

Finding behavioral parameterization for a 1-D water balance model by multi-criteria evaluation

Markus C. Casper¹, Hadis Mohajerani^{1*}, Sibylle Hassler^{2,3}, Tobias Herdel¹, Theresa Blume²

¹ University of Trier, Faculty VI, Dep. of Physical Geography, Universitätsring 12, 54286 Trier, Germany.

² GFZ German Research Centre for Geosciences, Section 5.4 Hydrology, Telegrafenberg, 14473 Potsdam, Germany.

³ Karlsruhe Institute of Technology (KIT), Institute for Water and River Basin Management, Chair of Hydrology, Kaiserstr. 12, 76131 Karlsruhe, Germany.

* Corresponding author. Tel.: +49 651 2014557. FAX: +49 651 2013976. E-mail: s6hsmoha@uni-trier.de

Abstract: Evapotranspiration is often estimated by numerical simulation. However, to produce accurate simulations, these models usually require on-site measurements for parameterization or calibration. We have to make sure that the model realistically reproduces both, the temporal patterns of soil moisture and evapotranspiration. In this study, we combine three sources of information: (i) measurements of sap velocities; (ii) soil moisture; and (iii) expert knowledge on local runoff generation and water balance to define constraints for a “behavioral” forest stand water balance model. Aiming for a behavioral model, we adjusted soil moisture at saturation, bulk resistance parameters and the parameters of the water retention curve (WRC). We found that the shape of the WRC influences substantially the behavior of the simulation model. Here, only one model realization could be referred to as “behavioral”. All other realizations failed for a least one of our evaluation criteria: Not only transpiration and soil moisture are simulated consistently with our observations, but also total water balance and runoff generation processes. The introduction of a multi-criteria evaluation scheme for the detection of unrealistic outputs made it possible to identify a well performing parameter set. Our findings indicate that measurement of different fluxes and state variables instead of just one and expert knowledge concerning runoff generation facilitate the parameterization of a hydrological model.

Keywords: Forest evapotranspiration; Water balance simulation; Soil parameterization; Behavioral model.

INTRODUCTION

Extraction of water from the soil by the root system and return of water to the atmosphere as plant transpiration are important processes in the global circulation of water (Kramer and Boyer, 1995). Quantitative means of describing transpiration are essential for an improved understanding of water and energy exchange processes between the land surface and the atmosphere. Transpiration is controlled by a combination of biotic factors (e.g. stomatal functions; leaf area; root depth and distribution, and hydraulic characteristics) and abiotic factors (e.g. soil water availability; climate, and depth to groundwater) (Durigon et al., 2016).

There is a variety of techniques to measure transpiration at different scales such as direct measurements of sap flow on individual trees (Lu et al., 2004), eddy flux gradient analyses (Saugier et al., 1997), or gauged watersheds (Wilson et al., 2001). Alternatively, simulation models are used to estimate transpiration. However, to produce accurate simulations, these usually require on-site parameterization or calibration (Durigon et al., 2016; Vose et al., 2003). Recently, a simple approach was developed by Ayyoub et al (2017) relating the normalized daily sap velocities and the daily reference evapotranspiration (ET₀). This method used both, FAO-Penman-Monteith (FAO-PM) method and Hargreaves-Samani (HARG) method to estimate ET₀. The FAO-PM method produced the highest correlations to daily sap velocities (Ayyoub et al., 2017).

System state (“soil moisture”) and actual evapotranspiration are known to be highly correlated. Therefore, all water balance models directly couple these two components. Wrong estimates of temporal or spatial patterns of soil moisture result in erroneous temporal or spatial patterns of transpiration (Casper and Vohland, 2008; Koch et al., 2017). Therefore, soil parameteri-

zation, and especially the representation of the Water Retention Curve - as one of the most important soil-physical characteristics - strongly influence the simulation of evapotranspiration. A similar effect can be observed when canopy resistances are wrongly estimated (Bie et al., 2015). In order to find an appropriate model parametrization, we have to verify that the model realistically reproduces both, the temporal patterns of soil moisture and evapotranspiration. This has been done in a study carried out by Holst et al. (2010) where the water balance of two beech stands in Southwest Germany was investigated using two different forest hydrological models (DNDC and BROOK90). They demonstrated that both models were able to reproduce the observed dynamics of the soil water content in the uppermost 30 cm and the transpiration estimates from sap flow measurements (Holst et al., 2010).

To analyze different assumptions on catchment behavior and hydrological processes, it is necessary to evaluate the model performance with respect to multiple indicators that evaluate the contribution of different sources of data (Gupta et al., 1998). The value of these additional data sources has been demonstrated by Fenicia et al. (2008a). They evaluated the accuracy of a hydrological simulation with respect to the observed discharge, groundwater level dynamics, and isotope signatures. If appropriate data is lacking, incorporation of expert knowledge (as an alternative source of information) into hydrological modeling and water management issues becomes more important (Bromley et al., 2005; Cash et al., 2003; Mohajerani et al., 2017). As recent studies suggest, use of expert knowledge in choosing parameter sets and introducing constraints by forcing the model to reproduce the processes observed in the real system, can also improve the model performance even without traditional calibration (Bahremand, 2016; Gharari et al., 2014; Hrachowitz et al., 2014). For instance,

having expert knowledge on local runoff generation processes, as a potential source of information in every hydrologic unit, can considerably improve hydrological simulations (Antonetti and Zappa, 2018; Casper et al., 2015; Franks et al., 1998; Seibert and McDonnell, 2002). Modelers need to consider a proper balance between parameter identifiability and the model's ability to precisely represent the observed system response. This has prompted the development of alternative approaches to hydrological modeling including the dominant process concept (Fenicia et al., 2008b; Grayson and Blöschl, 2001; Wagener et al., 2001). The concept of dominant runoff generation process (DRGP) assumes that at a particular location one particular runoff generation mechanism is dominant (Blöschl, 2001). In most of the studies, however, modelers have evaluated the model performance with respect to discharge data alone. This may cloud model realism and hamper understanding of catchment behavior. In fact, to accurately evaluate hydrological models, one has to carefully look at the hydrological responses that a model is challenged to represent rather than just being satisfied with a simple calibration. This interestingly takes us back to what Fenicia et al. (2008a) call the "Art of Modeling" and what Gupta et al. (2005) call the "Behavioral Model". The former says: "...modeling is both an Art and a Science. The science lies in the use of fundamental scientific principles and the formality of analysis; the art accounts for professional experience, insight, creativity and intuition. The latter is particularly important in developing a perceptual and conceptual model that captures the main processes at play, while maintaining minimum levels of complexity...". A "behavioral" model has the following characteristics: (i) the input-state-output behavior of the model is consistent with the measurements, (ii) the model predictions are accurate (i.e. they have negligible bias) and (iii) model structure and behavior are consistent with our hydrologic understanding of reality (Fenicia et al., 2008a; Gupta et al., 2005).

The present study combines three sources of information: measurements of sap velocities; soil moisture data; and expert knowledge of local runoff generation and water balance to define constraints for a "behavioral" forest stand water balance model. We evaluated the model by defining multi-criteria performance measures according to the constraints that data are supposed to impose on model behavior. In particular, we investigated the following research questions: (i) How can we combine different sources of information to modify the parameterization scheme in order to achieve a "behavioral model"? (ii) How does the implementation of expert knowledge of site-specific dominant runoff generation processes affect the simulation results? (iii) What are the impacts of model setup, i.e. the parameterization approach and the parameter allocation strategy on the simulated soil moisture and evapotranspiration dynamics (e.g. the effect of different parameterizations of the water retention curve of the soil)?

To address the research questions listed above, we used a one-dimensional (1-D) hydrological model (WaSiM-ETH) to simulate the soil water content as well as the actual transpiration at stand level. The basic motivation of 1-D models is often to simulate soil water content, and water balance components such as evapotranspiration, deep drainage and runoff. In the 1-D models, no groundwater flow is simulated and the upper and lower limits are soil water content at field capacity and permanent wilting point, and upper and lower loss of soil water is caused by evapotranspiration and deep percolation, respectively (Walker and Zhang, 2002). WaSiM-ETH was selected due to its highly differentiated 1-D model structure. The model represents all relevant hydrological processes at the point scale in a

physically meaningful way (Schulla, 2017). As all measured data including soil moisture and sap velocity are point measurements, setting up a "1-D model" is sufficient for our purpose. A site in the sandstone region of western Luxembourg was used as a test case. On sandstone, we expect neither stream channels nor surface runoff due to the high hydraulic conductivities of the sandy soils. The headwaters start at springs on top of the less permeable marls underlying the sandstone. While this work is not going to provide new insights of the behavior of the study catchment, it arguably is going to contribute to understanding of the value of different sources of data and information for hydrological modeling. The test case is used as a "proof-of-concept" location to investigate how different parameterization with different content of information can affect the model behavior. Our investigation is subdivided into four scenarios, i.e. different soil parameterizations. For each scenario the simulation results are evaluated by the model performance criteria defined in the section 2.4. In scenario A, the soil parameterization is taken from Tepee et al. 2003. In the two scenarios B1 and B2, we parameterize the water retention curve with three different variations of the van Genuchten parameters according to (Sauer, 2007). In a last step (scenario C), we evaluate the model performance using the soil parameter set provided by Sprenger et al. (2016). All scenarios are summarized in Table 6.

METHODS

Site description

The study area is the Huewelerbach, a sub-catchment (2.7 km² in area, ranging from 280 to 400 m in elevation) of the Attert River basin located in the west of Luxembourg (for detailed information see Martínez-Carreras et al. (2010)). The whole area is part of the "Catchments As Organized Systems" (CAOS) observatory investigating landscape-scale structures, patterns and interactions in hydrological processes for model development (Zehe et al., 2014). The catchment is mainly forested, but the alluvial section of the area is dominated by grassland. The mean annual precipitation of the area is approximately 850 mm (Pfister et al., 2000). In terms of lithology, the Huewelerbach catchment consists of jurassic Luxembourg sandstone which is underlain by marls (Martínez-Carreras et al., 2012, 2010). According to previous studies, the sandstone bedrock and the underlying marls produce a very stable base flow regime (Juilleret et al., 2012). Rainfall-runoff reaction is delayed on the deep sandy soils on hillslopes (deep percolation and subsurface flow). Siltation and compaction in the valley bottom may cause sporadic surface runoff (Sprenger et al., 2016). Measurements at sites Sa_G and Sa_K include meteorological variables such as air temperature, humidity and solar radiation and soil moisture at three depths in three different profiles. At the forested site Sa_G there are also measurements of sap velocity at 4 trees, two of them European Beech (*Fagus sylvatica* L.) and two hornbeams (*Carpinus betulus* L.). Figure 1 shows the study area and the location of the selected sites. Dominant vegetation at site Sa_G is a relatively young beech forest with a basal area of 16 m²/ha. Within the measurement plot we find 34 stems with a mean diameter of 19 cm (median: 14 cm).

Hydrological model

To simulate the actual evapotranspiration (ET_a), we applied a hydrological model – WaSiM-ETH (Schulla, 1997). This model is a distributed, deterministic, mainly physical and grid-based hydrological model running with variable time steps

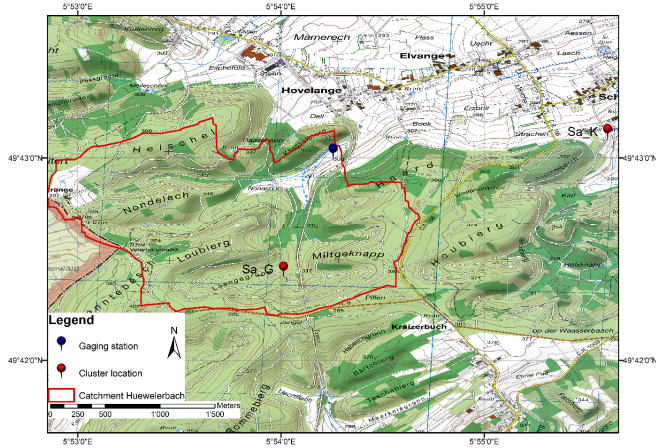


Fig. 1. Location of the study area Huewelerbach catchment with the test sites Sa_G (forest) and Sa_K (grassland).

(Schulla, 2017). The WaSiM-ETH model has performed well in sub-alpine and alpine catchments (Cullmann et al., 2006; Gurtz et al., 2003; Jasper, 2001; Jasper et al., 2002; Klok et al., 2001; Verbunt et al., 2003), also in middle-mountain (Bie et al., 2015; Grigoryan et al., 2010; Middelkoop et al., 2001), and lowland catchments (Elfert and Bormann, 2010). The model is documented in both English and German and can be used free of cost (<http://www.wasim.ch>). The model comprises different components (e.g. evapotranspiration model, soil model, snow model, glacier model, silting up, surface routing, groundwater model, discharge routing model, lake model etc.). In our case only the modules related to the soil model and evapotranspiration were used and all other components were disregarded.

Calculation of evapotranspiration

There are three main steps to estimate the ETa (plant transpiration as well as evaporation from the soil separately) in WaSiM-ETH. First, estimation of potential evapotranspiration (ETp) on the basis of the ground-measured meteorological data; second, simulation of soil water content in vertical direction via Richards equation (Richards, 1931). In the third step, the amount of ETa is simulated at every time step by reducing ETp according to the actual soil water content.

There are four different methods available in WaSiM-ETH model to calculate the ETp rates: Penman-Monteith approach (Monteith, 1981; Monteith et al., 1965); Wendling (Wendling, 1975); Haude (Haude, 1955) and Hamon (Federer and Lash, 1978). In this study, we choose Penman-Monteith equation (Monteith et al., 1965) (see equation 1). It is the most widely used and recommended method for ETp estimation, first developed for agricultural contexts and later also applied to other land covers such as forests (Allen et al., 1998; Droogers and Allen, 2002). This method is based on simulated potential transpiration and the available water content. In our case, actual plant transpiration is simulated in hourly time steps. However, the Penman-Monteith approach has some limitations in practical terms, as a large number of environmental variables are required to determine ETa. This is particularly challenging especially when there is a lack of appropriate atmospheric data (Allen et al., 1998).

$$\lambda E = \frac{3.6 \frac{\Delta}{\gamma_p} (Rn - G) + \frac{\rho c_p}{\gamma_p} (e_s - e) t_i}{\frac{\Delta}{\gamma_p} + 1 + \frac{r_s}{r_a}} \quad (1)$$

where λ is the latent vaporization heat, $\lambda = (2500.8 - 2.372 \cdot T)$ kJ kg⁻¹, with T is the temperature in °C; E is the latent heat flux in mm m⁻² \equiv kg m⁻²; Δ is the tangent to the saturated vapor pressure curve in hPa K⁻¹; Rn is the net radiation in Wh m⁻² and $G = 0.1 \cdot Rn$ is the soil heat flux in Wh m⁻², the factor 3.6 is used to convert both fluxes from Wh m⁻² to kJ m⁻²; ρ is the density of dry air, $\rho = p/(R \cdot T)$, at 0 °C and 1013.25 hPa, $\rho = 1.29$ kg m⁻³; c_p is the specific heat capacity of the dry air at constant pressure, $c_p = 1.005$ kJ kg⁻¹ K⁻¹; e_s is the saturation vapor pressure at temperature T , in hPa; e is the observed actual vapor pressure in hPa; t_i is the number of seconds within a time step; γ_p is the psychrometric constant in hPa K⁻¹; r_a and r_s are the bulk-aerodynamic resistance and the bulk-surface resistance in s m⁻¹, respectively.

Resistances for evapotranspiration

The two resistance parameters in the Penman-Monteith equation: the bulk aerodynamic resistance r_a and the bulk surface resistance r_s play an important role. However, r_s (with diurnal and seasonal variations) is more important than r_a in a forested area for ETa estimation (Beven, 1979). The bulk surface resistance r_s can be divided into two terms, the soil surface resistance r_{se} for evaporation from bare soil; and the canopy surface resistance r_{sc} describing the plant resistances in the transpiration process. There are default values of bulk surface resistance parameters in WaSiM-ETH. The maximum amount of canopy surface resistance r_{sc} is in November to February, whereas in May to September, it reduces to its annual lowest level (Bie et al., 2015; Schulla, 2017). The soil surface resistance r_{se} remains constant for the entire year. Table 1 shows the standard values applied for surface resistances parameters in the WaSiM-ETH model (Schulla, 2017).

Table 1. Canopy surface resistance r_{sc} (s m⁻¹) and soil surface resistance r_{se} (s m⁻¹).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
r_{sc}	100	100	95	75	65	65	65	65	65	85	100	100
r_{se}	230	230	230	230	230	230	230	230	230	230	230	230

Simulation of soil water content

For estimation of actual transpiration, WaSiM-ETH simulates soil moisture in the root zone. The soil module in WaSiM-ETH uses the van Genuchten method (Van Genuchten, 1980) for parameterization of the water retention curve to solve the Richard Equation. Water fluxes are simulated vertically in one dimension. Soil moisture in the root soil layer can potentially limit transpiration (Paço et al., 2014). In WaSiM-ETH, soil moisture simulation and ETa are linked, reduction of ETa would result in more water availability in the soil whereas increase of ETa will decrease the soil moisture. The Penman-Monteith equation implicitly includes the influence of soil moisture on plant transpiration through parameter r_{sc} (canopy surface resistance). Water content in soil profiles changes with time and values of the r_{sc} also show diurnal and seasonal variations. In dry periods, r_{sc} is very sensitive to soil moisture. When soil moisture content falls below a given point, the plants start decreasing transpiration to prevent internal water losses. Below that point, soil water availability becomes a key factor in obtaining ETa. ETa is gradually reduced until soil moisture reaches the wilting point at which water is no longer available for transpiration (Allen et al., 1998; Anderson et al., 2007).

Soil parameterization

Certain predefined parameters of the WaSiM-ETH model are specific for the area where the model was developed. Thus, these parameters should be modified for each new study area. In the investigated area, the predominant soil type was described as ‘‘Podzolic Cambisol’’. The soil texture is loamy sand. It developed on a sandstone bedrock. The maximum rooting depth for the soil was observed at approx. 100 cm. The stone content is relatively low and the unaltered parent sandstone is usually not reached within the first 200 cm below soil surface (Sprenger et al., 2016). These quite sandy soils show a high permeability resulting in deep percolation as the dominant hydrological process.

In the model, the van Genuchten parameters α (m^{-1}) and n (–) are empirical constants that determine the shape of the WRC, and therefore influence substantially the behavior of the simulation model. We chose three different methods to determine the parameters of the WRC.

(i) Baseline parameterization after Teepe et al. 2003

We derived the corresponding van Genuchten parameters in the different soil horizons based on soil texture and bulk density classification obtained by Teepe et al. (2003). This formed our baseline parameterization of the soil in our study area (Table 2).

Table 2. Baseline soil parameterization of the WaSiM-ETH soil model (based on Teepe et al. (2003)).

PMacroThresh	20					
MacroCapacity	4					
CapacityRedu	0.5					
MacroDepth	1					
Horizon	1	2	3	4	5	6
Name	Ahe	Ae	Bvs	Bsv	IIBvs	IIBvs
Ksat	1.01E-4	7.95E-05	1.65E-04	1.29E-04	4.84E-05	4.84E-05
K_recession	1	1	1	1	1	1
Theta_sat	0.41	0.41	0.41	0.41	0.41	0.41
Theta_res	0.11	0.05	0.06	0.06	0.13	0.13
Alpha	0.3	0.3	0.26	0.41	0.2	0.2
Par_n	1.17	1.17	1.203	1.191	1.191	1.191
Par_tau	0.5	0.5	0.5	0.5	0.5	0.5
Thickness	0.1	0.1	0.1	0.1	0.1	1
Layers	1	1	1	3	4	7

PMacroThresh (mm/h) is given by the precipitation threshold value and if is reached or exceeded, water can infiltrate into the macropore; MacroDepth (m) is depth of the macropores; MacroCapacity (mm/h) is capacity of the macropores; CapacityRedu (m^{-1}) defines the reduction of the macropore capacity per meter soil depth; Ksat ($m s^{-1}$) is saturated hydraulic conductivity that can be given for each soil layer; K_recession (–) is specified for each soil type describing the recession of the saturated conductivity with depth; theta_sat (m^3/m^3) is saturated water content; theta_res (–) is the residual water content which cannot be extracted by transpiration; α (m^{-1}) and Par_n (–) are empirical van-Genuchten parameters; Par_tau is Mualem parameter; thickness (m) is the thickness of every single numerical layer in the given horizon, and layers defines the number of layers in the given horizon.

(ii) Parameterization after Sauer (2007)

Sauer (2007) proposes three different methods to derive the van-Genuchten parameters α and n :

Variation 1: Fitting of WRC based on grain size fractions, bulk density and water content at pF 2.5 and 4.2 using the software ‘‘Rosetta Lite’’ (Schaap et al., 2001).

Variation 2: Fitting of WRC based on water content at pF 1.8, 2.5, 4.2 and Theta_sat (= 41%) using the software ‘‘RET’’ (Van Genuchten et al., 1991).

Variation 3: Fitting of WRC based on water content at pF 1.8, 2.5, 4.2 using the software ‘‘RET’’ (Van Genuchten et al., 1991). See Table 3 for the three variations of parameters α and n .

Table 3. Variations of van Genuchten parameters α (m^{-1}) and n (dimensionless) in different soil horizons as re-parameterization of the baseline (Table 2).

Horizon		Ahe	Ae	Bvs	Bsv	IIBvs	IIBvs
Variation 1	alpha	0.83	0.83	0.58	0.58	0.88	1.83
	n	1.5653	1.5653	1.6416	1.6416	1.4974	1.4553
Variation 2	alpha	2.86	2.86	3.97	3.97	4.96	1.83
	n	1.3656	1.3656	1.3965	1.3965	1.4598	1.4553
Variation 3	alpha	25.73	25.73	35.87	35.87	29.34	1.83
	n	1.2138	1.2138	1.2506	1.2506	1.3009	1.4553

(iii) Parameterization after Sprenger et al. (2016)

Sprenger et al. (2016) list soil parameters for the same site (Sa_G). These parameters were obtained by fitting the simulation results to observed soil moisture and pore water stable isotope data. In this case the soil profile was divided into three different horizons (Table 4).

Table 4. Parameterization of WRC for the site Sa_G (Sprenger et al., 2016).

Horizon	Ah	B	II_B
width	11 cm	110 cm	> 80 cm
theta_sat	0.546	0.319	0.470
alpha	0.033	0.005	0.005
n	1.228	1.194	1.194
k _{sat}	6.11E-04	1.53E-04	6.16E-04

Data description

To simulate transpiration and soil water content at the forest site Sa_G, climate data from the grassland site Sa_K were used as input for the model (Figure 1). These data better represent the atmospheric conditions above the trees which mainly drive the transpiration of the trees. In contrast, climate data from site Sa_G represents the conditions inside the forest and therefore this data cannot be used in our simulation study. To run the model, climate data between 2013 and 2016 is available. All subsequent model evaluation is done for the year 2015. The years 2013 and 2014 are used as spin-up period until stabilization of the model. Climate data includes air temperature, relative humidity, wind speed, global radiation as five-minute measurements, and precipitation as hourly data. All data were checked for errors and the data gaps were filled. Soil moisture was measured in three profiles per site at 10 cm, 30 cm and 50 cm depth. For our analyses we took the average across all depths and profiles estimating the average soil moisture in the top 60 cm for each site (Hassler et al., 2018). Precipitation data (station Useldange) are available as hourly values with annual mean value of 791 mm for the year 2015 (Agrarmeteorologie Luxemburg: /http://www.agrimeteo.lu). Therefore, all other climate variables and the soil moisture measurements are averaged to hourly values.

Based on the soil moisture and grain size distribution characteristics of the study area, deep percolation is usually observed as dominant runoff generation process. Saturation excess flow or Hortonian overland flow can be excluded.

For the year 2015, transpiration of the adult beech overstorey was analyzed by determining sap velocities using the heat ratio method with a central heater needle and two thermistor needles located upstream and downstream of the heater (Köstner et al., 1996). The sap velocity sensors, manufactured by East30Sensors in Washington, were installed at breast height on the north-facing side of the stem and protected with a reflective cover (Hassler et al., 2018). Sap velocities at each of those locations were calculated.

ed from the temperatures measured at the corresponding thermistor pairs according to Equation (2) (Campbell et al., 1991):

$$V_{sap} = \frac{2k}{C_w(r_u + r_d)} \ln \left(\frac{\Delta T_u}{\Delta T_d} \right) \quad (2)$$

where V_{sap} is the sap velocity (m s^{-1}), k is the thermal conductivity of the sapwood, set to $0.5 \text{ W m}^{-1} \text{ K}^{-1}$, C_w is the specific heat of water ($\text{J m}^{-3} \text{ K}^{-1}$), r is the distance (m) from the heater needle to the sensor (in our case 6 mm) and ΔT is the temperature difference (K) before heating and 60 seconds after the heat pulse. Subscripts u and d stand for location upstream and downstream of the heater.

These values were corrected to account for wounding of the xylem tissue because of the drilling according to the numerical model solutions for the heat pulse velocity method as suggested by Burgess et al. (2001):

$$V_c = bV_{sap} + cV_{sap}^2 + dV_{sap}^3 \quad (3)$$

where V_c is the corrected sap velocity (m s^{-1}) and b , c and d are correction coefficients; for the 2-mm-wounds we have set $b = 1.8558$, $c = -0.0018 \text{ s m}^{-1}$, $d = 0.0003 \text{ s}^2 \text{ m}^{-2}$ (Burgess et al., 2001).

We selected a dataset of continuous sap velocity measurements from four trees. Daily mean values of the sap velocities were used for the photosynthetically active period from May to October 2015 in which there was a complete time series of sap flow measurements available. Simulated daily sums of actual transpiration from the model were then compared with the average sap velocity of the four trees at the site for the same period (growing season).

For better comparison sap velocities and simulated transpiration were normalized.

Evaluation of model behavior

In our definition, a model is “behavioral”, when it is able to simulate runoff generation, water balance and the temporal pattern of soil moisture and evapotranspiration consistently with the reality. Therefore, we propose a scheme including four qualitative performance evaluation criteria to check the simulated output. This scheme allows excluding simulations that are not realistic in terms of the four sources of information mentioned above (see Table 5). Sap velocity (SV) and soil moisture (SM) criteria define the necessity of temporal consistency between observed and simulated time series of transpiration and soil water content (by comparing stand transpiration simulations with sap velocity measurements, and by comparing simulated and observed temporal pattern of soil moisture, re-

spectively). Therefore, all simulated time series that would be less consistent with the temporal variability of observations will be rejected. Since actual evapotranspiration is usually less than precipitation in the water budget (Hasenmueller and Criss, 2013), the RETa (“Realistic amount of actual evapotranspiration”) criterion eliminates simulations in which the total amount of evapotranspiration exceeds 750 mm/year. According to our knowledge of local terrain properties and field surveys, RRGP (“Realistic Runoff Generation Process”) criterion was set to deep percolation as the most plausible hydrological process at our site.

In addition to the criteria mentioned above, three widely used statistical goodness-of-fit measures complement the qualitative evaluation of model performance: Mean absolute error (MAE), correlation (R^2) and Nash-Sutcliffe efficiency index (NSE) provide additional information on the goodness-of-fit between normalized simulated transpiration and normalized sap velocity (SV) and simulated and observed soil moisture (SM). MAE (Eq. 4) is a basic index (McKeen et al., 2005; Savage et al., 2013) derived from the mean error (difference) between simulated variable and observed variable with the same length and dimensions. This measure is recommended for model performance evaluation (Fox, 1981). It is calculated as follows:

$$MAE = N^{-1} \sum_{i=1}^N |P_i - O_i| \quad (4)$$

where N is the number of the cases, $i = 1, 2, 3, \dots, N$; P is the simulation time series, and O is the observation time series.

While MAE estimates the size of difference, the correlation index R^2 quantitatively estimates the agreement between observations and simulations. R^2 can be expressed as the squared ratio between the covariance and the multiplied standard deviations of the predicted and observed values. Higher R^2 value indicates higher correlation (Legates and McCabe, 1999; Willmott, 1982).

The Nash-Sutcliffe efficiency index (NSE), is dimensionless describing the relative error between simulations and measured data (Nash and Sutcliffe, 1970). It is calculated as:

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} = 1 - N \frac{RMSE^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (5)$$

where the NSE index demonstrates the normalized ratio of residual variance (noise) to the observation variance ranging between $-\infty$ and 1. An NSE value is considered to be acceptable when it ranges between 0 and 1. Fewer errors between simulations and observations always lead to a bigger NSE value and a better model performance. It is important to mention that a negative NSE value ($NSE < 0$) indicates a bad model performance that is even worse than the mean of the observed variable.

Table 5. Model performance evaluation criteria.

<i>Evaluation element</i>	<i>Description</i>	<i>Evaluation criterion</i>
Sap velocity measurements (SV)	Temporal pattern of sap velocities in terms of normalized values	SV criterion : There should be similar variability and no high deviations between the sap velocities and simulated transpiration amounts
Soil moisture measurements (SM)	Temporal pattern of soil moisture measurements in terms of mean values (%) for uppermost 50 cm of soil layer	SM criterion : There should be similar variability and no high deviations between the soil moisture measurements and simulated soil moisture amounts
Realistic amount of actual evapotranspiration (RETa)	Total amount of evapotranspiration as a component of the water budget in terms of mm/year	RETa criterion : Total evapotranspiration simulated should be between 450 to 750 mm/year
Realistic runoff generation process (RRGP)	Derived from runoff component of the water balance	RRGP criterion : The simulated runoff generation process should be deep percolation and no direct runoff as saturation or Hortonian overland flow

Based on the four evaluation criteria from Table 5 and three performance measures, unrealistic simulations will be eliminated from consideration to attain the best parameterization which provides an overall agreement among the combined performance criteria. Therefore, only under this condition, the simulation will be categorized as “behavioral”.

Applying three groups of scenarios (Table 6), we investigated the soil parameterization that reaches to the behavioral model. The simulation results of each scenario were evaluated by the model performance criteria and statistical goodness-of-fit measures. The soil parameterization in scenario A was taken from Tepee et al. (2003). In the two scenarios B1 and B2, the water retention curve was parameterized with three different variations of the van-Genuchten parameters according to (Sauer, 2007). In scenario C, the model performance was evaluated using the soil parameter set provided by Sprenger et al. (2016).

RESULTS

SCENARIO A: Model parameterization according to Tepee et al. (2003)

Since logged air or stone fraction may reduce maximum soil moisture at saturation by up to 30% (Mualem, 1974), θ_{sat} (saturated water content) is reduced in scenario A1 in three steps from 41% (baseline parameterization according to Tepee et al. (2003), see Table 2) to 35% and finally to 30%. Soil moisture simulated with the baseline parameterization of the soil Table ($\theta_{sat} = 41\%$) shows much higher values than the measurements (Figure 2). As the parameter θ_{sat} decreases, the simulated soil moisture values also decrease. Simulated soil water content with $\theta_{sat} = 30\%$ shows the highest similarity with the measurements. However, the simulated dynamics of the soil moisture simulations do not match the measured dynamics.

Sap velocity rose in the transition from spring to summer and it started to decrease again with the end of the summer (Figure 3). A rapid drop in the sap velocity was observed in June and August 2015 while there was a steep rise in July 2015 for all measuring points. The simulated transpiration with different θ_{sat} values and the sap velocity measurements have a similar temporal pattern. Changing the θ_{sat} value has only a negligible effect on transpiration (Figure 3).

Evaluation of the water balance (see Table A in the appendix) unveiled that total simulated evapotranspiration (775, 773 and 762 mm/year for $\theta_{sat} = 41, 35$ and 30 respectively) is too high. It is close to the annual precipitation amount (791 mm) which is not realistic. The dominant runoff

generation process was saturation excess flow or Hortonian overland flow which is not realistic according to landscape characteristics. Table 7 illustrates the model performance in scenario A1 evaluated by the three statistical efficiency measures as well as by four criteria. Meeting or not meeting a criterion is expressed in terms of “Yes” or “No” respectively. All simulations are highly correlated with the corresponding measurements ($R^2 \geq 0.73$ for all simulations). Model performances for transpiration show the same values for all θ_{sat} . While for the soil moisture, the simulation with $\theta_{sat} = 30$ Vol% shows the lowest bias (MAE = 0.02) and a positive NSE (0.55). This confirms the results obtained from visual inspection.

To investigate the effect of scaling the bulk surface resistance parameters (r_{sc} and r_{se}), in scenario A2, the parameters r_{sc} and r_{se} are adjusted in the evapotranspiration module of the WaSiM-ETH model. The applied percentage changes were 25, 50, 75, 150, 200 and 400% according to the standard values in the model for deciduous forest (Table 1). Parameter θ_{sat} was set to 30 Vol% due to the relatively satisfactory simulation results obtained from scenario A1.

Changing the bulk surface resistance parameters affects the simulated soil water content. The dynamics of the soil moisture simulations are now more consistent with measured values (Figure 2). The best fit could be obtained by decreasing the r_{sc} and r_{se} values to 75% and 50% of their standard values, respectively. By lowering the bulk surface resistance parameters, the (potential) evapotranspiration increases. This extracts more water through plant transpiration and soil evaporation. Hence, under these conditions, simulated soil moisture was reduced and became closer to the measured values. The dynamics of the simulated transpiration also corresponds well to the sap velocity measurements (Figure 3).

However, the amount of evapotranspiration losses (850 mm and 867 mm with $r_{sc} = 75\%$ and $r_{se} = 50\%$, respectively) exceeded precipitation input. The simulated runoff generation process was saturation excess flow or Hortonian overland flow which was unrealistic with regard to real soil characteristics at site Sa_G (Table A in appendix). Evaluation of the results obtained from the scenario A2 is shown in Table 8. Here, all measures indicate an almost perfect fit after scaling the bulk surface resistance (NSE = 0.74 for transpiration, and NSE = 0.85 or 0.91 for soil moisture). This confirms a substantial improvement of simulation accuracy. Nevertheless, runoff generation process and water balance are not correctly reproduced.

Table 6. Overview of different scenario combinations.

SCENARIO A (using soil parameterization after Tepee et al. (2003))	A1 (Scaling θ_{sat})	-41% -35% -30%	
	A2 Scaling bulk surface resistances $\theta_{sat} = 30\%$	A2-1 (Scaling soil surface resistance r_{se})	25%-50%-75%-100%- 150%-200%-400%
		A2-2 (Scaling canopy surface resistance r_{sc})	25%-50%-75%-100%- 150%-200%-400%
	SCENARIO B (using soil parameterization after Sauer (2007))	B1 Re-parameterization of Water Retention Curve with $\theta_{sat} = 30\%$	-Variation 1 -Variation 2 -Variation 3
B2 Re-parameterization of Water Retention Curve with $\theta_{sat} = 41\%$		-Variation 1 -Variation 2 -Variation 3	
SCENARIO C		Soil parameterization after Sprenger et al. (2016)	Comparison to best performing parameter set ($\theta_{sat} = 41\%$, Var1)

Table 7. Criteria evaluation and efficiency measures in scenario A1.

Scenario A1: scaling theta_sat						
Criterion	30 Vol%		35 Vol%		41 Vol%	
SV	Yes		Yes		Yes	
SM	No		No		No	
RETA	No		No		No	
RRGP	No		No		No	
Efficiency measure	Transpiration	Soil moisture	Transpiration	Soil moisture	Transpiration	Soil moisture
	R ²	0.73	0.95	0.73	0.95	0.73
MAE	0.42	0.02	0.42	0.08	0.42	0.13
NSE	0.71	0.55	0.71	-2.66	0.71	-8.92

Table 8. Criteria evaluation and efficiency measures in scenario A2.

Criterion	30 Vol% (baseline)		A2-1 ($r_{sc} = 50\%$)		A2-2 ($r_{sc} = 75\%$)	
SV	Yes		Yes		Yes	
SM	No		Yes		Yes	
RETA	No		No		No	
RRGP	No		No		No	
Efficiency measure	Transpiration	Soil moisture	Transpiration	Soil moisture	Transpiration	Soil moisture
	R ²	0.73	0.95	0.73	0.94	0.76
MAE	0.42	0.02	0.39	0.01	0.39	0.01
NSE	0.71	0.55	0.74	0.91	0.74	0.85

Table 9. Criteria evaluation and efficiency measures for model performances in scenarios B1 and B2.

Criterion	Re-parameterization of the water retention curve – scenario B															
	scenario B1 (theta_sat = 30 %)								scenario B2 (theta_sat = 41%)							
	Baseline		Var1		Var2		Var3		Baseline		Var1		Var2		Var3	
SV	Yes	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No
SM	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No
RETA	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RRGP	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Efficiency measure	SM	Tr	SM	Tr	SM	Tr	SM	Tr	SM	Tr	SM	Tr	SM	Tr	SM	Tr
	R ²	0.95	0.73	0.75	0.64	0.60	0.37	0.62	0.11	0.93	0.61	0.74	0.73	0.67	0.5	0.62
MAE	0.02	0.42	0.05	0.53	0.07	0.75	0.06	0.97	0.13	0.42	0.02	0.43	0.04	0.66	0.06	0.97
NSE	0.55	0.71	-0.44	0.6	-1.98	0.21	-1.64	-0.34	-8.92	0.71	0.65	0.71	-0.01	0.41	-1.64	-0.34

SCENARIO B: Re-parameterization of Water Retention Curve

In scenario B1, van Genuchten parameters of the baseline parametrization of the soil were re-parameterized according to Sauer (2007), where three variations of the parameters “alpha” and “n” were proposed for the same soil type “loamy sand”. Figure 2 depicts soil moisture of the three variations of van Genuchten parameters for a soil with theta_sat = 30%. All variants underestimate the measured values. Variation of van Genuchten parameters also affects simulated actual transpiration rates: In Figure 3, we clearly see that simulated transpiration does not match the temporal pattern of sap velocity measurements.

In all three variations, the soil was significantly dryer than the measured value. In variation 2 and 3 the soil water content came close to the residual water content. For the simulated transpiration, its temporal consistency with sap velocity decreased from variation 1 to 3. As all three variations performed worse than the best A1 scenario all three variants are rejected.

In scenario B2 we changed theta_sat from 30 to 41 Vol% and then repeated the three variations of alpha and n after Sauer (2007) and then checked both soil moisture (Figure 2) and

transpiration dynamics (Figure 3). Simulation results with theta_sat = 41% for the transpiration dynamics are relatively consistent with the observed sap velocities over the entire vegetation period for variation 1. In scenario B2, variation 1 provided sufficient soil water during the vegetation period for plant transpiration. This corresponds well to temporal patterns of sap velocities. Nevertheless, in variation 2 and 3, simulated transpiration did not reproduce the temporal patterns of the sap velocity data. There is a strong deviation in July and August 2015 and at some points the simulated transpiration drops to zero. This is the result of the low soil water content in summer (close to residual water content) for these two variations.

Table 9 provides all evaluation results related to the scenario B. Runoff generation process for all three variations with theta_sat = 30% and 41% is now deep percolation (Table A in appendix). Furthermore, the total amount of evapotranspiration was less than 750 mm for all variations. Variation 1 with theta_sat = 41% fulfills all four evaluation criteria. This is confirmed by the three statistical efficiency measures. Here, variation 1 with theta_sat = 41% clearly performs the best. In accordance with our definition we can label this model parameterization as “behavioral”.

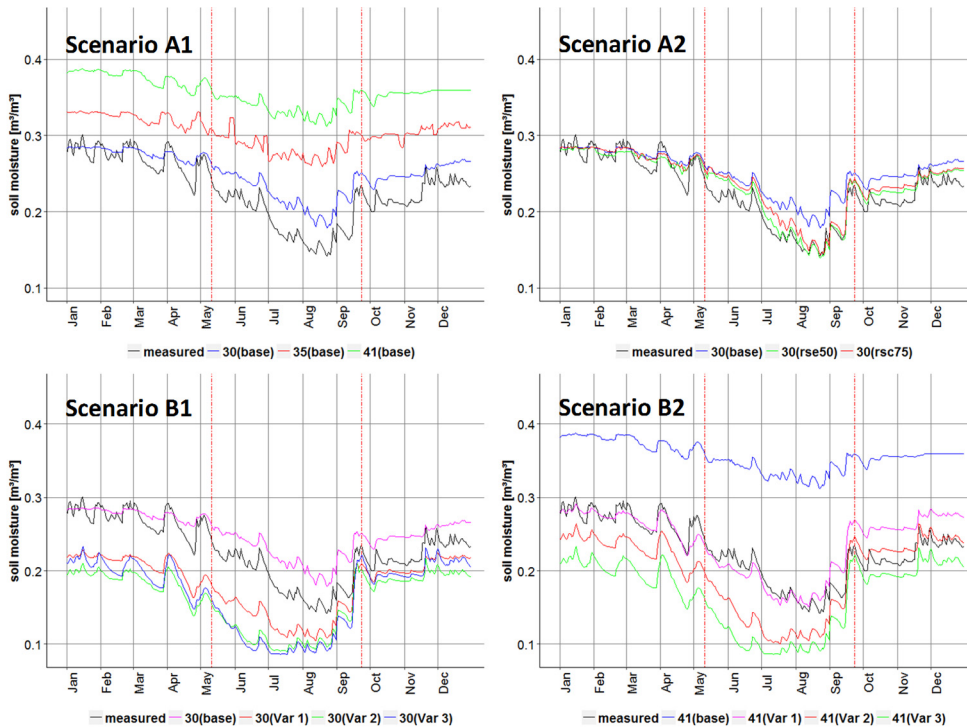


Fig. 2. Simulated and measured soil moisture in the root zone (Vol%) in 2015 for Scenarios A1, A2, B1 and B2. Red vertical lines indicate the growing season.

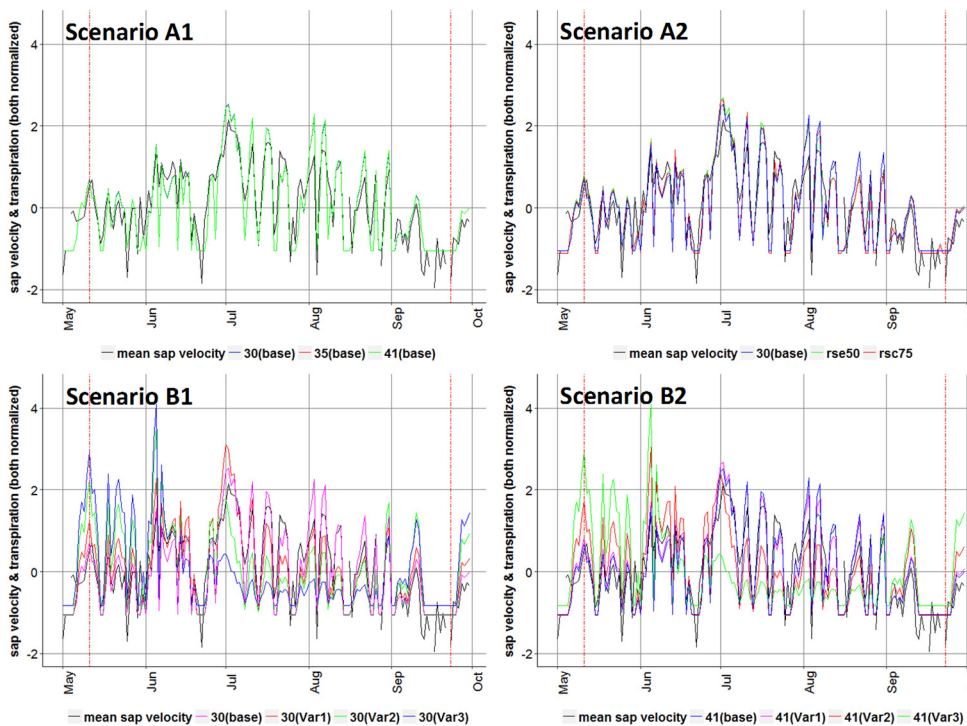


Fig. 3. Simulated transpiration (normalized) and normalized mean values of sap velocity (measured) for growing season 2015 (Scenarios A1, A2, B1, B2). In Scenario A1 all simulations show the same transpiration (identical lines).

SCENARIO C: Soil parameterization after Sprenger et al. (2016)

Sprenger et al. (2016) provided a soil parameter set for the Site Sa_G. In the scenario C, this parameter set was determined by automatic fitting to soil moisture measurements and stable isotope data. We compared this parameterization with our best performing model from the previous section (variation 1 from scenario B2 with $\theta_{sat} = 41\%$).

The simulated runoff generation process was deep percolation which is plausible (Table A in appendix). Total

evapotranspiration (602 mm) was estimated to be lower than precipitation (791 mm) which is also correct. But it can be seen in Figure 4 that soil moisture simulation does not show the correct dynamics compared to the measured time series. Here again, the parameterization of WRC is the reason that the soil is drying out in summer. This causes a significant reduction in transpiration during August which does not correspond to our sap velocity measurements. Additionally, the statistical efficiency measures (Table 10) reveal that the model performs very weak in simulating soil moisture (negative value for NSE). Simulated transpiration is therefore not consistent with the

corresponding sap velocity time series (Figure 5). It shows an overestimation in spring and a slight underestimation in summer, which is also indicated by lower model performance measures compared to the simulation with the optimal parameter set (B2, variation 1 with $\theta_{\text{sat}} = 41\%$) from the previous section.

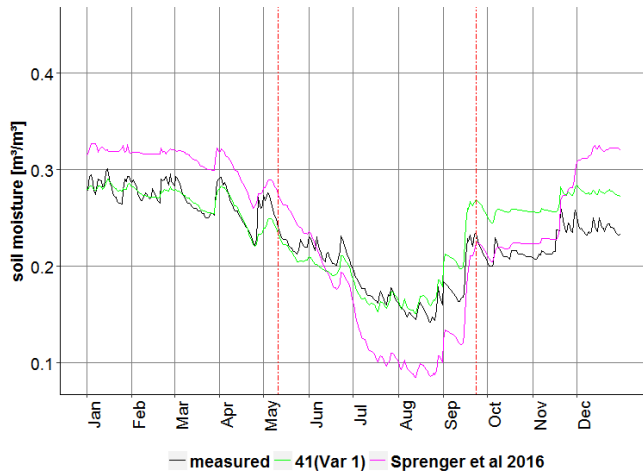


Fig. 4. Simulated soil water content for soil parameters according to Sprenger et al. 2016 compared to corresponding measured values and simulated soil moisture from variation 1 ($\theta_{\text{sat}} = 41\%$) in 2015. Red vertical lines indicate the growing season.

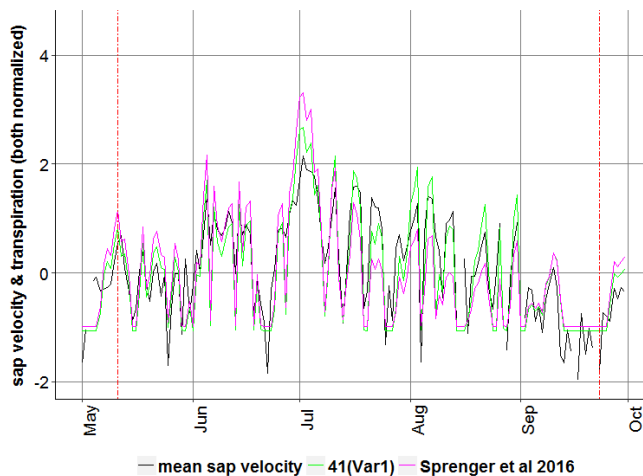


Fig. 5. Normalized simulated transpiration for soil parameters according to Sprenger et al. 2016 compared to normalized mean sap velocity (measured) and normalized simulated transpiration for variation 1 ($\theta_{\text{sat}} = 41\%$) in growing season 2015.

Table 10. Criteria evaluation and efficiency measures for soil parameterization according to Sprenger et al. (2016).

Criterion	Optimal parameter set ($\theta_{\text{sat}} = 41\%$, Variation 1)		Sprenger et al. 2016	
	Transpiration	Soil moisture	Transpiration	Soil moisture
SV	Yes		No	
SM	Yes		No	
REtr	Yes		Yes	
RRGP	Yes		Yes	
Efficiency measure				
	R ²	0.73	0.74	0.89
	MAE	0.43	0.02	0.55
	NSE	0.71	0.65	0.57

DISCUSSION

The main objective of this study was to build up a behavioral forest stand water balance model to characterize the temporal changes in hydrological components of water balance by making use of both observed soil moisture and sap velocities as well as expert knowledge of local runoff generation processes. A behavioral model was defined as a model in which simulation results have to be consistent with measurements of soil moisture and sap velocity and with our hydrologic understanding of runoff generation processes in the area of investigation. To accomplish the objectives, a multi-criteria evaluation scheme was developed. While 24 model realizations were tested, only one model realization could be categorized as “behavioral”.

Results of this study demonstrated that without the use of additional information (e.g. using sap velocity measurements for transpiration dynamics; different soil parameterizations, and expert knowledge), it is not possible to identify a model which captures these processes and dynamics adequately. This sheds light on the value of the contribution of different forms of data in representing the catchment behavior. In a case study in a Swiss Pre-Alpine catchment, it was also found that the application of expert knowledge and the concept of dominant processes can increase the realism of the hydrological models (Antonetti and Zappa, 2018). Taking into consideration that model evaluation would always be partly subjective, we looked at the model behavior from different perspectives through application of multi criteria evaluation that integrated this additional information. Therefore, we were able to select a behavioral parameter set from a number of equally likely soil parameterizations. The development of a multi-criteria approach for model evaluation is based on the consideration that a single measure of performance does not properly extract the information contained in the data (Gupta et al., 1998). This approach includes multiple performance measures and allows to evaluate if the hydrological model is able to represent the behavior of internal catchment processes (Fenicia et al., 2008b). Moreover, our results are also in line with Livneh (2012). He improved model performance significantly by the application of a multi-criteria scheme to evaluate multiple model outputs and by adding supplementary information in the parameterization process (Livneh, 2012). Another study showed that the introduction of constraints was efficient in reducing simulation uncertainty, in conditioning parameters, and in identifying critical parameters (Senapati et al., 2016).

All functions describing soil water retention imply a specific soil hydraulic behavior. Soil parameterization schemes according to Teepe et al. (2003) and Sauer (2007) use different amounts of soil information to derive pedo-transfer functions to translate soil information into van Genuchten parameters. Our results revealed that different parameterizations of the corresponding soil led to diverse simulation results. This issue is of great significance in all models applying the Richard’s equation (e.g. WaSiM-ETH). Therefore, finding a behavioral model for evapotranspiration is highly dependent on the identification of an appropriate WRC. This is consistent with the results of Garrigues et al. (2018). They compared the performance of two water transfer models in simulating evapotranspiration using different soil parameterizations. They found an unexpectedly high model sensitivity to soil moisture at field capacity, root extinction coefficient, and the proportion of homogeneous root distribution (Garrigues et al., 2018). In our proof-of-concept study based on a 1-D model, we took the measurements of sapflow and soil moisture as representative for the “sandstone area” where the dominant hydrologic process is deep percola-

tion. This made it possible to identify a behavioral model. It is known that models might work for the wrong reasons (i.e. reproducing discharge with incorrect process representations) (Beven, 2006; Walker and Zhang, 2002). This makes it advisable to implement expert knowledge to develop a proper parameterization to reflect our perceptions of the processes observed.

CONCLUSION

A considerable amount of effort is still being devoted to the development of hydrological models, and there is a continuing need to advance the techniques for their parameter estimation. It is also important to develop a good working knowledge of their sensitivity, and strengths and weaknesses.

This study underlines the importance of correctly setting up the 1-D water balance simulation model WaSiM-ETH in order to reproduce the dynamics of soil water fluxes and the physiological control of water loss through transpiration at a specific site (beech forest in Western Luxembourg). Adjustment of the parametrization of the WRC showed a high impact on simulation results. Our main finding was that: even though all parameter sets refer to the same soil (“loamy sand”), a slightly different parameterization of soil moisture at saturation (θ_{sat}), bulk surface resistance parameters and WRC may result in implausible model behavior. Even if transpiration and soil moisture are simulated consistently with our observations, runoff generation or total water balance may be wrongly estimated. Therefore, only the introduction of a multi-criteria evaluation scheme for exclusion of unrealistic outputs allowed finding a well performing parameter set for our test site. These findings suggest that using different sources of information such as expert knowledge on the dominant hydrological processes and the understanding of local controls facilitate parameterization and evaluation of a hydrological model. We should question the generally accepted procedure to parametrize soils using “default” parameter sets based on soil texture description or similar. Only if porosity and WRC for all soil horizons are correctly adjusted, a “physically based” model may simulate runoff processes and transpiration consistently with observations. Only in this case, we may refer to a model as “behavioral” (Gupta et al., 2005). We recommend finding “prototype soils” which are in accordance with soil description (e.g. texture) and expert knowledge on runoff processes in the area under investigation. This in turn implies that model parameterization, evaluation or calibration has to incorporate this “soft” knowledge.

Setups identified as optimal for 1-D simulations will go a long way of improving the application of WaSiM-ETH water balance model on catchment scale to answer questions about watershed characteristics and water resources management. Since point measurements are not valid on catchment scale, we may try to address the spatiotemporal distribution of evapotranspiration, soil moisture and runoff generation processes at catchment scale as well as the estimation of overall water balance at the corresponding gaging station(s) (Koch et al., 2016, 2015). The current study showed that soil parameterization affects not only the temporal distribution of soil moisture and transpiration, but also the runoff generation process. This also highlights the need to consider the incorporation of several data products to increase knowledge about the hydrological processes on catchment scale (Casper et al., 2015). Remotely sensed data will open up the possibility to analyze spatial patterns of actual evapotranspiration (ETa) or soil moisture (Koch et al., 2017). Together with additional knowledge of the spatial distribution of dominant runoff processes on catchment scale this

will facilitate the parameterization of the hydrological model WaSiM-ETH and its subsequent optimization by extending the traditional model evaluation procedure at gaging stations with the search for a best fit of spatial patterns of ETa and runoff processes on catchment scale. A number of automatic mapping approaches for delineation of dominant runoff process exist, which can be used to constrain the uncertainty of hydrological simulations (Antonetti et al., 2016; Behrens et al., 2010). The model RoGeR (Runoff Generation Research) demonstrated its ability to quantify runoff process in high spatial and temporal resolution without the need of parameter calibration (Steinbrich et al., 2016). This approach combines knowledge of runoff process gained through long term research with spatially distributed data sets and can thus be used to extend the here presented approach to the catchment scale.

Acknowledgements. The data was collected within the DFG Research Unit FOR 1598 “Catchments As Organized Systems (CAOS)” in a project led by Markus Weiler and Theresa Blume. All datasets can be obtained from Theresa Blume (blume@gfz-potsdam.de) upon request. We thank the technicians Britta Kattenstroth and Tobias Vetter for the maintenance of the sensor network.

REFERENCES

- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., others, 1998. Crop evapotranspiration-guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. FAO Rome 300, D05109.
- Anderson, M.C., Norman, J.M., Mecikalski, J.R., Otkin, J.A., Kustas, W.P., 2007. A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *J. Geophys. Res.-Atmospheres*, 112, Article Number: D10117.
- Antonetti, M., Buss, R., Scherrer, S., Margreth, M., Zappa, M., 2016. Mapping dominant runoff processes: an evaluation of different approaches using similarity measures and synthetic runoff simulations. *Hydrol. Earth Syst. Sci.*, 20, 2929–2945. <https://doi.org/10.5194/hess-20-2929-2016>
- Antonetti, M., Zappa, M., 2018. How can expert knowledge increase the realism of conceptual hydrological models? A case study based on the concept of dominant runoff process in the Swiss Pre-Alps. *Hydrol. Earth Syst. Sci.*, 22, 4425–4447.
- Ayyoub, A., Er-Raki, S., Khabba, S., Merlin, O., Ezzahar, J., Rodriguez, J., Bahlaoui, A., Chehbouni, A., 2017. A simple and alternative approach based on reference evapotranspiration and leaf area index for estimating tree transpiration in semi-arid regions. *Agric. Water Manag.*, 188, 61–68.
- Bahreman, A., 2016. HESS Opinions: Advocating process modeling and de-emphasizing parameter estimation. *Hydrol. Earth Syst. Sci.*, 20, 1433–1445.
- Behrens, T., Zhu, A.-X., Schmidt, K., Scholten, T., 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geoderma*, 155, 175–185. <http://dx.doi.org/10.1016/j.geoderma.2009.07.010>
- Beven, K., 2006. A manifesto for the equifinality thesis. *J. Hydrol.*, 320, 18–36.
- Beven, K., 1979. A sensitivity analysis of the Penman-Monteith actual evapotranspiration estimates. *J. Hydrol.*, 44, 169–190.
- Bie, W., Casper, M.C., Reiter, P., Vohland, M., 2015. Surface resistance calibration for a hydrological model using evapotranspiration retrieved from remote sensing data in Nahe catchment forest area. *Proc. Int. Assoc. Hydrol. Sci.*, 368, 81–86. <https://doi.org/10.5194/piahs-368-81-2015>
- Blöschl, G., 2001. Scaling in hydrology. *Hydrol. Process.*, 15, 709–711.
- Bromley, J., Jackson, N.A., Clymer, O., Giacomello, A.M., Jensen, F.V., 2005. The use of Hugin® to develop Bayesian networks as an aid to integrated water resource planning. *Environ. Model. Softw.*, 20, 231–242.

- Burgess, S.S., Adams, M.A., Turner, N.C., Beverly, C.R., Ong, C.K., Khan, A.A., Bleby, T.M., 2001. An improved heat pulse method to measure low and reverse rates of sap flow in woody plants. *Tree Physiol.*, 21, 589–598.
- Campbell, G., Calissendorff, C., Williams, J., 1991. Probe for measuring soil specific heat using a heat-pulse method. *Soil Sci. Soc. Am. J.*, 55, 291–293.
- Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Guston, D.H., Jäger, J., Mitchell, R.B., 2003. Knowledge systems for sustainable development. *Proc. Natl. Acad. Sci.*, 100, 8086–8091.
- Casper, M.C., Gronz, O., Gemmar, P., 2015. Process-oriented parameterisation and calibration of a water balance model. *Hydrol. Wasserbewirtsch.*, 59, 136–144.
- Casper, M.C., Vohland, M., 2008. Validation of a large scale hydrological model with data fields retrieved from reflective and thermal optical remote sensing data – A case study for the Upper Rhine Valley. *Phys. Chem. Earth Parts ABC*, 33, 1061–1067. <http://dx.doi.org/10.1016/j.pce.2008.06.001>
- Cullmann, J., Mishra, V., Peters, R., 2006. Flow analysis with WaSiM-ETH? model parameter sensitivity at different scales. *Adv. Geosci.*, 9, 73–77.
- Droogers, P., Allen, R.G., 2002. Estimating reference evapotranspiration under inaccurate data conditions. *Irrig. Drain. Syst.*, 16, 33–45.
- Durigon, A., Van Lier, Q.D.J., Metselaar, K., 2016. Forcing variables in simulation of transpiration of water stressed plants determined by principal component analysis. *Int. Agrophysics*, 30, 431–445.
- Elfert, S., Bormann, H., 2010. Simulated impact of past and possible future land use changes on the hydrological response of the Northern German lowland ‘Hunte’ catchment. *J. Hydrol.*, 383, 245–255.
- Federer, C.A., Lash, D., 1978. Simulated streamflow response to possible differences in transpiration among species of hardwood trees. *Water Resour. Res.*, 14, 1089–1097.
- Fenicia, F., McDonnell, J.J., Savenije, H.H., 2008a. Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resour. Res.*, 44, 6, Article Number: W06419.
- Fenicia, F., Savenije, H.H., Matgen, P., Pfister, L., 2008b. Understanding catchment behavior through stepwise model concept improvement. *Water Resour. Res.*, 44, 1, Article Number: W01402.
- Fox, D.G., 1981. Judging air quality model performance. *Bull. Am. Meteorol. Soc.*, 62, 599–609.
- Franks, S.W., Gineste, P., Beven, K.J., Merot, P., 1998. On constraining the predictions of a distributed model: the incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resour. Res.*, 34, 787–797.
- Garrigues, S., Boone, A., Decharme, B., Olioso, A., Albergel, C., Calvet, J.-C., Moulin, S., Buis, S., Martin, E., 2018. Impacts of the soil water transfer parameterization on the simulation of evapotranspiration over a 14-year Mediterranean crop succession. *J. Hydrometeorol.*, 19, 3–25.
- Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., Savenije, H., 2014. Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrol. Earth Syst. Sci.*, 18, 4839.
- Grayson, R., Blöschl, G., 2001. Summary of pattern comparison and concluding remarks. In: Grayson, R., Blöschl, G. (Eds): *Spatial Patterns in Catchment Hydrology – Observations and Modelling*. Cambridge University Press, Cambridge, UK, pp. 355–367.
- Grigoryan, G.V., Casper, M.C., Gauer, J., Vasconcelos, A.C., Reiter, P.P., 2010. Impact of climate change on water balance of forest sites in Rhineland-Palatinate, Germany. *Adv. Geosci.*, 27, 37–43. <https://doi.org/10.5194/adgeo-27-37-2010>
- Gupta, H.V., Beven, K.J., Wagener, T., 2005. Model calibration and uncertainty estimation. In: Anderson, M. (Ed.): *Encyclopedia of Hydrological Sciences*, Vol. 3, Chapter 131, pp. 2015–2032.
- Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.*, 34, 751–763.
- Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., Vitvar, T., 2003. A comparative study in modelling runoff and its components in two mountainous catchments. *Hydrol. Process.*, 17, 297–311.
- Hasenmueller, E.A., Criss, R.E., 2013. Water balance estimates of evapotranspiration rates in areas with varying land use. In: Alexandris, S. (Ed.): *Evapotranspiration-An Overview*. IntechOpen, DOI: 10.5772/52811.
- Hassler, S.K., Weiler, M., Blume, T., 2018. Tree-, stand- and site-specific controls on landscape-scale patterns of transpiration. *Hydrol. Earth Syst. Sci.*, 22, 13–30.
- Haude, W., 1955. Zur Bestimmung der Verdunstung auf möglichst einfache Weise. Dt. Wetterdienst, Bad Kissingen.
- Holst, J., Grote, R., Offermann, C., Ferrio, J.P., Gessler, A., Mayer, H., Rennenberg, H., 2010. Water fluxes within beech stands in complex terrain. *Int. J. Biometeorol.*, 54, 23–36.
- Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H., Gascuel-Oudou, C., 2014. Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resour. Res.*, 50, 7445–7469.
- Jasper, K., 2001. Hydrological modelling of Alpine river catchments using output variables from atmospheric models (PhD Thesis). ETH Zurich.
- Jasper, K., Gurtz, J., Lang, H., 2002. Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. *J. Hydrol.*, 267, 40–52.
- Juilleret, J., Iffly, J.-F., Hoffmann, L., Hissler, C., 2012. The potential of soil survey as a tool for surface geological mapping: a case study in a hydrological experimental catchment (Huewelerbach, Grand-Duchy of Luxembourg). *Geologica Belgica*, 15, 1–2, 36–41.
- Klok, E., Jasper, K., Roelofsma, K., Gurtz, J., Badoux, A., 2001. Distributed hydrological modelling of a heavily glaciated Alpine river basin. *Hydrol. Sci. J.*, 46, 553–570.
- Koch, J., Jensen, K.H., Stisen, S., 2015. Toward a true spatial model evaluation in distributed hydrological modeling: Kappa statistics, Fuzzy theory, and EOF-analysis benchmarked by the human perception and evaluated against a modeling case study. *Water Resour. Res.*, 51, 1225–1246. <https://doi.org/10.1002/2014WR016607>
- Koch, J., Mendiguren, G., Mariethoz, G., Stisen, S., 2017. Spatial sensitivity analysis of simulated land surface patterns in a catchment model using a set of innovative spatial performance metrics. *J. Hydrometeorol.*, 18, 1121–1142. <https://doi.org/10.1175/JHM-D-16-0148.1>
- Koch, J., Siemann, A., Stisen, S., Sheffield, J., 2016. Spatial validation of large-scale land surface models against monthly land surface temperature patterns using innovative performance metrics. *J. Geophys. Res. Atmospheres*, 121, 5430–5452.
- Köstner, B., Biron, P., Siegwolf, R., Granier, A., 1996. Estimates of water vapor flux and canopy conductance of Scots pine at the tree level utilizing different xylem sap flow methods. *Theor. Appl. Climatol.*, 53, 105–113.
- Kramer, P.J., Boyer, J.S., 1995. *Water Relations of Plants and Soils*. Academic Press.
- Legates, D.R., McCabe, G.J., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.*, 35, 233–241.
- Livneh, B., 2012. Development of a unified land model with multi-criteria observational data for the simulation of regional hydrology and land-atmosphere interaction. PhD Thesis. University of Washington, Seattle, USA.
- Lu, P., Urban, L., Zhao, P., 2004. Granier’s thermal dissipation probe (TDP) method for measuring sap flow in trees: theory and practice. *ACTA Bot. Sin. (Engl. Ed.)*, 46, 631–646.
- Martínez-Carreras, N., Krein, A., Gallart, F., Iffly, J.-F., Hissler, C., Pfister, L., Hoffmann, L., Owens, P.N., 2012. The influence of sediment sources and hydrologic events on the nutrient and metal content of fine-grained sediments (Attert River basin, Luxembourg). *Water. Air. Soil Pollut.*, 223, 5685–5705.
- Martínez-Carreras, N., Udelhoven, T., Krein, A., Gallart, F., Iffly, J.F., Ziebel, J., Hoffmann, L., Pfister, L., Walling, D.E., 2010. The use of sediment colour measured by diffuse reflectance spectrometry to determine sediment sources: application to the Attert River catchment (Luxembourg). *J. Hydrol.*, 382, 49–63.
- McKeen, S., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Hsie, E.-Y., Gong, W., Bouchet, V., Menard, S., Moffet, R., others, 2005. Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004. *J. Geophys. Res.-Atmospheres*, 110, Article Number: D21307.
- Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J.C.,

Lang, H., Parmet, B.W.A.H., Schädler, B., Schulla, J., Wilke, K., 2001. Impact of climate change on hydrological regimes and water resources management in the Rhine basin. *Clim. Change*, 49, 105–128.

Mohajerani, H., Kholghi, M., Mosaedi, A., Farmani, R., Sadoddin, A., Casper, M., 2017. Application of Bayesian decision networks for groundwater resources management under the conditions of high uncertainty and data scarcity. *Water Resour. Manag.*, 31, 1859–1879.

Monteith, J., 1981. Evaporation and surface temperature. *Q. J. R. Meteorol. Soc.*, 107, 1–27.

Monteith, J., Szeicz, G., Waggoner, P., 1965. The measurement and control of stomatal resistance in the field. *J. Appl. Ecol.*, 345–355.

Mualem, Y., 1974. A conceptual model of hysteresis. *Water Resources Research*, 10, 514–520.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.*, 10, 282–290.

Paço, T.A., Pôças, I., Cunha, M., Silvestre, J.C., Santos, F.L., Paredes, P., Pereira, L.S., 2014. Evapotranspiration and crop coefficients for a super intensive olive orchard. An application of SIMDualKc and METRIC models using ground and satellite observations. *J. Hydrol.*, 519, 2067–2080.

Pfister, L., Humbert, J., Hoffmann, L., 2000. Recent trends in rainfall-runoff characteristics in the Alzette river basin, Luxembourg. *Clim. Change*, 45, 323–337.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Physics*, 1, 318–333.

Sauer, T., 2007. Modellierung von Bodenwasserhaushalt und Abflussprozessen auf der Plotskale in Abhängigkeit von Substrat und Landnutzung (Dissertation). University of Trier, Trier, Germany.

Saugier, B., Granier, A., Pontailler, J., Dufrene, E., Baldocchi, D., 1997. Transpiration of a boreal pine forest measured by branch bag, sap flow and micrometeorological methods. *Tree Physiol.*, 17, 511–519.

Savage, N., Agnew, P., Davis, L., Ordóñez, C., Thorpe, R., Johnson, C., O’Connor, F., Dalvi, M., 2013. Air quality modelling using the Met Office Unified Model (AQUAM OS24-26): model description and initial evaluation. *Geosci. Model Dev.*, 6, 353.

Schaap, M.G., Leij, F.J., Van Genuchten, M.T., 2001. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.*, 251, 163–176.

Schulla, J., 2017. Model Description WaSiM (Water balance Simulation Model), completely revised version 2017. Zür. Switz. Hydrol. Softw. Consult., 347.

Schulla, J., 1997. Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen. PhD Thesis. ETH Zurich.

Seibert, J., McDonnell, J.J., 2002. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resour. Res.*, 38, 11, Article Number: 1241.

Senapati, N., Jansson, P.-E., Smith, P., Chabbi, A., 2016. Modelling heat, water and carbon fluxes in mown grassland under multi-objective and multi-criteria constraints. *Environ. Model. Softw.*, 80, 201–224.

Sprenger, M., Seeger, S., Blume, T., Weiler, M., 2016. Travel times in the vadose zone: Variability in space and time. *Water Resour. Res.*, 52, 5727–5754.

Steinbrich, A., Leister, H., Weiler, M., 2016. Model-based quantification of runoff generation processes at high spatial and temporal resolution. *Environ. Earth Sci.*, 75, 1423.

Teepe, R., Dilling, H., Beese, F., 2003. Estimating water retention curves of forest soils from soil texture and bulk density. *J. Plant Nutr. Soil Sci.*, 166, 111–119.

Van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils 1. *Soil Sci. Soc. Am. J.*, 44, 892–898.

Van Genuchten, M.T., Leij, F.J., Yates, S.R., Williams, J.R., 1991. The RETC code for quantifying the hydraulic functions of unsaturated soils. U.S. Salinity Laboratory, USDA, Riverside, California.

Verbunt, M., Gurtz, J., Jasper, K., Lang, H., Warmerdam, P., Zappa, M., 2003. The hydrological role of snow and glaciers in alpine river basins and their distributed modeling. *J. Hydrol.*, 282, 36–55.

Vose, J.M., Harvey, G.J., Elliott, K.J., Clinton, B.D., 2003. Measuring and modeling tree and stand level transpiration. In: Lehr, J.H., Keeley, J. (Eds.): *Water Encyclopedia*, Volume 3, Surface and Agricultural Water. Wiley, pp. 732–740.

Wagner, T., Boyle, D.P., Lees, M.J., Wheeler, H.S., Gupta, H.V., Sorooshian, S., 2001. A framework for development and application of hydrological models. *Hydrol. Earth Syst. Sci.*, 5, 13–26.

Walker, G.R., Zhang, L., 2002. Plot Scale Models and their Application to Recharge Studies. Part 10 of Basics of Recharge and Discharge Series. CSIRO Publishing.

Wendling, U., 1975. Zur Messung und Schätzung der potentiellen Verdunstung. *Z. Für Meteorol.*, 25, 103–111.

Willmott, C.J., 1982. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.*, 63, 1309–1313.

Wilson, K.B., Hanson, P.J., Mulholland, P.J., Baldocchi, D.D., Wullschleger, S.D., 2001. A comparison of methods for determining forest evapotranspiration and its components: sap-flow, soil water budget, eddy covariance and catchment water balance. *Agric. For. Meteorol.*, 106, 153–168.

Zehe, E., Ehret, U., Pfister, L., Blume, T., Schröder, B., Westhoff, M., Jackisch, C., Schymanski, S. J., Weiler, M., Schulz, K., Allroggen, N., Tronicke, J., van Schaik, L., Dietrich, P., Scherer, U., Eccard, J., Wulfmeyer, V., Kleidon, A., 2014. HESS Opinions: From response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments. *Hydrol. Earth Syst. Sci.*, 18, 4635–4655.

Received 30 April 2018
Accepted 14 November 2018

APPENDIX

Table A. Water balance for all simulation runs, DP = Deep Percolation, SOF = Saturation Overland Flow, HOF = Hortonian Overland Flow.

Scenario	A1			A2		B1:			B2			C
	base-line 41	baseline 35%	baseline 30%	r _{sc} 75%	r _{se} 50%	Var. 1 30%	Var. 2 30%	Var. 3 30%	Var. 1 41%	Var. 2 41%	Var. 3 41%	
Pot. Evaporation	270	270	270	281	471	270	270	270	270	270	270	270
Real Evaporation	261	259	250	254	372	174	167	179	183	168	179	160
Interception Evaporation	146	146	146	157	146	146	146	146	146	146	146	146
ETp	851	851	851	978	1051	851	851	851	851	851	851	851
ETr	776	774	762	850	867	613	536	497	676	569	497	602
ETr_Layer1 = Transpiration	368	368	365	438	348	293	222	172	347	255	172	296
Baseflow	145	166	107	117	73	212	278	316	251	267	316	196
Direct Runoff	161	11	82	63	73	0	0	0	0	0	0	0
Interflow	0	0	0	0	0	0	0	0	0	0	0	0
Total Runoff	306	177	190	180	146	212	278	316	251	267	316	196
GW recharge	68	114	94	103	64	190	268	307	130	248	307	156
Delta Storage	-290	-160	-161	-239	-222	-33	-22	-22	-136	-45	-22	-7
Precipitation	791	791	791	791	791	791	791	791	791	791	791	791
Total Balance Error	0	0	0	0	0	0	0	0	0	0	0	0
Runoff Process	SOF/HOF	SOF/HOF/DP	SOF/HOF	SOF/HOF	SOF/HOF	DP	DP	DP	DP	DP	DP	DP