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#3

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#3, 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#3 and marked in orange.

Answers to the general comments

This study is designed to investigate subsurface flow processes, and in specific preferential flow, at two different scales. The authors chose a multi-method approach including quantitative soil moisture measurements using TDR (time domain reflectometry) and qualitative soil moisture measurements using GPR (ground penetrating radar) during an irrigation experiment at the hillslope scale. The experiment was complemented with common hydrometric and tracers techniques (interpretation of the rainfall runoff graph and stable water isotopes of rainfall and runoff) at different catchment scales. The study addresses an important topic in the hydrological sciences. The manuscript is clearly structured following common scientific standards. The data are presented in (mostly) clearly arranged, high quality figures. The combination of TDR and GPR measurements provide a very promising approach to monitor subsurface processes. While this approach is relatively novel, it bears a number of uncertainties and disadvantages, which are discussed at length (although the methodological discussion lacks a bit in references).

However, I do have some concerns. First of all, clear objectives were not formulated in the beginning of the manuscript and it follows that the rest of the story misses a central theme.

We will formulate clear hypotheses (see above) and state them at the beginning of the revised manuscript. The story line will be revised and adapted accordingly.

I do understand though, that the catchment scale analyses revealed that the system responded with a double-peak hydrograph to water input with the first, steeply and quickly responding hydrograph consisting of a mixture of new and old water and the second, dampened hydrograph consisting dominantly of old water. The results from the irrigation experiment were then used to explain the hillslope processes that may potentially lead to this catchment scale response. This leads to my second and third concerns: (2) I feel that only one hydrograph separation to calculate the event and pre-event water fractions of a catchment is not representative. Many studies have shown in the past that the event water fraction may vary largely in one catchment depending on various factors, e.g. water input and/or antecedent moisture conditions [e.g. Munoz-Villers & McDonnell, 2012]. Hence, I suggest including more data and/or an uncertainty analysis to add representability to the HS.

The one hydrograph was shown to temporally relate the observed subsurface response after the natural rain event (the green portion of the GPR data shown in Fig. 8) to the discharge behavior. The timing of the two observations, the subsurface response on the one hand and the discharge response on the other, allows us to draw conclusions on the processes within the hillslope. This line of thought was not properly conveyed. We will improve the story line by presenting and discussing the responses to the natural rain event and the irrigation separately. A quantitative comparison of the GPR response patterns of both events (the natural and experimental) will emphasize the focus on subsurface flow processes rather than catchment-scale responses. We agree that the fraction of event water in the first peak is likely to vary with rainfall characteristics and antecedent conditions and will avoid general statements based on this single event. We will furthermore check, whether the inclusion of more events will benefit the revised story line.

And (3) the authors clearly point out that they did not intend to mimic natural conditions with the irrigation experiment, however, they use the results to explain the natural response at the catchment scale. While this shortcoming is mentioned in the methodological discussion, I don't feel that this justification is sufficient to link the observed processes in the irrigation experiment to the natural conditions at the catchment scale.
The conclusions about relevant processes under natural conditions was backed by the observations of the subsurface response to the rain event prior to the experiment. This was not clearly elaborated and will be improved as described above. We will also shift the focus of the manuscript towards the methodological aspects of GPR measurements for hydrological monitoring and the informative power of response observations for the characterization of hillslopes.

At last, I feel that the novelty of this study is not clearly conveyed to the reader. I think it is the combined use of TDR and GPR measurements which bear a large potential to move forward in subsurface flow process understanding. However, this message needs to be presented more clearly.
With a stronger focus on the methodological aspects of the experiment, we are confident to improve the presentation and convey the novelty of our findings.

Answers to the detailed comments

Objectives and experimental approach

- I think that this subsection needs a bit more structure and/or more precise phrasing. Maybe name all objectives at the beginning of the paragraph and then list the approaches that you chose to address them in addition to the explanations why you chose said approaches (advantages vs disadvantages of others).
  This will be done. Clear hypotheses and research questions will be stated at the end of the introduction of the revised manuscript. We furthermore developed a concept to organize and subsume the approaches and methods presented in the two companion papers. This will strongly improve the structure of the introduction as well as the story line of the manuscript.

- Page 3, line 24: the role of what? Please add information because otherwise this sentence and subsequent paragraph come a bit out of the blue, i.e. raise further questions, e.g. why is it necessary to use a multi-scale approach?
  We will fully revise this subsection and clarify that sentence.

- Page 3, line 25: What are the conventional hydrological methods that you mention? I suppose the TDR measurements that you chose as approach in your experiment? Please add this information to the text or rephrase.
  The term 'conventional hydrological methods' refers to TDR measurements and piezometers. We will clarify this in the revised manuscript and fully revise this subsection.

- Page 3, line 27: you only propose GPR, why do you not mention the other techniques? (Maybe use this 3 comments to elaborate your objectives.)
  We agree and will mention other techniques in the revised manuscript. We will also fully revise this subsection.

Methods

Hydrological response monitoring

- Page 6, line 22: How exactly was rainfall water sampled for isotope analysis? Was a sequential sampler used or is it a bulk sample? Please add information.
  This information will be provided.

- Line 24 (and everywhere else in the ms): I suggest to use 'hydrograph separation' instead of 'mixing model', the latter implies that an EMMA was performed which is not the case.
  We will make this correction, thank you for pointing this out.

Process monitoring

- Page 8, line 3: Please add information why different TDR sensors were used.
  We wanted to increase the measurement frequency by measuring with three sensors in parallel. As only these three sensors were available to us, we decided that the benefit of higher frequency measurements out-weighed the disadvantages caused by the different sampling volume of one of the sensors.

- Page 8, line 6: Which 3 TDR tubes were used? Sounds like they were mentioned before which they aren't. Please rephrase or add information for better understanding.
  The sentence refers to the three different sensors, which were mentioned above. As this was obviously not phrased clearly enough, we will rephrase the paragraph: 'Soil moisture was measured manually. To enable parallel measurements and increase temporal resolution of the measurements, three probes were used in parallel. These probes differed slightly in their sensor design: Two TDR probes had an integration depth (i.e. sensor length) of 0.12 m and one probe had an integration depth of 0.18 m. These sensors were successively lowered to different depths into the 16 access tubes, where they measured the dielectric permittivity of the surrounding soil in the time domain through the PVC of the access tubes. Given a mean penetration depth of 5.5 cm and a tube diameter of 4.2 cm this yields an integration volume of 0.72 L and 1.05 L, respectively. The manual measurements were conducted in 0.1 m depth increments and followed a flexible measuring routine to cover active soil profiles with higher frequency.'

Results

- Figure 3:

  there is too much information in this Figure. I suggest moving some of the information to e.g. a separate table. For example, the information of the catchment sizes and the % of the precipitation amount should be placed

somewhere else, maybe together in one table. You may mark the 7th and 72th hr in the RR-Graph and refer the reader to the table. It also remains unclear to me why the information about the irrigation experiment (i.e. the mean structural similarity attribute) are placed in the rainfall-runoff and chemographs of the headwater catchments. Which y-axis does the mean structural attribute refer to? And what do the vertical bars mean? Please provide this information in the legend. I suggest considering making two different graphs, for example below each other in the manner of Figure a) catchment scale and b) hillslope scale. This should help to understand this figure faster and identify/extract important information more easily.
We agree that this figure is very busy and loaded with information. To make it easier to grasp this information, we will divide the figure in two graphs, add another y-axis for the mean structural similarity attribute and more clearly explain the shown GPR dynamics. However, as we consider the timing of (the green portion of) the GPR signal relative to the hydrograph to be important for the overall interpretation, we would prefer to keep this information together in one figure. We will furthermore elaborate the importance of the timing in more detail in the revised manuscript.

Information from the legend will be provided in a separate table.

Please also provide information about where groundwater was sampled for d18O, e.g. in section 2.2.
We will provide this information.

- Page 12, lines 7: Since there are several peaks in Figure 3, it would make it easier for the reader if you added the date of the described rain/hydrograph event.
The events will be labeled for clarification.

- Page 13, line 5: What is an 'overshoot' in mass recovery? Please add a bit more information to explain this.
We will rephrase the sentence and explain the observation: 'All profiles showed a mass recovery of more than 100 % (i.e. higher storage increase than water input at measuring time, see Fig. 4) in the first 60 min of the irrigation period.'

- Page 14, lines 19-20 and Figure 5: The weak signal of soil moisture dynamics in TDR 13, 6 and 14 is apparent in Figure 5. While an explanation is provided for the signal in TDR 6, I am missing explanation for TDR 13 and 14. How do you explain this? I suggest including this in the discussion section 4.1.
The behavior of TDR14 and TDR13 will be mentioned here ('TDR13 and TDR14 show a slight decrease in soil moisture.') and discussed in Section 4.

- Figure 5: Soil moisture change is interpolated over time for visual reasons (according to figure caption), I think it would be helpful to add this information somewhere in the methods section (e.g. 2.4.1. TDR data analysis).
We agree and will add this information Section 2.4.1. 'While the data

was interpolated in time for better visualization, all data analyses were performed with the uninterpolated data.'

Discussion

- Process interpretation: What is the novelty about this study? I suggest elaborating this e.g. add a paragraph to emphasize this a bit more clearly. The strength of our two-part study is the discussion of various data sources and their contribution to process understanding in the context of the interplay of form and function (a biological concept here transferred to hillslope hydrology). In the first part (MS1), the focus is on the observation of response patterns with GPR and TDR measurements. The combined use of these methods and their evaluation in the context of form and function of a hillslope as well as an alternative to the classical hillslope trench is the novelty of this study. To emphasize this aspect more clearly, we developed the above-mentioned concept for the two companion papers, highlighting two different methodological approaches. Both approaches include GPR measurements, which allows discussing the method under different aspects. According to this new concept, the story line of the manuscript will be revised and better structured along our research objective and hypotheses.

- Page 22, lines 22-25: I do not understand the link between delayed signal in the intermediate depth and the network of preferential flow paths. Please add a bit more text.
The term 'delayed signal in the intermediate depths' refers to the fact, that the soil moisture signal arrived in greater depth (i.e. below approximately 1 m) before it appeared in the intermediate depth (approximately 0.5 to 1 m). Thus, the water bypassed this intermediate layer without being detected. Two scenarios can explain this observation: a) The water passes through the intermediate depth outside of the monitored soil volume. b) The water passes through the intermediate depth through small preferential flow paths. If the volume of these flow path is comparably small in comparison with the monitored soil volume, they will only become visible (by means of soil moisture changes) if the water leaks into the surrounding matrix. While we can not distinguish between these processes by means of the data presented in MS1, both are preferential flow processes acting at different scales. This will be clarified in the revised manuscript.

- Page 23, lines 22-29: So, does that mean that the first peak is generated by overland flow (as observed by Wrede et al.) AND shallow subsurface flow? Or is all of the event water conveyed via subsurface flow? Please express your interpretation a bit more clearly.
We can not make any statement about the importance of overland flow for runoff generation beyond our observation, that surface runoff is of minor importance at the hillslopes, and that the areal share of saturated areas in the catchments is small. We therefore assume that the fraction of surface runoff is small. On the other hand, our results from the hillslope experiment show that the slope response is indeed fast and a potential source for mobile water. We thus conclude that preferential flow in the shallow

subsurface to some extent contributes to the first peak. To address the question about the fraction of event and pre-event water from the hillslopes in more detail, we will also include the soil water isotope data in MS1 (currently presented in MS2).
We will clarify these interpretations and include the above-mentioned argumentation. We will also revise the story line in a manner to stronger focus it on the hillslope response and its investigation instead of catchment-scale behavior.

Technical corrections

- Figure 4: Can you move the y-axis name and unit away from the numbers, e.g. to the very top or remove 140 and insert 120 instead?
  Will be done.

Typing errors

- Throughout the whole manuscript use either one of Fig. X, Fig X or Figure X.
  Will be checked and adapted to the journal's specifications.
- Page 3, line 8: empirically
  Will be done.
- Page 9, line 23: Add 'of' in 'In case of the vertical profiles...'
  Will be done.
- Page 12, line 19: use 'to' instead of 'of' in '...up to 67.6 % to the event runoff...'
  Will be done.
- Page 13, Figure 3, legend: 'attribute' instead of 'atribute'
  Will be done.
- Page 14, line 3: delete 'to' in '...returned to similar to initial conditions...'
  Will be done.
- Page 14, line 15: insert '(' in 'Fig.5)'
  Will be done.
- Page 15, line 2: use plural instead of singular: 'flow paths'
  Will be done.
- Page 19, Figure 6 caption, line 1: time-lapse GPR
  Will be done.
- Page 22, line 13: I suggest to use '...match...' instead of '...fitted with...'
  Will be changed to 'agreed'.
- Page 24, line 22: delete redundant 'the'
  Will be done.

- Page 27, line 17: substitute 'is' with 'are' in 'The dynamics of preferential flow are often characterized...'
  Will be done.
- Page 29, line 14: stable isotopes instead of staple isotopes
  Will be done.