The role of surface–subsurface interactions in the distributed hydrological response of a headwater-dominated catchment

Ilhan Özgen-Xian[1,2,‡], Sergi Molins[1], Rachel M. Johnson[1], Ralf Loritz[3], Zexuan Xu[1], Dipankar Dwivedi[1], Utkarsh Mital[1], Craig Ulrich[1], and Carl I. Steefel[1]

[1] *Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, USA*
[2] *Institute of Geoecology, Technische Universität Braunschweig, Germany*
[3] *Institute of Water and River Basin Management, Karlsruhe Institute of Technology, Germany*
[‡] *Corresponding author at Institute of Geoecology, Technische Universität Braunschweig, Langer Kamp 19c, 38106 Brunswick, Germany. E-mail:* i.oezgen@tu-braunschweig.de

Submitted to *Hydrological Processes — Special Issue: Towards more credible models in catchment hydrology to enhance hydrological process understanding* on February 27, 2022.

Abstract

We numerically study how the interactions between surface, subsurface, and rainfall create complex runoff behavior in a headwater-dominated high elevation, mountainous catchment in East River Watershed, Colorado, USA. In order to understand the effect of surface–subsurface interactions on the hydrological response, we compare model variations with different soil permeability. We compute hydrographs, hydrological indices, and spatio-temporal distributions of hydrological variables. The indices and distributions are then linked to the hydrograph at the outlet of the domain. Our results show that changing surface–subsurface interactions trigger different flow regimes, connectivity dynamics, and runoff generation mechanisms inside the catchment, and hence, affect the distributed hydrological response. Increased surface–subsurface interaction leads to a higher degree of connectivity in the catchment. Temporal dynamics of both distributed and aggregated hydrological response did not change significantly, which may be because the structural connectivity of the catchment remained untouched. Our results indicate that in higher order catchments, the effect of changes in distributed hydrological response may not always be observable through aggregated hydrological signatures.

Keywords. Integrated hydrology; Catchment hydrology; Preferential flow paths; Hydrological connectivity; East River Watershed, Colorado

Contents

1 Introduction

Runoff formation, in the sense of runoff generation and resulting overland flow, results from the interactions among rainfall, surface topography and roughness, and soil physical and hydrological properties [Caviedes-Voullième et al., 2021]. These interactions result in preferential flow paths on the surface and in the subsurface, which route the water towards the catchment's outlet [Berkowitz and Zehe, 2020, Zehe et al., 2021] and along which spatial patterns of infiltration and exfiltration form [Mirus and Loague, 2013]. Runoff formation processes occur at small spatio-temporal scales (below the meter scale), but manifest themselves in hydrological signatures across multiple scales up to the catchment scale. The mechanisms of this *flux re-scaling* are still not completely understood [Khosh Bin Ghomash et al., 2019, Ries et al., 2017].

Runoff formation and preferential flow paths are interlinked with other environmental processes [Loague et al., 2006]. For example, the increased infiltration along the preferential flow paths creates local pressure gradients, biogeochemical hotspots [Frei and Fleckenstein, 2014], and vegetation patterns in water-limited ecosystems [Caviedes-Voullième and Hinz, 2020, and cited references within].

Further, spatial dynamics of infiltration and exfiltration, controlled by the interactions between surface and subsurface, feed back into the runoff generation and formation in the catchment itself. Thus, understanding the development of runoff formation and preferential flow paths can give insight into spatial patterns of hydrology-driven biogeochemical and ecological processes [Maxwell et al., 2019, Zhi et al., 2019].

A common approach to understand runoff formation is to analyze *aggregated hydrological responses* sensu [Ebel and Loague, 2008], such as the catchment hydrograph, with a top-down approach and look at individual contributions of each component—see, for example, [Ebel and Loague, 2008, Kirchner, 2009, Stephenson and Freeze, 1974, Stoelzle et al., 2020]. However, because different combinations of individual contributions can generate similar aggregated hydrological responses, this approach yields non-unique descriptions of the catchment system [Beven, 1993, Mueller et al., 2007]. Sometimes referred to as *equifinality* in hydrology, this poses a significant limitation to predicting hydrology-driven biogeochemical and ecological processes [Ebel and Loague, 2006, Maxwell et al., 2019].

The analysis of runoff formation can be further constrained through the inclusion of *distributed hydrological responses*, which are point or profile measurements of hydrological variables and processes inside the catchment [Ebel and Loague, 2008]. The inclusion of such spatio-temporal distributions of hydrological variables and linking them to aggregated hydrological response can overcome the aforementioned issue of equifinality. In this context, hydrological connectivity is a promising conceptual development [Ares et al., 2020, Bracken et al., 2013]. Hydrological connectivity is multiply defined, but the emerging consensus is that it describes how different parts of the catchment connect through mass fluxes. The hydrological connectivity of a catchment is typically considered to consist of two parts: (i) the structural connectivity, which is static and can be deduced by catchment characteristics and (ii) the dynamic connectivity, which is a transient and emergent property of interactions between rainfall and the catchment [Bracken et al., 2013]. Indices of hydrological connectivity—see, for example, [Antoine et al., 2011, Bracken et al., 2013, Mays et al., 2002]—aim to relate these spatio-temporal distributions to hydrological signatures. As pointed out in [Khosh Bin Ghomash et al., 2019], an ongoing effort in the study of runoff formation is to entangle the effects of interactions between topography, rainfall, and geology, in order to improve process understanding.

In this work, we aim to understand (i) to what extent surface–subsurface interactions drives changes in spatio-temporal distributions of hydrological variables and (ii) how these changes manifest themselves in the catchment's aggregated hydrograph. In other words: to what extent does the surface–subsurface interaction affect the distributed and aggregated hydrological response of a catchment? In the spirit of the concept-development simulations in [Caviedes-Voullième et al., 2021, Fiori et al., 2007, Loague et al., 2006, Mirus et al., 2011, Stephenson and Freeze, 1974], we perform simulations of runoff formation in a headwater-dominated high order catchment using an integrated hydrological model. In order to trigger changes in the surface–subsurface interactions, we

change the permeability of the subsurface layers. This has the advantage that the topography—and hence, the structural connectivity—of the catchment remain unperturbed. Thus, it isolates the surface–subsurface interaction from other potential drivers of change in the distributed hydrological response.

In the literature, modelling studies on runoff formation processes have focused on exploring (micro)topographical controls [Caviedes-Voullième et al., 2021, Dunne et al., 1991, Frei et al., 2010, Khosh Bin Ghomash et al., 2019, Thompson et al., 2010, Viero and Valipour, 2017], model complexity [Loritz et al., 2019, Mirus et al., 2011, Mueller et al., 2007], and identifying runoff generation mechanisms [Loague et al., 2010, Mirus and Loague, 2013]. Besides a handful of studies such as [Fiori et al., 2007, Fiori and Russo, 2007, Meyerhoff and Maxwell, 2011, Weill et al., 2013], process-based modelling studies on the effect of surface–subsurface interactions, especially exfiltration, are rare. In order to understand the effect of surface–subsurface interactions on runoff formation, we compare distributed model results obtained by model variations with different surface–subsurface interactions and link them to catchment hydrograph dynamics. We further compute and compare the number of disconnected water clusters in the catchment, which is an index of hydrological connectivity [Khosh Bin Ghomash et al., 2019, Weill et al., 2013], and the spatial distribution of runoff generation mechanisms in the catchment from different model variations.

2 Materials and Methods

2.1 Study site

The East River Watershed is a mountainous headwater catchment in the Upper Colorado River Basin, see Fig. 1. As one of two major contributors to the Gunnison River, the East River Watershed is a crucial source of water for the western USA [Hubbard et al., 2018]. The majority of this water is stored as snowpack and released over the year. Our study site is the Lower Triangle (LT), a headwater-dominated, 15 km^2 subcatchment of the East River Watershed, see the LT labeled region in Fig. 1.

The geology of the LT is characterized by marine shale of the Mancos formation, with Cenozoic igneous formations intruding the Mancos formation in the western part and the Palaezoic and Mesozoic sedimentary strata in the eastern part, where it forms steeply dipping beds [Carroll et al., 2018].

The elevation of the LT ranges from 2759 m at its south-western part to 3787 m at its northern part. The lower parts of LT feature a meandering riparian corridor. The stream is fed by high-energy mountain streams from subcatchments on the LT's north, namely Copper Creek, Rustlers Gulch, Quigley Creek (EAQ), and the Middle East River (ME) catchment [Maavara et al., 2021]. The LT is one of the U.S. Department of Energy's intensively monitored sites to understand the impact of perturbations on watershed systems [Hubbard et al., 2018]. A pumphouse (PH), located at the outlet of the LT, extracts water for the municipality of Mt. Crested Butte, Colorado, USA. Daily discharge mea-

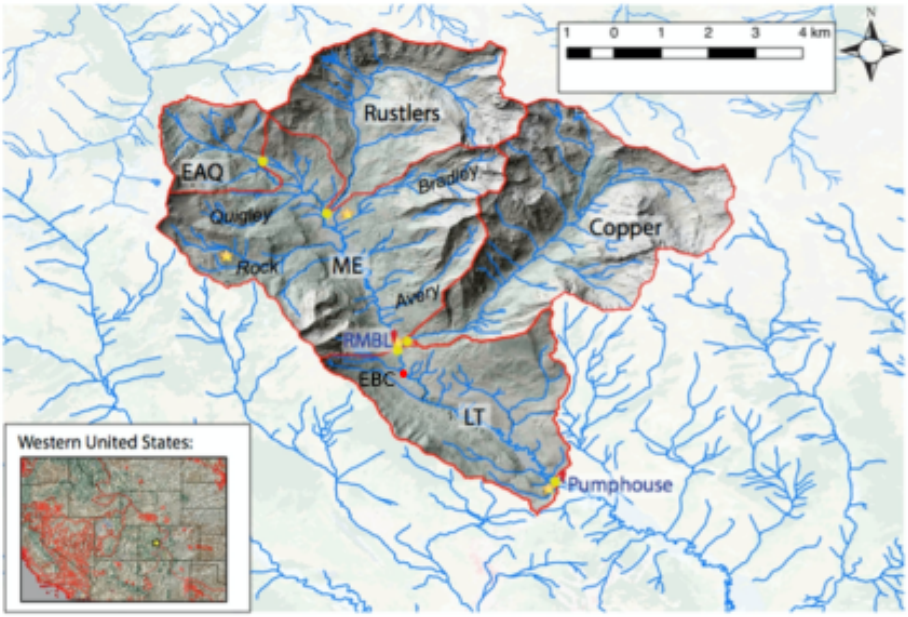

Figure 1: Map of East River Watershed, Colorado, with sub-watersheds (red boundaries), reproduced with permission from [Maavara et al., 2021]. The study region Lower Triangle (LT) is located at the downstream of the watershed. At the outlet of the LT, a pumphouse is located where discharge data is measured daily. The red point labeled with EBC denotes the location of discharge measurements that were used as inflow boundary condition.

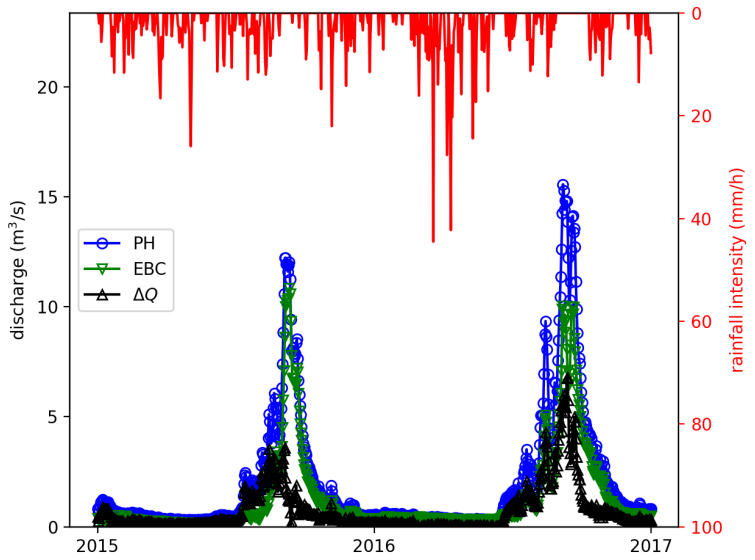

Figure 2: Inflow at the northern boundary (EBC), the discharge measured at the Pumphouse (PH), and difference between the inflow and the outflow, calculated as $\Delta Q = Q_{\mathrm{PH}} - Q_{\mathrm{EBC}}$. The majority of contributed discharge occurs during peak flow.

surements are provided at this location. In the northern part of the domain, close to the catchment boundary, the EBC station measures daily discharges in the river, which we use as the inflow hydrograph at the northern boundary of the LT in this work.

The hydrology in the region is snow-dominated, due to the continental subarctic climate with long, cold winters and short, cool summers. The annual precipitation ranges between 670 and 1200 mm, where up to 70% is in the form of snow and the majority of the remaining precipitation is in the form of monsoonal rains [Carroll et al., 2018, Carroll et al., 2020]. Snowfall occurs from October through May and monsoonal rains occur from July through September. Snowmelt starts in April and continues through the year. The discharge at the PH station usually peaks in June to July, when snowmelt from the northern subcatchments arrives at the outlet [Carroll et al., 2019], see also Fig. 2.

Fig. 2 shows the inflow and outflow hydrographs as well as the difference between these hydrographs for the water years 2015 and 2016. In this time interval, the LT contributes about 36% of the total discharge volume at the PH, while the rest is headwater. While the LT contribution to the outflow hydrograph is significant at some times during the high flow periods, it is diminished significantly during the low flow periods. This results in a peculiar situation, where changes in spatial flow dynamics may not always be observable from the

hydrograph alone.

2.2 Integrated hydrological solver

We use the physically-based integrated hydrological model *The Advanced Terrestrial Simulator* (ATS) [Coon et al., 2020] to compute the flow and the spatial patterns of different state variables—for example, snow depth, and soil moisture. ATS solves the two-dimensional zero-inertia equation for overland flow and three-dimensional Richards equation for subsurface flow in a fully coupled manner, using an implicit finite-volume solver. Surface and subsurface flow are coupled with a pressure continuity assumption across the surface and subsurface interface. Further processes considered in this study are snow storage and evapotranspiration. Snow storage is modeled through a temperature-based approach and evapotranspiration is modeled using the Priestley-Taylor equation [Priestley and Taylor, 1972].

2.3 Study setup

2.3.1 Model of the Lower Triangle

Computational mesh The domain is discretized using a multiresolution, unstructured triangular mesh that is extruded in the vertical. The cell size on the surface varies between 10 and 160 m. The resolution is decided through a wavelet analysis-based meshing strategy [Özgen-Xian et al., 2020]. This meshing strategy can use different data as input, such as topography or slopes. Using different data results in different refinement patterns. In this study, we use a pre-computed flow field—obtained by analyzing the topography data with the algorithm in [Tarboton, 1997]—as input data for the wavelet analysis. This results in refinement around streams and preferential flow paths in the domain. The surface mesh is extruded vertically with increasing mesh spacing, which ranges from 0.1 m at the top to 30 m at the bottom of the domain, spanning 100 m in total. The resulting mesh—shown in Fig. 3 (left)—contains 894,740 cells.

Geology data In contrast to previous model studies in the region that extrapolated the geology from USGS geology maps, the geological layer information of the present model is based on the results of a forward simulation by a three-dimensional geological model (C. Ulrich, LBNL, unpublished). The geological model simulation is constrained by the aforementioned USGS geology maps. The model results provide a complete high-resolution description of the geology in the domain, including heterogeneity that cannot be obtained by extrapolating the USGS geology maps. The geology predicted by this model is mapped onto the computational mesh using a nearest neighbor interpolation, see Fig. 3 (right). The weathered shale is layered on top of fractured shale and sand. Localized quartz intrusions pinch through the weathered and fractured shale.

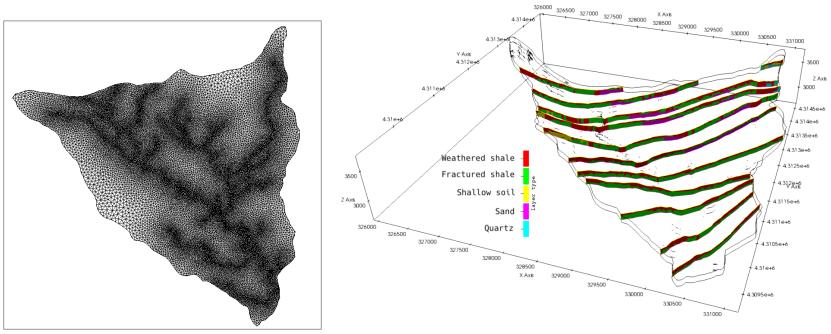

Figure 3: Top view of the computational mesh (left) and a three-dimensional view of selected cross sections showing the geological layer structure in the LT as interpolated onto the computational mesh (right)

Boundary conditions At the north boundary of the domain, we impose a discharge boundary condition to account for the streams entering the domain from this boundary. The boundary condition is informed by measurement data in the stream at the station EBC, which is located close to the north boundary. The imposed hydrograph is plotted in Fig. 2. At the downstream boundary, a free outflow boundary condition is set. Precipitation is applied using the Daymet dataset [Thornton et al., 2020]. The precipitation time series is plotted in Fig. 2. We use spatially uniform precipitation.

Initial conditions Initial conditions for the subsequent simulations were generated by a so-called "spin-up" simulation. Starting with initially saturated soil and a dry surface, the domain was drained for 10 years. Afterwards, a constant precipitation with an intensity of 10^{-8} ms^{-1}, which corresponds to the annual mean, was applied for 5 years. Finally, a 1 year simulation using the precipitation data for the water year 2015 was carried out to obtain the initial conditions.

2.3.2 Model calibration

The model is not calibrated in the traditional sense. The van Genuchten soil parameters were set equal to the values reported in [Xu et al., 2022]. The saturated permeabilities were set based on the values reported in [Tokunaga et al., 2019], where permeability values at a hillslope in the region are quantified through field observations. These were set individually for each geological group. We summarize these values in Tab. 1. The surface rougness is accounted for by Manning's law. The Manning coefficient is set to 1.0 sm$^{-1/3}$ throughout the domain and was not used for calibration. Thus, all model parameters were derived from available data sets and were not used for calibration. The model performance is assessed by comparing model results with the measured daily

Geology	Permeability (m^2)	α (Pa^{-1})	m	θ_s	θ_r
Shallow soil	9.9×10^{-13}	0.0003	0.6	0.53	0.2
Weathered shale	1.13×10^{-13}	0.0004	0.5	0.5	0.1
Fractured shale	1.6×10^{-14}	0.0004	0.5	0.5	0.1
Quartz	1×10^{-15}	0.0003	0.4	0.67	0.1
Sand	2.26×10^{-12}	0.0002	0.6	0.2	0.1

Table 1: Calibrated permeability and van Genuchten soil parameters

Geology	P−	P0	P+
Shallow soil	1.98×10^{-13}	9.9×10^{-13}	4.95×10^{-13}
Weathered shale	2.26×10^{-14}	1.13×10^{-13}	5.65×10^{-13}
Fractured shale	3.2×10^{-15}	1.6×10^{-14}	8×10^{-14}
Quartz	2×10^{-16}	1×10^{-15}	5×10^{-15}
Sand	4.52×10^{-13}	2.26×10^{-12}	1.13×10^{-11}

Table 2: Permeability values (m^2) of the investigated model variations

discharge at the PH. Here, we used the hydrographs for the water years 2015 and 2016.

2.3.3 Model variations

As discussed in the introduction, we change the permeability of the subsurface layers as a convenient way to trigger changes in the surface–subsurface interactions, without changing other catchment characteristics. Thus, changes in the hydrological response can be directly attributed to the change in the surface–subsurface interactions. Hereinafter, the model variation P0 corresponds to the calibrated model. For the model variation P−, all permeability values are reduced by a factor of 5, while for the model variation P+, all permeability values are increased by a factor of 5. The resulting permeability values are summarized in Tab. 2. We note that the reduction and increase by the chosen factor of 5 is rather small when compared to the range of permeability values typically observed in the field. However, as we will see in the following section, even such small changes can yield significantly different runoff behavior.

3 Results and discussion

3.1 Spatio-temporal distributions

We quantitatively assess the interaction between surface and subsurface through the exchange mass flux across the surface–subsurface interface. Here, negative values of the exchange flux imply exfiltration and positive values imply infiltration. Based on the sign of the exchange flux, spatial patterns of infiltration and

exfiltration can be deduced. Fig. 4 showcases the differences in these spatial patterns obtained from different model variations at selected time steps. In all model variations, water exfiltrates in the riparian areas, especially in the downstream parts, which has flat topograpy and saturated soils—see Fig. 5. This is in agreement with our conceptual understanding of the hydrological cycle, where rainfall infiltrates in the upslope areas, is routed through the subsurface, and exfiltrates in the riparian areas and river channel—see, for example, [Freeze and Cherry, 1979]—and the modelling results in [Weill et al., 2013]. During low flow conditions ($T = 50$ d and $T = 500$ d, compare also the hydrograph in Fig. 2), infiltration is observed in the upstream parts of the river channel, because the soil water in the top layers is drained by subsurface flow. As the hydrograph begins to rise, the soil in the river channel gets quickly saturated and switches from infiltration to exfiltration. Here, we assume that the simulated time scales are small, such that the influence of deep groundwater is negligible.

As expected, surface–subsurface interaction increases as the permeability is increased. This has significant influence on the spatio-temporal distributions of the exchange flux. As surface–subsurface interaction increases, patterns of infiltration and exfiltration appear along ephemeral streams in the northeast of the catchment. Thus, increased surface–subsurface interaction leads to a higher degree of connectivity inside the catchment.

The degree of surface hydrological connectivity, which is relevant to the runoff formation, can be visualized through the ponding water depth distribution. Spatial distributions of ponding water depths can be thresholded into binary wet/dry distributions, which are visually easier to interpret. Here, we use a threshold of 0.01 m as a threshold to separate between wet and dry, considering locations with water depths lower than this threshold as dry. Fig. 6 shows such wet/dry distributions for all model variations at selected times. In [Thompson et al., 2010], four separate stages of flow regimes are identified in the context of microtopography at the hillslope scale, which can be transferred to the catchment scale to categorize the flow regimes. These flow regime stages are namely (i) local flow, where water flows mostly in disconnected depressions and fills them, (ii) channel flow, where local streams start to develop between filled depressions, (iii) mixed flow, where filled depressions start to merge as the topography gets inundated, with higher parts remaining dry, and finally, (iv) sheet flow, where all depressions are filled and the surface is completely inundated.

The spin-up simulations ensure that at the initial state, the flow regime at the riparian areas reaches sheet flow. In the P− model variation, upslope areas showcase local flow regimes with disconnected puddles that showcase a mixture of in- and exfiltration. P0 model variation with increased surface–subsurface interaction results in localized channel flow regimes in the upslope areas, in combination with local flow regime in parts. The wetland area is extended and in addition to sheet flow showcases mixed flow regimes at its edges. Further increase in surface–subsurface interaction leads to a connection between depressions and leads to a mixed flow regime in the upslope areas, that are then partially connected to the riparian areas through one-dimensional

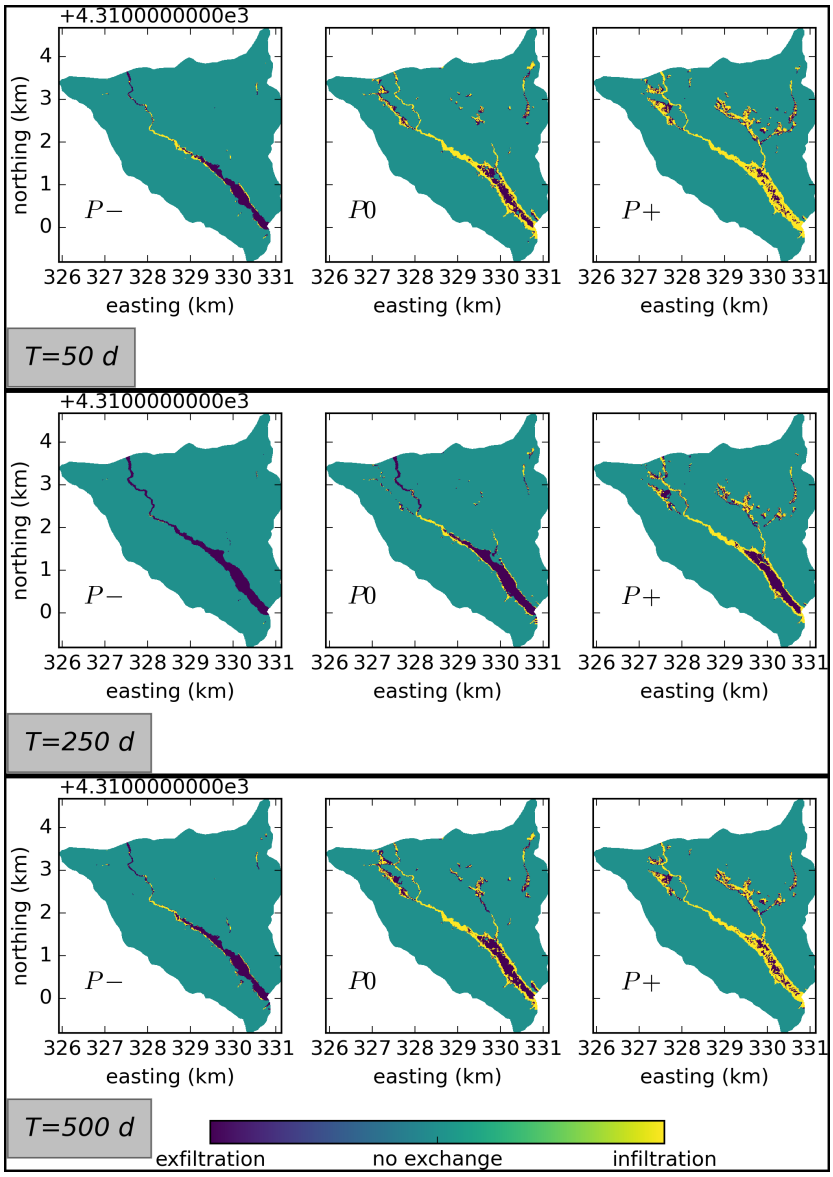

Figure 4: Snapshots of the spatial patterns of infiltration and exfiltration at the surface. $T = 50$ d: low flow conditions in the first water year, $T = 250$ d: rising limb in the first water year, $T = 500$ d: low flow conditions in the second water year, shortly before the rising limb.

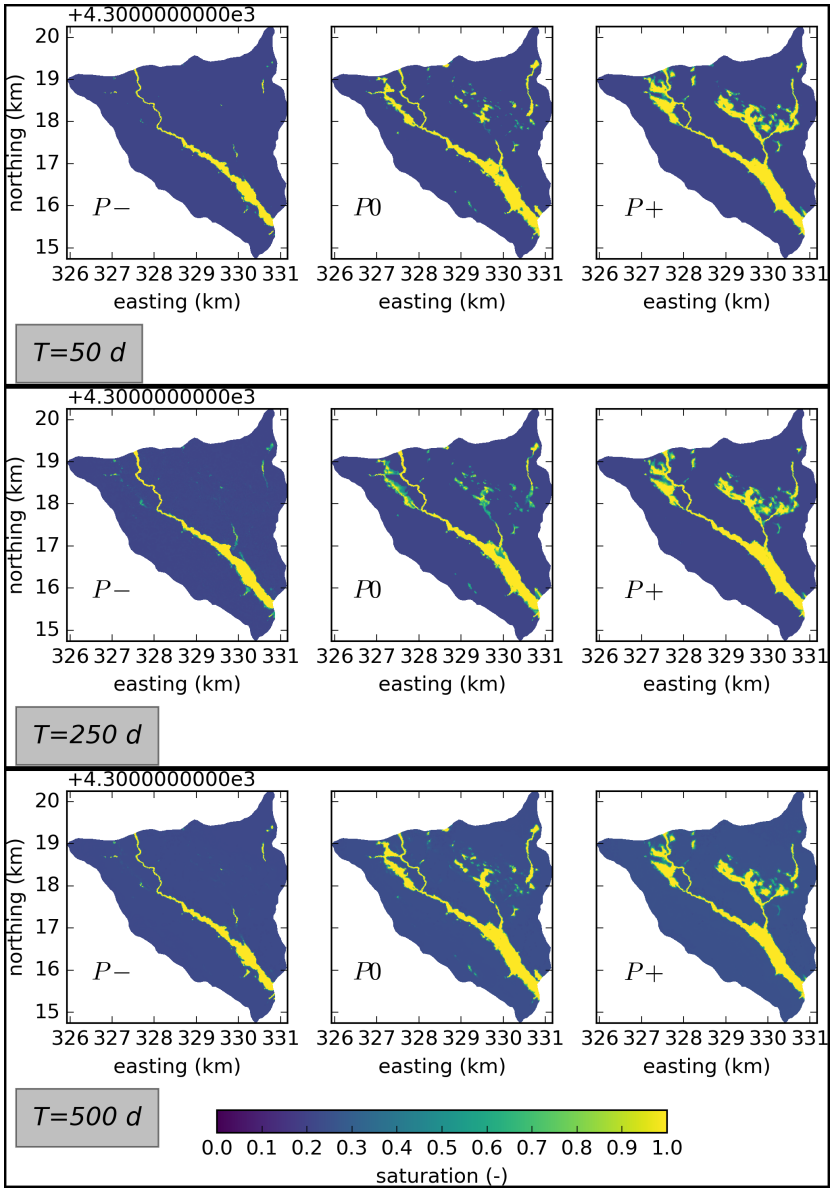

Figure 5: Snapshots of the spatial patterns of soil saturation in the top layer. $T = 50$ d: low flow conditions in the first water year, $T = 250$ d: rising limb in the first water year, $T = 500$ d: low flow conditions in the second water year, shortly before the rising limb.

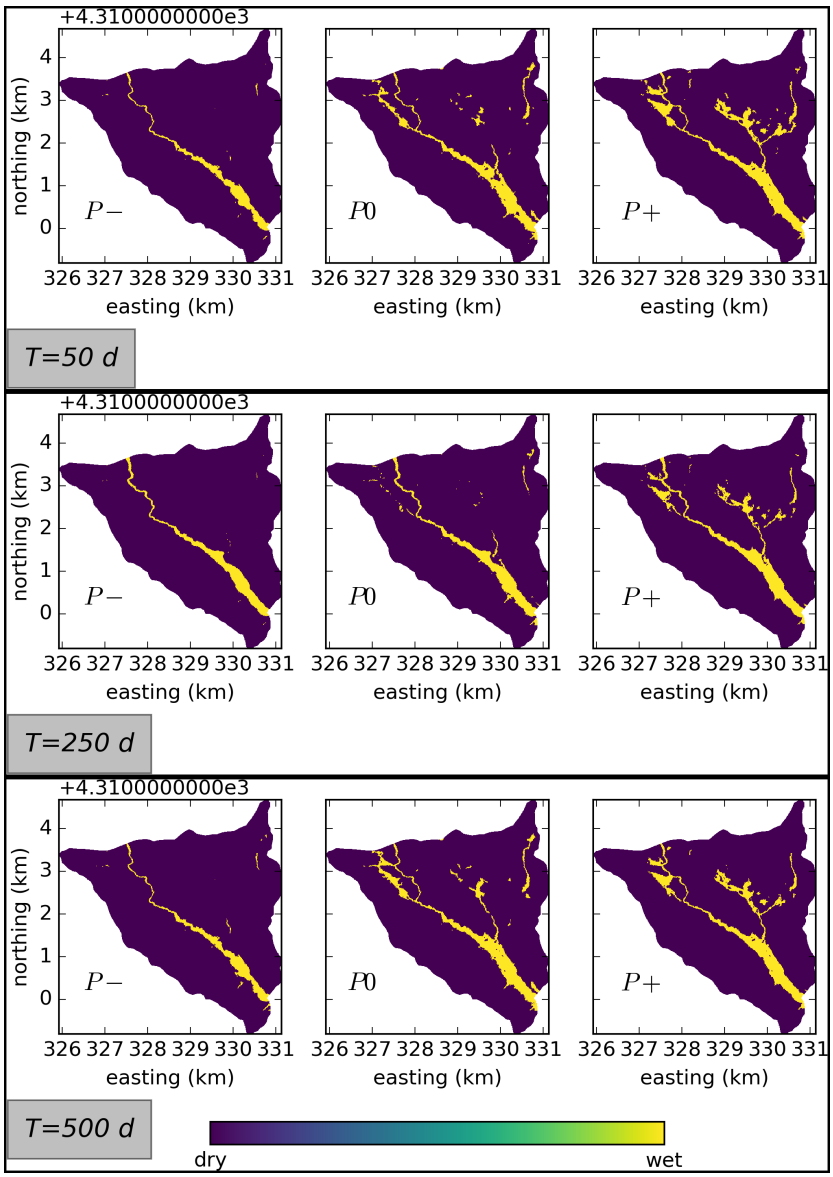

Figure 6: Snapshots of the spatial patterns of wetting and drying at the surface. $T = 50$ d: low flow conditions in the first water year, $T = 250$ d: rising limb in the first water year, $T = 500$ d: low flow conditions in the second water year, shortly before the rising limb.

channels. The riparian areas are dominated by sheet flow. Temporally speaking, the low flow conditions ($T = 50$ d and $T = 500$ d) showcase a higher degree of connectivity between the depressions than the conditions at the rising limb ($T = 250$ d), where the ephemeral streams in the upslope areas seem to dry out, reverting the flow regime from channel flow to local flow in these areas. Comparison between $T = 50$ d and $T = 500$ d reveals that the catchment is not fully drained, some of the initially dry localized ponds and ephemeral streams that appear on the rising limb of the hydrograph ($T = 250$ d) persist during the low flow conditions.

Comparing spatial distributions of permeability with soil saturation reveals that permeability distribution does not significantly affect spatial distributions of soil saturation. In all model variations throughout the subsurface, the Pearson correlation coefficient is the range of -0.01 at the top to -0.13 at the bottom of the subsurface, indicating no significant linear relationship between these two distributions. Meanwhile, the Spearman correlation coefficient is in the range of -0.1 at the top to -0.4 at the bottom of the subsurface, indicating that no significant monotonicity can be detected in the relationship, that is to say, locally higher permeability does not lead to locally higher saturation. This indicates that in our case, the topography and surface hydrology are the dominant control on the spatio-temporal distribution of soil saturation throughout the subsurface. In general, the correlations are slightly worse for the P$-$ model variation, where the surface–subsurface interaction is decreased. The calculation of these correlation coefficients is given in App. A.

3.2 Aggregated hydrological response

We have observed that the different model variations result in different spatial distributions of hydrological variables. We will now proceed to analyze whether these different spatial distributions can be observed in the hydrographs at the outlet of the catchment. Fig. 7 compares the computed hydrographs of all model variations and the observed hydrograph (PH). Qualitatively, the hydrographs from all model variatons showcase similar dynamics and accurately capture the peaks. Indeed, the hydrograph produced by P0 is difficult to distinguish visually in this figure. The reason for this might be that during peak flow conditions the soil below the flow paths in the riparian areas quickly becomes saturated, which increases the dependency of the hydrograph on the headwater. This is partially supported by the fact that smaller fluctuations on the rising limb of the first peak are not reproduced, which is one of the rare times when the LT's contribution to the hydrograph exceeds the contribution of the headwater catchments significantly—see Fig. 2. This is similar to the findings in [Caviedes-Voullième et al., 2021], where it is observed that for high rainfall intensities, the effect of topography on the hydrograph is reduced, because the surface storage is saturated.

Discharge volume The volume of the discharge is calculated as the integral of the hydrograph, see Tab. 3. All discharge volumes are in the same order

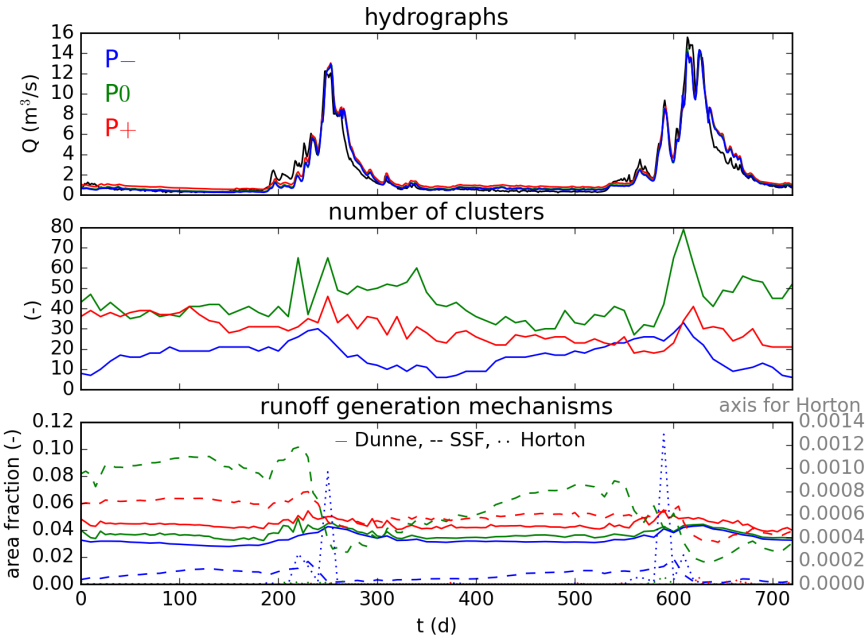

Figure 7: *Top*: Model discharge predictions at PH for different model variations. The observed discharge at PH is plotted with a black continuous line. *Middle*: Number of disconnected wet clusters. *Bottom*: Runoff generation mechanism dynamics in the catchment. The figure plots the fraction of the catchment area that showcases a certain runoff generation mechanism over time.

	Volume (m^3)	$\Delta V/V_o$ (%)	NSE	KGE
Observation	$125,983,524$	-	-	-
P+	$142,524,628$	+13.1	0.94	0.86
P0	$129,078,697$	+2.5	0.95	0.96
P-	$124,933,050$	−0.8	0.95	0.97

Table 3: Summary of discharge volumes, NSE and KGE for the observed discharge data and the simulations. ΔV denotes the difference in discharge volumes with respect to the observed discharge volume and V_o is the observed discharge volume.

of magnitude. The volume of the hydrograph belonging to P+ model is the largest, the difference is about 13% of the observed discharge volume. This can be attributed to the increased base flow accumulating throughout the years. The best match for the discharge volume is achieved by the P− model, but the P0 model's discharge volume does not deviate significantly.

Coefficients of efficiency The Nash-Sutcliffe [Nash and Sutcliffe, 1970] and Kling-Gupta [Gupta et al., 2009] efficiency coefficients are calculated for each simulated hydrograph with respect to the measured time series. In general, higher values of both of these coefficients are considered to suggest a better agreement between simulated and observed hydrographs. The calculation of these coefficients is shown in Appendix B.

The computed coefficients for all model variations are summarized in Tab. 3. The deviation in discharge volumes is reflected in both the NSE and KGE coefficients. While all hydrographs have very similar NSE values, the P+ model variation with the highest deviation in the discharge volume yields the lowest NSE. Nevertheless, the difference might be considered negligible, even though the base flow in this model variation is overestimated significantly. This is a known limitation of the NSE, which tends to place more emphasis on the peaks of the hydrograph [Knoben et al., 2019]. The KGE takes the deviation in base flow into account in a more balanced way. Consequently, the difference in the discharge volumes is better reflected in the corresponding KGE values. P+ has the lowest KGE value and the difference is clearly distinguishable. P− has the highest KGE value, followed by P0 by a small margin.

Summary Overall, we see that the change in the distributed hydrological response triggered by a change in surface–subsurface interactions is reflected in the difference in discharge volumes as well as the KGE. The NSE is not sufficient to detect these changes. This suggests that the most significant impact of surface–subsurface interactions on the aggregated hydrograph is the rainfall partitioning into overland flow and subsurface flow during low flow conditions, which is better reflected in the discharge volumes and the KGE. Arguably, the significant differences in spatio-temporal distributions between the P− and the P0 model variations is not directly visible in any of the aggregated indices we

used.

3.3 Linking hydrological connectivity to hydrological response

Surface connectivity The binary wet and dry distribution in Fig. 6 showcases disconnected clusters of wet cells. As the flow regime shifts from local flow towards sheet flow these clusters connect. Consequently, a decrease in the number of disconnected clusters implies an increase in the degree of connectivity and vice versa. The evolution of the number of these disconnected clusters is plotted in Fig. 7 (middle). The rainfall does not affect the connectivity in the catchment, which is in contrast to the findings in [Han et al., 2020], where rainfall characteristics were found to dominate connectivity in a zero order catchment. This difference is most likely due to the LT being headwater dominated. The connectivity of the catchment decreases during the rising limb of the hydrographs and gradually increases as the hydrographs decline. This is roughly the behavior reported in [Khosh Bin Ghomash et al., 2019], where the decrease in connectivity is related to the microtopography getting partially inundated. Similarly, in the LT the wetland area extends as the hydrograph rises, causing local flow with disconnected clusters of water at its edges. Increasing the surface–subsurface interactions alters the connectivity in the catchment but the relationship is non-linear. The P0 model variation results in the least degree of connectivity, and increasing and decreasing the surface–subsurface interaction both results in an increased connectivity.

Fig. 8 shows a hysteretic relationship between the number of disconnected wet clusters plotted and the discharge at the outlet. As the hydrograph rises, the number of disconnected wet clusters increases and thus, the connectivity decreases, reaching its minimum on the receding limbs of the hydrograph. We observe a thresholding behavior at about 2 m^3/s, after which the surface connectivity does not decrease significantly. However, in the second water year, the P0 model variation leads to a rapid disconnection of the catchment during the receding limb of the hydrograph, leading to pronounced new maxima in the disconnected wet cluster number.

Subsurface connectivity Similar to the thresholding to create wet/dry distributions, the subsurface saturation can be thresholded into a binary saturated/unsaturated distribution. Similar to the wet clusters on the surface, such saturated cluster distributions can be used to derive metrics of subsurface connectivity [Mays et al., 2002, Weill et al., 2013]. Fig. 9 shows the number of disconnected saturated clusters at the top subsurface layer and its relation to the discharge at the outlet of the domain. The subsurface connectivity follows the same dynamics as the surface connectivity. The subsurface displays a higher degree of connectivity during low flow conditions and disconnects as the hydrograph rises and falls. This might be related to the widening of the wetland area, which inundates previously dry areas and causes localized saturated clusters.

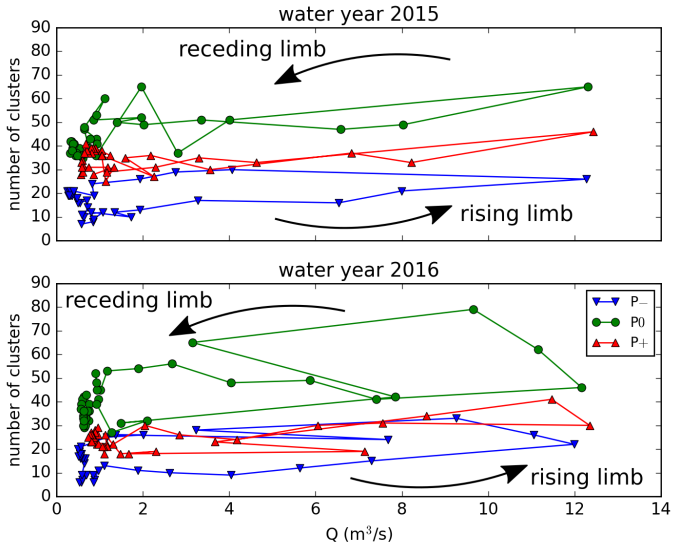

Figure 8: Relationship between number of disconnected wet clusters and discharge at the outlet for both simulated water years.

We identify the same threshold of 2 m^3/s, after which the connectivity for most model variations does not decrease significantly. This kind of thresholding behavior is also reported in [Weill et al., 2013]. However, the dynamics in our case are the opposite of what they report in [Weill et al., 2013]: the highest degree of connectivity is reached during peak flow.

Summary Both surface and subsurface connectivity dynamics follow the same pattern. High connectivity during low flow conditions, gradual disconnectivity until a threshold whereinafter either no significant change or rapid increase in disconnectivity is observed. The thresholding behavior is consistent with [Weill et al., 2013], but the relationship between connectivity and discharge is the opposite. The catchment in [Weill et al., 2013] is a headwater catchment, while our study considers a higher order catchment. This might be the reason for this difference. The differences in the connectivity dynamics between the model variations do not manifest in the hydrographs.

3.4 Linking runoff generation mechanisms to hydrological response

Runoff generation mechanisms are commonly classified as (i) saturation excess (Dunne-type) runoff, (ii) infiltration excess (Hortonian) runoff, and (iii) subsurface storm flow [Dunne, 1978, Mirus and Loague, 2013, Loague et al., 2010].

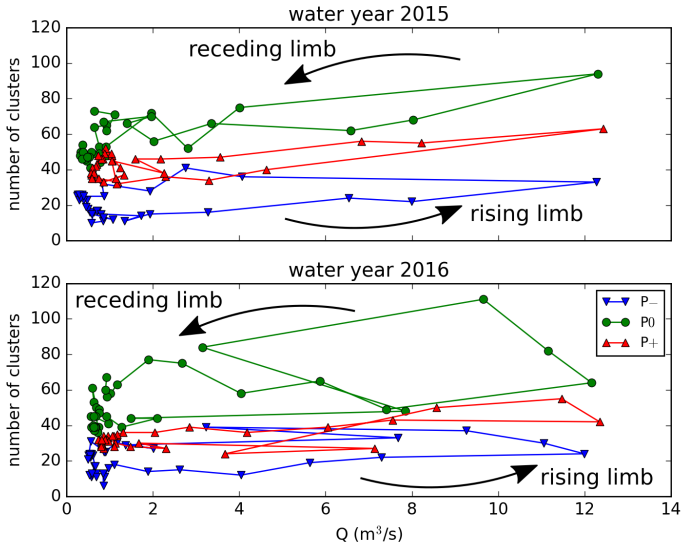

Figure 9: Relationship between number of disconnected saturated clusters and discharge at the outlet for both simulated water years.

Using the spatial distributions discussed above, we classify the runoff generation mechanism in a spatially distributed manner. Starting from the binary wet/dry distribution (Fig. 6), we decide for each wet cell, whether the runoff is of Dunne-type or Hortonian. If the underlying subsurface is fully saturated, the runoff is classified as Dunne-type and otherwise it is classified as Hortonian. For dry cells, we classify the runoff as subsurface stormflow if the underlying subsurface features flow.

Fig. 7 (bottom) compares the temporal evolution of the areal fraction controlled by different runoff generation mechanisms for all model variations. Dunne flow stays nearly constant during low flow conditions and slightly increases during the hydrographs's peaks. All model variations result in similar Dunne flow, especially the P− and P0 model variations. The subsurface stormflow shows more variation across the model variations. For all model variations, the subsurface stormflow starts high, rises slightly during the hydrograph's first rising limb and suddenly drops during the hydrograph's receding limb. During the low flow period, subsurface flow gradually increases until it drops again during the hydrograph's second rising limb. The highest subsurface stormflow results from the P0 model, followed by the P+ model. The lowest subsurface stormflow results from the P− model. Hortonian flow is only observed during both peaks of the hydrograph. P− model yields the highest Hortonian flow by orders of magnitude. Overall, P0 and P− models result in similar hydrographs, but result from different combinations of runoff generation mechanisms. The

subsurface stormflow is inversely related to the surface connectivity in the domain. This is reasonable, since increased subsurface stormflow implies decreased surface runoff, which results in less ponding on the surface.

4 Conclusions and implications

Surface–subsurface interactions cause complex spatial runoff and connectivity dynamics at the catchment scale, which may not be observable through aggregated hydrological signatures. Indeed, changing surface–subsurface interactions change the catchment behavior—as observed both in the connectivity and the localized runoff generation mechanisms—in a non-linear manner that still result in similar hydrographs. Here, the surface topography was identified as the dominant control on spatio-temporal distributions of hydrological variables both at the surface and in the subsurface.

In headwater-dominated catchments, the effects of spatiotemporal distributions on the catchment's hydrograph may be overshadowed by the headwater dynamics. Thus, we expect that as the catchment order increases, the effect of changes in spatiotemporal distributions on aggregated signatures diminishes. However, localized high intensity rainfall events may amplify the effect of localized runoff dynamics on the hydrograph. The effect of spatially heterogeneous precipitation on the hydrological response of headwater-dominated catchments remains to be investigated.

In this study, the temporal evolution of connectivity in the headwater-dominated high order catchment is the opposite of what has been reported for a headwater catchment in [Weill et al., 2013]. This may be related to the widening of the riparian areas due to increased headwater inflow during peak flow conditions, which inundate previously dry areas and create disconnected local flow. This widening of the riparian areas may not be as significant in headwater catchments, where the emergence of ephemeral streams connecting disconnected clusters dominates the connectivity. However, further studies are necessary to understand these contrasting catchment responses. The inverted temporal evolution implies that geochemical hotspots associated with hydrological (dis)connectivity are different in high order catchments compared to headwater catchments.

Differences in the rising limbs of the hydrographs from different model variations, which have been reported for microtopographical controls on runoff [Caviedes-Voullième et al., 2021], have not been observed. This indicates that the rising limb of the hydrograph is strongly related to the structural connectivity, which is changed by varying microtopography but has been kept constant in this study. This is also reported in [Khosh Bin Ghomash et al., 2019].

From a modelling perspective, the aggregated hydrograph is not very sensitive to the permeability, except for the base flow. In terms of hydrograph analysis, the KGE is found to detect the discrepancy in the base flow better than the NSE, which is in agreement with the literature [Knoben et al., 2019]. This concept of equifinality has been known for long time in the hy-

drology literature—see, for example [Beven, 1993]. Depending on the aim of the study, it may be seen as beneficial that changes in connectivity do not affect the aggregated hydrograph, however, the correct representation of connectivity may be significant when modelling morphodynamics, ecological processes, or (bio)geochemistry [Ackerer et al., 2020, Caviedes-Voullième and Hinz, 2020, Sivapalan, 2009, Refsgaard et al., 2022, Zhi et al., 2019]. A promising development in hydrograph analysis is the use of water isotope-based modeling approaches [Gillefalk et al., 2021, Smith et al., 2021, Sterte et al., 2021], which allow constructing a more precise relationship between spatiotemporal distributions and the hydrograph that can be directly supported by field measurements. The different initial conditions that result from the individual spin-up simulations carried out for each model variation amplifies the non-linearity in the surface–subsurface interactions, highlighting the sensitivity of initial conditions on model results [Seck et al., 2015].

A limitation of this study is that only three model variations with a relatively narrow permeability range have been compared. In order to gain a more complete understanding of the effect of surface–subsurface interactions on the hydrologial response, a more systematic study spanning a wider range of permeability values is needed to detect possible thresholding behavior in the catchment's response. Further, permeability values were varied in bulk, which does not result in new patterns of heterogeneity. The effect of changing the permeability of individual layers was not explored. Another limitation is the spatially uniform precipitation that has been used for the studies, which does not allow the study of localized rainfall events.

Future research may focus on the reversed temporal evolution of hydrological connectivity displayed by the high order catchment in this study. It would be desirable to see if this behavior is generalizable to other high order catchments. Another future study may test the findings of this study for a wider range of permeability values, including changing the permeability of a single geological layer to create localized heterogeneity.

Acknowledgments

This work is supported as part of the Watershed Function Scientific Focus Area funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-AC02-05CH11231. This work used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy, Office of Science, user facility operated under Contract Number DE-AC02-05CH11231. This work used the Lawrencium computational cluster resource provided by the IT Division at the Lawrence Berkeley National Laboratory, supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract Number DE-AC02-05CH11231. The geology of the LT was modeled using Emerson-Paradigm's SKUA-GOCAD.

A Pearson and Spearman correlation

Pearson's correlation coefficient (r) between two data sets x and y is calculated as

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_i - \bar{y})^2}}, \tag{1}$$

where x_i and y_i denote the i-th entry of the data set x and y, respectively, and $\bar{x}$ and $\bar{y}$ denote the mean value of the data set x and y, respectively.

Spearman's correlation coefficient (ρ) between two data sets x and y is calculated as

$$\rho = 1 - \frac{6\sum_i (x_i - y_i)^2}{n(n^2 - 1)}, \tag{2}$$

where n is the size of the data sets.

B Nash-Sutcliffe and Kling-Gupta efficiency

The Nash-Sutcliffe efficiency (NSE) is calculated through

$$\mathrm{NSE} = 1 - \frac{\sum_t (Q_s(t) - Q_o(t))^2}{\sum_t \left(Q_o(t) - \bar{Q}_o(t)\right)^2}, \tag{3}$$

where t is the time, Q_s is the simulated discharge, Q_o is the observed discharge, and $\bar{Q}_o$ is the mean observed discharge.

The Kling-Gupta efficiency KGE is calculated using statistical moments as

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_s}{\sigma_o} - 1\right)^2 + \left(\frac{\bar{Q}_s}{\bar{Q}_o} - 1\right)^2}, \tag{4}$$

where r is the linear correlation between observation and simulation data, σ_s and σ_o are the standard deviation in simulation and observation data, respectively, and $\bar{Q}_s$ is the mean the simulation discharge.

References

[Ackerer et al., 2020] Ackerer, J., Jeannot, B., Delay, F., Weill, S., Lucas, Y., Fritz, B., Viville, D., and Chabaux, F. (2020). Crossing hydrological and geochemical modeling to understand the spatiotemporal variability of water chemistry in a headwater catchment (Strengbach, France). *Hydrology and Earth System Sciences*, 24:3111–3133.

[Antoine et al., 2011] Antoine, M., Javaux, M., and Bielders, C. L. (2011). Integrating subgrid connectivity properties of the micro-topography in distributed runoff models, at the interrill scale. *Journal of Hydrology*, 403:213–223.

[Ares et al., 2020] Ares, M. G., Varni, M., and Chagas, C. (2020). Runoff response of a small agricultural basin in the argentine Pampas considering connectivity aspects. *Hydrological Processes*, 34:3102–3119.

[Berkowitz and Zehe, 2020] Berkowitz, B. and Zehe, E. (2020). Surface water and groundwater: unifying conceptualization and quantification of the two "water worlds". *Hydrology and Earth System Sciences*, 24:1831–1858.

[Beven, 1993] Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, 16:41–51.

[Bracken et al., 2013] Bracken, L. J., Wainwright, J., Ali, G. A., Tetzlaff, D., Smith, M. W., Reaney, S. M., and Roy, A. G. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews*, 119:17–34.

[Carroll et al., 2018] Carroll, R. W. H., Bearup, L. A., Brown, W., Dong, W., Bill, M., and Williams, K. H. (2018). Factors controlling seasonal groundwater and solute flux from snow-dominated basins. *Hydrological Processes*, 32:2187–2202.

[Carroll et al., 2019] Carroll, R. W. H., Deems, J. S., Niswonger, R., Schumer, R., and Williams, K. H. (2019). The importance of interflow to groundwater recharge in a snowmelt-dominated headwater basin. *Geophysical Research Letters*, 46:5899–5908.

[Carroll et al., 2020] Carroll, R. W. H., Gochis, D., and Williams, K. H. (2020). Efficiency of the summer monsoon in generating streamflow within a snow-dominated headwater basin of the Colorado River. *Geophysical Research Letters*, 47:e2020GL090856.

[Caviedes-Voullième et al., 2021] Caviedes-Voullième, D., Ahmadinia, E., and Hinz, C. (2021). Interactions of microtopography, slope and infiltration cause complex rainfall-runoff behavior at the hillslope scale for single rainfall events. *Water Resources Research*, 57:e2020WR028127.

[Caviedes-Voullième and Hinz, 2020] Caviedes-Voullième, D. and Hinz, C. (2020). From nonequilibrium initial conditions to steady dryland vegetation patterns: how trajectories matter. *Ecohydrology*, 13:2199.

[Coon et al., 2020] Coon, E. T., Berndt, M., Jan, A., Svyatsky, D., Atchley, A. L., Kikinzon, E., Harp, D. R., Manzini, G., Shelef, E., Lipnikov, K., Garimella, R., Xu, C., Moulton, J. D., Karra, S., Painter, S. L., Jafarov, E., and Molins, S. (2020). Advanced Terrestrial Simulator. US Department of Energy, USA. Version 1.0.

[Dunne, 1978] Dunne, T. (1978). Field studies of hillslope flow processes. In Kirkby, M. J., editor, *Hillslope Hydrology*, pages 227–293. John Wiley, NY, USA.

[Dunne et al., 1991] Dunne, T., Zhang, W., and Aubry, B. F. (1991). Effects of rainfall, vegetation, and microtopography on infiltration and runoff. *Water Resources Research*, 27:2271–2285.

[Ebel and Loague, 2006] Ebel, B. A. and Loague, K. (2006). Physics-based hydrologic-response simulation: Seeing through the fog of equifinality. *Hydrological Processes*, 20:2887–2900.

[Ebel and Loague, 2008] Ebel, B. A. and Loague, K. (2008). Rapid simulated hydrologic response within the variably saturated near surface. *Hydrological Processes*, 22:464–471.

[Fiori et al., 2007] Fiori, A., Romanelli, M., Cavalli, D. J., and Russo, D. (2007). Numerical experiments of streamflow generation in steep catchments. *Journal of Hydrology*, 339:183–192.

[Fiori and Russo, 2007] Fiori, A. and Russo, D. (2007). Numerical analyses of subsurface flow in a steep hillslope under rainfall: the role of the spatial heterogeneity of the formation hydraulic properties. *Water Resources Research*, 43:2006WR005365.

[Freeze and Cherry, 1979] Freeze, R. A. and Cherry, J. A. (1979). *Groundwater*. Prentice-Hall, Inc., NJ, USA.

[Frei and Fleckenstein, 2014] Frei, S. and Fleckenstein, J. H. (2014). Representing effects of micro-topography on runoff generation and sub-surface flow patterns by using superficial rill/depression storage height variations. *Environmental Modelling and Software*, 52:5–18.

[Frei et al., 2010] Frei, S., Lischeid, G., and Fleckenstein, J. H. (2010). Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland – a modelling study. *Advances in Water Resources*, 33:1388–1401.

[Gillefalk et al., 2021] Gillefalk, M., Tetzlaff, D., Hinkelmann, R., Kuhlemann, L.-M., Smith, A., Meier, F., Maneta, M. P., and Soulsby, C. (2021). Quantifying the effects of urban green space on water partitioning and ages using an isotope-based ecohydrological model. *Hydrology and Earth System Sciences*, 25:3635–3652.

[Gupta et al., 2009] Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *Journal of Hydrology*, 377:80–91.

[Han et al., 2020] Han, X., Liu, J., Srivastava, P., Mitra, S., and He, R. (2020). Effects of critical zone structure on patterns of flow connectivity induced by rainstorms in a steep forested catchment. *Journal of Hydrology*, 587:125032.

[Hubbard et al., 2018] Hubbard, S. S., Williams, K. H., Agarwal, D., Banfield, J., Beller, H., Bouskill, N., Brodie, E., Carroll, R., Dafflon, B., Dwivedi, D., Falco, N., Faybishenko, B., Maxwell, R. M., Nico, P., Steefel, C. I., Steltzer, H., Tokunaga, T., Tran, P. A., Wainwright, H., and Varadharajan, C. (2018). The East River, Colorado, Watershed: A mountainous community testbed for improving predictive understanding of multiscale hydrological-biogeochemical dynamics. *Vadose Zone Journal*, 17:180061.

[Khosh Bin Ghomash et al., 2019] Khosh Bin Ghomash, S., Caviedes-Voullième, D., and Hinz, C. (2019). Effects of erosion induced changes to topography on runoff dynamics. *Journal of Hydrology*, 573:811–828.

[Kirchner, 2009] Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, 45:W02429.

[Knoben et al., 2019] Knoben, W. J. M., Freer, J. E., and Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23:4323–4331.

[Loague et al., 2010] Loague, K., Heppner, C. S., Ebel, B. A., and VanderKwaak, J. E. (2010). The quixotic search for a comprehensive understanding of hydrologic response at the surface: Horton, Dunne, Dunton, and the role of concept-development simulation. *Hydrological Processes*, 24:2499–2505.

[Loague et al., 2006] Loague, K., Heppner, C. S., Mirus, B. B., Ebel, B. A., Ran, Q., Carr, A. E., BeVille, S. H., and VanderKwaak, J. E. (2006). Physics-based hydrologic-response simulation: foundation for hydroecology and hydrogeomorphology. *Hydrological Processes*, 20:1231–1237.

[Loritz et al., 2019] Loritz, R., Kleidon, A., Jackisch, C., Westhoff, M., Ehret, U., Gupta, H., and Zehe, E. (2019). A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation. *Hydrology and Earth System Sciences*, 23:3807–3821.

[Maavara et al., 2021] Maavara, T., Siirila-Woodburn, E. R., Maina, F., Maxwell, R. M., Sample, J. E., Chadwick, K. D., Carroll, R., Newcomer, M. E., Dong, W., Williams, K. H., Steefel, C. I., and Bouskill, N. J. (2021). Modeling geogenic and atmospheric nitrogen through the East River Watershed, Colorado, Rocky Mountains. *PLoS ONE*, 16:e0247907.

[Maxwell et al., 2019] Maxwell, R. E., Condon, L. E., Danesh-Yazdi, M., and Bearup, L. A. (2019). Exploring source water mixing and transient residence time distributions of outflow and evapotranspiration with an integrated hydrologic model and lagrangian particle approach. *Ecohydrology*, 12:e2042.

[Mays et al., 2002] Mays, D. C., Faybishenko, B. A., and Finsterle, S. (2002). Information entropy to measure temporal and spatial complexity of unsaturated flow in heterogeneous media. *Water Resources Research*, 38:1313.

[Meyerhoff and Maxwell, 2011] Meyerhoff, S. B. and Maxwell, R. M. (2011). Quantifying the effects of subsurface heterogeneity on hillslope runoff using a stochastic approach. *Hydrogeology Journal*, 19:1515–1530.

[Mirus et al., 2011] Mirus, B. B., Ebel, B. A., Heppner, C. S., and Loague, K. (2011). Assessing the detail needed to capture rainfall-runoff dynamics with physics-based hydrologic response simulation. *Water Resources Research*, 47:W00H10.

[Mirus and Loague, 2013] Mirus, B. B. and Loague, K. (2013). How runoff begins (and ends): characterizing hydrologic response at the catchment scale. *Water Resources Research*, 49:2987–3006.

[Mueller et al., 2007] Mueller, E. N., Wainwright, J., and Parsons, A. J. (2007). Impact of connectivity on the modeling of overland flow within semiarid shrubland environments. *Water Resources Research*, 43:W09412.

[Nash and Sutcliffe, 1970] Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - a discussion of principles. *Journal of Hydrology*, 10:2147–2171.

[Özgen-Xian et al., 2020] Özgen-Xian, I., Kesserwani, G., Caviedes-Voullième, D., Molins, S., Xu, Z., Dwivedi, D., Moulton, J. D., and Steefel, C. I. (2020). Wavelet-based local mesh refinement for rainfall-runoff simulations. *Journal of Hydroinformatics*, 22:1059–1077.

[Priestley and Taylor, 1972] Priestley, C. H. B. and Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100:81–92.

[Refsgaard et al., 2022] Refsgaard, J. C., Stisen, S., and Koch, J. (2022). Hydrological process knowledge in catchment modelling – lessons and perspectives from 60 years development. *Hydrological Processes*, 36:1099–1085.

[Ries et al., 2017] Ries, F., Schmidt, S., Sauter, M., and Lange, J. (2017). Controls on runoff generation along a steep climatic gradient in the Eastern Mediterranean. *Journal of Hydrology: Regional Studies*, 9:18–33.

[Seck et al., 2015] Seck, A., Welty, C., and Maxwell, R. M. (2015). Spin-up behavior and effects of initial conditions for an integrated hydrologic model. *Water Resources Research*, 51:2188–2210.

[Sivapalan, 2009] Sivapalan, M. (2009). The secret to 'doing better hydrological science': change the question! *Hydrological Processes*, 23:1391–1396.

[Smith et al., 2021] Smith, A., Tetzlaff, D., Kleine, L., Maneta, M., and Soulsby, C. (2021). Quantifying the effects of land use and model scale on water partitioning and water ages using tracer-aided ecohydrological models. *Hydrology and Earth System Sciences*, 25:2239–2259.

[Stephenson and Freeze, 1974] Stephenson, G. R. and Freeze, R. A. (1974). Mathematical simulation of subsurface flow contributions to snowmelt runoff, Reynolds Creek Watershed, Idaho. *Water Resources Research*, 10:284–294.

[Sterte et al., 2021] Sterte, E. J., Lidman, F., Lindborg, E., Sjöberg, Y., and Laudon, H. (2021). How catchment characteristics influence hydrological pathways and travel times in a boreal landscape. *Hydrology and Earth System Sciences*, 25:2133–2158.

[Stoelzle et al., 2020] Stoelzle, M., Schuetz, T., Weiler, M., Stahl, K., and Tallaksen, L. M. (2020). Beyond binary baseflow separation: A delayed-flow index for multiple streamflow contributions. *Hydrology and Earth System Science*, 24:849–867.

[Tarboton, 1997] Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33:309–319.

[Thompson et al., 2010] Thompson, S. E., Katul, G. G., and Porporato, A. (2010). Role of microtopography in rainfall-runoff partitioning: an analysis using idealized geometry. *Water Resources Research*, 46:W07520.

[Thornton et al., 2020] Thornton, M. M., Shrestha, R., Wei, Y., Thornton, P. E., Kao, S., and Wilson, B. E. (2020). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4, ORNL DAAC, Oak Ridge, Tennessee, USA.

[Tokunaga et al., 2019] Tokunaga, T. K., Wan, J., Williams, K. H., Brown, W., Henderson, A., Kim, Y., Tran, A. P., Conrad, M. E., Bill, M., Carroll, R. W. H., Dong, W., Xu, Z., Lavy, A., Gilbert, B., Carrero, S., Christensen, J. N., Faybishenko, B., Arora, B., Siirila-Woodburn, E. R., Versteeg, R., Raberg, J. H., Peterson, J. E., and Hubbard, S. S. (2019). Depth- and time-resolved distributions of snowmelt-driven hillslope subsurface flow and transport and their contributions to surface waters. *Water Resources Research*, 55:9479–9499.

[Viero and Valipour, 2017] Viero, D. P. and Valipour, M. (2017). Modeling anisotropy in free-surface overland and shallow inundation flows. *Advances in Water Resources*, 104:1–14.

[Weill et al., 2013] Weill, S., Altissimo, M., Cassiani, G., Deiana, R., Marani, M., and Putti, M. (2013). Saturated area dynamics and streamflow generation from coupled surface–subsurface simulations and field observations. *Advances in Water Resources*, 59:196–208.

[Xu et al., 2022] Xu, Z., Molins, S., Özgen-Xian, I., Dwivedi, D., Svyatsky, D., Moulton, J. D., and Steefel, C. O. (2022). Integrated hydrology and reactive transport modeling for mountainous watershed. *Water Resources Research*, submitted:n/a.

[Zehe et al., 2021] Zehe, E., Loritz, R., Edery, Y., and Berkowitz, B. (2021). Preferential pathways for fluid and solutes in heterogeneous groundwater systems: self-organization, entropy, work. *Hydrology and Earth System Sciences*, 25:5337–5353.

[Zhi et al., 2019] Zhi, W., Li, L., Dong, W., Brown, W., Kaye, J., Steefel, C. I., and Williams, K. H. (2019). Distinct source water chemistry shapes contrasting concentration-discharge patterns. *Water Resources Research*, 55:4233–4251.