study of mean transit time in the granitic Strengbach catchment (Vosges massif, France): application of the FlowPC model with modified input function, Hydrol. Process., 20, 1737–1751, doi:10.1002/hyp.5950, 2006.

Weiler, M., McGlynn, B. L., McGuire, K. J. and McDonnell, J. J.: How does rainfall become runoff?, a combined tracer and runoff transfer function approach, Water Resour. Res., 39, 1315-1327, doi:10.1029/2003WR002331, 2003.

Wilcke, W., Yasin, S., Abramowski, U., Valarezo, C. and Zech, W.: Nutrient storage and turnover in organic layers under tropical montane rain forest in Ecuador, Eur. J. Soil Sci., 53, 15–27, doi:10.1046/j.1365-2389.2002.00411.x, 2002.

Willems, P.: A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models, Environ. Model. Softw., 24, 311–321, doi:10.1016/j.envsoft.2008.09.005, 2009.

Windhorst, D., Waltz, T., Timbe, E., Frede, H.-G. and Breuer, L.: Impact of elevation and weather patterns on the isotopic composition of precipitation in a tropical montane rainforest, Hydrol. Earth Syst. Sci., 17, 409–419, doi:10.5194/hess-17-409-2013, 2013.

Wolock, D. M., Fan, J. and Lawrence, G. B.: Effects of basin size on low-flow stream chemistry and subsurface contact time in the Neversink River Watershed, New York, Hydrol. Process., 11, 1273–1286, doi:10.1002/(SICI)1099-1085(199707)11:9<1273::AID-HYP557>3.0.CO;2-S, 1997.

Zimmermann, B., Elsenbeer, H. and De Moraes, J. M.: The influence of land-use changes on soil hydraulic properties: Implications for runoff generation, For. Ecol. Manag., 222, 29–38, doi:10.1016/j.foreco.2005.10.070, 2006.

Table 1. Main characteristics of the San Francisco catchment and its tributaries.

Parameter	Units	Outlet	Sub-catchment						
		PL	FH	QZ	QN	QR	QP	QM	QC
Catchment physical characteristics									
Drainage area	[km^2]	76.9	34.9	11.2	9.8	4.7	3.4	1.3	0.7
Mean elevation	[m a.s.l.]	2,531	2,615	2,615	2,591	2,472	2,447	2,274	2,290
Altitude range	[m]	1,325	1,133	991	975	1,424	975	772	516
Mean slope	[%]	63	63	63	60	69	67	57	56
Hydrological parameters									
Discharge	[mm]	2,959	2,691	-	1,291	-	-	3,315	2,742
Baseflow	[mm]	2,520	2,152	-	1,044	-	-	2,118	2,268
	[%]	85.2	80.0	-	80.8	-	-	63.9	82.7
Land use									
Forest	[%]	68	67	72	65	80	63	90	22
Sub-páramo	[%]	21	29	15	17	18	10	9	10
Pasture/Bracken	[%]	9	3	12	16	2	26	1	67
Others	[%]	2	1	1	2	0	1	0	1
Soil type									
Histosols	[%]	74	74	70	71	70	62	57	54
Regosols	[%]	15	15	18	16	18	21	25	24
Cambisols	[%]	7	7	8	8	8	11	13	14
Stagnasols	[%]	4	4	4	5	4	6	5	8

Table 2. Applied sampling strategy in the San Francisco catchment.

Sample type	Collection method	Sampled since [a]	Site Name	Site code	Altitude m a.s.l.	Samples Number (Weeks)
Rainfall	Collector	AUG 2010	Estación San Francisco	ECSF	1900	99
Main river	Manually	AUG 2010	Planta (outlet)	PL	1725	104
			San Francisco	SF	1825	104
Tributaries	Manually	AUG 2010	Francisco Head	FH	1917	98
			Zurita	QZ	2047	103
			Navidades	QN	2050	104
			Ramon	QR	1726	104
			Pastos	QP	1925	103
			Milagro	QM	1878	104
			Cruces	QR	1978	102
Creeks	Manually	DEC 2010	Pastos tributary	TP	1950	88
			Q3	Q3	1907	88
Springs	Manually	AUG 2010	PL Spring	PLS	1731	98
			SF Spring	SFS	1826	100
			QR Spring	QRS	1900	100
Pastures soil water	Wick-sampler	NOV 2010	Pastos alto [b]	A1 / A2 / A3	2025	60 / 58 / 45
			Pastos medio [b]	B1 / B2 / B3	1975	70 / 70 /63
			Pastos bajo [b]	C1 / C2 / C3	1925	67 /71 / 55
Forest soil water	Wick-sampler	SEP 2010	Bosque alto [b]	D1 / D2 / D3	2000	78 / 74 / 62
			Bosque medio [b]	E1 / E2 / E3	1900	86 / 80 / 62
			Bosque bajo [b]	F1 / F2 / F3	1825	55 / 53 / 36

[a] Sampling campaign was completed until mid-August 2012.

[b] There are three wick-samplers per site (i.e. A1= 0.10 m, A2 = 0.25 m and A3 = 0.40 m below surface).

Table 3. Lumped parameter models used for the calculation of the transit time distribution.

Model	Transit time distribution g(τ)	Parameter(s) range
Exponential Model (EM)	$\frac{1}{\tau}\exp\left(\frac{-t}{\tau}\right)$	τ [1-400]
Linear Model (LM)	$\frac{1}{2\tau}$ for $t \le 2\tau$ 0 for $t > 2\tau$	τ [1-400]
Exponential Piston flow Model (EPM)	$\frac{\eta}{\tau}\exp\left(-\frac{\eta}{\tau}+\eta-1\right)$ for $t \ge \tau(1-\eta^{-1})$ 0 for $t < \tau(1-\eta^{-1})$	τ [1-400] η [0.5-4]
Linear Piston flow Model (LPM)	$\frac{\eta}{2\tau}$ for $\tau-\frac{\tau}{\eta} \le t \le \tau+\frac{\tau}{\eta}$ 0 for other t	τ [1-400] η [0.5-4]
Dispersion Model (DM)	$\left(\frac{4\pi D_p t}{\tau}\right)^{-1/2} t^{-1}\exp\left[-\left(1-\frac{t}{\tau}\right)^2\left(\frac{\tau}{4D_p t}\right)\right]$	τ [1-400] D_p [0.5-4]
Gamma Model (GM)	$\frac{\tau^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp^{-\tau/\beta}$	α [0.0001-10] τ [1-400] $\beta = \alpha/\tau$
Two Parallel Linear Reservoirs (TPLR)	$\frac{\phi}{\tau_f}\exp\left(-\frac{t}{\tau_f}\right)+\frac{1-\phi}{\tau_s}\exp\left(\frac{-t}{\tau_s}\right)$	τ_s [1-400] τ_f [1-40] ϕ [0-1]

τ = tracer's mean transit time; η = parameter that indicates the percentage contribution of each flow type distribution; D_p = fitting parameter; α and β = shape parameters; τ_s, τ_f = transit time of fast and slow flows, ϕ = flow partition parameter between fast and slow flow reservoirs. Units for parameters and their respective ranges are a-dimensional except for τ, which has units of time (for our case it is given in weeks).

Table 4. Main statistical parameters of observed $\delta^{18}O$ and predicted results for soil waters using a LPM distribution function. Statistical parameters of modeled results: RMSE, bias, mean and σ correspond to the best matching value of the objective function NSE. Uncertainty bounds of modeled parameters (τ and η), calculated through Generalized Likelihood Uncertainty Estimation (GLUE) are showed in parenthesis.

Site	Sampling depth	Observed $\delta^{18}O$, ‰, VSMOW			Modeled $\delta^{18}O$, ‰, VSMOW						
					Mean	σ	NSE	RMSE	Bias	τ	η
	m	Mean	N	σ	‰	‰	-	‰	‰	weeks	-
Pastures transect											
A1	0.10	-6.70	60	3.65	-6.80	3.06	0.87	1.32	-0.099	3.5 (2.8 - 4.4)	1.40 (0.93-2.23)
A2	0.35	-6.79	58	3.33	-6.87	2.46	0.73	1.72	-0.084	5.3 (4.6 - 6.3)	0.99 (0.90-1.28)
A3	0.60	-7.13	45	3.98	-7.31	3.18	0.86	1.46	-0.181	4.9 (3.6 - 5.3)	1.11 (0.88-1.37)
B1	0.10	-6.84	70	3.71	-6.91	3.01	0.83	1.52	-0.069	4.7 (3.4 - 5.1)	1.10 (0.93-1.47)
B2	0.35	-7.03	70	3.41	-7.02	2.71	0.78	1.57	0.007	4.3 (3.9 - 5.3)	0.98 (0.90-1.33)
B3	0.60	-6.76	63	3.41	-6.77	2.97	0.79	1.54	-0.006	4.5 (3.4 - 5.2)	1.03 (0.89-1.45)
C1	0.10	-6.65	67	3.66	-6.74	3.15	0.84	1.44	-0.090	3.3 (2.3 - 4.2)	0.96 (0.87-1.82)
C2	0.35	-7.06	71	3.49	-7.10	3.11	0.87	1.27	-0.043	3.1 (2.7 - 4.4)	0.89 (0.84-1.55)
C3	0.60	-6.52	55	3.07	-6.53	2.56	0.80	1.36	-0.015	5.4 (4.4 - 5.8)	1.09 (0.88-1.32)
Forest transect											
D1	0.10	-7.38	78	3.12	-7.26	2.56	0.78	1.44	0.122	5.7 (4.8 - 6.4)	1.27 (0.97-1.60)
D2	0.35	-7.06	74	2.59	-6.97	2.56	0.78	1.19	0.087	6.8 (5.5 - 9.2)	1.04 (0.86-1.19)
D3	0.60	-6.80	62	2.75	-6.73	2.56	0.80	1.22	0.062	6.0 (4.8 - 6.7)	0.99 (0.86-1.28)
E1	0.10	-6.65	86	3.14	-6.58	2.56	0.80	1.40	0.070	5.1 (4.8 - 6.3)	1.15 (0.93-1.61)
E2	0.35	-6.63	78	2.94	-6.64	2.56	0.78	1.37	-0.016	6.4 (5.7 - 7.3)	1.01 (0.93-1.45)
E3	0.60	-6.44	62	2.57	-6.48	2.56	0.76	1.24	-0.036	8.3 (7.2 - 9.2)	1.03 (0.88-1.18)
F1	0.10	-6.75	55	3.16	-6.79	2.56	0.89	1.05	-0.039	4.3 (3.8 - 5.5)	0.96 (0.87-1.38)
F2	0.35	-6.45	53	3.15	-6.54	2.56	0.89	1.03	-0.089	4.3 (3.7 - 5.5)	0.94 (0.83-1.58)
F3	0.60	-8.09	36	2.56	-8.05	2.56	0.66	1.46	0.045	6.0 (6.0 - 7.8)	0.80 (0.76-0.94)

N = number of samples, σ = standard deviation, RMSE = Root Mean Square Error, NSE = Nash-Sutcliffe Efficiency.

Table 5. Main statistical parameters of observed δ^{18}O and predicted results for surface and spring waters using an EPM distribution function. Statistical parameters of modeled results: RMSE, Bias, Mean and σ correspond to the best matching value of the objective function NSE. Uncertainty bounds of modeled parameters (τ and η), calculated through Generalized Likelihood Uncertainty Estimation (GLUE) are showed in parenthesis.

Site	Drainage area	Outlet altitude	Recharge altitude	Observed δ^{18}O, ‰, VSMOW			Modeled δ^{18}O, ‰, VSMOW						
				Mean	N	σ	Mean	σ	NSE	RMSE	Bias	τ	η
	km^2	m a.s.l.	m a.s.l.				‰	‰	-	‰	‰	yr	-
Stream													
PL	76.93	1,725	2,488	-8.25	97	0.54	-8.25	0.42	0.56	0.36	0.003	2.0(1.8 - 2.2)	1.84(1.73 - 1.98)
SF	65.09	1,825	2,437	-8.12	88	0.56	-8.11	0.43	0.55	0.37	0.001	2.0(1.9 - 2.2)	1.85(1.71 - 1.97)
Streamwater tributaries													
FH	34.92	1,917	2,492	-8.28	83	0.55	-8.28	0.42	0.48	0.39	0.000	2.1(2.0 - 2.3)	1.84(1.70 - 1.93)
QZ	11.25	2,047	2,565	-8.41	93	0.47	-8.42	0.36	0.55	0.32	-0.004	2.2(2.1 - 2.5)	1.72(1.61 - 1.82)
QN	9.79	2,050	2,503	-8.28	92	0.50	-8.28	0.40	0.57	0.33	-0.002	2.1(2.0 - 2.3)	1.78(1.67 - 1.90)
QR	4.66	1,726	2,350	-7.96	97	0.48	-7.96	0.16	0.56	0.32	0.000	2.2(2.0 - 2.4)	1.73(1.62 - 1.84)
QP	3.42	1,925	2,418	-8.07	98	0.34	-8.07	0.26	0.57	0.22	-0.001	3.7(3.5 - 4.1)	2.06(1.91 - 2.21)
QM	1.29	1,878	2,310	-7.81	90	0.59	-7.81	0.44	0.51	0.41	0.005	2.0(1.8 - 2.2)	1.85(1.73 - 1.98)
QC	0.70	1,978	2,197	-7.62	95	0.30	-7.62	0.24	0.58	0.19	0.000	3.9(3.8 - 4.4)	1.97(1.81 - 2.06)
Creeks													
TP	0.14	1,950	2,213	-7.66	80	0.25	-7.66	0.20	0.49	0.17	0.000	4.5(4.2 - 5.1)	1.74(1.61 - 1.82)
Q3	0.10	1,907	2,165	-7.67	88	0.54	-7.67	0.45	0.65	0.32	-0.002	2.1(1.9 - 2.2)	1.84(1.72 - 2.01)
Springs													
PLS	-	1,731	2,377	-8.03	101	0.50	-8.04	0.43	0.69	0.28	-0.009	2.0(1.9 - 2.2)	1.85(1.70 - 1.94)
SFS	-	1,826	2,187	-7.61	101	0.29	-7.61	0.23	0.47	0.21	-0.002	3.3(3.0 - 3.6)	1.42(1.36 - 1.47)
QRS	-	1,900	2,285	-7.80	97	0.17	-7.79	0.09	0.28	0.14	0.005	9.6(8.8 - 10.1)	1.70(1.65 - 1.82)

N = number of samples, σ= standard deviation, RMSE = Root Mean Square Error, NSE = Nash-Sutcliffe Efficiency.

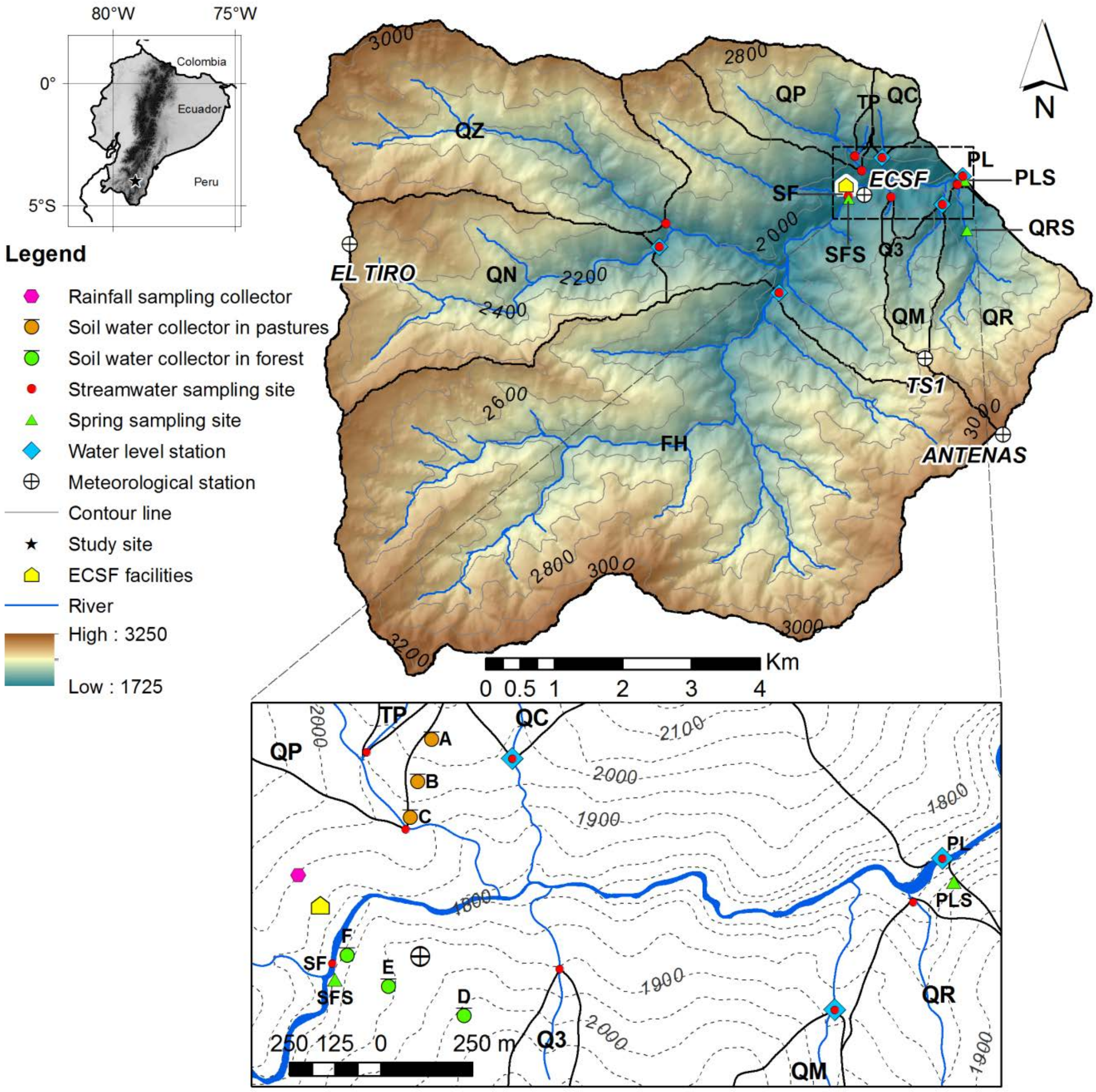


Fig. 1. San Francisco catchment with sampling locations and delineation of drainage area. Acronyms in bold are defined in Table 1. Framed image shows the zoomed area of the lower part of the catchment.

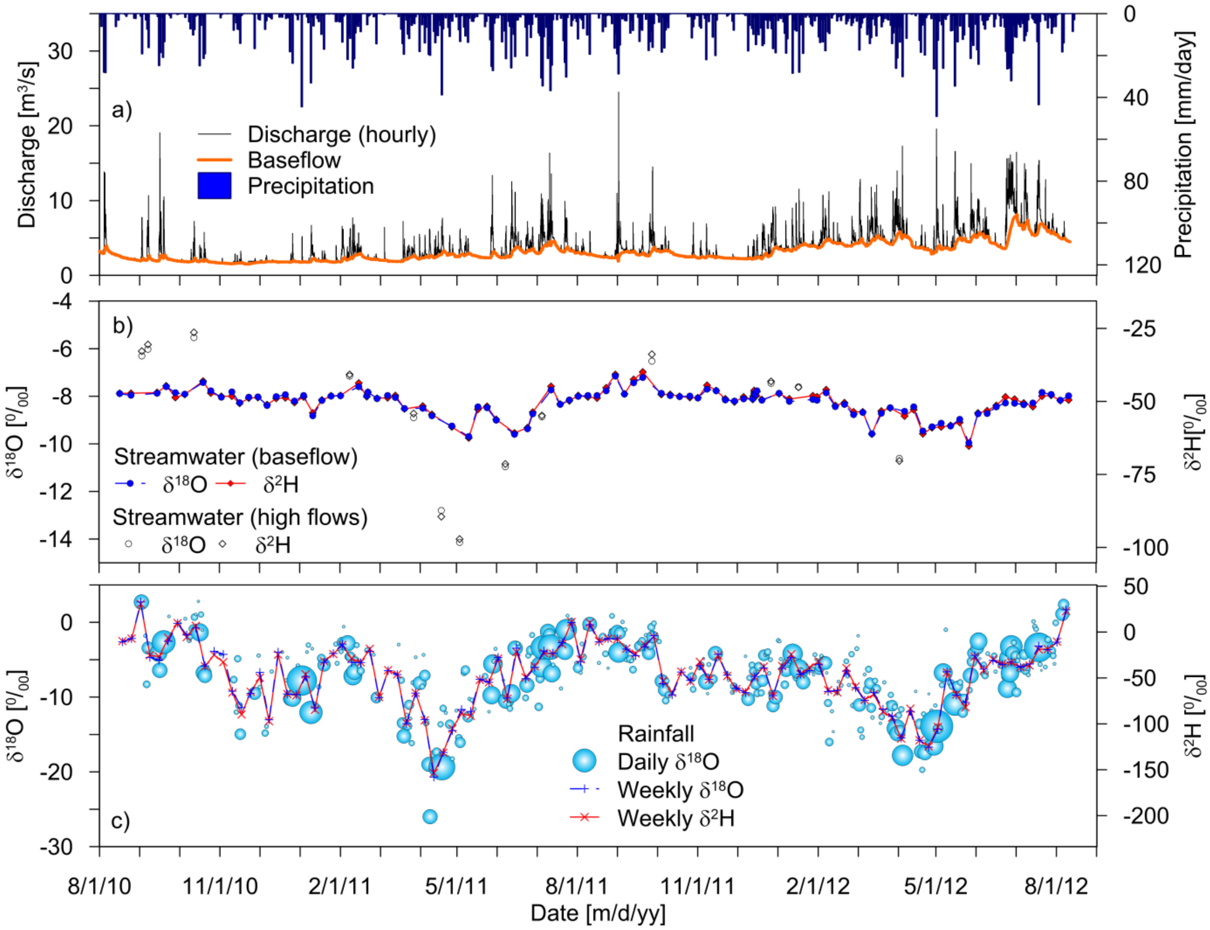


Fig. 2. (a) Time series of rainfall for ECSF meteorological station, hourly discharge and baseflows at the catchment outlet (PL); **(b)** weekly $\delta^{18}O$ and $\delta^{2}H$ of streamwater at PL for baseflow and high flow conditions; and **(c)** weekly $\delta^{18}O$ and $\delta^{2}H$ at the ECSF rainfall sampling collector; light blue bubbles indicate daily $\delta^{18}O$ and relative volume of daily rainfall.

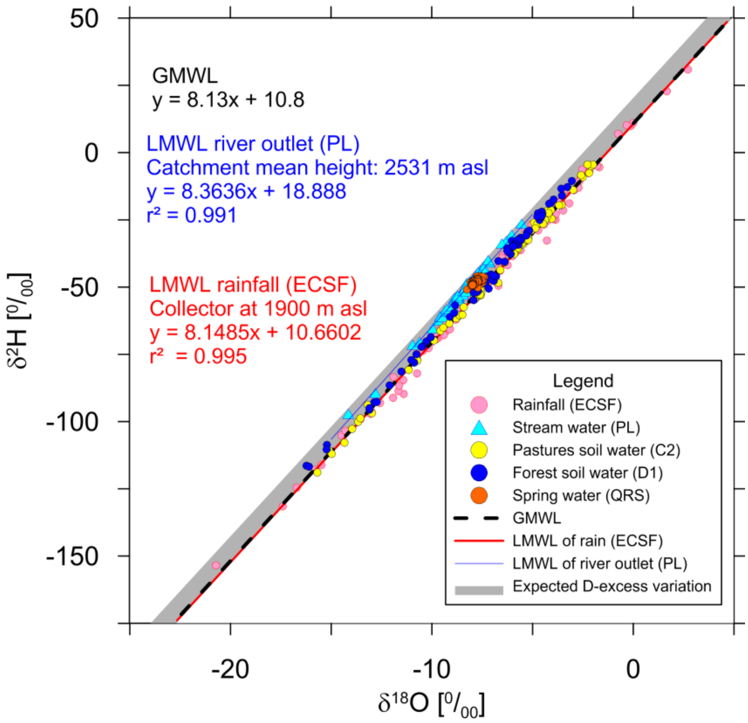


Fig. 3. Shaded area depicts the expected variation range of the Local Meteorological Water Line of rainfall (LMWL) considering the altitudinal range of the catchment (1,725-3,150 m a.s.l.) and estimated d-excess gradient. Symbols in colors depict weekly values of some of the catchment's waters. Acronyms are defined in Table 1.

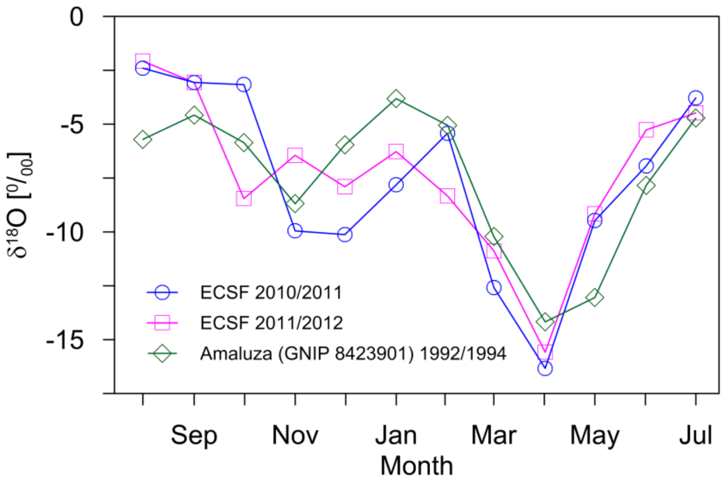


Fig. 4. Monthly isotopic $\delta^{18}O$ signals between two consecutive years (2010-2012) at ECSF (1,900 m a.s.l.) and averaged monthly values (1992-1994) at Amaluza GNIP station (latitude -2.61, longitude -78.57, altitude 2,378 m a.s.l.).

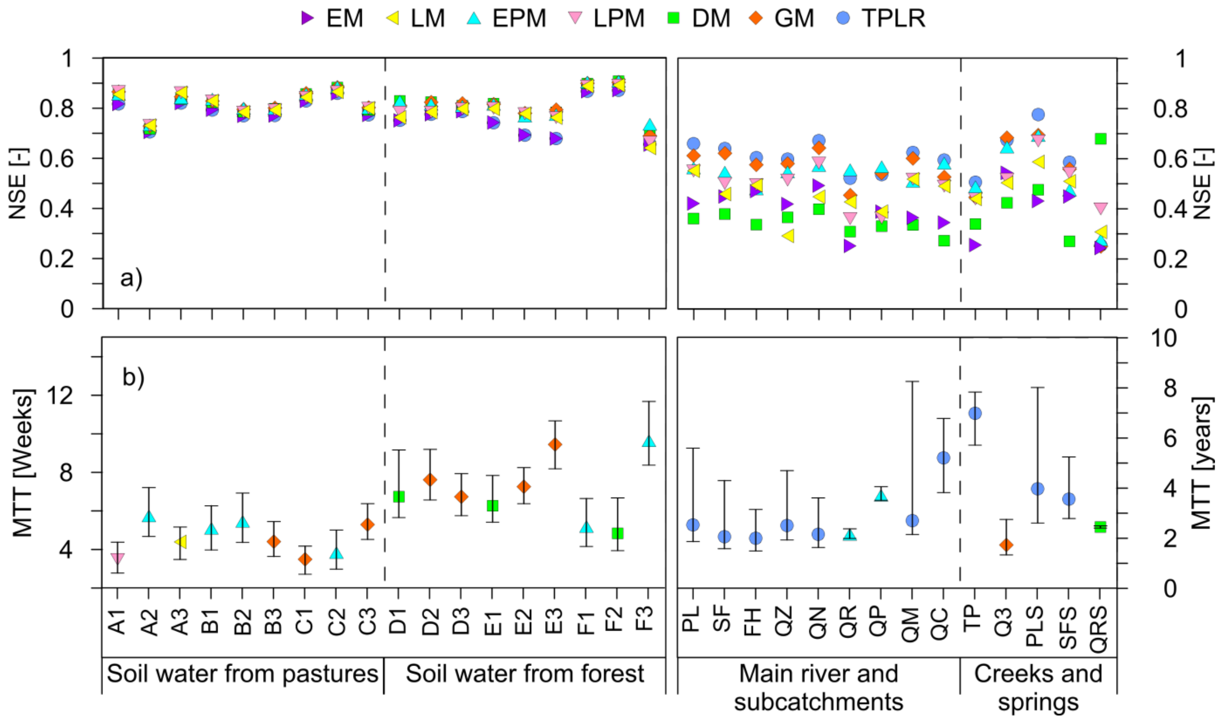
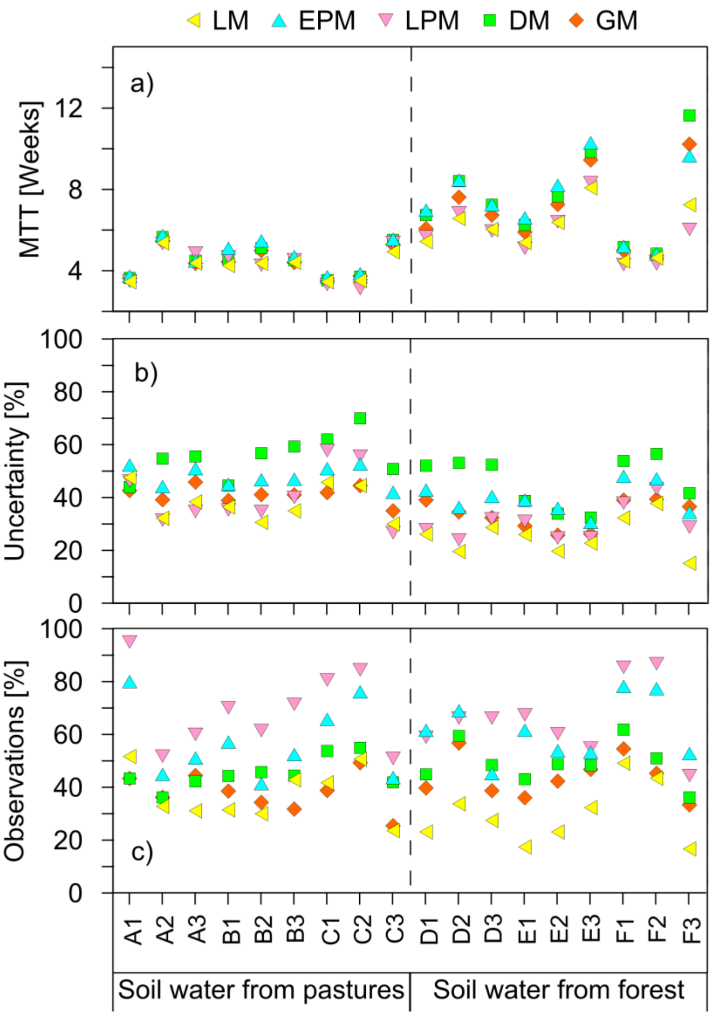


Fig. 5. (a) Best NSE for each of the seven lumped parameter models; **(b)** MTT estimation according the best NSE per site: symbols represent MTT corresponding to the best matching result among 7 models considering the NSE criteria showed in **(a)**, vertical line represents uncertainty bounds according the GLUE methodology for the selected model.

892



893 **Fig. 6.** Intercomparison of models for soil sites according to their: **(a)** estimated mean transit
894 times; **(b)** uncertainty ranges expressed in percentage of its respective MTT estimation; and
895 **(c)** number of observations inside the range of behavioral solutions.

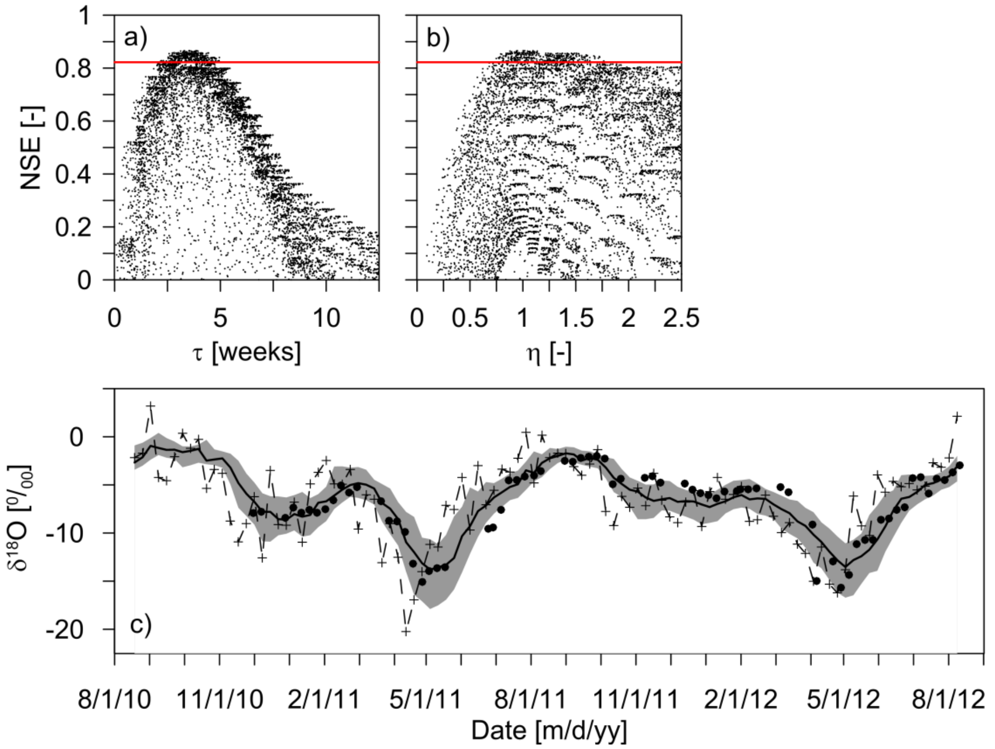


Fig. 7. Fitted results of the LPM model compared to observed data for soil water of a pastures site (C2). Sub-plots **(a)** and **(b)** show the uncertainty analysis of 10,000 simulations and the feasible range of behavioral solutions of model parameters as a 5% of the top best prediction. Black filled circles in sub-plot **(c)** represents the observed data; the black line and the shaded area represent the best possible solution and its range of variation according to the 5-95% confidence limits of the behavioral solutions shown in **(a)**; and the gray dashed line with crosses represents the weekly rainfall variation as input function for the model.

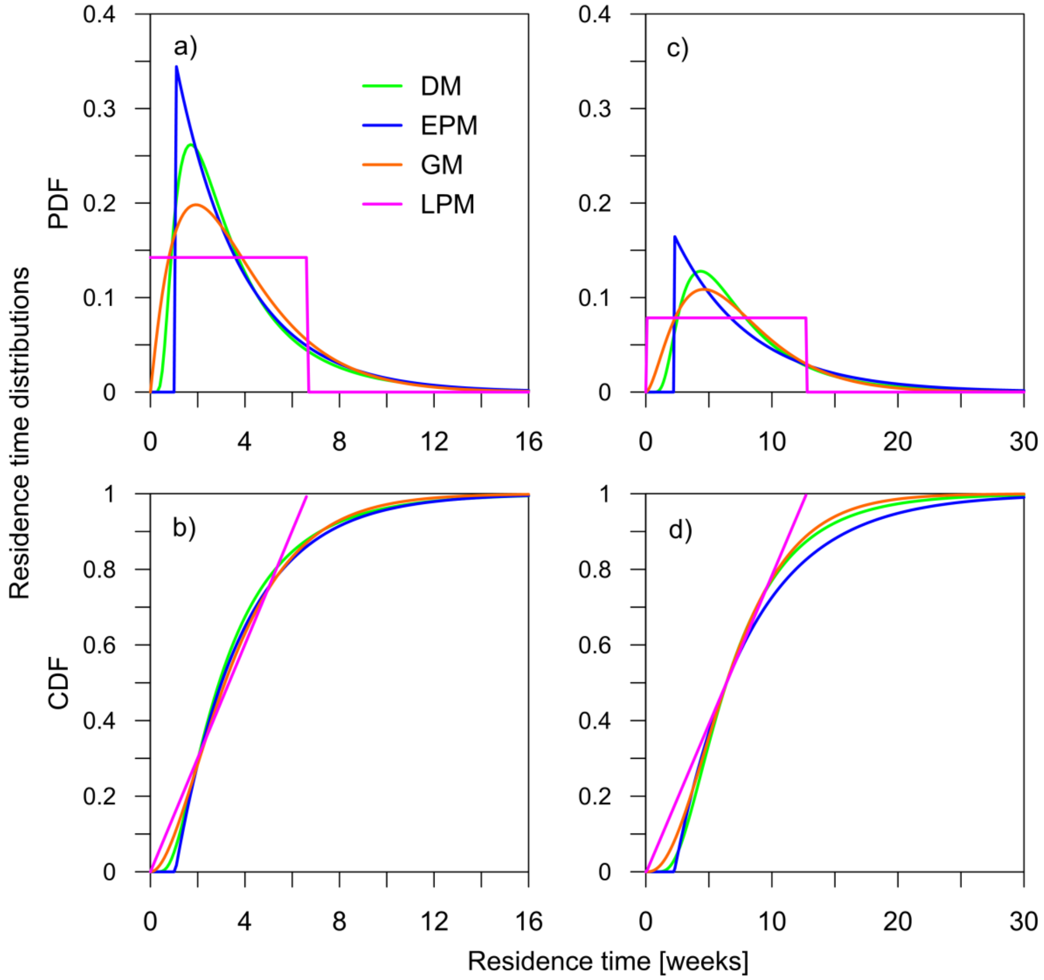


Fig. 8. Comparative characteristic shapes of residence time distribution functions corresponding to the best NSE using four lumped parameter models (DM, EPM, GM and LPM): **(a)** and **(b)** for the soil site C2 located in a pastures land cover; **(c)** and **(d)** for the soil site E2 located in a forest land cover.

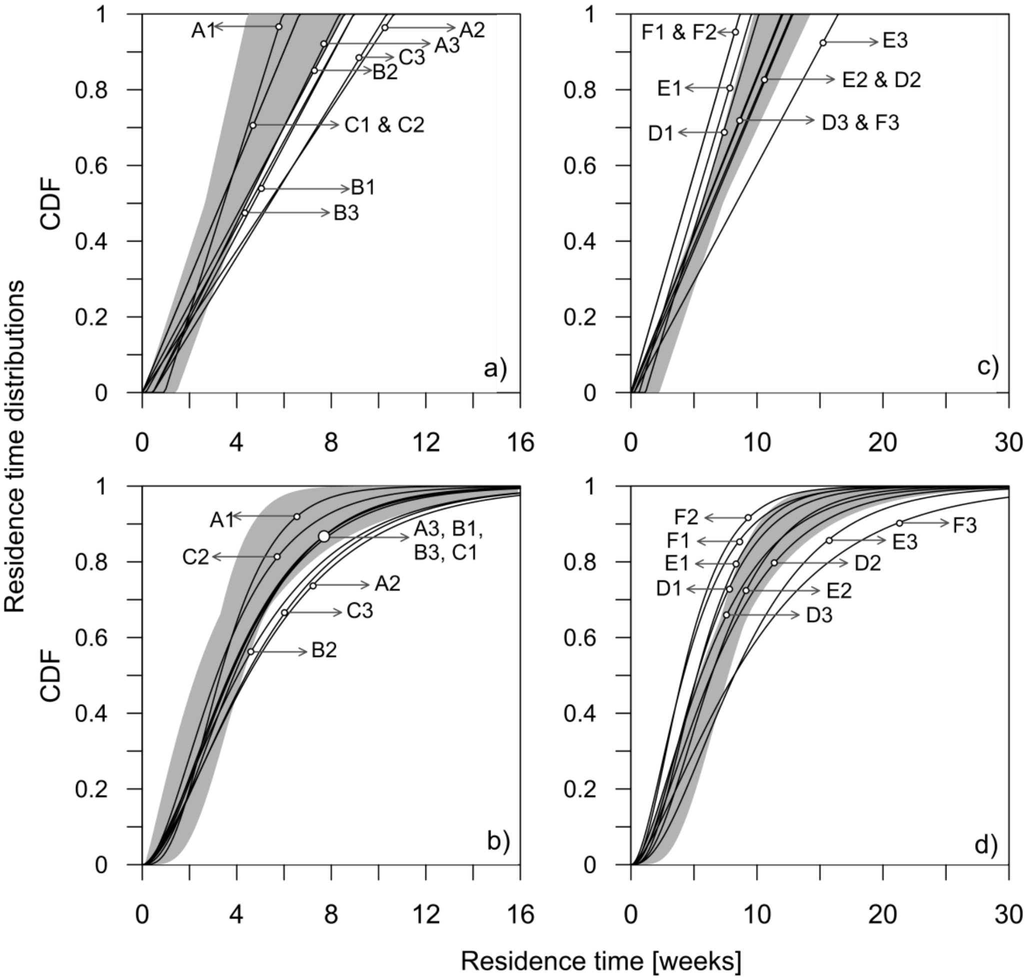


Fig. 9. Comparative results between LPM and GM models of soil water residence time distributions functions corresponding to the best NSE for every sampling site: **(a)** pastures sites using LPM; **(b)** pastures sites using GM; **(c)** forest sites using LPM; **(d)** forest sites using GM. Gray shaded area in each plot corresponds to the range of possible shapes of the distribution function for one of the sampling sites: C2 in sub-plots **(a)** and **(b)**, and E2 in sub-plots **(c)** and **(d)**.

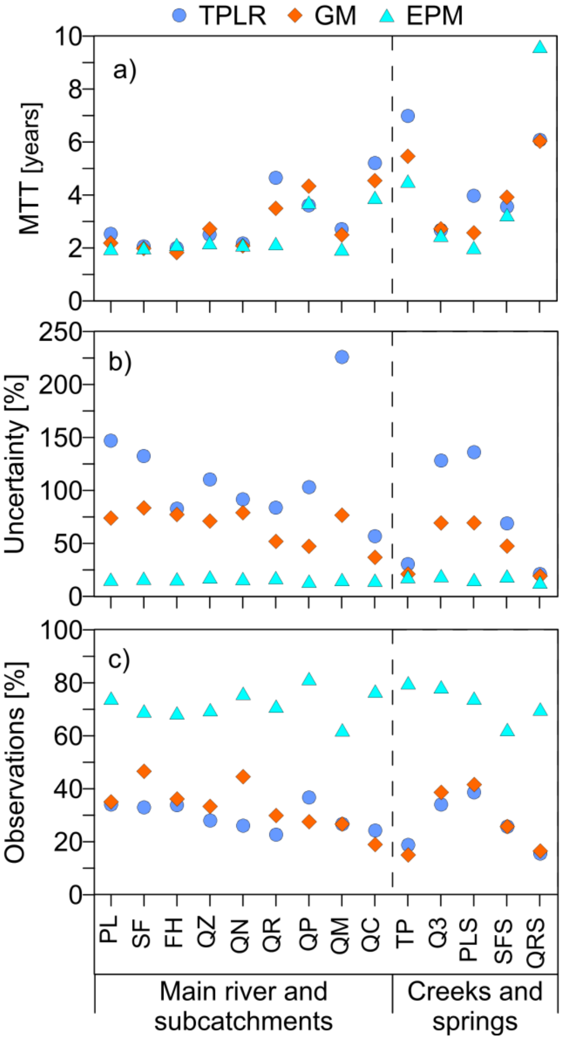


Fig. 10. Intercomparison of models for surface waters and springs according to their: **(a)** estimated mean transit times; **(b)** uncertainty ranges expressed in percentage of its respective MTT estimation; and **(c)** number of observations inside the range of behavioral solutions.

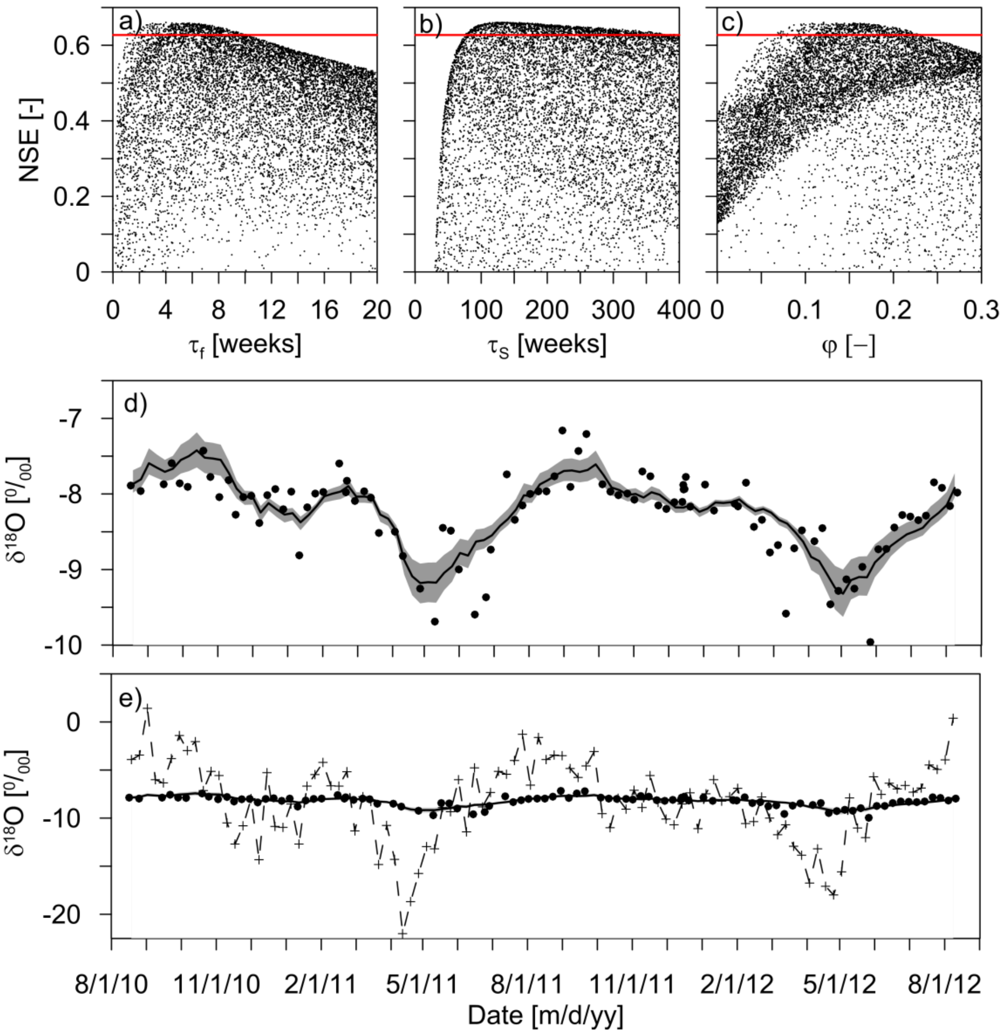


Fig. 11. Uncertainty ranges for outlet stream water (PL site) using a TPLR distribution function. Sub-plots **(a)**, **(b)** and **(c)** show the modeled parameter uncertainties of 10,000 random simulations and the feasible range of behavioral solutions taking a lower limit of 5% from the best solution. Black filled circles in the sub-plots **(d)** and **(e)** represents the observed data, the black line and shaded area depict the best possible solution and its range of variation according to the 5-95% confidence limits of the behavioral solutions shown in sub-plot **(b)**; and the gray dashed line with crosses in sub-plot **(e)** represents the weekly rainfall variation as input function for the model.

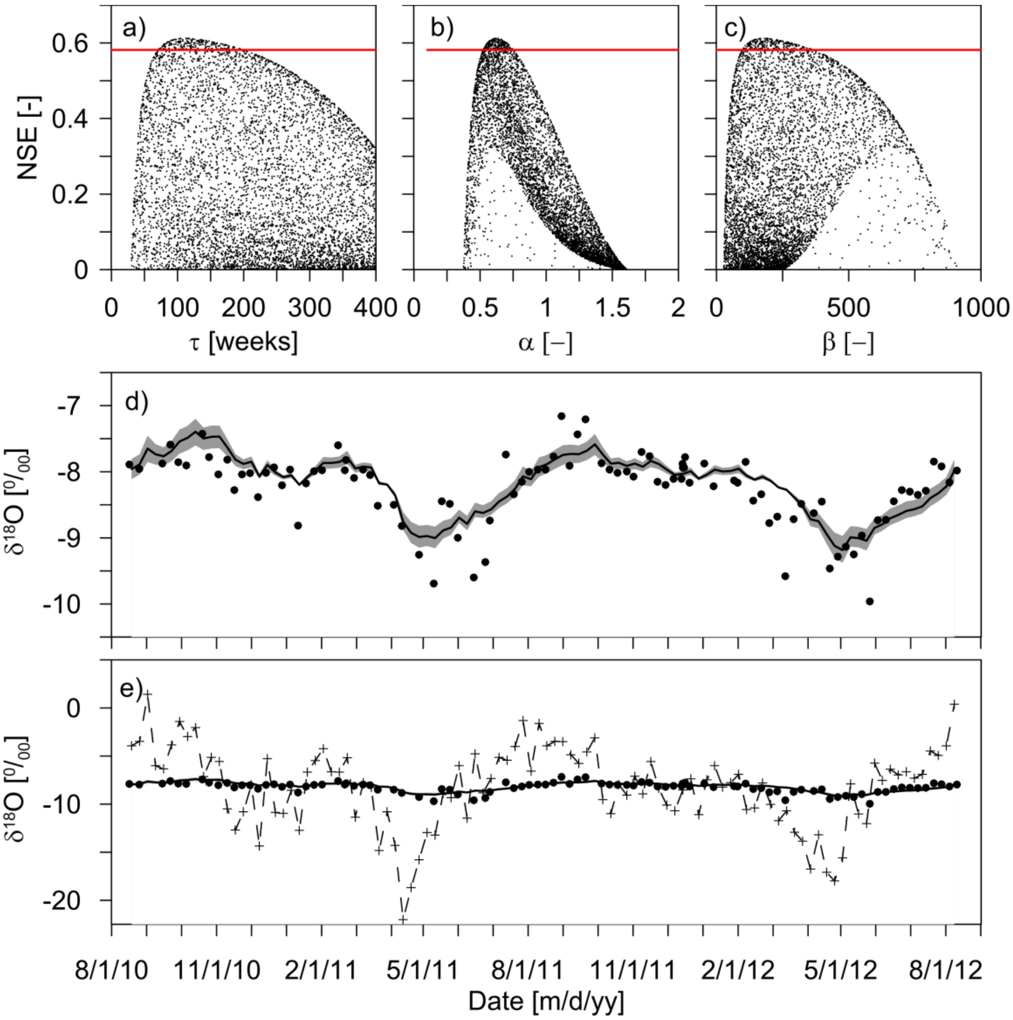


Fig. 12. Uncertainty ranges for outlet stream water (PL site) using a GM distribution function. Sub-plots **(a)**, **(b)** and **(c)** show the modeled parameters uncertainties of 10,000 simulations and the feasible range of behavioral solutions taking a lower limit of 5% from the best solution. Black filled circles in the sub-plots **(d)** and **(e)** represents the observed data, the black line and the shaded area represent the best possible solution and its range of variation according to the 5-95% confidence limits of the behavioral solutions shown in sub-plot **(a)**; and the gray dashed line with crosses in sub-plot **(e)** represents the weekly rainfall variation as input function for the model.

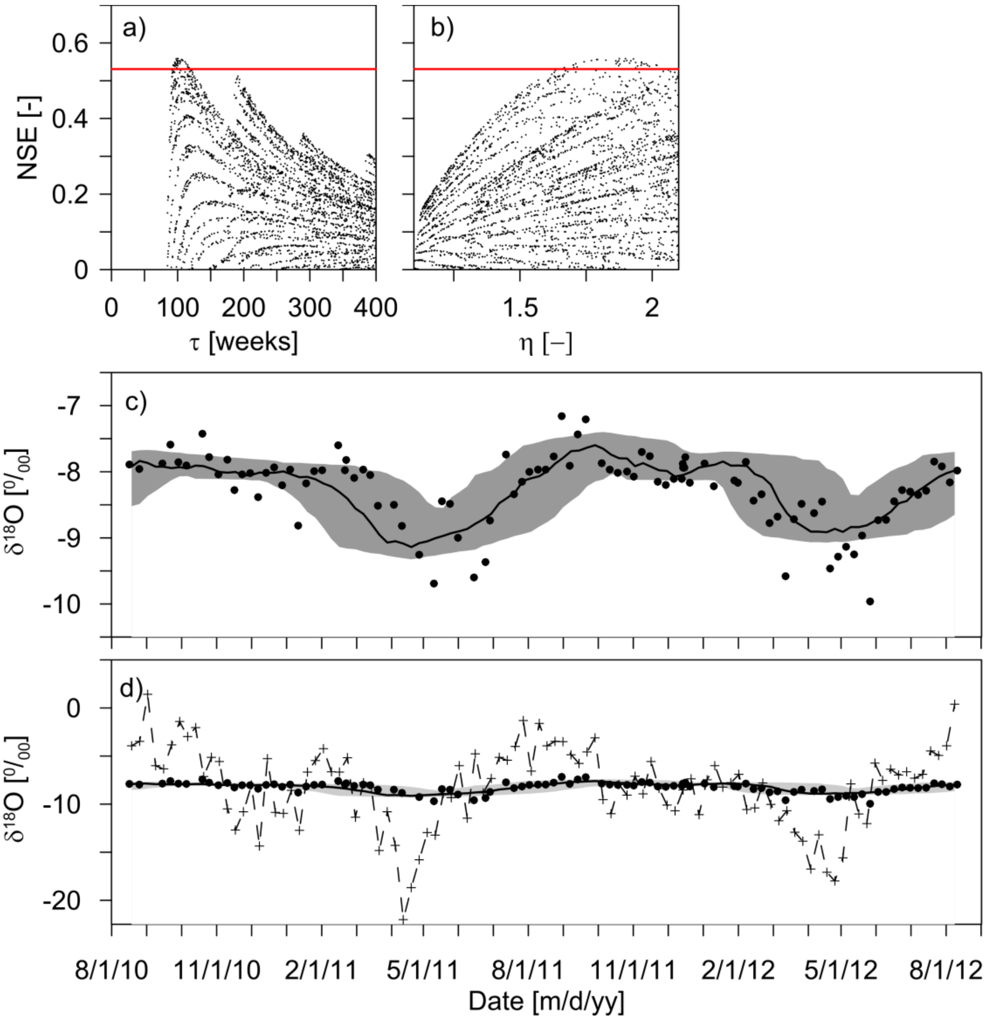


Fig. 13. Uncertainty ranges for outlet stream water (PL site) using an EPM distribution function. Sub-plots **(a)** and **(b)** show the modeled parameters uncertainties of 10,000 simulations and the feasible range of behavioral solutions taking a lower limit of 5% from the best solution. Black filled circles in the sub-plots **(c)** and **(d)** represent the observed data, the black line and the shaded area represent the best possible solution and its range of variation according the 5-95% confidence limits of the behavioral solutions shown in sub-plot **(a)**; and the gray dashed line with crosses in sub-plot **(d)** represents the weekly rainfall variation as input function for the model.

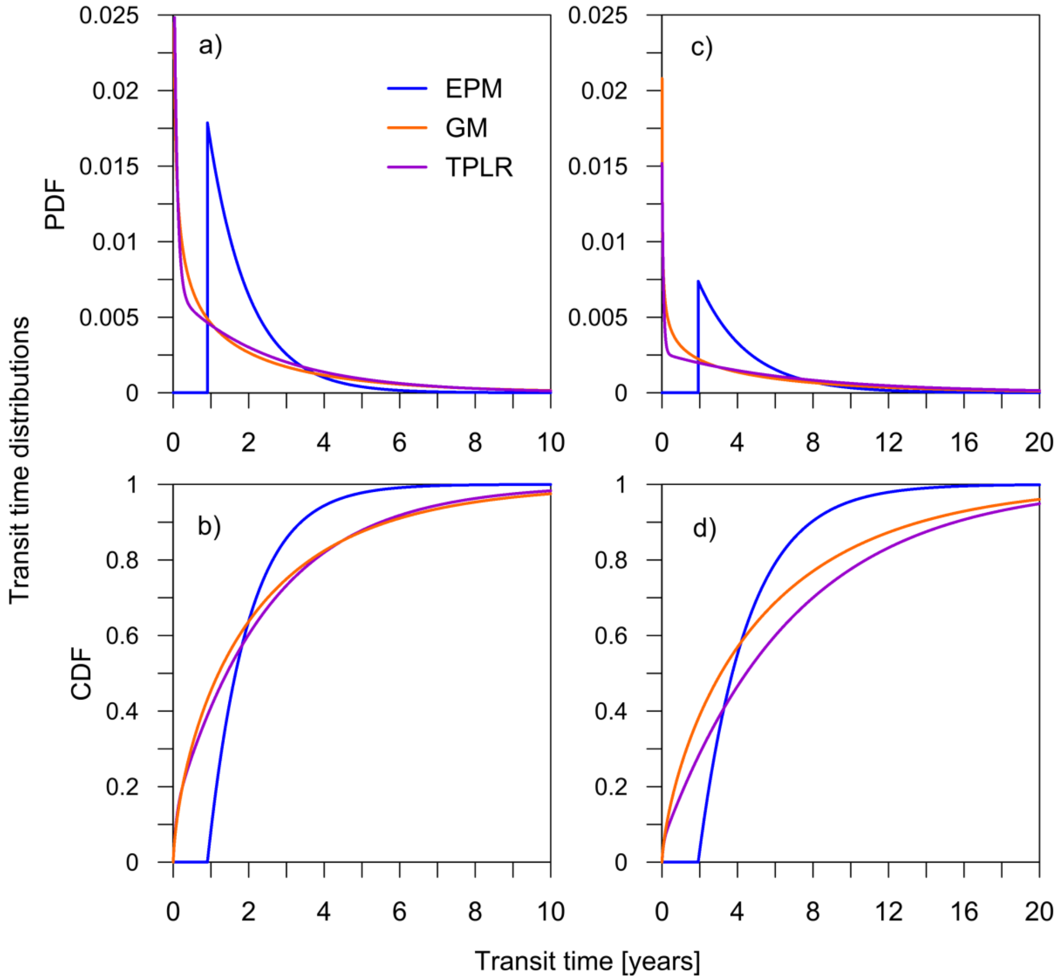


Fig. 14. Comparative characteristic shapes of the transit time distribution functions corresponding to the best NSE using three lumped parameter models (EPM, GM and TPLR): **(a)** and **(b)** for the stream water sampled at the main outlet PL; **(c)** and **(d)** for the small creek TP.

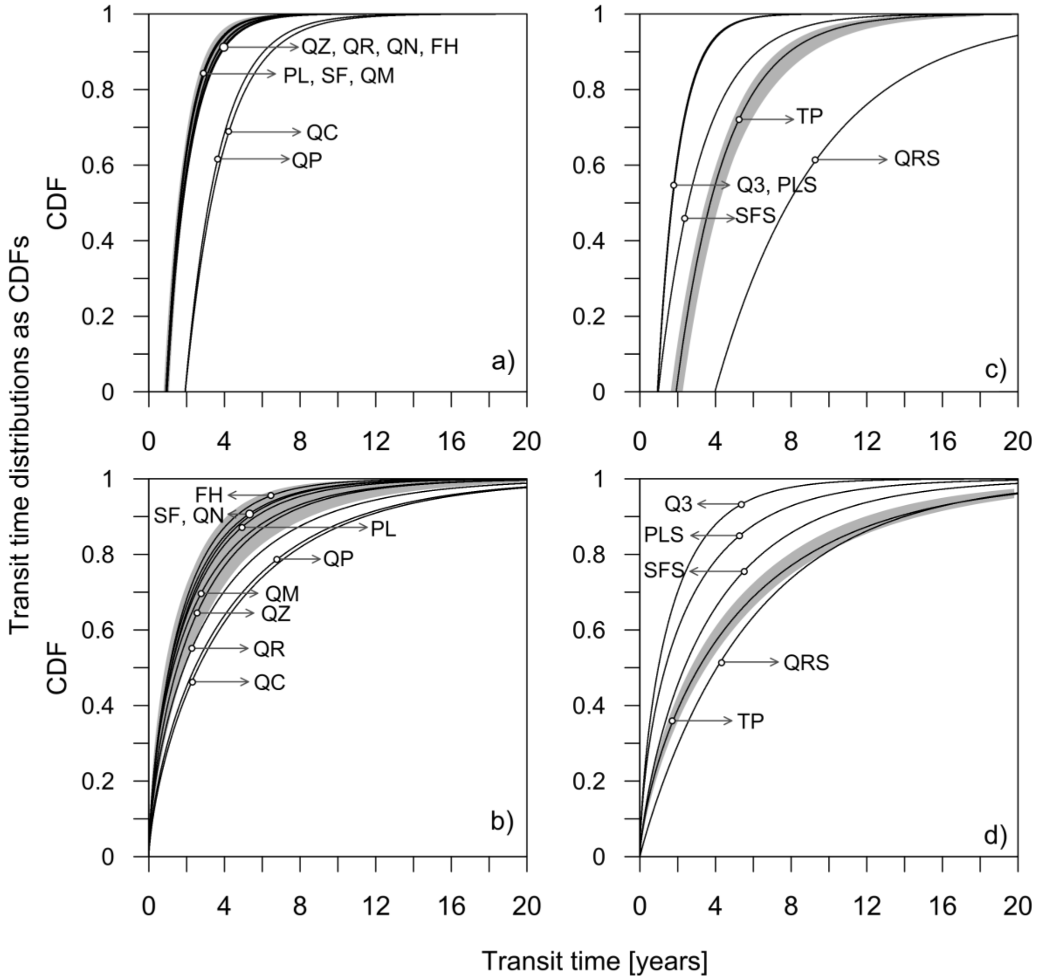


Fig. 15. Comparative results between EPM and GM models of soil water transit time distribution functions corresponding to the best NSE for every sampling site: **(a)** stream water of main outlet and sub-catchments using EPM, and **(b)** using GM; **(c)** spring waters and creeks using LPM, and **(d)** using GM. Gray shaded area in each plot corresponds to the range of possible shapes of the distribution function for one of the sampling sites: the main outlet (PL) in sub-plots **(a)** and **(b)** and TP creek in sub-plots **(c)** and **(d)**.

Annex 1. Predicted results of soil waters for the lumped models: Gamma, Exponential-Piston Flow, Dispersion and Linear.

Table 1. Best predicted results for the Gamma model parameters (τ, α) and corresponding uncertainty ranges.

Site	Mean $^{0}/_{00}$	σ $^{0}/_{00}$	NSE -	RMSE $^{0}/_{00}$	Bias $^{0}/_{00}$	τ weeks	α -
Pastures transect							
A1	-6.74	3.06	0.87	1.33	-0.04	3.6(2.9-4.4)	3.6(2.0-13.4)
A2	-6.72	2.46	0.73	1.72	0.07	5.5(4.5-6.7)	1.8(1.2-3.4)
A3	-7.17	3.18	0.85	1.54	-0.04	4.4(3.5-5.5)	2.0(1.4-8.3)
B1	-6.58	3.01	0.83	1.53	0.27	4.4(3.6-5.3)	2.0(1.4-7.0)
B2	-6.88	2.71	0.80	1.53	0.15	5.0(4.1-6.1)	1.7(1.2-3.5)
B3	-6.72	2.97	0.80	1.51	0.04	4.4(3.6-5.4)	2.1(1.3-5.1)
C1	-6.68	3.15	0.86	1.36	-0.04	3.5(2.7-4.2)	2.4(1.6-9.0)
C2	-7.19	3.11	0.88	1.19	-0.14	3.7(2.9-4.6)	2.1(1.2-5.4)
C3	-6.53	2.56	0.80	1.35	-0.01	5.3(4.5-6.4)	1.9(1.3-3.8)
Forest transect							
D1	-7.26	2.79	0.81	1.35	0.12	6.1(5.1-7.5)	2.4(1.5-4.9)
D2	-7.03	2.35	0.82	1.08	0.03	7.6(6.6-9.2)	1.9(1.3-3.2)
D3	-6.82	2.40	0.82	1.16	-0.02	6.7(5.8-7.9)	1.8(1.2-3.6)
E1	-6.54	2.79	0.82	1.34	0.10	5.9(5.1-6.8)	2.9(1.8-7.1)
E2	-6.52	2.44	0.78	1.37	0.11	7.3(6.4-8.2)	2.7(1.8-5.6)
E3	-6.43	1.97	0.79	1.16	0.02	9.4(8.2-10.7)	2.5(1.8-4.0)
F1	-6.81	2.72	0.90	0.99	-0.06	5.0(4.2-6.1)	1.9(1.3-4.7)
F2	-6.74	2.79	0.90	0.97	-0.29	4.7(3.8-5.7)	2.4(1.4-6.4)
F3	-8.50	1.87	0.69	1.41	-0.41	10.2(8.7-12.5)	1.6(1.2-2.2)

σ= standard deviation, NSE = Nash-Sutcliffe Efficiency, RMSE = Root Mean Square Error

Table 2. Best predicted results for the Exponential Piston flow model parameters (τ, η) and corresponding uncertainty ranges.

Site	Mean	σ	NSE	RMSE	Bias	τ	η
	$^0/_{00}$	$^0/_{00}$	-	$^0/_{00}$	$^0/_{00}$	weeks	-
Pastures transect							
A1	-6.88	3.00	0.86	1.38	-0.18	3.7(2.9-4.8)	1.40(1.28-1.59)
A2	-6.91	2.53	0.73	1.71	-0.12	5.7(4.7-7.2)	1.26(1.18-1.34)
A3	-7.31	3.23	0.84	1.57	-0.18	4.5(3.5-5.7)	1.33(1.21-1.48)
B1	-6.99	3.11	0.84	1.49	-0.14	5.1(4.0-6.3)	1.33(1.24-1.43)
B2	-7.14	2.82	0.80	1.51	-0.10	5.5(4.4-6.9)	1.28(1.20-1.36)
B3	-6.82	2.94	0.80	1.52	-0.05	4.7(3.8-6.0)	1.30(1.21-1.40)
C1	-6.75	3.15	0.86	1.38	-0.10	3.7(2.9-4.7)	1.40(1.29-1.57)
C2	-7.15	3.09	0.88	1.18	-0.09	3.8(3.0-5.0)	1.36(1.25-1.51)
C3	-6.59	2.54	0.80	1.36	-0.08	5.5(4.5-6.9)	1.25(1.17-1.33)
Forest transect							
D1	-7.40	2.79	0.83	1.28	-0.02	7.0(5.6-8.6)	1.44(1.33-1.56)
D2	-7.06	2.36	0.82	1.11	0.00	8.5(7.2-10.2)	1.32(1.26-1.39)
D3	-6.84	2.36	0.81	1.19	-0.05	7.2(6.0-8.9)	1.18(1.12-1.23)
E1	-6.67	2.75	0.82	1.34	-0.03	6.6(5.5-8.1)	1.47(1.37-1.63)
E2	-6.69	2.38	0.77	1.40	-0.07	8.2(6.9-9.8)	1.37(1.29-1.46)
E3	-6.54	1.99	0.78	1.21	-0.09	10.3(8.9-12.1)	1.45(1.32-1.58)
F1	-6.88	2.73	0.90	0.97	-0.13	5.2(4.2-6.6)	1.27(1.19-1.36)
F2	-6.61	2.65	0.91	0.95	-0.16	4.8(3.8-6.1)	1.25(1.16-1.37)
F3	-8.14	2.02	0.74	1.30	-0.05	9.6(8.4-11.7)	1.37(1.22-1.47)

σ= standard deviation, NSE = Nash-Sutcliffe Efficiency, RMSE = Root Mean Square Error

Table 3. Best predicted results for the Dispersion model parameters (τ, D_p) and corresponding uncertainty ranges.

Site	Mean $^0/_{00}$	σ $^0/_{00}$	NSE -	RMSE $^0/_{00}$	Bias $^0/_{00}$	τ weeks	D_p -
Pastures transect							
A1	-6.77	3.10	0.86	1.33	-0.07	3.6(3.1-4.7)	0.13(0.07-0.53)
A2	-6.63	2.45	0.72	1.76	0.16	5.7(4.7-7.8)	0.33(0.22-0.99)
A3	-7.15	3.23	0.84	1.59	-0.02	4.5(3.6-6.1)	0.22(0.11-0.97)
B1	-6.56	3.08	0.82	1.55	0.29	4.6(3.8-5.8)	0.21(0.11-0.78)
B2	-6.79	2.77	0.79	1.57	0.24	5.1(4.3-7.2)	0.31(0.21-1.06)
B3	-6.64	2.94	0.80	1.52	0.12	4.5(3.8-6.5)	0.28(0.17-0.87)
C1	-6.61	3.17	0.86	1.37	0.03	3.5(2.9-5.1)	0.19(0.10-0.85)
C2	-7.00	3.06	0.88	1.19	0.06	3.7(3.2-5.8)	0.29(0.17-0.97)
C3	-6.46	2.53	0.79	1.38	0.06	5.5(4.7-7.5)	0.32(0.20-0.84)
Forest transect							
D1	-7.24	2.68	0.83	1.28	0.14	6.7(5.7-9.2)	0.31(0.18-0.64)
D2	-6.99	2.33	0.82	1.08	0.07	8.4(7.2-11.7)	0.34(0.23-0.76)
D3	-6.77	2.40	0.81	1.19	0.03	7.2(6.2-10.0)	0.32(0.19-0.82)
E1	-6.55	2.75	0.82	1.33	0.10	6.3(5.4-7.8)	0.21(0.12-0.46)
E2	-6.51	2.45	0.77	1.39	0.11	7.6(6.8-9.4)	0.20(0.13-0.43)
E3	-6.41	2.00	0.78	1.19	0.03	9.8(8.7-11.8)	0.22(0.15-0.39)
F1	-6.72	2.72	0.90	1.00	0.04	5.2(4.3-7.1)	0.29(0.18-0.83)
F2	-6.66	2.73	0.91	0.95	-0.21	4.8(3.9-6.7)	0.29(0.15-0.73)
F3	-8.49	1.90	0.70	1.39	-0.40	11.6(9.8-14.6)	0.41(0.29-0.75)

σ= standard deviation, NSE = Nash-Sutcliffe Efficiency, RMSE = Root Mean Square Error

Table 4. Best predicted results for the Linear Model parameter (τ) and corresponding uncertainty ranges.

Site	Mean $^0/_{00}$	σ $^0/_{00}$	NSE -	RMSE $^0/_{00}$	Bias $^0/_{00}$	τ weeks
Pastures transect						
A1	-6.85	3.06	0.86	1.37	-0.15	3.5(2.8-4.5)
A2	-6.87	2.63	0.73	1.72	-0.08	5.4(4.5-6.2)
A3	-7.32	3.30	0.86	1.46	-0.19	4.4(3.5-5.2)
B1	-6.89	3.19	0.83	1.52	-0.04	4.3(3.3-4.9)
B2	-7.03	3.02	0.78	1.57	0.00	4.4(3.8-5.2)
B3	-6.77	3.03	0.79	1.54	0.00	4.4(3.4-4.9)
C1	-6.72	3.17	0.84	1.44	-0.07	3.5(2.5-4.1)
C2	-7.10	3.16	0.87	1.27	-0.04	3.5(2.9-4.5)
C3	-6.54	2.71	0.80	1.36	-0.02	4.9(4.4-5.9)
Forest transect						
D1	-7.31	2.91	0.76	1.50	0.07	5.4(4.8-6.2)
D2	-6.97	2.56	0.78	1.19	0.09	6.6(5.9-7.1)
D3	-6.74	2.61	0.80	1.22	0.05	6.0(4.9-6.6)
E1	-6.65	2.84	0.80	1.41	0.00	5.4(4.8-6.1)
E2	-6.64	2.55	0.78	1.37	-0.01	6.4(5.8-7.1)
E3	-6.48	2.14	0.76	1.24	-0.04	8.1(7.3-9.2)
F1	-6.79	2.90	0.89	1.05	-0.03	4.5(4.0-5.5)
F2	-6.52	2.79	0.89	1.03	-0.08	4.6(3.9-5.6)
F3	-8.42	2.37	0.64	1.51	-0.33	7.2(7.1-8.2)

σ= standard deviation, NSE = Nash-Sutcliffe Efficiency, RMSE = Root Mean Square Error

Annex 2. Predicted results of stream, creek and spring waters for the lumped models Gamma and Two Parallel Linear Reservoirs.

Table 1. Best predicted results for the Gamma model parameters (*τ, α*) and corresponding uncertainty ranges.

Site	Mean	σ	NSE	RMSE	Bias	τ	α
	$^{0}/_{00}$	$^{0}/_{00}$	-	$^{0}/_{00}$	$^{0}/_{00}$	yr	-
Stream							
PL	-8.16	0.42	0.61	0.34	0.0909	2.2(1.6-3.2)	0.62(0.55-0.71)
SF	-8.03	0.43	0.62	0.34	0.0836	2.0(1.5-3.1)	0.63(0.56-0.72)
Streamwater tributaries							
FH	-8.21	0.42	0.58	0.36	0.0765	1.8(1.5-2.9)	0.71(0.60-0.78)
QZ	-8.35	0.36	0.58	0.31	0.0596	2.7(2.0-3.9)	0.63(0.57-0.72)
QN	-8.21	0.40	0.64	0.30	0.0681	2.1(1.6-3.2)	0.66(0.58-0.75)
QR	-7.86	0.16	0.45	0.35	0.0915	3.5(2.6-4.4)	0.60(0.56-0.67)
QP	-8.04	0.26	0.54	0.23	0.0240	4.3(3.3-5.4)	0.65(0.62-0.73)
QM	-7.74	0.44	0.60	0.37	0.0706	2.5(1.8-3.7)	0.57(0.51-0.64)
QC	-7.57	0.24	0.53	0.21	0.0508	4.5(3.7-5.4)	0.68(0.64-0.74)
Creeks							
TP	-7.63	0.20	0.45	0.18	0.0249	5.5(4.8-5.9)	0.68(0.64-0.73)
Q3	-7.66	0.45	0.68	0.30	0.0126	1.7(1.3-2.8)	0.65(0.55-0.74)
Springs							
PLS	-7.94	0.43	0.69	0.28	0.0945	2.6(1.9-3.7)	0.58(0.53-0.66)
SFS	-7.57	0.23	0.56	0.19	0.0432	3.9(3.0-4.9)	0.74(0.68-0.81)
QRS	-7.78	0.09	0.25	0.14	0.0146	6.0(5.3-6.5)	0.94(0.91-1.00)

σ= standard deviation, NSE = Nash-Sutcliffe Efficiency, RMSE = Root Mean Square Error

Table 2. Best predicted results for the Two Parallel Reservoir model parameters (τ_s, φ) and corresponding uncertainty ranges. A fixed range from 4 to 4.5 weeks was maintained for τ_f in all cases.

Site	Mean $^0/_{00}$	σ $^0/_{00}$	NSE -	RMSE $^0/_{00}$	Bias $^0/_{00}$	τ_s yr	φ -
Stream							
PL	-8.24	0.44	0.66	0.32	0.0176	2.5(1.9-5.6)	0.622(0.554-0.706)
SF	-8.10	0.44	0.64	0.33	0.0117	2.1(1.6-4.3)	0.631(0.555-0.721)
Streamwater tributaries							
FH	-8.24	0.43	0.60	0.34	0.0383	2.0(1.5-3.1)	0.708(0.605-0.782)
QZ	-8.41	0.37	0.60	0.30	0.0000	2.5(1.9-4.7)	0.632(0.570-0.717)
QN	-8.27	0.41	0.67	0.29	0.0141	2.2(1.6-3.6)	0.660(0.582-0.749)
QR	-7.93	0.23	0.52	0.33	0.0280	4.6(3.1-7.0)	0.603(0.562-0.672)
QP	-8.09	0.24	0.54	0.23	-0.0207	3.6(2.8-6.5)	0.653(0.620-0.728)
QM	-7.84	0.48	0.63	0.36	-0.0307	2.7(2.1-8.3)	0.565(0.506-0.636)
QC	-7.60	0.23	0.59	0.19	0.0183	5.2(3.8-6.8)	0.685(0.642-0.741)
Creeks							
TP	-7.65	0.17	0.51	0.17	0.0054	7.0(5.7-7.8)	0.680(0.642-0.726)
Q3	-7.71	0.43	0.67	0.31	-0.0428	1.7(1.3-2.7)	0.648(0.554-0.742)
Springs							
PLS	-8.04	0.44	0.78	0.24	-0.0045	4.0(2.6-8.0)	0.581(0.526-0.659)
SFS	-7.58	0.23	0.59	0.19	0.0255	3.6(2.8-5.2)	0.735(0.684-0.813)
QRS	-7.79	0.09	0.25	0.14	0.0119	6.1(5.3-6.6)	0.945(0.911-0.997)

σ= standard deviation, NSE = Nash-Sutcliffe Efficiency, RMSE = Root Mean Square Error