Spatially Distributed Characterization of Soil Dynamics Using Travel-Time Distributions

Falk Heße[1], Matthias Zink[1], Rohini Kumar[1], Luis Samaniego[1], and Sabine Attinger[1,2]

[1]Department of Computational Hydrosystems, UFZ – Helmholtz Centre for Environmental Research, Leipzig, Germany.
[2]Institute of Geosciences, Faculty for Chemistry and Earth Sciences, Friedrich-Schiller-University Jena, Jena, Germany.

Correspondence to: Falk Heße(falk.hesse@ufz.de)

Abstract.

Travel-time distributions are a comprehensive tool for the characterization of hydrological system dynamics. Unlike streamflow hydrographs, they describe the movement and storage of water inside and through the hydrological system. Until recently, studies using such travel-time distributions have generally either been applied to simple (artificial toy) models or to real-world catchments using available time series, e.g. stable isotopes. Whereas the former are limited in their realism, the latter are limited in their use of available data sets. In our study, we employ a middle ground by using the mesoscale Hydrological Model (mHM) and apply it to a catchment in Central Germany. Being able to draw on multiple large data sets for calibration and verification, we generate a large array of spatially distributed states and fluxes. These hydrological outputs are then used to compute the travel-time distributions for every grid cell in the modeling domain. A statistical analysis shows the general soundness of the upscaling scheme employed in mHM and reveal precipitation, saturated soil moisture and potential evapotranspiration as important predictors for explaining the spatial heterogeneity of mean travel times. In addition, we demonstrate and discuss the high information content of mean travel times for characterization of internal hydrological processes.

Keywords

Travel-time distributions,mesoscale Hydrological Model (mHM), multiscale modeling, catchment hydrology, stochastic hydrology, model calibration

1 Introduction

The description of storage and transport of both water and dissolved contaminants in catchments is a challenging subject due to the high heterogeneity of the subsurface properties that govern their fate (*Dagan*, 1989). This heterogeneity, combined with a limited knowledge about the subsurface, results in high degrees of uncertainty. As a result, stochastic methods are often applied, where the relevant processes are modeled as being random (*Dagan*, 1986; *Rubin*, 2003). Amongst these methods, a powerful tool is the use of travel-time distributions (TTDs), where storage and transport inside the catchment are modeled from a Lagrangian perspective (*Rinaldo and Marani*, 1987; *Rinaldo et al.*, 1989). This means that the catchment itself

or meaningful parts of it is treated as a control volume (CV). The spatially complex array of different flow paths inside such a CV is consequently ignored and only inlet and outlet fluxes are used for the analysis (*Botter et al.*, 2010; *Rinaldo et al.*, 2011; *Botter*, 2012). This observation-based description of catchment dynamics makes TTDs a very robust tool. Although the application of TTDs goes back many decades (*Danckwerts*, 1953; *Niemi*, 1977), recent developments have strongly improved their

5 theoretical foundations turning them into a versatile and coherent tool to characterize catchment dynamics (*Bertuzzo et al.*, 2013; *Benettin et al.*, 2015; *Rinaldo et al.*, 2015; *Porporato and Calabrese*, 2015). Owing to this progress, *McMillan et al.* (2012) and *McDonnell and Beven* (2014) have opined that TTDs should be used routinely for hydrological model calibration, a notion that has been picked up with tremendous speed (*Windhorst et al.*, 2014; *Vereecken et al.*, 2015; *McGuire and McDonnell*, 2015). Independently but somewhat parallel to that, *Kitanidis* (2015) has recently pointed out, that the key to subsurface character-

10 ization is to use all available information. From this information-centered perspective, using TTDs have several advantages. First, the travel-time behavior is controlled by different factors than the hydrograph response. Whereas the latter is relating rainfall–runoff events the former is relating rainfall–runoff water (*McDonnell and Beven*, 2014; *Birkel and Soulsby*, 2015). Second, spatially distributed tracer experiments may dramatically increase the information content available for catchment characterization (*Birkel and Soulsby*, 2015).

15 This range of advantages have lead to a steady increase in both theoretical and practical studies using TTDs for the description of catchment dynamics. Such studies can roughly be divided into two major categories. First, theoretical studies, where the catchments is represented by very simple computational models (*Rinaldo et al.*, 2006; *Duffy*, 2010; *Botter et al.*, 2010; *van der Velde et al.*, 2012; *Benettin et al.*, 2015; *Porporato and Calabrese*, 2015). Such studies allow a thorough and detailed analysis of the involved processes and their interdependence but may suffer from an oversimplified model setup for

20 in- and outflux generation. In the second category we find field studies using data from real-world sites (*McGuire et al.*, 2005; *Cardenas*, 2007; *Broxton et al.*, 2009; *Tetzlaff et al.*, 2011; *Dunn et al.*, 2012; *Hrachowitz et al.*, 2013, 2015; *Harman*, 2015). Here the advantage is that the data used for the analysis do not suffer from model errors or other conceptual limitations. However, such data are generally limited in amount (e.g. tracer or isotope time series limited to a few years only, although *Hrachowitz et al.* (2009) used time series of up to 17 years) and variety (only a limited number of data types are available).

25 As a result, such studies might fail to find long-term trends, establish connections between travel-time behavior and specific catchment properties or to investigate the impact of certain hydraulic regimes that are only rarely occurring (e.g. drought or extremely rainy months).

In this study we therefore attempt to close this gap by using a large array of fluxes and states generated through the well constrained mesoscale Hydrological Model (mHM), a spatially distributed hydrological model (*Samaniego et al.*, 2010a;

30 *Kumar et al.*, 2013a). Using detailed data of precipitation, land cover, morphology and soil type as inputs, mHM is able to provide a continuous simulations of spatially distributed fluxes (e.g. recharge or evapotranspiration) and states (e.g. soil moisture) as outputs.

As a case study, we use a ca. $1000\ km^2$, catchment in Central Germany for which detailed morphological and climatological data are available to parametrize mHM. In addition, the chosen catchment is the location of the Hainich Critical Zone

35 Exploratory, a comprehensive monitoring network used within the Collaborative Research Center AquaDiva (*Küsel et al.*,

2016). AquaDiva seeks to elucidate the critical role of water fluxes connecting surface conditions with biogeochemical functions in the subsurface. One of the goal of this project is to understand how far signal of surface properties, like landcover or land management, can be traced into the subsurface water and solute dynamics. Spatially explicit travel-time distributions are the perfect analytical tool to investigate such questions.

5   By virtue of using the modeled data of mHM, we are able to take a mediating position between the two aforementioned categories, i.e. simple (toy) models and data-limited real-world field studies. Using the process based hydrological model mHM, we aim to describe the storage and release of water with good physical plausibility and consistency, therefore avoiding the artificial nature of simplified models. On the other hand, a large set of continuous simulations available from mHM with a high spatial and temporal resolution avoids limitations of field studies. These advantages facilitate to address several

10   types of investigation. First, it allows for a comprehensive description of the flow and transport dynamics taking place in the catchment. The spatial distribution of such dynamics can then be related to e.g. land cover and physical properties of the soil as well as to driving forces like precipitation to determine dominant predictive factors. Apart from the general relevance of establishing such a relationship, we anticipate important ramifications for the application of such results in the aforementioned Collaborative Research Center. In addition, it allows to investigate how certain parametrizations of the mHM

15   model are related to the travel-time behavior of the catchment. This opens the way for a more robust model calibration of hydrological models using additional datasets (*McDonnell and Beven*, 2014; *Birkel and Soulsby*, 2015; *Kitanidis*, 2015).

To present our results on these questions, the rest of the paper is organized as follows: In Section 2 we describe the numerical and analytical tools used in this study. Thus comprises the framework of travel-time distributions, as applied in this study, as well as the relevant features of mHM. In Section 3, we present the results of our study and demonstrate how

20   they relate to the questions raised above. Finally, in Section 4 we summarize our main findings in light these questions and draw some conclusions.

2 Methods

In the following, we provide a short overview of the analytical and numerical tools and methods used in this study. We start by introducing the concept of travel-time distributions. In the following, we use the nomenclature as given by *Benettin et al.*

25   (2015) and the theoretical framework by *Botter et al.* (2010). In addition to that, we give a short overview of the numerical model (mHM) which was used for the calculation of the states and fluxes. Finally, we introduce the catchment used in our study.

2.1 Travel-time distributions for a single control volume

Travel-time distributions are a stochastic description of the dynamic of a water parcel moving through a given control volume

30   (CV). The definition of such a control volume for real-world situation is often arbitrary to some extent (see e.g. the schematic in Figure 1). Within the context of this study, we used a spatially distributed model where the catchment is partitioned in

regular grid cells (for more details see Section 2.2 below). Consequently, the boundaries of our CV were given by the grid cells of the model.

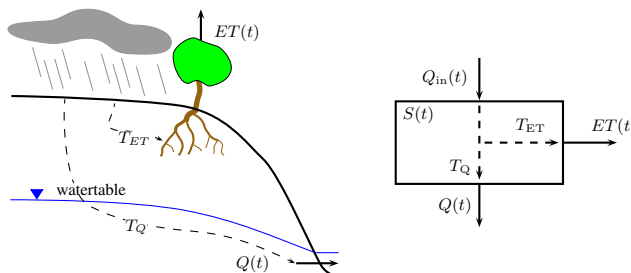



Figure 1. Water movement inside a hillslope (physical schematic on the left and conceptual schematic on the right).

Now, given that such a CV can be reasonably defined, it is clear that the dynamics of a water parcel is determined by the in- and out-fluxes, that are changing the water content inside it. The time evolution of the water content S inside such a CV

5   is then given by the following balance equation

$$\frac{d}{dt}S(t) = Q_{in}(t) - Q_{out}(t) = J(t) - (ET(t) + Q(t)). \tag{1}$$

Equation (1) is a simple initial-value problem with the in-flux $Q_{in}(t)$ given by the effective precipitation $J(t)$ whereas the out-flux $Q_{out}(t)$ given by evapotranspiration $ET(t)$ and discharge $Q(t)$.

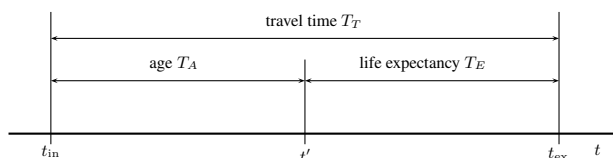



Figure 2. Interpretation of different times associated with dynamics of a water parcel.

To denote the different times involved in the dynamic of a water parcel, we followed the notation of *Benettin et al.* (2015).

10   Chronological time was accordingly denoted with t, whereas the water parcel entered the CV at t_{in} and left at t_{ex}. At any given time t' in between these two points, any water parcel can therefore be characterized by two different properties; its age T_A as well as its (remaining) life expectancy T_E (see Figure 2).

In their paper, *Benettin et al.* (2015) emphasize the two interpretations that originate from these two points of view. Age is a backwards concept referring to the time past since beginning. The associated travel time distribution is therefore called the

15   backward TTD. The concept of backward TTDs is of particular interest for the characterization of e.g. a water sample, since its composition is determined by the age of the water in the CV. Life expectancy, on the other hand, is a forward concept since it is referring to the time still left until exit from the CV. The associated travel time distribution is therefore called the

forward TTD. Such forward TTDs are relevant e.g. for tracer test, since the concentration of an ideal tracer at the outlet is given by the TTD of its associated water parcel.

In order to derive the TTDs associated with the forward and backward formulation, *Botter et al.* (2011) presented a derivation using only the states and fluxes inside the CV. They stated for the forward formulation

$$\overrightarrow{p}_Q(T_E, t_{\mathrm{in}}) = \frac{Q(t)}{\theta(t_{\mathrm{in}})S(t)} \exp\left(-\int\limits_{T_E} \frac{Q(t') + ET(t')}{S(t')} dt'\right) \tag{2}$$

with $T_E = t - t_{\mathrm{in}}, t > t_{\mathrm{in}}$, i.e. the time from the moment the water parcel entered the reservoir until now. The function θ in Equation (2) is called the partition function (*Botter et al.*, 2010, 2011) and can be derived using the following formula

$$\theta(t_{\mathrm{in}}) = \int\limits_{t_{\mathrm{in}}}^{\infty} \frac{Q(t)}{S(t)} \exp\left(-\int\limits_{T_E} \frac{Q(t') + ET(t')}{S(t')} dt'\right) dt. \tag{3}$$

This partition function describes the portion of the water parcel, entering the CV at t_{in}, that is contributing eventually to discharge as opposed to evapotranspiration. It is consequently a dimensionless number between 0 and 1.

For the backward formulation, *Botter et al.* (2011) stated the following

$$\overleftarrow{p}_Q(T_A, t_{\mathrm{ex}}) = \frac{J(t)}{S(t)} \exp\left(-\int\limits_{T_A} \frac{Q(t') + ET(t')}{S(t')} dt'\right) \tag{4}$$

with $T_A = t_{\mathrm{ex}} - t, t < t_{\mathrm{ex}}$, i.e. the time from now until the moment the water parcel leaves the reservoir.

Both these formulations determine the travel time of the water leaving as discharge. The TTDs for the water leaving as evapotranspiration can be determined in an analogous way.

2.2 Numerical model

We used a spatially distributed grid-based mesoscale Hydrological Model (mHM; *Samaniego et al.* (2010a); *Kumar et al.* (2013a)) to generate the states and fluxes needed for the TDD analysis described above. The model uses the grid cell as a primary hydrological unit and models the following dominant hydrological process: interception, snow accumulation and melting, root zone soil moisture dynamics, evapotranspiration, surface flow, interflows, recharge and baseflow. The total runoff generated at each grid cell is routed to the neighbouring downstream cell following the river network using the Muskingum-Cunge routing algorithm. Interested reader may refer to (*Samaniego et al.*, 2010a) for further details on the model components. The model code is open source and can be downloaded from www.ufz.de/mhm. The model has been successfully applied to a number of river basins across Germany, USA and Europe (*Samaniego et al.*, 2010a, b; *Kumar et al.*, 2010, 2013a, b; *Samaniego et al.*, 2013; *Livneh et al.*, 2015; *Thober et al.*, 2015; *Rakovec et al.*, 2016).

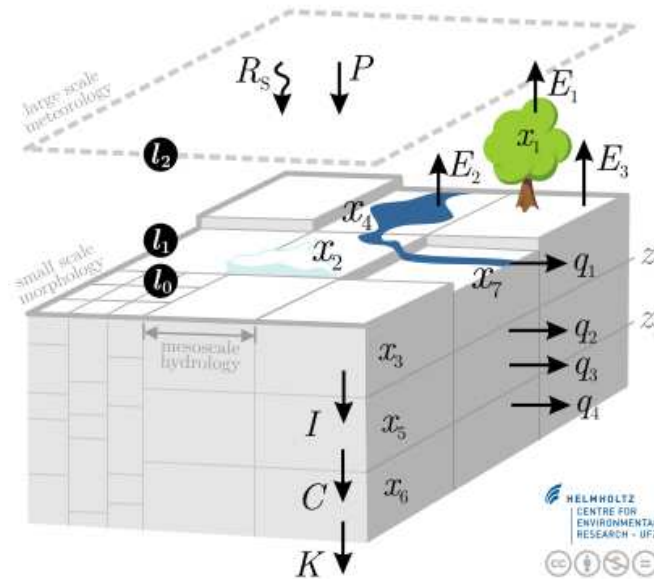



Figure 3. Schematic of the mesocale hydrological model used in the study, depicting the different scales as well as the states and fluxes represented in a single cell.

An important and unique feature of mHM is its Multiscale Parameter Regionalization (MPR), that explicitly accounts for subgrid variability of basin physical characteristics such as terrain, soil, vegetation, and geological properties (*Samaniego et al.*, 2010a; *Kumar et al.*, 2013a). The model considers different level of spatial resolution to better account for spatial heterogeneity of inputs, forcings and the modeled hydrological processes(see schematic in Figure 3). The smallest scale (called l_0 within the mHM nomenclature) is representing morphological factors, like elevation, soil type, land cover etc. On the other hand, meteorological inputs can be represented on a larger scale (called l_2 within mHM). The modeling of the hydrology is done on a third scale (called l_1 within mHM) that can vary depending, e.g. on catchment size or computational resources. Based on the MPR technique, morphological inputs are linked to internal model parameters (e.g. through the use of pedo-transfer functions) and a set of regional coefficients (or global parameters, γ). In a second step, these parameters are upscaled to the resolution of the hydrological processes, i.e. l_1. A set of global parameters (γ) is usually inferred via a calibration procedure. Further details on MPR can be found in *Samaniego et al.* (2010a); *Kumar et al.* (2013a).

Relevant to this study is near-surface and root-zone hydrological process, which are computed using different conceptualizations. In the upmost layer (x_3 in Figure 3) soil moisture is estimated using the infiltration excess approach similar to the HBV model (*Bergström*, 1995), but enhanced to account of multiple soil layers. Within this layer, the water is either percolating into deeper layers or evapotranspirates to the atmosphere. Evapotranspiration is estimated based on potential evapotranspiration, root water uptake and water availability in layer x_3. In the second layer (x_5 in Figure 3) two different

types of interflow take place. Slow interflow IF_s is implemented using a power-law model, whereas fast interflow IF_f is triggered when a threshold value TV is reached, i.e

$$IF_s = \gamma_s x_5^{(1+\alpha)}, \tag{5a}$$

$$IF_f = \gamma_f(x_5 - TV). \tag{5b}$$

In the third level (x_5 in Figure 3) baseflow BF is generated using a simple reservoir model, i.e.

$$BF = \gamma_b x_6. \tag{6}$$

These runoff generation processes are represented at every grid cell of mHM. The sum of interflows and baseflow constitutes the grid specific total runoff which is then routed through a river network. Interested readers may refer to *Samaniego et al.* (2010a) or www.ufz.de/mhm (user manual) for further details on mHM.

2.3 Study area and model set-up

In this study, we used a mesoscale catchment in Central Germany with a drainage area of approximately $1000\ km^2$ to the gauging station at Nägelstedt (see Figure 4). The catchment is the headwaters of the Unstrut river basin, and was selected in this study for its relevance to the Collaborative Research Center AquaDiva (*Küsel et al.*, 2016). The terrain elevation within the catchment ranges between $170m$ and $520m$ with the higher regions in the west and south being the forested hill chain of the Hainich (see Figure 4). The forested area covers approximately 17% of the catchment, while 78% of the area is covered by crop/grassland. The remaining 5% is urban/build up area. The area is characterised by the continental climatic conditions with a mean annual precipitation of approximately $660\ mm$ and a mean temperature of approximately $8\ ^{\circ}$C.

We established mHM over the study catchment and performed numerical simulations on several resolutions ranging from $200\ m$ to $2\ km$. The model was forced using the daily gridded fields of precipitation, air temperature and potential evapotranspiration. The point datasets for the precipitation and air temperature at several raingauges and weather stations located in and around the catchment were acquired from the German Meteorological Service (DWD). These points stations were then interpolated on regular grids using an external drift kriging interpolation procedure wherein the terrain elevation was used as an external drift (*Samaniego et al.*, 2013). The potential evapotranspiration was estimated using the *Hargreaves and Samani* (1985) method. Other datasets required to set-up the model include digital elevation model (DEM) and derived terrain properties like slope, aspect, flow direction, catchment boundary; soil and geological maps were provided by the Federal Institute for Geosciences and Natural Resources (BGR) and meta data such as sand and clay contents, bulk density, horizon depths, dominant hydrogeological classes; CORINE land cover information for the years 1990, 2000 and 2005 available from the European Environment Agency (EEA); and runoff dataset provided by the European Water Agency (EWA) and the Global Runoff Data Centre (GRDC).

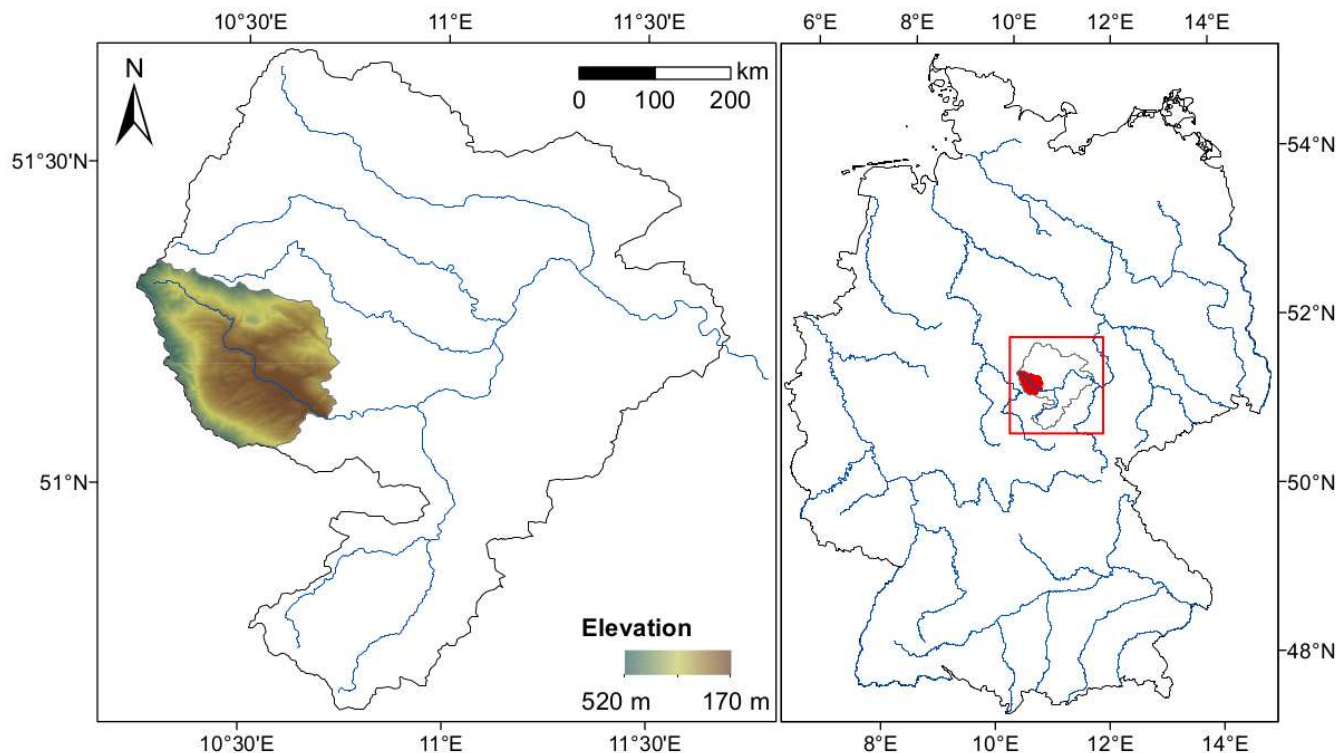



Figure 4. Left: Catchment (highlighted) used in the study shown within the larger confines of the Unstrut catchment (area enclosed by continuous line). The larger rivers of the catchment are shown in blue. The colorbar shows the elevation (in m) of the study catchment. Right: Unstrut catchment within the larger confines of Germany. The axis descriptions denote the latitude and longitude values.

The model simulations were performed for the period 1950 to 2005. The first five years of the data were used to warm-up the model to acquire plausible initial conditions. We therefore discarded the first five years of simulations and the further analyses were performed using model outputs for the period $1955 - 2005$. The model showed quite good performance with $NSE > 0.8$ for the daily discharge simulations at the Nägelstedt station. Other statistics like bias and correlations were also

5   within a satisfactory range.

3   Results and discussion

In this section, we present and discuss the results which have been derived using the methods described above. We will begin in the following by demonstrating and exemplify our general research procedure by virtue of a singly yet representative example.

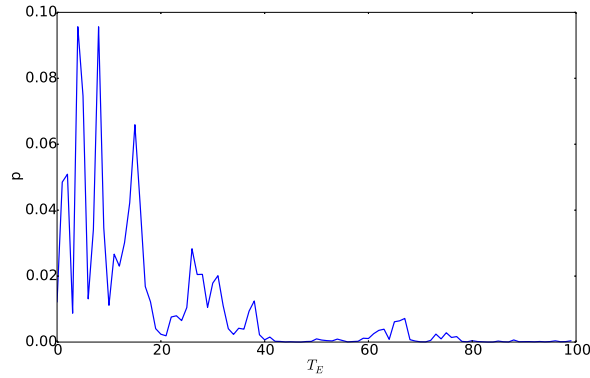
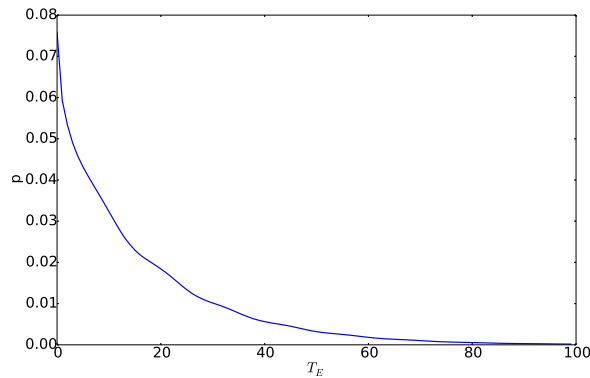
10   The starting point for the derivation of soil travel times were the states and fluxes as computed through mHM. Using the time series of soil moisture, evapotranspiration, interflow and recharge, we used Equation (2) to compute the travel time

(a) time-dependent TTD

(b) stationary TTD

Figure 5. Forward TTD for a single cell within the Nägelstedt catchment. Left hand side: time-dependent TTD derived using Equation (2) for a given t_{in}. Right hand side: stationary TTD derived using Equation (7).

distribution (see Figure 5 a). The shape of these time-dependent distributions varied strongly, depending in particular on rainfall events that triggered the mobilization of older water stored within the soil. Another factor, although not apparent from Figure 5, was the water content, i.e. the state of the soil itself. As been demonstrated by *Niemi* (1977), soil response to rain events is strongly different between wet or dry conditions.

5   To disentangle this event-driven as well as state-dependent effects from other factors that influence the water movement in the soil, we averaged these time dependent distributions. As a result we got the stationary TTDs for every cell

$$\overrightarrow{p}_Q(T_E) = \int \frac{Q(t)}{\theta(t_{in})S(t)} \exp\left(-\int_{T_E} \frac{Q(t') + ET(t')}{S(t')} dt'\right) dt_{in} \tag{7}$$

with $T_E = t - t_{in}, t > t_{in}$.

In all investigated cases, these stationary TTDs could be well approximated by an exponential-like behavior (see Figure 5 b). Such behavior is often assumed to be valid for TTDs in general (*Małoszewski and Zuber*, 1982; *Rodriguez-Iturbe and Valdés*, 1979). Recent works however, have questioned this generalization by emphasizing the time-dependent nature of TTDs (*Duffy*, 2010; *Botter et al.*, 2011). The examples given in Figure 5 exemplify these concerns by illustrating their respective origins. Consequently, we acknowledged the inherent differences between these two TTDs. Furthermore, the study area falls within a humid region with soils being generally wet and rainfall being evenly distributed throughout the year. Under these conditions the assumption of (quasi) stationary TDDs is reasonable (*Tetzlaff et al.*, 2007; *Hrachowitz et al.*, 2009). These stationary TTDs provided the basis for all following analysis, due to allowing the description of the average hydrological response of the catchment. In addition, we also focused on travel-time behavior under specific hydrological regimes, i.e. wet and dry conditions, providing a more detailed understanding of the catchment.

For our statistical analysis, we used these stationary TTDs, which, due to their exponential-like behavior, can be characterized by its expected value τ. We will call this value mean life expectancy (or mean age respectively) in the following. Estimating this value for every mHM cell provided a single measure for the travel-time behavior in the soil without the otherwise dominating impact of single precipitation events (see Figure 6). One feature that became immediately apparent were the long travel times in urban areas (see Figure 6 a). This can be explained by the fact that these areas are largely sealed off, resulting in low infiltration rates and consequently low turnover rates inside the soil. To disentangle this sealing effect from the soil behavior, we discarded cells inside urban regions from our analysis (see Figure 6 b). This allowed us to investigate the interplay between soil properties and travel-time behavior apart from such artificial influences.

3.1 Impact of modeling resolution

Due to its multiscale parameterization, mHM is able to model catchment dynamics at different spatial resolution with a similar set of calibration parameters (see e.g., *Samaniego et al.* (2010a) or *Kumar et al.* (2013a)). Within the context of TTDs, this feature is helpful to investigate the potential influence of StorAge Selection (SAS) functions (*Rinaldo et al.*, 2015; *Harman*, 2015). These function describe the sampling behavior of the catchment with respect to the age distribution of the stored water when discharge is generated. Discharge from a catchment may e.g. be primarily composed of younger or older water or it may show no preference to age whatsoever. SAS functions are therefore a concise mathematical representation of this behavior.

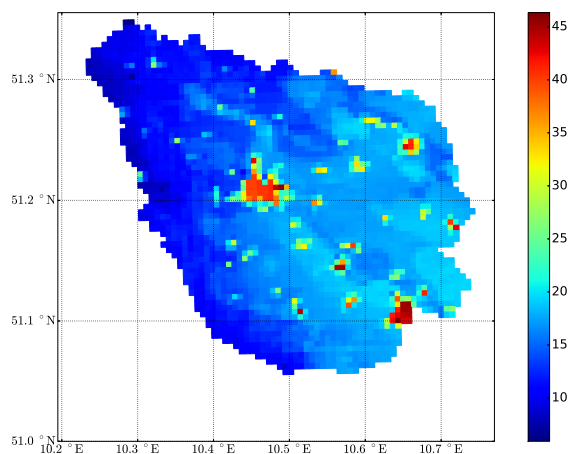
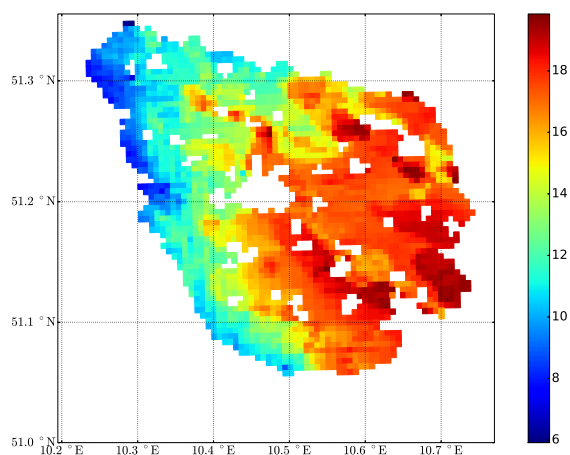
On a physical basis, such preference for different water age should be interpreted as the result of complex mixing processes taking place in the subsurface of the catchment (*Botter*, 2012; *Benettin et al.*, 2013; *van der Velde et al.*, 2012). In order to determine the appropriate SAS function for a given catchment, predictions using different functions could be compared with measurements. Alternatively, the form of the SAS function can be determined by using a physically based catchment model (*Cornaton and Perrochet*, 2006a, b). Within the context of our study, we could not directly infer, which form of a SAS function would be the best choice for our catchment. Instead we calculated the mean life expectancy for our catchment on different scales using the most simple SAS function. Such a function means that discharge is generated proportionally to the age distribution of the stored water.

(a) all mHM-cells

(b) non-urban mHM-cells

Figure 6. Mean life expectancy for the Nägelstedt catchment (see also Figure 1 for comparison) once for all mHM cells (left) and for all non-urban cells.

We hypothesized that the presence of a strong scale dependency of these life expectancy would mean that relevant mixing processes are unaccounted for in our model. On the other side, the lack of such a scale dependency would indicate that mixing processes inside our catchment have only minor relevance. Results from our simulations show no discernible differences in the statistical distribution of mean life expectancy (see Figure 7). Using a smaller resolution had positive effect on statistical estimation procedure due to the increase in data points. In addition, we saw more extreme values due to small scale features

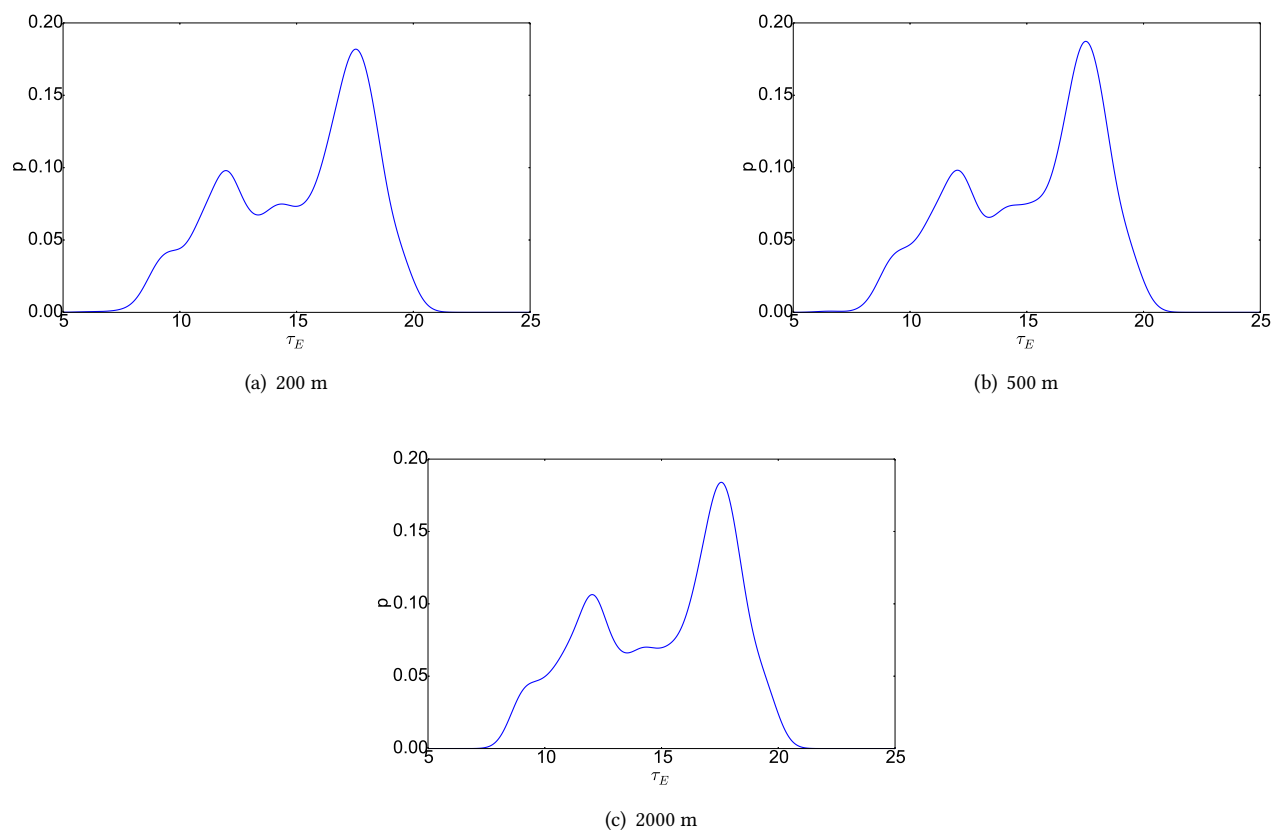



Figure 7. Kernel density estimate of the mean life expectancy (in months) for several grid sizes in the Nägelstedt catchment.

that were smeared out on coarser resolutions. Other than these two changes, we noted only minor changes in the statistics of mean life expectancy. We therefore concluded that, within the limits of the spatial scales tested here, mixing processes inside our catchment have no major impact on mean life expectancy.

Being able to find a good trade-off between computational costs and data amount, we used for our analysis a data set from
5   simulations using a grid size of $500\ m$.

3.2   Statistical analysis of mean life expectancy

The mean life expectancy τ of a water parcel inside a catchment is the result of a complex interplay of morphological and climatological factors. Several recent studies have therefore tried to determine their relative importance under varying condi-tions (*McGuire et al.*, 2005; *Cardenas*, 2007; *Broxton et al.*, 2009; *Tetzlaff et al.*, 2009, 2011). Contrary to these studies where field
10   measurements were used, we used results from computational simulations only. This gave us a much larger dataset, both in time and space, from which we could infer the relative impact of different factors, in particular meteorological (precipitation), land surface (land cover, leaf-are index) and subsurface (soil) properties. Notably, our approach differs from *Hrachowitz et al.*

(2009) such that our analysis is based on model-derived gridded simulations of TDDs as compared to the observation-based basin-wise quantification of TDDs.

In the first step, we determined for every cell the statistical relationship between the mean life expectancy τ and a number of potential predictors like average precipitation, soil depth, soil type or leaf-area index (LAI). Similar to *Hrachowitz et al.*
5 (2009), we used the coefficient of determination R^2 to quantify the strength of the statistical relationship. This quantity equals to one minus the ratio of the remaining variance vs. the total variance of the data themselves. It is therefore a measure of the variance explained by the model (which was always assumed to be linear in our study).

3.2.1 Precipitation

The analysis above showed the strong impact of precipitation on the event-based TTDs (see Figure 5). It is therefore to be
10 expected to exert strong control on the steady-state TTDs as well. In our model two different quantities can be distinguished: first, the precipitation itself as well as, second, the effective precipitation. The latter value is here defined as the water flux that is actually entering the soil, i.e. corrected by surface runoff (through sealing), canopy interception and snowmelt. While the precipitation can be measured with high accuracy, it is the effective precipitation that directly impacts the soil dynamics.

Scatterplot of both data sets against the mean life expectancy show a significant negative correlation between them (see
15 Figure 8). This negative relationship can be explained such that larger amounts of precipitation lead to a 'flushing out' of the soil, therefore resulting in a shorter life expectancy.

3.2.2 Terrain elevation

In our next analysis, we used the physical elevation as a variable for our regression model. The height can simply be derived from the digital elevation model (DEM), which, in mHM, is represented using data obtained from the Shuttle Radar
20 Topography Mission.

Using a scatter plot for visualizing the statistical relationship between mean life expectancy and the DEM showed a negative correlation (see Figure 9 a), i.e. longer life expectancy correlated with lower heights of the terrain, and with a linear coefficient of determination of $R^2 = 0.668$. Since no direct causal connection can be drawn between physical elevation and travel-time behavior, such a high value is indicative of underlying mechanisms. One of these is the aforementioned precipitation, since
25 higher altitudes are correlated with stronger mean precipitation levels (linear coefficient of determination of $R^2 = 0.812$). Performing a multiple linear regression, including precipitation, the linear coefficient of determination for elevation drops sharply to $R^2 = 0.250$. Including additional factors, like potential evapotranspiration and saturated soil moisture (discussed below) lead to a further reduction of the coefficient to $R^2 = 0.041$.

3.2.3 Evapotranspiration

30 Evapotranspiration is directly influencing the form of a TTD (see e.g. Equation 2). Consequently, we anticipated a strong correlation between mean evapotranspiration rates and mean life expectancy.

(a) precipitation

(b) effective precipitation

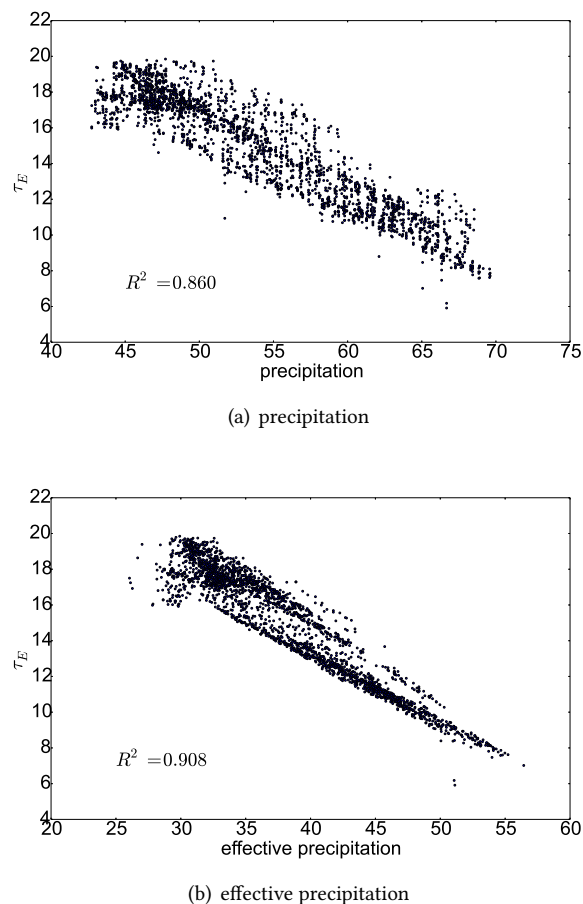



Figure 8. Scatterplot of precipitation and effective precipitation (in mm) vs. mean life expectancy (in months).

With respect to evapotranspiration, two different definitions are typically distinguished: potential evapotranspiration (PET) and actual evapotranspiration (AET). As implied by its name, PET describes the maximum rate of evapotranspiration at a given site. This value is dependent on quantities like solar radiation and temperature that can generally be measured with good accuracy (*Samani*, 2000). Using theoretical models, good estimates can therefore be provided for PET at a given site (*Almorox et al.*, 2015). On the other side, AET is a real quantity that can be measured. In principle, in situ measurements can therefore provide good estimates. In practice, however, exact measurements are hampered by a series of factors (*Wang and Dickinson*, 2012). As a consequence, PET can often be estimated with higher accuracy than AET.

Scatterplots of both PET and AET show a positive correlation between evapotranspiration and mean life expectancy in general (see Figure 10). This correlation is more pronounced for AET with a coefficient of determination of $R^2 = 0.496$ vs. only $R^2 = 0.259$ for PET.

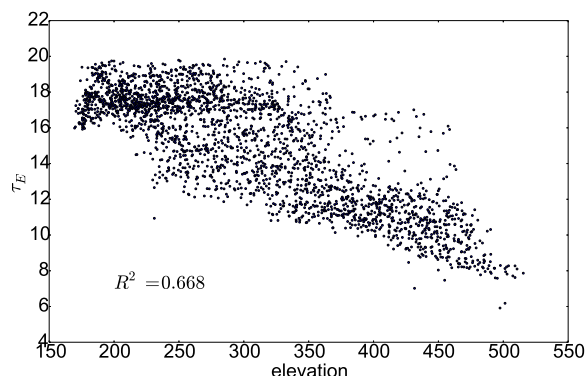



Figure 9. Scatterplot of elevation (in m) vs. mean life expectancy (in months).

Both the strength of the correlation as well as the tendency (AET stronger compared to PET) could be anticipated. The actual evapotranspiration is directly used in Equation (2) for the calculation of TTDs, therefore leading to the comparably stronger impact of this quantity.

As explained above, for real-world situations, better estimates can often be provided for PET. The higher explanatory power of AET has therefore to be balanced with its often less accurate estimate. Depending on the accuracy of measurements of AET, PET estimates may be a better predictor of mean life expectancy.

3.2.4 Land cover properties

Land cover is an important interface controlling the strength of incoming fluxes through artificial and natural sealing. In mHM, three different landcover types are distinguished: forest, grassland and urban area. As explained above, we excluded mHM cells inside urban area from our analysis in order to better focus on the soil properties themselves. To further elucidate possible influence of the remaining landcover types, we separated the catchment into forest and grassland and calculated the mean travel times separately.

Estimating the PDF of the mean life expectancy for both land cover types separately, revealed strong differences between them both in shape of the respective PDF and the range of values (see Figure 11). As shown above, results for the combined data set showed a distinct bimodal behavior (see Figure 7). In contrast to that, the PDF for both land cover types were almost unimodal. The most dominant peaks of every singly PDF coincided with the two peaks of the combined PDF. The behavior of the latter can therefore – to some degree – be considered to be a superposition of the former.

In addition to this classification scheme, mHM uses the leaf area index (LAI) to describe land cover properties. The LAI describes the ratio of the cell that is effectively covered by plant canopy. Due to the already established influence on evapotranspiration (see above), it stands to reason to expect an influence on the mean life expectancy as well. Comparing LAI class and land cover reveals a strong overlap between both (see Figure 11 (a) and Figure 12 (a)). Roughly, forest land cover

(a) PET

(b) AET

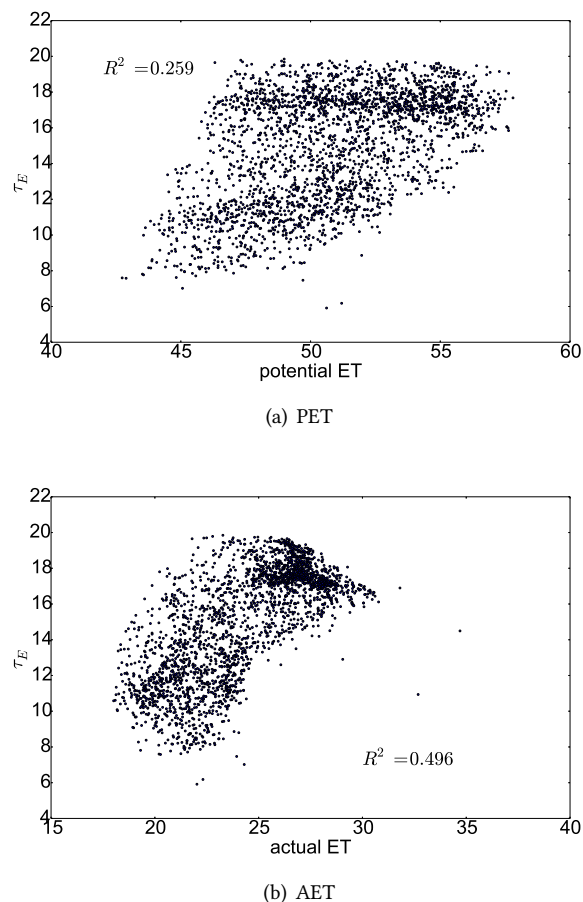



Figure 10. Scatterplot of mean life expectancy (in months) vs. potential as well as actual evapotranspiration (in mm).

corresponds with LAI class $1-4$, urban land cover corresponds with LAI class 5 and grassland corresponds with LAI class $6-10$.

Using the same approach as above, i.e. investigating the mean life expectancy for every LAI class independently, consequently revealed the same overall tendency for LAI classes compared to land cover types (data not shown). This was anticipated due to the aforementioned overlap between the two classification schemes. In addition, we saw little diversity for LAI classes within the same land cover class (data not shown).

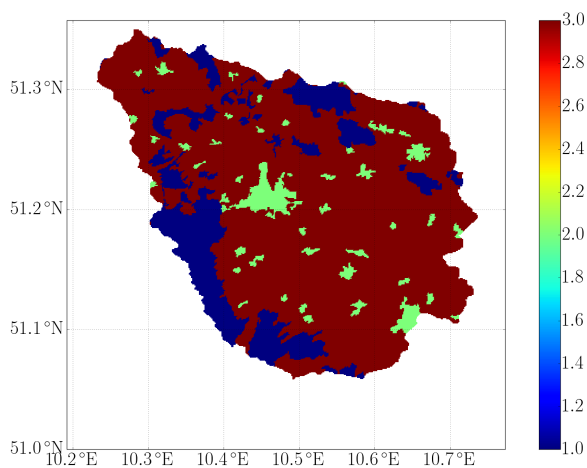
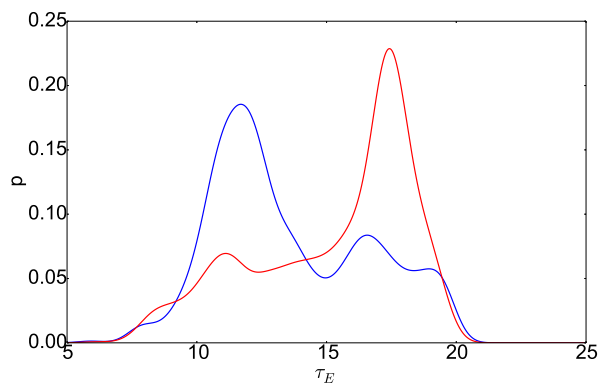
However, this tendency was not present when using the actual leaf-area values associated with every LAI class. These values could be constant over the year e.g. in case of coniferous forest or vary strongly e.g. in case of deciduous forest. To make values from different LAI classes comparable, we averaged the respective values year-wise. A scatter plot of leaf area index vs. mean life expectancy does not show any strong correlation between the two with similar ranges of values being found for almost all LAI values (see Figure 12 (b)). This discrepancy can be explained by the implementation of the

(a) spatial distribution

(b) kernel density estimate

Figure 11. Land cover in the Nägelstedt catchment. Spatial distribution and kernel density estimate of the mean life expectancy (in months). The blue curve represents forest and the red curve represents grassland.

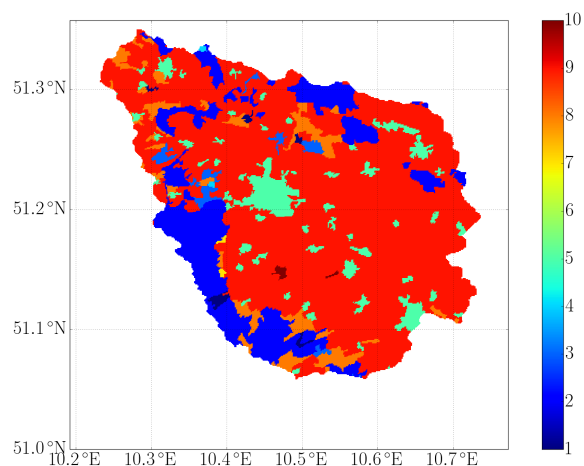
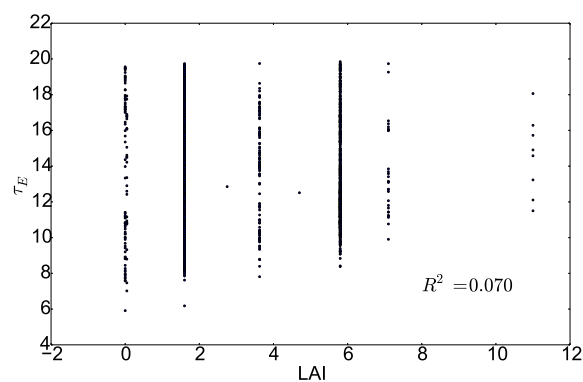
LAI in mHM. In contrast to the land cover type, that is used for the determination of ET processes in the upmost soil layer, LAI values are only used for interception and do consequently not directly influence travel-time behavior. As a result, any possible relationship between LAI and TTDs is therefore biased and conclusions from our results must take into account this limitation critically.

(a) spatial distribution

(b) scatter plot

Figure 12. Leaf area index (LAI) in the Nägelstedt catchment. Spatial distribution and scatter plot of LAI vs. mean life expectancy (in months).

3.2.5 Soil properties

An important input parameter in mHM is the soil type inside every cell. This property is implemented in mHM using the German soil data base Bodenübersichtskarte 1:1.000.000 (BÜK 1000) (*Federal Institute for Geosciences and Natural Resources (BGR)*, 1998).

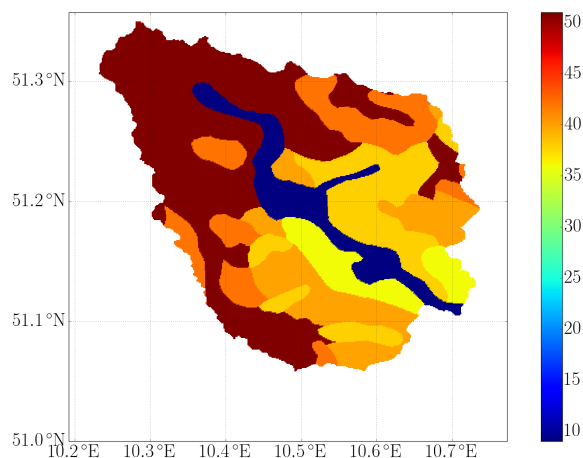
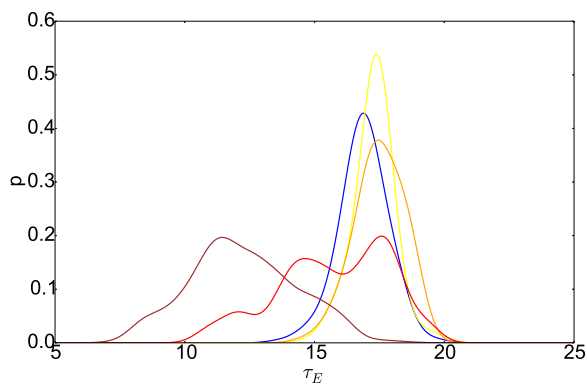
5   Due to this relevance in the model, we anticipated a strong impact of the soil type in a cell on the resulting mean life expectancy. Estimating the PDF of mean travel times for every soil type individually, did indeed show significant differences between them (see Figure 13). Soil classes found in the geographically lower regions of the catchment generally show longer

(a) spatial distribution

(b) kernel density estimate

Figure 13. Soil classes in the Nägelstedt catchment. Spatial distribution and kernel density estimate of the mean life expectancy (in months) for selected soil classes. Blue curve represents soil class 9, yellow curve represents soil class 38, orange curve represents soil class 40, red curve represents soil class 42 and , brown curve represents soil class 51.

mean travel times with a unimodal distribution shape, whereas soil types in the geographically higher regions correspond with generally shorter mean travel times with the shape of the distributions being less regular. This qualitative analysis reveals some overlap with the land cover distributions as well as mean precipitation rates. It is consequently not possible to directly infer causal correlation from statistical correlation.

In addition, the soil class is a symbolic variable, i.e. its values only indicates a certain type of soil but does not directly relate to any numerical quantity associated with this soil type. Consequently, we could not infer any quantitative connection between soil types and resulting travel-time behavior.

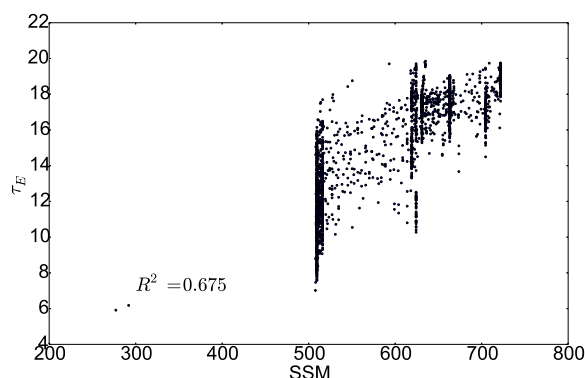



Figure 14. Statistical relationship between mean life expectancy (in months) and saturated soil moisture (in mm).

To address this problem, we used the saturated soil moisture of the soil. This quantity is the amount of pore space per cell
5 that can be potentially filled with water (porosity times the depth of root-zone soil layer). Its value is determined in mHM through pedo-transfer functions using the soil textural information on percentage of sand, clay and bulk density. Comparing these values in every single cell with the mean life expectancy shows a very strong statistical relationship with a coefficient of determination $R^2 = 0.675$.

3.3 Statistical analysis of mean age

10 As described above, the difference between the forward and backward formulations of travel time has long been acknowledged (*Niemi*, 1977) and many studies have investigated their relationship (*Cornaton and Perrochet*, 2006a; *Botter*, 2012; *Benettin et al.*, 2013; *Harman*, 2015; *Benettin et al.*, 2015). Both these formulations are linked by virtue of the so called Niemi relation

$$J(t_{\text{in}})\theta(t_{\text{in}})\overrightarrow{p}(t - t_{\text{in}}|t_{\text{in}}) = Q(t)\overleftarrow{p}(t - t_{\text{in}}|t),$$

which can be derived by considering a water parcel entering the CV at t_{in} and leaving at t. Consequently, mean life
15 expectancy and age only coincide in case of steady-state conditions. As a result, we also investigated the behavior of mean age to elucidate connections and differences between forward and backward formulations for our catchment.

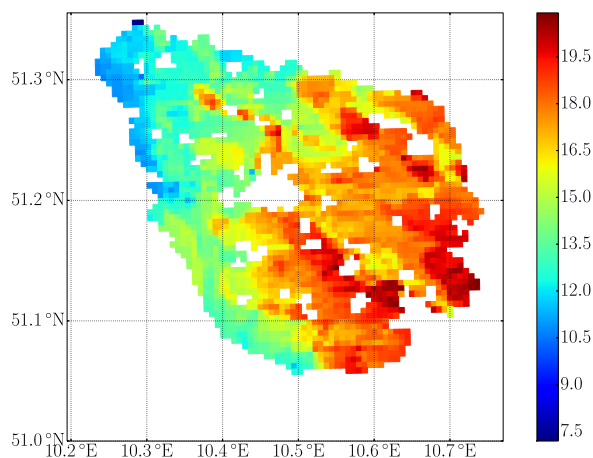
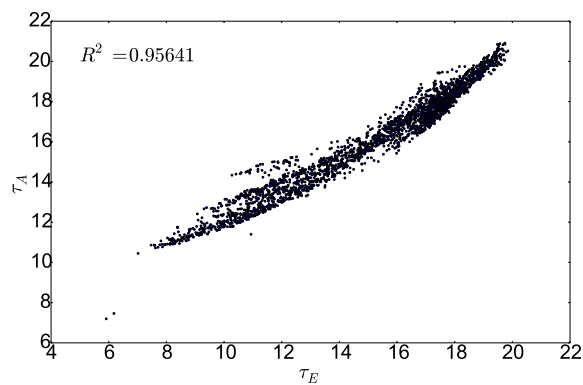
Visually comparing mean age (see Figure 15) and mean life expectancy (see Figure 6 b) in the Nägelstedt catchment showed strong qualitative and quantitative similarities. Accordingly, we also got a very strong statistical relationship between these two quantities with a coefficient of determination of $R^2 = 0.956$. Overall, the relationship was very linear with mean age
20 values falling short of mean life expectancy for both small and large values.

(a) spatial distribution

(b) scatter plot

Figure 15. Mean age in the Nägelstedt catchment.

	Pre	PET	SSM
mean life expectancy	0.860	0.260	0.675
mean age	0.728	0.143	0.711

Table 1. R^2 values for several predictors of mean travel time

Due to the mathematical and physical similarities, such a strong connection was anticipated. To further investigate possible origins of their respective differences, we performed the same statistical analysis for mean age.

	Pre	Pre + SSM	Pre + SSM + PET
mean life expectancy	0.860	0.911	0.913

Table 2. R^2 values of for several regression models of increasing complexity.

To that end, we considered proxy variables that have already been shown to have a considerable impact on travel-time behavior. As demonstrated by the analysis above, these were precipitation (Pre), potential evapotranspiration (PET) and saturated soil moisture (SSM) as proxies for influx, outflux and state respectively. Results showed overall the same trend for mean age and life expectancy with respect to these predictors (see Table 1). Precipitation was the most dominant factor for both quantities with the saturated soil moisture being a close second. This is in contrast to e.g. *Benettin et al.* (2015), who emphasized the role of the outfluxes for the time evolution of both age and life expectancy. In our analysis, we saw that proxy variables for influx and state showed strongest correlations with mean travel-time behavior. On the other hand, PET, which is a good proxy for one of the two outfluxes, showed only moderately strong correlations with said behavior. In case of mean age, this relationship was even weaker compared to the other two (precipitation and saturated soil moisture). Since we could not provide a proxy variable for the other outflux, i.e. discharge, we excluded this quantity from our analysis.

3.4 Joint impact of multiple variables on mean travel times

In the analysis above, the statistical relationship between mean travel-time behavior and a number of variables was presented and discussed. This was done for every variable individually to elucidate its possible impact on mean travel times. In addition to this simple analysis, we also investigated the joint impact of several variables. Such results can be of relevance for prediction, i.e. using a set of variables to predict travel times in a given CV.

To that end, we used the variables that had been shown to have the highest impact individually, i.e. precipitation, saturated soil moisture and potential evapotranspiration, and performed a multiple linear regression. Simple linear regression had already demonstrated that both precipitation and saturated soil moisture could explain a significant amount of the variability contained in the dataset. Combining these factors could therefore improve the predictability even further. We therefore applied Forward Stepwise Selection to generate a series of models with increasing complexity. The fist single-variable model consequently used only precipitation as the variable with the highest single R^2 value. Next, the double-variable model used both precipitation and saturated soil moisture and the most complex three-variable model used precipitation, saturated soil moisture and potential evapotranspiration jointly.

Results for the default case, showed that, compared to using only one variable (precipitation), using two variables for the regression (precipitation and saturated soil moisture) improved the predictability of mean travel times (see Table 2). This was expected since both variables alone provided already high R^2 values. In addition, precipitation and saturated soil moisture did only show moderate correlation ($R^2 = 0.451$), so adding the latter variable added new information to the prediction model. The correlation that existed between precipitation and saturated soil moisture is explained by the an orographic effect,

i.e. hilly regions in the catchment, with typically lower values of saturated soil moisture, also show higher precipitation values. In contrast, using three variables (precipitation, saturated soil moisture and potential evapotranspiration) resulted in almost negligible improvement (see Table 2). This is due to the already lower impact of PET compared to precipitation and saturated soil moisture. In addition, PET showed comparably stronger correlation with both precipitation and saturated

5   soil moisture (data not shown), therefore adding only little new information compared to the other two variables. Such low impact of outgoing fluxes compared to precipitation has already been reported before, for the case of synthetic toy models (*Daly and Porporato*, 2006). Moreover our results agree with the findings of *Hrachowitz et al.* (2009), who also reported similarly strong explanatory power of climatic variables like precipitation as well as soil and land surface properties.

3.5   Impact of hydrological regime on travel-time behavior

10   The analysis above revealed the strong impact of the influx (i.e. precipitation) as well as the state variable (i.e. saturated soil moisture) on the travel-time behavior. To further elucidate their impact, we investigated travel-time behavior independently for different hydrological regimes during the considered period of time, i.e. from 1955 - 2005. To that end we partitioned the available time series into regimes based on soil moisture (state variable) and precipitation events (influx).

    In the first case, we averaged the time series of mean saturated soil moisture in the whole Nägelstedt catchment for every

15   year, i.e. 50 years in total. Next, we divided the resulting time series such that years with an average soil moisture content above 85th percentile of the time series were labeled as wet years. In contrast, years with an average soil moisture content below 15th percentile of the time series were labels as dry years. This annual partitioning was seen necessary due to the strong annual fluctuations of this variable. Finally, we performed the same analysis as describe above for both – now smaller – datasets.

20   Using results from dry years only (see Figure 16), showed a similar qualitative travel-time behavior but stark quantitative contrast compared the mean travel-time behavior discussed above (see Figure 5). Compared to the general case, mean life expectancy was much larger in dry years. In addition, dry years exhibit a wider range of possible values with the largest one (over 50 months) being almost 4 times a large as the smallest one (approximately 12 months).

    Wet years on the other side, exhibit a very small range with the smallest value (approximately 5 months) being roughly

25   only half as large as the largest value (approximately 11 months) (see Figure 17). Compared to the general case, where the largest value (approximately 20 months) were roughly 3 times as large as the smallest value (approximately 6 months), these two scenarios fall on either side of this spectrum. This stark discrepancy demonstrates again the strong impact of the state variable (soil moisture) on travel-time behavior. Another difference between the mean travel-time behavior in wet years and the general case is the unimodal distribution of the former.

30   In addition, results showed different statistical dependency of travel-time behavior with respect to precipitation, PET and SSM (see Table 3). Dry years showed very similar correlation values compared to the general case (see Table 1 a). On the other side, correlation values for wet years were remarkably smaller.

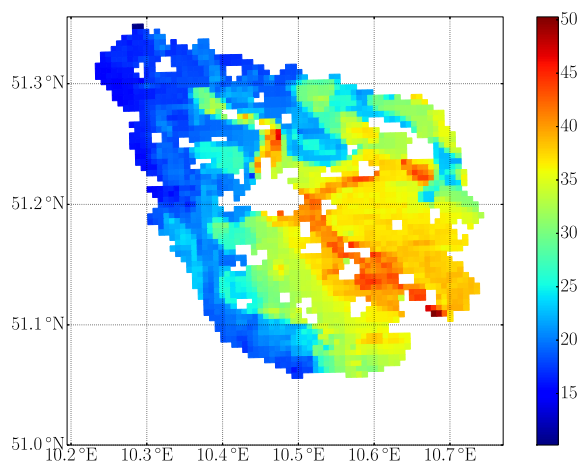
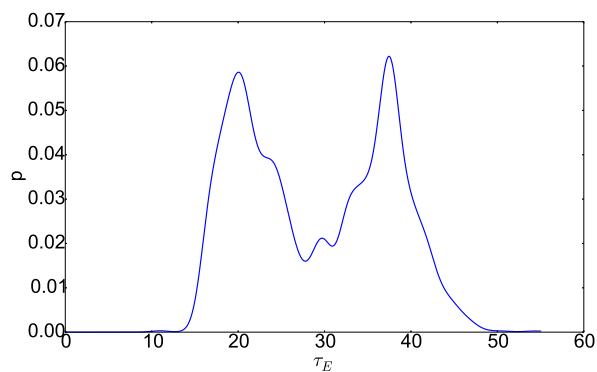
    In the second case, we also investigated travel-time behavior depending influx, i.e. in case of months having above average precipitation values (rainy months). To that end, we constrained our analysis to forward travel-time distributions which were

(a) spatial distribution

(b) kernel density estimate

Figure 16. Mean life expectancy in the Nägelstedt catchment in dry years.

	Pre	PET	SSM
wet years	0.374	0.084	0.388
dry years	0.781	0.223	0.834

Table 3. R^2 values for several predictors of mean travel time (as caused by wet and dry years)

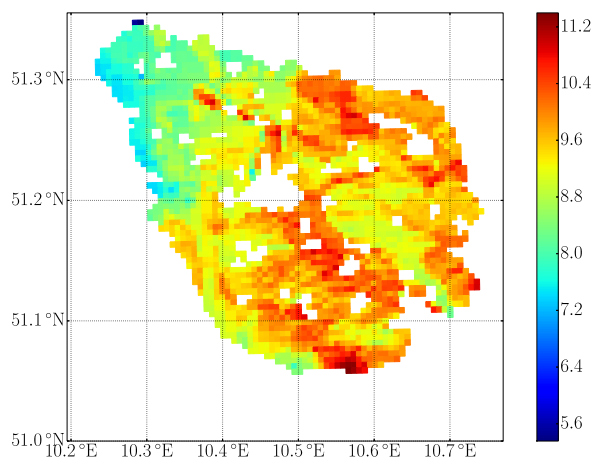
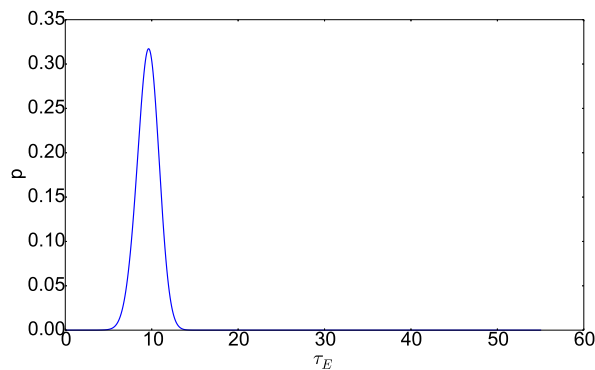
triggered by heavy rain events. This means that, in analogy to the analysis above we only used months with precipitation values above the 97th percentile and performed again the same analysis for the reduced dataset.

(a) spatial distribution

(b) kernel density estimate

Figure 17. Mean life expectancy in the Nägelstedt catchment in wet years.

	Pre	PET	SSM
mean life expectancy	0.736	0.221	0.857

Table 4. R^2 values for several predictors for mean travel time (as caused by rainy months)

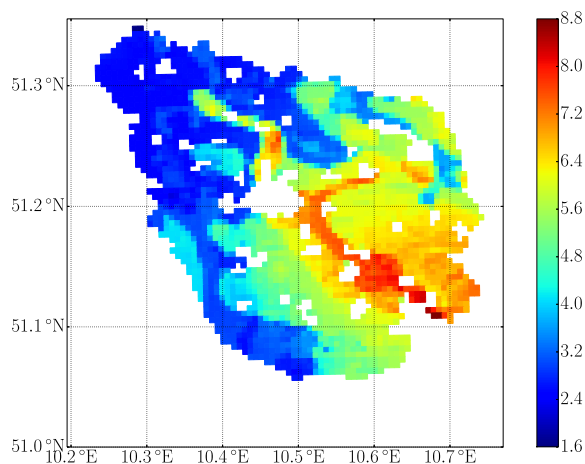
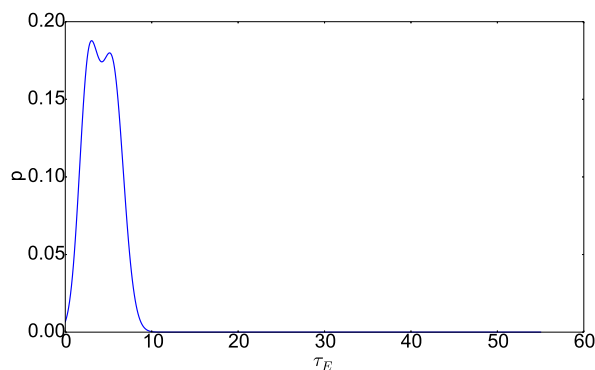
Results showed strong differences in mean life expectancy during rainy months compared to the scenarios discussed above (compare Figure 18 with Figures 6 and 16). Compared to wet years, we saw even lower mean life expectancy. This can be explained by the strong impact of the rain on soil moisture leading to a flushing of the soil. We also saw a similarly small variance and a nearly unimodal distribution of mean travel-time values.

(a) spatial distribution

(b) kernel density estimate

Figure 18. Mean life expectancy τ_E in the Nägelstedt catchment caused by rainy months.

In addition to that, we saw differences for the statistical correlation of mean life expectancy with precipitation, potential evapotranspiration and saturated soil moisture (see Table 4). Compared to the standard travel-time behavior, precipitation was slightly less explanatory with mean life expectancy. This was caused by lower overall variation in precipitation values, due to constraining our analysis to large values therefore excluding low and medium range rain events. In contrast to that, R^2 values for PET and SSM increased.

3.6 Relevance of TTDs for hydrological inference

The above results demonstrated the impact of certain soil properties, as implemented in mHM, on mean travel times using the R^2 metric as a measure. In addition to that statistical analysis, their relationship can further be elucidated by analyzing

	all years	dry years	wet years
QI_f	0.150	0.061	0.173
QI_s	0.298	0.239	0.317
QB	0.552	0.700	0.512

Table 5. Relative contribution of the different fluxes to discharge

	all years	dry years	wet years
R^2	0.6059	0.6954	0.3619

Table 6. R^2 values for recharge vs. mean travel times for different regimes.

Equation (2) or (3). Assuming for example a very simple linear relationship for both Q and ET with respect to S we get for Equation (2) the following

$$\overrightarrow{p}_Q(T_E, t_{in})_Q = \frac{\alpha_Q}{\theta(t_{in})} \exp\left(-\alpha_{ET}T_E\right)\exp\left(-\alpha_Q T_E\right). \tag{8}$$

Equation (8) shows how under such simplified assumptions, the TTD of such a CV would follow an exponential distribution with its mean travel time being related to the recession constants α_Q and α_{ET}. As shown above, such an exponential behavior is visible in the mean behavior (see Figure 5 right), whereas non-stationary TTDs show this exponential behavior to be superimposed by the event-based nature of the governing fluxes (see Figure 5 left).

In addition to these differences, we also saw different mean travel-time behavior for different regimes (see above). These differences can be explained by the actual implementation of Q and ET in mHM, which is generally non linear (see Section 2.2). In order to assess the different roles of each soil process on discharge generation, we calculated the relative contribution of each outflow mechanism for each regime. The data in Table 5 show how much of the water that entered the soil during a given time and left eventually as discharge was leaving as baseflow QB, slow interflow QI_s or fast interflow QI_f. On average, baseflow contributed the most to discharge with fast interflow having the smallest share. This overall distribution became stronger pronounced during dry years with baseflow taking the largest share of outflow generation and fast interflow becoming negligible. For wet years this trend is reversed, with water entering the soil during rainy months having an almost equal distribution. These different weighs show the relative impact and therefore the relative information content that travel-time distributions could contain, i.e. travel times in dry years are mostly the results of the successive processes leading eventually to base flow (see Figure 3), whereas travel times during storm events contain information on all discharge processes combined.

To further elucidate the relationship between the resulting mean travel times and certain model parameters, we performed a regression analysis comparing the recession constant for recharge with the mean travel times for different regimes. Results

confirmed the relationship described above with mean travel times during dry years showing the strongest correlation (see Table 6).

Such a high interdependency between certain model parameters and data from different flow regimes is not unique for TTDs. Using discharge alone would reveal similar overall tendencies, e.g. discharge data from droughts is more informative

5   for calibrating baseflow recession constants. What is, however, new is the additional information content, which is not contained in discharge data alone. Not only can this improve calibration efforts, it allows the inference of additional system states. This is in particular relevant, but not confined to, the total amount of stored groundwater. Discharge data are not sensitive to, and therefore not informative for, groundwater levels, but only to its relative changes. TTDs on the other hand, strongly depend on the total amount of water stored in every CV. Using both data types for inference would therefore allow to

10 provide reasonable estimates of this quantity. Similarly, the estimation of water in the root and vadose zone can be improved.

In addition *Birkel and Soulsby* (2015) highlight the temporal aspects of travel times on model calibration. They point out, how the sampling frequency of the time series should match the expected travel times of the underlying process. Our results above revealed different time scales for different hydrological regimes, that different by almost an order of magnitude. Despite this heterogeneity, all travel times in our study remained in the range of months. Under such circumstances, a fine-scale

15 measurement campaign with daily or even hourly intervals would for example not be necessary.

Although the above explanations provide only a limited perspective on the relationship between TTDs and model parameters, it can be said that the strong interlink between the travel-time behavior and outflow generation indicates the high information content of the former with respect to the latter. As a result, travel-time distributions should be regarded as highly informative for the calibration of hydrological models. As mentioned in the Introduction, *McDonnell and Beven* (2014) have

20 made the case for the usefulness of TTDs for the parametrization of such models. The above presentations provide empirical support for this notion.

4   Summary and Conclusions

In this study, we investigated soil behavior in the Nägelstedt catchment by virtue of travel-time distributions. The states and fluxes, needed for the derivation of the travel times, were numerically computed using the mesoscale Hydrological Model

25 (mHM), which was calibrated against 55 years of discharge data as well as using detailed data on soil properties, land cover and precipitation. We performed a statistical analysis of mean travel times to describe the soil response decoupled from the event-driven impact of precipitation.

Comparing the derived mean travel times for several modeling scales (spanning over one order of magnitude), we did not see any significant difference in their distribution. This indicates the general soundness of the parametrization scheme of

30 mHM used for the calculation of the states and fluxes on the different modeling scales. Our analysis shows that precipitation, saturated soil moisture and potential evapotranspiration are strong statistical predictors of mean travel time behavior. We also note that on average shorter mean travel times correspond to forested area and larger ones to crop/grassland.

We also investigated the travel-time behavior for different hydrological regimes, i.e. for dry and wet conditions (using soil moisture and precipitation as indicators). Our analysis revealed significantly different travel-time behavior for each of these regimes. Despite the strong heterogeneity of soil properties as well as (to a lesser extent) precipitation values, we could discriminate these regimes also in the resulting distribution of mean travel times.

Under dry conditions, we saw mean travel times having a pronounced bimodal distribution with long mean travel times and large variance. Such long travel times reveal the strong impact of baseflow on the generated outflow, whereas the large variance shows the variety of soil responses under dry conditions. Such conditions are therefore suited for inferring soil properties relating to baseflow generation. In addition, due to the large variance of soil responses, such conditions would allow to infer the spatial origin of solutes found in discharge streams. Such inferences are however, hampered by the long travel times involved. Not only are long time series needed, measurements must also being performed during such dry conditions.

Under wet conditions, we saw mean travel times having a unimodal distribution with shorter mean travel times and a smaller variance. This shorter travel times are caused by a larger influence of the slow and fast interflows on the total discharge behavior. As a result, TTDs derived under such conditions may be suited for inferring the parameters relating to these hydrological processes.

In case of rainy months, which overlap with wet conditions to a significant degree, we saw a similar distribution of travel times, but with even shorter mean values. This indicates a stronger impact of fast interflow on the total discharge behavior. Such information can therefore be valuable for improving the parametrization of the fast interflow related processes.

It is important to emphasize that our results have been derived with respect to a single hydrological model, i.e. mHM, only. However, due to the well established ability of mHM to take subgrid heterogeneity into account, we have good confidence in the physical plausibility of the spatially explicit soil moisture states and fluxes. The framework we used here is general and can be used in conjunction with other hydrological/land-surface models applied in other catchments. Moreover, future studies may incorporate other explanatory variables for the spatial heterogeneity of TTDs.

Having established a comprehensive description for the storage and release of water in the investigated catchment, the natural next step is the integration of reactive solute transport. As demonstrated by e.g. *Botter et al.* (2010), the concept of travel-time distributions can directly be adapted to account for the transport of both conservative and reactive solutes. This extension would facilitate to compare our predictions with the wealth of data that has been and continues to be collected within the AquaDiva center at the Hainich Critical Zone Exploratory (*Küsel et al.*, 2016). Thereby, we will be able to test our predictions by virtue of a large data set as well as initiate the collection of additional new data.

Appendix A: Forward and backward formulation of travel times

Both the forward and backward formulations for TTDs can be derived from Equation (1) by additionally associating each term with its distribution, so

$$\frac{d}{dt}[S(t)p_S(T,t)] = J(t)p_J(T,t) - ET(t)p_{ET}(T,t) + Q(t)p_Q(T,t). \tag{A1}$$

Here T is a placeholder for either the age or life expectancy of the water parcel. The total derivative in Equation (A1) can be reformulated using the material derivative, so

$$\left(\frac{\partial}{\partial t} + \frac{d}{dt}T\frac{\partial}{\partial T}\right)[S(t)p_S(t,T)] = J(t)p_J(t,T) - ET(t)p_{ET}(t,T) + Q(t)p_Q(t,T). \tag{A2}$$

Equation (A2) is the general PDE describing the time evolution of the age of the water in the CV. It is worthwhile to note that there is a significant inconsistency in the literature with respect to this equation. *Botter et al.* (2011) discuss the backward formulation of Equation (A2), while referring to it as the Master Equation (ME). This is certainly justified given that the ME is describing the time evolution of the PDF of any Markov process, i.e. a stochastic process that is local in (chronological) time. This condition is true for Equation (A2). In addition, Equation (A2) is not only local with respect to chronological time t but also with respect to the travel time T. Interpreting T as x, it becomes obvious that Equation (A2) is analogous to the much simpler Fokker-Planck equation or, since there is no 'diffusion', the even simpler advection-reaction equation. On the other side, *Porporato and Calabrese* (2015) are careful to trace this equation back to the seminal work of both *M'Kendrick* (1925) and *von Förster* (1959) in population dynamics. Consequently, they call this equation the McKendrick-von Förster (MKVF) equation.

One problem with Equations A1 and A2 is the lack of closure, i.e. $p_S(T,t)$, $p_J(T,t)$, $p_{ET}(T,t)$ and $p_Q(T,t)$ are different variables. The solution to this problem is the formulation and/or derivation of a dependency, i.e. closure, between the storage and the fluxes through

$$p_F(T,t) = \omega(T,t)p_S(T,t). \tag{A3}$$

This closure function $\omega(T,t)$ must follow some properties to ensure the normality of both $p_S(T,t)$ and $p_F(T,t)$, with the latter being the PDF of a flux, i.e. effective precipitation, evapotranspiration or discharge. These closure functions are called StorAge Selection (SAS) functions in the literature (*Rinaldo et al.*, 2015). Several different formulations exist with the one given above being based on the work of *Botter* (2012).

The shape of the SAS function is determining the preference of the fluxes, e.g. discharge, for several ages of the water stored in the CV. In the backward formulation a flat function would correspond to no preference with respect to age, a monotonously decreasing function would correspond to a preference for younger water and a monotonously increasing function would correspond to a preference for older water

Acknowledgements. This project was financially supported by the Deutsche Forschungsgemeinschaft via the Sonderforschungsbereich CRC 1076 AquaDiva. We furthermore kindly acknowledge our data providers: the German Weather Service (DWD), the Joint Research Center of the European Commission, the European Environmental Agency, the Federal Institute for Geosciences and Natural Resources (BGR), the European Water Archive and the Global Runoff data Centre.

References

Almorox, J., V. H. Quej, and P. Martí (2015), Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Köppen climate classes, *Journal of Hydrology*, *528*, 514 – 522, doi:http://dx.doi.org/10.1016/j.jhydrol.2015.06.057.

Benettin, P., A. Rinaldo, and G. Botter (2013), Kinematics of age mixing in advection-dispersion models, *Water Resources Research*, *49*(12), 8539–8551, doi:10.1002/2013WR014708.

Benettin, P., A. Rinaldo, and G. Botter (2015), Tracking residence times in hydrological systems: forward and backward formulations, *Hydrological Processes*, pp. n/a–n/a, doi:10.1002/hyp.10513.

Bergström, S. (1995), The HBV Model, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 443–476, Water Resour. Publ., Colorado, USA.

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

Birkel, C., and C. Soulsby (2015), Advancing tracer-aided rainfall-runoff modelling: a review of progress, problems and unrealised potential, *Hydrological Processes*, *29*(25, SI), 5227–5240, doi:10.1002/hyp.10594.

Botter, G. (2012), Catchment mixing processes and travel time distributions, *Water Resources Research*, *48*(5), n/a–n/a, doi:10.1029/2011WR011160, w05545.

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

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

Broxton, P. D., P. A. Troch, and S. W. Lyon (2009), On the role of aspect to quantify water transit times in small mountainous catchments, *Water Resources Research*, *45*(8), n/a–n/a, doi:10.1029/2008WR007438, w08427.

Cardenas, M. B. (2007), Potential contribution of topography-driven regional groundwater flow to fractal stream chemistry: Residence time distribution analysis of tóth flow, *Geophysical Research Letters*, *34*(5), n/a–n/a, doi:10.1029/2006GL029126, l05403.

Cornaton, F., and P. Perrochet (2006a), Groundwater age, life expectancy and transit time distributions in advective–dispersive systems: 1. generalized reservoir theory, *Advances in Water Resources*, *29*(9), 1267 – 1291, doi:http://dx.doi.org/10.1016/j.advwatres.2005.10.009.

Cornaton, F., and P. Perrochet (2006b), Groundwater age, life expectancy and transit time distributions in advective–dispersive systems; 2. reservoir theory for sub-drainage basins, *Advances in Water Resources*, *29*(9), 1292 – 1305, doi:http://dx.doi.org/10.1016/j.advwatres.2005.10.010.

Dagan, G. (1986), Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale, *Water Resources Research*, *22*(9S), 120S–134S, doi:10.1029/WR022i09Sp0120S.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer.

Daly, E., and A. Porporato (2006), Impact of hydroclimatic fluctuations on the soil water balance, *Water Resources Research*, *42*(6), n/a–n/a, doi:10.1029/2005WR004606, w06401.

Danckwerts, P. (1953), Continuous flow systems, *Chemical Engineering Science*, *2*(1), 1 – 13, doi:http://dx.doi.org/10.1016/0009-2509(53)80001-1.

Duffy, C. J. (2010), Dynamical modelling of concentration–age–discharge in watersheds, *Hydrological Processes*, *24*(12), 1711–1718, doi:10.1002/hyp.7691.

Dunn, S. M., W. G. Darling, C. Birkel, and J. R. Bacon (2012), The role of groundwater characteristics in catchment recovery from nitrate pollution, *Hydrology Research*, *43*(5), 560–575, doi:10.2166/nh.2012.020.

Federal Institute for Geosciences and Natural Resources (BGR) (1998), Digital soil map of Germany 1:1\,000\,000 (BUEK 1000), *Tech. rep.*, Hannover and Berlin, Germany.

5 Hargreaves, G. H., and Z. A. Samani (1985), Reference crop evapotranspiration from temperature, *Appl. Engrg. in Agri.*, *1*, 96–99.

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

Hrachowitz, M., C. Soulsby, D. Tetzlaff, J. J. C. Dawson, and I. A. Malcolm (2009), Regionalization of transit time estimates in montane catchments by integrating landscape controls, *Water Resources Research*, *45*(5), n/a–n/a, doi:10.1029/2008WR007496, w05421.

10 Hrachowitz, M., H. Savenije, T. A. Bogaard, D. Tetzlaff, and C. Soulsby (2013), What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, *Hydrology and Earth System Sciences*, *17*(2), 533–564, doi:10.5194/hess-17-533-2013.

Hrachowitz, M., O. Fovet, L. Ruiz, and H. H. G. Savenije (2015), Transit time distributions, legacy contamination and variability in biogeochemical 1/f(alpha) scaling: how are hydrological response dynamics linked to water quality at the catchment scale?, *Hydrological Processes*, *29*(25, SI), 5241–5256, doi:10.1002/hyp.10546.

15 Kitanidis, P. K. (2015), Persistent questions of heterogeneity, uncertainty, and scale in subsurface flow and transport, *Water Resources Research*, *51*(8), 5888–5904, doi:10.1002/2015WR017639.

Kumar, R., L. Samaniego, and S. Attinger (2010), The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics, *J. Hydrol.*, *392*(1–2), 54–69, doi:10.1016/j.jhydrol.2010.07.047.

Kumar, R., L. Samaniego, and S. Attinger (2013a), Implications of distributed hydrologic model parameterization on water fluxes at multiple

20 scales and locations, *Water Resources Research*, *49*(1), 360–379, doi:10.1029/2012WR012195.

Kumar, R., B. Livneh, and L. Samaniego (2013b), Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme, *Water Resour. Res.*, *49*(9), 5700–5714, doi:10.1002/wrcr.20431.

Küsel, K., K. U. Totsche, S. E. Trumbore, R. Lehmann, C. Steinhäuser, and M. Herrmann (2016), How deep can surface signals be traced in the critical zone? merging biodiversity with biogeochemistry research in a central german muschelkalk landscape, *Frontiers in Earth*

25 *Science*, *4*(32), doi:10.3389/feart.2016.00032.

Livneh, B., R. Kumar, and L. Samaniego (2015), Influence of Soil Textural Properties on Hydrologic Fluxes in the Mississippi River Basin, *Hydrol. Processes*, *29*, 4638–4655, doi:10.1002/hyp.10601.

Małoszewski, P., and A. Zuber (1982), Determining the turnover time of groundwater systems with the aid of environmental tracers, *Journal of Hydrology*, *57*(3), 207 – 231, doi:http://dx.doi.org/10.1016/0022-1694(82)90147-0.

30 McDonnell, J. J., and K. Beven (2014), Debates – The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph, *Water Resources Research*, *50*(6), 5342–5350, doi:10.1002/2013WR015141.

McGuire, K. J., and J. J. McDonnell (2015), Tracer advances in catchment hydrology, *Hydrological Processes*, *29*(25, SI), 5135–5138, doi:10.1002/hyp.10740.

35 McGuire, K. J., J. J. McDonnell, M. Weiler, C. Kendall, B. L. McGlynn, J. M. Welker, and J. Seibert (2005), The role of topography on catchment-scale water residence time, *Water Resources Research*, *41*(5), n/a–n/a, doi:10.1029/2004WR003657, w05002.

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

M'Kendrick, L.-C. A. G. (1925), Applications of mathematics to medical problems, *Proceedings of the Edinburgh Mathematical Society*, *44*, 98–130.

Niemi, A. J. (1977), Residence time distributions of variable flow processes, *The International Journal of Applied Radiation and Isotopes*, *28*(10–11), 855 – 860, doi:http://dx.doi.org/10.1016/0020-708X(77)90026-6.

Porporato, A., and S. Calabrese (2015), On the probabilistic structure of water age, *Water Resources Research*, *51*(5), 3588–3600, doi:10.1002/2015WR017027.

Rakovec, O., R. Kumar, J. Mai, M. Cuntz, S. Thober, M. Zink, S. Attinger, D. Schäfer, M. Schrön, and L. Samaniego (2016), Multiscale and multivariate evaluation of water fluxes and states over European river basins, *J. Hydrometeorol.*, *17*, 287–307, doi:10.1175/JHM-D-15-0054.1.

Rinaldo, A., and A. Marani (1987), Basin scale model of solute transport, *Water Resources Research*, *23*(11), 2107–2118, doi:10.1029/WR023i011p02107.

Rinaldo, A., A. Marani, and A. Bellin (1989), On mass response functions, *Water Resources Research*, *25*(7), 1603–1617, doi:10.1029/WR025i007p01603.

Rinaldo, A., G. Botter, E. Bertuzzo, A. Uccelli, T. Settin, and M. Marani (2006), Transport at basin scales: 1. theoretical framework, *Hydrology and Earth System Sciences*, *10*(1), 19–29, doi:10.5194/hess-10-19-2006.

Rinaldo, A., K. J. Beven, E. Bertuzzo, L. Nicotina, J. Davies, A. Fiori, D. Russo, and G. Botter (2011), Catchment travel time distributions and water flow in soils, *Water Resources Research*, *47*(7), n/a–n/a, doi:10.1029/2011WR010478, w07537.

Rinaldo, A., P. Benettin, C. J. Harman, M. Hrachowitz, K. J. McGuire, Y. van der Velde, E. Bertuzzo, and G. Botter (2015), Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes, *Water Resources Research*, *51*(6), 4840–4847, doi:10.1002/2015WR017273.

Rodriguez-Iturbe, I., and J. B. Valdés (1979), The geomorphologic structure of hydrologic response, *Water Resources Research*, *15*(6), 1409–1420, doi:10.1029/WR015i006p01409.

Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, 416 pp., Oxford University Press, USA.

Samani, Z. A. (2000), Estimating solar radiation and evapotranspiration using minimum climatological data, *Journal of Irrigation and Drainage Engineering*, *126*(4), 265–267, doi:10.1061/(ASCE)0733-9437(2000)126:4(265).

Samaniego, L., R. Kumar, and S. Attinger (2010a), Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, *Water Resources Research*, *46*(5), n/a–n/a, doi:10.1029/2008WR007327, w05523.

Samaniego, L., A. Bárdossy, and R. Kumar (2010b), Streamflow prediction in ungauged catchments using copula-based dissimilarity measures, *Water Resources Research*, *46*.

Samaniego, L., R. Kumar, and M. Zink (2013), Implications of parameter uncertainty on soil moisture drought analysis in Germany, *J. Hydrometeorol.*, *14*(1), 47–68, doi:10.1175/JHM-D-12-075.1.

Tetzlaff, D., I. Malcolm, and C. Soulsby (2007), Influence of forestry, environmental change and climatic variability on the hydrology, hydrochemistry and residence times of upland catchments, *Journal of Hydrology*, *346*(3–4), 93 – 111, doi:http://dx.doi.org/10.1016/j.jhydrol.2007.08.016.

Tetzlaff, D., J. Seibert, K. J. McGuire, H. Laudon, D. A. Burns, S. M. Dunn, and C. Soulsby (2009), How does landscape structure influence catchment transit time across different geomorphic provinces?, *Hydrological Processes*, *23*(6), 945–953, doi:10.1002/hyp.7240.

Tetzlaff, D., C. Soulsby, M. Hrachowitz, and M. Speed (2011), Relative influence of upland and lowland headwaters on the isotope hydrology and transit times of larger catchments, *Journal of Hydrology*, *400*(3–4), 438 – 447, doi:http://dx.doi.org/10.1016/j.jhydrol.2011.01.053.

Thober, S., R. Kumar, J. Sheffield, J. Mai, D. Schäfer, and L. Samaniego (2015), Seasonal Soil Moisture Drought Prediction over Europe Using the North American Multi-Model Ensemble (NMME), *Journal of Hydrometeorology*, *16*(6), 2329–2344.

van der Velde, Y., P. J. J. F. Torfs, S. E. A. T. M. van der Zee, and R. Uijlenhoet (2012), Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, *Water Resources Research*, *48*(6), n/a–n/a, doi:10.1029/2011WR011310, w06536.

Vereecken, H., J. A. Huisman, H.-J. H. Franssen, N. Brueggemann, H. R. Bogena, S. Kollet, M. Javaux, J. van der Kruk, and J. Vanderborght (2015), Soil hydrology: Recent methodological advances, challenges, and perspectives, *Water Resources Research*, *51*(4), 2616–2633, doi:10.1002/2014WR016852.

von Förster, H. (1959), Some remarks on changing populations, in *The Kinetics of Cellular Proliferation*, edited by F. Stohlman, pp. 382–407, Grune and Stratton, N. Y.

Wang, K., and R. E. Dickinson (2012), A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Reviews of Geophysics*, *50*(2), n/a–n/a, doi:10.1029/2011RG000373, rG2005.

Windhorst, D., P. Kraft, E. Timbe, H.-G. Frede, and L. Breuer (2014), Stable water isotope tracing through hydrological models for disentangling runoff generation processes at the hillslope scale, *Hydrology and Earth System Sciences*, *18*(10), 4113–4127, doi:10.5194/hess-18-4113-2014.