Towards a better process representation of catchment hydrology in conceptual runoff modelling

Jan Seibert, Jeff McDonnell
Oregon State University, Department of Forest Engineering, Corvallis OR 97331, U.S.A.

Abstract

Hydrologists often have a detailed yet highly qualitative understanding of dominant runoff processes—thus we usually know much more about a catchment than we use for calibration of a model. We present a new method where soft data (*i.e.*, qualitative knowledge from the experimentalist that cannot be used directly as exact numbers) is made useful through fuzzy measures of model-simulation and parameter-value acceptability. A three-box lumped conceptual model was developed for the Maimai catchment in New Zealand, a particularly well-studied process-hydrological research catchment. The boxes represent the key hydrological reservoirs that are known to have distinct groundwater dynamics, isotopic composition and solute chemistry. The model was calibrated against hard data (runoff and groundwater-levels) as well as a number of criteria derived from the soft data (*e.g.* percent new water, reservoir volume etc). We achieved very good fits for the three-box model when optimizing the parameter values with only runoff (R_{eff}=0.92). However, parameter sets obtained in this way, showed in general a poor goodness-of-fit for other criteria such as the simulated new-water contributions to peak runoff. Including soft-data criteria lead to lower R_{eff}-values (around 0.86 when including all criteria) but led to better overall performance. We argue that accepting lower model efficiencies for runoff is worth it if one can develop a more “real” model of catchment behavior.

Introduction

Catchment hydrology is at a cross-roads. While complex descriptions of the age, origin and pathway of subsurface stormflow abound in the literature (reviewed recently by Bonell, 1998), most catchment modeling studies do not fully use this information for model development, calibration and testing. As a result, process hydrological studies of dominant runoff producing processes and model studies of runoff generation are often poorly linked. Recently there has been a tendency away from fully-distributed, physically-based models back to conceptual models due to concerns overparameterisation, parameter uncertainty and model output uncertainty. While conceptual models may be much more simplified and lumped they offer the potential for development based on process understanding of key zones or reservoirs of catchment response. As interest in the geochemical dimensions of streamflow modeling increases, reservoir (or box model) conceptual approaches that explicitly treat volume-based mixing and water (and ultimately tracer) mass balance become increasing useful (Harris et al., 1995; Hooper et al., 1998; Seibert, 1999).

A major obstacle in moving forward with conceptual modeling approaches is how to fully utilize experimental data for internal calibration and validation. Currently, the use of this field data for model calibration is often limited beyond simple streamflow information despite the general acceptance that internal state information is necessary for ensuring model consistency. The usefulness of having various criteria for assessment of model performance is widely accepted (Freer et al., 1998). When multi-criteria are used for calibration or validation, this has often meant only the use of two or three criteria (*e.g.* runoff and groundwater levels) as compared to only one criterion (*i.e.* runoff). Clearly, more criteria are desirable but in most cases there is no suitable data available. The dilemma in conceptual modeling of catchment hydrology is that parsimonious models, which may allow identification of parameter values through calibration against runoff, in general are too simple to allow a realistic representation of the main hydrological processes and, thus, provide only limited possibilities for internal model testing. This paper explores a new philosophy and approach for development of more realistic models of catchment behavior using "soft data" where multiple criteria can now be used to constrain the model in various aspects.

Hydrologists often have a detailed yet qualitative understanding of dominant runoff processes and we usually know much more about a catchment than we use for calibration of a model. This soft data is a qualitative knowledge from the experimentalist, that cannot be used as exact numbers but is made useful through fuzzy measures of model-simulation and parameter-value acceptability. We argue that this method is the necessary dialog that must occur between the modeler and the experimentalist to enable a better process representation of catchment hydrology in conceptual runoff models. We use the well-characterized Maimai watershed as the testing ground for these new ideas. Thus, this paper: (1) presents a new 3-box model of headwater catchment response based on an extension of ideas developed in Seibert *et al.* (this issue), (2) incorporates a number of soft-data measures from experimental studies at the catchment for model calibration, and (3) assesses the value of soft data relative to traditional hard information measures.

Study site and perceptual model

The Maimai watershed

Maimai M8 is small 3.8 ha study catchment located to the east of the Paparoa Mountain Range on the South Island of New Zealand. Slopes are short (<300 m) and steep (average 34^o) with local relief of 100-150 m. Stream channels are deeply incised and lower portions of the slope profiles are strongly convex. Areas that could contribute to storm response by saturation overland flow are small and limited to 4-7 % (Mosley, 1979; Pearce et al., 1986). Mean annual precipitation is approximately 2600 mm, producing an estimated 1550 mm of runoff. The summer months are the driest; monthly rainfall from December to February averages 165 mm and for the rest of the year between 190 to 270 mm. On average, there are 156 rain days per year with little temperature extreme and only about 2 snow days per year (Rowe et al., 1994). In addition to being wet, the catchments are highly responsive to storm rainfall. Quickflow comprises 65% of the mean annual runoff and 39% of annual total rainfall (Pearce et al., 1986).

The watershed is underlain by a firmly compacted poorly impermeable conglomerate--seepage losses to deep groundwater are estimated at 100 mm/yr. The wet and humid climatic environment, in conjunction with topographic and soil characteristics, result in the soils normally remaining within 10% of saturation (Mosley, 1979). As a result, the soils are strongly weathered and leached, with low natural fertility. The thin nature of the soils promotes the lateral development of root networks and channels. Soil profiles reveal extensive macropores and preferential flow pathways at vertical pit faces which form along cracks and holes in the soil and along live and dead root channels (Mosley 1979). Lateral root channel networks are evident in the numerous tree throws that exist throughout the catchments. Preferential flow also occurs along soil horizon planes and the soil-bedrock interface.

Perceptual model of the Maimai watershed

M8 has been the site of ongoing hillslope research by several research teams since the late 1970s. These studies have facilitated the development of a very detailed yet qualitative perceptual model of hillslope hydrology, reviewed recently by McDonnell *et al.* (in press). While dye tracer studies by Mosley (1979) showed that storm rainfall follows preferential flow pathways at the hillslope scale, subsequent water isotopic tracing studies in the catchment by Pearce *et al.* (1986) and Sklash *et al.* (1986) showed (paradoxically) that there was little if any "event" water in the stream during stormflow periods. Thus, stored soil water and groundwater comprise the majority of channel stormflow. McDonnell (1990) developed a perceptual model to explain the mechanism of stormflow generation by constraining the dominant processes using recording tensiometer observations, isotope tracing and various other chemical and hydrometric approaches. For small events of less than about 15 mm rainfall, McDonnell *et al.* (1991) found that the riparian zone could account for the volume of old water in the channel hydrograph. During larger events, McDonnell (1990) found that hillslope hollows were the dominant runoff producing zones where new water moved to depth and created a perched water table at the soil-bedrock interface. Lateral pipeflow then formed along the soil bedrock interface (McDonnell *et al.*, 1997), conveying quantities of old water laterally downslope sufficient in quantity and quality to explain measured old water volumes. Topographic convergence of flowpaths from hillslopes to the hollows enabled hollows to be well-primed for rapid conversion of matrix to pressure potentials. Soil water isotopic composition (McDonnell *et al.*, 1991) and chemical composition (Grady *et al.*, 2000) all followed a similar pattern of evolution along three major runoff response zones and inter-storm reservoirs: hillslopes, hollows and riparian zones.

Model theory

Conceptual three-box model

The conceptual model is based on the three reservoirs identified from the experimental studies at M8: riparian, hollow and hillslope zones (Fig. 1). Water is simulated to flow from the hillslope zone into the hollow zone and from the hollow zone into the riparian zone. Outflow from the riparian zone is taken as catchment runoff. No lateral flow is assumed to take place from the unsaturated reservoirs and no bypass flow from hillslope to the stream is allowed. Tensiometer data has shown that shallow groundwater in the hillslope reservoir (groundwater levels 0 – 1.5 m below the ground surface) exhibits considerable interaction between saturated and unsaturated storage. Thus, a coupled formulation of the saturated and unsaturated storage is used (for description see *Seibert et al.*, this issue).

Outflow from the hillslope and riparian reservoir is computed as a simple linear function. The hollow reservoir is given an additional threshold-based linear function based on the McDonnell (1990; pp. 2830 Fig 10) perceptual model (Eq. 1-3):

$$Q_{hillslope} = k_{1,hillslope}\ S_{hillslope} \tag{1}$$

$$\begin{aligned} Q_{hollow} &= k_{1,hollow}\ S_{hollow} && if\ \ S_{hollow} \le S_{threshold} \\ Q_{hollow} &= k_{1,hollow}\ S_{threshold} + k_{2,hollow}\left(S_{hollow} - S_{threshold}\right) && if\ \ S_{hollow} > S_{threshold} \end{aligned} \tag{2}$$

$$Q_{riparian} = k_{1,riparian}\ S_{riparian} \tag{3}$$

This use of a threshold in the hollow reservoir is also motivated by field observations reported by McDonnell *et al.* (1998) that indicate large fluxes through macropores along the bedrock-soil interface.

Utilization of soft information

Given the relatively large number of parameters (20) in the three-box model, the information contained in the hard data (runoff and two groundwater-level series) is insufficient for calibration of the parameter values. Soft data can be used in two ways to constrain the calibration: (1) to assess how reasonable the parameter values might be and (2) to evaluate aspects of the model simulations for which there is no hard data available (Table 1). A general characteristic of comparing parameter values or model simulations with such soft data is that there is a relatively wide range of similar acceptable values. Furthermore, there might be a range of values that fall between fully acceptable and not acceptable. Fuzzy measures of acceptance are used to consider these ranges. For each soft data a trapezoidal function (Equation 4) is defined to compute the degree of acceptance from the corresponding simulated quantity or parameter value.

$$\mu(x)=\begin{cases}0 & if\ \ x\le a_1\\ \dfrac{x-a_1}{a_2-a_1} & if\ \ a_1\le x<a_2\\ 1 & if\ \ a_2\le x<a_3\\ \dfrac{a_4-x}{a_4-a_3} & if\ \ a_3\le x<a_4\\ 0 & if\ \ x>a_4\end{cases} \qquad (4)$$

Soft data enables the judging of model simulations in more ways than those for which hard data are available. For instance, the field hydrologist might have an idea about the range in which groundwater levels fluctuate or the contribution of new water to peak flow. In this study, we assessed model performance by comparing simulated contributions of new water to catchment runoff with results from hydrograph separations reported in and McDonnell *et al.* (1991) for a number of events. For each event, ranges of degree of acceptance were computed based on the simulated new-water contribution and pre-defined ranges of for acceptable and perfect simulations (Table 2).

Model parameters in conceptual models are not directly measurable. Parameters may be related to measurable quantities but they are effective values for a much larger scale than the measurement scale. In general parameter values for a conceptual model are found by calibration. However, for many parameters the field hydrologist might reject or prefer values within certain ranges. Usually the search of parameter values is constrained by feasible ranges. Based on the perceptual model of the catchment runoff response we added a desirable range, which was smaller than the feasible range, for a number of parameters. For each of these parameters a degree of acceptance was computed. This value varied from one, if the value was within the desirable ranges and decreased towards zero with increasing deviations from this range (Table 2). For example, we allowed searching values from 5 to 20 percent for the spatial fraction of the hollow zone, but the degree of acceptance was one only for values between 10 and 15 percent. Based on the individual parameters the acceptability of a certain parameter set was computed as the geometric mean of the respective degrees of acceptance.

Table 1. The three different ways of evaluating model acceptability based on hard data (A1) and soft (A2 and A3) data..

	Acceptability according to …	**Example**	**Measure**
A_1	Fit between simulated and observed data	Runoff	Efficiency
A_2	Agreement with perceptual (qualitative) knowledge	New water contribution	Percentage of peak flow for certain events
A_3	Reasonability of parameter values	Spatial extension of riparian zone	Fraction of catchment area

The acceptability considering hard data (A1) was computed from the efficiency (*Nash & Sutcliffe*, 1970) of the runoff simulation, R_{eff}, the relative volume error, V, and r^2 values for the groundwater levels in the riparian and the hollow zone (Eq. 5).Following Lindström (1997) a value of 0.1 is chosen for the weighing coefficient ω.

$$A_1 = \frac{1}{2}\left(R_{eff} - \omega|V| + \sqrt{r^2_{gwhollow}\ r^2_{gw\ riparian}}\right) \qquad (5)$$

A_2 was computed as arithmetic mean of the 15 evaluation rules of the soft data regarding groundwater levels and contribution of new water (Table 2). The geometric mean of the nine evaluation rules of the different parameters was taken for A_3 (Table 2). Generally the geometric mean was used whenever all measures were larger zero.

The overall acceptability, A, of a parameter set was computed as a weighted geometric mean (Eq. 6). Values of 2, 2 and 1 were chosen for n_1, n_2, and n_3 respectively to place more emphasize on the acceptability with regard to the simulations.

$$A = \sqrt[n]{A_1^{n_1} A_2^{n_2} A_3^{n_3}} \quad with \ \ n = n_1 + n_2 + n_3 \qquad (6)$$

We also tested the worth of adding additional criteria by calibrating the model with a varying set of criteria. The genetic algorithm, as described by Seibert (2000), includes stochastic elements. Thus, calibrated parameter values may differ, especially when there is a significant parameter uncertainty. To address the parameter uncertainty 15 calibration trials were performed for each goodness-of-model measure and the best 12 parameter sets were used for further analysis. The period of record for calibration was August-December, 1987.

Table 2. Evaluation rules based on soft data used for model calibration (the values for a_i define the trapezoidal function used to compute the degree of acceptance, see Eq. 6)

Type of soft information	**Specific soft information**	a_1	a_2	a_3	a_4	**Motivation**
New water contribution to peak runoff [-]	870930 18.00	0.03	0.06	0.12	0.15	McDonnell et al 1991
	871008 3.00	0.05	0.13	0.31	0.40	“
	871010 17.00	-	0	0.03	0.06	“
	871013 11.00	0.17	0.23	0.35	0.41	“
	871113 19.00	-	0	0.03	0.06	“
	871127 8.00	0.04	0.07	0.13	0.16	“
Range of groundwater levels, min./max. fraction of saturated soil [-]	Maximum hillslope	0	0.2	0.5	0.7	McDonnell (1989)
	Maximum hollow	0	0.5	0.75	1	McDonnell (1990)
	Minimum hollow	0	0.05	0.1	0.2	“
	Minimum riparian	0.05	0.1	0.3	0.5	“
Frequency of groundwater levels above a certain level (as fraction of soil [-]) [-]	Hillslope, above 0.5 during events	-	0	0.1	0.3	McDonnell et al. (1997)
	Hollow above 0.7 during events	-	0	0.1	0.2	McDonnell (1990)
	Hollow above 0.9 during events	-	-	0	0.1	McDonnell (1990)
	Riparian above 0.2	0.6	0.8	1	1	McDonnell (1990)
	Riparian above 0.9 during events	0	0.25	0.75	1	McDonnell (1990)
Parameter values	Fraction of riparian zone [-]	0.01	0.03	0.07	0.10	Mosley (1979)
	Fraction of hollow zone [-]	0.05	0.10	0.15	0.20	McDonnell (1990)
	Porosity in hillslope zone [-]	0.45	0.6	0.7	0.75	(McDonnell, 1989)
	Porosity in hollow zone [-]	0.45	0.55	0.65	0.75	McDonnell (1989)
	Porosity in riparian zone [-]	0.45	0.5	0.6	0.75	“
	Soil depth for hillslope zone [m]	0.1	0.3	0.8	1.5	McDonnell et al. (1997)
	Soil depth for hollow zone [m]	0.5	1	2	2.5	“
	Soil depth for riparian zone [m]	0.15	0.4	0.75	1	“
	Threshold level in hollow zone, fraction of soil depth [-]	0	0.1	0.4	1	McDonnell (1990) McDonnell et al. (1991)

Results and Discussion

Model output

The model was able to reproduce observed runoff. Model simulations calibrated with only runoff values led to very good fits with R_{eff} of 0.92. Efficiency values were also satisfactory (around 0.86) when the model was calibrated with respect to all criteria (Fig.2). The decrease of unsaturated storage decreased during events is a result of the coupled formulation of saturated and unsaturated storage.

Parameter uncertainty

Adding different criteria in general reduced parameter uncertainty, but results were mixed among the parameters. The reduction of parameter uncertainty was most obvious for the outflow coefficients. The range of calibrated parameter values decreased when using all criteria compared to when only using the hard data (runoff and groundwater levels). The coefficients of variation computed from the 12 values obtained by the different calibration trials were used as a measure of parameter uncertainty. On average, using all criteria helped to reduce parameter uncertainty to a third relative to a single criteria calibration against only runoff. Including hard groundwater data or soft data for new-water contribution to peak runoff also reduced parameter uncertainty, but not as significant as the combination of different criteria.

Overall performance

The model performance with regard to the various criteria varied of course between the parameter sets, which had been calibrated using different combinations of these criteria (Figure 4). Calibration against only one or two criteria led to poor simulations according to the other criteria. For example, the best parameter sets according to runoff (median efficiency 0.93) were poor in their ability to correctly reproduce hard and soft groundwater levels (r^2=0.41 and μ_{gw}=0.29). While the calibration against all criteria did not provide the best fits according to single criteria, the best overall performance was obtained in this way. Compared to the calibration using only hard data (A1, runoff and groundwater) the efficiency dropped (median values from 0.91 to 0.86), but the contribution of new water to peak runoff was much better reproduced (median $\mu_{new\ water}$=0.8 compared to 0.67).

On the value of soft information

Runoff simulation from the Maimai watershed is relatively easy by comparison to many catchments since there is only minimal seasonality and soils are transmissive and underlain by impermeable substrate. Previous TOPMODEL simulations (Beven and Freer, 2000) at the site and the present study have all achieved good fits for streamflow. However, simply modeling runoff with a high efficiency is not a challenging test of model performance. This work shows that sometimes lower R_{eff}-values are “the price we have to pay” to obtain a better overall model performance and better adherence to the perceptual model of runoff generation. The question then becomes: Is this reduction worth accepting to achieve a better conceptualization with respect to the soft data available? We argue from data presented in this paper that it is indeed worth accepting lower runoff-efficiency values if one can develop a more “real” model of the catchment. The parameter set determined by using several criteria for calibration (based on hard and soft data) will in most cases lead to a poorer fit of simulated and observed catchment, but move the model to one that better captures the key processes that the experimentalist feels is important in controlling catchment response.

Concluding remarks

This study has attempted to use multi-criteria soft data for model development and for internal calibration and validation. The study shows that conceptual modeling of catchment hydrology can include identification of parameter values through calibration against hard and soft data. We believe that this approach is the way forward for development of more realistic models of catchment behavior using soft data where multiple criteria can now be used to constrain the model in various ways. This soft data is a qualitative knowledge from the experimentalist, which cannot be used as exact numbers but is made useful through fuzzy measures of model-simulation and parameter-value acceptability. We argue that the necessary dialog that must occur between the modeler and the experimentalist can be made explicit in this way. We propose that this approach is also useful for comparing the value of different field measurements that experimentalists might make in support of modeling. We are currently exploring other types of soft data (*e.g.* mean residence time data) as we move to larger watershed scales and begin to incorporate conservative mixing between reservoirs. Our main message in this work is that rather than being “right for the wrong reasons”, a better process representation of catchment hydrology in conceptual runoff modeling should be “less right, for the right reasons.”

References

Bonell, M. (1998). Selected challenges in runoff generation research in forests from the hillslope to headwater drainage scale. Journal of the American Water Resources Association, 34: 765-785.

Freer, J., J.J. McDonnell, D. Brammer, K. Beven, R. Hooper, D. Burns (1997). Topographic controls on subsurface stormflow at the hillslope scale for two hydrologically distinct catchments. Hydrological Processes, Vol 11(9): 1347-1352.

Grady, E, H. Elsenbeer, B. McGlynn and J. McDonnell (2000). The variability of soil water chemistry and its effect on EMMA-based hydrograph separation. Hydrological Processes, in review.

Harris, D.M., J.J. McDonnell and A. Rodhe (1995). Hydrograph separation using continuous open-system isotopic mixing. Water Resources Research., 31(1): 157-171.

Hooper, R., B. Aulenbach, D. Burns, J.J. McDonnell, J. Freer, C. Kendall and K. Beven (1998). Riparian control of streamwater chemistry: Implications for hydrochemical basin models. International Association of Hydrological Sciences, Publication 248: 451-458.

Lindström, G. (1997) A simple automatic calibration routine for the HBV model. *Nordic Hydrology*, **28**(3): 153-168

McDonnell, J.J., B. McGlynn and D. Brammer (2000). An evolving perceptual model of hillslope flow in a steep forested humid catchment: A review of the Maimai catchment. Journal of Hydrology, in press.

McDonnell, J.J., M.K. Stewart and I.F. Owens (1991). Effects of catchment-scale subsurface watershed mixing on stream isotopic response. Water Resources Research, 26:3065-3073.

McDonnell, J.J. (1990). A rationale for old water discharge through macropores in a steep, humid catchment. Water Resources Research, 26(11)2821-2832.

McDonnell, J.J., D. Brammer, C. Kendall, N. Hjerdt, L. Rowe, M. Stewart and R. Woods (1998). Flow pathways on steep forested hillslopes: The tracer, tensiometer and trough approach, In Tani et al. (eds). Environmental Forest Science, Kluwer Academic Publishers, pp. 463-474.

Mosley, M. (1979). Streamflow generation in a forested watershed, New Zealand. Water Resources Research, 15: 795-806.Pearce, A., M Stewart and M. Sklash (1986). Storm runoff generation in humid headwater catchments, 1: Where does the water come from? Water Resources Research, 22: 1263-1272.

Rowe, L., A. Pearce and C. O'Loughlin (1994). Hydrology and related changes after harvesting native forest catchments and establishing Pinus radiata plantations. Part 1: Introduction to the study. Hydrological Processes, 8: 263-279.

Seibert, J. (1999). Regionalisation of parameters for a conceptual rainfall-runoff model. Agricultural and Forest Meteorology, 98-99: 279-293.

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

Sklash, M., M. Stewart and A. Pearce (1986). Storm runoff generation in humid headwater catchments, 2, A case study of hillslope and low-order stream response, Water Resources Research, 22: 1273-1282.

Figures

1. The 3-box model developed for the Maimai watershed in New Zealand including hillslope, hollow and riparian zone reservoirs.
2. Modeled outflow, storage and groundwater levels for the period September-December 1987. Measured runoff is shown by the hatched line.
3. Goodness of fit measures for runoff efficiency, groundwater levels, new water ratios, soft groundwater measures, and parameter-value acceptability for calibrations agains various combinations hard and soft information. The point is the median of all calibration trials and the lines indicate the range.

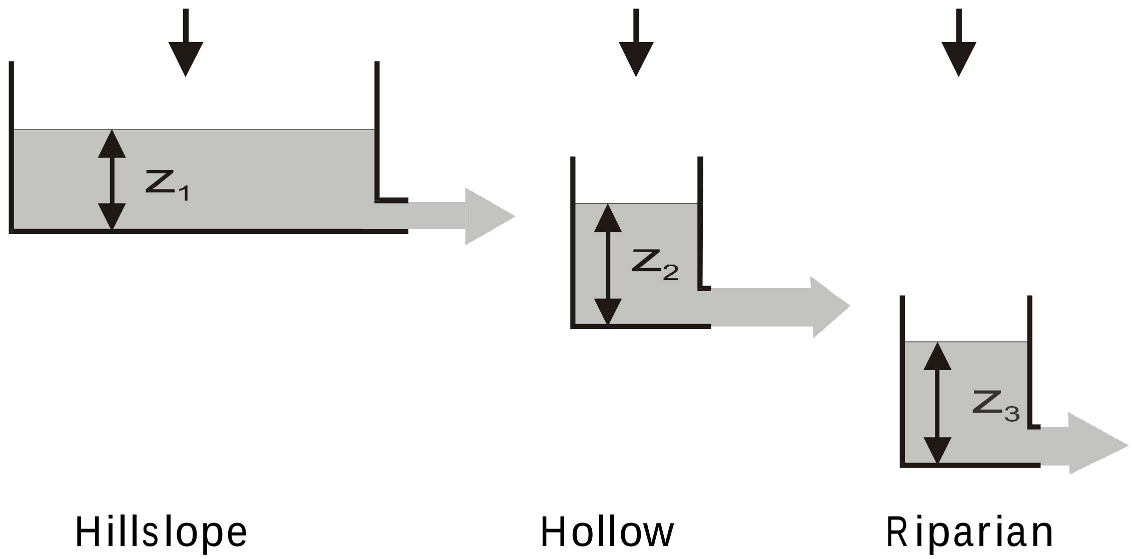
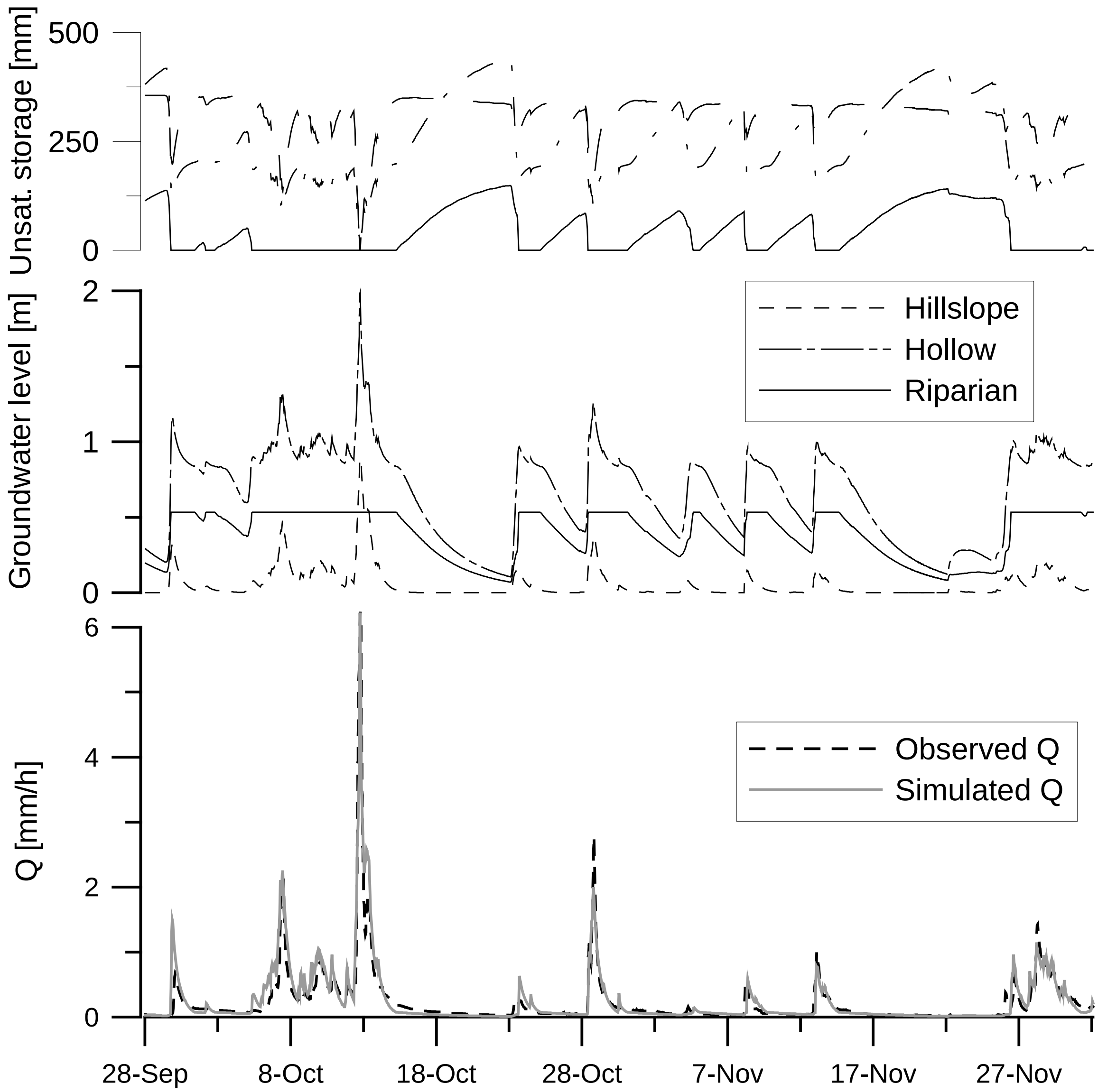
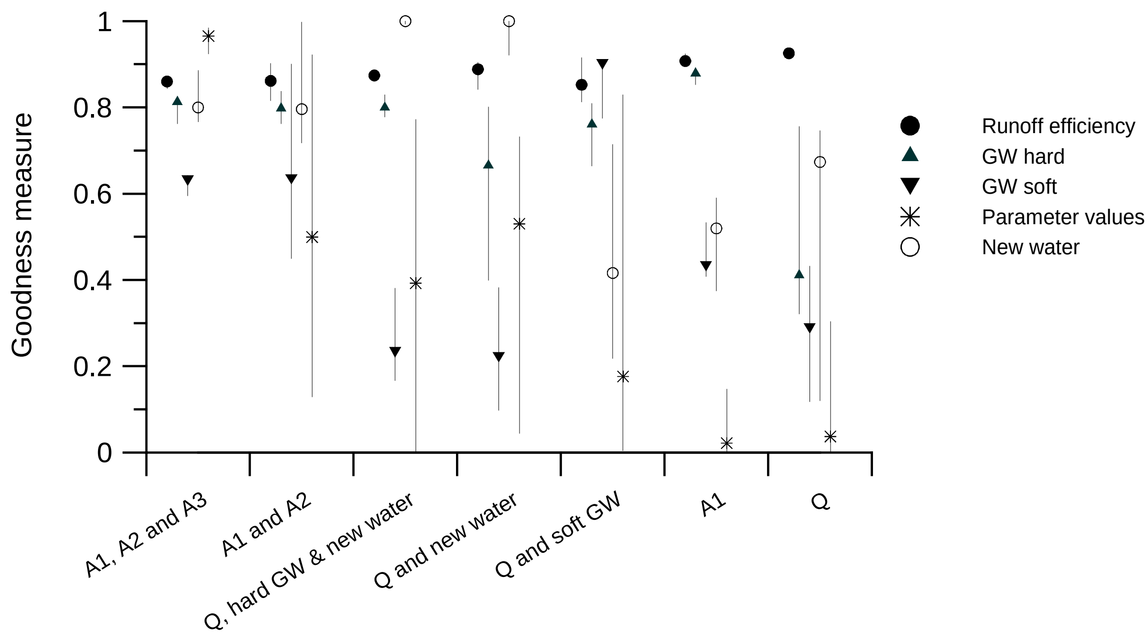


Fig. 1



500
250
0
Unsat. storage [mm]
2
1
0
Groundwater level [m]
Hillslope
Hollow
Riparian
6
4
2
0
Q [mm/h]
Observed Q
Simulated Q
28-Sep
8-Oct
18-Oct
28-Oct
7-Nov
17-Nov
27-Nov

Fig. 2



Goodness measure
1
0.8
0.6
0.4
0.2
0
A1, A2 and A3
A1 and A2
Q, hard GW & new water
Q and new water
Q and soft GW
A1
Q
Runoff efficiency
GW hard
GW soft
Parameter values
New water

Fig. 3