Characterizing hillslope-stream connectivity with a joint event analysis of stream and groundwater levels

Daniel Beiter[1], Markus Weiler[2], and Theresa Blume[1]

[1]GFZ German Research Centre for Geosciences, 4.4 Hydrology, Potsdam, Germany
[2]University of Freiburg, Chair of Hydrology, Freiburg, Germany

Correspondence: Daniel Beiter (daniel.beiter@gfz-potsdam.de)

Abstract. Hillslope-stream connectivity controls runoff generation, both during events and baseflow conditions. However, assessing subsurface connectivity is a challenging task, as it occurs in the hidden subsurface domain where water flow cannot be easily observed. We therefore investigated if the results of a joint analysis of rainfall event responses of near-stream groundwater levels and stream water levels could serve as a viable proxy for hillslope-stream connectivity. The analysis focuses on the

5 extent of response, correlations, lag times and synchronicity. A newly developed data analysis scheme of separating the aspects of a) response timing and b) extent of water level change provides new perspectives on the relationship between groundwater and stream responses. In a second step we investigated if this analysis can give an indication of hillslope-stream connectivity at the catchment scale.

Stream- and groundwater levels were measured at five different hillslopes over 5 to 6 years. Using a new detection algorithm

10 we extracted 706 rainfall response events for subsequent analysis. Carrying out this analysis in two different geological regions (schist and marls) allowed us to test the usefulness of the proxy under different hydrological settings while also providing insight into the geologically-driven differences in response behaviour.

For rainfall events with low initial groundwater level, groundwater level responses often lag behind the stream with respect to the start of rise and the time of peak. This lag disappears at high antecedent groundwater levels. At low groundwater levels the

15 relationship between groundwater and stream water level responses to rainfall are highly variable, while at high groundwater levels, above a certain threshold, this relationship tends to become more uniform. The same threshold was able to predict increased likelihood for high runoff coefficients, indicating a strong increase in connectivity once the groundwater level threshold was surpassed.

The joint analysis of shallow near-stream groundwater and stream water levels provided information on the presence or absence

20 and to a certain extent also on the degree of subsurface hillslope-stream connectivity. The underlying threshold processes were interpreted as transmissivity feedback in the marls and fill-and-spill in the schist. The value of these measurements is high, however, time series of several years and a large number of events are necessary to produce representative results. We also find that locally measured thresholds in groundwater levels can provide insight into catchment-scale connectivity and event response. If the location of the well is chosen wisely, a single time series of shallow groundwater can indicate if the catchment

25 is in a state of high or low connectivity.

1 Introduction

Hillslope-stream connectivity controls both runoff generation (Detty and McGuire, 2010; Jencso et al., 2010; Penna et al., 2015; Scaife and Band, 2017) and export of solutes, pesticides (Ocampo et al., 2006; Jackson and Pringle, 2010) and particulate matter (Thompson et al., 2013). Understanding patterns, controls and dynamics of hillslope-stream connectivity is

5   therefore of interest not only for flood prediction but also for water quality management and policy making. Unfortunately, the investigation of this connectivity is notoriously difficult, for a number of reasons: it is variable in space and time (much more than our catchment models generally account for) and it is often controlled by thresholds, either in wetness state or in forcing (rainfall amounts and intensity) (Detty and McGuire, 2010b; McGuire and McDonnell, 2010; Scaife and Band, 2017; Oswald et al., 2011; Graham et al., 2010). Full connectivity is usually established only during brief periods of time (Freer et al., 2002;

10   Ocampo et al., 2006; Haught and Meerveld, 2011; van Meerveld et al., 2015). Identifying and measuring hillslope-stream connectivity becomes even more challenging as we are dealing with extensive along-stream interfaces which makes identification/pinpointing of hot spots difficult. While surface connectivity at least often leaves visible traces, subsurface connectivity is usually invisible and therefore hard to localise and measure (Blume and van Meerveld, 2015).

Standard approaches for the investigation of hillslope-stream connectivity include hillslope trench studies (often combined

15   with piezometers) (Bachmair and Weiler, 2014; van Meerveld and McDonnell, 2006b) and tracer-based analyses (McGuire and McDonnell, 2010; McGlynn and McDonnell, 2003; Anderson et al., 1997). While the first approach gives detailed information about (usually) a single hillslope (Graham et al., 2010) it requires considerable effort in the field (both with respect to time and finances), the second approach provides an integral assessment at the catchment scale, but offers little information on spatial patterns or spatial extent of connectivity. At the stream bed interface distributed temperature sensing (DTS) can provide

20   spatially highly resolved information of stream bed temperatures and under favourable conditions information about groundwater inflow points (Krause et al., 2012). While these datasets can be very informative, DTS systems are expensive, require continuous power supply and are time-intensive in installation. All of these methods are often employed on a short-term basis only: a few events, a season, possibly a year. As a result, one is left with the question how representative these snapshots are.

Even though state variables such as soil moisture or groundwater level do not provide actual water fluxes they are often used

25   to asses hydrologic subsurface connectivity (Detty and McGuire, 2010; Haught and Meerveld, 2011; Freer et al., 2002; van Meerveld and McDonnell, 2006b; Ali et al., 2011; Anderson et al., 2010), and using many repeated snap-shots allows to at least infer flow processes (Bracken et al., 2013). Shallow groundwater levels can provide information about catchment state and a joint analysis of groundwater and streamflow dynamics in response to rainfall events offers basic information on runoff generation processes and hillslope-stream connectivity. The relationship of pre-event groundwater levels and streamflow re-

30   sponse is often governed by a threshold in groundwater level above which streamflow responds much more strongly than below (Anderson et al., 2010; Detty and McGuire, 2010b; van Meerveld and McDonnell, 2006b). This threshold indicates a sudden increase in contributing area which directly translates to an increase in hillslope-stream connectivity (Anderson et al., 2010; Detty and McGuire, 2010b; van Meerveld and McDonnell, 2006b).

In this study we go for a targeted as well as pragmatic approach: we target specifically the footslope and the riparian zone as

the essential interface between slope and stream. We hypothesise that monitoring shallow groundwater tables in the riparian zone over longer periods of time and thus not only a few, but a large number of events will provide not full, but representative information on hillslope-stream connectivity at low cost. Previous use of piezometers for this purpose often extended over the entire hillslope (Bachmair and Weiler, 2014; van Meerveld and McDonnell, 2006b) which increased financial and maintenance

5   efforts. While this can be very informative, we suggest that our pragmatic approach focusing only on the footslope and a joint analysis of shallow groundwater and streamflow response to rainfall events will still allow us to develop a general picture of when and where connectivity is established and how often this occurs. Analysing the relationship between responses in near-stream shallow groundwater and stream thus permits a rough interpretation of the dominant processes. We will investigate the potential and limitations of this approach by comparing 5 footslopes covering two distinct geologies. A newly developed data

10   analysis scheme which separates the aspects of response timing and extent of water level change opens up new perspectives on these interactions.

We thus target the following research questions: Can the measurements of shallow near-stream groundwater and stream water level and their targeted analysis provide information about subsurface hillslope-stream connectivity? When does hillslope-stream connectivity occur and what are the underlying hydrological processes? How does this differ between geologic regions?

15   Can the state of connectivity of the catchment be assessed through shallow groundwater measured at a single location?

2 Methods

2.1 Study catchment

This investigation targets the 244km^2 Attert catchment in western Luxembourg, with altitudes between 243 and 549m.a.s.l. (Figure 1, top left). It is driven by a runoff regime with generally low discharge in summer and high discharge in winter. Despite

20   the seasonal differences in runoff, precipitation events are distributed over the entire year, with a mean annual precipitation of 760mm.

The catchment can be divided into three main geologies – marls, schist and sandstone – and two geologies of lower significance (alluvials and buntsandstein), shown in Figure 1 (top right). Most of the catchment is characterised by marls and Stagnasols with high clay content (20-60%), an undulating landscape and mostly agricultural land use (Sprenger et al. (2016)). The high

25   contents of clay lead to low hydraulic conductivities and a limited drainage capacity. The north-western area (Figure 1) consists of schist bedrock and Cambisols with a texture between loam, silty loam and clayey loam which can drain freely until the soil-bedrock interface (Sprenger et al. (2016)). The landscape is here governed by elevated plateaus with mostly agricultural land use and steep forested hillslopes leading to perennial headwater streams.

A monitoring network with 45 stations was installed in the Attert catchment, recording environmental data such as climate data,

30   soil moisture, groundwater and stream level, amongst others (Zehe et al., 2014; Demand et al., 2019). For the investigation of hillslope-stream connectivity we selected those monitoring sites which were situated at a stream and thus allow a comparison between near-stream shallow groundwater level and the associated stream water level. Unfortunately no such site was available in the sandstone due to its very low drainage density, so the investigation focused on the two geologies marls and schist (Table

1 and Figure 1, bottom). The five selected stations were put into operation between June 2012 and July 2013 and the time span until end of Juli 2017 was considered for the analysis. The spatial distribution of the 3-4 piezometers at each site can be seen in
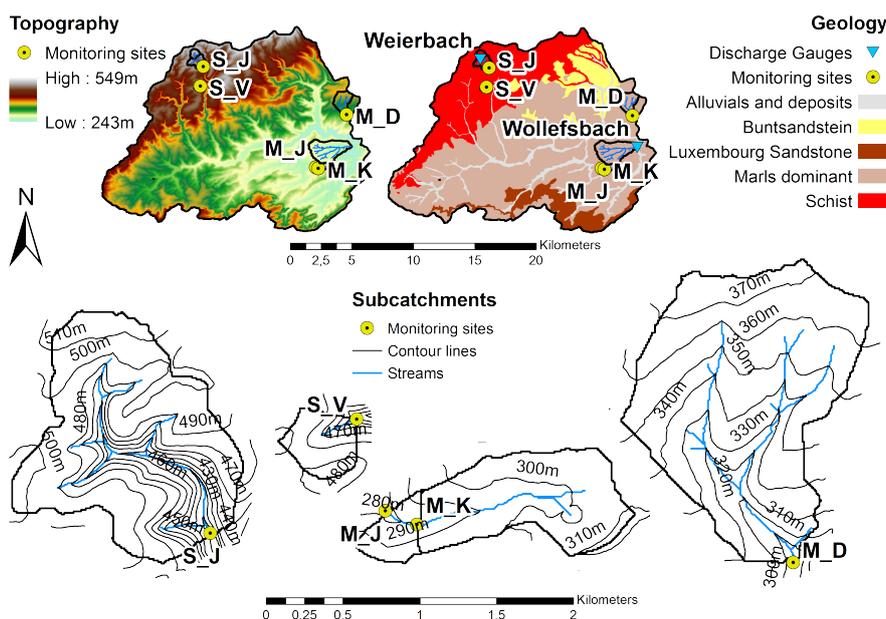
Figure 1. The Attert catchment in western Luxembourg and the five monitoring sites: M_D, M_J, M_K (marls), S_J and S_V (schist). Top left: catchment topography, top right: geology, bottom: subcatchments.

Figure 2. M_D is located on a wide meadow with gentle inclination and Piezometers 1-3 have a distance to the stream between 2m and 10m, while Piezometer 4 is on the steep opposite hillslope directly below a road cut (subsurface probably disturbed

5   during road construction). Piezometer depths extend to about a meter below the stream bed. The other two marls sites – M_J and M_K – are located on two nearby forested plains with the stream incised to about 2.5m and piezometer depths of around 2m. The horizontal distances between stream and piezometers are between 4m and 13m for both sites. S_J is located on a small meadow flood plain, flanked by steep forested hillslopes on both sides of the stream. Piezometer depths are here around 1.5m and reach below the stream bed. Piezometers 1-3 are situated on one side of the stream with distances of about 4 - 8m,

10   while Piezometer 4 is located on the other side at a distance of 6m. S_V is located at a steep forested hillslope. The distance to the stream is between 2m (Piezometer 4) and 15m (Piezometer 1) and only the lower piezometers (3 and 4) extend to depths below the stream bed.

2.2   Monitoring data

Each of the five sites described in section 2.1 was equipped with three to four piezometers to measure shallow groundwater
15   level and one sensor for stream water level. Vertical boreholes were drilled until refusal using the Cobra TT jackhammer with
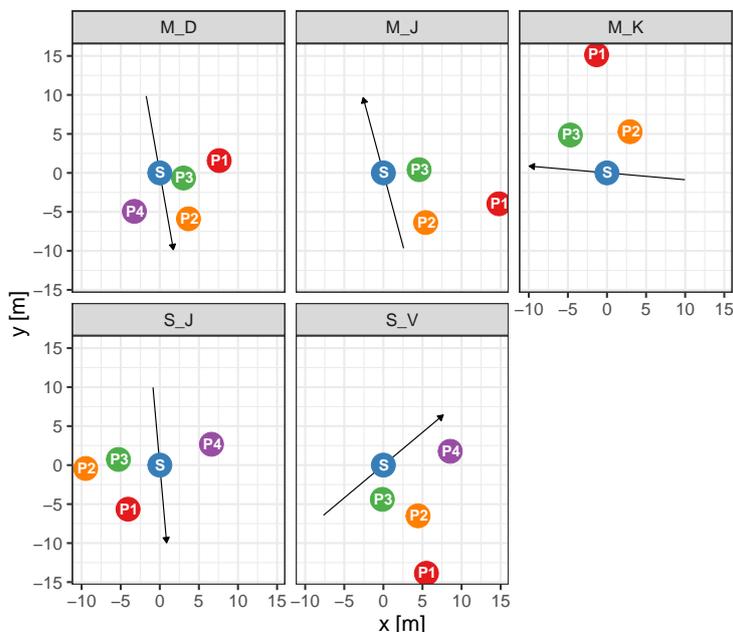
Figure 2. Schematic maps of the five sites. "P" stands for piezometer with the corresponding number while "S" stands for stream and is located at (0,0). The arrows point into the direction of stream flow. The coordinates are relative distances to the stream water level sensor (positive y-axis points north).

Table 1. Monitoring sites with basic attributes.

Site name	Geology	Soil	Land use	Drainage[1]	Slope quartiles	Upstream area	# of Piezometers
[-]	[-]	[-]	[-]	[-]	[°]	[ha]	[-]
M_D	Marls	Stagnosol	Pasture	Limited	2.3 / 3.3 / 4.5	200	4
M_J	Marls	Stagnosol	Forest	Limited	1.3 / 2.3 / 3.6	80	3
M_K	Marls	Stagnosol	Forest	Limited	1.3 / 2.2 / 3.3	68	3
S_J	Schist	Cambisol	Pasture	Free	2.3 / 4.8 / 12.2	154	4
S_V	Schist	Cambisol	Forest	Free	2.7 / 5.2 / 8.0	17	4

[1] According to Sprenger et al. (2016)

a hollow boring head of 75mm diameter. Perforated PVC tubings of 50mm diameter were wrapped into non-woven fabric, installed and packed with filter gravel between 4 and 8mm diameter. The uppermost 30cm below ground level were packed with sealing clay to prevent infiltration bypassing the soil. Depth of refusal was in most cases below 2m and the water level sensors were installed around 2cm above the bottom.

5 The sensors used were CTD temperature corrected pressure transducers by METER (formerly Decagon), measuring electric

conductivity, temperature and water depth. Full scale is 10m, with a resolution of 2mm and an accuracy of ±0.05% of full scale. Connection cables provide ventilation to the transducer and compensate for air pressure. Automated data loggers (CR1000 by Campbell Scientific) were taking measurements from all sensors with a temporal resolution of 5min. Hourly precipitation data from Roodt station was obtained from the "Administration de la gestion de l'eau de Luxembourg". Discharge data with 15

5   min temporal resolution were provided from the Luxembourg Institute for Science and Technology (LIST) for the Weierbach station (for schist) and the Wollefsbach station (for marls).

2.3   Event definition

Automatic event detection is essential when working with long time series and a large number of events. To this end, it is necessary to define a generic response pattern (Figure 3). The general response pattern begins with a pre-event minimum

10   (h_{preMin}). When a precipitation event starts, the water level increases until it reaches its peak ($h_{maximum}$). After that peak, water level decreases and the event ends with a post-event minimum ($h_{postMin}$) that might differ from the pre-event minimum. These three points are used to describe water level changes during the event. However, the time period between the two minima (pre- and post-event) is not a robust measure for the event duration. Before or after events water levels are often not stable but subject to small but misleading trends (e.g. wetting-up phase or recession). While searching for the two minima a minimal

15   decline has almost no effect on the water level but inappropriately increases the extracted event duration. To compensate for that, two threshold points ($h_{riseThreshold}$ and $h_{fallThreshold}$) were introduced – one on each limb – that allow for a better temporal representation of each event. Both are defined as a certain percentage of $h_{preAmplitude}$ and $h_{postAmplitude}$. In the case of the rising limb the time where the water table exceeds $h_{riseThreshold}$ is called t_{rise} (see Figure 3). Analogously, the moment the water level falls below $h_{fallThreshold}$ is defined as t_{fall}. The distances to $t_{maximum}$ are described as the $t_{riseInterval}$ and

20    $t_{fallInterval}$, respectively. So for time-related analyses these two intervals are used as they are not prone to pre- and post-event trends, but capture the actual event response dynamics. A percentage of 10% of $h_{preAmplitude}$ and $h_{postAmplitude}$ was found to be suitable for that task.

2.4   Event detection

The purpose of the event detection is to parse the entire water level time series and extract those intervals during which the

25   water level shows a response to rainfall. Algorithm 1 specifies the necessary steps for the event detection. At first, minimum amplitudes and search intervals need to be defined. Both parameters are subject to a compromise: The minimum amplitudes are used to prevent measurement noise from being mistakenly detected as events, with the drawback of possibly excluding actual low-amplitude events from detection. Search intervals are used to discriminate between subsequent events, which involves the risk of not completely capturing a very long event. In a second step, all local maxima of the stream water level are located.

30   Thirdly, for each maximum the pre-defined search intervals are used in order to determine the global minima in the rising and falling limb. Defining these search intervals depends on the catchment size. Generally speaking, the search interval for the rising limb should be approximately equal to the concentration time of the subcatchment to guarantee that the complete rising limb is covered. Therefore, shorter search intervals are suited for headwaters (several hours to a day) and longer ones for
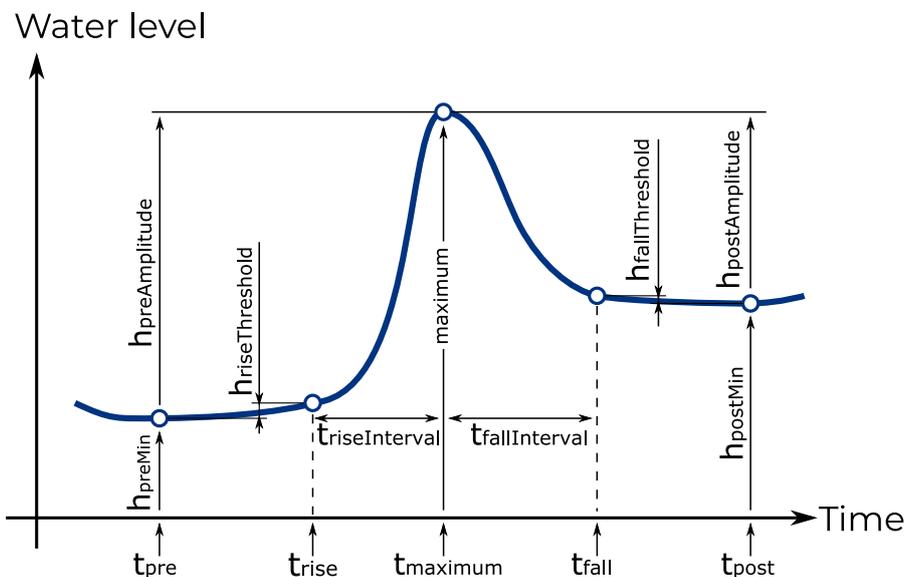
Figure 3. Event definition and characteristic variables for event response analysis.

Algorithm 1 Event detection algorithm

DEFINE: amplitude thresholds for rising and falling limb ($h_{preAmplitudeMin} = 20$mm, $h_{postAmplitudeMin} = 10$mm)

DEFINE: search interval for rising and falling limb ($t_{preSearchInterval} = 24$h, $t_{postSearchInterval} = 48$h)

FIND: all local maxima in time series

for each maximum **do**

    FIND: absolute minima on rising and falling limb within defined search intervals

    CALCULATE: $h_{preAmplitude}$ and $h_{postAmplitude}$

    if $h_{preAmplitude} < h_{preAmplitudeMin}$ or $h_{postAmplitude} < h_{postAmplitudeMin}$ **then**

        DISCARD: current maximum

    end if

end for

MERGE: overlapping events

RETURN: t_{pre} and t_{post} of each merged event

---

lowland basins (several days). Also, the rise interval is shorter than the fall interval as such events are generally right-skewed due to retention behaviour. If two or more events overlap, they are merged into one (Figure 4) and the characteristic variables are recalculated for the entire event.

The event detection was first applied to the stream water level time series which returns t_{pre} and t_{post} for each detected stream event. For each of these stream events a subsequent event detection is performed on the shallow groundwater level time series.

Thus, we only include events in the analysis where stream water levels showed a response. Using each stream event for the detection of a possible groundwater event implies that the maximum temporal extent of the groundwater event is equal to the stream event. This is a shortcoming of this method, as a time lag between shallow groundwater and stream or drawn-out groundwater recession might lead to the predefined search window clipping the drawn-out event in the shallow groundwater.

5   However, in the case of multiple subsequent events a clear definition must exist in order to keep a one-to-one relation between stream and ground water events. If no temporal boundaries were applied for subsequent event detections, an event in the shallow groundwater might overlap with two or more stream events which would drastically increase the complexity of the analysis. Because of the relatively small distances of less than 15m between stream and piezometers, and the small headwater catchments, response delays between stream and piezometer are presumed to be rather short, reducing the risk of clipping.

10   Also, taking t_{pre} and t_{post} as the temporal extent for subsequent detections in groundwater provides a buffer for potential lag times. This one-to-one approach is considered most appropriate as it is a trade-off between good operability of the detection algorithm and a high coverage of stream and groundwater events.

Amplitude thresholds were chosen via trial and error to prevent diurnal stream water fluctuations caused by root water uptake from provoking (erroneous) events. The threshold for the rising limb (20mm) is greater than for the falling limb (10mm)

15   because during the wetting-up phase (in autumn) post-event water levels are very often higher than the pre-event water levels as the catchment becomes more saturated. Using the same threshold for rising and falling limb would lead to the rejection of small events with such a behaviour.

Search intervals were estimated by testing a range of values. A time of 24h for the rising limb performed satisfactorily in our catchments even for long precipitation events and did not merge several subsequent events into one bulk event. With 48h for the

20   falling limb the retention behaviour of the catchment was taken into account allowing for a long-tailed recession in comparison to the rise. The detection algorithm was run for each site individually, as a result the number and selection of detected events is site-specific.

2.5 Event type

Introducing event type descriptors allows to infer specific characteristics of a site and its experimental setup. The total number

25   of events for a certain site is defined by its stream response, regardless of whether or not the shallow groundwater responds during the stream events. Event types are: *Complete* detections arise when the water level sensor was initially submerged and the occurring event fulfils the stated detection criteria. For *Partial* detections the criteria are met but the piezometer is initially dry, so it is unknown how far below the sensor level the event started. *Dry* events are events where the piezometer is dry during the stream event and does not record any response. If no local maximum could be found in the groundwater during a

30   stream event, the type was set to *noLocalMaximum*. *lowAmplitudes* means that the rise and/or fall amplitude thresholds are not surpassed. This might be due to a very low-amplitude response but can also cover events with a high rise amplitude but low fall amplitude, in particular when the peak is very close to the t_{post} boundary, which signals a long time lag between stream and groundwater. *All NA* indicates technical sensor problems in the piezometers during the detected streamflow event. While only *complete* events contain valid state and timing variables that can be put into relation with the stream (and are subsequently
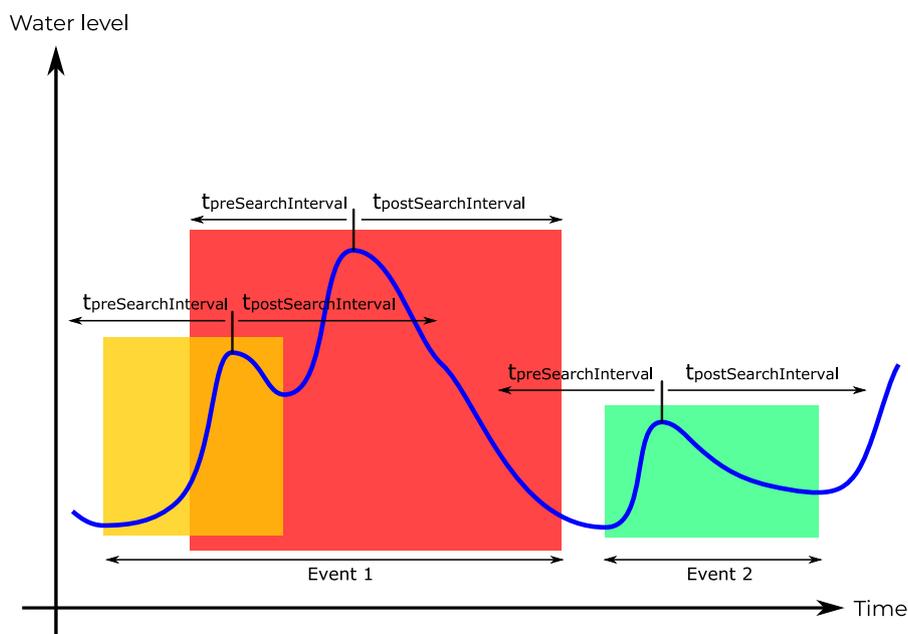
Figure 4. Merging conditions of consecutive events. Time series of water level showing three local maxima and the corresponding search windows for the minima. The coloured boxes mark the independently detected events as the interval between the two absolute minima around each peak within the respective search interval. The yellow and red events overlap and are merged into one. The green event is an independent second event.

used for the detailed analyses), all non-complete events also contain relevant information. Knowing about the frequency of occurrence of these other event types helps to characterise each piezometer and site.

2.6 Event analysis

The event analysis aims for a better understanding of how and under which conditions the shallow groundwater connects to the
5   stream or disconnects from it. Observing the relation of water table dynamics between stream and shallow groundwater, can reveal connectivity patterns which in turn give insight in the underlying processes. This simultaneous view on groundwater and stream is what is defined as the groundwater-stream (response) relation. A many-event approach ensures that a high variability of catchment conditions and response behaviours is incorporated into the analysis to cover the entire bandwidth of hydrologic system behaviour. Analyses covering single or a low number of events lack the ability of estimating variability and do not allow
10 to deduce how "typical" or "extreme" the event is and if it is representative.

Because the problem is multi-dimensional and considerably complex, a strategy was chosen that allows to examine various aspects of the hydrologic responses independently. Combining the information of these different aspects should then give a deeper insight into the occurring processes that control the various hillslope-stream-systems.

The hillslope-stream connectivity can be investigated for periods before an event starts where underlying hydrologic processes

take place on more long term (seasonal) time scales and are represented by the baseflow. As a measure for this connectivity during baseflow conditions (between events) the rank correlation of all pre-event minima (h_{preMin}) between each piezometer and the corresponding stream was used. To visually compare before-event relations across piezometers and sites, each sensor's water level was normalised by its minimum and maximum h_{preMin} value.

$$h'_{preMin} = \frac{h_{preMin} - \min(h_{preMin})}{\max(h_{preMin}) - \min(h_{preMin})} - \begin{cases} 0 & \text{for stream} \\ 1 & \text{for groundwater} \end{cases} \tag{1}$$

Equation 1 results in a value between 0 and 1 for the stream. For groundwater the value 1 was subtracted from the normalised h'_{preMin} to obtain values between -1 and 0 emphasising water levels below ground surface.

When analysing the response relation during an event, highly dynamic infiltration and runoff processes govern the hydrologic hillslope-stream system on short time scales (hours and days). To analyse this relation during events, timing and state changes were not examined jointly (e.g. by plotting water table time series against each other) but as two separate aspects. This approach provides a separate view of the temporal behaviour on the one hand and changes in the state variables (water levels) of the hydrologic system on the other.

Relative timing and lags between groundwater and stream responses are hints for a causal relationship. To investigate the variability of this relative timing across all events, piezometers and sites the response behaviour was reduced to timing effects only. A very similar normalisation approach as in Equation 1 was used to compare timings of groundwater responses with those of the stream. Equation 2 uses the time at which the stream exceeds the 10% threshold $t_{rise\ stream}$ and the time where it reaches its peak $t_{max\ stream}$ to normalise groundwater and stream event timing.

$$t' = \frac{t - t_{rise\ stream}}{t_{max\ stream} - t_{rise\ stream}} \tag{2}$$

This stream-based normalisation leads to a value of 0 for the t_{rise} in the stream and 1 for the $t_{maximum}$. A corresponding groundwater event that starts at 0 and reaches its maximum at 1 has the exact same timing as the stream. Values below 0 correspond to a time before the stream responded while values above 1 correspond to a time where the stream already is in recession. By applying this normalisation it is possible to compare relative time lags between stream and groundwater as well as differences in the duration.

The extent of water level increases in stream and groundwater and the relationship of the two can provide useful information on the dominant runoff generation processes. We would expect that a given increase in groundwater level at a given depth would result in a more or less deterministic increase of stream water level (assuming the groundwater fluctuations are representative of the catchment). As this also assumes that there is a connection between groundwater and stream and that runoff generation is controlled by shallow groundwater contributions, deviations from deterministic relationships are an indication of other runoff generation processes or flow path variability. Removing the temporal component and only focusing on the extent of the increase between pre-event water level and peak water level enables inspecting this relationship.

To investigate if shallow groundwater observations at a given hillslope can be used as a proxy for the state of connectivity in the entire catchment we analysed the relationship between event runoff coefficients and the depth. The runoff coefficient

describes the ratio of accumulated event discharge at the catchment outlet and accumulated catchment precipitation (Equation 3). Even though each experimental site monitors stream level, no reliable discharge information is available since rating curves are fragmentary and thus uncertain, or do not exist. Therefore, runoff coefficients (C) are calculated for nearby subcatchments (Wollefsbach and Weierbach; see Figure 1). The spatial proximity ensures that detected stream water level events coincide

5 with discharge events. Baseflow ($Q_{baseflow}(t)$) was defined as the area below the straight line connecting t_{rise} and t_{fall} and was subtracted from the total discharge ($Q(t)$) to calculate the actual stormflow. Precipitation ($P(t)$) from Roodt station was considered sufficiently representative across the Attert Catchment to be used for all runoff coefficient calculations.

$$C = \frac{\int_{t_{rise}}^{t_{fall}} Q(t) - Q_{baseflow}(t)\, dt}{A \int_{t_{rise}}^{t_{fall}} P(t))\, dt} \tag{3}$$

Relating the shallow groundwater information to the event runoff coefficients can help us to assess how representative the local

10 measurements are for the entire catchment upstream.

3 Results

3.1 Event detection

Our event detection algorithm identified between 119 and 159 stream runoff events per site and covered a period of five to six years. Not all of these were also detected in all piezometers (Figure 5 and 6). This can be due to data gaps as a result of

15 technical failure of the sensor or data gaps as the piezometer fell dry or because the response in the groundwater was strongly dampened and thus did not fulfil the criteria of the algorithm. In general, the temporal distribution of the detected events shows similar patterns across all sites (Figure 5). It also allows to identify M_D_Piezo4 and M_K_Piezo3 as behaving very differently with many *lowAmplitudes* and *allNA* events. In the case of *lowAmplitudes* we found that many events were clipped by the pre-defined time-window due to very long delays in relation to the stream, which were longer than in the other piezometers

20 at these sites.

Referring to the response type two main patterns can be distinguished (Figure 5). Sites where the sensors remain submerged throughout the observation period thus producing many *complete* events (M_D and S_J) and sites with piezometers falling dry in summer and autumn (M_K, M_J and S_V). While at the two marls sites these dry periods occur at all piezometers concurrently, at S_V the number of dry events increases in upslope direction (from Piezometer 3 to Piezometer 1). The aggregated

25 values in Figure 6 also reveal two response types with low occurrences – namely *noLocalMaximum* and *partial* events. The *noLocalMaximum* response is very rare with only 11 occurrences and is thus considered negligible.

3.2 Before-event hillslope-stream connectivity

The rank correlation coefficients were found to be lower in marls than in schist sites (background colour in Figure 7). In schist only the two upslope piezometers (Piezo1 and Piezo2) of S_V show lower correlation values, while the others remain above

30 0.75. For the three marls sites rank correlation coefficients are generally lower (between 0.25 and 0.75) with higher variation.
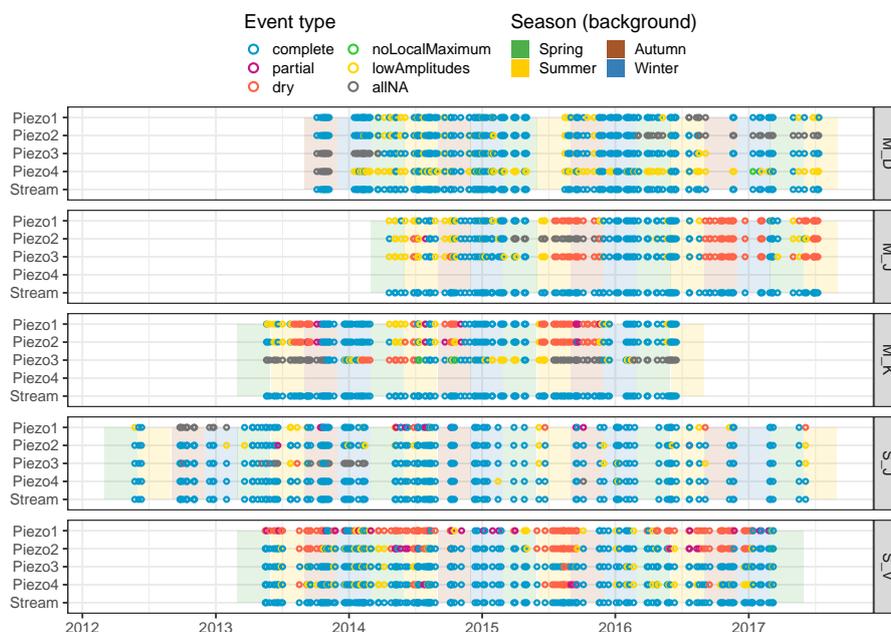
Figure 5. Spatiotemporal distribution of detected events for marls (M_) and schist (S_) between June 2012 and July 2017. Seasons are defined as the periods Dec-Feb (Winter), Mar-May (Spring), Jun-Aug (Summer) and Sep-Nov (Autumn)

In marls most pre-event groundwater levels cluster in the the shallow depths above -0.4 (M_K) and -0.3 (M_D and M_J) which correspond to absolute depths of -0.5m (Figure 9). Schist groundwater levels are more evenly distributed over the entire range (Figure 7). The point colours representing the seasons illustrate that groundwater levels are high in winter and spring and low in summer and autumn.

5   ### 3.3   Comparison of relative response timing between stream and groundwater

The relative timing between groundwater and stream is illustrated in Figure 8. The two black vertical lines represent the timing of the stream event with t_{rise} at $x = 0$ and $t_{maximum}$ at $x = 1$. Each horizontal bar depicts a groundwater response event with its own t_{rise} at the left end and $t_{maximum}$ at the right end. Groundwater responses that start at 0 and end at 1 have the exact same timing as the stream response. Starting values below 0 reveal a groundwater response before the stream, while an end

10   value above 1 indicates that the stream is already in recession before the groundwater reaches its maximum. The events are sorted on the y-axis by the normalised rise time in the groundwater from delayed groundwater response at the bottom to early groundwater response at the top. Additionally, the bar colours display the normalised pre-event water levels with high pre-event groundwater level in blue and low pre-event groundwater level in red.

At M_D (Piezo1 to Piezo3), S_J (Piezo1 to Piezo3) and S_V (Piezo3 to Piezo4) a strong relation between pre-event groundwa-

15   ter levels and event timing can be observed. Events occurring at high pre-event groundwater levels (bluish) correspond with a
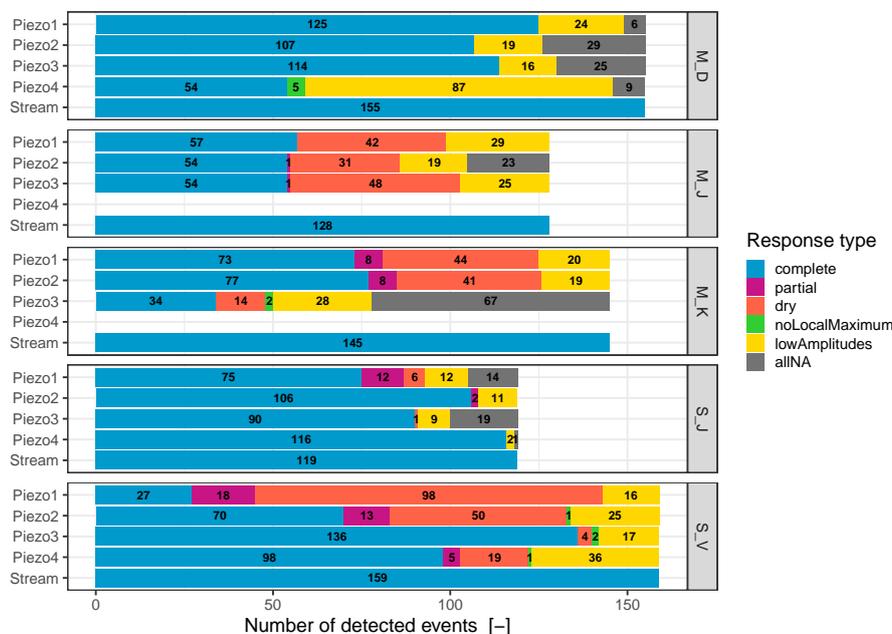
Figure 6. Number of detected events in the streams and groundwater at each stream gauge and piezometer, including types of event responses

mostly simultaneous rise in groundwater and stream while for events at low groundwater levels (reddish) the groundwater rise lags behind the stream. Considering the peak, high groundwater events reach their maximum before or simultaneously with the stream while during low groundwater the maximum is reached significantly after the stream. At sites M_J, M_K and S_V (Piezo1 to Piezo2) this separation of high (bluish) pre-event groundwater on top and low (reddish) pre-event groundwater at

5   the bottom is visible but not quite as pronounced as for the other sites.

3.4   Event-induced increases in stream and groundwater levels

The extent of water level increases in stream and groundwater and the relationship of the two is illustrated in Figure 9. Both pre-event water levels (stream and groundwater) are used as coordinates for the beginning of an event line (lower left point) and the maxima as the coordinates for the end (upper left), with stream water levels on the x-axis and groundwater levels on

10   the y-axis. As we removed the temporal component it is important to keep in mind that peak values not necessarily occur at the same time. We observe a change in response behaviour between stream and groundwater marked by a threshold which was derived visually (dotted horizontal lines) in Figure 9. At low groundwater levels amplitudes in the rising limb are large in the groundwater and low in the stream (steep slope of lines), while above the threshold the amplitudes in groundwater are capped at a certain depth below the surface, and stream amplitudes can become large (low slope of lines in Figure 9). Also, the

15   variability of pre-event conditions and event responses is larger below the threshold, while above, the lines are more likely to fall on top of each other and become more deterministic. This is particularly the case for M_D (except Piezo4), M_K (except
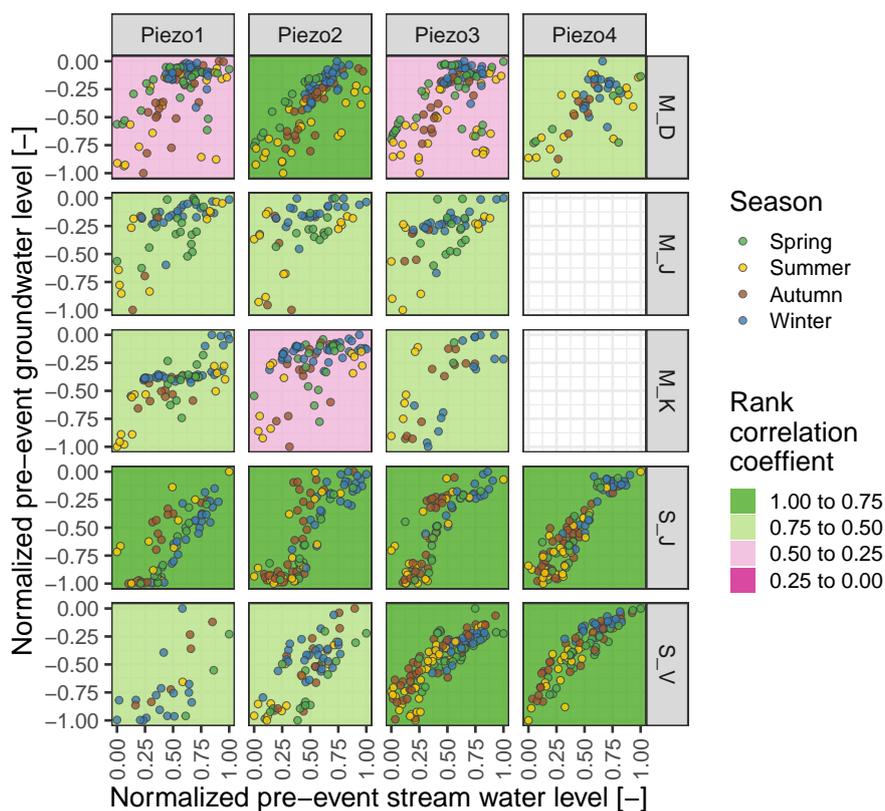
Figure 7. Normalised stream and groundwater levels before the investigated precipitation events (h_{preMin}). The background colour represents the rank correlation coefficient, the point colours illustrate the season. Both axes are normalised by minimum and maximum h_{preMin} (see equation 1). The negative range on the y-axis indicates depths below ground (groundwater), the positive range on the x-axis depths above ground (stream).

Piezo3) and S_J. Winter events cluster above the threshold and the other three seasons below the threshold and in the transition zone.

3.5 Runoff coefficient

The relation between local pre-event groundwater levels and the event runoff coefficients is displayed in Figure 10. The dotted
5 horizontal lines represent the same individual shallow groundwater thresholds for each piezometer identified in Figure 9 (but here with the normalised pre-event water level on the y-axis). Colours indicate whether the groundwater responded before the stream (red) or after the stream (blue). At M_D, S_J and S_V the pattern is very similar: below the individual pre-event groundwater thresholds runoff coefficients are very small, but increase significantly both in value as well as in variability when
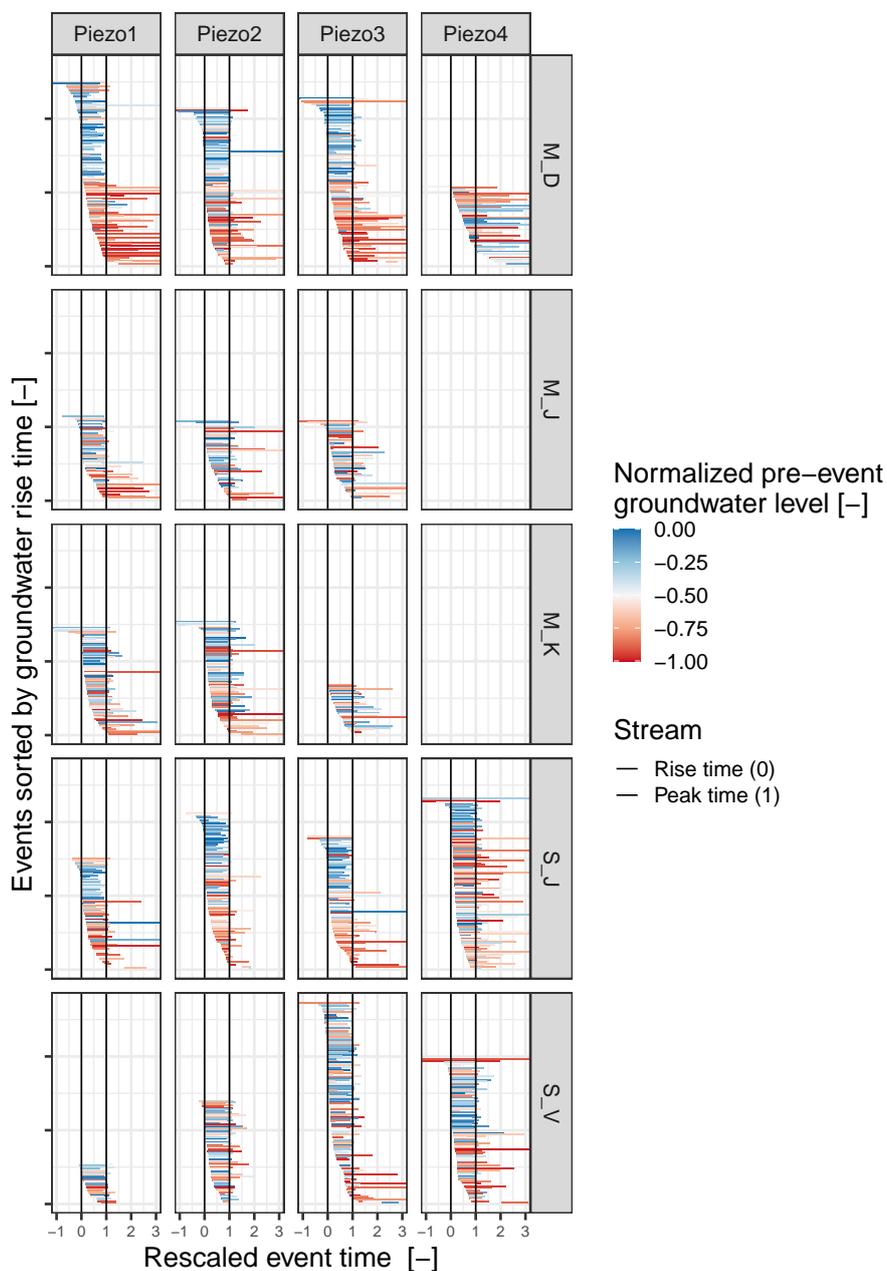
Figure 8. Timing of groundwater response relative to stream response. The two vertical lines at 0 and 1 represent the normalised rise and maximum time of each individual stream flow event. Horizontal bars each represent a groundwater event with its individual normalised rise and maximum time. Bluish colours indicate high and reddish colours indicate low pre-event groundwater levels.
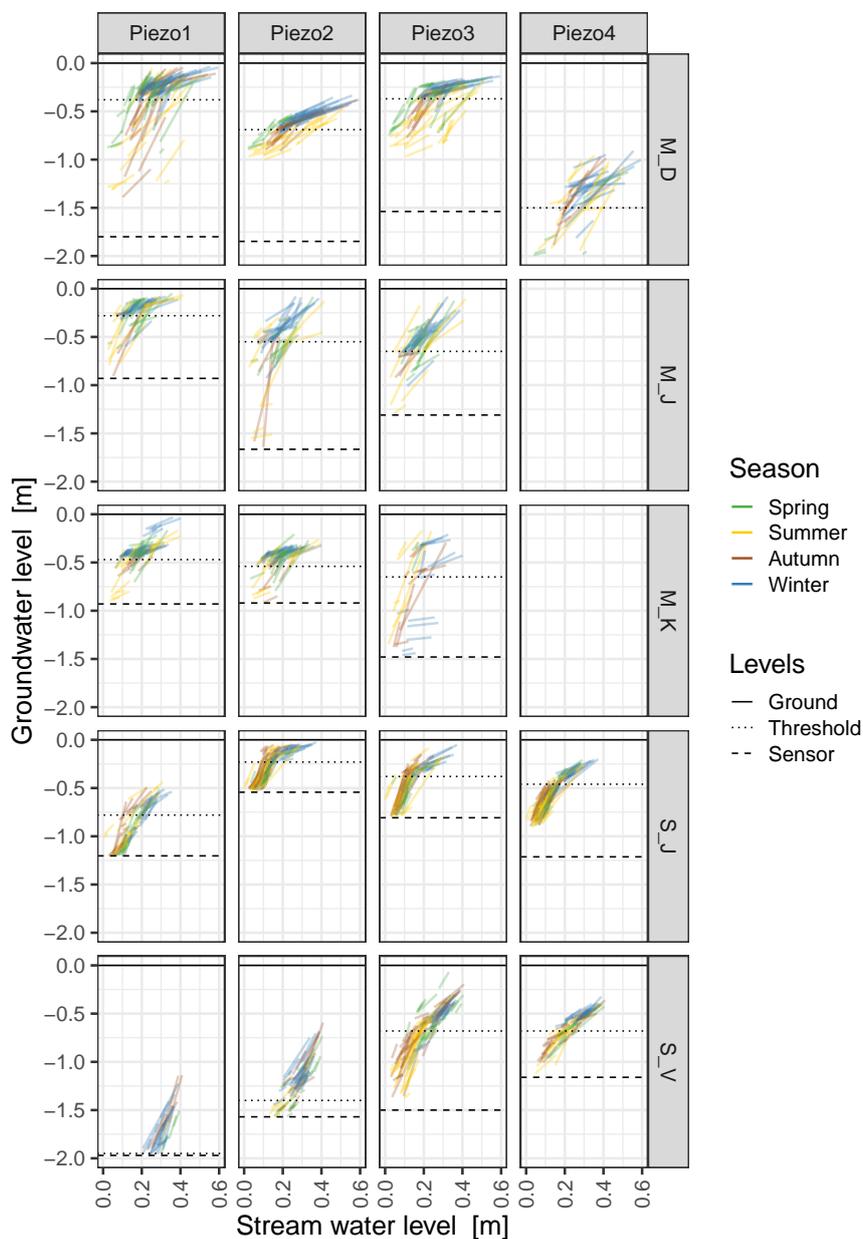
Figure 9. Responses in water tables of groundwater (y-axis) and the corresponding stream (x-axis). The lines connect the pre-event minimum with the event maximum. Note that these water levels do not necessarily occur at the same point in time as this visualisation removes the temporal dimension. The y-axis of the plot ends at a depth of 2m for purpose of comparison, thus omitting 2 events occurring below this groundwater level at M_D_Piezo4. Dotted horizontal lines illustrate the threshold between the lower (more variable but mainly steep sloping lines) and upper (less variable with shallow slopes) hydrologic response behaviour.

pre-event groundwater levels rise above the threshold. For the two forest sites in the marls region – M_J and M_K – the pattern is less clear, with some larger runoff coefficients also occurring below the threshold. A separation with regards to relative response timing (red vs blue) can be observed at M_D and S_J where groundwater responds before the stream for most events above the pre-event water level threshold. At the other three sites M_J, M_K and S_V no clear distinction can be made.

3.6 Catchment state

When looking at the dynamics of catchment state across all events at all sites (see Figure 11) the general pattern clearly shows a common shift in hydrologic connectivity with higher probabilities of catchment states above the threshold from late autumn until early spring. However, below threshold states can occur in winter (see for example the winter of 2016) and
10 above threshold states can also occur in summer (see for example summer of 2014). There is no clear distinction between the geological regions but there are periods where system state varies across the different sites (e.g. fall 2014). However, for most events the below/above threshold state identification is similar across many piezometers.

4 Discussion

4.1 Event detection

15 The events summarised in Figures 5 and 6 allow us to identify erratic sensors but also reveal topographic characteristics of the various sites. Topography can explain the occurrence of *dry* events, with a deeply incised stream at M_J and M_K, where we observe the lowest fraction of complete events in the groundwater with 50% or less of the streamflow events, and the steep hillslope at S_V leading to a gradient in water level depths and thus differing responses among the piezometers as well as seasonally more strongly fluctuating groundwater levels. The low numbers of *partial* events at sites with high numbers
20 of *complete* and *dry* events (M_J and M_K) signal that the seasonal transition between low and high groundwater levels is very abrupt, skipping intermediate levels. This might be due to pronounced capillarity fringes reaching into the very shallow subsurface. In that case, infiltrating water would reach the upper end of the fringe very quickly and only little water volume would be necessary to lift the groundwater level significantly (e.g. Cloke et al. (2006)).

4.2 Before-event hillslope-stream connectivity

25 Cross-correlation has been used in previous studies to assess different aspects of hydrologic connectivity, such as lag time analysis between stream and groundwater table (Allen et al., 2010; Bachmair and Weiler, 2014), relating water table connectivity to topographic indices (Jencso et al., 2009) or comparing groundwater levels with runoff coefficients (Seibert et al., 2003). Assuming well coupled hydrologic systems, high correlation coefficients would be expected which applies to the two schist sites. Low correlation coefficients indicate a streamflow (baseflow) response that is decoupled from the groundwater. This applies
30 to all three marls sites. Visually comparing the point cloud patterns of the piezometers at each single site (Figure 7) reveals a
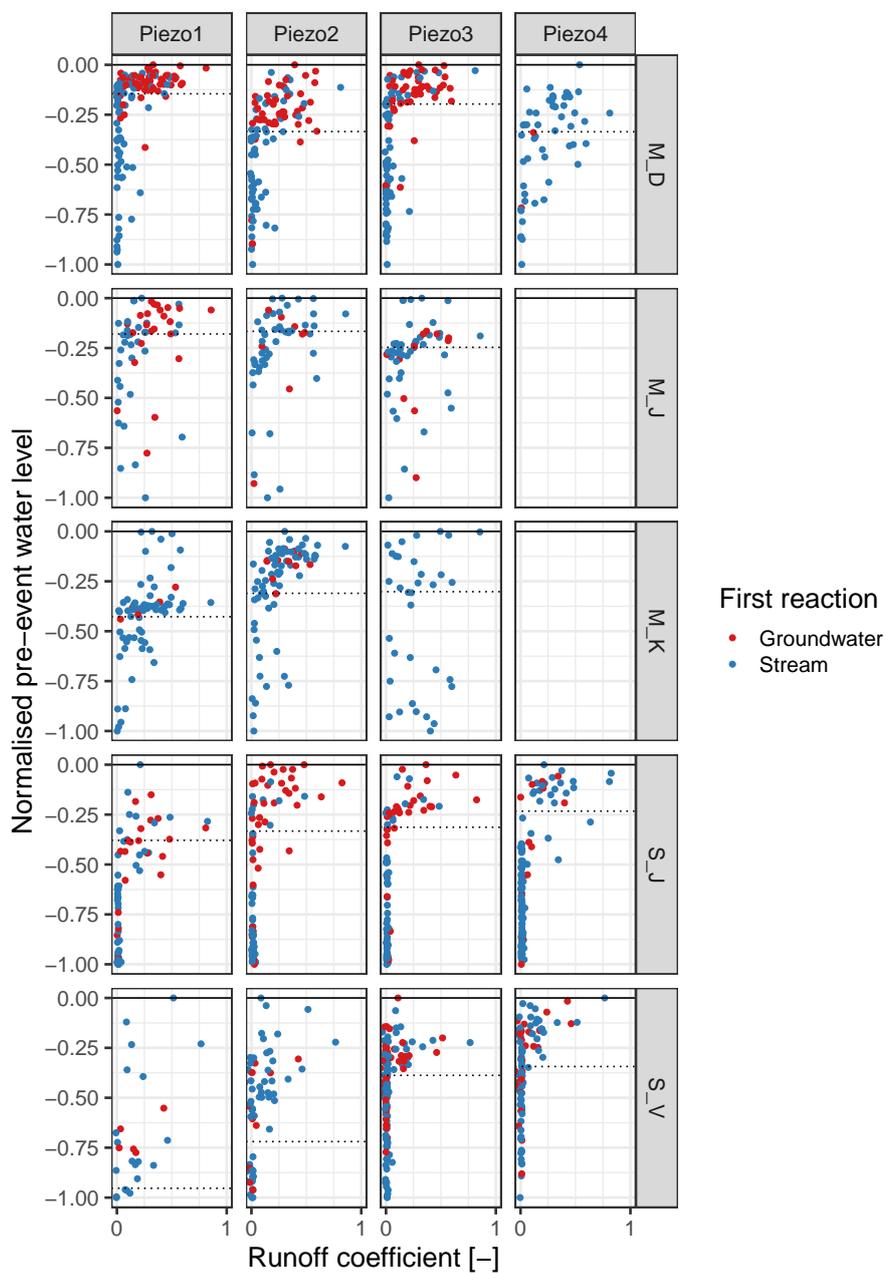
Figure 10. Event runoff coefficients versus shallow groundwater levels. Event runoff coefficients were determined for the Weierbach (schist) and Wollefsbach (marls) catchments where discharge data is available. The dotted horizontal lines illustrate the individual thresholds obtained from Figure 9. The point colours indicate whether the groundwater levels responded first (red) or the stream (blue).
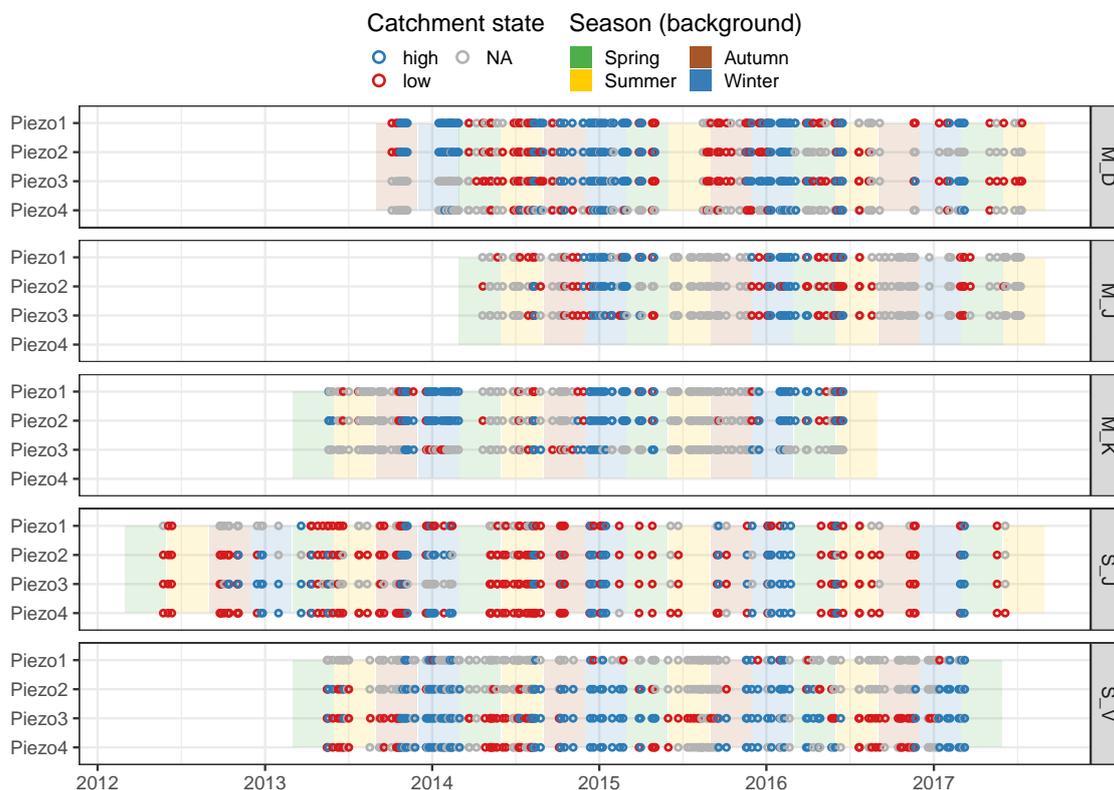
Figure 11. Catchment states at the beginning of events. In contrast to Figure 5 it shows whether or not the groundwater levels are above (high/blue) or below (low/red) the locally defined groundwater threshold levels.Only complete events are included, all other events are shown in grey (NA).

high site-internal similarity (a site-specific fingerprint) among the piezometers (with the previously mentioned exceptions of M_D_Piezo4 and M_K_Piezo3) which indicates that well-placed groundwater observation points can be representative for the given footslope, at least for pre-event conditions. The observed differences between the geologies suggest that soil texture and bedrock structure might control regional similarities.

4.3 Comparison of relative response timing between stream and groundwater

Identical response timing or groundwater rising and peaking just before the stream suggests that hillslope groundwater is driving streamflow response and thus that hillslope-stream connectivity is high (Haught and Meerveld, 2011; Rinderer et al., 2016). That this occurs under high groundwater levels further supports this conclusion. Groundwater rising and peaking after streamflow indicates that streamflow response is probably not caused by hillslope shallow groundwater and that hillslope-stream subsurface connectivity is low. The depth of groundwater level (and its capillary fringe) directly influences the delay in groundwater response as this depth defines the infiltration distance. During events with low groundwater levels, precipitation

falling onto or very close to the stream might generate a rise in the stream before the groundwater response (McGuire and McDonnell, 2008). Triggering an early response in stream compared to groundwater can also be the result of infiltration excess overland flow where surface runoff connects faster to the stream than it infiltrates towards the groundwater. However, this can be ruled out for schist as the high infiltration capacity makes overland flow unlikely, while it can not be ruled out for the clayey

5   soils in the marls region (Wrede et al., 2015). No pronounced differences in timing can be observed between marls and schist. The large variability in response timing confirms the need for monitoring over extended time periods as few or single event analyses run the risk of not being representative. Temporal relationship and water level responses are intertwined in a time series which makes it very intricate focusing on one while looking at both at the same time, e.g. by plotting two time series against each other and interpreting the resulting hysteresis (Kendall et al., 1999; McGuire and McDonnell, 2010; Zuecco et al.,

10   2016). Choosing to separate the analysis of the temporal response from the water level changes allowed to better reveal the temporal relationship of the hillslope-stream system on the one hand, and water level changes on the other.

4.4   Event-induced increase in stream and groundwater levels

Previous studies observed transmissivity feedback as a key mechanism controlling subsurface runoff (Bishop et al., 2011; Detty and McGuire, 2010b). Transmissivity feedback has previously been observed directly via piezometers (Bishop et al., 2011) or

15   indirectly through stable isotope composition in stream runoff (Bishop et al., 2004; Laudon et al., 2004) and tracer transport rates (Laine-Kaulio et al., 2014). In our study the capped response of groundwater events above a certain threshold is a strong indication of transmissivity feedback as one controlling mechanism (M_D and M_K). At low groundwater levels, infiltrating water results in a substantial increase of the groundwater level, suggesting that lateral conductivities are low as water is added more quickly than it can flow away laterally. This changes when the water level reaches a certain level or soil horizon. Now

20   infiltrating water is no longer increasing groundwater level substantially but instead fast lateral transport is likely to be causing the observed pronounced rise in stream water levels. This sudden fast lateral transport of the shallow groundwater is likely due to substantially higher lateral hydraulic conductivity of the upper soil horizons compared to the lower soil horizons. This fits well with the findings by Sprenger et al. 2016 who at site M_K found a strong increase in saturated hydraulic conductivity by a factor of 40 at a depth of 36 cm, while the increase for site M_J (where we did not observe a strong capping of the response)

25   is less than 20% (the other 3 sites were unfortunately not included in the analysis by Sprenger et al.). A raise in the hydraulic gradient in a more uniform depth profile of hydraulic conductivities, on the other hand, would only lead to a gradual increase in lateral flow. At S_V transmissivity feedback does not seem to occur as the slopes of the lines do not change as abruptly (Figure 9). This is in accordance with the findings of Angermann et al. (2017) at the same hillslope: During sprinkling experiments they observed that relatively high vertical and lateral hydraulic conductivities ($10^{-3} m/s$) lead to fast lateral responses in subsurface.

30   The high variability of the trajectories for events with low pre-event groundwater levels can be explained by a higher variability in possible flow paths compared to conditions of high groundwater levels. Another possibility is the complete disconnection between hillslope and stream which allows varying stream water levels to occur with similar groundwater levels. Whether or not a precipitation event can activate certain flow paths depends on the spatial distribution of pre-event water and the characteristics of bedrock topography (van Meerveld et al., 2015). Demand et al. (2019) found that preferential flow is present in particular

during dry conditions. When the groundwater level is high, the majority of flow paths is already activated and the degrees of freedom to activate new flow paths are limited. Therefore, the relation between stream and groundwater converges and shifts from variable to more uniform (Figure 9). Investigating rainfall characteristics and their effect on event responses did not help explaining the underlying mechanisms. While correlation coefficients between precipitation sum and $h_{preAmplitude}$ reached values of 0.5 and above, they stayed below 0.4 for $t_{riseInterval}$ showing that the precipitation had no clearly identifiable effect on event timing.

4.5 Runoff coefficient

Threshold behaviour is a common observation in runoff generation (Ali et al., 2013), for example Scaife and Band (2017) and Detty and McGuire (2010b) observed a threshold effect of antecedent precipitation and soil moisture on stormflow, and Latron and Gallart (2008) identified a threshold behaviour between groundwater level and runoff coefficient depending on seasonal catchment conditions (dry, wetting-up and wet). In our study the groundwater threshold marking the change in event runoff coefficients (Figure 10) coincides with the regime shift of water table responses (Figure 9). At M_D, S_J and S_V the pattern is very similar: below the individual pre-event groundwater thresholds runoff coefficients are very small, but increase significantly both in value as well as in variability when pre-event groundwater levels rise above the threshold. For the two forest sites in the marls region – M_J and M_K – the pattern is less clear, with some larger runoff coefficients also occurring below the threshold. A possible explanation could be that Wollefsbach gauge used to determine the runoff coefficients is less representative for these forest sites, as the Wollefsbach Catchment consists almost entirely of pasture and agricultural areas. In addition, the morphology of slopes and stream channel at the two marls forest sites is very distinct (and different to the Wollefsbach), with very low gradients in the slopes but a deeply incised stream bed. As the probability of high runoff coefficients increases above the groundwater threshold it seems that local observations of groundwater levels can give a good indication of catchment state with respect to connectivity and storage and release behaviour. This is true even for neighbouring catchments within the same geological region (M_D and S_V for example are not located in or downstream of the catchments used for the determination of the runoff coefficients). We also find that especially the regime shift and the corresponding threshold can be more clearly identified by groundwater level observations than by antecedent stream water level (Figure 9). This implies that near-stream groundwater observations hold significant predictive power to estimate whether or not an upcoming precipitation event is likely to produce major runoff at the outlet of the subcatchment.

4.6 Catchment state

The previously obtained groundwater thresholds allow us to split all events into two groups: Events with catchment states above the threshold are considered to generate substantial lateral subsurface stormflow and therefore have high hillslope-stream connectivity, while catchment states below the threshold generate only minor lateral flow and hillslope-stream connectivity is not well-established or does not extend far up the slopes. Just taking season as a predictor for the expected event response and hillslope-stream connectivity would be too simple: while summer events are likely to be below threshold and winter events above, this is not a general rule and spring and fall events can also not be classified just by their season (Figure 11). However,

Table 2. Observations and corresponding process interpretations

Observation	Process interpretation
low correlation of pre-event stream and groundwater levels (Figure 7)	low or only temporary hillslope-stream connectivity (M_D, M_J, M_K)
stream response and peak prior to groundwater response and peak, mainly under dry conditions (Figure 8)	runoff generated by near-stream overland flow, unsaturated zone preferential flow or direct rainfall (M_D, S_J, S_V)
at high groundwater levels: little to no event-induced increase in groundwater levels but high increase in stream water levels (Figure 9)	transmissivity feedback (M_D, M_K) and fill-and-spill (S_J)
Schist: very low runoff coefficients at low groundwater levels. Marls: higher runoff coefficients also occur at low groundwater levels (Figure 10)	different processes are active in the two geologies at low groundwater levels, surface runoff or preferential flow paths above the shallow groundwater table can produce significant runoff in the marls
Marls: groundwater levels cluster at high values, only few data points at low levels, few points in-between (Figure 9 and Figure 10)	groundwater ridging due to capillary fringe effects in the clayey soils

our study results suggest that a single well chosen piezometer can already provide substantial information on catchment state and the potential for high connectivity and thus high runoff events.

4.7 Synthesis: Process deductions

The joint analysis of shallow near-stream groundwater and stream water levels allows us to identify several runoff generation
5   mechanisms. Observations and the corresponding interpretations are listed in Table 2. The observations described in Table 2 require a large number of events. Only if the number of events is sufficiently high we can capture the variability in responses, the frequency of different response types, the dominant responses and then interpret the underlying processes. The 150 events identified for this study are relatively high compared to many studies (e.g. Detty and McGuire (2010b) 15 events / 3 months;Ali et al. (2011) 50 events / 1 year; Penna et al. 2015 63 events / 3 years; Anderson et al. (2010) 99 events / 19 months; van Meerveld
10   and McDonnell (2006a) 147 events / 2 years 4 months; and, exceptionally: Scaife and Band (2017) 811 events / 15 years). Events in marls cluster at high pre-event groundwater levels with 60 to 80% of events found in the upper half of the total range and only few events at low levels or in-between. At the same time the piezometers at M_J and M_K experience a considerably high number of dry events (Figure 6). Groundwater transitions fast from very low levels to levels near the surface, with only few events in-between. This fast transition hints towards extended capillary fringes where only low volumes of water are necessary
15   to rise the groundwater table (Cloke et al., 2006). As a result of the transmissivity feedback, runoff coefficients significantly

increase when groundwater levels reach the threshold as the hillslope connects to the stream (Figure 10. This behaviour can be observed in particular at the largest catchment in marls (M_D) with an undulating landscape and mostly pasture and to a lesser degree at smaller catchments with very flat topography and forest. As several characteristics are different between this catchments, this behaviour can not be assigned to one single attribute with confidence. In schist, events are spread over

5   the whole range of pre-event groundwater levels with no clear difference between low, in-between and high events. Since hydraulic conductivities in schist are generally very high, the sudden increase in runoff coefficient above the threshold can not be explained by transmissivity feedback being the governing process. Nevertheless, capping of groundwater response was observed at S_J. Anderson et al. (2010) found that in watersheds with lateral preferential flow the fill-and-spill mechanism was responsible for capped groundwater responses. This observation can be transferred to the schist site to explain the inhibited

10 groundwater response making its soil-bedrock interface responsible for the threshold relationship.

5   Conclusions

We found that a multi-event analysis approach including a large number of events is suitable for characterising the hydrologic response behaviour of the hillslope-stream-system and the dynamics of its connectivity. A more selective and exemplary analysis of only a few events (e.g. one for winter and one for summer) would lead to misinterpretation of the results. If these two

15 events represented the two extremes (high and low pre-event water levels) one would need to assume a functional relation for all the potential events in-between (which is likely to be flawed). Detecting threshold behaviour and identifying the correct threshold would be very unlikely and the lack of information on event variability would significantly reduce the confidence of the findings. Our analyses identified patterns that are representative for the site or hillslope, i.e. which were shown by all or most piezometers at these sites. However, piezometers can also be located at points where very localised features drastically

20 influence the response behaviour which is why at least two, or better three, piezometers should be used when investigating hillslope-stream relation to secure redundant information and identify the most representative and informative monitoring location for the hillslope or even catchment. Then a single, well chosen, piezometer can already provide substantial information on catchment state and the potential for high connectivity and thus high runoff events.

Differences between the response behaviour of the two geologies were less pronounced than expected for some of the analyses,

25 but the observed results showed that both hydrologic systems are subject to a threshold behaviour where dominating hydrologic processes change when the system state variable (groundwater level) crosses a certain threshold. Above this threshold larger increases in stream water levels and also larger runoff coefficients are observed. While both geologies show this threshold behaviour the underlying processes are likely to be different, with transmissivity feedback occuring in the marls and a more fill-and-spill-like process in the schist. The fact that at low groundwater levels runoff coefficients in the marls tend to be higher

30 than in the schist, in some cases even by an order of magnitude, suggests that also at low groundwater levels different processes are active in the two geological regions and that in the marls surface runoff or lateral preferential flow above the shallow groundwater must provide sufficient connectivity to enable runoff generation while saturated subsurface connectivity is low. The proposed separation of the temporal component and the extent of water level responses for certain aspects of the data anal-

ysis proved to be useful in visualising, analysing and interpreting the event response and its variability across a large number of events. Even though the installation and monitoring of piezometers in the near-stream zone is pragmatic and much less cost- and labour-intensive than the installation of hillslope trenches, local near-stream shallow groundwater observations do hold significant predictive power for the potential catchment response. They possibly provide more information than piezometer- or

5 trench observations located further upslope would, as the footslope and riparian zone are both link and gate-keeper, controlling connectivity between hillslopes and streams. Due to the lower cost of piezometer installation and monitoring compared to trenches it is possible to instrument a larger number of sites which in turn makes it possible to systematically investigate subsurface hillslope-stream connectivity in different hydrologic response units instead of focusing on within-slope connectivity on single hillslopes. While we focused on 5 hillslopes in this study it would easily be possible to extend this monitoring design

10 to a larger number of sites thus even better capturing the spatial variability in responses and allowing a thorough investigation into which sites tend to be most representative of the catchment and if these sites can be identified a-priori based on topography or other landscape characteristics. The application of our data analysis to other sites where data is already available might open up new ways of systematic site-intercomparison as our analysis provides a novel way of visualizing event responses and thus making the information contained in a large number of events more easily accessible.

15 *Competing interests.* None of the authors have financial competing interests, however, authors Weiler and Blume are members of the editorial board of the journal.

Data availability. Data will be made available in a corresponding data publication in ESSD.

Acknowledgements. We gratefully acknowledge DFG for the research funding (FOR 1598) and thank the group of Laurent Pfister at the Luxembourg Institut of Technology (LIST) for providing the discharge data, and Tobias Vetter and Britta Kattenstroth for their untiring

20 maintenance of the monitoring sites.

References

Ali, G., Oswald, C. J., Spence, C., Cammeraat, E. L., McGuire, K. J., Meixner, T., and Reaney, S. M.: Towards a unified threshold-based hydrological theory: necessary components and recurring challenges, Hydrological Processes, 27, 313–318, 2013.

Ali, G. A., L'Heureux, C., Roy, A. G., Turmel, M.-C., and Courchesne, F.: Linking spatial patterns of perched groundwater storage and storm-flow generation processes in a headwater forested catchment, Hydrological Processes, 25, 3843–3857, https://doi.org/10.1002/hyp.8238, http://dx.doi.org/10.1002/hyp.8238, 2011.

Allen, D. M., Whitfield, P. H., and Werner, A.: Groundwater level responses in temperate mountainous terrain: regime classification, and linkages to climate and streamflow, Hydrological Processes, 24, 3392–3412, https://doi.org/10.1002/hyp.7757, http://dx.doi.org/10.1002/hyp.7757, 2010.

Anderson, A. E., Weiler, M., Alila, Y., and Hudson, R. O.: Piezometric response in zones of a watershed with lateral preferential flow as a first-order control on subsurface flow, Hydrological Processes, 24, 2237–2247, https://doi.org/10.1002/hyp.7662, http://dx.doi.org/10.1002/hyp.7662, 2010.

Anderson, S. P., Dietrich, W. E., Montgomery, D. R., Torres, R., Conrad, M. E., and Loague, K.: Subsurface flow paths in a steep, unchanneled catchment, Water Resources Research, 33, 2637–2653, https://doi.org/10.1029/97WR02595, http://dx.doi.org/10.1029/97WR02595, 1997.

Angermann, L., Jackisch, C., Allroggen, N., Sprenger, M., Zehe, E., Tronicke, J., Weiler, M., and Blume, T.: Form and function in hillslope hydrology: characterization of subsurface flow based on response observations, Hydrology and Earth System Sciences, 21, 3727–3748, https://doi.org/10.5194/hess-21-3727-2017, https://www.hydrol-earth-syst-sci.net/21/3727/2017/, 2017.

Bachmair, S. and Weiler, M.: Interactions and connectivity between runoff generation processes of different spatial scales, Hydrological Processes, 28, 1916–1930, https://doi.org/10.1002/hyp.9705, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.9705, 2014.

Bishop, K., Seibert, J., Köhler, S., and Laudon, H.: Resolving the double paradox of rapidly mobilized old water with highly variable responses in runoff chemistry, Hydrological processes, 18, 185–189, 2004.

Bishop, K., Seibert, J., Nyberg, L., and Rodhe, A.: Water storage in a till catchment. II: Implications of transmissivity feedback for flow paths and turnover times, Hydrological Processes, 25, 3950–3959, 2011.

Blume, T. and van Meerveld, H. I.: From hillslope to stream: methods to investigate subsurface connectivity, Wiley Interdisciplinary Reviews: Water, 2, 177–198, https://doi.org/10.1002/wat2.1071, https://onlinelibrary.wiley.com/doi/abs/10.1002/wat2.1071, 2015.

Bracken, L., Wainwright, J., Ali, G., Tetzlaff, D., Smith, M., Reaney, S., and Roy, A.: Concepts of hydrological connectivity: Research approaches, pathways and future agendas, Earth-Science Reviews, 119, 17 – 34, https://doi.org/10.1016/j.earscirev.2013.02.001, http://www.sciencedirect.com/science/article/pii/S0012825213000202, 2013.

Cloke, H., Anderson, M., McDonnell, J., and Renaud, J.-P.: Using numerical modelling to evaluate the capillary fringe groundwater ridging hypothesis of streamflow generation, Journal of Hydrology, 316, 141–162, 2006.

Demand, D., Blume, T., and Weiler, M.: Relevance and controls of preferential flow at the landscape scale, Hydrology and Earth System Sciences Discussions, 2019, 1–37, https://doi.org/10.5194/hess-2019-80, https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-80/, 2019.

Detty, J. M. and McGuire, K. J.: Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment, Hydrological Processes, 24, 2222–2236, https://doi.org/10.1002/hyp.7656, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.7656, 2010.

Detty, J. M. and McGuire, K. J.: Threshold changes in storm runoff generation at a till-mantled headwater catchment, Water Resources Research, 46, https://doi.org/10.1029/2009WR008102, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008102, 2010b.

Freer, J., McDonnell, J. J., Beven, K. J., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B., and Kendall, C.: The role of bedrock topography on subsurface storm flow, Water Resources Research, 38, 5–1–5–16, https://doi.org/10.1029/2001WR000872, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001WR000872, 2002.

Graham, C. B., Woods, R. A., and McDonnell, J. J.: Hillslope threshold response to rainfall: (1) A field based forensic approach, Journal of Hydrology, 393, 65 – 76, https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.12.015, http://www.sciencedirect.com/science/article/pii/S0022169409007938, soil Architecture and Preferential Flow across Scales, 2010.

Haught, D. R. W. and Meerveld, H. J.: Spatial variation in transient water table responses: differences between an upper and lower hillslope zone, Hydrological Processes, 25, 3866–3877, https://doi.org/10.1002/hyp.8354, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.8354, 2011.

Jackson, C. R. and Pringle, C. M.: Ecological Benefits of Reduced Hydrologic Connectivity in Intensively Developed Landscapes, BioScience, 60, 37–46, https://doi.org/10.1525/bio.2010.60.1.8, https://doi.org/10.1525/bio.2010.60.1.8, 2010.

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and Marshall, L. A.: Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale, Water Resources Research, 45, https://doi.org/10.1029/2008wr007225, http://onlinelibrary.wiley.com/doi/10.1029/2008WR007225/abstract, w04428, 2009.

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Bencala, K. E., and Wondzell, S. M.: Hillslope hydrologic connectivity controls riparian groundwater turnover: Implications of catchment structure for riparian buffering and stream water sources, Water Resources Research, 46, https://doi.org/10.1029/2009WR008818, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008818, 2010.

Kendall, K., Shanley, J., and McDonnell, J.: A hydrometric and geochemical approach to test the transmissivity feedback hypothesis during snowmelt, Journal of Hydrology, 219, 188 – 205, https://doi.org/https://doi.org/10.1016/S0022-1694(99)00059-1, http://www.sciencedirect.com/science/article/pii/S0022169499000591, 1999.

Krause, S., Blume, T., and Cassidy, N. J.: Investigating patterns and controls of groundwater up-welling in a lowland river by combining Fibre-optic Distributed Temperature Sensing with observations of vertical hydraulic gradients, Hydrology and Earth System Sciences, 16, 1775–1792, https://doi.org/10.5194/hess-16-1775-2012, https://www.hydrol-earth-syst-sci.net/16/1775/2012/, 2012.

Laine-Kaulio, H., Backnäs, S., Karvonen, T., Koivusalo, H., and McDonnell, J. J.: Lateral subsurface stormflow and solute transport in a forested hillslope: A combined measurement and modeling approach, Water Resources Research, 50, 8159–8178, 2014.

Latron, J. and Gallart, F.: Runoff generation processes in a small Mediterranean research catchment (Vallcebre, Eastern Pyrenees), Journal of Hydrology, 358, 206 – 220, https://doi.org/https://doi.org/10.1016/j.jhydrol.2008.06.014, http://www.sciencedirect.com/science/article/pii/S0022169408002837, 2008.

Laudon, H., Seibert, J., Köhler, S., and Bishop, K.: Hydrological flow paths during snowmelt: Congruence between hydrometric measurements and oxygen 18 in meltwater, soil water, and runoff, Water resources research, 40, 2004.

McGlynn, B. L. and McDonnell, J. J.: Quantifying the relative contributions of riparian and hillslope zones to catchment runoff, Water Resources Research, 39, https://doi.org/10.1029/2003wr002091, http://onlinelibrary.wiley.com/doi/10.1029/2003WR002091/abstract, 1310, 2003.

McGuire, K. and McDonnell, J.: Stable Isotope Tracers in Watershed Hydrology, chap. 11, pp. 334–374, John Wiley & Sons, Ltd, https://doi.org/10.1002/9780470691854.ch11, https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470691854.ch11, 2008.

McGuire, K. J. and McDonnell, J. J.: Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlineari-ties, Water Resources Research, 46, https://doi.org/10.1029/2010WR009341, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009341, 2010.

Ocampo, C. J., Sivapalan, M., and Oldham, C.: Hydrological connectivity of upland-riparian zones in agricul-tural catchments: Implications for runoff generation and nitrate transport, Journal of Hydrology, 331, 643 – 658, https://doi.org/https://doi.org/10.1016/j.jhydrol.2006.06.010, http://www.sciencedirect.com/science/article/pii/S0022169406003234, 2006.

Oswald, C. J., Richardson, M. C., and Branfireun, B. A.: Water storage dynamics and runoff response of a boreal Shield headwater catchment, Hydrological Processes, 25, 3042–3060, https://doi.org/10.1002/hyp.8036, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.8036, 2011.

Penna, D., Mantese, N., Hopp, L., Dalla Fontana, G., and Borga, M.: Spatio-temporal variability of piezometric response on two steep alpine hillslopes, Hydrological Processes, 29, 198–211, https://doi.org/10.1002/hyp.10140, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10140, 2015.

Rinderer, M., van Meerveld, I., Stähli, M., and Seibert, J.: Is groundwater response timing in a pre-alpine catchment controlled more by topography or by rainfall?, Hydrological Processes, 30, 1036–1051, 2016.

Scaife, C. I. and Band, L. E.: Nonstationarity in threshold response of stormflow in southern Appalachian headwater catchments, Water Resources Research, 53, 6579–6596, https://doi.org/10.1002/2017WR020376, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR020376, 2017.

Seibert, J., Bishop, K., Rodhe, A., and McDonnell, J. J.: Groundwater dynamics along a hillslope: A test of the steady state hypoth-esis, Water Resources Research, 39, https://doi.org/10.1029/2002wr001404, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002WR001404, 2003.

Sprenger, M., Seeger, S., Blume, T., and Weiler, M.: Travel times in the vadose zone: Variability in space and time, Water Resources Research, 52, 5727–5754, https://doi.org/10.1002/2015wr018077, http://dx.doi.org/10.1002/2015WR018077, 2016.

Thompson, J., Cassidy, R., Doody, D. G., and Flynn, R.: Predicting critical source areas of sediment in headwater catchments, Agriculture, ecosystems & environment, 179, 41–52, 2013.

van Meerveld, H. J., Seibert, J., and Peters, N. E.: Hillslope–riparian-stream connectivity and flow directions at the Panola Mountain Research Watershed, Hydrological Processes, 29, 3556–3574, https://doi.org/10.1002/hyp.10508, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10508, 2015.

van Meerveld, H. J. T. and McDonnell, J. J.: Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hill-slope, Water Resources Research, 42, https://doi.org/10.1029/2004wr003778, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004WR003778, 2006a.

van Meerveld, H. J. T. and McDonnell, J. J.: Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis, Water Resources Research, 42, https://doi.org/10.1029/2004wr003800, http://dx.doi.org/10.1029/2004WR003800, w02411, 2006b.

Wrede, S., Fenicia, F., Martínez-Carreras, N., Juilleret, J., Hissler, C., Krein, A., Savenije, H. H. G., Uhlenbrook, S., Kavetski, D., and Pfister, L.: Towards more systematic perceptual model development: a case study using 3 Luxembourgish catchments, Hydrological Processes, 29, 2731–2750, https://doi.org/10.1002/hyp.10393, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10393, 2015.

Zehe, E., Ehret, U., Pfister, L., Blume, T., Schröder, B., Westhoff, M., Jackisch, C., Schymanski, S. J., Weiler, M., Schulz, K., All-roggen, N., Tronicke, J., van Schaik, L., Dietrich, P., Scherer, U., Eccard, J., Wulfmeyer, V., and Kleidon, A.: HESS Opinions: From

response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments, Hydrology and Earth System Sciences, 18, 4635–4655, https://doi.org/10.5194/hess-18-4635-2014, https://www.hydrol-earth-syst-sci.net/18/4635/2014/, 2014.

Zuecco, G., Penna, D., Borga, M., and van Meerveld, H.: A versatile index to characterize hysteresis between hydrological variables at the

5   runoff event timescale, Hydrological Processes, 30, 1449–1466, 2016.