Hydrological hysteresis and its value for assessing process consistency in catchment conceptual models

Fovet O.[1,2], Ruiz L.[1,2], Hrachowitz M.[3] , Faucheux M. [1,2],Gascuel-Odoux C. [1,2]

[1] INRA, UMR1069 SAS, 65 route de Saint Brieuc, 35042 Rennes, France,

[2] Agrocampus Ouest, UMR1069 SAS, 65 route de Saint Brieuc, 35042 Rennes, France

[3] Delft University of Technology, Water Resources Section, Faculty of Civil Engineering and Applied Geosciences, Stevinweg 1, 2600 GA Delft, The Netherlands

Correspondance to: Ophelie Fovet, *ophelie.fovet@rennes.inra.fr*

Abstract

While most hydrological models reproduce the general flow dynamics, they frequently fail to adequately mimic system internal processes. In particular, the relationship between storage and discharge, which often follows annual hysteretic patterns in shallow hard-rock aquifers, is rarely considered in modelling studies. One main reason is that catchment storage is difficult to measure and another one is that objective functions are usually based on individual variables time series (e.g. the discharge). This reduces the ability of classical procedures to assess the relevance of the conceptual hypotheses associated with models.

We analyzed the annual hysteric patterns observed between stream flow and water storage both in the saturated and unsaturated zones of the hillslope and the riparian zone of a headwater catchment in French Brittany (ORE AgrHys). The saturated zone storage was estimated using distributed shallow groundwater levels and the unsaturated zone storage using several moisture profiles. All hysteretic loops were characterized by a hysteresis index. Four conceptual models, previously calibrated and evaluated for the same catchment, were assessed with respect to their ability to reproduce the hysteretic patterns.

The observed relationship between stream flow, saturated, and unsaturated storages led to identify four hydrological periods and emphasized a clearly distinct behaviour between riparian and hillslope groundwaters. Although all the tested models were able to produce an

annual hysteresis loop between discharge and both saturated and unsaturated storage, integration of a riparian component led to overall improved hysteretic signatures, even if some misrepresentation remained. Such systems-like approach is likely to improve model selection.

Keywords: hydrological model, hysteresis index, model consistency, groundwater, soil moisture, catchment storage.

1 Introduction

Rainfall-runoff models are tools that mimic the low-pass filter properties of catchments. Specifically they aim at reproducing observed stream flow time series by routing time series of meteorological drivers through a sequence of mathematically formalized processes that allow a temporal dispersion of the input signals in a way that is consistent with the modeller's conception of how the system functions. The core of most models, in particular in temperate, humid climates, dominated by some type of subsurface flow, is a series of storage-discharge functions that, in most general terms, express system output (i.e. discharge and evaporation) as function of the system state (i.e. storage), thereby generating a signal that is attenuated and lagged with respect to the input signal (i.e. precipitation).

However, modelling efforts on the catchment scale typically face the problem that on that scale neither integrated internal fluxes nor the integrated storage and the partitioning between different storage components at a given time, can be easily observed within limited uncertainty. Indeed, indicators of catchment storage such as groundwater levels and soil water content can be highly variable in space, and exhibiting heterogeneous spatio-temporal dynamics. While spatial aggregation of storage estimates (e.g. catchment averages) in lumped models may lead to a loss of crucial information and thus to overly-simplistic representations of reality, allowing for explicit incorporation of spatial storage heterogeneity in (semi) distributed models may prove elusive in the presence of data error and the frequent absence of detailed spatial knowledge of the properties of the flow domain. A time-series of groundwater table level from a single piezometer is not representative of the behaviour of the groundwater, even at the hillslope scale; therefore it is difficult to link it with either a reservoir volume simulated by a lumped model or an average water table level of a grid point simulated by a fully distributed model. These problems were recently addressed in some studies that intended to assess catchment storage using all available data (McNamara et al., 2011; Tetzlaff et al., 2011) and showing the importance of this storage in thresholds observed in the response of discharge to precipitation in catchments. For example, Spence (2010) argued that the observed non-linear relationships between stream flow and catchment storage (i.e. no unique storage-discharge relations) are the manifestation of thresholds occurring in catchment runoff generation. Thus, depending on the structure of the system, storage-discharge dynamics can exhibit hysteretic patterns, i.e. the system response depends on the history and the memory of the system (e.g. Everett and Whitton, 1952; Ali et al., 2011; Gabrielli et al., 2012; Haught and van Meerveld, 2011). Andermann et al. (2012) found a hysteretic relationship between precipitation and discharge in both glaciated and unglaciated catchments in the Himalaya Mountains that was shown to be due to groundwater storage rather than to snow or glacier melt. Hrachowitz et al. (2013a), demonstrating the presence of hysteresis in the distribution of water ages, highlighted the importance of an adequate characterization of all system-relevant internal states at a given time, to predict the system response within limited uncertainty as flow can be generated from different system components depending on the wetness state of the system.

In catchment-scale rainfall-runoff models, the need for calibration remains inevitable (Beven, 2001) due to the presence of data errors (e.g. Beven, 2013), and to the typically oversimplified process representations (e.g. Gupta et al., 2012). In spite of their comparatively high degrees of freedom, such models are frequently evaluated only against one single observed output variable, e.g. stream flow. Although the calibrated models may then adequately reproduce the output variable, model equifinality (e.g. Savenije, 2001) will lead to many apparently feasible solutions that do not sufficiently well reproduce system internal dynamics as they are mere artefacts of the mathematical optimization process rather than suitable representations of reality (Gharari et al., 2013; Hrachowitz et al., 2013b; Andréassian et al., 2012; Beven, 2006; Kirchner, 2006). The understanding for the need for multi-variable and -objective model evaluation strategies to identify and discard solutions that do not satisfy all evaluation criteria applied is therefore gaining ground (e.g. Freer et al., 1996; Gupta et al., 1998; Gupta et al., 2008, Gascuel-Odoux et al., 2010), as this will eventually lead to models that are not only capable of reproducing the observed output variables (e.g. stream flow) but that also represent the system internal dynamics in a more realistic way (Euser et al., 2013). The value of such multi-variable and/or -objective evaluation strategies has been demonstrated in the past, for example using groundwater levels (e.g. Fenicia et al., 2008; Molenat et al., 2005, Giustolisi and Simeone, 2006; Freer et al., 2004; Seibert, 2000; Lamb et al., 1998), soil moisture (Kampf and Burges, 2007; Parajka et al., 2006), saturated areas extension (Franks et al., 1998), snow cover patterns (e.g. Nester et al., 2012), remotely sensed evaporation, (e.g. Mohamed et al., 2006; Winsemius et al., 2008), stream flow at sub-catchment outlets (e.g. Moussa et al., 2007), and even water quality data such as e.g. chloride concentrations (Hrachowitz et al., 2011), atmospheric tracers (Molenat et al., 2013) or nitrates and sulfate concentrations (Hartmann et al, 2013 a), and water isotopes such as $\delta^{18}O$ (Hartmann et al., 2013 b). However, most studies using multiple response variables only evaluate them individually to identify Pareto optimal solutions. This practice may result in the loss of critical information such as the timing between the multiple variables. In other words it is conceivable that model calibration leads to Pareto-optimal solutions with adequate model performance for all variables, while at the same time misrepresenting the dynamics between these variables. Rather, using a synthetic catchment property (Sivapalan et al., 2005) or a hydrological signature (Wagener and Montanari, 2011; Yadav et al., 2007), combining different variables into one function, may potentially serve as a instructive diagnostic tool or as a calibration objective or even as a metric for catchment classification (Wagener, 2007).

Hysteretic patterns between hydrological variables are potentially good candidates to build such tools. The objective of this paper is to explore i) the potential of using annual hysteric patterns observed between stream flow and water storage both in the saturated and unsaturated zones of the hillslope and of the riparian zone for characterizing the hydrological functioning of a small headwater catchment in French Brittany (ORE AgrHys) and ii) to which degree a suite of conceptual rainfall-runoff models with increasing complexity, which were calibrated and evaluated for this catchment in previous work, using a flexible modelling framework (Hrachowitz et al., 2014), can reproduce the observed storage-discharge hysteresis and iii) if the use of the storage-discharge hysteresis can provide additional information for model diagnostics compared to traditional model evaluation metrics.

2 Materials and Methods

2.1 Study sites

Kerrien (10.5 ha) is a headwater catchment located in South-western French Brittany (47°, 35' N; 117°52' E, see Figure 1). Elevations range from 14 to 38 m a.s.l., slopes are less than 8.5%. The climate is oceanic, with mean annual temperature of 11.9°C, minimum of 5.9°C in winter and maximum of 17.9°C in summer. Mean annual rainfall over the period 1992-2012 is 1113 mm (+/-20%) and mean annual Penman potential evapotranspiration (PET) is 700 mm (+/- 4%). Mean annual drainage is 360 mm (+/- 60%) at the outlet. There is a high water deficit in the annual budget almost each year due to underflows below the outlet (Ruiz et al., 2002). The catchment is laying under granite called leucogranodiorite of Plomelin, which upper part is weathered on 1 to more than 20 m deep. Soils are mainly sandy loam with an upper horizon rich in organic matter, depths are comprised between 40 and 90 cm. Soils are well drained except in the bottomlands which represent 7% of the total area. Agriculture dominates the land use with 86% of the total area covered by grassland, maize and wheat, none of them irrigated. The base flow index is about 80 to 90%, thus the hillslope aquifer is the main contributor to stream (Molenat et al., 2008; Ruiz et al., 2002). Both stream flow and shallow groundwater tables exhibit a strong annual seasonality in this catchment (Fig. 2 and 3a.)

2.2 Data

Meteorological data were recorded in an automatic weather station (CIMEL, Figure 1) which provides hourly rainfall and variables required to estimate daily Penman PET (net solar radiation, air and soil temperatures, wind speed and direction). Discharge was calculated from water level measurements at the outlet (Figure 1) using a V-notch weir equipped with a shaft encoder with integrated Data Logger (OTT Thalimedes) recorded every 10 min since 2000 (E3). Groundwater levels were monitored every 15 min since 2001 in 3 piezometers F1b, F4, and F5b (Figure 1) using vented pressure probe sensors (OTT Orpheus Mini).

Moisture in the unsaturated zone was recorded every 30 minutes since July 2010, at 7 depth (25, 55, 85, 125, 165, 215, and 265 cm), on 2 profiles sB1 and sB2 (Figure 1), using capacitive probes which provide volumic humidity based on Frequency Domain Reflectometry (EnvironScan SenteK). Due to technical problems, data are missing in December 2012 and January 2013, then only two complete water years were available (2010-2011 and 2011-2012). In summary, stream discharge, water table levels and soil moisture were considered for the years 2002-2012, 2002-2012, and 2010-2012 respectively.

2.3 Catchment storage estimates

In order to obtain a proxy for the saturated zone storage at the catchment scale, the time series of groundwater level were normalized between their minimal and maximal values over the 10 years of records so that the normalized value is comprised between 0 and 1. The resulting normalized variable exhibited very similar dynamics among all the piezometers (see Figure 2 a). However, the piezometer located in the riparian zone (F1b) exhibited variations at a higher frequency, especially during the winter. Therefore, in the following, we used the average of the normalized level in the two hillslope piezometers (F5b, F4) as a proxy for the hillslope groundwater storage dynamics, and the normalized level in the riparian piezometer as a proxy for the riparian groundwater storage dynamics.

In order to obtain a proxy for the unsaturated zone storage, moisture time series were also normalized using the minimal and maximal values observed in all the sensors of the two profiles over the two water years with complete records, setting the minimal value as 0 and the maximal value as 1. As the obtained normalized unsaturated storage variables were following very similar trends and dynamics, we used in the following an average of the normalized unsaturated zone storage among all the measurement points (depths and profiles) (Figure 2 b). The two profiles are located on the upslope and downslope parts of the hillslope respectively. Thus, we assumed that averaging their normalized values will allow us to build a proxy for the dynamics of the unsaturated zone storage on the whole hillslope.

2.4 Hysteresis Indices

Studies on hysteretic relationships in catchments generally focus on qualitative descriptions of patterns associated with a cross correlation analysis between the two variables (Frei et al., 2010; Hopmans and Bren, 2007; Jung et al., 2004; Salant et al., 2008; Schwientek et al., 2013; Spence et al., 2010; Velleux et al., 2008). Some authors proposed a typology of hysteretic loops based of their rotational direction, curvature and trend to identify solute controls during storm events (Butturini et al., 2008; Evans and Davies, 1998). For storage-discharge hysteresis at the annual scale, this approach is not sufficient as the same type of hysteretic loop is likely to happen for almost all the years, when a strong seasonality exists and its pattern is repeated across years. This is the case in your study, where seasonality of groundwater level and discharge was showing a strong unimodal pattern for all years, except 2011-2012 which was bimodal (Figure 2 and 3a). Moreover a preliminary cross correlation analysis revealed that storage and stream flow are strongly correlated, and cross correlation value is the greatest for lag time of 0 day (results not shown).

Quantitative descriptions of the hysteretic loop are also found in the literature, and various ways of computing hysteresis indices (HI) have been proposed, for example using the relative difference between extreme concentration values (Butturini et al., 2008) or using the ratio of turbidity values in rising and falling limbs of the storm hydrograph at the mid-point discharge value (Lawler et al., 2006). The latter authors argue that computing HI by using mid-point discharge usually allows avoiding the small convolutions which are frequently observed at both ends of the hysteretic loop.

In this paper, as the hydrological variables exhibit a strong annual uni-modal cycle, we calculated the hysteresis index (HI) each year as the difference between water storages at the dates of mid-point discharge in the two phases of the hydrological year, during the recharge period (R) and the recession period (r), i.e. respectively before and after reaching the maximal discharge $Q_{\max}$, as follow:

$$\begin{cases} HI = S(t_{R,mid}) - S(t_{r,mid}) \\ Q(t_{R,mid}) = Q_{mid} \text{ and } t_{R,mid} < t_{Q_{max}} \\ Q(t_{r,mid}) = Q_{mid} \text{ and } t_{r,mid} > t_{Q_{max}} \\ Q(t_{Q_{max}}) = Q_{max} \end{cases} \quad (1)$$

where *S(t)* is the storage value at time t and *Q(t)* the stream flow value at time t. The mid-point discharge Q_{mid}, is defined as the mean value of discharge between Q_0 the initial value at the beginning of the hydrological year (October), and $Q_{\max}$ the maximal value reached during that year:

$$Q_{mid} = \frac{Q_0 + Q_{\max}}{2} \quad (2)$$

In order to reduce the impact of the quick variations of discharge or groundwater level due to individual storm events, we smoothed the time series using 7-day moving averages. The strong seasonal discharge cycle led to identify two occurrences of Q_{mid} per year only: during the recharge period (t_R) and during the recession period (t_r), while high and low stream flow values are taken several times per year as explained by Lawler et al. (2006). Computing HI using the difference in storage was possible here because storage and stream flow values are varying among years within a narrow range of magnitude, while Lawler et al. (2006) used the ratio because turbidity can differ of several orders of magnitude from one storm to the other.

Computing HI with the difference between the values of storage and not their ratio allowed maintaining its sensitivity to the year to year variations of the width of the hysteretic loop. The difference in water storage dynamics in the unsaturated and saturated zones were approximated by the difference in normalized soil moisture content and by the difference in normalized groundwater level, respectively.

HI gives 2 types of information i) its sign indicates the direction of the loop (anticlockwise loop induces a negative value of HI whereas a clockwise loop leads to a positive value of HI) and ii) its absolute value is proportional to the magnitude of the hysteresis (i.e. the width of the hysteretic loop). HI is a proxy for the importance of lag time response between variations in catchment storages (unsaturated and saturated) and stream discharge, its sign indicates if storage reacts before or after the stream flow. Therefore, it can be used for comparing the capacity of the different models to reproduce to some extent the observed storages/discharge relationships. The normalization of the observed variables related to the storages (here either groundwater level or soil moisture) has no effect on the sign of HI, the HI values are being only divided by the maximal amplitude observed in the storage during the whole period. Therefore, as long as the normalization is applied for the whole period (for all years and for both measurements and simulations), it does not affect the interpretation related to absolute values of HI.

2.5 Models

In previous work, a range of conceptual models was calibrated and evaluated for the Kerrien catchment in a stepwise development using a flexible modelling framework (see Hrachowitz et al., 2014). This section aims at summarizing the results of this previous study as they are used as a basis for the present work. In this previous study, adopting a flexible stepwise modelling strategy, eleven models with increasing complexity, i.e. allowing for more process heterogeneity, were calibrated and evaluated for the study catchment. Four of these 11 models (hereafter referred to as M1 to M4; details given in Tables 1 and 2) were selected for the present work, as they correspond to the sequence of model architectures that provide the most significant performance improvements among the tested set-ups. As a starting point and benchmark, Model M1 with 7 parameters, resembling many frequently used catchment models, such as HBV, was used(e.g. Bergström, 1995). The three boxes represent respectively an unsaturated zone, a slow responding and a fast responding reservoir. In Model M2, additional deep infiltration losses are integrated from the slow store to take into account the significant groundwater export to adjacent catchments in this study catchment as indicated by the observed long-term water balance (Ruiz et al., 2002). This is done by adding a second outlet together with threshold into this storage to allow for continued groundwater export from a storage volume below the stream during 0-flow conditions, i.e. when the stream runs dry. As riparian zones frequently exhibit a distinct hydrological functioning, (e.g. Molénat et al., 2005; Seibert et al., 2003), indicated in the study catchment by distinct response dynamics in the riparian piezometers (Martin et al., 2006), models M3 and M4 additionally integrate a wetland/riparian zone component, composed of an unsaturated zone store and a fast responding reservoir, parallel to the other boxes. The riparian unsaturated zone generates flow using a linear function in M3 and a non-linear function in M4. The complete set of water balance and constitutive model equations of the four models is listed in Table 1 while the model structures are schematized in Table 2.

2.6 Calibration and Evaluation

This section is also a summary of the findings of Hrachowitz et al. (2014) that served as a basis for this study, and not results of the current study. The models have been calibrated for the period 01/10/2002 – 30/09/2007 after a one-year warm-up period, using a multi-objective calibration strategy (e.g. Gupta et al., 1998), based on Monte-Carlo sampling (10^7 realizations). The uniform prior parameter distributions used for M1 – M4 are provided in Table 3. To reduce parameter and associated predictive uncertainty, the models were calibrated using a total of four calibration objective functions (see Table 4), i.e. the Nash-Sutcliffe efficiencies (Nash and Sutcliffe, 1970) respectively for stream flow ($E_{NS,Q}$), for the logarithm of the stream flow ($E_{NS,log(Q)}$), and for the flow duration curve ($E_{NS,FDC}$) as well as the volumetric efficiency for stream flow ($V_{E,Q}$; Criss and Winston, 2008). To facilitate clearer assessment, the calibration objective functions (n = 4) were combined in a single calibration metric: the Euclidean Distance to the perfect model ($D_{E,cal}$, e.g. Hrachowitz et al., 2013a; Gascuel-Odoux et al., 2010):

$$D_E = \sqrt{\frac{\left(1-E_{NS,Q}\right)^2+\left(1-E_{NS,\log(Q)}\right)^2+\left(1-E_{V,Q}\right)^2+\left(1-E_{NS,FDC}\right)^2}{n}} \qquad (5)$$

As mathematically optimal parameter sets are frequently hydrologically sub-optimal, i.e. unrealistic (e.g. Beven, 2006), all parameter sets within the 4-dimensional space spanned by the calibration Pareto fronts, as approximated by the cloud of sample points, were retained as feasible.

The calibrated models were then evaluated against their respective skills to predict the system response with respect a selection of 13 catchment signatures (described in Table 4) in a multi-criteria posterior evaluation strategy. Figure 3 and Table 4 show the global performance D_E of the 4 models in terms of the Euclidean Distance to the perfect model constructed from all calibration objective functions and evaluation signatures. Model M1 provided good performance in calibration on the objective functions while its validation performances were considerably decreased. Its ability to reproduce the different signatures showed that it failed in particular to reproduce flow in wet periods (such as the evaluation period in Fig. 3a) and groundwater dynamics. Model M2 led to calibration performances slightly lower than model M1 but higher validation performances. The hydrological signatures simulated by M2 exhibited lower uncertainties both in validation and calibration periods because of better simulation of low flow conditions and groundwater dynamics. Model M3 provided similar performances as M2 for calibration and for validation but with clearly reduced uncertainty bounds. Overall signatures reproduction was improved because of clear improvement of low flow and groundwater related signatures even if performance in calibration objective functions remained lower than that for model M1. Model M4 exhibited similar performances than the previous models both in calibration and validation periods but better performance for the whole set of signatures and lower uncertainties.

More details on the model calibration and evaluation with respect to hydrological signatures can be found in Hrachowitz et al., (2014; note that M1, M2 , M3 and M4 presented in this study correspond respectively to M1, M6, M8 and M11 in the original paper). Within the obtained range of parameter uncertainty, the types of simulated hysteresis patterns were not affected by the parameter values but only by the model structures. Note that we restricted the following analysis only to the optimal parameter set in each case, first for the sake of clarity and also because at this stage, our interest was on assessing the ability of model structures to reproduce the observed general features in hysteresis patterns and not on quantifying their performances in fitting the observations.

In the present work the sensitivity of the Hysteresis Indices to parameter uncertainty is investigated by computing the HI values for the all sets of feasible parameters.

3 Results and discussion

3.1 Hysteretic pattern on the groundwater storage/discharge relationship

3.1.1 Observations in hillslope and riparian zones: saturated storage vs. flow

The 2-dimensional observed relationship between saturated storage in the hillslope (HSS) or in the riparian zone (RSS) and stream discharge (Q) for each year was hysteretic, highlighting the non-uniqueness of the response of discharge to storage depending on the initial conditions and a lag time between both variables dynamics in particular during the recharge period as illustrated in Figure 4 for two contrasted water years.

The direction of the hysteretic loop was different depending on the topographic position of the piezometer: loops were always anticlockwise (leading to negative values of HI) for the piezometer located at the top of the hillslope HSS-F5b(Q), mostly anticlockwise for the mid-slope piezometer HSS-F4(Q) and mostly clockwise (positive values of HI) for the piezometer in the riparian zone RSS-F1b(Q) (Figure 5).

In the riparian zone, storage at Q_{mid} was usually lower in the recession period than in the recharge period, especially in dry years, leading to positive HI. This is due to the fact that riparian groundwater level increased early at the beginning of the recharge period, before the stream discharge, due to the limited storage capacity of the narrow unsaturated layer in bottom lands reinforced by groundwater ridging which is linked to the extent of the capillary fringe. However, the hysteretic loops were narrow and for wet years, the storage value during the recession period occasionally exceeded the value in the recession period without modifying the general direction of the hysteresis when looking at the whole pattern (e.g. in 2003-2004, see Figure 4 a). When this occurred at the time of Q_{mid}, it led to negative HI although absolute values remained small (Figure 5)

The hillslope groundwater responded later than the stream, due to the deeper groundwater levels and higher unsaturated storage capacity (Rouxel et al. 2011), both introducing a time lag for the recharge and thus for the groundwater response. This led to negative values of HI as groundwater levels in recession periods were higher than in recharge periods for the same level of discharge (in particular at Q_{mid}). The loops were also wider in the hillslope, leading to high absolute values of HI (Figure 5).

The intermediate behavior of the mid-slope piezometer (F4), exhibiting varying patterns throughout the years, reflects the fact the riparian zone extends spatially towards the hillslope and reaches a larger spatial extension during wet years.

Similar observations have been reported by other authors. For example, anticlockwise hysteresis between groundwater tables and discharge are observed by Gabrielli et al. (2012) in the Maimai catchment, while studies on riparian groundwater or river bank groundwater report clockwise hysteresis at the storm event scale (Frei et al., 2010; Jung et al., 2004). Similar pattern were also observed by Jung et al. (2004) who found that in the inner floodplain and in river bank piezometers, the hysteresis curve between water table and river stage exhibit a synchronous response, while in the hillslope hysteresis curves are relatively open, as the water table is higher during the recession than during the rising limb.

3.1.2 Observations in hillslope: saturated and unsaturated storages vs. flow

Figure 6 shows the 3-dimensional relationship between hillslope saturated storage (HSS), unsaturated storage (HUS), and stream flow (Q) for the year 2010-2011. Four main periods can be identified, similar to what was outlined in recent studies (e.g. Heidbuechel et al., 2012; Hrachowitz et al., 2013a): three are characterizing the recharge period and the last one the recession period. First, stream flow was close or equal to zero and was almost exclusively sustained by drainage of the saturated storage, while the unsaturated zone exhibited a significant storage deficit and only minor fluctuations due to transpiration and small summer rain events (dry period). As more steady precipitation patterns set in, here typically around November, the unsaturated zone storage relatively quickly reached its maximal value, rapidly establishing connectivity of fast responding flow pathways (wetting up period). This led to a relatively rapid increase in stream flow while the saturated storage did not change much until the end of this period as incoming precipitation first had to fill the storage deficit in the unsaturated zone before significant increase in percolation could occur. A further lag was introduced by the time taken for water to percolate and eventually recharge the relatively deep groundwater. As soon as conditions were wet enough to allow for established percolation, the saturated storage eventually also responded, increasing faster than the stream flow (wet period) while unsaturated storage remained full. During the wet period (or high flow period), no pattern appeared clearly because all storage elements were almost full and the response of all the compartments were more directly linked to the short term dynamics of rain events. Finally during the recession period (drying period), unsaturated storage decreased comparatively quickly by drainage and transpiration while the saturated storage may keep increasing for a while by continued percolation from the unsaturated zone before decreasing through groundwater drainage at a relatively slow rate. A similar pattern was also observed for 2011-2012 (not shown).

The unsaturated zone storage followed a clockwise hysteresis loop with the stream flow and with the saturated zone storage. The hysteresis indices (Figure 5, years 2010-2011 and 2011-2012) reflected these directions, and showed that the hysteresis loops were narrower for unsaturated storage than for saturated storage, inducing smaller absolute values of the hysteresis indices due to the small size of the unsaturated storage compartment compared to the saturated storage compartment.

3.1.3 Interpretation

There are 3 main hypotheses generally proposed to interpret storage-discharge hystereses in hydrology. The first one is related to the increase of transmissivity with the groundwater level due to the frequently observed exponential decrease of hydraulic conductivity with depth. However, this would lead to systematic clockwise hysteresis loops and cannot explain the anticlockwise patterns observed between hillslope saturated storage and stream flow. The second hypothesis proposed by (Spence et al., 2010) is that during the recharge period, the groundwater storage is not only increasing locally (as measured by the piezometric variations) but also the spatial extension of connected storage increases gradually, while during the recession period, the storage is decreasing homogenously across the entire contribution area. This is likely for riparian groundwater and could explain the clockwise hysteresis observed on this piezometer but cannot explain the anticlockwise hysteresis observed in the hillslope groundwater. The third hypothesis is that dominant hydrological processes are different between recharge and recession periods. For instance, (Jung et al., 2004) interpret their clockwise hysteresis in peatlands groundwater as the results of a stepwise filling process during the rising flows (fill and spill mechanism) opposed to a more gradual drainage of the groundwater during the recession combined with the first hypothesis result, similar to what was found by Hrachowitz et al. (2013a). This hypothesis of different hydrological pathways allows an adequate interpretation of the opposite directions of the observed hystereses. Recharge period is characterized by a quick filling of the unsaturated and saturated storages in the riparian zone which is always close to the saturation while the saturated storage on the hillslope is not yet filling up (wetting up period). Thus, the wetting up period is characterized by an increase of stream flow, here mainly generated in the riparian zone, and eventual quick flows in the hillslope while hillslope unsaturated zone is reaching the storage capacity volume. At the beginning of the wet period, hillslope saturated storage is filling and starts to contribute to the stream along with riparian and fast flows. During the recession period (drying period), hillslope saturated zone is the only compartment which continues to sustain stream flow. If so there are three contributions to stream flow in the wet period while during the recession period, hillslope groundwater remains the only contributor to stream flow (cf. Hrachowitz et al., 2013a, see Figure 7). This can explain the difference between storages values between recharge and recession periods. Finally, the hysteretic hydrological signature is not only related to the amount of stored water in the catchment but rather to where it is stored.

These results are consistent with previous studies: the distinction between riparian groundwater and hillslope groundwater components has also been identified in similar catchments by Molenat et al. (2008) based on nitrate concentrations analysis and by Aubert et al. (2013) based on a range of solutes, and in other site as by Haught and van Meerveld (2011) using such Q-S relationships and lag time analysis.

3.1.4 Sensitivity of HI to initial conditions

Sensitivity to antecedent soil moisture conditions are often cited as an explanation for observed storage-discharge hysteresis and its variability between years. Initial levels of each store will obviously influence the time required to fill them and consequently the duration of the successive periods identified in the whole recharge period. As only 2 years of data were available, it was not possible to define a relationship between the initial average soil moisture and the magnitude of the hysteresis indices. However, the magnitude of HI was lower for high initial values of average unsaturated zone storage for both the saturated and unsaturated zones in 2011-2012 (Table 5). The HI for midslope saturated zone (F4b) seemed to be more sensitive too these initial moisture conditions than HI for upslope saturated zone and unsaturated zone. Similarly, the width of the loop (absolute value of HI) was little sensitive to initial groundwater levels in the hillslope: although the larger absolute values of HI were observed for the lower initial water table levels, no clear correlation was observed (Figure 8).

3.1.5 Sensitivity of HI to annual rainfall

For the saturated zone, observed values of HI were negatively correlated with the total annual rainfall for both the hillslope and the riparian zone, with a more negative slope for the hillslope (Figure 9). Wet years (i.e. large values of annual rainfall) are generally associated with large values of annual maximal and mid-point stream flows and also to large values of groundwater table level, leading to larger saturated storage values during the recession period, while the storage values during the recharge period do not change much from year to year. Thus, larger storage values at the time of mid-point discharge in the recession period led to smaller values of HI (i.e larger absolute values for the hillslope where hytereses are anticlockwise and smaller absolute value of HI for the riparian zone where hystereses are clockwise). In the riparian zone, when rainfall and maximal drainage reached very high value, it could lead to saturated storage value at the time of mid-point discharge in the recession period larger than the corresponding value during the recharge period, explaining the inversion of the sign of HI for RSS(Q) in very wet years.

3.2. Models assessment based on their ability to reproduce the observed hysteresis

3.2.1 Hysteresis simulations

For all years, all models (M1-M4) exhibited a hysteretic relationship between stream flow and the storages, as shown in Figure 10 for the years 2003-2004 and 2007-2008, pertaining to the calibration and validation periods respectively. This means that all tested models introduced a lag time between catchment stores and the stream dynamics. The Figure 11a presents the observed and modelled average and standard deviation of annual hysteresis Indices, for Hillslope saturated storage vs. discharge HSS(Q), Hillslope unsaturated storage vs. discharge HUS(Q), Hillslope unsaturated storage vs. Hillslope saturated storage HUS(HSS), and Riparian saturated storage vs. discharge RSS(Q). As Riparian saturated storage (RSS) is not modelled in M1 and M2, simulated RSS(Q) was available only for M3 and M4.

For M1, the shape of the simulated hysteresis showed an overestimation of hillslope saturated storage (HSS) and of flow during dry years (e.g. the year 2007-2008 shown in Figure 10). This was expected as we have seen that the model was unable to reproduce groundwater dynamics and the low signatures during the validation period (Figure 3 and supplementary material). Simulated HI values were close to the observed ones for HSS(Q) (Figure 11a). The simulated hysteresis indices were small and negative for HUS(Q) while the observed values were large and positive. Simulated HI values for HUS(HSS) were also overestimated. These results show that in model M1, the overestimation of the hillslope saturated storage was partially compensated by the underestimation of the hillslope unsaturated storage. This reveals the poor consistency of the model and explains why it was able to reach good performance in the calibration period but not in the validation period (Figure 3).

For the model M2, the shape of the hysteresis loops showed a considerable underestimation of HSS and a large underestimation of stream flow in wet years (Figure 10). Compared to M1, although the introduction of deep losses in M2 led to higher validation performances and better simulation of hydrological signatures (Figure 3), the simulated HI (Figure 11a) were worsened, suggesting a poorer model consistency with respect to internal hydrologic processes.

For both models M3 and M4, the introduction of a riparian compartment improved the simulated hysteretic loops, due to a better simulation of stream flow in wet years, but HSS was still largely underestimated (Figure 10) . The mean HI values for HSS(Q) were close to the observed one, but the range of variation were smaller, indicating a reduced sensitivity to climate (Figure 9). The mean values for HUS(Q) were clearly improved compared to M1 and M2, as the direction of the loop were clockwise as for the observed ones, although still underestimated. The mean HI values for HUS(HSS) were also greatly improved. The shape of the simulated hysteresis loops between riparian saturated storage (RSS) and stream flow (Q) showed a large underestimation of RSS especially during the recession period (Figure 10 c, d). This led to simulated HI for RSS(Q) which are positive, like the observed ones, but also largely overestimated (Figure 11a). Overall, these results suggest that for models including a riparian component, the underestimation of the hysteresis between HUS and Q was compensated by an overestimation of the hysteresis between RSS and Q. This highlights that despite a significant improvement in performances and improved hydrological signature reproduction, these models still involve a certain degree of inconsistency with respect to internal processes. However, M4 provided the most balanced performance considering hysteretic signatures between all storage components and strongly underlines the limitations of overly simplistic model architectures (e.g. M1) and the need for more complete representations of process heterogeneity. The hysteresis index sensitivity to parameter uncertainty increases with the number of parameters from M1 to M2 and then stay in the same range from M2 to M4 (Figure 11b). This analyse confirms the importance of considering the Hysteresis Indices both between saturated and unsaturated storage (HSS and

HUS) to avoid accepting a wrong model. For example, considering only the performance on the HSS(Q) relationship could lead to accept model M1 while its performance on HUS is lower and it is not able to reproduce the Riparian compartment hysteresis. For readability purposes, Figure 11b illustrates this sensitivity for the different HI in the year of 2011-2012 only but similar behaviour is observed every year. It showed that best behavioural parameters sets (bbp) lead to modelled HI values closer to the observed values than average modelled HI values. Using an additional calibration criterion related to the hysteresis could reduce the sensitivity of HI to parameter uncertainty and lead to narrow range of feasible parameter sets.

3.2.2 Sensitivity of modelled hysteresis indices to annual rainfall

All models were able to represent the decrease of the hysteresis indices with annual rainfall in the hillslope, the slope of the correlation getting closer to the observed one from M1 to M4 (Figure 9). The introduction of deep groundwater losses (M2) led to smaller saturated storage during recharge periods and increased the difference between saturated storage during recharge and recession periods at the time of mid-point discharge. However, as all models tended to overestimates low stream flow values, the slopes of the correlations between annual rainfall and simulated HI were smaller than for the observed one.

In the riparian zone, the modelled trends were the inverse of the observed one. The modelled recessions were always very sharp (see Hrachowitz et al., 2014), and the simulated riparian storage dried up every year, explaining that saturated storage at the time of mid-point discharge during the recession periods were much greater than during the recharge periods. This led to a general overestimation of HI values which were even stronger for wet years. This overestimation may be related to an improper conceptualization of the riparian zone functioning, which is never connected to hillslope reservoir in the tested models. In reality, during high flow periods, observed hydraulic gradient increased along the hillslope inducing a connection between riparian and hillslope reservoirs which are disconnected during low flow periods.

3.2.3 Value of such internal signatures for model evaluation

The use of hydrological hysteretic signatures in model assessments led to conclusions that were consistent with the classical hydrological signatures used in Hrachowitz et al. (2014). However, model M2 was less able to reproduce the different hysteretic signatures whereas it led to a real improvement regarding to the classical signatures in low flows. Considering only the distance between observed and simulated hysteresis indices on hillslope saturated storage and stream flow would lead to select model M1. This highlights the fact that using saturated storage dynamics alone can be deceptive for understanding the system response behaviour and that it is thus crucial to also consider the hysteretic signatures of unsaturated and riparian zones in a combined approach to develop a more robust understanding of the system. Here, hysteretic signatures of the unsaturated and riparian zones provided valuable additional assessment metrics regarding the performance of models M3 and M4 to represent the riparian zone. It was possible to identify when the model failed to represent processes and which processes are mostly compensating for missing ones and therefore why the model may provide some good performance for wrong reasons. To do so, the hysteresis index proved to be a useful proxy of hystereses themselves as it exhibited contrasted patterns sensitive to climate and localization within the catchment.

3.3 Perspectives: toward integrated hydrological-signatures-based modelling?

A general issue in model calibration is that because of over-parameterization of hydrological models and because the objective functions integrate generally only one variable, like the stream flow, automatic calibration techniques may lead to parameter sets which compensate for internal model errors. These parameter sets are mathematically correct but wrong in a hydrological point of view. The subsequent model is then to be considered non behavioural (Beven, 2006). For instance, if storage properties are not well taken into account by the model, this is likely to lead to a wrong simulation of storage dynamics in response to precipitation and thus the parameterization using traditional objective functions can lead to compensation of these errors in order to simulate a discharge value close to observed one while the storage is wrong. In such case, a model able to represent the internal catchment behaviour will generate a wrong discharge value but consistent with the storage value and will be rejected in traditional calibration procedures. To handle this issue and in order to select behavioural models, one can use multiple objective functions (Gupta et al., 1998; Seibert and McDonnell, 2002) Freer et al., 2003), including a range of hydrological signatures to be reproduced or additional realism constraints (Kavetski and Fenicia, 2011; Yadav et al., 2007; Yilmaz et al., 2008; Euser et al., 2013; Gharari et al., 2013; Hrachowitz et al., 2014). We argue that rather than increasing the number of constraints or objective functions to satisfy, an alternative could be to use some objective functions based on a combination of different variables as stream flow and the groundwater level, soil moisture, or stream concentrations. Among the possible combination of variables, objective functions based on the relative dynamics of storage in different spatial locations, such as riparian versus hillslope, might provide new insights about the catchment internal processes. We suggest that such combined objective function would be more constraining for model selection. Therefore, the present study is a first step which aims at highlighting the still underexploited potential of hydrological hysteresis. The next step would be to quantify these relationships through functions or several indices usable in calibration criteria such as the Hysteresis Index proposed in this study. Moreover, such criteria could be used in classification studies. Indeed, some studies on the literature present storage/discharge relationships for different catchments that show patterns that are similar or not to the ones we observed in the Kerrien catchment (Ali et al., 2011; Gabrielli et al., 2012). This signature may help to classify catchments in terms of dominant processes driving their behaviour.

A remaining difficulty to integrate storage in calibration or evaluation procedure in hydrological modelling is how to measure this storage. McNamara et al. (2011 and Tetzlaff et al. (2011) proposed to use all available data from groundwater level monitoring, soil moisture records, water budget, modelling results, and so on, to estimate the storage in catchments. In this study, we used a quite dense network of piezometers and soil moisture measurements relatively to the small size of the catchment. Promising ways to estimate spatial quantification of storage in catchments include remote sensing of soil moisture (Sreelash et al., 2013; Vereecken et al., 2008), gravimetric techniques (Creutzfeldt et al., 2012), geodesy and geophysical methods. The interest of such techniques would be to provide a spatially integrated vision of the catchment water content.

As for the different hydrological variables, the combination of hydrological and chemical variables appears relevant to investigate the hydrochemical behaviour of catchment. Hysteresis patterns between concentration and discharge have been largely documented for storm events characterization (Evans and Davies, 1998; Evans et al., 1999; Taghavi et al., 2011). Some studies are also reporting similar patterns at the annual scale (*e.g.* Aubert et al., 2013 b). Such hysteretic relationships have been observed also between water and chemistry in groundwater (Rouxel et al., 2011; Hrachowitz et al., 2013a) emphasizing a disconnection between water and solutes dynamics that simple diffusion or partial mixing processes cannot explain. Stream water chemistry exhibits also particular seasonal cycles with different phasing with discharge depending on the solutes (Aubert et al., 2013). This provides extra information on the water pathways within the catchment. These relationships appears also powerful to constraint hydrochemical modelling.

4 Conclusion

A method to characterize and quantify partially the relationship between storages in a headwater catchment and stream flow along the year has been proposed. It allowed us to then assess the ability of a range of conceptual lumped models to reproduce this catchment internal signature. Catchment storage has been approximated using a network of piezometric data and several unsaturated zone moisture profiles to consider the storage in the saturated as well as in the unsaturated zones.

The observations showed that storage/discharge relationships in catchments can be hysteretic highlighting a successive activation of different hydrological components during the recharge period while the recession exhibits a fast decrease of unsaturated and riparian storages and a slow decrease of hillslope saturated storage which sustains the stream flow. Four periods have been identified along the hydrological year: 1) first, at the end of the dry period, rainfall starts to refill unsaturated storages; 2) in the wetting period, riparian unsaturated storage is filled and the saturated storage starts to supply the stream while hillslope unsaturated storage is still being replenished; 3) during the wet period, unsaturated storage in the hillslope is also filled and the saturated hillslope storage also feeds the stream. Finally when rainfall declines, flow from the riparian groundwater recedes and during the recession period, the stream discharge is sustained only by hillslope groundwater. Stream discharge and riparian and hillslope saturated storages exhibited different patterns of hysteresis, with opposite directions of the hysteretic loops.

The tested models were characterized by an increasing degree of complexity and also an increasing consistency, as shown in a previous study using classical hydrologic signatures. In this study, we showed that if all of them simulated a hysteretic relationship between storages and discharge, their ability to reproduce hysteresis index also increased with model complexity. In addition, we suggest that if classical hydrological signatures help to assess model consistency, the hysteretic signatures help also to identify quickly when and why the models give “right answer for wrong reasons” and can be used as a descriptor of the internal catchment functioning.

Acknowledgments

The investigations benefited from the support of INRA and CNRS for the Research Observatory ORE AgrHys, and from Allenvi for the SOERE RBV. Data are available on http://geowww.agrocampus-ouest.fr/web/.

References

Ali, G.A., L'Heureux, C., Roy, A.G., Turmel, M.-C., Courchesne, F., 2011. Linking spatial patterns of perched groundwater storage and stormflow generation processes in a headwater forested catchment. Hydrological Processes, 25(25): 3843-3857.

Andermann, C. et al., 2012. Impact of transient groundwater storage on the discharge of Himalayan rivers. Nature Geosci, 5(2): 127-132.

Andréassian, V. et al., 2012. All that glitters is not gold: the case of calibrating hydrological models. Hydrological Processes, 26(14): 2206-2210.

Aubert, A.H. et al., 2013. Solute transport dynamics in small, shallow groundwater-dominated agricultural catchments: insights from a high-frequency, multisolute 10 yr-long monitoring study. Hydrology and Earth System Sciences, 17(4): 1379-1391.

Aubert, A.H., Gascuel-Odoux, C., Merot, P., 2013. Annual hysteresis of water quality: A method to analyse the effect of intra- and inter-annual climatic conditions. Journal of Hydrology, 478: 29-39.

Bergström, S. (1995), The HBV model. In: Computer Models of Watershed Hydrology, Singh, V.P. (ed.), Water Resources Publications: Highlands Ranch, CO, 443-476.

Beven, K., 2001. How far can we go in distributed hydrological modelling? Hydrology and Earth System Sciences, 5(1): 1-12.

Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology, 320(1–2): 18-36.

Beven, K., 2013. So how much of your error is epistemic? Lessons from Japan and Italy. Hydrological Processes, 27(11): 1677-1680.

Butturini, A., Alvarez, M., Bernal, S., Vazquez, E., Sabater, F., 2008. Diversity and temporal sequences of forms of DOC and NO3-discharge responses in an intermittent stream: Predictable or random succession? Journal of Geophysical Research-Biogeosciences, 113(G3).

Creutzfeldt, B. et al., 2012. Total water storage dynamics in response to climate variability and extremes: Inference from long-term terrestrial gravity measurement. Journal of Geophysical Research: Atmospheres, 117(D8): D08112.

Criss, R.E., Winston, W.E., 2008. Do Nash values have value? Discussion and alternate proposals. Hydrological Processes, 22(14): 2723-2725.

Euser, T. et al., 2013. A framework to assess the realism of model structures using hydrological signatures. Hydrology and Earth System Sciences, 17(5): 1893-1912.

Evans, C., Davies, T.D., 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. Water Resources Research, 34(1): 129-137.

Evans, C., Davies, T.D., Murdoch, P.S., 1999. Component flow processes at four streams in the Catskill Mountains, New York, analysed using episodic concentration/discharge relationships. Hydrological Processes, 13(4): 563-575.

Everett, D.H., Whitton, W.I., 1952. A general approach of hysteresis. Trans. Faraday Soc, pp. 749-757.

Fenicia, F., McDonnell, J.J., Savenije, H.H.G., 2008. Learning from model improvement: On the contribution of complementary data to process understanding. Water Resources Research, 44(6).

Franks, S.W., Gineste, P., Beven, K.J., Merot, P., 1998. On constraining the predictions of a distributed moder: The incorporation of fuzzy estimates of saturated areas into the calibration process. Water Resources Research, 34(4): 787-797.

Freer, J., Beven, K. and Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, Water Resour. Res., 32(7): 2161-2173.

Freer, J., Beven, K. and Peters, N., 2003. Multivariate seasonal period model rejection within the generalised likelihood uncertainty estimation procedure, Calibration of watershed models, 69-87.

Freer, J.E., McMillan, H., McDonnell, J.J., Beven, K.J., 2004. Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures. Journal of Hydrology, 291(3-4): 254-277.

Frei, S., Lischeid, G., Fleckenstein, J.H., 2010. Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland - A modeling study. Advances in Water Resources, 33(11): 1388-1401.

Gabrielli, C.P., McDonnell, J.J., Jarvis, W.T., 2012. The role of bedrock groundwater in rainfall–runoff response at hillslope and catchment scales. Journal of Hydrology, 450–451(0): 117-133.

Gascuel-Odoux, C., Weiler, M., Molenat, J., 2010. Effect of the spatial distribution of physical aquifer properties on modelled water table depth and stream discharge in a headwater catchment. Hydrology and Earth System Sciences, 14(7): 1179-1194.

Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H. and Savenije, H.H.G. (2013), Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration, Hydrol. Earth Syst. Sci. Disc., 10, 14801-14855.

Giustolisi, O., Simeone, V., 2006. Optimal design of artificial neural networks by a multiobjective strategy: groundwater level predictions. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 51(3): 502-523.

Gupta, H.V., Clark, M.P., Vrugt, J.A., Abramowitz, G., Ye, M., 2012. Towards a comprehensive assessment of model structural adequacy. Water Resources Research, 48.

Gupta, H.V., Wagener, T., Liu, Y.Q., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrological Processes, 22(18): 3802-3813.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4): 751-763.

Hartmann, A., Wagener, T. , Rimmer, A. , Lange, J., Brielmann, H., and Weiler, M. 2013 a. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. Water Resources Research, 49, 3345–3358, doi:10.1002/wrcr.20229.

Hartmann, A., Weiler, M., Wagener, T., Lange, J., Kralik, M., Humer, F., Mizyed, N., Rimmer, A., Barberá, J. A., Andreo, B., Butscher, C., and Huggenberger, P. 2013b. Process-based karst modelling to relate hydrodynamic and hydrochemical characteristics to system properties, Hydrol. Earth Syst. Sci., 17, 3305-3321, doi:10.5194/hess-17-3305-2013.

Haught, D.R.W., van Meerveld, H.J., 2011. Spatial variation in transient water table responses: differences between an upper and lower hillslope zone. Hydrological Processes, 25(25): 3866-3877.

Hopmans, P., Bren, L.J., 2007. Long-term changes in water quality and solute exports in headwater streams of intensively managed radiata pine and natural eucalypt forest catchments in south-eastern Australia. Forest Ecology and Management, 253(1-3): 244-261.

Hrachowitz M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Freer, J., Savenije, H.H.G., and Gascuel-Odoux, C. 2014. Process consistency in models: The Importance of System Signatures, Expert Knowledge and Process Complexity. accepted in Water Resources Research in September 2014. doi: 10.1002/2014WR015484.

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

Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E. and Cudennec, C., 2013 b. A decade of predictions in ungauged basins (PUB)—a review, Hydrol. Sci. J., 58(6): 1198-1255.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., Malcolm, I. A., 2011. Sensitivity of mean transit time estimates to model conditioning and data availabilityle. Hydrological Processes, 25 (6), 980-990.

Jung, M., Burt, T.P., Bates, P.D., 2004. Toward a conceptual model of floodplain water table response. Water Resources Research, 40(12).

Kampf, S.K., Burges, S.J., 2007. Parameter estimation for a physics-based distributed hydrologic model using measured outflow fluxes and internal moisture states. Water Resources Research, 43(12).

Kavetski, D., Fenicia, F., 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. Water Resources Research, 47: 19 p.

Kirchner, J.W., 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research, 42(3): W03S04.

Lamb, R., Beven, K., Myrabo, S., 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. Advances in Water Resources, 22(4): 305-317.

Lawler, D.M., Petts, G.E., Foster, I.D.L., Harper, S., 2006. Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK. Science of the Total Environment, 360(1–3): 109-126.

Martin, C. et al., 2006. Modelling the effect of physical and chemical characteristics of shallow aquifers on water and nitrate transport in small agricultural catchments. Journal of Hydrology, 326(1-4): 25-42.

McNamara, J.P. et al., 2011. Storage as a Metric of Catchment Comparison. Hydrological Processes, 25(21): 3364-3371.

Mohamed, Y.A., Savenije, H.H.G., Bastiaanssen, W.G.M., van den Hurk, B., 2006. New lessons on the Sudd hydrology learned from remote sensing and climate modeling. Hydrology and Earth System Sciences, 10(4): 507-518.

Molenat, J., Gascuel-Odoux, C., Ruiz, L., Gruau, G., 2008. Role of water table dynamics on stream nitrate export and concentration. in agricultural headwater catchment (France). Journal of Hydrology, 348(3-4): 363-378.

Molénat, J., Gascuel-Odoux, C., Davy, P., Durand, P., 2005. How to model shallow water-table depth variations: the case of the Kervidy-Naizin catchment, France. Hydrological Processes, 19(4): 901-920.

Moussa, R., Chahinian, N., Bocquillon, C., 2007. Distributed hydrological modelling of a Mediterranean mountainous catchment - Model construction and multi-site validation. Journal of Hydrology, 337(1-2): 35-51.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I -A discussion of principles. Journal of Hydrology, 10(3): 282-290.

Nester, T., Kirnbauer, R., Parajka, J., Bloschl, G., 2012. Evaluating the snow component of a flood forecasting model. Hydrology Research, 43(6): 762-779.

Parajka, J. et al., 2006. Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale. Hydrology and Earth System Sciences, 10(3): 353-368.

Rouxel, M. et al., 2011. Seasonal and spatial variation in groundwater quality along the hillslope of an agricultural research catchment (Western France). Hydrological Processes, 25(6): 831-841.

Ruiz, L. et al., 2002. Effect on nitrate concentration in stream water of agricultural practices in small catchments in Brittany : I. Annual nitrogen budgets. Hydrology and Earth System Sciences, 6(3): 497-505.

Salant, N.L., Hassan, M.A., Alonso, C.V., 2008. Suspended sediment dynamics at high and low storm flows in two small watersheds. Hydrological Processes, 22(11): 1573-1587.

Savenije, H.H.G., 2001. Equifinality, a blessing in disguise? Hydrological Processes, 15(14): 2835-2838Schwientek, M., Osenbruck, K., Fleischer, M., 2013. Investigating hydrological drivers of nitrate export dynamics in two agricultural catchments in Germany using high-frequency data series. Environmental Earth Sciences, 69(2): 381-393.

Seibert, J., 2000. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrology and Earth System Sciences, 4(2): 215-224.

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 Resources Research, 38(11).

Seibert, J., Rodhe, A., Bishop, K., 2003. Simulating interactions between saturated and unsaturated storage in a conceptual runoff model. Hydrological Processes, 17(2): 379-390.

Sivapalan, M., Bloschl, G., Merz, R., Gutknecht, D., 2005. Linking flood frequency to long-term water balance: Incorporating effects of seasonality. Water Resources Research, 41(6).

Spence, C., 2010. A Paradigm Shift in Hydrology: Storage Thresholds Across Scales Influence Catchment Runoff Generation. Geography Compass, 4(7): 819-833.

Spence, C. et al., 2010. Storage dynamics and streamflow in a catchment with a variable contributing area. Hydrological Processes, 24(16): 2209-2221.

Sreelash, K., Sekhar, M., Ruiz, L., Buis, S., Bandyopadhyay, S., 2013. Improved Modeling of Groundwater Recharge in Agricultural Watersheds Using a Combination of Crop Model and Remote Sensing. Journal of the Indian Institute of Science, 93(2): 189-207.

Taghavi, L., Merlina, G., Probst, J.L., 2011. The role of storm flows in concentration of pesticides associated with particulate and dissolved fractions as a threat to aquatic ecosystems Case study: the agricultural watershed of Save river (Southwest of France). Knowledge and Management of Aquatic Ecosystems(400).

Tetzlaff, D., McNamara, J.P., Carey, S.K., 2011. Measurements and modelling of storage dynamics across scales Preface. Hydrological Processes, 25(25): 3831-3835.

Velleux, M.L., England, J.F., Julien, P.Y., 2008. TREX: Spatially distributed model to assess watershed contaminant transport and fate. Science of the Total Environment, 404(1): 113-128.

Vereecken, H. et al., 2008. On the value of soil moisture measurements in vadose zone hydrology: A review. Water Resources Research, 44.

Wagener, T., 2007. Can we model the hydrological impacts of environmental change? Hydrological Processes, 21(23): 3233-3236.

Wagener, T., Montanari, A., 2011. Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. Water Resources Research, 47.

Winsemius, H.C., Savenije, H.H.G., Bastiaanssen, W.G.M., 2008. Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins? Hydrology and Earth System Sciences, 12(6): 1403-1413.

Yadav, M., Wagener, T., Gupta, H., 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. Advances in Water Resources, 30(8): 1756-1774.

Yilmaz, K.K., Gupta, H.V., Wagener, T., 2008. A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. Water Resources Research, 44(9).

Figure 1. Study site in West Brittany (square near Quimper) and location of the monitoring equipments. The weather station is located 500 m north of the catchment.

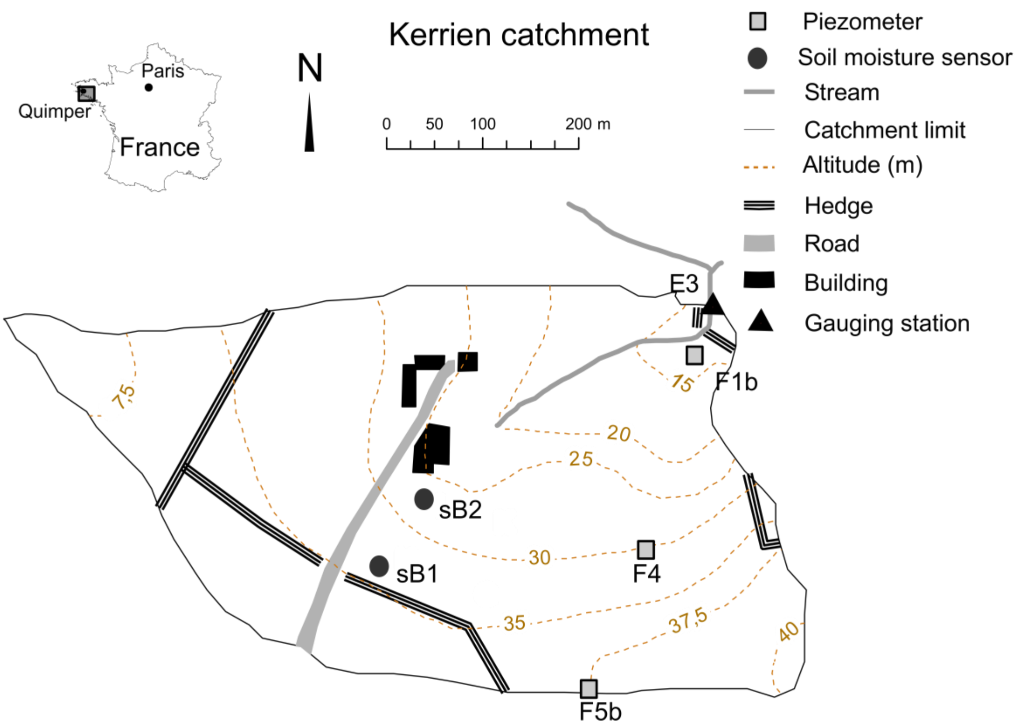
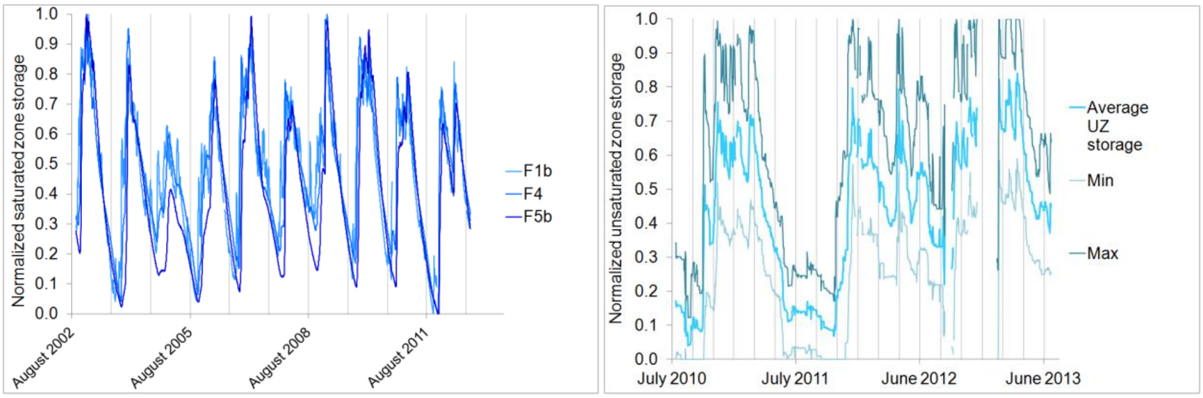
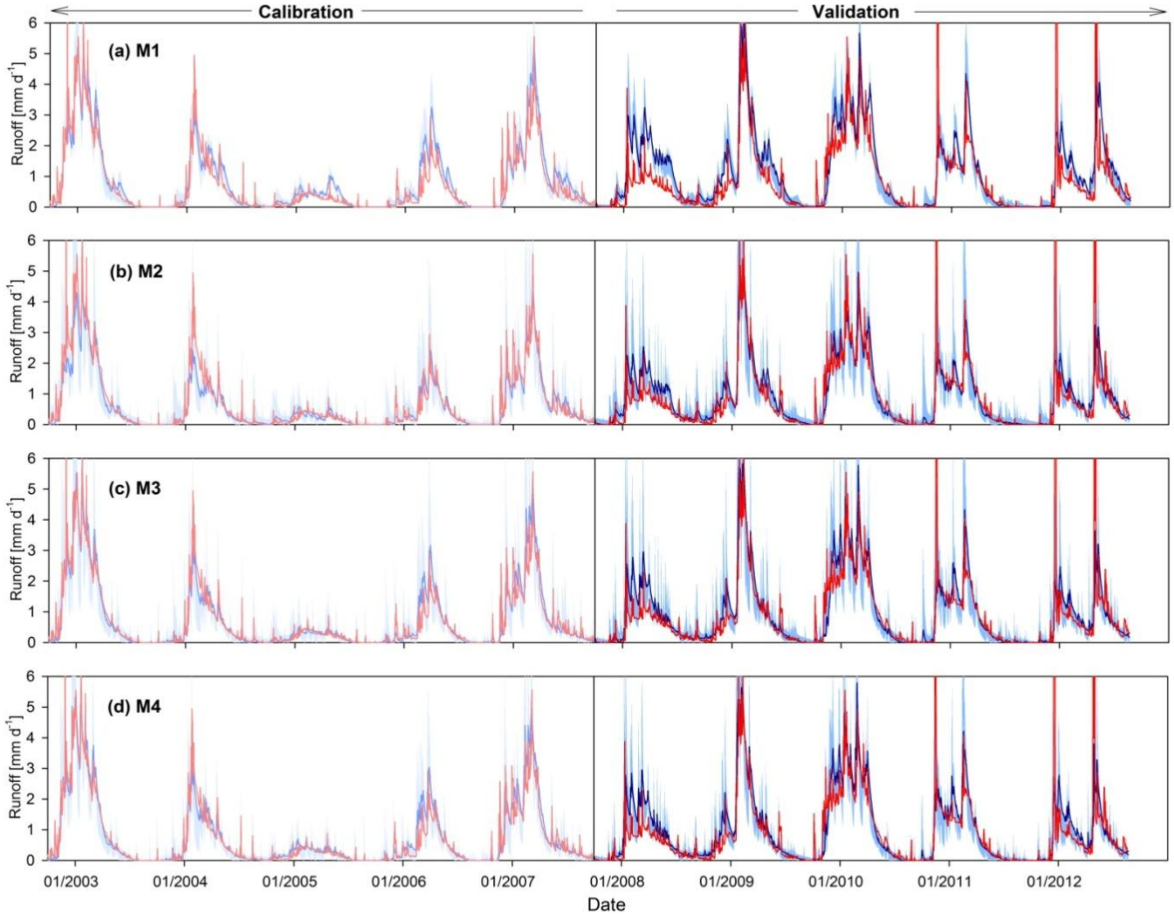


Figure 2. Normalized (a) groundwater levels for piezometers in the hillslope (F4 and F5b) and in the riparian zone (F1b) and (b) average, maximum and minimum unsaturated zone storages for all the sensors on the two profiles, on the Kerrien catchment.



a b

Figure 3. a. Observed (red line) and modelled runoff for model set-ups (a) M1, (b) M2, (c) M3 and (d) M4 in calibration and independent evaluation (validation) periods. Modelled runoff shown as most balanced solution (dark blue line) and the 5/95th uncertainty bounds (light blue shaded area). Adapted from Hrachowitz et al. (2014).



b. Overall model performance for all model set-ups (M1-M4) expressed as Euclidean Distance from the "perfect model" computed from all calibration objectives and signatures with respect to calibration and validation periods. Triangles represent the optimal solution, i.e. the solution obtained from the parameter set with the lowest Euclidean Distance during calibration. Box plots represent the Euclidean Distance for the complete sets of all feasible solutions (the dots indicate 5/95th percentiles, the whiskers 10/90th percentiles and the horizontal central line the median). From Hrachowitz et al. (2014).

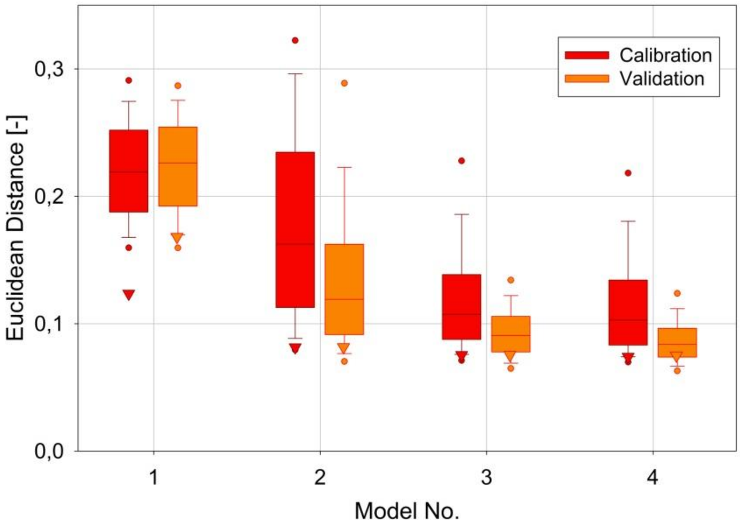
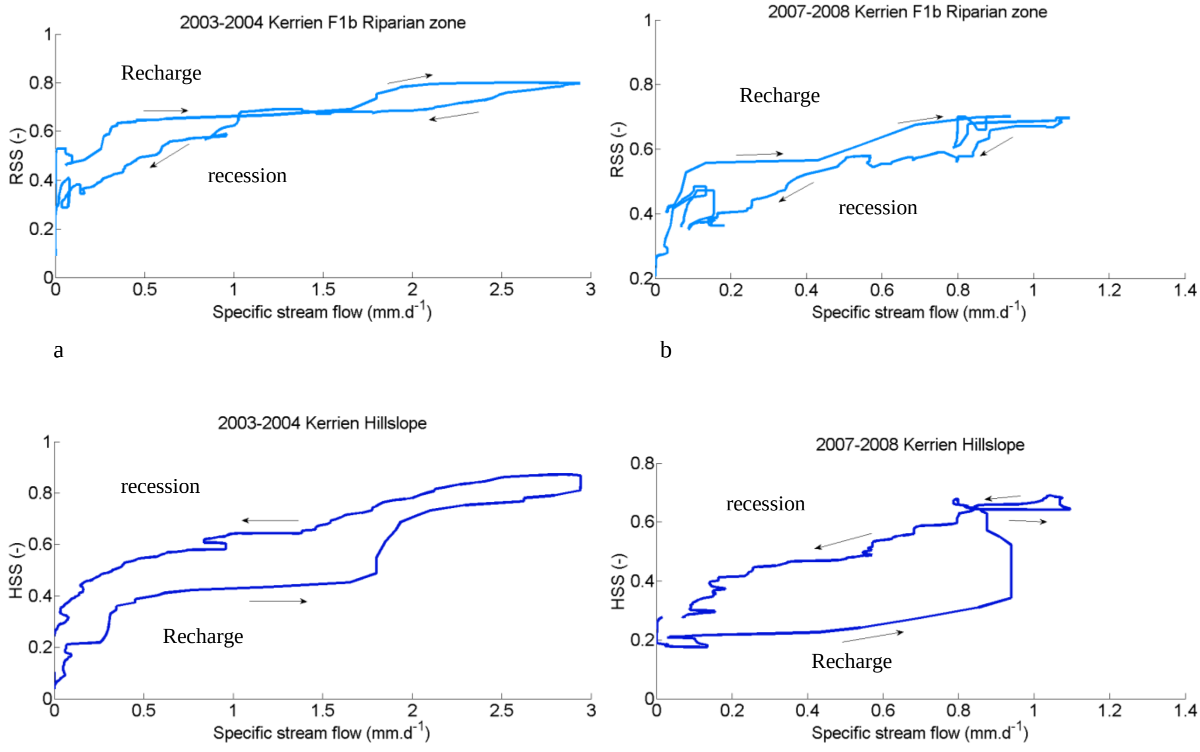


Figure 4. Examples of annual hysteretic loops for saturated zone storage vs. stream flow which are clockwise on the riparian zone (a, b) and anticlockwise on the hillslope (c, d) for the wet year 2003-2004 (a, c) and the dry year 2007-2008 (b, d).



c d

Figure 5. Annual Hysteresis Indices (HI) computed for the piezometers in the Kerrien catchment from 2002 to 2012. F5b is located upslope, F4 midslope and F1b downslope in the riparian area. RRS(Q) is the hysteresis between stream flow and riparian saturated zone storage (measured at F1b). HSS-F5b(Q), HSS-F4(Q) and HSS(Q) are hystereses between stream flow and upslope (at F5b), midslope (at F4) and hillslope (average of F5b and F4) Saturated Storages respectively. HUS(Q) is the hysteresis between stream flow and Hillslope Unsaturated Storage (HUS) (computed from the average of normalized volumic moisture sensors in profiles sB1 and sB2), and HUS(HSS) between the hillslope unsaturated and saturated zone storages (average of F5b and F4).

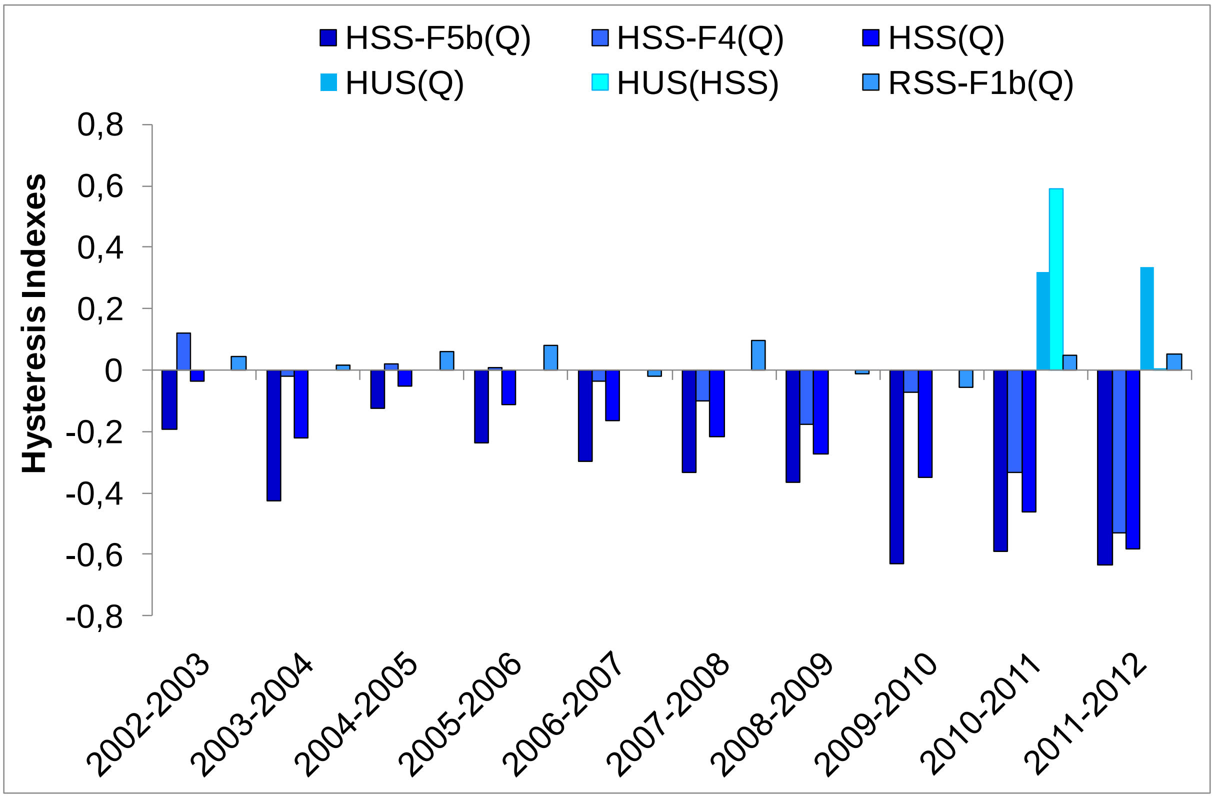
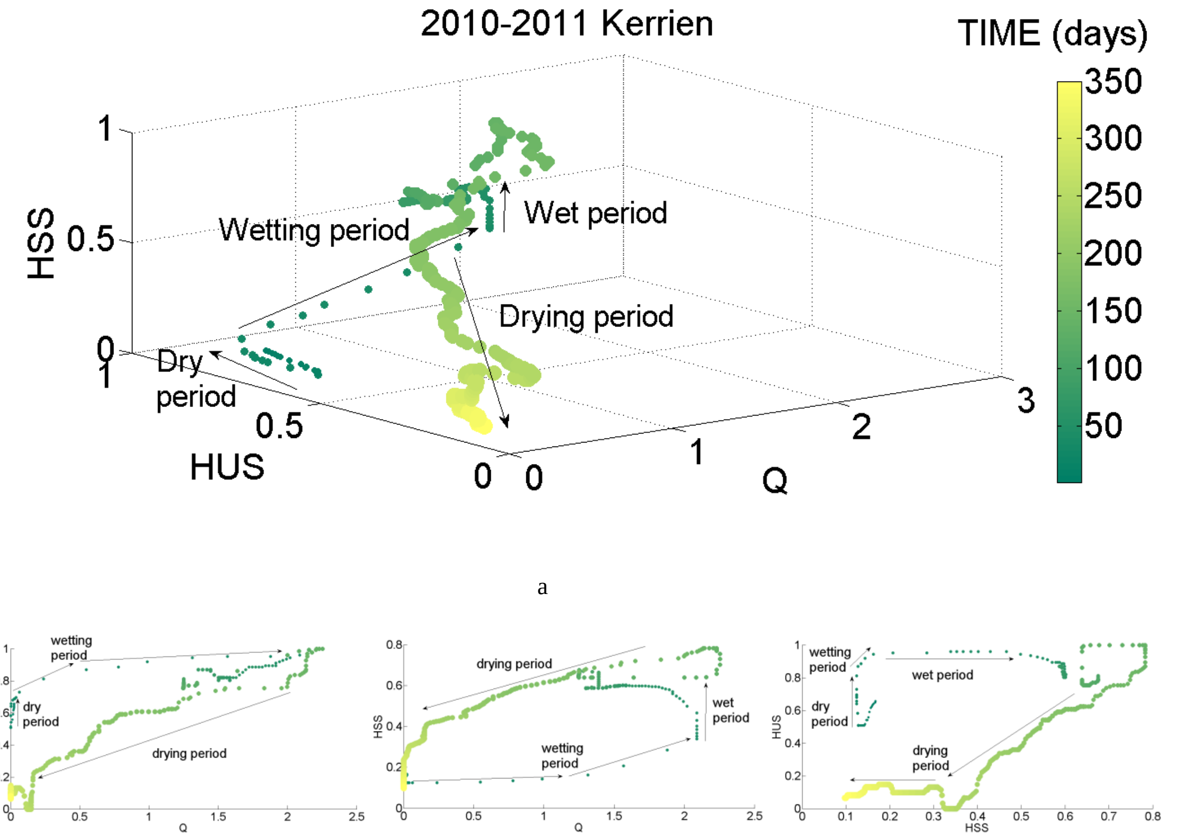


Figure 6. Evolution of stream flow (Q in mm.d^{-1}) and normalized Hillslope Unsaturated Storage (HUS) and Hillslope Saturated Strorage (HSS) for the water year 2010-2011 (October to September). The size of the dots is increasing with time. Unsaturated storage (HUS) is computed from the moisture sensors in profiles sB1 and sB2), saturated storage (HSS) is represented using normalized groundwater table level (computed from 2 piezometers in the hillslope). (a) is the 3D plot and (b, c, d) are the respective 2D projections of (a) on the three plans.



a

b c d

Figure 7: Conceptual scheme of successive mechanisms explaining the annual hysteresis between storages and stream flows. HUS: Hillslope unsaturated storage, HSS: hillslope saturated storage, RUS: riparien unsaturated storage, RSS: riparian saturated storage, Q: stream flow, bold characters indicate compartments with varying storage, grey arrows indicate if the compartment is filling or emptying, black arrows indicate the water flow paths.

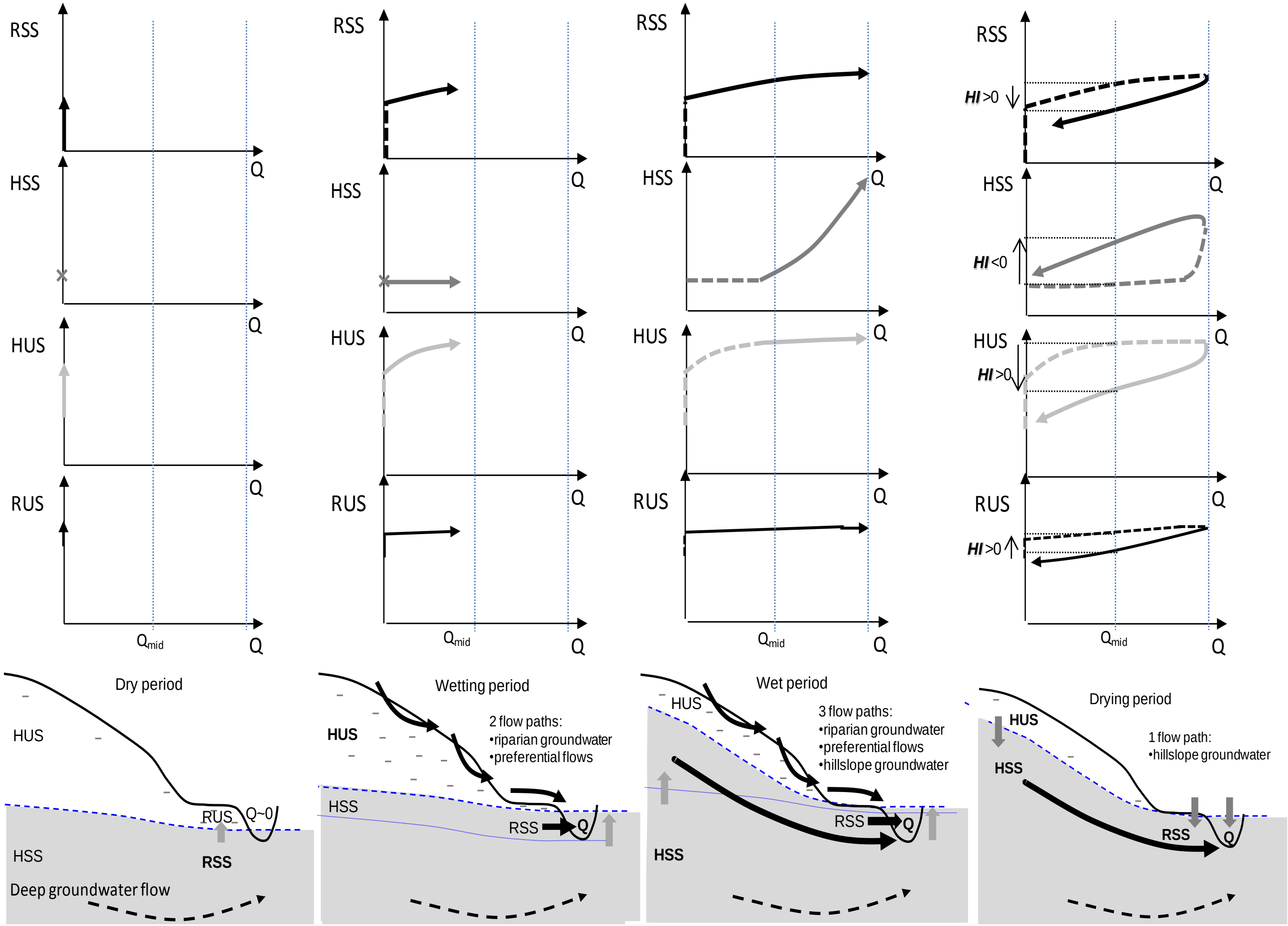
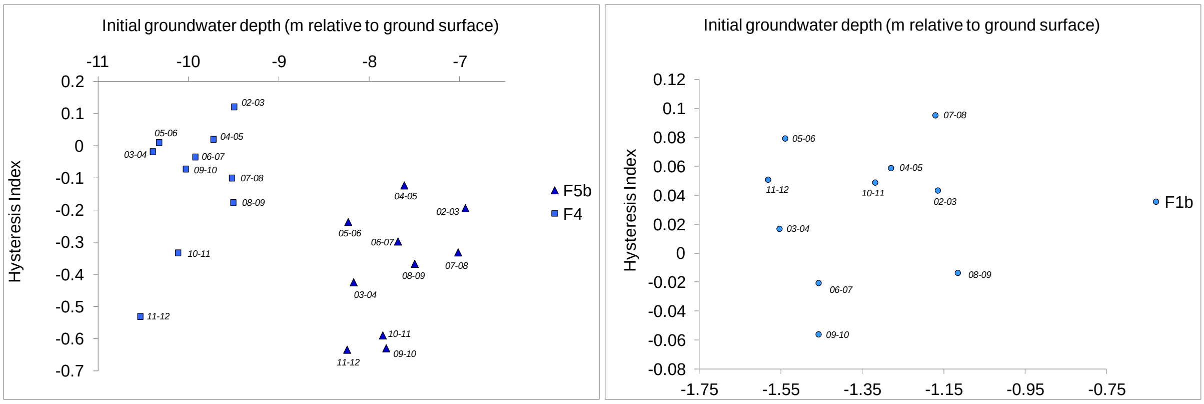
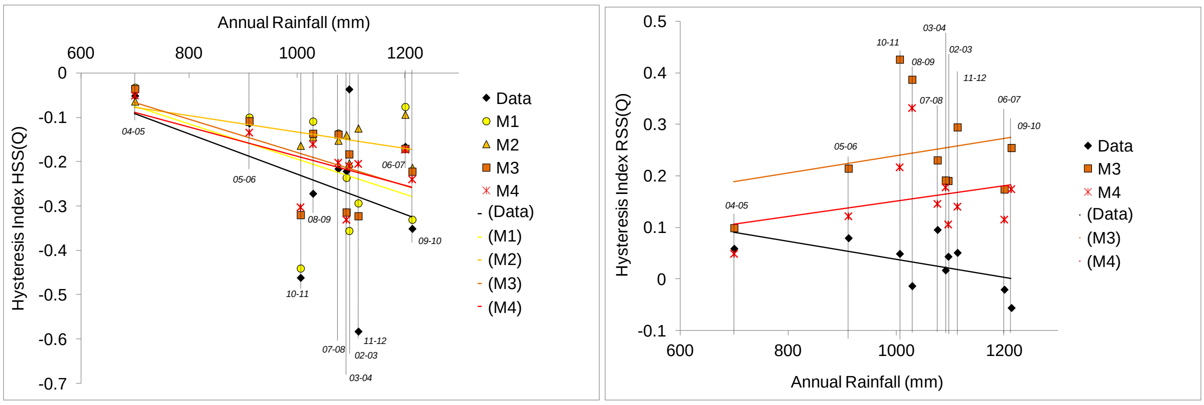


Figure 8. Year to year variations, for the 10 monitoring years, of the hysteresis indices a) HSS-F5b(Q) and HSS-F4(Q) (HI) versus the initial groundwater table level depth in the corresponding hillslope piezometer (F5b or F4) and b) HSS-F1b(Q) versus the initial groundwater table level depth in the piezometer in the riparian area (F1b).



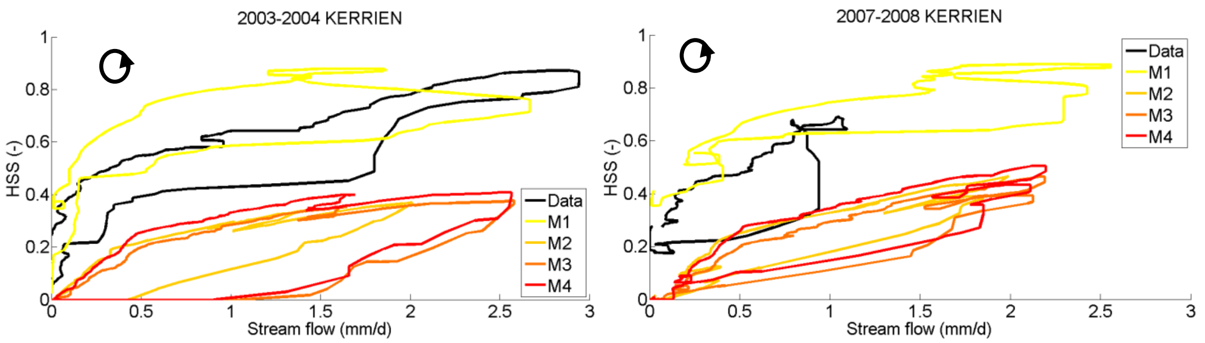
a b

Figure 9. Variations of observed (Data) and simulated (M1 to M4) hysteresis Index versus annual rainfall for the 10 monitored water years for (a) Hillslope Saturated Storage versus discharge HSS(Q), (b) Riparian Saturated Storage vs. discharge RSS(Q). Solid lines indicate the linear regressions



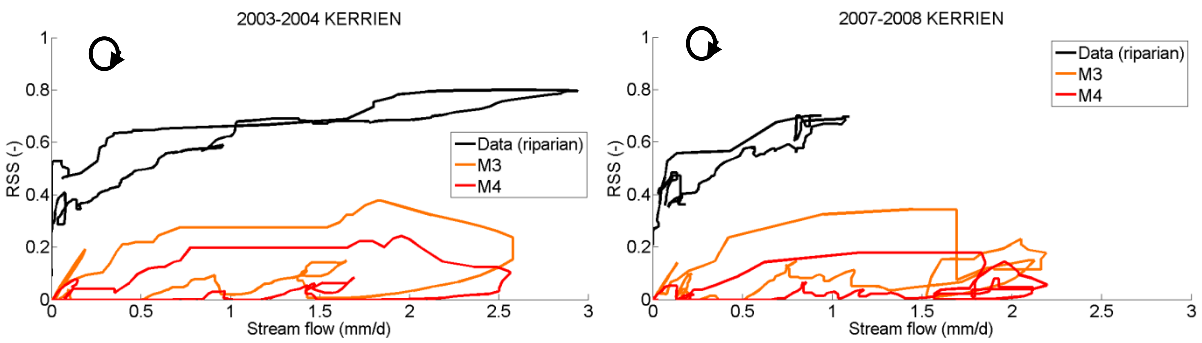
a b

Figure 10. Observed and simulated annual hysteresis between stream flow (Q) and (a, b) Saturated Storage in the hillslope HSS (for observed, HSS is the average of F5b and F4) and (c, d) Saturated Storage in the riparian area RSS (for simulated, only M3 and M4 represent the riparian area), for the water years (a, c) 2003-2004 (wet year, calibration period) and (b, d) 2007-2008 (dry year, validation period).



a

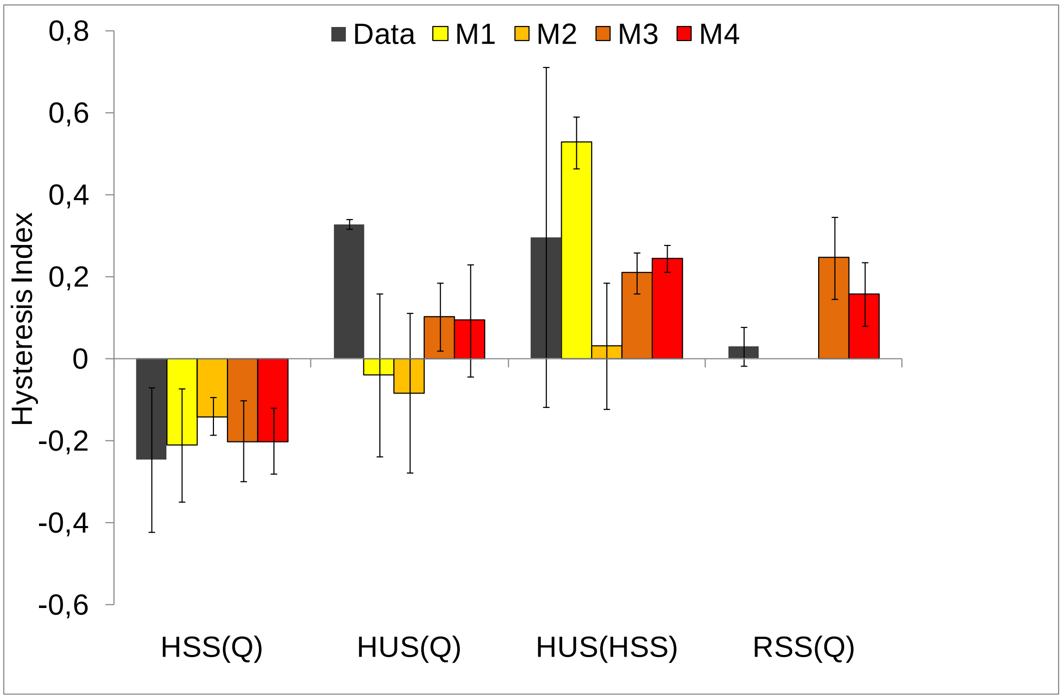
b



c

d

Figure 11. a.Mean annual hysteresis Indices observed and simulated with the 4 models M1 to M4, for Hillslope saturated storage vs. discharge HSS(Q), Hillslope unsaturated storage vs. discharge HUS(Q), Hillslope unsaturated storage vs. Hillslope saturated storage HUS(HSS), and Riparian saturated storage vs. discharge RSS(Q). RSS is simulated only in models M3 and M4. Error bars show the standard deviation for the 10 years for HSS(Q) and RSS(Q), and the values for the two available years for HUS(Q) and HUS(HSS).



b.Sensitivity of Hysteresis Index values to parameter uncertainty for the year 2011-2012. Mx bbp indicates the value for best behavioural parameter sets, the circles, triangles, squares ,and diamonds indicate the mean HI value for the all the behavioural parameter sets, and the corresponding bars its range of variation.

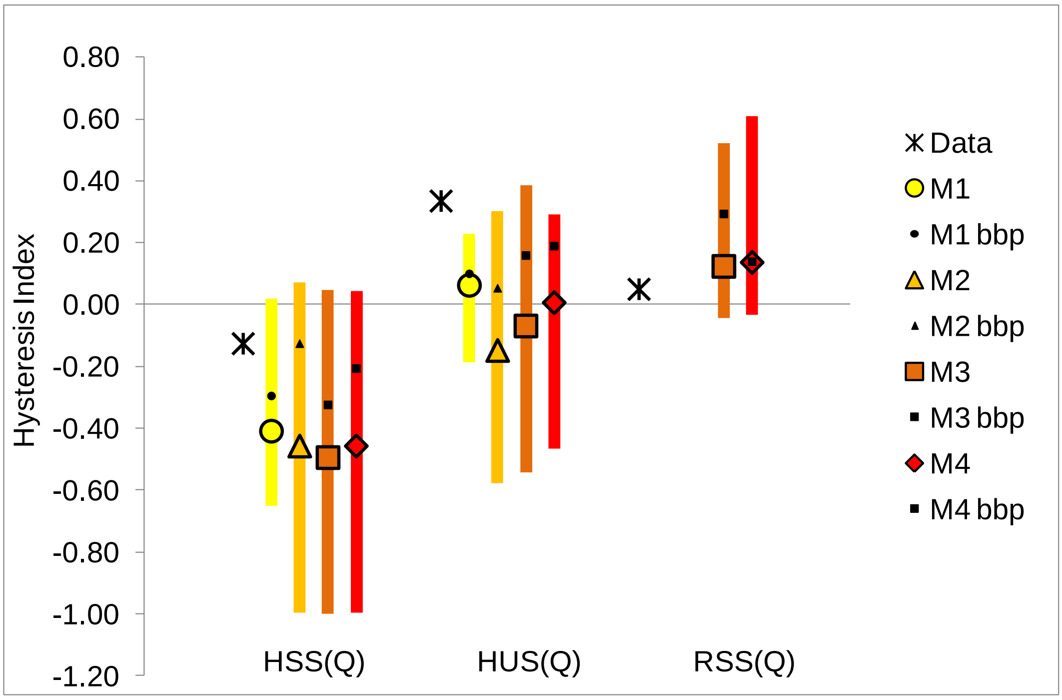


Table 1. Water balance, state and flux equations of the models used.

Process	Water balance	Eq.	Models	Flux and state equations	Eq.	Models
Unsaturated zone	$dS_U/dt = P - E_U - R_F - R_P - R_S$	(1.1)	M1,2,3,& 4	$E_U = E_P \mathrm{Min}\left(1, \frac{S_U}{S_{Umax,H}} \frac{1}{L_P}\right)$	(1.2)	M1,2,3,& 4
				$R_U = (1 - C_R)P$	(1.3)	M1,2,3,& 4
				$R_F = C_R(1 - C_P)P$	(1.4)	M1,2,3,& 4
				$R_S = P_{max}\left(\frac{S_U}{S_{Umax,H}}\right)$	(1.5)	M1,2,3,& 4
				$C_R = \frac{1}{1 + exp\left(\frac{-S_U/S_{Umax,H} + 0.5}{\beta}\right)}$	(1.6)	M1,2,3,& 4
Fast reservoir	$dS_F/dt = R_F - Q_F - E_F$	(2.1)	M1,2,3,& 4	$S_{F,in} = S_F + R_F$	(2.2)	M1,2,3,& 4
				$Q_F = S_{F,in}(1 - e^{-k_F t})$	(2.3)	M1,2,3,& 4
				$E_F = \mathrm{Min}(E_P - E_U, S_{F,in} - Q_F)$	(2.4)	M1,2,3,& 4
Slow reservoir	$dS_S/dt = R_S + R_P - Q_S$	(3.1)	M1	$S_{S,in} = S_S + R_S + R_P$	(3.2)	M1
				$Q_S = S_{S,in}(1 - e^{-k_S t})$	(3.3)	M1
	$dS_{s,a}/dt = \begin{cases} S_{s,a} - \mathrm{Max}(0, S_{S,tot,out}), S_{S,tot,in} > 0 \\ 0, \quad S_{S,tot,in} \leq 0 \end{cases}$	(3.4)	M2, 3 & 4	$Q_S = \mathrm{Max}(0, Q_{S,tot} - Q_{L,cst})$	(3.7)	M2, 3 & 4
	$dS_{s,p}/dt = \begin{cases} S_{s,p} + \mathrm{Min}(0, S_{S,tot,out}), S_{S,tot,in} > 0 \\ S_{s,p} + S_{S,tot,out}, \quad S_{S,tot,in} \leq 0 \end{cases}$	(3.5)	M2, 3 & 4	$S_{S,tot,in} = S_{s,a} + S_{s,p} + R_S + R_P$	(3.8)	M2, 3 & 4
	$dS_s/dt = dS_{s,a}/dt + dS_{s,p}/dt = R_S + R_P - Q_{L,cst}$	(3.6)	M2, 3 & 4	$S_{S,tot,out} = \begin{cases} S_{S,tot,in} e^{-k_S t} - \frac{Q_{L,cst}}{k_S}(1 - e^{-k_S t}), \quad S_{S,tot,in} > 0 \\ S_{S,tot,in} - Q_{L,cst}, \quad S_{S,tot,in} \leq 0 \end{cases}$	(3.9)	M2, 3 & 4
				$Q_{L,cst} = \text{constant}$	(3.10)	M2, 3 & 4
Unsaturated riparian zone	$dS_{U,R}/dt = P - E_{U,R} - R_R$	(4.1)	M3 & 4	$E_{U,R} = E_P \mathrm{Min}\left(1, \frac{S_{U,R}}{S_{Umax,R}} \frac{1}{L_P}\right)$	(4.2)	M3 & 4
				$R_R = C_{R,R} P$	(4.3)	M3 & 4
				$C_{R,R} = \mathrm{Min}\left(1, \frac{S_{U,R}}{S_{Umax,R}}\right)$	(4.4)	M3
				$C_{R,R} = \mathrm{Min}\left(1, \left(\frac{S_{U,R}}{S_{Umax,R}}\right)^{\beta_R}\right)$	(4.5)	M4
Riparian reservoir	$dS_R/dt = R_R - Q_R - E_R$	(5.1)	M3 & 4	$S_{R,in} = S_R + R_R$	(5.2)	M3 & 4
				$Q_R = S_{R,in}(1 - e^{-k_R t})$	(5.3)	M3 & 4
				$E_R = \mathrm{Min}(E_P - E_{U,R}, S_{R,in} - Q_R)$	(5.4)	M3 & 4
Total runoff	$Q_T = Q_F + Q_S$	(6.1)	M1 & 2			

	$Q_T = (1-f)(Q_F + Q_S) + fQ_R$	(6.2)	M3 & 4
Total evaporative fluxes	$E_A = E_U + E_F$	(7.1)	M1 & 2
	$E_A = (1-f)(E_U + E_F) + f\left(E_{U,R} + E_R\right)$	(7.2)	M3 & 4

List of Symbols:

C_P: Preferential recharge coefficient [-]
C_R: Hillslope runoff generation coefficient [-]
$C_{R,R}$: Riparian runoff generation coefficient [-]
k_F: Storage coefficient of fast reservoir [T^{-1}]
k_S: Storage coefficient of slow reservoir [T^{-1}]
k_L: Storage coefficient for deep infiltration loss [T^{-1}]
k_R: Storage coefficient of riparian reservoir [T^{-1}]
f: Proportion wetlands in the catchment [-]
L_P: Transpiration threshold [-]
P_{max}: Percolation capacity [L T^{-1}]
$S_{Umax,H}$: Unsaturated hillslope storage capacity [L]
$S_{Umax,R}$: Unsaturated riparian storage capacity [L]
β: Hillslope shape parameter for C_R [-]
β_R: Riparian shape parameter for $C_{R,R}$ [-]

P: Total precipitation [L T^{-1}]
E_F: Transpiration fast responding reservoir [L T^{-1}]
E_P: Potential evaporation [L T^{-1}]
E_R: Transpiration from riparian reservoir [L T^{-1}]
E_U: Transpiration from unsaturated reservoir [L T^{-1}]
$E_{U,R}$: Transpiration unsaturated riparian reservoir [L T^{-1}]
Q_R: Runoff from riparian reservoir [L T^{-1}]
Q_S: Runoff from slow reservoir [L T^{-1}]
Q_F: Runoff from fast reservoir [L T^{-1}]
$Q_{L,const}$: Constant deep infiltration loss [L T^{-1}]
R_F: Recharge of fast reservoir [L T^{-1}]
R_P: Preferential recharge of slow reservoir [L T^{-1}]
R_R: Recharge of riparian reservoir [L T^{-1}]
R_S: Recharge of slow reservoir [L T^{-1}]
R_U: Infiltration into unsaturated reservoir [L T^{-1}]

S_F: Storage in fast reservoir [L]
S_R: Storage in riparian reservoir [L]
S_S: Storage in slow reservoir [L]
$S_{S,a}$: Active storage in slow reservoir [L]
$S_{S,p}$: Passive storage in slow reservoir [L]
$S_{S,tot}$: Total storage in slow reservoir [L]
S_U: Storage in unsaturated reservoir [L]
$S_{S,tot,in}$: Total storage incoming in slow reservoir [L]
$S_{S,tot,out}$: Total storage outcoming from slow reservoir [L]

Table 2. Model structures and parameters.

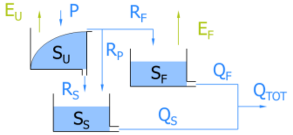
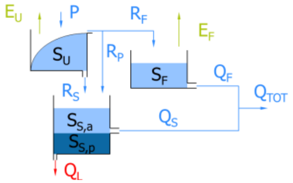
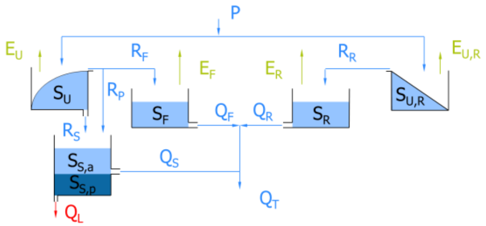
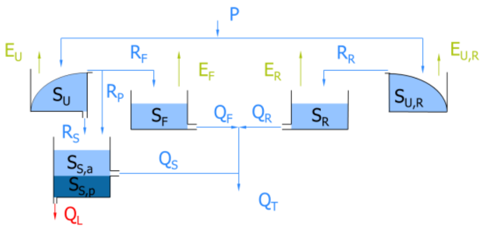
Model structure	Name	Parameters	Equations
	M1	k_F, $k_{S,}$ P_{max}, L_P, $S_{Umax,H}$, β, C_P	(1.1) to (1.6); (2.1) to (2.4); (3.1) to (3.3); (6.1) & (7.1)
	M2	k_F, k_S,P_{max}, L_P, $S_{Umax,H}$, β, $C_{P,}$ $Q_{L,cst}$	(1.1) to (1.6); (2.1) to (2.4); (3.4) to (3.10); (6.1) & (7.1)
	M3	k_F, k_S, P_{max}, L_P, $S_{Umax,H}$, β, C_P, $Q_{L,cst}$, k_R, f, $S_{Umax,R}$	(1.1) to (1.6); (2.1) to (2.4); (3.4) to (3.10); (4.1) to (4.4); (5.1) to (5.4); (6.2) & (7.2)
	M4	k_F, k_S, P_{max}, L_P, $S_{Umax,H}$, β, C_P, $Q_{L,cst}$, f, k_R, $S_{Umax,R}$, $β_R$	(1.1) to (1.6); (2.1) to (2.4); (3.4) to (3.10); (4.1) to (4.3); (4.5); (5.1) to (5.4); (6.2) & (7.2)

Table 3. Prior and posterior distribution of the model parameters.

		C_p	f	k_F	k_R	k_S	L_p	$Q_{L,const}$	P_{max}	$S_{S,p,max}$	$S_{Umax,H}$	$S_{Umax,R}$	β	β_R
		[-]	[-]	[d^{-1}]	[d^{-1}]	[d^{-1}]	[-]	[mm d^{-1}]	[mm d^{-1}]	[mm]	[mm]	[mm]	[-]	[-]
Prior distribution		0-1	0.1	0.025-1	0.05-2	0.001-0.05	0-1	0.37	0-4	0-2000	0-1500	0-750	0-100	0-2
Posterior distribution	M1	0.12/0.63		0.042/0.094		0.031/0.049	0.00/0.07		0.03/0.29		637/1446		10.5/61.5	
	M2	0.14/0.55		0.054/0.627		0.041$^{*)}$	0.05/0.34	0.37$^{*)}$	0.27/1.98		722/1461		2.4/36.9	
	M3	0.15/0.64	0.1$^{*)}$	0.054/0.619	0.333/1.863	0.041$^{*)}$	0.04/0.27	0.37$^{*)}$	0.34/2.29		686/1442	132/725	13.6/69.7	
	M4	0.19/0.64	0.1$^{*)}$	0.054/0.466	0.318/1.857	0.041$^{*)}$	0.04/0.27	0.37$^{*)}$	0.29/2.18		683/1444	120/730	13.0/69.2	0.13/1.86

Table 4: Hydrological calibration criteria and evaluation signatures.

The performance metrics include the Nash-Sutcliffe Efficiency (E_{NS}), the Volume Error (E_V) and the Relative Error (E_R). For all variables and signatures, except for Q, Qlow and GW, the long-term averages were used.

$$E_{NS,X} = 1 - \frac{\sum_{i=1:n}\sqrt{\left(X_{obs,i} - X_{sim,i}\right)^2}}{\sum_{i=1:n}\sqrt{\left(X_{obs,i} - \frac{1}{n}\sum_{i=1:n}X_{obs,i}\right)^2}}$$

$$E_{V,X} = 1 - \frac{\sum_{i=1:n}\left|X_{obs,i} - X_{sim,i}\right|}{\sum_{i=1:n}X_{obs,i}}$$

$$E_{R,X} = \frac{X_{obs} - X_{sim}}{X_{obs}}$$

	Variable/Signature	ID	Performance metric	Reference
Calibration	Time series of flow	O1	$E_{NS,Q}$	Nash and Sutcliffe (1970)
		O2	$E_{NS,log(Q)}$	
		O3	$E_{V,Q}$	Criss and Winston (2008)
	Flow duration curve	O4	$E_{NS,FDC}$	Jothityangkoon et al. (2001)
			$D_{E,cal}$	Schoups et al. (2005)
Evaluation	Flow during low flow period	S1	$E_{NS,Q,low}$	Freer et al. (2003)
	Groundwater dynamics[a)]	S2	$E_{NS,GW}$	Fenicia et al. (2008a)
	Flow duration curve low flow period	S3	$E_{NS,FDC,low}$	Yilmaz et al. (2008)
	Flow duration curve high flow period	S4	$E_{NS,FDC,high}$	Yilmaz et al. (2008)
	Groundwater duration curve[a)]	S5	$E_{NS,GDC}$	-
	Peak distribution	S6	$E_{NS,PD}$	Euser et al. (2013
	Peak distribution low flow period	S7	$E_{NS,PD,low}$	Euser et al. (2013)
	Rising limb density	S8	$E_{R,RLD}$	Shamir et al. (2005)
	Declining limb density	S9	$E_{R,DLD}$	Sawicz et al. (2011)
	Auto-correlation function of flow[b)]	S10	$E_{NS,AC}$	Montanari and Toth (2007)
	Lag-1 auto-correlation of high flow period	S11	$E_{R,AC1,Q10}$	Euser et al. (2013)
	Lag-1 auto-correlation of low flow period	S12	$E_{R,AC1,low}$	Euser et al. (2013)
	Runoff coefficient[c)]	S13	$E_{R,RC}$	Yadav et al. (2007)
			D_E	Schoups et al. (2005)

[a)]Averaged and normalized time series data of the five piezometer were compared to normalized fluctuations in model state variable SS

[b)]Describing the spectral properties of a signal and thus the memory of the system, the observed and modelled auto-correlation functions with lags from 1-100d where compared

[c)]Note that in catchments without long-term storage-changes and inter-catchment groundwater flow, long-term average RC equals the long-term average 1-EA

Table 5. Hysteresis indices (HI) and initial hillslope unsaturated storage values (HUS) at the beginning of the water year

Year.	initial HUS	Hysteresis Index (HI)			
		HSS-F5b(Q)	HSS-F4(Q)	HSS(Q)	RSS-F1b(Q)
2010-2011	0.148	-0.591	-0.334	-0.462	0.590
2011-2012	0.026	-0.635	-0.532	-0.583	0.003