Interactive comment on 'In situ investigation of rapid subsurface flow: Temporal dynamics and catchment-scale implication' by L. Angermann et al.

Response to comments by reviewer#1

We highly appreciate the thorough review and constructive comments of the anonymous reviewer. We are happy to implement and discuss the cogent remarks on data analysis and presentation of our findings and developed a concept to tackle the basic criticism regarding the story line and the conclusiveness of our results. This concept is based on the comments by all three reviewers and provides a common theme as well as overarching yet standalone story lines for the two companion papers. We will shortly present the main idea of the new concept, the revised hypotheses as well as the restructured story line of the first manuscript (MS1). As these improvements are relevant for our response to the comments by all reviewers, this part is included in all replies. Afterwards, we will address the general and specific comments made by reviewer#1, with our replies inserted in the original review.

Conceptual framework

To better elaborate the methodological aspects of the study and to provide a common theme for the two companion papers, we want to employ the concept of form vs. function. The original term 'form follows function' was first established in architecture and soon was adopted by biologists. It refers to the idea, that form and functionality are closely correlated, influence each other and co-evolve. We suggest to transfer the same idea to hydrological systems. This allows us to separate and analyze their two main characteristics: Their form, which is equivalent to the spatial structure and static properties, and their function, equivalent to internal responses and hydrological behavior. While this approach itself is not particularly new to hydrological field research, we want to employ this concept to explicitly pursue the question of what information is most advantageous to understand a hydrological system.

Accordingly, we developed different categories to organize and describe the data presented in the two manuscripts: Structural data summarizes all sorts of data which focus on direct exploration of form, e.g. soil cores. Response dynamics, on the other hand, are observations of function. They represent processes deprived of their spatial context and include soil moisture dynamics and discharge responses. In between these two categories are flow-relevant structures and response patterns, which may contain information on both, form and function of the system.

In the presented study, we apply this concept to subsurface flow within a hillslope. The first part of the study (MS1) methodologically focuses on function: We observed response patterns and dynamics from a natural rainfall event and during an irrigation experiment. The results are used to infer hydrological processes and the spatial organization of the monitored system. Based on these findings, the informative power and conclusiveness of the data will be discussed.

The second manuscript (MS2) focuses on form and starts off with a thorough structural exploration of the subsurface. It then proceeds towards observations of flow-relevant structures and response patterns and analyzes the information gain along this path.

Hypotheses and story line

The hypotheses of both manuscripts can be aligned according to the form/function framework and will clearly be stated at in the beginning. The hypotheses of MS1 will focus on the potential of response observations for hillslope hydrological field research and the application of time-lapse GPR measurements in this context:

- **H1.1** Response observations (discharge, TDR & GPR data) are sufficient to characterize subsurface flow ywithin the hillslope.
- **H1.2** Response patterns can be used to deduce flow-relevant structures in the subsurface.
- **H1.3** Time-lapse GPR measurements visualize subsurface flow dynamics and patterns and can replace hillslope trenches.

The story line will be streamlined and arranged along these hypotheses. This will help to make the manuscript easier to follow and to better elaborate the important and novel key points of our study. In the following, new or restructured sections are marked in brown.

1 Introduction

- Concept of form and function: form and function are the main defining features of a system. Applied to catchments, the concept describes any kind of spatial structure (from topography to macropores) as form, and the sum of all processes defining the hydrological behavior of the catchment as function. Both strongly influence and determine each other.
- Example subsurface flow at the hillslope: structures and heterogeneity control flow patterns and velocities and thus the occurrence of preferential flow. Which one is better suited to characterize a hillslope, form or function?
- Focus on function: What does it need to describe subsurface flow and preferential flow at the hillslope? How to observe subsurface flow processes? What do response patterns and dynamics tell us about form and function of a hillslope?
- Methodological challenge of preferential flow: former approaches at different scales: plot, hillslope- and catchment-scale.
- Hypotheses as stated above

2 Methods

- Study site description
- Hydrological response monitoring: Hydrograph and surface water isotopes
- Hillslope-scale irrigation experiment
  - Setup
  - Process monitoring
  - Piezometer isotope sampling
- Data analysis
  - TDR data analysis
  - GPR data analysis
  - Comparison natural event vs. irrigation: Distinguish the signals and calculate areal share of activated cross section
  - Response velocity calculation

3 Results

- Response to the natural rainfall
  - Hydrograph and surface water isotopes
  - Subsurface response patterns (green GPR reflection patterns)
- Irrigation experiment
  - Core area water balance
  - Soil moisture dynamics
  - 2D time-lapse GPR
  - Soil and piezometer isotopes
  - Combination of TDR and GPR
  - Response velocities
- Comparison of natural event and irrigation: Areal share and signal strength of GPR measurements before and after irrigation

4 Discussion

- Process interpretation
  - Interpretation of artificially induced response observations during and after the irrigation
  - Interpretation of the natural response observations after the rainfall event prior to the irrigation experiment
  - Identification of (flow-relevant) structures from response patterns
- Methodological discussion

- Conclusiveness of function observations without structural knowledge with regard to process identification and the characterization of a system
- Conclusiveness of function observations without structural knowledge with regard to transferability and regionalization of results
- Evaluation of GPR as trench replacement

5 Conclusions

- **H1.1** 'Response observations (discharge, TDR & GPR data) are sufficient to characterize subsurface flow within the hillslope.'

  → Processes can be identified and characterized without any concrete information about spatial structures. However, observations are limited in spatial (and temporal) resolution and interpretations beyond observation scale remain speculative.

- **H1.2** 'Response patterns can be used to deduce flow-relevant structures in the subsurface.'

  → Response patterns allowed to develop a conceptual description of the flow paths network, which is linked to subsurface structures. However, actual structural features, such as the deposit layer or the bedrock interface, could not be located.

- **H1.3** Time-lapse GPR measurements visualize subsurface flow dynamics and patterns and can replace hillslope trenches.'

  → Time-lapse GPR measurements lack the quantitative power and the direct link to structures (obtained by excavation) of trenches. Their spatial and temporal flexibility as well as their non-invasive character, however, are very advantageous for the investigation of highly dynamics and spatially distributed flow processes. Thus, the application of time-lapse GPR measurements in combination with soil moisture measurements are a powerful tool for the observation of hydrometric responses. Depending on the research question, the method can replace labor-intense trenches and even increase the observation density due to its spatial flexibility.

The titles of the two manuscripts will be adapted accordingly. They will be rephrased to emphasize the methodological aspects of the two papers, while keeping the focus on hillslope processes. The final versions of the titles are still subject to discussion. A possible suggestion for the title of this manuscript is:

FORM AND FUNCTION IN HILLSLOPE HYDROLOGY:
IN SITU CHARACTERIZATION OF SUBSURFACE FLOW BASED ON RESPONSE OBSERVATIONS

We hope to have given a good overview over the anticipated revisions of the manuscript. While the elaboration above was meant to provide the 'big picture', our answers to the first reviewer's general and specific comments will illustrate how this concept will help to mitigate the reviewer's concerns. In the following, our replies are inserted into the original text by reviewer#1 and marked in blue.

Answers to the general comments

This paper presents very interesting data from a plot/hillslope scale sprinkling experiment combining TDR observations with time-lapse GPR to investigate soil moisture dynamics. Sadly, the promised (in the title) investigation of subsurface flow and the link to catchment scale is not convincingly presented, interpretations are often overdrawn and not supported by the results, and the paper currently reads more like a patchwork of data and ideas and not like a well aligned story. During the read, it felt as if the authors were not sure of what to present and what the key message of the work can/should be.
To set the frame for a more clear story line, we elaborated a common theme for the two companion papers: The theme focuses on the informative value of observations of 'form and function' for the hydrological investigation of hillslopes. While the first manuscript (MS1, at hand) uses process observations (i.e. response patterns and dynamics) to characterize the investigated hillslope, the second manuscript (MS2) starts with structural exploration of the subsurface only and complements these with the observed responses. This frame allows us to formulate clear and distinct hypotheses for both manuscripts and to align the story lines along the overarching and coherent theme without relying on the other manuscript too much.

For MS1 this means a thorough revision of the methodological discussion: The general aspects of subsurface process monitoring as presented in the current version will be drastically shortened. Instead, the discussion will focus on the value and shortcoming of response observations and the lack of structural information. Furthermore, the potential of time-lapse GPR measurements will be discussed in this context.

Finally, we will restructure the presentation and discussion of the hydrological response data in a way to to more clearly elaborate the logical link between observations and conclusions. Namely, the current separation in hillslope-scale and catchment-scale flow dynamics will be transferred into a separation of observations of the irrigation experiment and of the natural rainfall event. These changes will help to focus on the relevant aspects of the study and the novelty of the presented approach.

The data is visually very well presented, but the actual analysis of the experimental data remains poor, eventually supplying temporal evolution of soil moisture (although the method combination gives an interesting depth distribution to 4 m below surface) and flow velocities for the first respond[a)]. Moreover, these velocities are not really used[b)]. From a water balance perspective it is interesting, if the occurring amounts of PF can be, by any means, linked to the generated runoff volume in events[c)]. I also think that the inclusion of one event based hydrograph separation is a little bit thin[d)].
a) In this manuscript (MS1), we originally focused on the temporal evolution of the responses and deliberately minimized the discussion of spatial response patterns, as this was to be discussed in the companion paper. While we would like to keep the general separation of the two parts, the foci will be adapted to the new concept of 'form and function' of a hydrological system. This allows us to discuss certain aspects of the response patterns in more detail in MS1 already, without impairing the story line and arguments of MS2. We will for example provide a quantitative comparison between the responses to the natural rain event and the irrigation.

b) The response velocities were calculated to evaluate the observations of the response to the natural rainfall event: If the experimentally quantified response velocities are also occurring in flow paths, which were observed under natural conditions, these flow processes could be linked to the observed runoff response. In the revised manuscript, we will elaborate the line of argumentation and discuss its explanatory power as well as the calculated response velocities in more detail.
c) Unfortunately, the GPR signal can not be translated into quantitative soil moisture changes. The structural similarity attributes indicate areas of structural changes and can only be used to obtain spatial patterns and qualitative estimates of soil moisture dynamics. Although strong dissimilarity speaks for large changes in soil moisture, a quantitative relationship cannot be derived. To comply with the reviewers remark, we will improve the interpretation of the GPR data, e.g. by quantifying the spatial portions of the GPR transects activated during by the natural rainfall event and the irrigation experiment. We will also include isotope data sampled from piezometers at the irrigation site, which provide a basis for the discussion of flow proceses within the hillslope and the link between hillslope observations and the hydrograph.
d) The one hydrograph was shown to compare it with the subsurface response we could (partially) observe with our GPR measurements after the event and prior to the irrigation. As we do not have information on subsurface response for other events, we did not consider to include more hydrographs, and referred to former studies instead, reporting similarly fast runoff responses in the area. We will check if the inclusion of more hydrographs will improve the story line of the revised manuscript.

Stated by authors, the runoff ratio of 4% during the rapid peak (that consists of event water) cannot be explainable by direct rainfall on saturated areas. Nevertheless, the experimental observations and the fact that rapid flow downslope of the sprinkling site can be detected do not, in my opinion, convincingly explain the rapid stream response that is observed during rainfall in the study catchment. And this is why: I think that the method of the sprinkling experiment, with the use of a sprinkling intensity of 30mm/h, is simply not appropriate for investigating if preferential flow reaches the stream and generates the hydrograph under natural conditions.

I do understand that the purpose was to initiate the preferential flow with this rate, but I disagree with the authors that any explanation for the stream response can be done in this way. Such high intensities will of course lead lateral flow out of the sprinkling site as soon non-vertically oriented preferential flow paths are filled, or some low permeable lense or layer is reached. It would be interesting to report the natural rainfall intensities in the area and for the particular shown runoff event. As only a total of around 20 mm occurred. I am wondering if natural intensities would be indeed able to establish, supply, and connect a preferential flow system from the hillslope to the stream under unsaturated conditions on that time scale. This would be very important, since initiation of preferential flow strongly depends on intensities. I do not doubt that PF flow might be active at the hillslope and connecting to the stream in events with 30 mm/h rainfall intensity over hours, but when do such events occur? I think that there are no data and results provided in this manuscript that showed that preferential flow at natural rainfall intensities, amounts, and duration can explain the observed behaviour.

We agree with the reviewer, that the observations of the irrigation experiment alone are not sufficient to extrapolate our findings to natural rainfall events. The conclusion, that the experimentally initiated preferential flow processes are also relevant under natural conditions, was mainly based on the strong subsurface response after the natural rain event seen in the GPR data. The stronger subsurface response despite lower intensity is an indicator, that lateral preferential flow occurs in the shallow subsurface of the hillslope during natural events. To better convey this line of argumentation, we will present and discuss the observations from the natural event and the irrigation event separately. We will also include the intensities of natural rain events in the area and quantitatively compare the observed subsurface responses with regard to the areal share of activated flow paths. Finally, we will discuss the validity of these finding more carefully under a methodological aspect.

Furthermore, early work of McDonnell and various follow up studies showed that much of the water in preferential flow actually is not event water. This adds further doubt to the authors' interpretation, since the observed peaks in the stream in the area are dominated by event water.

We will reference and discuss the work of McDonnell et al. and others reporting on that topic in the revised manuscript. We will also add data on the isotopic composition of water seeping into the installed piezometers. In the current version of the two companion papers, the data were presented in the second part (MS2) to discuss the interaction between preferential flow paths and matrix. In the revised version, the data will be shifted to MS1 to support the discussion of the event water and pre-event water fractions. We will furthermore clarify, that we cannot be sure that the first peak is to be attributed solely to lateral preferential flow.

Leaving the interpretation of the catchment scale interpretation and the lateral preferential flow aside, the paper does not provide much novelty from a process perspective. The depth distribution data of soil moisture is interesting from methodological perspective, but work on this was published previously (see the Allroggen references). The discussion about the processes is not supported by the results of the work, and the methodological discussion is overly long, repetitive, lacks nearly any citations, and I was not struck by a major scientific message. That just leaves a case study of a sprinkling experiment, which is - in the current form - not sufficient to be published in HESS. That said, the dataset is very interesting and could supply interesting insights, but a clear idea of what should be presented, and what conclusions can be supported by the data is needed.

We understand, that our results and conclusions were not convincingly presented in the current version of the manuscript. The comments of the three reviewers, however, are helpful in pointing out shortcomings in the story line and we are confident to convey our key points and the novelty of our results more clearly in the revised manuscript.

As stated earlier, we developed a common theme linking the two companion papers. This theme or concept follows the idea, that a system is defined by its form (i.e. spatial structure) and function (i.e. processes and behavior). Based on this concept, the questions to pursue are: What information is needed or most suitable to understand the investigated system. And how much information on both, form and function, is necessary?

The irrigation experiment will be presented as an exemplary study on hillslope hydrology. In this context, the power of time-lapse GPR measurements as a non-invasive, temporally and spatially flexible method for hydrological monitoring will be discussed.

Following, an in-depth analysis is needed that goes beyond the very nice visual presentation of soil moisture data and calculated flow velocities.

We had originally hoped to obtain soil moisture changes from GPR data. However, due to the highly heterogeneous radar reflections and the absence of defined horizons, only a qualitative interpretation was possible. Due to the open system and the unknown fluxes from the irrigation area towards the GPR-monitored downhill area, assumption about mass input or the ratio between diffusive and advective soil water flow remained highly ambiguous. To avoid over-interpretation, we chose to remain at the level of recorded arrival times of the signal, response patterns and dynamics.

We will elaborate the analysis of the GPR response patterns to quantitatively compare the natural event and the irrigation with regard to the spatial share of the activated hillslope cross-section. Furthermore, the dynamics of the TDR and GPR data will be analyzed more thoroughly.

Answers to line-by-line comments

- The syntax is often awkward (except in the introduction) and could highly benefit from a very careful revision (or editing by a native speaker)
  Will be done

- P1L10: 'hillslope-scale connectivity'. There is no evidence provided for connectivity. There is evidence that some water moves downward from the sprinkling site. This neither indicates connectivity in this case nor connectivity under natural conditions.
  Will be changed:
  - before: The experiment revealed a fast establishment of hillslope-scale connectivity despite unsaturated conditions...
  - now: The experiment revealed fast establishment of preferential flow paths at the hillslope-scale despite unsaturated conditions.

- P1L11-12: 'These processes' is somewhat vague.
  Will be changed to 'These preferential flow processes'

- P2L13: Change: 'is on' to 'was on'
  Will be done

- P2L26: Change: 'At minimal...' to: 'GPR provides...'
  Will be changed:
  - before: At minimal invasive cost it provides information on subsurface structures in electrically resistive soils.
  - now: GPR provides information on subsurface structures in electrically resistive soils at minimal invasive cost.

- P3L7ff. I am not sure, if introducing multi-modal transit time distributions here and in the following parts is beneficial for the manuscript. If the authors introduce it and eventually mention it at several points, they should start calculating them for the mentioned events, rather than doing

a simple two component end member mixing (although this seems not to be possible with the isotope sampling). References are needed after 'flow process'.
In the revised manuscript, we will avoid the term 'multi-modal transit time distributions' to keep the focus on hillslope processes.

- P3L16ff. References for the claims are needed. Why vague?
This refers to the fact that many studies do not specify subsurface flow processes in more detail, while the term itself subsumes a wide variety of different processes. The sentence was meant to introduce the following paragraph, where we state that the specific characteristics of subsurface flow processes are a crucial component in the hydrological behavior of a landscape. A less vague formulation will be used.

- P3L16ff. The research need that is outlined here, is not convincingly derived and link to the majority of the introduction. The introduction exclusively deals with preferential flow processes, before some double-peaked hydrograph behaviour is mentioned. After that, the research needs is only related to these hydrographs. As mentioned earlier, I miss the well thought out story line. I also missed references to important connectivity work, since the connectivity (hillslope-stream) part seemed to be an important part of the paper.
The introduction will be revised and focused on investigation of subsurface flow in the context of the concept of 'form and function' applied to hydrological systems. These two aspects will be merged to elaborate the research question of how the observation of responses can be used to characterize a hillslope with regard to its form and function. The use of concepts like connectivity, double-peak hydrographs and multi-modal transit time distributions will be carefully checked and adapted to the revised story line.

- P6L10: What is the uncertainty that is associated with this very simple hydrograph separation, especially when accounting for the high variability of stable isotopes in rainfall?
Uncertainties will be provided and discussed

- P7L3: Please provide the relative location of the irrigation site, when hillfoot is 0 and hilltop is 1. What is the length of the hillslope?
Will be provided

- P7L3: Language and syntax are improvable.
We will improve the language of the manuscript and consult a native speaker for final editing.

- P7L10: What is the return period of the sprinkling event (intensity, amount)?
Will be provided

- P7L11: Here it is stated that it is not intended to mimic natural conditions, but then the authors used the results to infer flow processes under natural conditions. This does not fit together.
The conclusion, that similar processes also occur under natural conditions

is based on the observation of similar GPR reflection patterns and dynamics in response to the natural storm event. In the revised manuscript, we will elaborate this line of thought and clarify the causal connection between observations and conclusions. This will be done by stating clear hypotheses in the beginning. We will furthermore discuss the natural and irrigation responses separately, compare the results, and put the observations in relation to each other.

- P7L17ff: Information about size and height above ground of the samplers are needed. What were the intensities of the natural rainfall events? This could be a good comparison to the experiment.
  Will be provided

- P9L26-28: 'The...' Better delete? Not needed. Another suggesting: better not start sentences with 'Figure 2 shows/presents/' or 'This/that is presented in Figure 2'; rather report the finding or what the reader show see and then reference the figure.
  We are not sure what this comment is referring to, as there is no reference to a figure in the annotated paragraph. We will, however, take this comment as a general recommendation to be applied to the whole manuscript.

- P10L1: 'Similar'. So what is the difference to the cited work?
  The same procedure was used and the text will accordingly be changed to 'as presented by...'.

- P10L14: 'suggested' Is certain or uncertain?
  The GPR data certainly show changes in GPR reflection patterns, which is interpreted as changes in water content. Soil moisture is not the only possible factor causing differences in GPR reflection patterns, but under the given circumstances/time scale the only reasonable one. The term will be rephrased for clarification.

- P10L14: I am also wondering where the separation between mobile and stagnant water was made?
  In contrast to the TDR data, GPR data does not provide absolute soil moisture values, and all interpretations are based on differences between different measurements. As the last GPR measurement was chosen as reference, water which was still there during the last measurement will not be visible in the structural similarity patterns. We therefore can only observe changes in soil moisture, and infer information on advectively and non-uniformly flowing mobile water only. The sentence will be rephrased for clarification.

- P10L31: Why was the 2% value chosen?
  The 2% threshold was chosen based on the variability within presumably constant soil moisture measurements. This will be clarified in the revised manuscript.

- P10L4ff: 'were calculated for the entire depth': Do you mean, that if you give the value for 10 cm it is the velocity for 0-10 cm, if you report the velocity for 20 cm depth, it is the velocity from 0-20 cm?
  That is right. This will be clarified in the revised manuscript.

- P10L9: 'main stem'. Delete?
Will be changed.

- P12L13ff: How are these hydrographs separated? What method was chosen? This needs to be mentioned in the methods. What are the uncertainties associated with the isotope hydrograph separation?
The required method descriptions and uncertainties will be provided.

- P12L13ff: Why was only one event chosen for analysis?
As mentioned earlier, we chose to show only one hydrograph, because the emphasis of this part of the discussion was on the timing between rain event and the observed subsurface GPR reflection patterns (prior to the irrigation experiment). We chose to show hydrographs of several subcatchments to prove that the behavior is not specific to the investigated Holtz river, and referenced former studies to show that the observed response was not a unique event. In the revised manuscript, we will emphasize the results and discussion of the subsurface responses rather than the hydrograph. However, we will also check if the presentation of more hydrographs benefits the story line.

- P13: Strange font size/type of '20 mm' in the figure captions of fig.3.
Will be changed.

- P13L2: 0.5 L. This was over the whole experiment? So more or less unimportant in the water balance.
This is what we conclude later on in the manuscript. We also consider this an important finding, as surface runoff is often mentioned as a possible reason for an immediate runoff response or the event water signal in the river.

- P13L2 'relatively': Please provide a objective description, rather than using subjective terms like small, relative, best with numbers. Do that throughout the manuscript.
Thank you for pointing this out. We will comply with this request.

- P13L8: Can these 20% be the uncertainty of the water balance? Or what is the actual uncertainty and how does that influence the results and interpretation? I am not sure what 'average storage deficit of 20%' means.
The average storage deficit refers to difference between the mean storage increase in the four core area TDR profiles and the water input, which will be clarified. There is an uncertainty in this average water balance, which is mainly due to the low spatial coverage of four TDR profiles at the core area. It is likely that four profiles are not sufficient to fully represent the water balance of the entire irrigation site, which is why we chose to show single profiles instead of average values. We will mention and discuss this aspect in the revised manuscript.
The water balance uncertainty of single profiles, however, is smaller, yet not representative for the entire core area. Here, the change between mass balance overshoot in the beginning and deficit later on is an indicator for a high portion of mobile water in the soil. This differentiation and the respective uncertainties will be mentioned in the revised manuscript.

- P14: Fig4. Uncertainty needed.
  Will be provided.

- P15L4-7: Move to methods.
  Will be changed.

- P15L13: 'generally low dynamics'. Please change.
  Will be changed to 'The temporally and spatially low dynamics'

- P15L31: 'all fast dynamics'
  Will be changed to 'no changes in GPR signal could be observed anymore'

- P15L32-33: Is there a possibility to differentiate if it is in the matrix or left the monitored area? How about working more with the water balance?
  There are several reasons why we can not calculate a water balance from GPR data. The most important one is that we can only observe structural differences between measurements. We can not tell how much water is left in the mapped transects at the end of the monitoring period, nor can we quantify the differences in terms of water content.
  In the revised manuscript, we will include the calculation of storage changes for the TDR profiles at the downhill monitoring area, to provide an estimation on how much water remains in the matrix. A water balance can not be calculated here either, though. Reasons are the unknown water input, i.e. lateral flow from the irrigation area, and the sparse spatial coverage vs. high heterogeneity of the monitored soil volume.

- P17L2: This is more discussion of processes then results. How was this inferred? It seems like a statement, but not supported by data. There is no evidence for this.
  The paragraph will be moved to the Discussion and the line of thought will be elaborated more clearly. Nevertheless we consider it as one of the results of the GPR measurements, that the natural rain event as well as the irrigation experiment caused soil moisture changes in the entire monitored soil profile down to a depth of 4.2 m. These patterns were shown to be patchy and do not indicate a water table.

- P17L4: Again, there is no proof for this. It is just speculative. How deep below the soil surface is the bedrock?
  This statement is based on the comparison of the GPR reflection patterns caused by the natural rain event and the artificial irrigation. It will be moved to the Discussion section and discussed more carefully.
  We do not have reliable information on the depth of the bedrock interface and it is presumably characterized by a gradual transition between the deposit layer and solid bedrock. At the investigated hillslope, drilling was usually inhibited by high stone content at a depth of 1.7 to 2m. This information will be added to the methods section.

- P18: Can velocity distributions be estimated? They might be more interesting/informative than just the first response time. One could see how the majority of the water moves, as the response time of the preferential flow should be somehow consistent with the time of the hydrograph response, especially if the preferential flow should explain the delivery of

event water to the stream.
Unfortunately, soil moisture changes on their own are not sufficient to calculate or estimate velocity distributions. To do so, tracer applications and the recording of breakthrough curves are required. We agree, that the estimation of velocity distributions would be very helpful in the investigation of the role of preferential flow at the catchment-scale. However, response velocities also bear some very important information, especially in such high spatial coverage as provided by the GPR measurements.

- P22L6: This is commonly observed. Soil moisture variability increases with decreasing moisture contents
  The statement was meant to explain the observation, rather than being a finding of the experiment. An adequate citation will be added.

- P22L12: 'high portion of mobile water'. Please quantify. This is important.
  Will be provided in the Results section.

- P22L13: 'Velocities over...' But this only relates to the fastest component. The water itself should travel with a velocity distribution, which was not estimated. I would argue that the median or average speed to the 'mobile' water (the matrix water is mobile, too) is less than 10-3 m/s. Give the numbers for MaiMai.
  We agree that the median or average *real* velocity of the water surely is much lower than the given maximum *response* velocity of $2.2 \cdot 10^{-3}\, ms^{-1}$. The response velocities shown in Figure 8 range between $10^{-4}\, ms^{-1}$ and $10^{-3}\, ms^{-1}$ (based on GPR and TDR measurements). Any water arriving after the first response cannot be used to quantify volumes or velocities, because (a) we can not reliably quantify the newly arriving water based on the soil moisture state of the monitored soil volume and (b) we do not know the 'start time' of the water, which might be somewhere between irrigation start and measuring time.
  The intention here was to investigate the potential of preferential flow for the immediate runoff response. We use the maximum response velocities to show that the response velocities occurring in preferential flow paths are sufficient to route water from the hillslopes to the river within the short time between rain event and discharge response. The fast decrease in soil moisture after irrigation stop shows that a large portion of the water (approx. 50% at the core area, see Figure 4) is highly mobile and leaves the monitored depth at velocities that exceed the potential matrix flow velocities.

  This will be clarified in the revised manuscript and the line of argumentation will be elaborated.

  Numbers for MaiMai will be provided.

- P22L16ff. This is true, but it still does not show water transport of volumes to the stream that might explain the hydrograph response. Is there impermeable/less permeable material below the field site? How deep is the bedrock? Was an attempt made to observe the subsurface stormflow there?

We cannot exactly locate the depth of the bedrock interface, but maximum possible drilling is usually at depths between 1.5m to 2m at the investigated hillslope. The application of GPR measurements was clearly motivated by our will to identify preferential flow layers if they existed. This will be discussed in the revised manuscript.

- P22L16ff. In addition: The activation of the preferential flow is highly dependent on the rainfall intensities. They might never be active under natural conditions when rainfall intensities are an order of magnitude lower (over event scale).
  Based on the strong signal observed in the GPR measurements between natural rain event and irrigation, we conclude that there was an even stronger response of preferential flow caused by the natural event despite lower intensity and shorter duration. This finding was thus not solely based on the experimental observations. The line of argumentation was obviously not clear enough and will be elaborated in the revised manuscript. This will be done by a separate discussion of the experimental results and the natural response (including the GPR responses) and the quantitative comparison of the two different responses.

- P22L24-25: Connectivity at hillslope scale was not shown in this experiment. The data simply showed that some preferential flow out of the sprinkling site occurred.
  We agree that the experimental results alone do not prove hillslope-scale connectivity. As mentioned above, this statement is based on a synthesis of experimental results and the observed response triggered by the natural rain event. Unfortunately, we can not translate the GPR signal into quantitative soil moisture changes and provide a mass balance of the monitored transects. But the strength of the natural signal in relation to the experimentally initiated response, clearly indicates a higher amount of water despite lower local input. We will qualitatively compare the observations from natural rainfall event and irrigation and also discuss the effect of additional vertical input during the natural event in contrast to the exclusively lateral input during irrigation. We will also check whether the use of the term connectivity is appropriate in this context.

- P22L30: Every hydrograph is likely a result of multi-modal distribution of travel times...
  The catchment-scale response, as well as double-peak hydrographs and transit time distributions in particular, will be discussed more carefully in the revised manuscript. The entire section will therefore be restructured.

- P23L20: How were the riparian zones and wetland patches delineated? What is the uncertainty? Is there a contraction/extension during events that could change this?
  The wetland patches are very small and restricted to the source area of the river. A larger patch which is not in direct proximity of the river was mapped in the field once. The rest was estimated based on field observations, by applying a buffer of 2 m to the river sections associated with wetland patches. This is just an estimate, and we can not exclude contraction/extension of this area. However, consider its impact minor

with regard to the overall areal share. This will be clarified in the revised manuscript.

- P23L22ff. Again, there is no evidence shown in this work that there was any connectivity between hillslope and stream. This is even more true when natural conditions are considered.
  As mentioned above, we will elaborate the line of argumentation regarding the observed hillslope processes more clearly in the revised manuscript (see replies to comments on P22L16ff and P22L24-25). We will also focus more on hillslope processes rather than catchment-scale interpretations. By doing so, we will stay within the actual scope of our dataset and avoid over-interpretation.

- P24L26ff. This methodological discussion is overly long, repetitive, and lacking references to work of others.
  The methodological discussion will be significantly shortened and streamlined to emphasize the novelty of the method rather than discussing general aspects of soil moisture measurements and subsurface exploration.

- P24L27-P25L3: Here I disagree. As said, the experiment was not designed to mimic natural conditions. This is completely fine, since it seems to me that the original idea was to employ the GPR system. Nevertheless, the setup does not allow inferring natural flow processes that explain hydrograph response.
  The experiment was designed to visualize potential preferential flow paths. The natural rain event, however, allowed us to put our observations in the context of naturally occurring responses. The analysis of similarities and differences between the response patterns and dynamics allows for conclusions on naturally occurring processes.

  We agree that the current presentation is misleading and needs clarification. We will revise the discussion of the general experimental setup. Furthermore, we will shift the focus of the entire manuscript towards an evaluation of the methods and the approach.

- P25L10: Explain: 'sufficiently steady state'
  The phrase refers to the preceding paragraph, stating that natural systems barely reach steady state conditions. For the investigation of preferential flow processes, however, the absence of advection can be assumed as steady state, even though comparably slow diffusive movement is still going on. In other words: To identify the dynamics of advective flow processes initiated by a temporally limited input, the amount of water moving slowly through the system does not matter. While the first one causes fast changes in soil moisture, the latter is visible as a trend and can easily be ignored.

  The paragraph will be rephrased and the term avoided.

- P25L24: Sat. conditions are most likely occurring at the bedrock interface. Technically the periglacial layers are not 'soils' anymore. Try to have a look at the work of Uhlenbrook et al. on periglacial slopes.
  We agree that the sentence 'Saturated conditions are very improbable in such soils.' needs to be rephrased or even deleted as indeed saturation is possible above a confining layer. However, the GPR data did not show

such a saturated area within the top 4 m, while on the other hand the rock became too hard to further penetrate with our power drill at a depth of max. 2 m. We therefore struggle to identify both the form/structure and the function/response of such an impermeable layer in the investigated hillslope.
We will also be more careful with the use of the word 'soil'.

- P27L31ff. This seems to be unnecessary - no important point was made in this section.
  The methodological discussion will be shortened and the section removed.

- P29L4: The resulting connectivity was not shown and remains highly speculative.

  P29L5ff: This is speculative. Furthermore, the link to transit time distributions is somewhat strange.

  P29L20: Again, I think this was not shown here
  Summarizing the three comments above, we are very sorry that some key points of the manuscript were not well conveyed in its current version. As mentioned earlier, the manuscript will be revised and restructured to better elaborate these points. Furthermore, we will check the conclusiveness of our findings and conclusions and formulate them more carefully.