Response to the Editor

Dear Authors:
Your original manuscript was evaluated by three reviewers who rated it from "good" to "excellent" as for the scientific significance of your study, but even rather "fair" with respect to the overall presentation quality and the robustness of a few scientific parts. Some comments from Refs.#1 and #3 require careful considerations and adequate responses. Ref.#2's suggestions also help improve the presentation of your study. Please, give adequate attention the Ref.#3's comments related to P15, L32.
As my personal comment, I understand that your definition of "surface saturation" is functional (and somewhat qualitative) to the major objective of your study, namely to exploit the TIR imagery. However, with a view to the use of a very comprehensive, distributed hydrological model (such as the Richards-based HGS model), I do think that a more physical definition would have been desirable (even to the benefit of a wider readership), such as that full saturation at soil surface is attained when soil-water matric pressure head is zero. A more qualitative than quantitative definition of soil saturation may perhaps then require some words about the distinction between the hortonian and dunnian overland flow generation processes.
Please upload your revised manuscript, together with detailed point-by-point replies to all of the comments received so far. An additional reviewing step may be required.

Dear Nunzio Romano,

We wish to thank you for the handling of our manuscript and the provided feedback. We revised our manuscript following the comments and suggestions of the three referees and our changes are explained in detail in the point-by-point replies to the referees. Below, we shortly describe the main changes and provide a reply to your comment on the definition of surface saturation.
The main criticism of referee #1 and #3 related to the structure of the manuscript and the broad, unspecific research objectives. Thus, we revised the title and the research questions, specified that we focus our study on 4 different characteristics of surface saturation, and added a small description on the structure of the study at the end of the introduction. Moreover, we moved the last section of the results and revised the conclusions to be consistent with the structure of the presented research objectives (cf. our responses to referee #1 and #3).
All three referees missed some information in the abstract and in order to include the different missing aspects within 250 words we rewrote the abstract completely. Furthermore, we added more information on the model parameterization and TIR image processing as asked for by referee #2 and #3.
Finally, we improved the description and discussion on the possibilities of the model to simulate surface saturation as a result of different processes, i.e. infiltration excess, saturation excess, overland flow, or groundwater water exfiltration. This includes a better clarification that we estimated simulated surface saturation based on the simulated water depth in the surface domain (i.e. we classified the cells of the surface domain as saturated if simulated water depths were >10-4 m, being consistent with the mapping of the TIR imagery). The definition is purely physical and, in contrast to a definition based on the matric potential of the soil surface (i.e. subsurface domain of the model), it includes all different types of surface saturation and not only surface saturation generated by saturation excess (which would be the case when defining the saturation based on the matric potential of the soil surface). Moreover, we included some sentences to discuss that infiltration excess overland flow was unlikely to be simulated based on our model setup and to explain why the analysis of the model output with regard to the generation processes of surface saturation is nonetheless of interest (cf. our response to the comment on P15L32 by referee #3).

We kindly ask you to reassess our revised manuscript and consider it for publication.
On behalf of all co-authors
Barbara Glaser

Response to anonymous referee #1

General comments

The article by Glaser et al presents an interesting analysis of the spatial and temporal variability of surface saturation dynamics in a small forested catchment in Luxembourg. The study of surface saturation patterns and development is certainly a relevant topic in hydrological research and I really appreciate the contribution of the authors in terms of the large amount of field data and observations. The authors use an impressive dataset of hydrometric information (continuous data of discharge, water table, soil moisture (at different profiles)) at several locations within the catchment. Additionally, they include an interesting and somehow novel methodology to map surface saturation over specific areas (through thermal infrared images). The high frequency of these images (a total of 291) allows them to nicely catch the temporal dynamics of surface saturation patterns over several years. All these is very valuable. They also use a hydrological model to investigate the generation and development of surface saturation. The model is validated not only with the stream discharge, but also and, interestingly, with "internal" information (groundwater, soil moisture and saturation patterns) within the catchment.

We wish to thank the reviewer for the supportive assessment of our manuscript and the given comments and suggestions that helped to improve the quality of the manuscript.

1. My main criticism is that the presentation of the results + discussion is somehow a bit mixed up and I think a better structure will help the reader to follow more smoothly the story (which is complex in terms of the quantity of data used, number of riparian areas and variables analyzed). An easy way to ensure a good structure is once the objectives are clearly stated, used them to structure the Results, Discussion and Conclusion sections. Eg, Objective 1, 2, 3 –> corresponding Results 1,2,3 –> corresponding Discussion 1,2,3 etc. Here I really think a better correspondence Results-Discussion-Conclusion is needed.

We revised the research objectives to make the research questions more specific (cf. also our response to the comment below and to referee #3). The reshaped presentation of the research objectives should also clarify the structure of the results and discussion section. Namely, we now clearly express two research questions regarding the intra-catchment variability of surface saturation characteristics (RQ 1-> model performance, RQ 2 -> identification of key controls). Moreover, we specifically indicate that we consider four different surface saturation characteristics: i) the temporal dynamics of surface saturation extent, ii) the relationship between saturation extent and catchment discharge, iii) the spatial patterns of surface saturation occurrence, and iv) the spatial patterns of saturation frequency. The results and discussion section of the original manuscript were in large parts structured following these four different characteristics. Only results section 4.1 and 4.5 did not directly relate to one of the characteristics i)-iv) and thus also did not have a direct correspondence in the discussion section. However, the two sections provide necessary general information about model performance and functioning over the entire catchment. In order to improve the structure, we moved the results section 4.5 directly after section 4.1. By this, the results now start with an analysis of the simulation results for the entire catchment (section 4.1 and 4.2 in the revised manuscript). Then the results continue with the specific analysis of the surface saturation characteristics i)-iv) in the different riparian areas (i.e. section 4.3 focuses on the temporal dynamics (i), section 4.4 focuses on the saturation-discharge relationship (ii) and section 4.5 focuses on the spatial patterns (iii and iv)). We specified this structure at the end of the introduction and adapted the titles of the results sections to further clarify that the focus of sections 4.1 and 4.2 is on the entire catchment simulation and that the focus of sections 4.3 to 4.5 is on the variability of the different surface saturation characteristics between the distinctly investigated riparian areas. Finally, we revised the conclusion section so that it follows the order of the presented research objectives by first answering the two research questions and then specifically picking up on the surface saturation characteristics i)-iv).

2. Also, Iom not sure whether the main objectives are sufficiently clearly formulated. As far as I understand, the ultimate objectives of the paper are to analyze the variability of surface saturation dynamics and patterns within a forested catchment (this is clearly stated), but also to explore (identify and discuss) the possible factors controlling the generation of surface saturation (this is done through the analysis of the matches and mismatches of observations and model simulations). This is not specified in the objectives and I think it is important.

We revised the presentation of the research objectives to make the research questions more specific (cf. also our response to referee #3). We removed the first part of research question #1 so that the question only focusses on the model performance and does not mix this with the investigation of the intra-catchment variability of different surface saturation characteristics. Instead, we specify that we base our study on an existing TIR imagery data set and that we perform additional analyses of the data (compared to previous studies) in order to fully characterize the intra-catchment variability of the surface saturation dynamics and spatial patterns of surface saturation occurrence and frequency. In research question #2 we replaced 'What can we learn about the reasons for …' with the more specific formulation 'What key controls …can we identify …' to state more precisely that we aim to

explore the key controls on the generation of surface saturation and that this is done based on the identified match and mismatch between observations and simulation. Moreover, we now clearly point out that we address both questions with regard to the four different characteristics of surface saturation: i) temporal dynamics, ii) relationship between saturation extent and catchment discharge, iii) spatial patterns, and iv) spatial patterns of saturation frequency (cf. our response to the comment above).

3. My last main criticism is that the authors do not sufficiently justify the novelty of this study with respect to previous studies (eg., Glaser 2016, 2019). There is only one sentence in L42-43 but is not very convincing. It is important they reinforce this idea (eg, why it is important to do the analysis in a larger catchment? What do you expect as a new finding?).

The novelty of the study is that we investigate the spatial and temporal variability of surface saturation in a range of saturated riparian areas with different characteristics, location, and potentially different underlying hydrological processes. This goes beyond previous work elsewhere and also beyond previous simulation studies in the Weierbach catchment (Glaser et al. 16, Glaser et al. 19) that only investigated one specific riparian area. We now clarify in the introduction that the previous 6 ha model included the source area of the stream of the 6 ha headwater and that the simulation of surface saturation in this area was also validated with TIR imagery, but that the expansion of the model extent from the 6 ha headwater to the entire 42 ha catchment in this study allows the simulation and investigation of the spatial and temporal variability of surface saturation occurrence within an entire catchment. Moreover, we now present more details on the previous TIR imagery studies in the Weierbach (Glaser et al. 2018, Antonelli et al. 2019) to better justify why it is of interest to explore the intra-catchment variability between the different riparian areas with a combined observation – simulation approach.

Given this, my recommendation is moderate revisions are needed.

Specific/minor comments:

Title: I do not think we can we consider 2 years as “long term” period of observations and simulations. Consider the possibility of changing the title according to my comment 2.

We removed ‘long-term’ and revised the title also considering the comments of referee #2 and #3.

Abstract: you do not mention that you also analyze the relationships between surface saturation and discharge and groundwater

The other two referees saw a lack of information in the abstract as well. In order to include the different missing aspects in the abstract, we rewrote the abstract completely. The abstract now better specifies the identified differences and similarities of surface saturation dynamics and patterns between the simulations and observations as well as the different investigated areas, including the analyzed relationship between surface saturation and discharge. The comparison of the extent of simulated surface saturation and simulated groundwater reaching the surface over the entire catchment (results section 4.2, previously 4.5) was not a main focus of our study but was supposed to support the analysis of the model functioning to identify possible key controls on the surface saturation generation (see also our response to the main comment on the structure of the manuscript). Therefore, we do not explicitly mention this analysis in the abstract but only generally mention the finding of a large control of groundwater exfiltration on surface saturation generation.

p.3L27. Is it not the hydrological year 2016-17 (starting Oct 2016)?

Yes, we specified it in the text.

p.5L7. can you give an aprox surface of the riparian areas (and/or TIR images)?

We now indicate, as a reference, the areal extent of the smallest and largest of the seven investigated riparian areas.

p.6Fig2. for a clear reading in the text it would be helpful to add a), b) and c). Is it mC or Cm? (homogenize with the figure caption, text and tabl 1). Define LP?

We homogenized the label for the fresh bedrock to mC, added a definition of LP in the figure caption and included labels a)-c) for reference in the figure caption and text.

p.9fig3. Is there a way to indicate at which swc is there saturation?

It would be possible to indicate for the simulated soil water content where saturation of the soil is reached. However, it is not possible to indicate the same for the measured soil water content, since we do not know the exact porosity of the soil at the measurement locations. We actually calculated the simulated volumetric soil water content from the simulated relative soil saturation and the model porosity (given in Table 1) in order to be able to compare the simulation results with the observations. We do not think that it would be useful to indicate at which swc the soil is saturated for the simulation if we cannot do the same for the observations.

P11fig4. Can you remind in the fig caption that the colors correspond to that of Fig 1?

We added a sentence explaining the colours in the caption of Fig. 5 (previously Fig. 4) and also Fig. 6 (previously Fig. 5).

P17L23-30. I believe microtopography can be an important factor at small scales (small study areas) so be sure (and indicate) that the examples you give to discuss this (L24-27) are also carried out at the small scale. I would say at larger scales the main factors explaining the extent/variability of surface saturation are catchment slope, climate, land use:

It is true that the influence of microtopography may vary with the size of the study area and that other factors such as physiography or climate are likely to be more relevant at larger scales. We added a sentence to specify this and we now indicate at which scales the discussed studies were carried out.

P18L3. “A simulate surface extension of 1.6%”. As far as I remember from section 2.1, you include an exceptionally dry year (?). You should mention this.

This is correct, the investigated hydrological year 2016-2017 was exceptionally dry. Yet the maximum extent of surface saturation was estimated from the entire simulation period (October 2015 – January 2018) and it was actually reached at the beginning of the hydrological year 2017-2018, where conditions were very wet. Therefore, we think it is rather misleading to mention the exceptionally dry year in the context of the maximum extent of surface saturation. Instead, we now specify in the text that the maximum extent of surface saturation was simulated for the very wet conditions in winter 2017/2018.

P18L24. “the frequency maps of surface saturation”. Specify you are referring to the images in fig 6 (otherwise the reader may think it is Fig 7).

The first sentence of section 5.4 is also valid for the map of Fig. 4a (previously Fig. 7a). We now nonetheless refer directly to Fig. 7 (previously Fig. 6), since it is true that the following discussion mainly refers to Fig. 7 (previously Fig. 6).

P19L14. “observed frequencies” of surface saturation (?)

We completely rephrased the part on the saturation frequencies in the conclusion, also ensuring to specify that we refer to the observe

Response to anonymous referee #2

The paper manuscript presents a rich dataset on Thermal Infrared Image collected for a large number of snapshots and derived maps of saturated areas of 7 stream sections in the Weierbach (L). These images are compared to saturated areas simulated based on a physically-based, spatially distributed model. The text is well written, structured, gives credit to the relevant literature. The methods used are described and results presented in a clear and appropriate way. I have only minor suggestions on how to make the text and the figures clearer (see details in the pdf uploaded).

We wish to thank the reviewer for the positive and thorough assessment of the manuscript and the helpful comments and suggestions for improvement.

The most important suggestions are:

1) Consider to include “Thermal Infrared Image observations” in the title

As suggested, we included ‘thermal infrared image’ in the title. In addition, we further revised the title considering the comments of referee #1 and #3.

2) In the introduction I would encourage the authors to better justify why surface runoff can be an important component of runoff generation.

We added three sentences in the introduction that differentiate the importance of surface runoff in comparison to subsurface flow for discharge generation in different environments and that better explain why the investigation of surface saturated areas in forested catchments is of interest.

3) I would also suggest to better explain the abbreviations used in the paper; check if they are used consistently in the text and the figures and include the abbreviations in the figure legend or caption.

We specified in the text that the different riparian areas are named according to their location along the different tributaries by indicating their location together with the used abbreviations in brackets (‘seven distinct riparian areas along the left (L1, Fig. 1), middle (M1-M3, Fig. 1), and right (R2, R3, Fig. 1) tributary and in the central stream valley (S2, Fig. 1)’). In the same way we added an explanation of the abbreviations used in Figure 1 in the figure caption. Moreover, we carefully rechecked the consistent use of the abbreviations throughout the text, figures, and tables.

4) I suggest to include some more information on how the TIR pictures were taken (see my detailed comment in the pdf), how the temperature range to classify pixels to be saturated or un-saturated was determined (and not only refer to an earlier paper), and state that the model parameters were based on field and lab measurements (instead of refereeing to earlier literature).

We added more information on the acquisition and processing of the TIR images following the detailed demands and suggestions in the PDF (see our detailed replies in the PDF). Furthermore, we now give more information on the data that were used in Glaser et al. 2016 for the model setup, parametrization and calibration, following also the comments in the PDF (see our reply there) and the questions of Referee #3.

5) Clarify to the reader what the difference is between the data shown in Figure 7a and 7b.
(see also my detailed comments in the pdf attached).

We defined simulated surface saturation based on a simulated water depth $>10^{-4}$ m in the surface domain of the model, independently of the saturation of the subsurface. This is shown in Fig. 7a.

'Groundwater reaching the surface' indicates that the subsurface domain of the model was fully saturated, i.e. we marked a cell of the surface domain as cell where groundwater reached the surface when the subsurface domain below the surface cell was fully saturated from the bottom to the top. This is shown in Fig. 7b.

We admit that the two different definitions were originally explained only very briefly in the methods section of the manuscript. We revised the methods section and now include an entire paragraph to explain more clearly the difference between the two definitions. Moreover, we now specify the difference between Fig. 7 a and b in the figure caption and in the text of the according results section by repeating the definition criteria for simulated surface saturation (Fig. 7a) and simulated groundwater reaching the surface (Fig. 7b).

These are all minor suggestions that can be address easily by the reviewers in a short time. The content of the article is relevant to the hydrological community and meets the focus of the selected journal very well.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-203/hess-2019-203-RC2-supplement.pdf

Please note our responses to the commented PDF attached after the redlined version of the manuscript.

Response to anonymous referee #3

The authors use a novel and large dataset of hydrometric data and thermal infrared images to describe and simulate the extent of surface saturation at seven locations across a small catchment in Luxembourg. The work is very interesting and novel and the simulations are well described. Based on the title and the first research question, I had expected more analysis and discussion of the observed patterns but instead the manuscript mainly focuses on how well these patterns of saturation can be simulated with a physics-based model. I highly command the authors for this study but think that the impact of the manuscript can be improved by focusing more on specific research questions. Currently, the questions (particularly question 2) are very broad and the reader can't escape the occasional thought that the manuscript 'just' describes a model application. That is a real petty because the dataset and the work are very novel. I, therefore, recommend that the authors adjust the title, rewrite the research questions and focus the discussion and also the conclusion on these research questions so that it is clearer what the take home messages are. Overall, the manuscript is well written and particularly well-illustrated but the wording can be clearer and more to the point at some locations (see attached annotated pdf for some suggestions).

We wish to thank the reviewer for the supportive assessment of our manuscript and the useful comments and given suggestions for improvement.
We swapped the order of 'simulations' and 'observations' in the title to put a higher focus on the modelling aspect rather than on the observation data. In addition, we revised the title accounting for the comments of referee #1 and #2. We think that these adaptations will avoid misleading expectations on the content of the study from the title.
Furthermore, we made the research questions more specific (cf. also our response to referee #1). We removed the first part of research question 1 so that the question only focusses on model performance and does not include the investigation of the intra-catchment variability of different surface saturation characteristics. Instead, we specify that we based our study on an existing TIR imagery data set. We also specify that we performed additional analyses of the data compared to previous studies in order to fully characterize the intra-catchment variability of the surface saturation dynamics and spatial patterns of surface saturation occurrence and frequency. In research question 2, we replaced 'What can we learn about the reasons for …' with the more specific formulation 'What key controls …can we identify …' to state more precisely that we aim to explore the key driving controls on the generation of surface saturation based on the identified match and mismatch between observations and simulation. Moreover, we now clearly point out that we address both questions with regard to four different surface saturation characteristics: i) temporal dynamics of surface saturation extent, ii) relationship between saturation extent and catchment discharge, iii) spatial patterns of surface saturation occurrence, and iv) spatial patterns of saturation frequency.
The results and discussion section of the original manuscript were in large parts structured following the four different saturation characteristics. Only results section 4.1 and 4.5 did not directly relate to one of the characteristics since these two sections provide necessary general information about model performance and functioning over the entire catchment. Considering the comments of referee #1 on the structure of the manuscript, we moved the results section 4.5 directly after section 4.1. By this, the results now start with an analysis of the simulation results for the entire catchment (section 4.1 and 4.2 in the revised manuscript). Then the results continue with the specific analysis of the surface saturation characteristics i)-iv) in the different riparian areas (i.e. section 4.3 focuses on the temporal dynamics (i), section 4.4 focuses on the saturation-discharge relationship (ii) and section 4.5 focuses on the spatial patterns (iii and iv)). The discussion section remained unchanged, since it was already structured according to the saturation characteristics i)-iv), always discussing the aspects for both research questions.
Finally, we revised the conclusion section so that it follows the order of the presented research objectives by first answering the two research questions and then specifically picking up on the surface saturation characteristics i)-iv).

Specific comments:

P1L1: Adjust the title so that it is clearer that the manuscript mainly focuses on the simulations and what "part" of the patterns can be simulated based on topography and climate inputs. Now the reader may expect a more in-depth analysis of the observed patterns of surface saturation.

We swapped the order of 'simulations' and 'observations' in the title to clarify that the study focusses on 'simulations in comparison with … observations' rather than on an in-depth analysis of the observation data. In addition, we further revised the title considering the comments of referee #1 and #2 to be more precise about the applied methodology. We think that these adaptations will help to avoid misleading expectations on the content of the study from the title.

P1L23: It is unclear from the abstract which part of the observed variability is important but not explained by topography or groundwater exfiltration. Please add additional info.

The other two referees saw a lack of information in the abstract as well. In order to include the different missing aspects in the abstract, we rewrote the abstract completely. The abstract now better specifies the identified differences and similarities of surface saturation dynamics and patterns between the simulations and observations as well as the different investigated areas.

P2L5-15: This is a nice list of what has been done previously but it would be more useful if a short summary of the main findings from these studies (and thus also the remaining open questions) is given as well.

The different studies listed in this paragraph have very different focuses and in some of them surface saturation was only a side aspect of the investigation. Thus, it is very difficult to summarize the findings of the different studies in a concise way in one or two sentences. Adding a longer description of the different focusses and findings of the studies would have been possible. However, we think that this would have been rather misleading, since the main message of the paragraph should be that there are some modelling studies that dealt with surface saturation, but that there is none that investigated the spatio-temporal saturation dynamics within a catchment in a distributed way. For this reason, we decided not to add any additional information on the listed studies.

P3L8-9: Make the research questions more specifc! Now they are very general and not so helpful, particularly question 2. Then focus the discussion and conclusion more on these questions, rather than the different model steps or the overall discussion of the goodness of fit of the model.

We removed the first part of research question 1 so that the question only focusses on model performance and does not include the investigation of the intra-catchment variability of different surface saturation characteristics. In research question 2, we replaced 'What can we learn about the reasons for …' with the more specific formulation 'What key controls …can we identify …' to state more precisely that we aim to explore the key driving controls on the generation of surface saturation based on the identified match and mismatch between observations and simulation. Moreover, we now clearly point out that we address both questions with regard to four different surface saturation characteristics: i) temporal dynamics of surface saturation extent, ii) relationship between saturation extent and catchment discharge, iii) spatial patterns of surface saturation occurrence, and iv) spatial patterns of saturation frequency. The discussion section of the original manuscript was already structured following these four different saturation characteristics i)-iv). Therefore and because both research questions actually require some general discussions of the goodness of fit of the model (RQ1: model performance, RQ2: identification of key controls based on match and mismatch between simulations and observations), we did not change the structure and focus of the discussion section. The conclusion section was revised so that it now follows the order of the presented research objectives by first answering the two research questions and then specifically picking up on the surface saturation characteristics i)-iv).

P5 section 2.2: Add more information on the georeferencing of the images. How was it done and how accurate is it? What is the size of a pixel and how many pixels do you think you are off?

We extended the description of the co-registration of the images. The images were not geo-referenced against coordinates but manually co-registered against a reference image. For details on the processing and the reasoning behind the reader is now referred to the previous study of Glaser et al (2018), a technical study that developed a detailed method for applying and processing TIR imagery for mapping surface saturation. A quantification of the accuracy of the co-registration is not possible since the overlapping / deviation of the pixels is not consistent within and between the images. The size of a pixel corresponded in average to 0.16 cm^2, but the average pixel size varied between 0.1 cm^2 and 0.2 cm^2 for the different investigated riparian areas and, even more importantly, the pixel size was not consistent within the images due to the low angles of view and distorted image perspectives. Thus, one pixel in the center and at the lower edge of the panorama images represented by tendency a smaller area than one pixel at the sides and upper edge of the images. We did not further specify the pixel sizes in the text since they were not really relevant for our study. Instead, the most important aspect for our study was that the TIR images of one area were consistent in the perspective and areal extent (and thus varying pixel sizes) between each other (cf. extended description of co-registration) and to the images extracted from the model output (what we ensured as explained in section 3.2 Assessment of the model performance).

P5L16: How many pictures (or what percentage of all pictures) were excluded?

We added the information that we monitored surface saturation on 63 monitoring dates. Based on this information and the given total number of 291 obtained binary panoramic images it is now possible for the interested reader to calculate the number / percentage of images that were excluded from the analysis (63 monitoring dates x 7 areas = a potential number of 441 images versus an actual number of 291 analyzed images).

P5: You show that topography and microtopography are very important for the patterns of saturation and frequency of occurrence of saturation. However, I seem to have missed how this topographic information was

obtained. Lidar data? What is the resolution and what smoothing algorithms did you use? What is the size of a mesh grid cell and how much of the microtopography does the mesh reflect?

This is a very good and important point. We actually did not include any information on the used digital elevation model (DEM) in the previous version of the manuscript. The topographic information for the model mesh nodes was interpolated from a 0.1 m digital elevation model (DEM). The DEM represented the combination of a coarse DEM of the hillslopes and plateau sites that was interpolated from 10 m contour lines of a topographic map and a very high resolved DEM for the stream valleys that was acquired with ground-based LiDAR (resolution around 5 cm). By merging and interpolating the two DEMs to a resolution of 0.1 m we ensured that most of the microtopographic information of the riparian zone and streambed was maintained in the model mesh (mesh resolution < 0.4 m in the stream valleys). We added this information in the manuscript.

P6L6: A bit more information on this previous calibration is needed. What data were used for it? Which data for the calibration are the same as used here for the validation? Which riparian zones correspond to this area and were these pictures of saturation used already for this calibration? Please provide more info for the reader who hasn't read this paper yet.

Following the questions here and the comments given by referee #2, we added information on the data that were used in Glaser et al. 2016 for the model setup, parametrization, and calibration. Specifically we clarified i) that the previous modelling work focused on the period from October 2010 to August 2014 (versus October 2015 to January 2018 in this study), ii) that the model included the riparian area M3, iii) that discharge measurements up- and downstream of the area M3 and soil moisture measurements at TSM1-5 were used for model calibration, and iv) that TIR imagery observations from winter 2010/2011 and spring 2014 were used for model validation for a part of area M3.

P7L3: The sentence on 'spatial heterogeneity' could be clearer. I think that you only added a different soil parameterization for the riparian zone but as it is written now the reader could think that you added spatial heterogeneity within the riparian zone. In that case, there are different parameters for different parts of the riparian zone, which would of course significantly influence your results.

This is correct, we only used a different subsurface structure in the riparian zone compared to the hillslopes and plateau sites and we did not apply spatially heterogeneous parameters within the riparian zone. We clarified this aspect in the text by now stating that the 'hillslopes and plateau sites' were parameterized in one way while 'a differing subsurface structure was implemented in the stream valleys' and we do not use the wording 'spatial heterogeneity' anymore.

P7L29-P8L1: This is unclear to me. Either you used the KGE as the evaluation criterion or the Pearson correlation. Did you add a different weight for the Pearson correlation in the KGE? If you did, then this needs to be described in the text.

We only validated the model performance and did not perform any model calibration. Therefore, it is possible to use different evaluation criteria for assessing the model performance and there is no need to apply different weights for different criteria. We simply calculated both the KGE and Pearson correlation for the groundwater levels to get a quantitative idea about the model performance with regard to both the performance in general (KGE as combined measure for correlation, bias, and relative variability) and the relative dynamics in particular (Pearson correlation). We clarified in the text that KGE and Pearson correlation were used to evaluate the model performance (not for calibration) and that they were both calculated for the groundwater level simulations in order to assess the general performance and the simulated dynamics in particular.

P11L5: I guess that you are showing the less reliable data with a different symbol but that is not clear from the caption or legend.

Exactly, the less reliable data are shown with triangular symbols. This is indicated in the legend on the right side of the figure, but we now additionally added the information to the figure caption.

P15L33: Provide a bit more details on the Antonelli et al (2019) study. How did they statistically analyse the images to show this?

We added the requested information in the text.

P15L32: While it is interesting that the model suggests that the ponding is due to saturation and groundwater exfiltration, your model parameters (i.e., the Ksat at the surface and near surface) would already have told you that infiltration excess overland flow is unlikely for this parameterization. How were these parameters derived or

calibrated? Make sure that you don't present results that are directly related to the parameter values without referring back to the choice of these parameter values.

This is actually an important point to discuss and clarify. We now mention in the discussion section that the model parameterization did not really allow for infiltration excess. Moreover, we explain how the parameterization was derived (referring back to the section 3.1 Model setup and parameterization and the study of Glaser et al. 2016). Finally, we also explain why the simulation result is nonetheless of interest (i.e. the match between observations and simulations indicates that infiltration excess is not relevant and the model helps to further distinguish between a general generation of surface saturation by saturation excess and groundwater exfiltration). In addition, we now present more information in section 3.1 (Model setup and parameterization) on the data that were used in Glaser et al. 2016 for the model setup, parametrization and calibration, following the comments on P6L6 (cf. our response there) and of referee #2.

P16L42: But isn't part of this variation due to micro-topography that is not included in your mesh?

The message of the sentence here is that different relationships between surface saturation and discharge do not only occur between catchments with differing physiographic characteristics, but that we also observed and simulated variable relationships (represented as convex, linear, or concave shapes of the relationship) for the different riparian areas within our catchment. This highlights the intra-catchment variability of surface saturation and it is of course possible that this intra-catchment variability is – besides other physiographic factors such as the width of the riparian area or location along the stream – due to micro-topography, which actually is included in the model mesh (cf our response to the comment on P5).

P17L16: So what was the resolution of the topography data that was used here? And the resolution of the model mesh?

We added the information on the topography data in the description of the model setup (cf. our response to the comment on P5). The resolution of the model mesh is described in the section on the model setup as well. How the model mesh actually looked like in the seven distinct riparian zones is shown in Fig. S2.

P18L6: Yes this is in line with this result but doesn't explain why this is the case.

We clarified the line of argumentation and the connection between the first four sentences of the paragraph.

P19L12: Why do you need additional water sources? Why are spatially variable soil properties (and thus parameters) not sufficient?

It is true that spatially variable soil properties (or other factors) might be sufficient to explain the spatially varying frequencies of surface saturation within the riparian areas. However, in section 5.4 we discuss the reason why we suggest that the different frequencies might also reflect the presence of additional water sources. We added an additional argumentation for this in section 5.4. Moreover, we rephrased the sentences on the frequencies of surface saturation in the conclusion section to clarify that the inference of additional water sources from the varying frequencies is only a guess and has not been tested within this study.

P19 conclusions: I find this study really nice but the conclusion doesn't seem to have a clear take home message – or perhaps simply too many messages. This is a petty as this reduces the impact of this manuscript. This may be due to the lack of clear well-defined research questions. I thus suggest that the authors reflect on these questions and to add clearer take home messages to the conclusion.

We revised the presentation of the research objectives to make the research questions more specific (cf. our responses to the general comment, the comment P3 and to referee #1). In line with the revised research objectives and questions, we revised the presentation of the conclusions so that it now strictly follows the order of the presented research objectives. First, we generally answer the two stated research questions concerning the intra-catchment variability of surface saturation (RQ1: model performance, RQ2: identification of key). Second, we specifically pick up on the different investigated surface saturation characteristics and separately describe the key findings regarding the intra-catchment variability of the i) temporal dynamics of surface saturation, ii) relationship between saturation extent and catchment discharge, iii) spatial patterns of surface saturation, and iv) spatial patterns of surface saturation frequency. We think that this new structure of the research objectives and conclusion will allow the reader to separate and understand the take home messages from the conclusion.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-203/hess-2019-203-RC3-supplement.pdf

Please note our responses to the commented PDF attached at the end of the document. If no reply is given to the comments or edits, we revised the manuscript exactly as suggested.

Intra-catchment variability of surface saturation – insights from physically-based simulations in comparison with biweekly~~long-term~~ thermal infrared image observations ~~and simulations~~

Barbara Glaser[1,2*], Marta Antonelli[3,1], Luisa Hopp[2], Julian Klaus[1]

[1]Catchment and Eco-Hydrology Research Group, Luxembourg Institute of Science and Technology, Esch/Alzette, 4362, Luxembourg
[2]Department of Hydrology, University of Bayreuth, Bayreuth, 95447, Germany
[3]Hydrology and Quantitative Water Management Group, Wageningen University & Research, Wageningen, 6700, The Netherlands
[*]Now at: Department of Geography, Ludwig-Maximilians-Universität München, Munich, 80333, Germany

Correspondence to: Barbara Glaser (~~barbara.~~b.glaser@lmu.de~~list.lu~~)

Abstract

In this study, we explored the spatio-temporal variability of surface saturation within a forested headwater catchment in a combined simulation-observation approach. We simulated the occurrence of surface saturation in the Weierbach catchment (Luxembourg) with the physically-based model HydroGeoSphere. We confronted the simulation with thermal infrared imagery observations that we acquired during a two-year mapping campaign for seven distinct riparian areas with weekly to biweekly recurrence frequency. Observations and simulations showed similar saturation dynamics across the catchment. The observed and simulated relation of surface saturation to catchment discharge resembled a power law relationship for all investigated riparian areas, but varied to a similar extent as previously observed between catchments of different morphological and topographical characteristics. The observed spatial patterns and frequencies of surface saturation varied between and within the investigated areas and the model reproduced these spatial variations well. The match between simulation and observations suggested that surface saturation in the Weierbach catchment was largely controlled by exfiltration of groundwater into local topographic depressions. However, the simulated surface saturation contracted faster than observed, the simulated saturation dynamics were less variable between the investigated areas than observed, and the match of simulated and observed saturation patterns was not equally good in all investigated riparian areas. These mismatches highlight that the intra-catchment variability of surface saturation must also result from factors that were not considered in the model setup, such as differing subsurface structures or a differing hysteretic behaviour between surface saturation and catchment discharge due to the local existence of perennial springs.

~~The inundation of flood-prone areas varies in space and time and can have crucial impacts on runoff generation and water quality when the surface saturated areas become connected to the stream. In this study, we aimed to investigate and explain the variability of surface saturation patterns and dynamics within a forested headwater catchment. On the one hand, we mapped surface saturation in seven distinct riparian areas of the Weierbach catchment (Luxembourg) with thermal infrared images, taken weekly to bi-weekly over a period of two years. On the other hand, we simulated the surface saturation generation in the catchment with the integrated surface subsurface hydrologic model HydroGeoSphere over the same period. Both the observations and simulations showed that the saturation dynamics were similar across the catchment, but that small differences between the dynamics at different areas occurred. Moreover, the model reproduced the observed saturation patterns well for all seasonal and hydrologic conditions and at all investigated locations. Based on the observations and simulation results and the matches and mismatches between them, we concluded that the generation of surface saturation in the Weierbach catchment was largely controlled by exfiltration of groundwater into local depressions. However, we also~~

~~illustrate that the entire variability of the patterns, dynamics and frequencies of surface saturation within the different riparian areas of the catchment can only result from additional controlling factors to microtopography and groundwater exfiltration, such as differing hysteretic behaviour, differing subsurface structures, or additional water sources.~~

1 Introduction

It is critical for flood risk assessment to understand where and when water is standing or flowing on the ground surface outside of perennial surface water bodies. ~~When such surface saturated areas connect to the stream via overland flow, they also become crucial for runoff generation and water quality.~~ In general, such surface saturated areas arise from 1) water ponding on the surface due to precipitation intensity exceeding~~ance of~~ the infiltration capacity of unsaturated soil, 2) water ponding on impermeable surfaces or saturated soil, 3) water exfiltrating from the subsurface or, 4) stream water ~~extending~~ flowing into the floodplain (e.g. Megahan and King, 1985). When ~~such~~ surface saturated areas connect to the stream via overland flow, they also become crucial for runoff generation and water quality. While overland flow from surface saturated areas has been considered the dominant runoff generation process in early years of catchment hydrology (e.g. Betson, 1964), surface runoff is nowadays known to only dominate in specific environments such as urban areas and Mediterranean or arid catchments (e.g. Latron and Gallart, 2007). Nonetheless, surface saturation and overland flow also occur in forested, humid temperate catchments and other environments where runoff generation is dominated by subsurface flow processes, mainly - but not exclusively - due to saturation excess in the vicinity of the stream (cf. variable source area concept, e.g. Dunne et al., 1975; Hewlett and Hibbert, 1967; Megahan and King, 1985). Such an occurrence of surface saturation and overland flow in the riparian zone can mediate a fast connection between the hillslopes and the stream, inducing quick responses of streamflow to rainfall events and influencing the mixing of water and water quality in the stream (cf. e.g. Ambroise, 2004; Birkel et al., 2010; Bracken and Croke, 2007; Tetzlaff et al., 2007; Weill et al., 2013).

Over the past years and decades, various field studies have mapped and analysed the spatial and temporal occurrence of surface saturation within different landscapes (e.g. Ambroise, 1986, 2016; Dunne et al., 1975; Gburek and Sharpley, 1998; Latron and Gallart, 2007; Silasari et al., 2017; Tanaka et al., 1988). From these field studies it is well recognized that surface saturation varies in space and time and that ~~its appearance~~this variability is affected by structural (e.g. topography) and dynamic factors (e.g. precipitation intensity, antecedent moisture). Yet there is limited understanding on how surface saturation evolves spatially and temporally between and within landscapes and how the interplay of different controlling factors and ~~generation~~ processes affects~~controls~~ the spatio-temporal variability of surface saturation.

Spatially distributed and dynamic hydrological models are potential tools for analysing the generation and development of surface saturation in space and time. While a model always represents a simplification of reality, the big asset of spatially distributed and dynamic hydrological models is that they ~~Such models~~ allow ~~a~~ detailed investigation of surface saturation at any desired location and time. This goes far beyond the information that can be gained by ~~any~~ field observations. Several simulation studies have systematically assessed the influence of static and dynamic factors on the temporal evolution, connectivity, and spatial distribution of surface saturation by performing virtual experiments with hillslope models (Ogden and Watts, 2000; Reaney et al., 2014) or by testing a range of terrain indices for predicting time-integrated saturation patterns (Güntner et al., 2004). Other studies relied on dynamic distributed and semi-distributed simulations for analysing the connectivity of surface saturation in relation to wetness conditions and catchment runoff (Mengistu and Spence, 2016; Qu and Duffy, 2007; Weill et al., 2013). Weill et al. (2013) and Partington et al. (2013) analysed the processes and water sources that generate surface saturation in a wetland and a pre-alpine grassland headwater, respectively. Both studies applied a model belonging to the group of integrated surface-subsurface hydrologic models (ISSHMs, Sebben et al., 2013), which can simulate the interplay of different surface and subsurface processes of surface saturation generation (e.g. ponding of precipitation from the surface, exfiltration from the subsurface). However, ~~Modelling~~ modelling studies that focus on a

comprehensive spatio-temporal analysis of surface saturation dynamics within a landscape by evaluating the spatially distributed model outputs rather than ~~aggregating~~ the aggregated outputs are scarce (e.g. Nippgen et al. (2015) for subsurface saturated areas).

When complementing field observations with simulations to analyse the generation and development of surface saturation in space and time, it is important to ensure that the model yields realistic results. Glaser et al. (2016) demonstrated for a small riparian area that a good match between modelled and observed discharge or soil moisture does not automatically imply a realistic simulation of saturation patterns. They concluded that a spatial validation of the dynamic saturation patterns itself is crucial. However, only few of the existing modelling studies explicitly checked the realism of their simulated surface saturation with field observations before using them for further analyses. Moreover, the few existing studies that performed an explicit validation of simulated surface saturation ~~These studies~~ focussed either on temporally integrated spatial patterns (Grabs et al., 2009; Güntner et al., 2004) or on temporal dynamics of overall catchment saturation (Birkel et al., 2010; Mengistu and Spence, 2016), but hardly~~barely~~ any study combined the observation and simulation of both surface saturation patterns and dynamics (Ali et al., 2014; Glaser et al., 2016). The lack of such studies is certainly explainable by the resources that are necessary for obtaining appropriate field data. Today, we still lack a standard method to map surface saturation and the different existing methods such as the 'squishy boot' method, the us~~ag~~e of 'on-off' surface saturation sensors, the mapping of soil morphology or vegetation as surrogates, or the usage of remote sensing techniques (e.g. Dunne et al., 1975; Gburek and Sharpley, 1998; Güntner et al., 2004; Latron and Gallart, 2007; Mengistu and Spence, 2016; Silasari et al., 2017) all have their own advantages and disadvantages.

A relatively new and powerful method for mapping surface saturation is thermal infrared (TIR) imagery. TIR mapping relies on the difference between the surface temperature of water and other materials ~~for~~ to identify~~ing~~ surface saturation. Previous work showed that recurrent mapping of surface saturation with high spatial resolution is possible with TIR imagery (Glaser et al., 2016; Pfister et al., 2010). Glaser et al. (2018) and Antonelli et al. (2019) applied TIR imagery mapping in the 42 ha forested Weierbach catchment in western Luxembourg and monitored the dynamics of surface saturation within several distinct riparian areas along the Weierbach stream with a weekly to biweekly mapping frequency over several seasons. While Glaser et al. (2018) focused on method development and image processing, Antonelli et al. (2019) analysed the saturation dynamics observed in various distinct riparian areas in comparison to meteorological and hydrological conditions. They found similar seasonal extension and contraction dynamics of surface saturation in their investigated areas. This was particularly related to near-stream groundwater level fluctuations, yet Antonelli et al. (2019) also identified some local differences of saturation dynamics depending on the location and morphological characteristics of the distinct riparian areas.

In this study, we explore the intra-catchment variability of temporal and spatial characteristics of surface saturation (dynamics, frequencies, patterns) based on~~with~~ a combination of field observation and modelling. We perform the study in the Weierbach catchment, where we can rely on existing TIR imagery data (cf. Antonelli et al., 2019; Glaser et al., 2018) and on previous modelling work for a 6 ha headwater of the catchment (Glaser et al., 2016, 2019) with the ISSHM HydroGeoSphere. Glaser et al. (2016, 2019) simulated the 6 ha ~~area of the catchment~~headwater by accounting for a layering of the subsurface, while spatial heterogeneity was only represented by microtopography and ~~by~~ a different sequence of subsurface layers in the riparian zone compared to the hillslopes and plateau. They set up, adapted, and assessed the simulation based on various distributed field data, including TIR imagery observations of surface saturation in the source area of the stream of the 6 ha headwater (Glaser et al., 2016). Here, we extend the model setup to the entire 42 ha catchment without introducing additional heterogeneity and without ~~performing~~ a re-calibration. The expansion of the model to the entire catchment allows the simulation and investigation of the spatial and temporal variability of surface saturation within the catchment, including various distinct riparian areas with a range of morphological characteristics (e.g. extent, location along the stream). Furthermore, this also enables to study the potential occurrence of different hydrological processes within the different areas that was suggested by the precedent analysis of the TIR imagery observations (cf. Antonelli et al., 2019).

By contrasting the simulation ~~We simulate surface saturation in the catchment and contrast the~~ results with ~~the~~ observed surface saturation ~~patterns~~ characteristics from ~~the~~ TIR imagery, we aim to use the model as learning tool and address the following research questions: ~~focussing on the long-term saturation dynamics over different seasons and wetness conditions (25 months with weekly to biweekly mapping resolution) and the spatial patterns of surface saturation occurrence and frequency at seven different riparian areas across the catchment. The two research objectives that we aim to address with this approach are:~~

1) ~~How variable are surface saturation dynamics and patterns within a catchment and to~~ To what extent can we reproduce the observed intra-catchment variability of ~~the~~ surface saturation characteristics ~~(dynamics, frequencies, patterns)~~ with a rather homogenously set~~-~~ up ISSHM?
2) ~~What do we learn about the~~What key control~~s reasons~~ for the intra-catchment variability of surface saturation characteristics can we identify ~~from~~ based on the match~~es~~ and mismatch~~es~~ between simulation results and observations?

The specific surface saturation characteristics that we consider for both questions are:

i) the temporal dynamics of surface saturation extent,
ii) the relationship between surface saturation extent and catchment discharge,
iii) the spatial patterns of surface saturation occurrence,
iv) the spatial patterns of surface saturation frequencies.

We base our analysis on the existing TIR imagery dataset (Antonelli et al., 2019; Glaser et al., 2018) and perform additional analyses of the field data in order to fully characterise the intra-catchment variability of the temporal dynamics of surface saturation extent and the spatial patterns of surface saturation occurrence and frequency. The analysis includes TIR images taken over different seasons and wetness conditions (25 months with weekly to biweekly mapping resolution) of seven different riparian areas across the Weierbach catchment. The catchment model is set up relying on the previous modelling work for the 6 ha headwater (Glaser et al., 2016, 2019). We first perform an evaluation of the 42 ha catchment model against distributed measurements of discharge, soil moisture, and groundwater levels and investigate the catchment-wide simulation of surface saturation patterns and dynamics. Then, we identify the different simulated surface saturation characteristics i) - iv) for the seven distinct riparian areas and compare them with the respective field observations for answering the two research questions.

2 Study site and data

2.1 Physiography, climate and hydrometry

The Weierbach catchment is an intensively studied headwater catchment (42 ha) in western Luxembourg. About half of the catchment area is characterized by gentle slopes <5°, forming a plateau landscape unit (Martínez-Carreras et al., 2016). The rest of the catchment is characterized by hillslopes with slopes > 5°, forming a central V-shaped stream valley from north to south and a V-shaped tributary valley in the east. A third, few metres long stream ~~branch~~ tributary is situated ~~in the~~ west of the central stream valley. Riparian zones along the stream account for 1.2 % of the catchment area (Antonelli et al., 2019). Large parts of the catchment are forested with deciduous trees (mainly European beech and Sessile oaks), the south-east and some other small parts of the catchment are forested with conifers (mainly Norway spruce and Douglas spruce). The riparian zones are free of tree ~~canopys~~ and covered with ferns, moss, and herbaceous plants. Siltic, sceletic Cambisols ~~Soil~~ developed from Pleistocene Periglacial Slope Deposits area~~s~~ shallow and highly-permeable ~~siltic, sceletic Cambisol~~ with a depth ranging between 0.4 and 0.9 m (Gourdol et al., 2018; Juilleret et al., 2011; Moragues-Quiroga et al., 2017). Beneath the solum, a 0.5 – 1 m thick basal layer with bedrock clasts oriented parallel to the slope overlies fractured Devonian slate and phyllites (Gourdol et al., 2018; Juilleret et al., 2011; Moragues-Quiroga et al., 2017; Scaini et al., 2017). In the riparian

zones, the Cambisol~~soil~~ and basal layer have been eroded and the fractured bedrock is overlain by shallow organic Leptosols (Glaser et al., 2016).

The climate is oceanic-continental without apparent seasonality in precipitation and with negligible amounts of snow (Carrer et al., 2019). Mean annual precipitation during the period from October 2013 to September 2017 was 955±53 mm. Mean annual discharge was 546±253 mm, with exceptionally dry conditions in the hydrological year 2016-2017. During wet periods, discharge is characterized by double peak hydrographs with the first peaks appearing ~~as~~ immediately in response to precipitation and the second pronounced peaks appearing 48h to 72h later (cf. Martínez-Carreras et al., 2016). During dry periods, only the first hydrograph peaks occur and the stream dries out intermittently starting from the source areas downstream.

Hydrological and meteorological data that were used in this study were measured from October 2013 to January 2018. Data from the period from October 2013 to September 2015 were used for spin-up simulations, data from the period from October 2015 to January 2018 were used to drive and validate the actual simulation (cf. Section 3). Discharge was measured with water pressure transducers (ISCO 4120 Flow Logger, 15 min logging intervals) at four v-notch weires, installed at the outlet of the catchment (SW1, Fig.1) and upstream of the confluences of the three ~~branches~~ tributaries (SW2-SW4). Groundwater levels were continuously recorded every 15 minutes with pressure sensors (OTT CTD) in five piezometers installed in different landscape units (riparian zone, hillslope, plateau) of the catchment (Fig. 1, GW1-3, GW5, GW7). Soil moisture was continuously monitored (30 min logging intervals) with water content reflectometers (CS650, Campbell Scientific) installed horizontally ~~in~~ at 10, 20, 40, and 60 cm depth at four different sites (Fig.1, SM3-SM5, SM7). At each site, two depth profiles were monitored. In addition, soil moisture ~~in~~ at 10 cm depth was monitored with water content reflectometers (CS616, Campbell Scientific, 30 min logging intervals) at five locations transecting~~crosscutting~~ the riparian zone of the stream source area of the middle ~~stream branch~~ tributary (Fig. 1, ~~SSM1~~TSM1-TSM5).

Cumulative precipitation was recorded every 5 minutes with a tipping bucket raingauge (Young 52203, unheated, 1 m height) at an open area within the catchment (data gaps were filled ~~by estimating~~based on a linear regression to data from a station approximately 4.5 km southward). Potential reference evapotranspiration was estimated based on measured air temperature, relative humidity, wind speed, and net radiation according to the FAO Penman-Monteith formulation (Allen et al., 1998). Air temperature and relative humidity data were recorded next to the soil moisture profile SM5 (Fig. 1, HMP45C-LC, Campbell Scientific, 15 min logging intervals, 2 m height). Wind speed and radiation data were recorded approximately 4.5 km southward of the study site. Wind speed (Young Wind Monitor 05103, Vector A100R Anemometer) was recorded every 15 minutes ~~in~~ at 3 m height and converted to wind speed in 2 m height (data gaps closed with data from a station approximately 11.5 km north-eastward) following the FAO guidelines (Allen et al., 1998). Net radiation was recorded every 15 minutes (Kipp & Zonen NR Lite net radiometer) until May 2017. From June 2017 on (and for closing other data gaps), we used net radiation data recorded every 5 minutes close to Luxembourg Airport (~40 km southeast of the study site), as these measurements were highly correlated (linear regression with an intercept of 7.6 W m^{-2}~~,~~ and a slope of 0.92, R^2 = 0.81) with the measurements close to the study site in the years before.

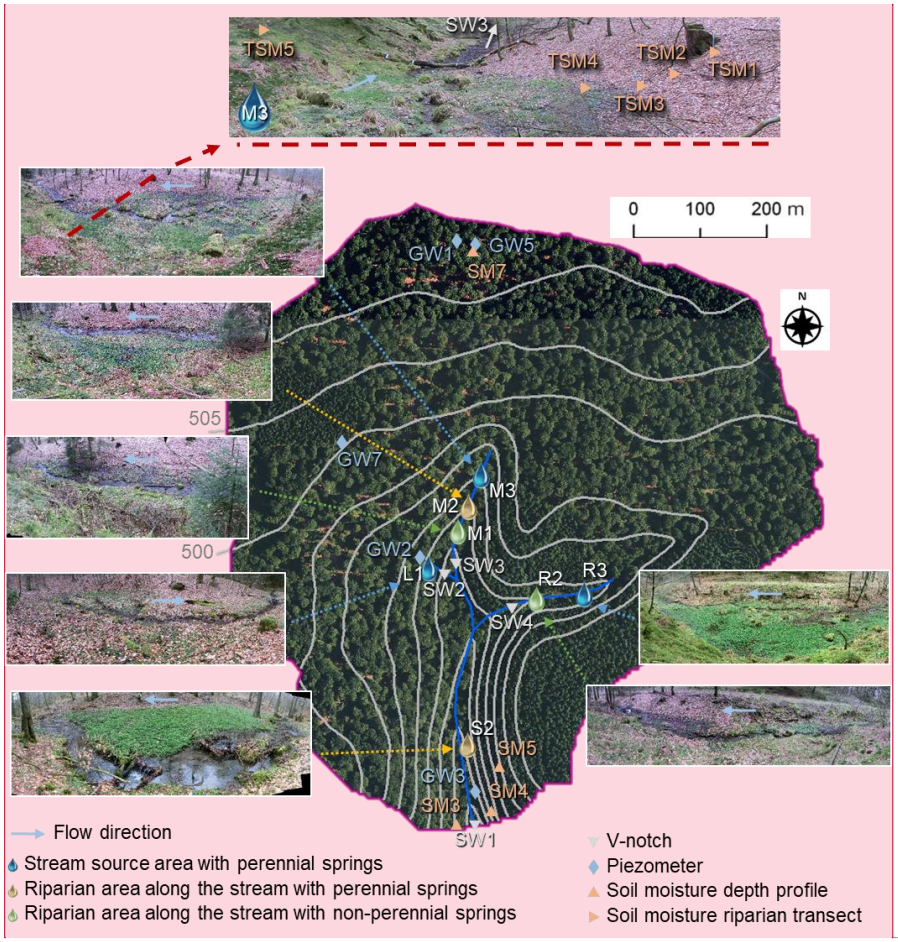

Figure 1: ~~The~~ Weierbach catchment with the locations of the installed v-notch weir~~s~~ for measuring discharge (SW1-SW4), piezometers for measuring groundwater levels (GW1-GW3, GW5, GW7), soil moisture sensors for measuring soil moisture in different depth profiles (SM3-SM5, SM7) and a riparian transect (TSM1-TSM5), and the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Orthophoto of the catchment: Administration du Cadastre et de la Topographie 2010 (geoportail.lu).

2.2 Surface saturation

Here, we define the surface~~s~~ as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This ~~involves~~ includes water bodies such as ~~lakes~~ponds and streams, but excludes mere saturation in the topsoil. According to this definition, surface saturation in the Weierbach catchment generally only occurs in the streambed and the adjacent riparian zones. Other areas that were occasionally observed to be surface saturated during very wet conditions or 'rain on snow' events are forest roads and the ~~prolongation~~ extension of the streambed above the source regions into the hillslopes. We focus in this study on seven distinct riparian areas along the left (L1, Fig. 1), middle (M1-M3, Fig. 1), and right (R2, R3, Fig. 1) tributary, and in the central stream valley (S2, Fig. 1)~~in the catchment~~. The seven investigated areas include about one quarter of the total stream network and area sizes range from 84 m^2 for the smallest monitored area (M1) to 232 m^2 for the largest monitored area (M3) (cf. Antonelli et al., 2019). According to their main hydro-morphological features, the seven areas ~~, which~~ can be classified into three different categories (cf. Antonelli et al., 2019): i) stream source areas with perennial

Kommentiert [B1]: We corrected the assignment of photographs of the riparian areas to their location along the stream for area M1 and M2

springs (L1, M3, R3, blue ~~areas~~ icons Fig. 1), ii) areas along the stream with perennial springs (M2, S2, yellow iconsareas Fig. 1), and iii) areas along the stream with non-perennial springs (M1, R2, green iconsareas Fig.1).

We mapped the surface saturation in the~~se~~ seven riparian areas weekly to biweekly from November 2015 to December 2017 with thermal infrared imagery (TIR). Details on the identification of surface saturation with TIR imagery and on the collected surface saturation dataset are presented and discussed in Glaser et al. (2018) and Antonelli et al. (2019). In brief, we took a panoramic snapshot of each of the distinct riparian areas with a handheld TIR camera (FLIR T640) on each mapping day. The TIR panoramas were taken each time from the same position in order to ensure a consistent areal coverage and angle of view. In addition, ~~W~~we manually co-registered the individual panoramas against a reference panorama for each area in an image post-processing step (cf. Glaser et al. 2018). Nonetheless,~~, but~~ slight ~~position~~ shifts in perspective~~shifts~~ between the panoramas of different dates were inevitable and the different features (e.g. streambed, stones, trees) in the images ~~. As a result,~~ the images ~~that were placed on top of each other~~ did not always overlap exactly pixel by pixel.

The ~~we created panoramic TIR images of the distinct areas and identified the~~ locations of surface saturation (including the stream) were identified from the TIR panoramas based on the acquired temperature information~~within the images~~. ~~To do this, e~~Each pixel in a~~n~~ TIR ~~image~~ panorama was assigned to be saturated or unsaturated based on the temperature range of locations that were obviously ~~to be~~ saturated according to~~from~~ field observations and visual images taken complementary to the TIR panoramas. The definition of the temperature range was done manually and individually for each mapping time and location (cf. Antonelli et al., 2019). While this is a subjective and laborious approach, precedent sensitivity and uncertainty analyses showed that it is a robust and reliable method, especially if the definition of the temperature range is done by only one person for all the images (cf. Antonelli et al., 2019; Glaser et al., 2018). In case the contrast between water temperature and temperature of surrounding materials was not sufficient for a reliable pixel classification, the images were excluded from the analysis. In case the pixel classification was affected by a poor temperature contrast or by pixels representing vegetation or snow cover ~~in the images~~, the images were analysed but flagged as less reliable. Altogether, we obtained from 63 monitoring dates a total of 291 binary panoramic images showing the temporal dynamics of surface saturation patterns in the seven studied riparian areas. The ~~with~~ total number~~s~~ of analysed images per site rang~~ed~~ing between 34 (L1) and 48 (M2).

Time series of saturation were created for each area by accounting for the percentage of saturated pixels within the individual panoramic images. We ~~normalized~~ normalised the saturation percentages to the maximum observed percentage of saturation in the distinct areas in order to allow a comparison of the saturation dynamics between the different riparian areas. Moreover, we compared the relationship between the normalised extent of surface saturation and catchment discharge for the different riparian areas with regard to monotonicity (quantified by Kendall correlation coefficients) and shape. ~~For picturing~~ In order to visualize the spatial surface saturation patterns and dynamics within a distinct riparian area, we created maps of saturation frequency. We counted for each area how often the individual pixels of the panoramic TIR images were classified as saturated and ~~normalized~~ normalised the resulting frequency numbers by the total number of TIR images analysed for that area.

The resulting maps of ~~normalized~~ normalised saturation frequency ~~rarely~~ suggest that very few~~showed~~ pixels ~~that~~ were always saturated (i.e. reaching a ~~normalized~~ normalised frequency of 1). Field experience and the analysis of individual TIR and visual images showed that ~~In~~ in reality, surface saturation was ~~more~~ per~~manents~~istent at more places than indicated by the frequency maps. The reason for this artefact is that the perspective and distortion ~~of~~ within the individual TIR panoramas was not 100% identical for all mapping instances (see above) and that vegetation sometimes covered parts of the saturated surface, especially ~~during near dry conditions~~for instances where the extent of surface saturation was narrow. ~~We co-registered the individual panoramas against a reference panorama for each area, but slight position shifts were inevitable. As a result, the images that were placed on top of each other did not always overlap exactly and~~ As a result, the generated saturation frequency maps are blurred. Nonetheless, the maps of ~~normalized~~ normalised saturation frequencies are very

useful to ~~understand at a glance~~quickly detect where surface saturation occurs more and less frequent within an area and ~~to be used~~ for model validation.

3 Catchment model

3.1 Model setup and parameterisation

We simulated the spatio-temporal dynamics of surface saturation across the Weierbach catchment with HydroGeoSphere (HGS, Aquanty Inc.). HGS is an integrated surface subsurface hydrological model and allows simultaneous simulation of transient surface and subsurface flow. Subsurface flow is simulated based on the 3D Richards equation. Surface flow is simulated based on the diffusive-wave approximation of the 2D Saint Venant equation. Evapotranspiration is simulated with a comparatively simple approach, following the mechanistic concept of Kristensen and Jensen (1975). The equations are linearized implicitly using the Newton-Raphson approach and solved in an unstructured finite element grid. HGS has been used in the past to~~for~~ address~~ing~~ diverse questions at various temporal and spatial scales (e.g. Ala-aho et al., 2015; Davison et al., 2018; Erler et al., 2019; Frei et al., 2010; Munz et al., 2017; Nasta et al., 2019; Partington et al., 2013; Schilling et al., 2017; Tang et al., 2018). It was also ~~has~~ already ~~been~~ applied for a 6 ha headwater region of the Weierbach catchment for simulations within the period October 2010 to August 2014 (Glaser et al., 2016, 2019). The 6 ha headwater model included the source area of the middle tributary and was set up, manually calibrated, and evaluated based on various distributed field data. The evaluation also included a validation of the simulated surface saturation based on 20 TIR images collected in winter 2010/2011 and spring 2014 for the western part of riparian area M3. In this study, we applied the parameterization of Glaser et al. (2016) to the entire 42 ha catchment without ~~performing~~ any additional parameter calibration.

The catchment was spatially discretized into 42,274 triangular elements (Fig. 2a), using the mesh generator AlgoMesh (HydroAlgorithmics Pty Ltd). Edge lengths of the mesh elements ranged from > 30 m at the plateau to < 0.4 m for the seven analysed riparian zones and the streambed (Fig. 2a-b). It was crucial to use such a fine mesh resolution in the riparian zone in order to enable a comparable spatial detail as obtained with the TIR imagery for the surface saturation patterns. The topographic information for the mesh nodes was interpolated from a 0.1 m digital elevation model (DEM). The DEM represented the combination of a coarse DEM of the hillslopes and plateau sites that was interpolated from 10 m contour lines of a topographic map and a highly resolved DEM for the stream valleys that was acquired with ground-based LiDAR (resolution around 5 cm). By merging and interpolating the two DEMs to a resolution of 0.1 m we ensured that most of the microtopographic information of the riparian zone and streambed was maintained in the model mesh. Vertically, the model grid comprised 5 m, ~~which were~~ divided into 14 layers with element depths ranging from 0.15 m for the top layers to 0.5 m for the bottom layers (Fig. 2c).

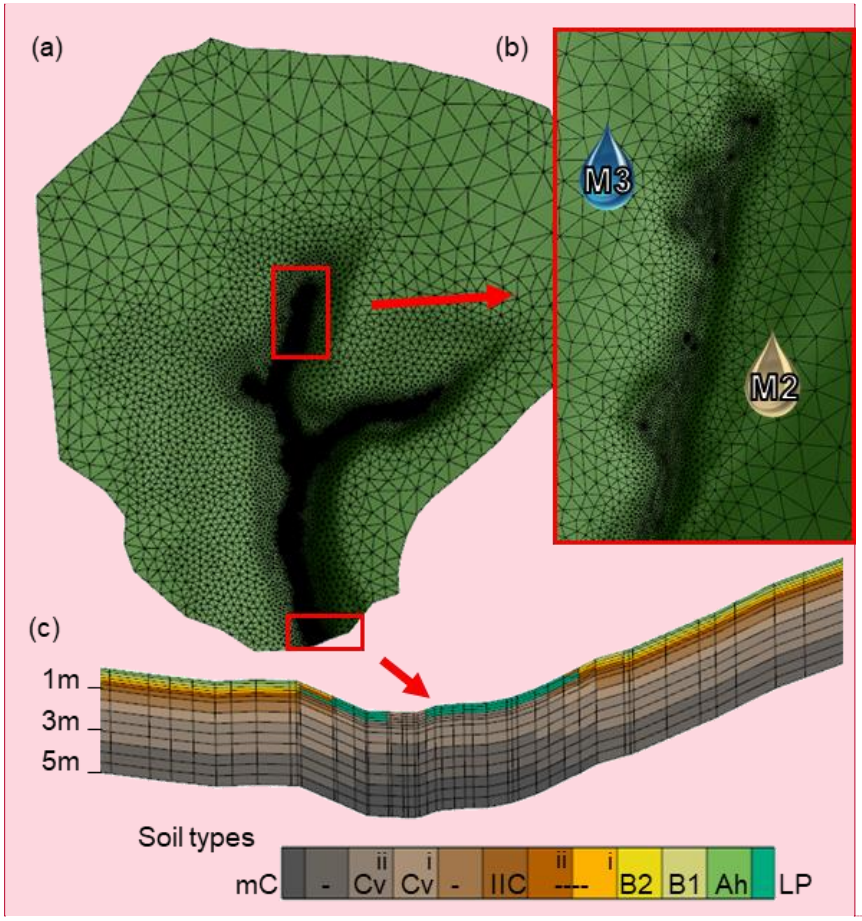

Figure 2: Setup of the model mesh (a) with a zoom on the fine horizontal resolution in the riparian areas and the streambed (~~inset on the right~~b) and a vertical cross section through the stream valley and adjacent hillslopes (~~c~~bottom) showing the vertical discretization and assignment of different soil properties (cf. Tab. 1). LP = riparian Leptosol, Ah = topsoil, B1 and B2 =subsoil, IIC = basal layer, Cv = fractured bedrock, ~~C~~mC = fresh bedrock.

The subsurface was parameterized based on information on the subsurface structure obtained from electrical resistivity tomography (ERT) measurements and a detailed description of soil properties from eight soil profiles distributed across the catchment (cf. Glaser et al. 2016). We parameterized the hillslopes and plateau sites homogeneously with 10 different property layers, representing top- and subsoil (Ah, B1, B2), the basal layer (IIC), fractured and fresh bedrock (Cv, ~~C~~mC), and transition layers ~~of transition~~ between subsoil, basal layer, and fractured bedrock (Fig. 2c). ~~We implemented~~ A differing subsurface structure was implemented ~~spatial heterogeneity~~ in the stream valleys because in this area~~, where~~ the soil and basal layers ~~were~~ are eroded and the outcropping fractured bedrock ~~was~~ is overlain with organic, stagnic Leptosol (LP) in the riparian zones (Fig. 2c). We used the Mualem-van Genuchten soil hydraulic functions ~~for~~ to ~~describing~~ describe the saturation-pressure relation. The necessary soil hydraulic parameter values for the different property layers (porosity, residual saturation, van Genuchten α, van Genuchten β, saturated hydraulic conductivity, Tab. 1) were assigned according to Glaser et al. (2016). They derived the parameter values based on in-situ field investigations (ERT profiles) and laboratory measurements. Soil samples were collected for the laboratory measurements from eight soil profiles distributed across the

Kommentiert [B2]: We added labels a), b) and c) for the different parts of the figure

catchment as well as from the shallow soil (5 cm and 35 cm) of nine locations in the headwater region, including six samples in the hillslope-riparian-stream zone of area M3. Furthermore, Glaser et al. (2016) relied on literature values for the parameterisation of the deeper property layers and performed minor manual value calibration for porosity and saturated hydraulic conductivity against stream discharge measured up- and downstream of area M3 and soil moisture measurements at locations TSM 1-5 (cf. Fig.1). In this study, ~~W~~we ~~only~~ added new parameter ~~parameterised~~ values for one additional layer for the fractured bedrock (Cv (ii)) in order to account for the adapted depth of 5 m in the catchment model compared to the depth of 3 m in the headwater model.

Table 1: Soil hydraulic parameters of the different soil property zones. Table adapted from Glaser et al. (2016)

Soil property zone	Residual saturation	van Genuchten parameter α [m^{-1}]	van Genuchten parameter β	Porosity	Saturated hydraulic conductivity [m d^{-1}]
Ah	0.12	6.6	1.46	0.74	1.71E+01
B1	0.10	22.1	1.42	0.61	1.71E+01
B2	0.10	22.1	1.42	0.45	4.59E+01
B2-IIC (i)	0.10	22.1	1.42	0.3	9.30E+02
B2-IIC (ii)	0.10	22.1	1.42	0.15	2.04E+03
IIC	0.02	6.0	1.50	0.20	8.40E+02
IIC-Cv	0.02	6.0	1.50	0.15	3.00E+00
Cv (i)	0.02	6.0	1.50	0.10	1.20E-02
Cv (ii)	0.02	6.0	1.50	0.07	1.20E-02
Cv-mC	0.02	6.0	1.50	0.05	9.00E-04
mC	0.02	6.0	1.50	0.01	2.40E-05
LP	0.10	22.1	1.42	0.61	7.80E+00

Surface and subsurface flow were coupled via a Darcy flux exchange through a thin coupling layer (10^{-4} m). We assumed different Manning's surface roughness values for the forested area ($1.24*10^{-6}$ d $m^{-1/3}$), the riparian zone ($9.41*10^{-7}$ d $m^{-1/3}$), and the stream bed ($4.4*10^{-7}$ d $m^{-1/3}$) (cf. Glaser et al., 2016). Evapotranspiration ~~properties~~ parameters (Tab. S1) were assigned individually for the deciduous forest, the coniferous forest in the southeast of the catchment, and the riparian zones including the streambed and values were based on the ~~calibrated~~ values of Glaser et al. (2016), who assigned the values according to literature values, estimates from field conditions, and calibration against stream discharge measured up- and downstream of area M3 and soil moisture measurements at locations TSM 1-5 (cf. Fig.1). The simulation was driven with daily sums of precipitation and reference evapotranspiration, which were treated as being spatially uniform. A critical depth boundary was assigned to the ~~The~~ outer edge of the surface domain ~~was assigned as critical depth boundary~~, allowing water to leave the model domain via surface flow. Side and bottom boundaries of the subsurface domain were no- flow boundaries. A spin-up simulation drained the catchment from full saturation to steady state conditions (for 1 mm d^{-1} of precipitation, no evapotranspiration) and subsequently repeated the period from October 2013 to October 2015 three times ~~for~~ to obtain~~ing~~ realistic initial conditions. The actual simulation spanned ~~over~~ the period from October 2015 to January 2018, which is the period where we mapped surface saturation with TIR imagery.

3.2 Assessment of model performance

We evaluated ~~benchmarked~~ the model performance with ~~against measured~~ discharge, groundwater level, and soil moisture measured ~~, and surface saturation patterns and dynamics~~ at various locations within the catchment (Fig. 1). We calculated the Kling Gupta Efficiency (KGE) as a combined measure for correlation, bias, and relative variability (Gupta et al., 2009)

between simulated and observed discharge. We also calculated KGEs as combined evaluation criteria for the simulated groundwater levels.~~, but particularly~~ Additionally, we particularly evaluated the simulated groundwater level dynamics rather than absolute values based on Pearson correlation coefficients. Soil moisture was also evaluated based on its dynamics with Pearson correlation coefficients, while absolute values were only compared visually. Since simulated soil moisture was extracted from model nodes whose depths did not exactly correspond with the measurement depths, we interpolated depth-weighted average values from the model output ~~for~~ to ~~calculating~~ calculate the correlation with the observations in the respective depths. The interpolated model values of volumetric water content were then correlated with the observations of water content, averaging the measurements of the ~~twin~~ two depth profiles at ~~the~~ each monitoring site~~s~~.

Simulated surface saturation was extracted from the surface domain of the model based on the simulated surface water depths. We classified the cells of the surface domain as saturated if simulated surface water depths were $>10^{-4}$ m, consistent with the definition used to determine surface saturation with the TIR images (i.e. surface saturation is water standing or flowing on the surface, cf. Section 2.2.). The depth of 10^{-4} m corresponds to the penetration depth of the used TIR camera for water columns and thus is the minimum depth that could be detected based on the pure water temperature signal with the camera. The applied definition of simulated surface saturation in combination with the explicit consideration of a subsurface and surface domain in the HGS model allows a differentiation between all water that is standing or flowing on the surface (i.e. surface saturation) and a fully saturated soil surface (i.e. soil water pressure head is zero, but water is not necessarily ponding or flowing on the surface). This implies that the simulated surface saturation can be the result of different processes, i.e. infiltration excess, saturation excess, subsurface water exfiltration, and overland flow. In order to qualitatively assess the importance of saturation excess and groundwater exfiltration in comparison to infiltration excess and overland flow, we compared the simulated frequency of surface saturation with the simulated frequency of groundwater reaching the surface across the entire catchment. The frequency map of surface saturation was generated based on simulated water depths $>10^{-4}$ m in the surface domain at noon of the days where TIR images were taken. Groundwater reaching the surface was identified based on the saturation characteristics of the subsurface domain of the model at noon of the days where TIR images were taken. We marked a cell of the surface domain as a cell where groundwater reached the surface if the subsurface domain below the surface cell was fully saturated from the bottom to the top. This information was then transformed into a frequency map analogous to the procedure for creating the surface saturation frequency maps.

For ~~comparing~~ comparison of the simulation output with the surface saturation information obtained with the TIR images, it was necessary to convert the model output into a comparable format ~~via several processing steps: First~~and perspective. ~~, we extracted the surface water depths in the surface domain of the model for noon of the days where TIR images were taken and analysed. Next, w~~First, we transformed the simulated surface water depths for the days with TIR images into ~~a~~ binary saturation maps of the entire catchment following the processing described above. ~~by classifying the surface domain cells as saturated if water depths were $>10^{-4}$ m. The depth of 10^{-4} m corresponds to the penetration depth of the used TIR camera for water columns and thus is the minimum depth that could be detected as pure water temperature signal with the camera. Finally, we~~ Then, we converted~~projected~~ the model output into jpeg images with the same perspective and extent of the TIR panoramic images by turning, bending, and cutting the modelled saturation maps according to each of the seven riparian areas individually. This model output processing allowed us to perform the same calculations for the model output as for the TIR images, i.e. to create time series of ~~normalized~~ normalised surface saturation extent, estimate the Kendall correlation for the relationship between normalised surface saturation extent and catchment discharge, and generate maps of ~~normalized~~ normalised saturation frequencies for the seven riparian areas with comparable perspectives and extents. Since it was not possible to project the model output identically to the perspectives of the TIR images, the calculation of quantitative performance metrics for the evaluation of the simulated time series of saturation and simulated frequency maps would have been biased by differences in image distortions and total area extent. Therefore, we evaluated the simulated surface saturation dynamics and patterns qualitatively only by visually comparing the observed and simulated time series of

normalised amounts of saturated pixels and saturation frequency maps generated from the TIR and model images. ~~compared the saturation dynamics and patterns of the model images with the observations qualitatively (visually) only. A quantitative comparison would have been biased by differences in image distortions and total area extent.~~

~~Furthermore, we compared the simulated frequency of surface saturation with the simulated frequency of groundwater reaching the surface. To do this, we marked the surface cells below which the subsurface domain was fully saturated from the bottom to the surface as cells where groundwater reached the surface. This binary information was transformed into a frequency map analogous to the procedure for creating the surface saturation frequency maps, using the same output times.~~

4 Results

4.1 Simulation of discharge, groundwater level, and soil moisture

The model reproduced the seasonal dynamics of measured discharge very well (Fig. 3, Fig. S1). The best fit was obtained at the outlet (SW1) with a KGE of 0.74. Discharge at SW2, SW3, and SW4 was reproduced ~~equally~~ reasonably well with KGEs of 0.49, 0.48, and 0.47. Groundwater levels were captured well with the model at the locations close to the riparian zone (KGE=0.57, r=0.78 for GW2; KGE=0.64, r=0.84 for GW3). At hillslopes and plateau sites, simulated groundwater levels were similar to the observed levels during the wet season, but during dry conditions the groundwater levels did not fall deep enough (Fig. 3, Fig. S1). This ~~level~~ discrepancy was reflected in low KGEs (0.30 for GW1, 0.21 for GW5, 0.02 for GW7). However, the general dynamics of levels increasing and decreasing were also captured at hillslopes and plateau sites (r = 0.66 for GW5, r = 0.62 for GW7, and r = 0.76 for GW1; note that the value for GW1 only includes data for wet periods, since the piezometer fell dry during summer months).

Simulated soil ~~moisture~~ water content generally showed a transition from higher to lower responsiveness from topsoil to subsoil layers consistent with the monitored soil moisture (Fig. 3, Fig. S1). ~~and~~ Pearson correlation coefficients indicated overall a good agreement between simulated and observed soil moisture dynamics (Tab. ~~1~~2). As for the groundwater levels at the hillslopes and plateau, soil moisture observations showed a distinct decrease in water content during dry periods, which the simulation could not reproduce to the same extent. The observed water content in the riparian zone was always close to saturation ~~(TSM4, Fig. 3)~~, while the simulation showed a decrease in water content during dry periods (TSM4, Fig. 3) ~~in the riparian zone~~. Yet the simulation also showed a spatial trend for more permanent soil saturation in the riparian zone (TSM4) and its vicinity (TSM3, Fig. S1) than at the hillslopes and plateau sites. The simulated values of water content were similar to the observed values at some locations (e.g. TSM2, SM4, Fig. 3) and clearly differed at other locations (e.g. SM7, Fig. 3), but the match~~es~~ and mismatch~~es~~ of the volumetric water content did not clearly depend on specific areas or landscape units. Moreover, we think that moisture dynamics and responsiveness are more informative for model performance than the absolute water content values, since also the measured values of volumetric water content differed ~~from each other~~ within small distances (e.g. measurements of water content in 10 cm depth at profile SM7, Fig. 3).

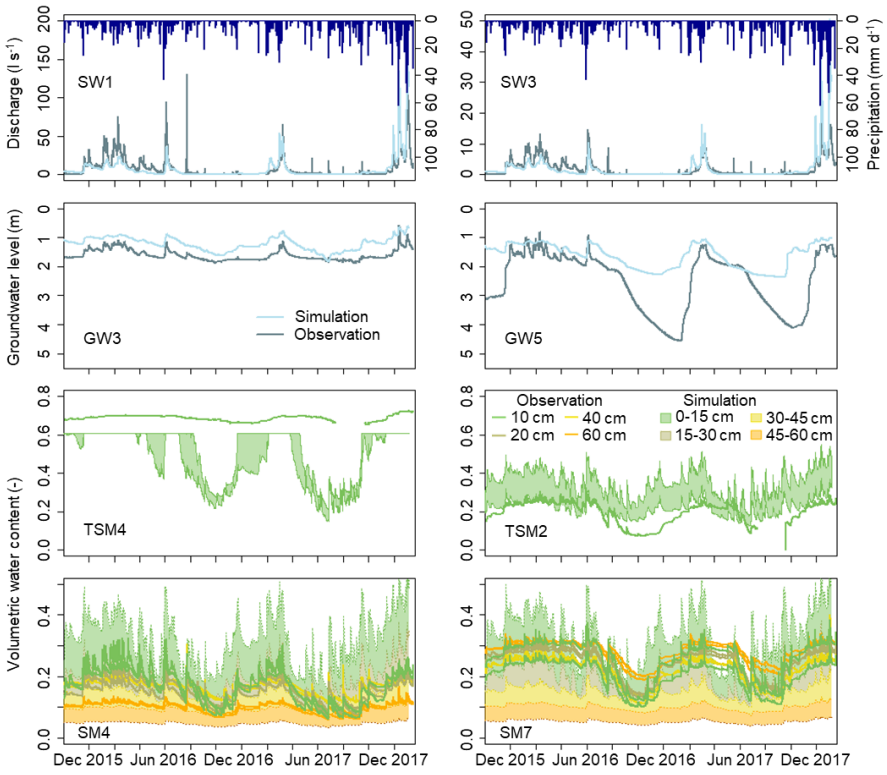

Figure 3: Simulated and observed time series of discharge, groundwater level below the surface, and volumetric water content. Colour bands indicate the possible span of simulated volumetric water contents in the depths between two model nodes. The time series of the observation locations (cf. ~~Figure~~ Fig. 1) that are not shown here, are shown in the supplemental material (Figure S1).

Table 2: ~~Coefficients of~~ Pearson correlation coefficients for the relation between simulated and observed dynamics of volumetric water content of the soil for the different measurement locations and depths (cf. Fig. 1).

	SM3	SM4	SM5	SM7	TSM1	TSM2	TSM3	TSM4	TSM5
10 cm	0.54	0.75	0.70	0.59	0.60	0.62	0.67	0.30	0.85
20 cm	0.67	0.82	0.76	0.62					
40 cm	0.82	0.89	0.88	0.79					
60 cm	0.85	0.92	0.91	0.82					

4.2~~5~~ Simulated patterns and dynamics of surface saturation versus groundwater reaching the surface at catchment scale

Simulated surface saturation (water depth $>10^{-4}$ m in the surface domain) generally occurred only in the streambed and adjacent riparian zones (Fig. ~~7a~~4a). During the wettest conditions of the study period (winter 2017/2018), surface saturation also occurred as extension~~prolongation~~ of the right ~~eastern stream branch~~tributary into the hillslope above the source area R3 (cf. Fig. 1). This simulated occurrence ~~behaviour~~ of surface saturation across the catchment is ~~in accordance~~ consistent with field evidence, ~~where~~since we observed surface saturation outside of the valley bottom only during very wet conditions or rain on snow events (cf. Section 2.2). The simulated patterns of where and how frequently groundwater reached the

ground surface (full saturation of the subsurface domain, Fig. ~~7b~~4b) ~~proved to be~~were very similar to the surface saturation frequency map of the catchment (Fig. ~~7a~~4a). The only obvious difference occurred in the area above the source area of the ~~eastern stream branch~~right tributary (R3), with a smaller extent of groundwater reaching the surface than the ~~extent of~~ surface saturation extent.

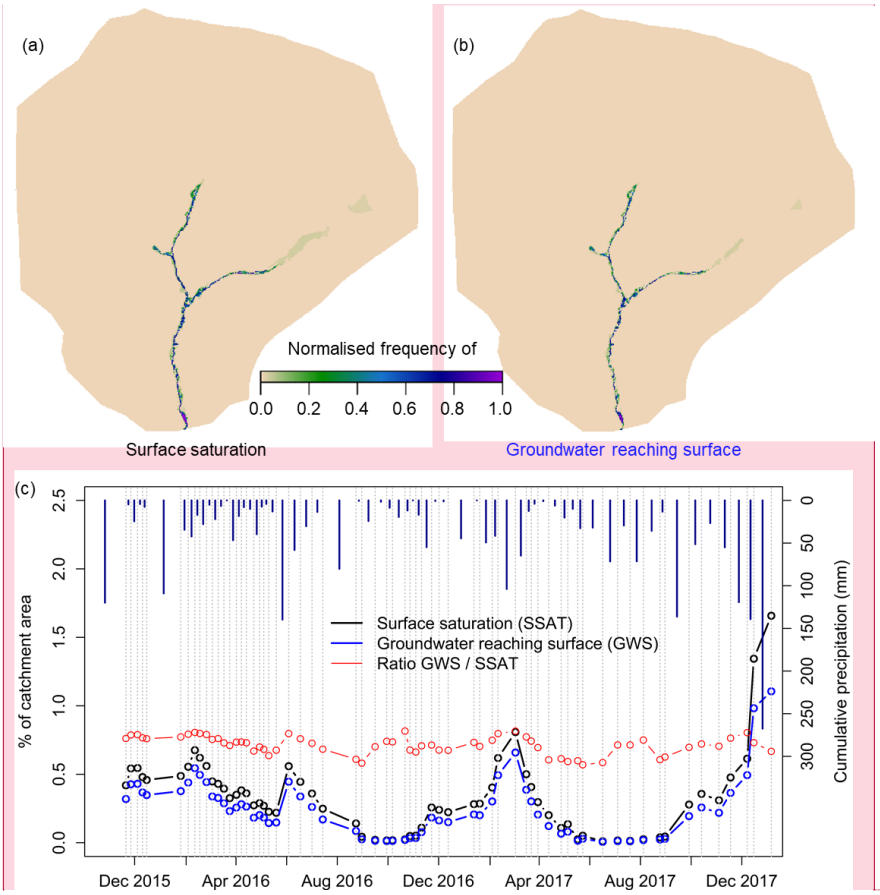

Figure ~~4~~7: Simulated frequency maps (a, b) and time series of percentage (c) of surface saturation and groundwater reaching the surface in the Weierbach catchment. Surface saturation (a, black line in c) corresponds to a simulated water depth >10^{-4} m in the surface domain of the model. Groundwater reaching the surface (b, blue line in c) corresponds to a complete saturation of the subsurface domain of the model, independent of the wetness state of the surface domain. Precipitation is given as cumulative amounts between the observation dates (grey dashed lines).

The time series of simulated percentage of catchment area with surface saturation and groundwater reaching the surface revealed that the area where groundwater reached the surface was always smaller in extent than the surface saturated area, even after dry condition (Fig. ~~4~~7c). The biggest absolute difference between the areal extent of surface saturation and groundwater reaching the surface was simulated during winter 2017/2018 (1.66 % vs 1.1 % of catchment area), where the conditions were very wet with high discharge and high cumulative precipitation and where the difference in areal extent was also visible in the frequency maps (Fig. ~~7a~~ 4a and b). However, the ratio between the extent of groundwater reaching the surface and the extent of surface saturation was not exceptionally high during winter 2017/2018. Instead, the ratio

Kommentiert [B3]: We replaced 'Fraction of time of' by 'Normalised frequency of'

variedscattered without a clear trend between 0.57 and 0.82 during the entire simulation period, apparently independent from the cumulative amount of precipitation or surface saturation.

4.23. Temporal dDynamics of surface saturation extent within distinct riparian areas

The observed dynamicstime series of observed normalized normalised surface saturation extent (Fig. 45, coloured lines) were similar for all seven investigated riparian areas and followed the same seasonal trend asof the catchment discharge. Yet some differences between the studied areas were discernible. For example, saturation was less persistent between February and April 2016 in the two areas without perennial springs (M1, R2, Fig. 45) than in the other areas. Maximum saturation was reached in December 2017 at M1, R2 and S2, but between February and April 2016 at the other locations (Fig. 45). As forSimilar to the observations, the simulated dynamics of normalizsed surface saturation (Fig 45, black lines) followed the general trend of the simulated discharge dynamic. The simulation showed a faster decrease and increase of the normalizsed saturation during dry periods compared to what than it was observed in most areas. However, simulated discharge also seemed to decrease and increase earlier than it was the observed discharge (c.f. Section 4.34). The simulated saturation dynamics did not clearly differ between the different locations and thus behaved more synchronous than the observations (e.g. maximum simulated saturation in December 2017 in all areas). As a result, the match between simulated and observed dynamics of normalizsed saturation was better for some areas (e.g. M1, R2, Fig. 45) than for others (e.g. S2, L1, Fig. 45).

The dynamic changes of normalizsed simulated saturation matched the normalizsed observations generally well, despite of under- and over-estimationed amounts of the minimum and maximum absolute saturation for all areas. The minimum amount number of saturated pixels in the TIR panoramas ranged between 0.02 % at M3 and R3 and 3.38 % at S2, while the model did not simulate any surface saturation during the driest period (Fig. 45). In addition, simulated normalizsed saturation stayed longer close to the minimum than the observed saturation for several areas (L1, S2, M1). These results show that the model simulated a stronger dry-out than observed in the Weierbach. At the same time, the simulation overestimated maximum saturation in the riparian zone (Fig. 45). The overestimation was not equally strong at the seven investigated areas and as a result, the distinction between areas showing with higher or lower maximum saturation was not the same for the observations and simulations (e.g. R3 showing one of the highest maximum saturation in the observation, but one of the lowest maximum saturation in the simulation compared to the other areas).

4.3 4 Discharge – surface saturation relationship Relationship between surface saturation extent and catchment discharge for distinct riparian areas

The PearsonKendall correlation between normalizsed surface saturation and discharge at the outlet SW1 was > 0.65 0.60 for both the simulation and the observation in almost all riparian areas, indicating a monotonic relationship between surface saturation and discharge in all areas . L1 was the only exception with r_{obs} = 0.54 (Fig. 56). The simulated relationships between normalizsed surface saturation and catchment discharge resembled the observed relationships in terms of value range and shape (Fig. 56), although the observation data were distinctly more scattered distinctly more than the simulation data. A power law relationship approximated the observed relationship between discharge and saturation for all seven areas, when data that were taken during rainfall or rising discharge were excluded (cf. Antonelli et al., 2019). For some areas, the simulation matched the trend lines of the observation data closely (e.g. L1, M2). For other areas, the visual fit of the model output to the observation data was less good (e.g. S2, R3), but still described a similar trend.

Despite the common shape of a power law function, the saturation – discharge relationships were slightly different between the different areas, both for observation and simulation data. For example, the power law functions fitted to the observations showed that saturation during high flow conditions (> 5 l s^{-1}) increased most strongly with discharge in the sources areas (especially M3 and R3). During low flow conditions (< 1 l s^{-1}), the source areas (L1, M3, R3) showed the lowest amount of the normalizsed saturation and its the least change relative to discharge was smallest in the source areas (L1, M3,

R3)~~compared to the other areas~~. In the simulated relationships, the increase in saturation for high discharge (> 5 l s^{-1}) was strongest for M3 and S2. The simulated relationship between discharge and surface saturation during low flow (< 1 l s^{-1}) was similar for all areas in terms of slope, but differed in the amount of normali~~s~~~~z~~ed saturation, being highest for areas in the right tributary~~east stream branch~~ (R2, R3), followed by the areas in the middle ~~upstream branch~~tributary (M1, M2, M3), and L1 and S2.

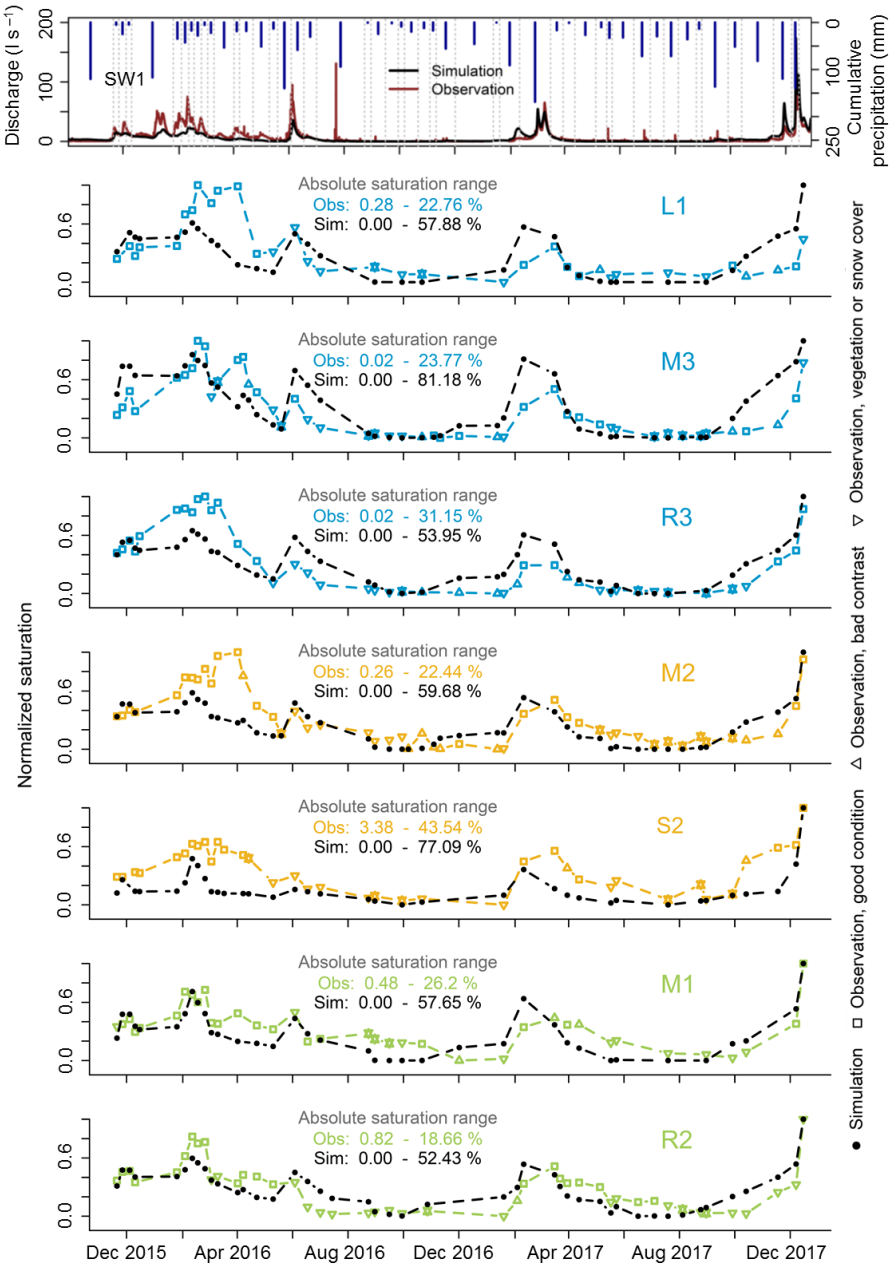

Figure 54: Time series of observed and simulated surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond to the colours of the icons in Fig. 1 and represent the different categories of riparian areas. Surface saturation is normalised to the minimum and maximum amount of saturation that was observed and simulated in the individual areas, respectively. Observations that were derived from TIR images with a poor temperature contrast or with influences of vegetation and snow cover are deemed less reliable and are marked with triangular symbols (see symbol legend displayed on the right). Cumulative precipitation between the measurement dates

(grey dashed lines) and discharge at catchment outlet SW1 are shown in the top panel ~~for~~ ~~to~~ ~~facilitating~~ facilitate the comparison to precipitation and flow conditions.

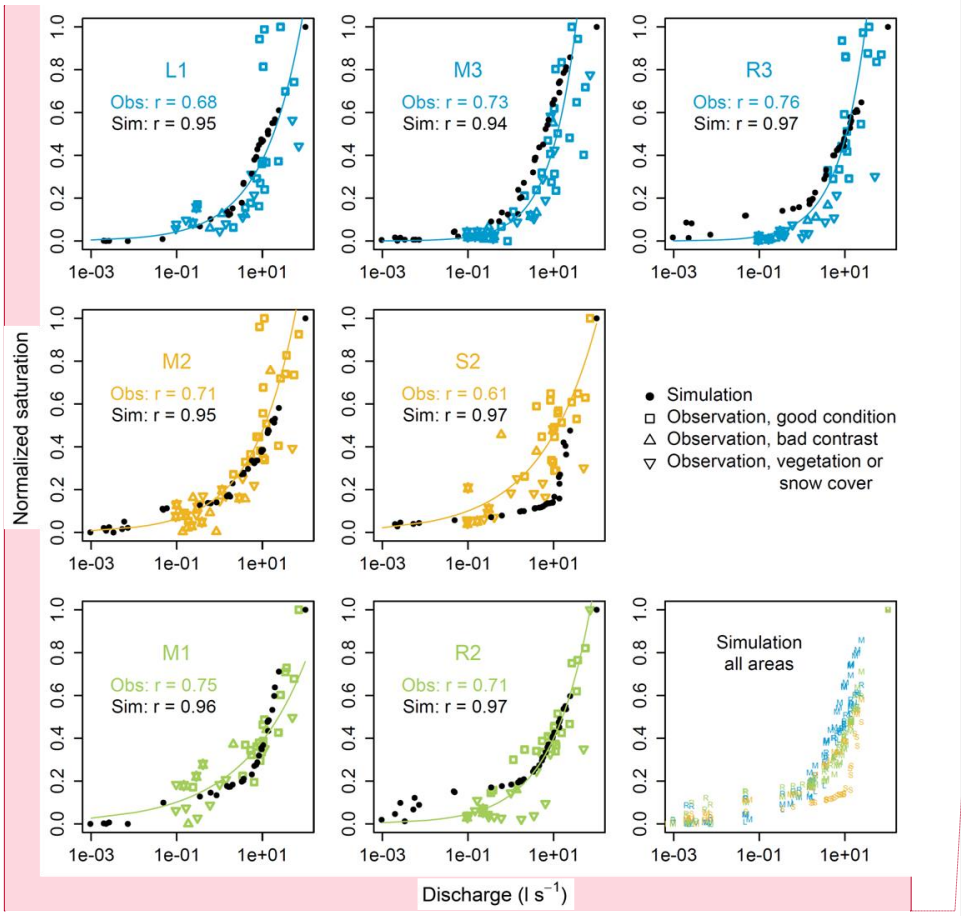

Figure ~~6~~5: Observed and simulated relationships and ~~Pearson~~ Kendall correlations between normali~~z~~ed surface saturation and discharge at the catchment outlet SW1 for the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond to the colours of the icons in Fig. 1 and represent the different categories of riparian areas. Observations that were derived from TIR images with a poor temperature contrast or with influences of vegetation and snow cover are deemed less reliable and are marked with triangular symbols. Solid lines are power law curves fitted to the observation data, excluding data taken during rainfall or rising discharge. ~~For~~ To ~~facilitating~~ facilitate the comparison between the seven areas, the ~~panel on the bottom right contains the~~ simulated data points from all seven areas ~~and~~are shown with the different colours and letters of the respective areas in the panel on the bottom right ~~the area affiliation is indicated with the respective colour and letter~~.

4.~~4~~ 5 Spatial patterns of surface saturation: Occurrence and frequencies within distinct riparian areas

The realism of simulated patterns of surface saturation was evaluated for each riparian area by visually comparing the surface saturation frequency maps obtained from the simulations and observations (Fig. ~~6~~7). The model captured the location of the stream and the locations that intermittently became surface saturated well for most of the seven investigated areas. For example, both observation and simulation showed that only the right side of the stream became saturated in M1, that the

Kommentiert [B4]: Values of Pearson correlation were replaced by Kendall correlation

riparian zone of the right streamside in M2 became saturated only in the upstream part, and that saturation mainly developed on the left streamside in R3, surrounding some permanently dry areas next to the stream (Fig. ~~6~~7). The only area with a clear mismatch between observed and simulated patterns of surface saturation was area L1, where surface saturation was simulated on the opposite streamside and at a clearly wrong position along the stream (upstream vs downstream).

The simulated surface saturation also reflected the observed saturation frequencies well. The simulation reproduced the general picture of more frequent surface saturation in the streambed than at the streamsides, but - as for the saturation patterns - simulated and observed frequencies corresponded better in some areas (e.g. S2, Fig. ~~6~~7) than in others (e.g. R3, Fig. ~~6~~7). For example, the observed frequency of surface saturation in the streambed was generally lower in the source areas (L1, M3, R3) than in the mid- and downstream areas (M2, S2, M1, R2), while the simulated frequency of surface saturation in the streambed was more similar between the areas and particularly overestimated in L1 and R3.

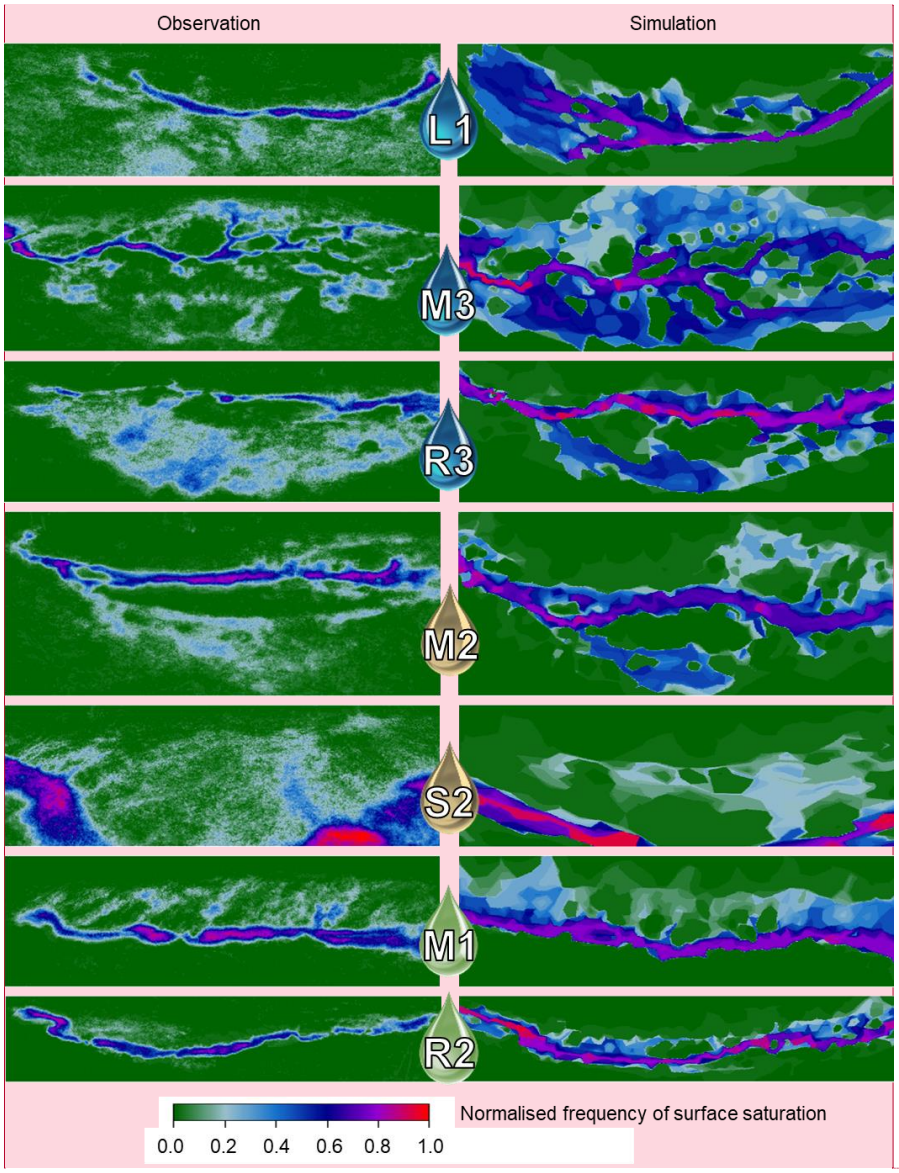

Figure 76: Observed (left) and simulated (right) frequencies of surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary, and the central stream (S2) (cf. Fig. 1). The maps were created by first counting how often the individual pixels were classified as saturated in the individual panoramic images and second normalizing the resulting frequency numbers by the total number of images analysed for the respective area.

Kommentiert [B5]: We replaced 'Fraction of time of saturation' by 'Normalised frequency of surface saturation'

~~4.5 Simulated patterns and dynamics of surface saturation versus groundwater reaching the surface at catchment scale~~

5 Discussion

The aim of this study was to use an ISSHM as complementary tool to field observations to analyse the spatio-temporal variability of surface saturation within the Weierbach catchment, with a focus on the stream valleys and riparian zones. ~~Even though~~ We found some discrepancies between observed and simulated discharge, groundwater levels, and soil moisture ~~showed some discrepancies to observations~~ in terms of absolute values. Particularly, the model had some problems to reproduce soil moisture and groundwater levels during the dry conditions at hillslopes and plateau. ~~, we~~ We ~~would~~nonetheless argue that the ~~performance of~~ match between the ~~different~~ observed and simulated time series of discharge, groundwater levels and soil moisture at different locations was quite good for a model that was not explicitly calibrated against the different time series distributed across the catchment but set up with uniform parameters. ~~and set up rather homogenously across the catchment. While the model had some problems to reproduce soil moisture and groundwater levels during the dry conditions at hillslopes and plateau,~~ Moreover, the simulated time series of soil moisture and groundwater levels matched the observations especially well in the riparian zone and vicinity. This gives us confidence that the model setup was valid for evaluating and analysing the spatio-temporal dynamics of surface saturation and its intra-catchment variability.

5.1 Temporal dynamics of surface saturation extent

The model reproduced well the observed ~~long-term~~ dynamics of surface saturation in the seven investigated riparian areas over different seasons and wetness conditions ~~well~~. Our work~~study~~ goes beyond previous studies~~works~~ that compared the simulation of surface saturation dynamics with observations (e.g. Ali et al., 2014; Birkel et al., 2010; Glaser et al., 2016; Mengistu and Spence, 2016) by relying on a longer study period and a ~~higher~~ larger number of observations in time. This allowed us to analyse and compare various hydrological conditions and the dynamic transition between them over all seasons with a ~~frequent~~large number of observations. Moreover, we accounted for spatial variability of saturated area dynamics within the catchment. Unlike the various quasi dynamic wetness indices presented in Ali et al. (2014), which could not satisfyingly reproduce the spatio-temporal variability of connected surface saturation observed in a catchment in the Scottish Highlands, our model reproduced the distributed dynamics of surface saturation well, without clear ~~performance~~ differences in performance for different wetness conditions.

Simulations and observations showed both that the temporal dynamics of the extent of surface saturation were mostly consistent across the catchment. Moreover, our simulations showed that the spatio-temporal development of surface saturation was very similar to the spatio-temporal dynamics of groundwater reaching the surface (cf. Fig. ~~7~~4). This suggests that the generation of surface saturation in the Weierbach catchment is largely driven by the synchronous exfiltration of groundwater in topographic depressions. The high hydraulic conductivities of the upper soil layers that we implemented in the model (cf. Tab. 1) already imposed that surface saturation due to infiltration excess was unlikely to be simulated. This model parameterisation was chosen based on field observations and the previous simulation of the 6 ha headwater around area M3 (cf. section 3.1, Glaser et al., 2016). In addition, the parameterisation is in line with the common assumption that surface saturation in forested catchments is mainly generated by saturation excess rather than infiltration excess (cf. e.g. Dunne et al., 1975; Hewlett and Hibbert, 1967; Latron and Gallart, 2007; Megahan and King, 1985; Weill et al., 2013). This assumption proved to be also valid for the Weierbach catchment, since the simulations with the chosen parameterisation captured the dynamics of surface saturation extent observed across the catchment. The simulation results furthermore helped to specify that surface saturation in the Weierbach is not the result of saturation excess on any (perched) saturated soil, but that it is in large parts controlled by a synchronous fluctuation of the groundwater levels across the catchment. Antonelli et

al. (2019) drew consistent conclusions with our simulation results based on a statistical analysis of the observation data. They analysed the relation between the observed surface saturation dynamics and various hydrometric measurements (i.e. discharge, groundwater levels, soil moisture, field-data-based estimates of catchment storage, precipitation, evapotranspiration) and found that the observed dynamics of surface saturation extent were particularly well correlated to the measured near-stream groundwater level fluctuations.

5.2 Relationship between surface saturation extent and catchment discharge

We found that the observed and simulated relationships between surface saturation and catchment discharge resembled a power law relationship for all areas~~s~~ (cf. Fig. ~~5~~6). This is consistent with earlier studies that showed power law relationships between contiguous connected surface saturated areas and discharge (Mengistu and Spence, 2016; Weill et al., 2013). In contrast to these studies, we did not observe clear hysteretic loops in the relationship between saturation and ~~high~~ streamflow. Nonetheless, the scatter in the observed discharge – surface saturation relationships might indicate that the development of surface saturation in the Weierbach catchment follows hysteretic loops, but that the hysteresis was not resolved with the available temporal resolution of the observations. For example, it is likely that surface saturation evolved in the riparian areas during high flow conditions and persisted on the ground surface during decreasing streamflow due to restricted infiltration capacities of the riparian soil (cf. Antonelli et al., 2019).

The lack of such a hysteretic process in the simulation could explain why the model showed the tendency for less persistent and faster contracting surface saturation. It may also explain why the simulated saturation dynamics differed less between the different investigated areas than the observed dynamics. It is likely that the observed saturation dynamics were not synchronous between the different areas due to a less persistent (and thus hysteretic) generation of surface saturation in the relatively narrow riparian areas without perennial springs (M1 and R2) compared to the wider riparian areas with perennial springs (cf. observation of less persistent saturation in M1 and R2 during February and April 2016, Fig. ~~4~~5). The model, instead, simulated a non-hysteretic saturation behaviour for all investigated riparian areas, which resulted in a better fit between simulated and observed dynamics in the areas M1 and R2 compared to the other areas.

At the same time, it might also be that the simulated relationship between saturation and discharge was correct in all riparian areas and that the scatter~~ing~~ of the observation data did not result from hysteretic behaviour, but from uncertainties in the TIR methodology. A good argument for a correct simulation of the discharge – surface saturation relationship is that not only simulated saturation but also simulated discharge seemed to be less persistent and to decrease and increase earlier than it was observed. In reality, the scatter of the observation data is likely related to both measurement uncertainties and hysteretic aspects and a future study with higher temporal resolution of field observations and corresponding simulation output could further analyse this.

Independently from the question on hysteretic loops, we found that the discharge – surface saturation relationships somewhat differed between the different areas. We could connect the main differences to different topographical and morphological features, yet we cannot decipher why the main controlling feature for the discharge – surface saturation relationship was different between observations (source areas vs non-source areas) and simulations (different ~~stream branches~~tributaries, cf. Section 4.~~3~~4). Nonetheless, our findings are in line with experimental studies that discussed that the relationships between baseflow discharge and total extent of contributing saturated areas differ between catchments with different physiographic characteristics (e.g. Dunne et al., 1975; Latron and Gallart, 2007).

~~By comparing our model results to the double logarithmic plot~~ The degree of variability of the discharge – surface saturation relationships for the different areas studied within the Weierbach catchment is comparable to the variability of the discharge – surface saturation relationships for different catchments presented by Latron and Gallart (2007) (Figure 8)~~, we could identify similar shape varieties of the discharge – surface saturation relationship for the different areas studied within the Weierbach catchment as observed for the different catchments presented in Latron and Gallart (2007)~~. We cannot compare

our results directly with the results shown in Latron and Gallart (2007), since we evaluated absolute discharge and normalis~~z~~ed saturation, while they evaluated connected saturated areas in percentage of catchment area, but discharge normalis~~z~~ed to the catchment area. In order to facilitate the comparison and to connect the two plots (Fig. 8a, 8b), we show the simulated relationship between discharge and surface saturation of the entire Weierbach catchment in both plots, once with normalis~~z~~ed discharge and absolute saturation (Fig. 8a), and once with absolute discharge and normalis~~z~~ed saturation (Fig. 8b). The shape of the relationship for the entire Weierbach catchment was nearly linear, similar to the relationship observed in the Can Vila catchment investigated by Latron and Gallart (2007) (Fig. 8a). The relationships of the seven studied riparian areas differed from the catchment relationship and between each other (Fig. 8b). For example, area S2 and M1 showed a convex shape similar to the observations in the Vermont W2 catchment made by Dunne et al. (1975), area M3 showed a rather concave shape similar to the relationships found for a sub-catchment of the Asker basin (Myrabø, 1986) and the Strengbach catchment (Latron, 1990), area M2 showed a rather linear shape similar to the Can Vila catchment studied by Latron and Gallart (2007). This clearly shows that differences in the relationship between surface saturation and discharge do not only occur between different catchments, but that they also occur within a catchment, highlighting ~~as~~ intra-catchment variability.

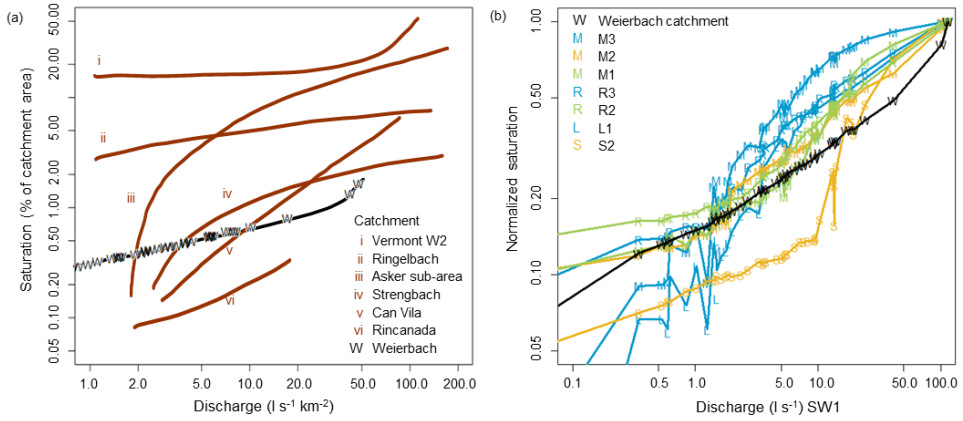

Figure 8: Simulated relationship between discharge and surface saturation of the entire Weierbach catchment (marked with W) in comparison to (a) the relationships observed in other catchments (Figure modified from Latron and Gallart (2007) and (b) the relationships simulated for the seven investigated riparian areas within the catchment. The presented relationships of the other catchments were investigated by i) Dunne et al. (1975), ii) Ambroise (1986), iii) Myrabø (Myrabø, 1986), iv) Latron (Latron, 1990), v) Latron and Gallart (2007), and vi) Martinez-Fernandez et al. (2005). Area affiliation for the investigated riparian areas of the Weierbach catchment is indicated with the respective colour and letter (cf. Fig.1, Fig. 4~~5~~-~~6~~7).

5.3 Spatial patterns of surface saturation occurrence

The observed spatial patterns of surface saturation occurrence were reproduced ~~with~~ well by the simulations ~~in great detail~~ for most of the investigated areas. We attribute the successful simulation of the spatial patterns to microtopography (local topographical features with extents of centimetres to few metres) since i) microtopography described the main spatial variability between the seven investigated areas in the model setup and ii) we observed that small changes in the setup and resolution of the model mesh in the riparian zones changed some details of the simulated surface saturation patterns (Fig. S2, especially area M2, S2). Therefore, we would like to stress that not only major topographic features of the catchment (e.g. hillslope shape, slope angle, valley width) but also its microtopography needs to be considered for identifying locations

where surface saturation may occur. This may sound trivial and several studies have already pointed out the importance of microtopography for the simulation of different hydrological aspects such as hydraulic heads, hyporheic surface-subsurface water exchange, bank storage and overbank flooding, water quality of shallow groundwater systems and runoff generation (e.g. Aleina et al., 2015; Frei et al., 2010; Käser et al., 2014; Van der Ploeg et al., 2012; Tang et al., 2018). Still, microtopography is not often considered in the simulation of surface saturation patterns.

When microtopography is not resolved ~~in~~detailed ~~enough~~ sufficient detail, it is more likely that the simulated surface water extends over a large area instead of being confined to occurring ~~in~~accumulating ~~in~~ topographic depressions and thus the model ~~overrates~~ overpredicts the extent of surface saturation. In this context it is interesting to note that there are studies that simulated maximum extents of surface saturation up to 80 % of the study area (Qu and Duffy, 2007; Weill et al., 2013), while field observations have only ~~reached~~ shown maximum extents up to 25 % - 50 % of catchment area (Ali et al., 2014; Birkel et al., 2010; Dunne et al., 1975; Mengistu and Spence, 2016) and often ~~show~~ suggest maximum extents around 10 % (Ambroise, 2016; Grabs et al., 2009; Güntner et al., 2004; Latron and Gallart, 2007; Tanaka et al., 1988). Microtopography might partly explain this discrepancy, even though the maximum extent of surface saturation certainly also depends on the climatic and physiographic conditions of the catchment and on the timing of the observations (e.g. baseflow conditions vs storm events). Moreover, the importance of microtopography for simulating surface saturation extent likely depends on catchment size. The two mentioned studies that simulated extremely high saturation extent (Qu and Duffy, 2007; Weill et al., 2013) were performed in small catchments (< 0.1 km^2), whereas ~~and there are some~~ studies that analysed the extent of surface saturation without considering microtopography in larger catchments (1 - 100 km^2) often simulated similar or ~~less~~ smaller maximum extents of surface saturation than observed ~~without considering the microtopography~~ (e.g. Ali et al., 2014; Birkel et al., 2010; Grabs et al., 2009; Güntner et al., 2004; Mengistu and Spence, 2016).

In our study, the simulated ~~maximum~~ extent of surface saturation reached a maximum of~~was~~ 1.6 % of catchment area during the very wet conditions in winter 2017/2018 (cf. Fig. 4c). This simulated maximum extent of surface saturation~~, which~~ is small compared to the estimation in other simulation studies, but ~~matches~~ is consistent with the observation that surface saturation commonly only occurs within the riparian zone and streambed (extent of 1.2 %) of the Weierbach catchment. Nonetheless, the realistically simulated small maximum extent of surface saturation for the entire catchment did not prevent that ~~also our~~ the maximum saturation within the individual areas was overestimated compared to the observations (cf. Fig. ~~5~~4). Besides the effect of microtopography, there are two other possible explanations for this. First, the largest simulated saturation extent occurred during winter 2017/2018, which is the same period where the model clearly overestimated discharge. This mismatch could partly explain the overestimation of saturation, assuming that the relationship between discharge and saturation was correctly captured with the model (cf. Section 5.2). Second, the overestimation of absolute saturation could result from different perspectives and extensions of model output and TIR images (cf. section 3.2, Fig. ~~6~~7). The TIR images included parts of the hillslopes around the riparian zones, which were not included to the same extent in the extracted model images. Since the hillslopes normally remained unsaturated, the maximum possible ~~amount~~ number of saturated pixels in the TIR images was thus lower than in the model images, while the minimum possible ~~amount~~ extent of saturation was not affected. This could also explain why overestimation of total ~~amounts~~ extent of saturation was different between the different areas.

Despite the importance of microtopography, the model results showed that microtopography alone was not sufficient to capture the spatial patterns of surface saturation occurrence correctly. The simulated patterns of surface saturation clearly did not match the observed patterns equally well in all seven investigated areas (cf. Fig. ~~6~~7), although the topographical information source and mesh resolution was consistent for the simulated riparian areas. This means that there are additional factors that control the spatial patterns of surface saturation occurrence that were not accounted for in the simulations. Such a factor could for example be the structure of the subsurface, which was treated as being homogeneous between all investigated riparian areas in the simulations. In reality, the subsurface structure may locally differ to some degree, for

example in the riparian area of the ~~western~~ left ~~stream branch~~tributary (L1), where saturation was simulated at the~~a~~ clearly wrong side along the stream.

5.4 Spatial patterns ~~Frequency maps~~ of surface saturation frequencies

The frequency maps of surface saturation (cf. Fig. 7) combine information on when and where surface saturation occurs. We do not think that the exfiltration of subsurface water into local depressions (cf. Section 5.1 and 5.2) can fully explain the spatial variability of saturation frequencies that was observed and simulated satisfactorily within the different riparian areas ~~(Fig. 6)~~. Instead, we assume~~suppose~~ that the differences in saturation frequency were controlled by additional water sources than exfiltrating groundwater, such as stream water or direct precipitation, and that the contribution of these additional water sources to surface saturation varied in space and time. For example, the lower frequencies of surface saturation ~~observations~~ observed at the streamsides compared to the streambed and the lower frequencies in the streambed of the source areas (L1, M3, R3) compared to the mid- and downstream areas (M2, S2, M1, R2) might reflect a lower and less frequent contribution of upstream water in these areas. The overestimation of simulated saturation frequencies in the streambed of R3 could thus indicate an overestimated upstream contribution due to simulating the stream extent too far upstream from the source area. Moreover, the fact that the simulated extent of groundwater reaching the surface was never as large as the simulated extent of surface saturation in the catchment (cf. Fig. 4) indicates that at least some locations are not exclusively surface saturated due to groundwater exfiltration. It is for example likely that the infrequently surface saturated area above the source area of the right tributary (R3) (cf. section 4.2) receives water from additional sources such as overland flow or direct precipitation. Future work should analyse potential water sources and generation processes of surface saturation with a suitable model framework (cf. Partington et al., 2013; Weill et al., 2013) in order to complement the interpretation of the observation data and to identify the mixture of different water sources of surface saturation (e.g. stream water, exfiltrating subsurface water, ponding precipitation), how the sources might vary in space and time, and how this might reflect in the surface saturation frequencies.

6 Summary and conclusions

We explored the intra-catchment variability of surface saturation in the 42 ha Weierbach catchment with joint observations and simulations. We showed that the model could reproduce the observed variability of the different surface saturation characteristics (dynamics, frequencies, patterns) with great detail, although the model setup was rather homogeneous and parameters were ~~not calibrated at catchment scale~~adopted without re-calibration from the 6 ha headwater model set up previously by Glaser et al. (2016). Our results demonstrated that a spatially distributed, physically-based, integrated hydrological model such as HGS is well-suited for reproducing and analysing the generation and development of surface saturation in space and time, if environmental conditions and characteristics are similar to those of the Weierbach catchment. Based on the identified match~~es~~ and mismatch~~es~~ between the simulation results and observations, we could identify groundwater exfiltration and microtopography as ~~some~~ key factors controlling the occurrence of surface saturation. Yet these two factors alone were not sufficient to explain the full variability of the different characteristics of surface saturation that were observed between the different areas.

The temporal dynamics ~~occurrence~~ of surface saturation extent ~~was~~ were observed and simulated to be similar across the catchment, which we related – based on the simulation results – to a large influence of synchronous groundwater level fluctuations across ~~that reacts synchronous across~~ the catchment. Still, we~~The~~ observed differences between the investigated riparian areas with regard to the seasonal dynamics of saturation extension and contraction and the surface saturation – discharge relationship. These differences likely ~~resulted from~~ relate to differing~~ent~~ morphological characteristics (width, existence of perennial springs) of the riparian areas that induce more or less persistent surface saturation and thus hysteretic

relationships between surface saturation extent and catchment discharge in some areas. ~~Although t~~The model could not fully reproduce a~~t~~he observed varying persistence of surface saturation and hysteretic relationships between surface saturation extent and catchment discharge. ~~occurrence and persistence of surface saturation in the different investigated areas, also~~ Nonetheless, the simulation results demonstrated that the shape of the relationship between surface saturation and discharge for different riparian areas ~~can differ~~ within a catchment can be as variable as it has been observed between different ~~in the same manner as between~~ catchments with different topographical and morphological conditions.

The spatial occurrence of ~~the~~ surface saturation differed between and within the seven investigated riparian areas, which we could mainly ~~could~~ relate to the influence of microtopography. Nonetheless, the model did not reproduce the spatial patterns of surface saturation occurrence equally well in all seven investigated areas. This suggests that some aspects that were not accounted for in the model setup, such as a spatial variability of subsurface structure, exhibit additional control on the spatial occurrence of surface saturation. Finally, the model could satisfactorily reproduce the observed patterns of surface saturation frequencies for the different riparian areas. We suggest that the spatially varying frequencies of surface saturation might reflect a locally varying relevance of different water sources. ~~Furthermore, we discussed that the full variability between the different areas and the mismatches between observations and simulation can only be explained with additional factors besides groundwater exfiltration and microtopography.~~

~~The spatially varying frequencies of surface saturation within the riparian areas indicated that there might be additional water sources than subsurface water that contribute to the generation of surface saturation.~~ Since the model could reproduce the observed frequencies, the model ~~can~~ may be used in a future study to analyse such a potential mixing of different water sources and their variation in space and time. ~~The observed differences between the investigated riparian areas with regard to the seasonal dynamics of saturation extension and contraction and the surface saturation – discharge relationship likely resulted from different morphological characteristics (width, existence of perennial springs) of the riparian areas. Although the model could not reproduce a varying hysteretic occurrence and persistence of surface saturation in the different investigated areas, also the simulation results demonstrated that the relationship between surface saturation and discharge can differ within a catchment in the same manner as between catchments with different topographical and morphological conditions.~~

Data availability. Data underlying the study are property of the Luxembourg Institute of Science and Technology. They are available on request from the authors.

Author contributions. BG, LH and JK designed and directed the study. BG and MA planned and carried out the field work and processed the TIR images. BG set up the simulation and processed the model output. BG, MA, LH and JK discussed and interpreted the results. BG prepared the manuscript with contributions from JK and LH.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. We wish to thank Jean-Francois Iffly, Jérôme Juilleret, the Observatory for Climate and Environment of LIST, and the Administration des Services Techniques de l'Agriculture (ASTA) for the collection and provision of the hydrometrical and meteorological data. We acknowledge deployment of a trial version of AlgoMesh by HydroAlgorithmics Pty Ltd. Barbara Glaser thanks the Luxembourg National Research Fund (FNR) for funding within the framework of the FNR-AFR Pathfinder project (ID 10189601). Marta Antonelli was funded by the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no. 607150 (FP7-PEOPLE-2013-ITN – INTERFACES – Ecohydrological interfaces as critical hotspots for transformation of ecosystem exchange

fluxes and bio-geochemical cycling). Three anonymous referees are thanked for their constructive comments and suggestions for improving the initial manuscript.

References

Ala-aho, P., Rossi, P. M., Isokangas, E. and Kløve, B.: Fully integrated surface–subsurface flow modelling of groundwater–lake interaction in an esker aquifer: Model verification with stable isotopes and airborne thermal imaging, J. Hydrol., 522, 391–406, doi:10.1016/j.jhydrol.2014.12.054, 2015.

Aleina, F. C., Runkle, B. R. K., Kleinen, T., Kutzbach, L., Schneider, J. and Brovkin, V.: Modeling micro-topographic controls on boreal peatland hydrology and methane fluxes, Bio, 12, 5689–5704, doi:10.5194/bg-12-5689-2015, 2015.

Ali, G., Birkel, C., Tetzlaff, D., Soulsby, C., McDonnell, J. J. and Tarolli, P.: A comparison of wetness indices for the prediction of observed connected saturated areas under contrasting conditions, Earth Surf. Process. Landforms, 39, 399–413, doi:10.1002/esp.3506, 2014.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration (guidelines for computing crop water requirements), FAO Irrig. Drain. Pap., 56, 1998.

Ambroise, B.: Rôle hydrologique des surfaces saturées en eau dans le basin du Ringelbach à Soultzeren (Hautes-Vosges), France, in Recherches sur l'Environnement dans la Région, Actes du 1er Colloque Scientifique des Universités du Rhîn Supérieur, edited by O. Rentz, J. Streith, and L. Ziliox, pp. 620–630, Université Louis Pasteur - Conseil de l'Europe, Strasbourg., 1986.

Ambroise, B.: Variable "active" versus "contributing" areas or periods: a necessary distinction, Hydrol. Process., 18, 1149–1155, doi:10.1002/hyp.5536, 2004.

Ambroise, B.: Variable water-saturated areas and streamflow generation in the small Ringelbach catchment (Vosges Mountains, France): the master recession curve as an equilibrium curve for interactions between atmosphere, surface and ground waters, Hydrol. Process., 30, 3560–3577, doi:10.1002/hyp.10947, 2016.

Antonelli, M., Glaser, B., Teuling, R., Klaus, J. and Pfister, L.: Saturated areas through the lens: 1. Spatio-temporal variability of surface saturation documented through Thermal Infrared imagery, Hydrol. Process., under review (HYP-19-0453), 2019.

Betson, R. P.: What is watershed runoff?, J. Geophys. Res., 69, 1541–1552, doi:10.1029/JZ069i008p01541, 1964.

Birkel, C., Tetzlaff, D., Dunn, S. M. and Soulsby, C.: Towards a simple dynamic process conceptualization in rainfall – runoff models using multi-criteria calibration and tracers in temperate, upland catchments, Hydrol. Process., 24, 260–275, doi:10.1002/hyp.7478, 2010.

Bracken, L. J. and Croke, J.: The concept of hydrological connectivity and its contribution to understanding runoff-dominated geomorphic systems, Hydrol. Process., 21(13), 1749–1763, doi:10.1002/hyp.6313, 2007.

Carrer, G. E., Klaus, J. and Pfister, L.: Assessing the Catchment Storage Function Through a Dual-Storage Concept, Water Resour. Res., 55, 476–494, doi:10.1029/2018WR022856, 2019.

Davison, J. H., Hwang, H.-T., Sudicky, E. A., Mallia, D. V. and Lin, J. C.: Full Coupling Between the Atmosphere, Surface, and Subsurface for Integrated Hydrologic Simulation, J. Adv. Model. Earth Syst., 10, 43–53, doi:10.1002/2017MS001052, 2018.

Dunne, T., Moore, T. R. and Taylor, C. H.: Recognition and prediction of runoff-producing zones in humid regions, Hydrol. Sci. Bull., 20, 305–327, 1975.

Erler, A. R., Frey, S. K., Khader, O., Orgeville, M., Park, Y.-J., Hwang, H.-T., Lapen, D. R., Peltier, W. R. and Sudicky, E. A.: Simulating Climate Change Impacts on Surface Water Resources Within a Lake-Affected Region Using Regional Climate Projections, Water Resour. Res., 55, 130–155, doi:10.1029/2018WR024381, 2019.

Frei, S., Lischeid, G. and Fleckenstein, J. H.: Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland --- A modeling study, Adv. Water Resour., 33(11), 1388–1401, doi:10.1016/j.advwatres.2010.07.006, 2010.

Gburek, W. J. and Sharpley, A. N.: Hydrologic Controls on Phosphorus Loss from Upland Agricultural Watersheds, J. Environ. Qual., 27, 267, doi:10.2134/jeq1998.00472425002700020005x, 1998.

Glaser, B., Klaus, J., Frei, S., Frentress, J., Pfister, L. and Hopp, L.: On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum, Water Resour. Res., 52, 8317–8342, doi:10.1002/2015WR018414, 2016.

Glaser, B., Antonelli, M., Chini, M., Pfister, L. and Klaus, J.: Technical note: Mapping surface-saturation dynamics with thermal infrared imagery, Hydrol. Earth Syst. Sci., 22(11), 5987–6003, doi:10.5194/hess-22-5987-2018, 2018.

Glaser, B., Jackisch, C., Hopp, L. and Klaus, J.: How meaningful are plot-scale observations and simulations of preferential

flow for catchment models ?, Vadose Zo. J., doi:10.2136/vzj2018.08.0146, 2019.

Gourdol, L., Clément, R., Juilleret, J., Pfister, L. and Hissler, C.: Large-scale ERT surveys for investigating shallow regolith properties and architecture, Hydrol. Earth Syst. Sci. Discuss., 1–39, doi:10.5194/hess-2018-519, 2018.

Grabs, T., Seibert, J., Bishop, K. and Laudon, H.: Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model, J. Hydrol., 373, 15–23, doi:10.1016/j.jhydrol.2009.03.031, 2009.

Güntner, A., Seibert, J. and Uhlenbrook, S.: Modeling spatial patterns of saturated areas: An evaluation of different terrain indices, Water Resour. Res., 40, W05114, doi:10.1029/2003WR002864, 2004.

Gupta, H. V, Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Hewlett, J. D. and Hibbert, A. R.: Factors affecting the response of small watersheds to precipitation in humid areas, in International Symposium on Forest Hydrology, edited by W. E. Sopper and H. W. Lull, pp. 275–290, Pergamon Press, Oxford. [online] Available from: http://coweeta.ecology.uga.edu/publications/851.pdf, 1967.

Juilleret, J., Iffly, J. F., Pfister, L. and Hissler, C.: Remarkable Pleistocene periglacial slope deposits in Luxembourg (Oesling): pedological implication and geosite potential, Bull. la Société des Nat. Luxemb., 112, 125–130 [online] Available from: http://www.snl.lu/publications/bulletin/SNL_2011_112_125_130.pdf, 2011.

Käser, D., Graf, T., Cochand, F., McLaren, R., Therrien, R. and Brunner, P.: Channel Representation in Physically Based Models Coupling Groundwater and Surface Water : Pitfalls and How to Avoid Them, Groundwater, 52(6), 827–836, doi:10.1111/gwat.12143, 2014.

Kristensen, K. J. and Jensen, S. E.: A model for estimating actual evapotranspiration from potential evapotranspiration, Nord. Hydrol., 6, 170–188, doi:10.2166/nh.1975.012, 1975.

Latron, J.: Caractérisation géomorphologique et hydrologique du bassin versant versant du Strengbach (Aubure), Université Louis Pasteur, Strasbourg I., 1990.

Latron, J. and Gallart, F.: Seasonal dynamics of runoff-contributing areas in a small mediterranean research catchment (Vallcebre, Eastern Pyrenees), J. Hydrol., 335, 194–206, doi:10.1016/j.jhydrol.2006.11.012, 2007.

Martínez-Carreras, N., Hissler, C., Gourdol, L., Klaus, J., Juilleret, J., Iffly, J. F. and Pfister, L.: Storage controls on the generation of double peak hydrographs in a forested headwater catchment, J. Hydrol., 543, 255–269, doi:10.1016/j.jhydrol.2016.10.004, 2016.

Martínez Fernández, J., Ceballos Barbancho, A., Morán Tejeda, C., Casado Ledesma, S. and Hernández Santana, V.: Procesos hidrológicos en una cuenca forestal del Sistema Central: Cuenca experimental de Rinconada, Cuad. Investig. Geográfica, 31, 7–25 [online] Available from: https://dialnet.unirioja.es/servlet/articulo?codigo=1975892, 2005.

Megahan, W. F. and King, P. N.: Identification of critical areas on forest lands for control of nonpoint sources of pollution, Environ. Manage., 9, 7–17, doi:10.1007/BF01871440, 1985.

Mengistu, S. G. and Spence, C.: Testing the ability of a semidistributed hydrological model to simulate contributing area, Water Resour. Res., 52, 4399–4415, doi:10.1002/2016WR018760, 2016.

Moragues-Quiroga, C., Juilleret, J., Gourdol, L., Pelt, E., Perrone, T., Aubert, A., Morvan, G., Chabaux, F., Legout, A., Stille, P. and Hissler, C.: Genesis and evolution of regoliths: Evidence from trace and major elements and Sr-Nd-Pb-U isotopes, Catena, 149, 185–198, doi:10.1016/j.catena.2016.09.015, 2017.

Munz, M., Oswald, S. E. and Schmidt, C.: Coupled Long-Term Simulation of Reach-Scale Water and Heat Fluxes Across the River-Groundwater Interface for Retrieving Hyporheic Residence Times and Temperature Dynamics, Water Resour. Res., (53), 8900–8924, doi:10.1002/2017WR020667, 2017.

Myrabø, S.: Runoff Studies in a Small Catchment, Nord. Hydrol., 17, 335–346, doi:10.2166/nh.1986.0025, 1986.

Nasta, P., Boaga, J., Deiana, R., Cassiani, G. and Romano, N.: Comparing ERT- and scaling-based approaches to parameterize soil hydraulic properties for spatially distributed model applications, Adv. Water Resour., 126, 155–167, doi:10.1016/j.advwatres.2019.02.014, 2019.

Nippgen, F., McGlynn, B. L. and Emanuel, R. E.: The spatial and temporal evolution of contributing areas, Water Resour. Res., 51, 4550–4573, doi:10.1002/2014WR016719, 2015.

Ogden, F. L. and Watts, B. a.: Saturated area formation on nonconvergent hillslope topography with shallow soils: A numerical investigation, Water Resour. Res., 36(7), 1795, doi:10.1029/2000WR900091, 2000.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., Dandy, G. C. and Fleckenstein, J. H.: Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, Water Resour. Res., 49, 5501–5519, doi:10.1002/wrcr.20405, 2013.

Pfister, L., McDonnell, J. J., Hissler, C. and Hoffmann, L.: Ground-based thermal imagery as a simple, practical tool for mapping saturated area connectivity and dynamics, Hydrol. Process., 24, 3123–3132, doi:10.1002/hyp.7840, 2010.

Van der Ploeg, M. J., Appels, W. M., Cirkel, D. G., Oosterwoud, M. R., Witte, J.-P. M. and Van der Zee, S. E. A. T. M.: Microtopography as a Driving Mechanism for Ecohydrological Processes in Shallow Groundwater Systems, Vadose Zo. J., 11(3), doi:10.2136/vzj2011.0098, 2012.

Qu, Y. and Duffy, C. J.: A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43(8), 1–18, doi:10.1029/2006WR005752, 2007.

Reaney, S. M., Bracken, L. J. and Kirkby, M. J.: The importance of surface controls on overland flow connectivity in semi-arid environments: results from a numerical experimental approach, Hydrol. Process., 28, 2116–2128, doi:10.1002/hyp.9769, 2014.

Scaini, A., Audebert, M., Hissler, C., Fenicia, F., Gourdol, L., Pfister, L. and Beven, K. J.: Velocity and celerity dynamics at plot scale inferred from artificial tracing experiments and time-lapse ERT, J. Hydrol., 546, 28–43, doi:10.1016/j.jhydrol.2016.12.035, 2017.

Schilling, O. S., Gerber, C., Partington, D. J., Purtschert, R., Brennwald, M. S., Kipfer, R., Hunkeler, D. and Brunner, P.: Advancing Physically-Based Flow Simulations of Alluvial Systems Through Atmospheric Noble Gases and the Novel 37Ar Tracer Method, Water Resour. Res., 53, 10,465-10,490, doi:10.1002/2017WR020754, 2017.

Sebben, M. L., Werner, A. D., Liggett, J. E., Partington, D. and Simmons, C. T.: On the testing of fully integrated surface – subsurface hydrological models, Hydrol. Process., 27, 1276–1285, doi:10.1002/hyp.9630, 2013.

Silasari, R., Parajka, J., Ressl, C., Strauss, P. and Blöschl, G.: Potential of time-lapse photography for identifying saturation area dynamics on agricultural hillslopes, Hydrol. Process., 1–18, doi:10.1002/hyp.11272, 2017.

Tanaka, T., Yasuhara, M., Sakai, H. and Marui, A.: The Hachioji experimental basin study -- storm runoff processes and the mechanism of its generation, J. Hydrol., 102, 139–164, 1988.

Tang, Q., Schilling, O. S., Kurtz, W., Brunner, P., Vereecken, H. and Hendricks Franssen, H.-J.: Simulating Flood-Induced Riverbed Transience Using Unmanned Aerial Vehicles , Physically Based Hydrological Modeling, and the Ensemble Kalman Filter, Water Resour. Res., (54), 9342–9363, doi:10.1029/2018WR023067, 2018.

Tetzlaff, D., Soulsby, C., Waldron, S., Malcolm, I. A., Bacon, P. J., Dunn, S. M., Lilly, A. and Youngson, A. F.: Conceptualization of runoff processes using a geographical information system and tracers in a nested mesoscale catchment, Hydrol. Process., 21(10), 1289–1307, doi:10.1002/hyp.6309, 2007.

Weill, S., Altissimo, M., Cassiani, G., Deiana, R., Marani, M. and Putti, M.: Saturated area dynamics and streamflow generation from coupled surface-subsurface simulations and field observations, Adv. Water Resour., 59, 196–208, doi:10.1016/j.advwatres.2013.06.007, 2013.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Intra-catchment variability of surface saturation – insights from long-term ~~observations and simulations~~

Barbara Glaser[1,2], Marta Antonelli[3,1], Luisa Hopp[2], Julian Klaus[1]

[1]Catchment and Eco-Hydrology Research Group, Luxembourg Institute of Science and Technology, Esch/Alzette, 4362, Luxembourg
[2]Department of Hydrology, University of Bayreuth, Bayreuth, 95447, Germany
[3]Hydrology and Quantitative Water Management Group, Wageningen University & Research, Wageningen, 6700, The Netherlands

Correspondence to: Barbara Glaser (barbara.glaser@list.lu)

Abstract

The inundation of flood-prone areas varies in space and time and can have crucial impacts on runoff generation and water quality when the surface saturated areas become connected to the stream. In this study, we aimed to investigate and explain the variability of surface saturation patterns and dynamics within a forested headwater catchment. On the one hand, we mapped surface saturation in seven distinct riparian areas of the Weierbach catchment (Luxembourg) with thermal infrared images, taken weekly to bi-weekly over a period of two years. On the other hand, we simulated the surface saturation generation in the catchment with the integrated surface subsurface hydrologic model HydroGeoSphere over the same period. Both the observations and simulations showed that the saturation dynamics were similar across the catchment, but that small differences between the dynamics at different areas occurred. Moreover, the model reproduced the observed saturation patterns well for all seasonal and hydrologic conditions and at all investigated locations. Based on the observations and simulation results and the matches and mismatches between them, we concluded that the generation of surface saturation in the Weierbach catchment was largely controlled by exfiltration of groundwater into local depressions. However, we also ~~illustrate~~ that the entire variability of the patterns, dynamics and frequencies of surface saturation within the different riparian areas of the catchment can ~~only result from additional controlling factors to microtopography and groundwater exfiltration,~~ such as differing hysteretic behaviour, differing subsurface structures, or additional water sources.

1 Introduction

It is critical for flood risk assessment to understand where and when water is standing or flowing on the ground surface outside of perennial surface water bodies. When such surface saturated areas connect to the stream via overland flow, they also become crucial for runoff generation and water quality. In general, surface saturated areas arise from 1) water ponding on the surface due to exceedance of the infiltration capacity of unsaturated soil, 2) water ponding on impermeable surfaces or saturated soil, 3) water exfiltrating from the subsurface or, 4) stream water extending into the floodplain (e.g. Megahan and King, 1985). Over the past years and decades, various field studies mapped and analysed the spatial and temporal occurrence of surface saturation within different landscapes (e.g. Ambroise, 1986, 2016; Dunne et al., 1975; Gburek and Sharpley, 1998; Latron and Gallart, 2007; Silasari et al., 2017; Tanaka et al., 1988). From the field studies it is well recognized that surface saturation varies in space and time and that its appearance is affected by structural (e.g. topography) and dynamic factors (e.g. precipitation intensity, antecedent moisture). Yet there is limited understanding on how surface saturation evolves spatially and temporally between and within landscapes and how the interplay of different controlling factors and generation processes controls the spatio-temporal variability of surface saturation.

Summary of comments: glaser-et-al_2019_hess-2019-203_commented.pdf

Page:1

Number: 1 Author: reviewer Subject: Replace Date: 2019-07-12 12:37:17

Thermal Infrared Image observations and physically-based modelling.

Number: 2 Author: reviewer Subject: Note Date: 2019-07-12 12:37:37

I would suggest to include TIR in the title

Number: 3 Author: reviewer Subject: Note Date: 2019-07-12 12:29:22

This sentence is not very speciffic. Can you better describe what the saturation dynamics was like and how it was different in certain areas? Where were the areas where the largest differences occured?

Number: 4 Author: reviewer Subject: Replace Date: 2019-07-12 12:31:04

conclude?

Number: 5 Author: reviewer Subject: Replace Date: 2019-07-12 12:34:26

not only be explained by microtopography and groundwater exfiltration, but also by ...

Number: 6 Author: reviewer Subject: Note Date: 2019-07-12 13:13:41

The introduction is well written and concise. I have mainly two suggestions
1) I would encourage the authors to better justify why surface runoff can be an important component of runoff generation (as in most forested catchments around the world it is not).
2) I would encourage the authors to be more critical about the power of models as they are simplifications of the real world and models can only be evaluated based on field data.

Number: 7 Author: reviewer Subject: Note Date: 2019-07-11 12:07:34

You mention two reasons why surface saturation is important. However in a lot of forested catchments subsurface flow is dominating runoff generation and surface runoff is a minor process. I encourage you to write two to three additional sentences that describe the relevance of surface saturation and surface runoff to justify why investigating surface saturation is really important.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Spatially distributed and dynamic hydrological models are potential tools for analysing the generation and development of surface saturation in space and time. Such models allow a detailed investigation of surface saturation at any desired location and time that goes far beyond the information that can be gained by any field observation. Several simulation studies systematically assessed the influence of static and dynamic factors on the temporal evolution, connectivity, and spatial distribution of surface saturation by performing virtual experiments with hillslope models (Ogden and Watts, 2000; Reaney et al., 2014) or by testing a range of terrain indices for predicting time-integrated saturation patterns (Güntner et al., 2004). Other studies relied on dynamic distributed and semi-distributed simulations for analysing connectivity of surface saturation in relation to wetness conditions and catchment runoff (Mengistu and Spence, 2016; Qu and Duffy, 2007; Weill et al., 2013). Weill et al. (2013) and Partington et al. (2013) analysed the processes and water sources that generate surface saturation in a wetland and a pre-alpine grassland headwater, respectively. Both studies applied a model belonging to the group of integrated surface-subsurface hydrologic models (ISSHMs, Sebben et al., 2013), which can simulate the interplay of different surface and subsurface processes of surface saturation generation (e.g. ponding of precipitation from the surface, exfiltration from the subsurface). Modelling studies that focus on a comprehensive spatio-temporal analysis of surface saturation dynamics within a landscape by evaluating the spatially distributed model outputs rather than aggregating the outputs are scarce (e.g. Nippgen et al. (2015) for subsurface saturated areas)

When complementing field observations with simulations to analyse the generation and development of surface saturation in space and time, it is important to ensure that the model yields realistic results. Glaser et al. (2016) demonstrated for a small riparian area that a good match between modelled and observed discharge or soil moisture does not automatically imply a realistic simulation of saturation patterns. They concluded that a spatial validation of the dynamic saturation patterns itself is crucial. However, only few of the existing modelling studies explicitly checked the realism of their simulated surface saturation with field observations before using them for further analyses. These studies focussed either on temporally integrated spatial patterns (Grabs et al., 2009; Güntner et al., 2004) or on temporal dynamics of overall catchment saturation (Birkel et al., 2010; Mengistu and Spence, 2016), but barely any study combined the observation and simulation of both surface saturation patterns and dynamics (Ali et al., 2014; Glaser et al., 2016). The lack of such studies is certainly explainable by the resources that are necessary for obtaining appropriate field data. Today, we still lack a standard method to map surface saturation and the different existing methods such as the ‘squishy boot’ method, the usage of ‘on-off” surface saturation sensors, the mapping of soil morphology or vegetation as surrogates, or the usage of remote sensing techniques (e.g. Dunne et al., 1975; Gburek and Sharpley, 1998; Güntner et al., 2004; Latron and Gallart, 2007; Mengistu and Spence, 2016; Silasari et al., 2017) all have their own advantages and disadvantages.

A relatively new and powerful method for mapping surface saturation is thermal infrared (TIR) imagery. TIR mapping relies on the difference between the surface temperature of water and other materials for identifying surface saturation. Previous work showed that recurrent mapping of surface saturation with high spatial resolution is possible with TIR imagery (Glaser et al., 2016; Pfister et al., 2010). Glaser et al. (2018) and Antonelli et al. (2019) applied TIR imagery mapping in the 42 ha forested Weierbach catchment in western Luxembourg and monitored the dynamics of surface saturation within several distinct riparian areas along the Weierbach stream with a weekly to biweekly mapping frequency over several seasons.

In this study, we explore the intra-catchment variability of temporal and spatial characteristics of surface saturation (dynamics, frequencies, patterns) with a combination of field observation and modelling. We perform the study in the Weierbach catchment, where we can rely on existing TIR imagery data (Antonelli et al., 2019; cf. Glaser et al., 2018) and on previous modelling work for a 6 ha headwater of the catchment (Glaser et al., 2016, 2019) with the ISSHM HydroGeoSphere. Glaser et al. (2016, 2019) simulated the 6 ha area of the catchment by accounting for a layering of the subsurface, while spatial heterogeneity was only represented by microtopography and by a different sequence of subsurface layers in the riparian zone compared to the hillslopes and plateau. Here, we extend the model setup to the entire 42 ha catchment without introducing additional heterogeneity and without performing a re-calibration. We simulate surface saturation in the catchment and contrast

Number: 1 Author: reviewer Subject: Note Date: 2019-07-11 12:08:29

I party agree, but models are only simplified representations of the real world and can only be as good as the observations used to calibrate them. I would rephrase this sentence a bit and be more critical.

Or in other words: If we do not have data at hand, we cannot say if models are performing well or not!

© Author(s) 2019. CC BY 4.0 License.

the results with the observed saturation patterns from the TIR imagery, focussing on the long-term saturation dynamics over different seasons and wetness conditions (25 months with weekly to biweekly mapping resolution) and the spatial patterns of surface saturation occurrence and frequency at seven different riparian areas across the catchment. The two research objectives that we aim to address with this approach are:

1) How variable are surface saturation dynamics and patterns within a catchment and to what extent can we reproduce the variability of the saturation characteristics (dynamics, frequencies, patterns) with a rather homogenously set-up ISSHM?
2) What do we learn about the reasons for the intra-catchment variability of surface saturation characteristics from the matches and mismatches between simulation results and observations?

2 Study site and data

2.1 Physiography, climate and hydrometry

The Weierbach catchment is an intensively studied headwater catchment (42 ha) in western Luxembourg. About half of the catchment area is characterized by gentle slopes <5°, forming a plateau landscape unit (Martínez-Carreras et al., 2016). The rest of the catchment is characterized by hillslopes with slopes > 5°, forming a central V-shaped stream valley from north to south and a V-shaped tributary valley in the east. A third, few metres long stream branch is situated in the west of the central stream valley. Riparian zones along the stream account for 1.2 % of the catchment area (Antonelli et al., 2019). Large parts of the catchment are forested with deciduous trees (mainly European beech and Sessile oaks), the south-east and some other small parts are forested with conifers (mainly Norway spruce and Douglas spruce). The riparian zones are free of tree canopy and covered with ferns, moss, and herbaceous plants. Soil developed from Pleistocene Periglacial Slope Deposits as shallow and highly-permeable siltic, sceletic Cambisol with a depth ranging between 0.4 and 0.9 m (Gourdol et al., 2018; Juilleret et al., 2011; Moragues-Quiroga et al., 2017). Beneath the solum, a 0.5 – 1 m thick basal layer with bedrock clasts oriented parallel to the slope overlies fractured Devonian slate and phyllites (Gourdol et al., 2018; Juilleret et al., 2011; Moragues-Quiroga et al., 2017; Scaini et al., 2017). In the riparian zones, soil and basal layer have been eroded and the fractured bedrock is overlain by shallow organic Leptosols (Glaser et al., 2016).

The climate is oceanic-continental without apparent seasonality in precipitation and with negligible amounts of snow (Carrer et al., 2019). Mean annual precipitation during the period from October 2013 to September 2017 was 955±53 mm. Mean annual discharge was 546±253 mm, with exceptionally dry conditions in the hydrological year 2017. During wet periods, discharge is characterized by double peak hydrographs with first peaks appearing as immediate response to precipitation and second pronounced peaks appearing 48h to 72h later (cf. Martínez-Carreras et al., 2016). During dry periods, only first hydrograph peaks occur and the stream dries out intermittently starting from the source areas downstream.

Hydrological and meteorological data that were used in this study were measured from October 2013 to January 2018. Data from the period from October 2013 to September 2015 were used for spin-up simulations, data from the period from October 2015 to January 2018 were used to drive and validate the actual simulation (cf. Section 3). Discharge was measured with water pressure transducers (ISCO 4120 Flow Logger, 15 min logging intervals) at four v-notches, installed at the outlet of the catchment (SW1, Fig.1) and upstream of the confluences of the three branches (SW2-SW4). Groundwater levels were continuously recorded every 15 minutes with pressure sensors (OTT CTD) in five piezometers installed in different landscape units (riparian zone, hillslope, plateau) of the catchment (Fig. 1, GW1-3, GW5, GW7). Soil moisture was continuously monitored (30 min logging intervals) with water content reflectometers (CS650, Campbell Scientific) installed horizontally in 10, 20, 40 and 60 cm depth at four different sites (Fig.1, SM3-5, SM7). At each site, two depth profiles were monitored. In addition, soil moisture in 10 cm depth was monitored with water content reflectometers (CS616, Campbell Scientific, 30 min logging intervals) at five locations crosscutting the riparian zone of the stream source area of the middle stream branch (Fig. 1, SSM1-5).

Number: 1 Author: reviewer Subject: Note Date: 2019-07-11 12:28:18

SSM seem to have different labels in Fig.1

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Cumulative precipitation was recorded every 5 minutes with a tipping bucket raingauge (Young 52203, unheated, 1 m height) at an open area within the catchment (data gaps were filled by estimating a linear regression to data from a station approximately 4.5 km southward). Potential reference evapotranspiration was estimated based on measured air temperature, relative humidity, wind speed, and net radiation according to the FAO Penman-Monteith formulation (Allen et al., 1998). Air temperature and relative humidity data were recorded next to the soil moisture profile SM5 (Fig. 1, HMP45C-LC, Campbell Scientific, 15 min logging intervals, 2 m height). Wind speed and radiation data were recorded approximately 4.5 km southward of the study site. Wind speed (Young Wind Monitor 05103, Vector A100R Anemometer) was recorded every 15 minutes in 3 m height and converted to wind speed in 2 m height (data gaps closed with data from a station approximately 11.5 km north-eastward) following the FAO guidelines (Allen et al., 1998). Net radiation was recorded every 15 minutes (Kipp & Zonen NR Lite net radiometer) until May 2017. From June 2017 on (and for closing other data gaps), we used net radiation data recorded every 5 minutes close to Luxembourg Airport (~40 km southeast of the study site), as these measurements were highly correlated (linear regression with an intercept of 7.6 W m^{-2}, a slope of 0.92, R^2 = 0.81) with the measurements close to the study site in the years before.

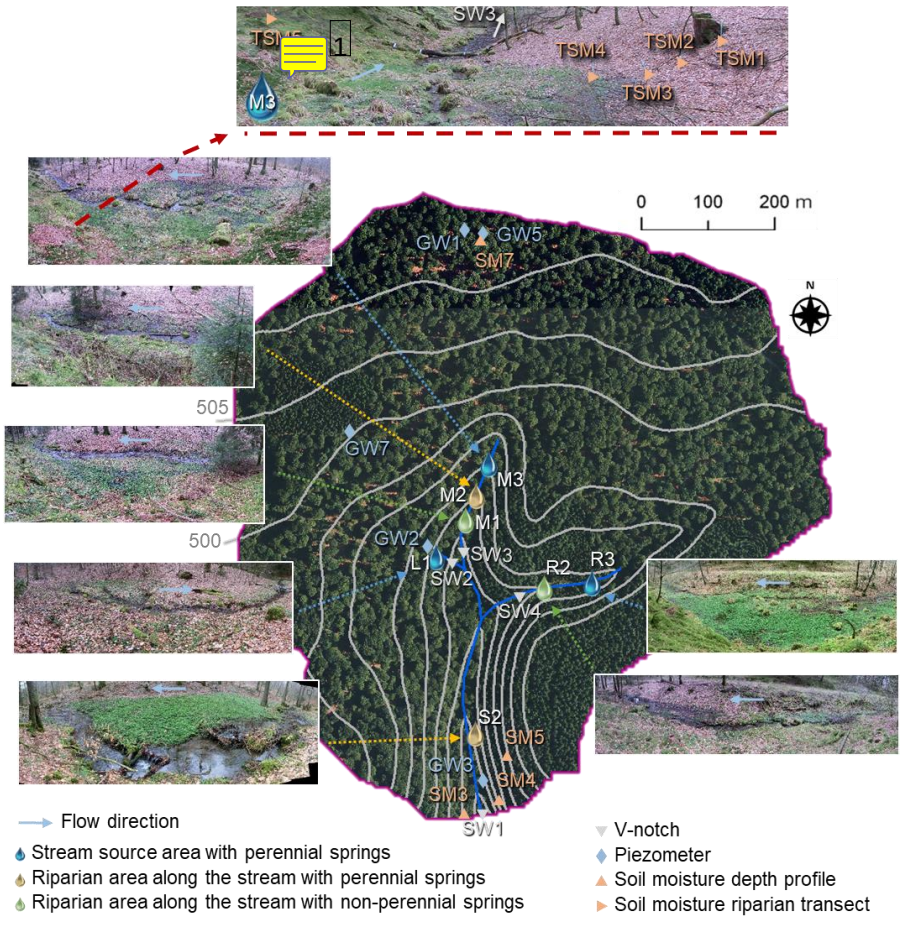

Figure 1: The Weierbach catchment with the locations of the installed v-notches, piezometers, soil moisture sensors and the seven investigated riparian areas.

Number: 1 Author: reviewer Subject: Note Date: 2019-07-11 12:37:24

TSM is not explained in the text or figure caption. I assume you call them SM in the text but please use a consistent abbreviations
For clarity you could include all abbreviations (e.g. in () in the figure caption or the legend

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

2.2 Surface saturation

Here, we define surfaces as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This involves water bodies such as lakes and streams, but excludes mere saturation in the topsoil. According to this definition, surface saturation in the Weierbach catchment generally only occurs in the streambed and the adjacent riparian zones. Other areas that were occasionally observed to be surface saturated during very wet conditions or 'rain on snow' events are forest roads and the prolongation of the streambed above the source regions into the hillslopes. We focus in this study on seven distinct riparian areas in the catchment, which can be classified into three different categories (cf. Antonelli et al., 2019): i) stream source areas with perennial springs (L1, M3, R3, blue areas Fig. 1), ii) areas along the stream with perennial springs (M2, S2, yellow areas Fig. 1), and iii) areas along the stream with non-perennial springs (M1, R2, green areas Fig.1).

We mapped the surface saturation in these seven riparian areas weekly to biweekly from November 2015 to December 2017 with thermal infrared imagery (TIR). Details on the identification of surface saturation with TIR imagery and on the collected surface saturation dataset are presented and discussed in Glaser et al. (2018) and Antonelli et al. (2019). In brief, we created panoramic TIR images of the distinct areas and identified the locations of surface saturation (including the stream) within the images. To do this, each pixel in an image was assigned to be saturated or unsaturated based on the temperature range of locations that were obvious to be saturated from field observations and visual images. In case the contrast between water temperature and temperature of surrounding materials was not sufficient for a reliable pixel classification, the images were excluded from the analysis. In case the pixel classification was affected by a poor temperature contrast or by pixels representing vegetation or snow cover in the images, the images were analysed but flagged as less reliable. Altogether, we obtained 291 binary panoramic images showing the temporal dynamics of surface saturation patterns in the seven studied riparian areas with total numbers of images per site ranging between 34 (L1) and 48 (M2).

Time series of saturation were created for each area by accounting for the percentage of saturated pixels within the individual panoramic images. We normalized the saturation percentages to the maximum observed percentage of saturation in the distinct areas in order to allow a comparison of the saturation dynamics between the different riparian areas. For picturing the spatial surface saturation dynamics within a distinct riparian area, we created maps of saturation frequency. We counted for each area how often the individual pixels of the panoramic TIR images were classified as saturated and normalized the resulting frequency numbers by the total number of TIR images analysed for that area.

The resulting maps of normalized saturation frequency rarely showed pixels that were always saturated (i.e. reaching a normalized frequency of 1). In reality, surface saturation was more persistent than indicated by the frequency maps. The reason for this artefact is that the perspective of the individual TIR panoramas was not 100% identical for all mapping instances and that vegetation sometimes covered parts of the saturated surface, especially during near dry conditions. We co-registered the individual panoramas against a reference panorama for each area, but slight position shifts were inevitable. As a result, the images that were placed on top of each other did not always overlap exactly and the generated saturation frequency maps are blurred. Nonetheless, the maps of normalized saturation frequencies are very useful to understand at a glance where surface saturation occurs more and less frequent within an area and to be used for model validation.

3 Catchment model

3.1 Model setup and parameterisation

We simulated the spatio-temporal dynamics of surface saturation across the Weierbach catchment with HydroGeoSphere (HGS, Aquanty Inc.). HGS is an integrated surface subsurface hydrological model and allows simultaneous simulation of transient surface and subsurface flow. Subsurface flow is simulated based on the 3D Richards equation. Surface flow is simulated based on the diffusive-wave approximation of the 2D Saint Venant equation. Evapotranspiration is simulated with a comparatively simple approach, following the mechanistic concept of Kristensen and Jensen (1975). The equations are

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 13:04:23

Most of this section 2.2 is clear. I would encourage the authors to add another two to three sentences on the how the TIR pictures were taken as it is not fully clear without reading other papers by the authors.

The authors should also include information on how (and on how may data) they determined the temperature range that defined a pixel to be classified :"satureted" or "un-saturated".

I would also suggest to better explain the abbreviations, check if they are used consistently in the text and the figures and include the abbreviations in the figure legend or caption.

Number: 2 Author: reviewer Subject: Replace Date: 2019-07-11 12:35:45

ponds?

Number: 3 Author: reviewer Subject: Note Date: 2019-07-12 11:47:29

Can you say how much of the total lenght of the stream network was monitored with these 7 sections?

Number: 4 Author: reviewer Subject: Note Date: 2019-07-11 12:39:30

What is L,M and R standing for. It is difficult to remember, when not knowing what the abbreviations stand for.

Number: 5 Author: reviewer Subject: Replace Date: 2019-07-11 12:40:34

icons?

Number: 6 Author: reviewer Subject: Replace Date: 2019-07-11 12:40:48

icons?

Number: 7 Author: reviewer Subject: Replace Date: 2019-07-11 12:40:41

icons?

Number: 8 Author: reviewer Subject: Note Date: 2019-07-11 12:53:05

Were the TIR pictures or panoramas taken from the same fixed position each time?
You refer to an earlier publication - and that's o.k. but please include one or two sentences how these TIR pictures were taken and how you can ensure, that you always investigated the same areal extent.

Number: 9 Author: reviewer Subject: Note Date: 2019-07-11 12:57:40

Can you say how much m2 aech pixel was representing?

Number: 10 Author: reviewer Subject: Note Date: 2019-07-13 17:34:25

Did you determine this temperature range in an earlier study? If so, please cite! On how much data did you determine this range? Was the same temperature range applied at each location and all sampling campaings?

How subjective was this temperature range?

Number: 11 Author: reviewer Subject: Note Date: 2019-07-11 12:55:47

It is not clear to me if the pictures were taken continuously with one IR-camera permanently installed at a fixed position at each location or during snapshot campains. Earlier you indicate a weekly to bi-weekly interval. I would then rather call the data snapshots.

Number: 12 Author: reviewer Subject: Note Date: 2019-07-11 13:00:13

On what information this statement is based? Did you also analyze the photos taken at the same time or map the saturated areas?

Number: 13 Author: reviewer Subject: Replace Date: 2019-07-11 12:59:07

frequent?

Number: 14 Author: reviewer Subject: Strikeout Date: 2019-07-11 13:00:37

Number: 15 Author: reviewer Subject: Note Date: 2019-07-11 13:02:35

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

2.2 Surface saturation

Here, we define surfaces as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This involves water bodies such as ~~lakes~~ and streams, but excludes mere saturation in the topsoil. According to this definition, surface saturation in the Weierbach catchment generally only occurs in the streambed and the adjacent riparian zones. Other areas that were occasionally observed to be surface saturated during very wet conditions or 'rain on snow' events are forest roads and the prolongation of the streambed above the source regions into the hillslopes. We focus in this study on seven distinct riparian areas in the catchment, which can be classified into three different categories (cf. Antonelli et al., 2019): i) stream source areas with perennial springs (L1, M2, R3, blue ~~areas~~ Fig. 1), ii) areas along the stream with perennial springs (M2, S2, yellow ~~areas~~ Fig. 1), and iii) areas along the stream with non-perennial springs (M1, R2, green ~~areas~~ Fig.1).

We mapped the surface saturation in these seven riparian areas weekly to biweekly from November 2015 to December 2017 with thermal infrared imagery (TIR). Details on the identification of surface saturation with TIR imagery and on the collected surface saturation dataset are presented and discussed in Glaser et al. (2018) and Antonelli et al. (2019). In brief, we created panoramic TIR images of the distinct areas and identified the locations of surface saturation (including the stream) within the images. To do this, each pixel in an image was assigned to be saturated or unsaturated based on the temperature range of locations that were obvious to be saturated from field observations and visual images. In case the contrast between water temperature and temperature of surrounding materials was not sufficient for a reliable pixel classification, the images were excluded from the analysis. In case the pixel classification was affected by a poor temperature contrast or by pixels representing vegetation or snow cover in the images, the images were analysed but flagged as less reliable. Altogether, we obtained 291 binary panoramic images showing the temporal dynamics of surface saturation patterns in the seven studied riparian areas with total numbers of images per site ranging between 34 (L1) and 48 (M2).

Time series of saturation were created for each area by accounting for the percentage of saturated pixels within the individual panoramic images. We normalized the saturation percentages to the maximum observed percentage of saturation in the distinct areas in order to allow a comparison of the saturation dynamics between the different riparian areas. For picturing the spatial surface saturation dynamics within a distinct riparian area, we created maps of saturation frequency. We counted for each area how often the individual pixels of the panoramic TIR images were classified as saturated and normalized the resulting frequency numbers by the total number of TIR images analysed for that area.

The resulting maps of normalized saturation frequency rarely showed pixels that were always saturated (i.e. reaching a normalized frequency of 1). In reality, surface saturation was more ~~persistent~~ than indicated by the frequency maps. The reason for this artefact is that the perspective of the individual TIR panoramas was not 100% identical for all mapping instances and that vegetation sometimes covered parts of the saturated surface, especially during ~~near~~ dry conditions. We co-registered the individual panoramas against a reference panorama for each area, but slight position shifts were inevitable. As a result, the images that were placed on top of each other did not always overlap exactly and the generated saturation frequency maps are blurred. Nonetheless, the maps of normalized saturation frequencies are very useful to understand at a glance where surface saturation occurs more and less frequent within an area and to be used for model validation.

3 Catchment model

3.1 Model setup and parameterisation

We simulated the spatio-temporal dynamics of surface saturation across the Weierbach catchment with HydroGeoSphere (HGS, Aquanty Inc.). HGS is an integrated surface subsurface hydrological model and allows simultaneous simulation of transient surface and subsurface flow. Subsurface flow is simulated based on the 3D Richards equation. Surface flow is simulated based on the diffusive-wave approximation of the 2D Saint Venant equation. Evapotranspiration is simulated with a comparatively simple approach, following the mechanistic concept of Kristensen and Jensen (1975). The equations are

I can imagine...
but why didn't you clip the images to the extent that was covered by all images and do the analysis on these pixels?

Number: 16 Author: reviewer Subject: Note Date: 2019-07-12 09:55:01

3.1 is clear. I only suggest to add one or two sentences that the parameters are based on field and lab investigations from 5 sites in teh riparian zone instead of only referring to Glaser et al, 2018.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

linearized implicitly using the Newton-Raphson approach and solved in an unstructured finite element grid. HGS has been used in the past for addressing diverse questions at various temporal and spatial scales (e.g. Ala-aho et al., 2015; Davison et al., 2018; Erler et al., 2019; Frei et al., 2010; Munz et al., 2017; Nasta et al., 2019; Partington et al., 2013; Schilling et al., 2017; Tang et al., 2018). It also has already been applied for a 6 ha headwater region of the Weierbach catchment (Glaser et al., 2016, 2019). In this study, we applied the parameterization of Glaser et al. (2016) to the entire 42 ha catchment without performing an additional parameter calibration.

The catchment was spatially discretized into 42,274 triangular elements, using the mesh generator AlgoMesh (HydroAlgorithmics Pty Ltd). Edge lengths of the mesh elements ranged from > 30 m at the plateau to < 0.4 m for the seven analysed riparian zones and the streambed (Fig. 2). It was crucial to use such a fine mesh resolution in the riparian zone in order to enable a comparable spatial detail as obtained with the TIR imagery for the surface saturation patterns. Vertically, the model grid comprised 5 m, which were divided into 14 layers with element depths ranging from 0.15 m for the top layers to 0.5 m for the bottom layers (Fig. 2).

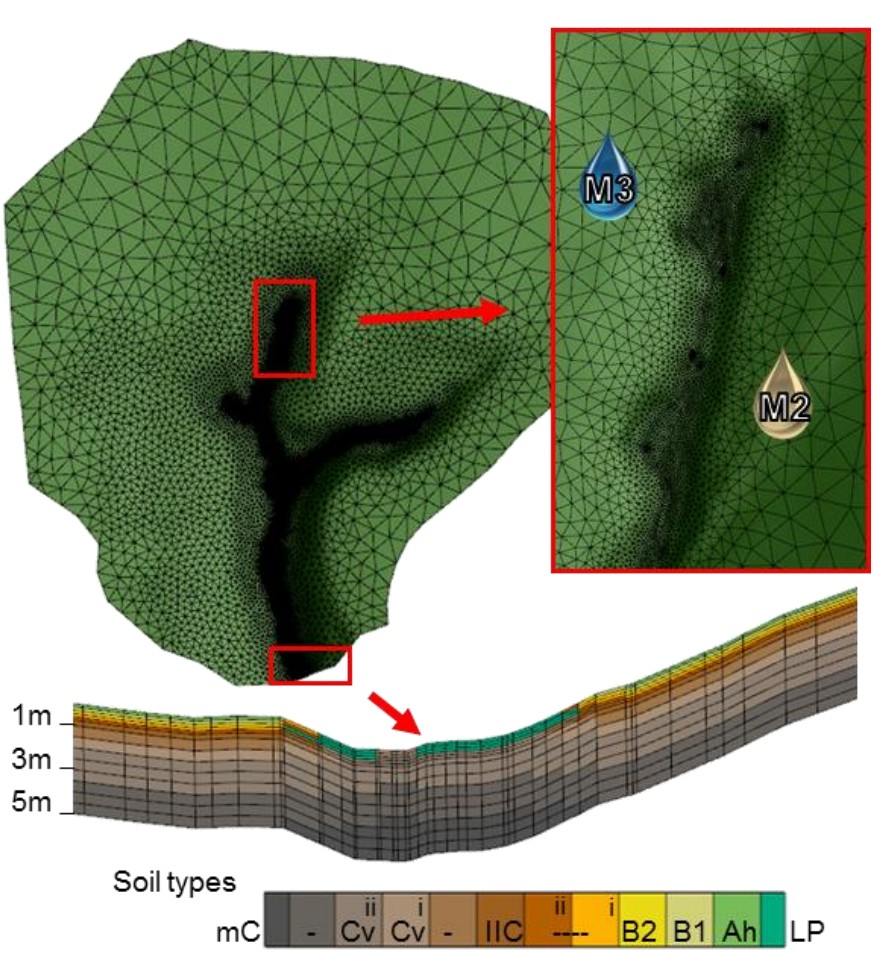

Figure 2: Setup of the model mesh with a zoom on the fine horizontal resolution in the riparian areas and the streambed (inset on the right) and a vertical cross section through the stream valley and adjacent hillslopes (bottom) showing the vertical discretization and assignment of different soil properties (cf. Tab. 1). Ah = topsoil, B1 and B2 =subsoil, IIC = basal layer, Cv = fractured bedrock, Cm = fresh bedrock.

No Comments.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

The subsurface was parameterized homogeneously with 10 different property layers, representing top- and subsoil (Ah, B1, B2), the basal layer (IIC), fractured and fresh bedrock (Cv, mC), and layers of transition between subsoil, basal layer, and fractured bedrock (Fig. 2). We implemented spatial heterogeneity in the stream valleys, where soil and basal layer were eroded and the outcropping fractured bedrock was overlain with organic, stagnic Leptosol in the riparian zones (Fig. 2). We used the Mualem-van Genuchten soil hydraulic functions for describing the saturation-pressure relation. The necessary soil hydraulic parameter values for the different property layers (porosity, residual saturation, van Genuchten α, van Genuchten β, saturated hydraulic conductivity, Tab. 1) were assigned according to Glaser et al. (2016). We only parameterised one additional layer for the fractured bedrock (Cv (ii)) in order to account for the adapted depth of 3.5 m in the catchment model compared to the depth of 3 m in the headwater model.

Table 1: Soil hydraulic parameters of the different soil property zones. Table adapted from Glaser et al. (2016)

Soil property zone	Residual saturation	van Genuchten parameter α [m^{-1}]	van Genuchten parameter β	Porosity	Saturated hydraulic conductivity [m d^{-1}]
Ah	0.12	6.6	1.46	0.74	1.71E+01
B1	0.10	22.1	1.42	0.61	1.71E+01
B2	0.10	22.1	1.42	0.45	4.59E+01
B2-IIC (i)	0.10	22.1	1.42	0.3	9.30E+02
B2-IIC (ii)	0.10	22.1	1.42	0.15	2.04E+03
IIC	0.02	6.0	1.50	0.20	8.40E+02
IIC-Cv	0.02	6.0	1.50	0.15	3.00E+00
Cv (i)	0.02	6.0	1.50	0.10	1.20E-02
Cv (ii)	0.02	6.0	1.50	0.07	1.20E-02
Cv-mC	0.02	6.0	1.50	0.05	9.00E-04
mC	0.02	6.0	1.50	0.01	2.40E-05
LP	0.10	22.1	1.42	0.61	7.80E+00

Surface and subsurface flow were coupled via a Darcy flux exchange through a thin coupling layer (10^{-4} m). We assumed different Manning's surface roughness values for the forested area ($1.24*10^{-6}$ d $m^{-1/3}$), the riparian zone ($9.41*10^{-7}$ d $m^{-1/3}$), and the stream bed ($4.4*10^{-7}$ d $m^{-1/3}$) (cf. Glaser et al., 2016). Evapotranspiration properties (Tab. S1) were assigned individually for the deciduous forest, the coniferous forest in the southeast of the catchment, and the riparian zones including the streambed and values were based on the calibrated values of Glaser et al. (2016). The simulation was driven with daily sums of precipitation and reference evapotranspiration, which were treated as being spatially uniform. The outer edge of the surface domain was assigned as critical depth boundary, allowing water to leave the model domain via surface flow. Side and bottom boundaries of the subsurface domain were no flow boundaries. A spin-up simulation drained the catchment from full saturation to steady state conditions (for 1 mm d^{-1} of precipitation, no evapotranspiration) and subsequently repeated the period from October 2013 to October 2015 three times for obtaining realistic initial conditions. The actual simulation spanned over the period from October 2015 to January 2018, the period where we mapped surface saturation with TIR imagery.

3.2 Assessment of model performance

We benchmarked the model against measured discharge, groundwater level, soil moisture, and surface saturation patterns and dynamics at various locations (Fig. 1). We calculated the Kling Gupta Efficiency (KGE) as a combined measure for correlation, bias, and relative variability (Gupta et al., 2009) between simulated and observed discharge. We also calculated KGEs for the simulated groundwater levels, but particularly evaluated the groundwater level dynamics rather than absolute values based on

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 08:53:42

Do you have actual measurements to parameterize 10 different soil layers? Otherwise this only increases parameter uncertainty.

Number: 2 Author: reviewer Subject: Note Date: 2019-07-12 09:50:29

It is fine to refer to earlier publications but I think that some information that is given in section 2.3 Pedological Data in Glaser et al., 2016 should be mentioned here to strengthen your approach. E.g. mention that the parameterization used is based on soil profile data (5 sites in the riparian zone) , lab experiments (n samples) and ERT measurements (and not taken from the literature.

© Author(s) 2019. CC BY 4.0 License.

Pearson correlation coefficients. Soil moisture was also evaluated based on its dynamics with Pearson correlation coefficients, while absolute values were only compared visually. Since simulated soil moisture was extracted from model nodes whose depths did not exactly correspond with the measurement depths, we interpolated depth-weighted average values from the model output for calculating the correlation with the observations in the respective depths. The interpolated model values of volumetric water content were then correlated with the observations of water content, averaging the measurements of the twin depth profiles at the monitoring sites.

For comparing the simulation output with the surface saturation information obtained with the TIR images, it was necessary to convert the model output into a comparable format via several processing steps: First, we extracted the surface water depths in the surface domain of the model for noon of the days where TIR images were taken and analysed. Next, we transformed the surface water depths into a binary saturation map of the entire catchment by classifying the surface domain cells as saturated if water depths were $>10^{-4}$ m. The depth of 10^{-4} m corresponds to the penetration depth of the used TIR camera for water columns and thus is the minimum depth that could be detected as pure water temperature signal with the camera. Finally, we projected the model output into jpeg images with the same perspective and extent of the TIR panoramic images by turning, bending, and cutting the modelled saturation maps according to each of the seven riparian areas individually. This model output processing allowed us to perform the same calculations for the model output as for the TIR images, i.e. to create time series of normalized saturation and maps of normalized saturation frequencies for the seven riparian areas with comparable perspectives and extents. Since it was not possible to project the model output identically to the perspectives of the TIR images, we compared the saturation dynamics and patterns of the model images with the observations qualitatively (visually) only. A quantitative comparison would have been biased by differences in image distortions and total area ext[illegible] 1

Furthermore, we compared the simulated frequency of surface saturation with the simulated frequency of groundwater reaching the surface. To do this, we marked the surface cells below which the subsurface domain was fully saturated from the bottom to the surface as cells where groundwater reached the surface. This binary information was transformed into a frequency map analogous to the procedure for creating the surface saturation frequency maps, using the same output times.

4 Results

4.1 Simulation of discharge, groundwater level and soil moisture 2

The model reproduced the seasonal dynamics of measured discharge very well (Fig. 3, Fig. S1). The best fit was obtained at the outlet (SW1) with a KGE of 0.74. Discharge at SW2, SW3, and SW4 was reproduced equally well with KGEs of 0.49, 0.48, and 0.47. Groundwater levels were captured well with the model at the locations close to the riparian zone (KGE=0.57, r=0.78 for GW2; KGE=0.64, r=0.84 for GW3). At hillslopes and plateau, simulated groundwater levels were similar to the observed levels during the wet season, but during dry conditions the groundwater levels did not fall deep enough (Fig. 3, Fig. S1). This level discrepancy was reflected in low KGEs (0.30 for GW1, 0.21 for GW5, 0.02 for GW7). However, general dynamics of level increasing and decreasing were also captured at hillslopes and plateau (r = 0.66 for GW5, r = 0.62 for GW7, and r = 0.76 for GW1; note that the value for GW1 only includes data for wet periods, since the piezometer fell dry during summer months).

Simulated soil moisture generally showed a transition from higher to lower responsiveness from topsoil to subsoil layers consistent with the monitored soil moisture (Fig. 3, Fig. S1) and Pearson correlation coefficients indicated overall a good agreement between simulated and observed soil moisture dynamics (Tab. 1). As for the groundwater levels at the hillslopes and plateau, soil moisture observations showed a distinct decrease in water content during dry periods, which the simulation could not reproduce to the same extent. The observed water content in the riparian zone was always close to saturation (TSM4, Fig. 3), while the simulation showed a decrease in water content during dry periods in the riparian zone. Yet the simulation also showed a spatial trend for more permanent soil saturation in the riparian zone (TSM4) and its vicinity (TSM3, Fig. S1)

Page:8

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 10:01:21

O.K. given the difficulty of maching the to souces of information a qualitative comparison is appropriate but can you describe what qualitative chriteria you used to do this? E.g., total number of saturated pixels in the TIR and the model output?

Number: 2 Author: reviewer Subject: Note Date: 2019-07-12 10:33:09

This is more a comment than a request for change of the text:

It is always difficult to directly compare real measurements at one groundwater well or one soil moisture pit with simulations. That's clear. Also your goodnes-of-fit-values and other statistical measure (person corr) are state of the art to compare measured and modeled time series.
However from a visual comparison of your measured and molded gw-levels and soil moisture data strong differences become aparent. My question is, how can a model simulate a detailed process such as surface saturation, if its gw-levels and soil moisture simulations are so far off from measurements?

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

than at the hillslopes and plateau. The simulated values of water content were similar to the observed values at some locations (e.g. TSM2, SM4, Fig. 3) and clearly differed at other locations (e.g. SM7, Fig. 3), but the matches and mismatches of the volumetric water content did not clearly depend on specific areas or landscape units. Moreover, we think that moisture dynamics and responsiveness are more informative for model performance than the absolute water content values, since also the measured values of volumetric water content differed from each other within small distances (e.g. measurements of water content in 10 cm depth at profile SM7, Fig. 3).

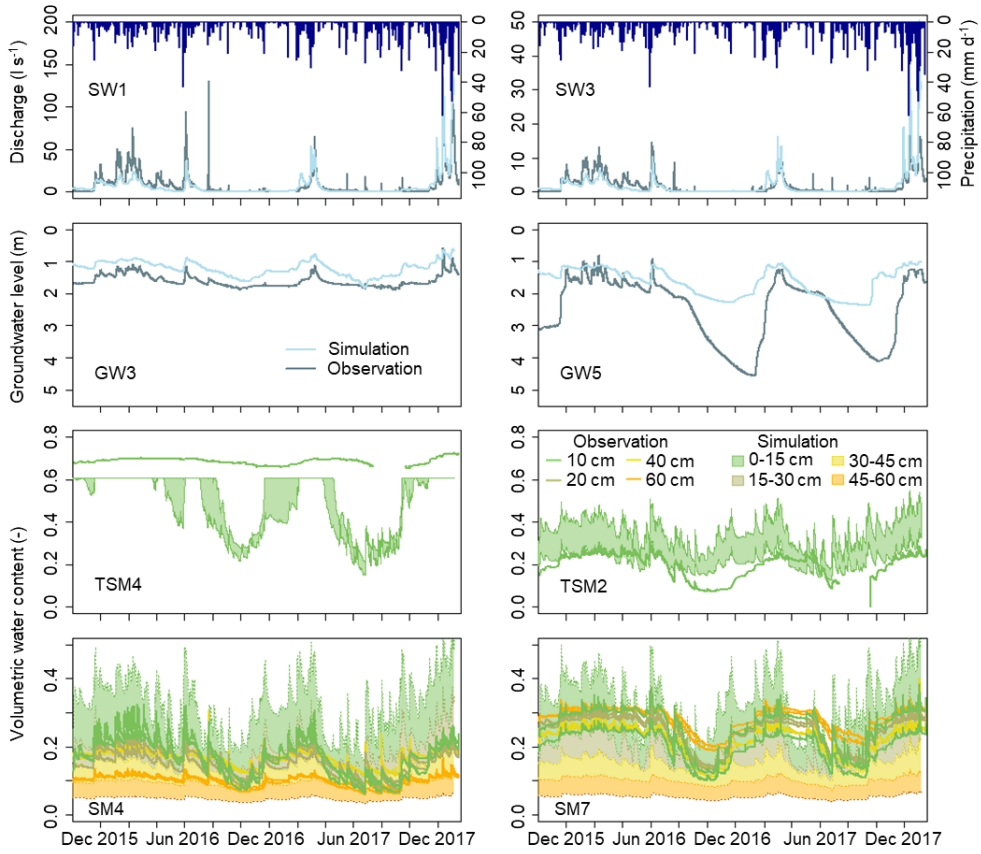

Figure 3: Simulated and observed time series of discharge, groundwater level below the surface, and volumetric water content. Colour bands indicate the possible span of simulated volumetric water contents in the depths between two model nodes. The time series of the observation locations (cf. Figure 1) that are not shown here, are shown in the supplemental material (Figure S1).

Table 2: Coefficients of Pearson correlation between simulated and observed volumetric water content of the soil for the different measurement locations and depths (cf. Fig. 1).

	SM3	SM4	SM5	SM7	TSM1	TSM2	TSM3	TSM4	TSM5
10cm	0.54	0.75	0.70	0.59	0.60	0.62	0.67	0.30	0.85
20cm	0.67	0.82	0.76	0.62					
40cm	0.82	0.89	0.88	0.79					
60cm	0.85	0.92	0.91	0.82					

No Comments.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

4.2. Dynamics of surface saturation

The observed dynamics of normalized surface saturation (Fig. 4, coloured lines) were similar for all seven investigated riparian areas and followed the seasonal trend of the catchment discharge. Yet some differences between the studied areas were discernible. For example, saturation was less persistent between February and April 2016 in the two areas without perennial springs (M1, R2, Fig. 4) than in the other areas. Maximum saturation was reached in December 2017 at M1, R2 and S2, but between February and April 2016 at the other locations (Fig. 4). As for the observations, the simulated dynamics of normalized surface saturation (Fig 4, black lines) followed the general trend of the simulated discharge dynamic. The simulation showed a faster decrease and increase of the normalized saturation during dry periods than it was observed in most areas. However, simulated discharge also seemed to decrease and increase earlier than it was observed (c.f. Section 4.3). The simulated saturation dynamics did not clearly differ between the different locations and thus behaved more synchronous than the observations (e.g. maximum simulated saturation in December 2017 in all areas). As a result, the match between simulated and observed dynamics of normalized saturation was better for some areas (e.g. M1, R2, Fig. 4) than for others (e.g. S2, L1, Fig. 4).

The dynamic changes of normalized simulated saturation matched the normalized observations generally well, despite of under- and over-estimated amounts of minimum and maximum absolute saturation for all areas. The minimum amount of saturated pixels in the TIR panoramas ranged between 0.02 % at M3 and R3 and 3.38 % at S2, while the model did not simulate any surface saturation during the driest period (Fig. 4). In addition, simulated normalized saturation stayed longer close to the minimum than the observed saturation for several areas (L1, S2, M1). These results show that the model simulated a stronger dry-out than observed in the Weierbach. At the same time, the simulation overestimated maximum saturation in the riparian zone (Fig. 4). The overestimation was not equally strong at the seven investigated areas and as a result, the distinction between areas showing higher or lower maximum saturation was not the same for observation and simulation (e.g. R3 showing one of the highest maximum saturation in the observation, but one of the lowest maximum saturation in the simulation compared to the other areas).

4.3 Discharge – surface saturation relationship

The Pearson correlation between normalized saturation and discharge at the outlet SW1 was > 0.65 for both the simulation and the observation in almost all riparian areas. L1 was the only exception with r_{obs} = 0.54 (Fig. 5). The simulated relationships between normalized saturation and discharge resembled the observed relationships in terms of value range and shape (Fig. 5), although the observation data scattered distinctly more than the simulation data. A power law relationship approximated the observed relationship between discharge and saturation for all seven areas, when data that were taken during rainfall or rising discharge were excluded (cf. Antonelli et al., 2019). For some areas, the simulation matched the trend lines of the observation data closely (e.g. L1, M2). For other areas, the visual fit of the model output to the observation data was less good (e.g. S2, R3), but still described a similar trend.

Despite the common shape of a power law function, the saturation – discharge relationships were slightly different between the different areas, both for observation and simulation data. For example, the power law functions fitted to the observations showed that saturation during high flow conditions (> 5 l s^{-1}) increased most strongly with discharge in the sources areas (especially M3 and R3). During low flow conditions (< 1 l s^{-1}), the source areas (L1, M3, R3) showed the lowest amount of normalized saturation and the least change relative to discharge compared to the other areas. In the simulated relationships, the increase in saturation for high discharge (> 5 l s^{-1}) was strongest for M3 and S2. The simulated relationship between discharge and surface saturation during low flow (< 1 l s^{-1}) was similar for all areas in terms of slope, but differed in the amount of normalized saturation, being highest for areas in the east stream branch (R2, R3), followed by the middle upstream branch (M1, M2, M3), and L1 and S2.

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 11:19:01

...but one of the assumptions of the person correlation coefficient is a linear relationship? So what is r that you present in the text and the figure 5?

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

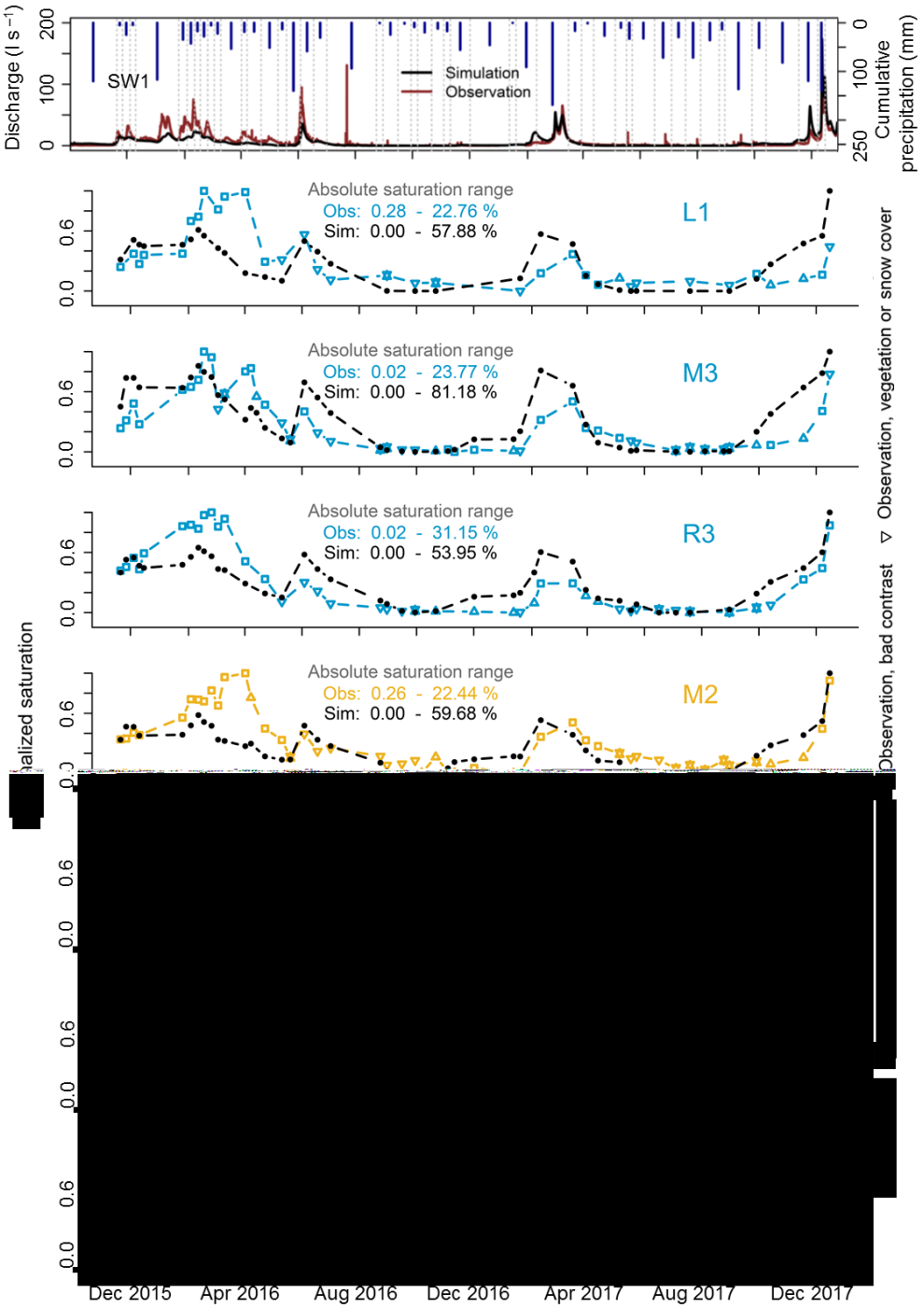

Figure 4: Time series of observed and simulated surface saturation in the seven investigated riparian areas. Surface saturation is normalized to the minimum and maximum amount of saturation that was observed and simulated in the individual areas, respectively. Observations that were derived from TIR images with a poor temperature contrast or with influences of vegetation and snow cover are deemed less reliable. Cumulative precipitation between the measurement dates (grey dashed lines) and discharge at catchment outlet SW1 are shown in the top panel for facilitating the comparison to precipitation and flow conditions.

No Comments.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

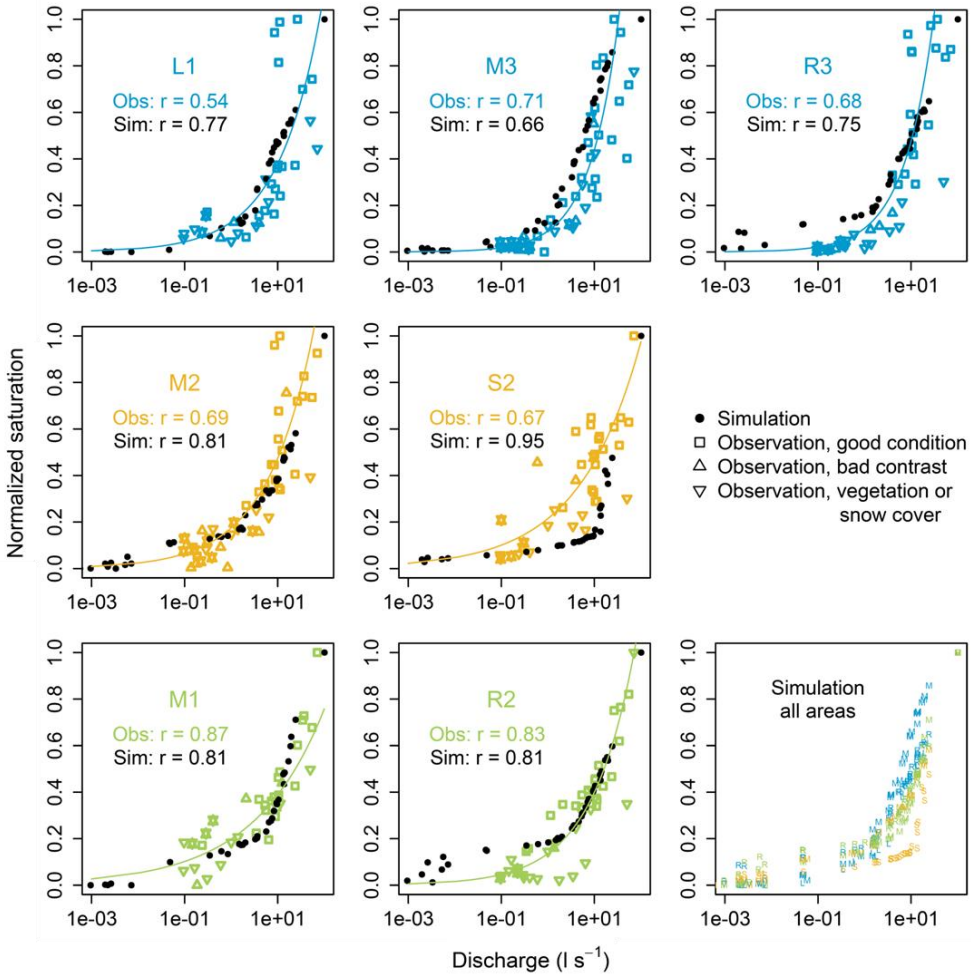

Figure 5: Observed and simulated relationships and Pearson correlations between normalized surface saturation and discharge at the catchment outlet SW1 for the seven investigated riparian areas. Observations that were derived from TIR images with a poor temperature contrast or with influences of vegetation and snow cover are deemed less reliable. Solid lines are power law curves fitted to the observation data, excluding data taken during rainfall or rising discharge. For facilitating the comparison between the seven areas, the panel on the bottom right contains the simulated data points from all seven areas and the area affiliation is indicated with the respective colour and letter.

4.4 Spatial patterns of surface saturation

The realism of simulated patterns of surface saturation was evaluated for each riparian area by visually comparing the surface saturation frequency maps obtained from the simulations and observations (Fig. 6). The model captured the location of the stream and the locations that intermittently became surface saturated well for most of the seven investigated areas. For example, both observation and simulation showed that only the right side of the stream became saturated in M1, that the riparian zone of the right streamside in M2 became saturated only in the upstream part, and that saturation mainly developed on the left streamside in R3, surrounding some permanently dry areas next to the stream (Fig. 6). The only area with a clear mismatch between observed and simulated patterns of surface saturation was area L1, where surface saturation was simulated on the opposite streamside and at a clearly wrong position along the stream (upstream vs downstream).

No Comments.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

The simulated surface saturation also reflected the observed saturation frequencies well. The simulation reproduced the general picture of more frequent surface saturation in the streambed than at the streamsides, but - as for the saturation patterns - simulated and observed frequencies corresponded better in some areas (e.g. S2, Fig. 6) than in others (e.g. R3, Fig. 6). For example, the observed frequency of surface saturation in the streambed was generally lower in the source areas (L1, M3, R3) than in the mid- and downstream areas (M2, S2, M1, R2), while the simulated frequency of surface saturation in the streambed was more similar between the areas and particularly overestimated in L1 and R3.

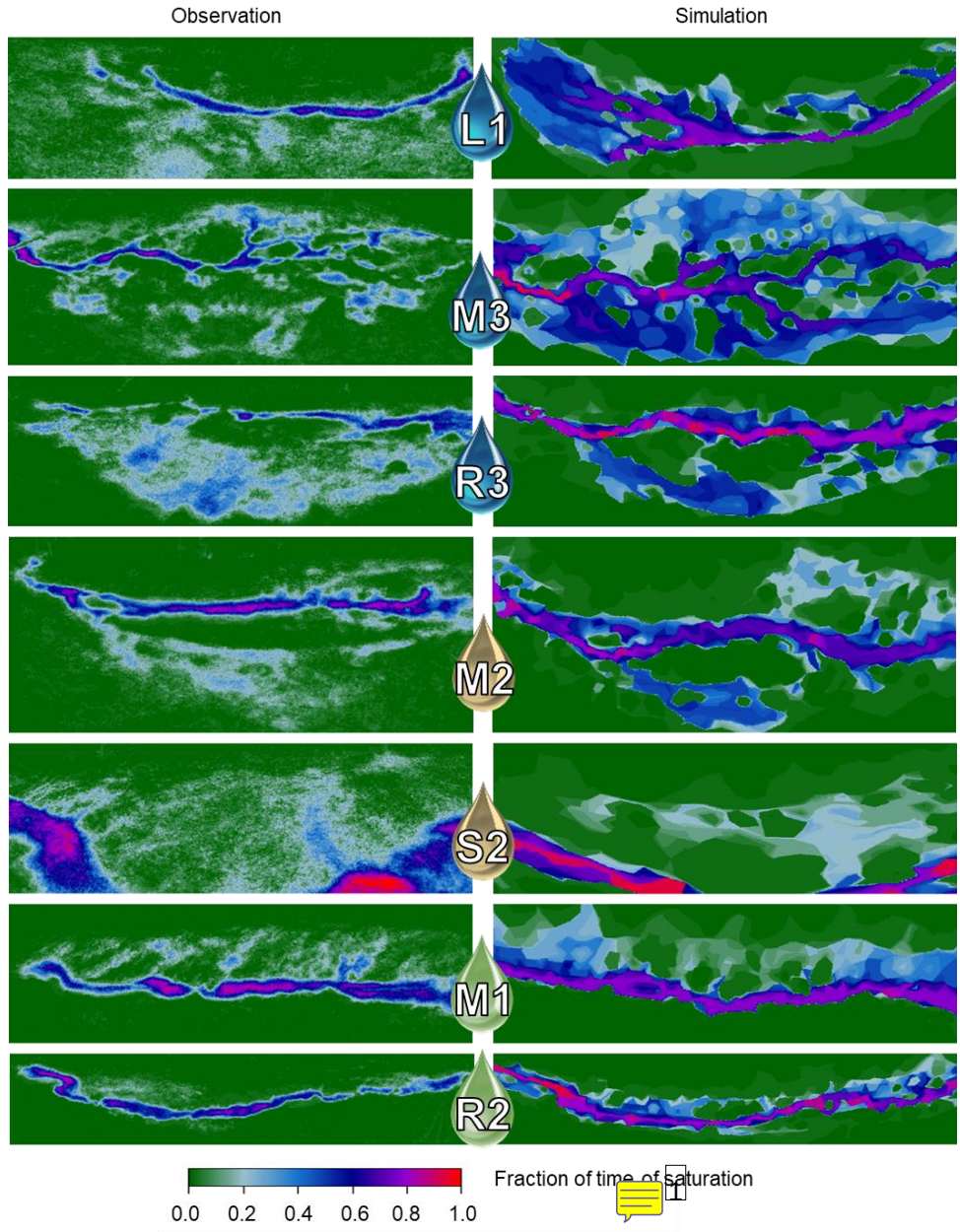

Figure 6: Observed (left) and simulated (right) frequencies of surface saturation in the seven investigated riparian areas. The maps were created by first counting how often the individual pixels were classified as saturated in the individual panoramic images and second normalizing the resulting frequency numbers by the total number of images analysed for the respective area.

Page:13

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 11:28:44

small detail: it is rather fraction of total number of TIR measuring campaigns because you do not have continuous information on TIR

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

4.5 Simulated patterns and dynamics of surface saturation versus groundwater reaching the surface at catchment scale

Simulated surface saturation generally occurred only in the streambed and adjacent riparian zones (Fig. 7a). During the wettest conditions of the study period (winter 2017/2018), surface saturation also occurred as prolongation of the eastern stream branch into the hillslope above the source area R3. This simulated occurrence behaviour of surface saturation across the catchment is in accordance with field evidence, where we observed surface saturation outside of the valley bottom only during very wet conditions or rain on snow events (cf. Section 2.2). The simulated patterns of where and how frequently groundwater reached the ground surface (Fig. 7b) proved to be very similar to the surface saturation frequency map of the catchment (Fig. 7a). The only obvious difference occurred in the area above the source area of the eastern stream branch (R3), with a smaller extent of groundwater reaching the surface than extent of surface saturation.

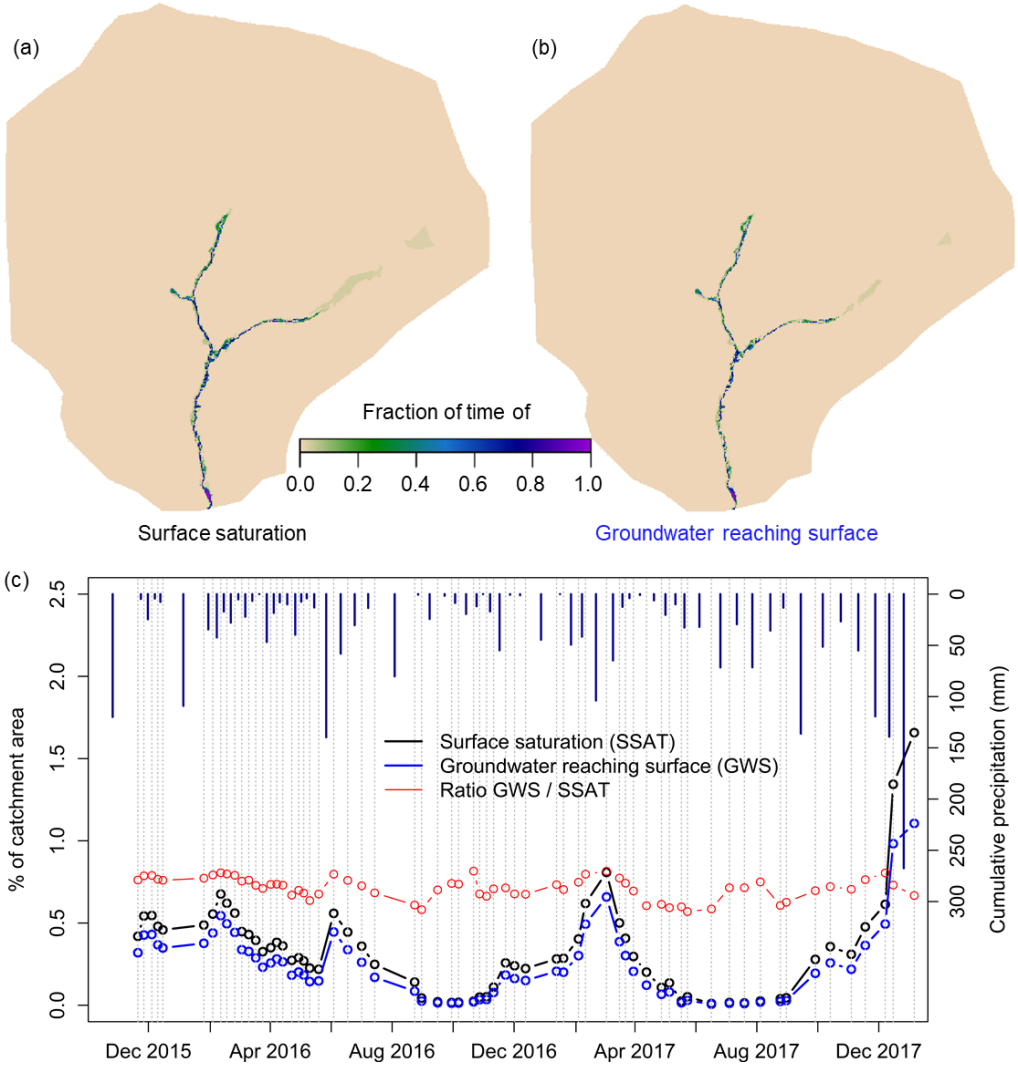

Figure 7: Simulated frequency maps (a, b) and time series of percentage (c) of surface saturation and groundwater reaching the surface in the Weierbach catchment. Precipitation is given as cumulative amounts between the observation dates (grey dashed lines).

Page:14

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 11:38:22

What is the difference between Fig 7 a and 7b?
As I understand; The modeled saturated area is always based on the criteria that gw reaches the soil surface. This is shown in Fig. 7b.

Fig 7b is modeled - O.K. but how is the surface saturation frequency map of the catchment (Fig 7a) generated?

Is Fig 7a based on soil moisture data?

© Author(s) 2019. CC BY 4.0 License.

The time series of simulated percentage of catchment area with surface saturation and groundwater reaching the surface revealed that the area where groundwater reached the surface was always smaller in extent than the surface saturated area, even after dry condition (Fig. 7c). The biggest absolute difference between the areal extent of surface saturation and groundwater reaching the surface was simulated during winter 2017/2018 (1.66 % vs 1.1 % of catchment area), where the conditions were very wet with high discharge and high cumulative precipitation and where the difference in areal extent was also visible in the frequency maps (Fig. 7a and b). However, the ratio between the extent of groundwater reaching the surface and the extent of surface saturation was not exceptionally high during winter 2017/2018. Instead, the ratio scattered without a clear trend between 0.57 and 0.82 during the entire simulation period, apparently independent from the cumulative amount of precipitation or surface saturation.

5 Discussion

The aim of this study was to analyse the spatio-temporal variability of surface saturation within the Weierbach catchment, with a focus on the stream valleys and riparian zones. Even though simulated discharge, groundwater levels and soil moisture showed some discrepancies to observations in terms of absolute values, we would argue that the ~~performance of~~ the different time series at different locations was quite good for a model that was not calibrated and set up rather homogenously across the catchment. While the model had some problems to reproduce soil moisture and groundwater levels during the dry conditions at hillslopes and plateau, the simulated time series matched the observations especially well in the riparian zone and vicinity. This gives us confidence that the model setup was valid for evaluating and analysing the spatio-temporal dynamics of surface saturation and its intra-catchment variability.

5.1 Temporal dynamics of surface saturation

The model reproduced the observed long-term dynamics of surface saturation over different seasons and wetness conditions well. Our study goes beyond previous works that compared the simulation of surface saturation dynamics with observations (e.g. Ali et al., 2014; Birkel et al., 2010; Glaser et al., 2016; Mengistu and Spence, 2016) by relying on a longer study period and a higher number of observations in time. This allowed us to analyse and compare various hydrological conditions and the dynamic transition between them over all seasons with a frequent number of observations. Moreover, we accounted for spatial variability of saturated area dynamics within the catchment. Unlike the various quasi dynamic wetness indices presented in Ali et al. (2014), which could not satisfyingly reproduce the spatio-temporal variability of connected surface saturation observed in a catchment in the Scottish Highlands, our model reproduced the distributed dynamics of surface saturation well, without clear performance differences for different wetness conditions.

Simulations and observations showed both that the temporal dynamics of surface saturation were mostly consistent across the catchment. Moreover, our simulations showed that the spatio-temporal development of surface saturation was very similar to the spatio-temporal dynamics of groundwater reaching the surface (cf. Fig. 7). This suggests that the generation of surface saturation in the Weierbach catchment is largely driven by the synchronous exfiltration of groundwater in topographic depressions. Antonelli et al. (2019) drew consistent conclusions based on a statistical analysis of the observation data.

5.2 Relation between surface saturation and discharge

We found that the observed and simulated relationships between surface saturation and discharge resembled power law relationships (cf. Fig. 5). This is consistent with earlier studies that showed power law relationships between contiguous connected surface saturated areas and discharge (Mengistu and Spence, 2016; Weill et al., 2013). In contrast to these studies, we did not observe hysteretic loops in the relationship between saturation and high streamflow. Nonetheless, the scatter in the observed discharge – surface saturation relationships might indicate that the development of surface saturation in the

Page:15

Number: 1 Author: reviewer Subject: Replace Date: 2019-07-12 11:41:01

agreement between

Number: 2 Author: reviewer Subject: Note Date: 2019-07-12 11:44:22

Wasn't it calibrated in Glaser et al., 2016?

Number: 3 Author: reviewer Subject: Note Date: 2019-07-12 11:41:41

which time series (GW, SM, frequency of saturation...)?

Number: 4 Author: reviewer Subject: Note Date: 2019-07-12 11:48:04

This refers to the stream sections you were investigating? Can you clarify in the text!

Number: 5 Author: reviewer Subject: Note Date: 2019-07-12 11:51:05

Again, can you better explain the difference between Fig 7a and 7b that you refer here in the text?

© Author(s) 2019. CC BY 4.0 License.

Weierbach catchment follows hysteretic loops, but that the hysteresis was not resolved with the available temporal resolution of the observations. For example, it is likely that surface saturation evolved in the riparian areas during high flow conditions and persisted on the ground surface during decreasing streamflow due to restricted infiltration capacities of the riparian soil (cf. Antonelli et al., 2019).

The lack of such a hysteretic process in the simulation could explain why the model showed the tendency for less persistent and faster contracting surface saturation. It may also explain why the simulated saturation dynamics differed less between the different investigated areas than the observed dynamics. It is likely that the observed saturation dynamics were not synchronous between the different areas due to a less persistent (and thus hysteretic) generation of surface saturation in the relatively narrow riparian areas without perennial springs (M1 and R2) compared to the wider riparian areas with perennial springs (cf. observation of less persistent saturation in M1 and R2 during February and April 2016, Fig. 4). The model, instead, simulated a non-hysteretic saturation behaviour for all investigated riparian areas, which resulted in a better fit between simulated and observed dynamics in the areas M1 and R2 compared to the other areas.

At the same time, it might also be that the simulated relationship between saturation and discharge was correct in all riparian areas and that the scattering of the observation data did not result from hysteretic behaviour, but from uncertainties in the TIR methodology. A good argument for a correct simulation of the discharge – surface saturation relationship is that not only simulated saturation but also simulated discharge seemed to be less persistent and to decrease and increase earlier than it was observed. In reality, the scatter of the observation data is likely related to both measurement uncertainties and hysteretic aspects and a future study with higher temporal resolution of field observations and corresponding simulation output could further analyse this.

Independently from the question on hysteretic loops, we found that the discharge – surface saturation relationships somewhat differed between the different areas. We could connect the main differences to different topographical and morphological features, yet we cannot decipher why the main controlling feature for the discharge – surface saturation relationship was different between observations (source areas vs non-source areas) and simulations (different stream branches, cf. Section 4.3). Nonetheless, our findings are in line with experimental studies that discussed that the relationships between baseflow discharge and total extent of contributing saturated areas differ between catchments with different physiographic characteristics (e.g. Dunne et al., 1975; Latron and Gallart, 2007).

By comparing our model results to the double logarithmic plot presented by Latron and Gallart (2007) (Figure 8), we could identify similar shape varieties of the discharge – surface saturation relationship for the different areas studied within the Weierbach catchment as observed for the different catchments presented in Latron and Gallart (2007). We cannot compare our results directly with the results shown in Latron and Gallart (2007), since we evaluated absolute discharge and normalized saturation, while they evaluated connected saturated areas in percentage of catchment area, but discharge normalized to the catchment area. In order to facilitate the comparison and to connect the two plots (Fig. 8a, 8b), we show the simulated relationship between discharge and surface saturation of the entire Weierbach catchment in both plots, once with normalized discharge and absolute saturation (Fig. 8a), and once with absolute discharge and normalized saturation (Fig. 8b). The shape of the relationship for the entire Weierbach catchment was nearly linear, similar to the relationship observed in the Can Vila catchment investigated by Latron and Gallart (2007) (Fig. 8a). The relationships of the seven studied riparian areas differed from the catchment relationship and between each other (Fig. 8b). For example, area S2 and M1 showed a convex shape similar to the observations in the Vermont W2 catchment made by Dunne et al. (1975), area M3 showed a rather concave shape similar to the relationships found for a sub-catchment of the Asker basin (Myrabø, 1986) and the Strengbach catchment (Latron, 1990), area M2 showed a rather linear shape similar to the Can Vila catchment studied by Latron and Gallart (2007). This clearly shows that differences in the relationship between surface saturation and discharge do not only occur between different catchments, but that they also occur as intra-catchment variability.

No Comments.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

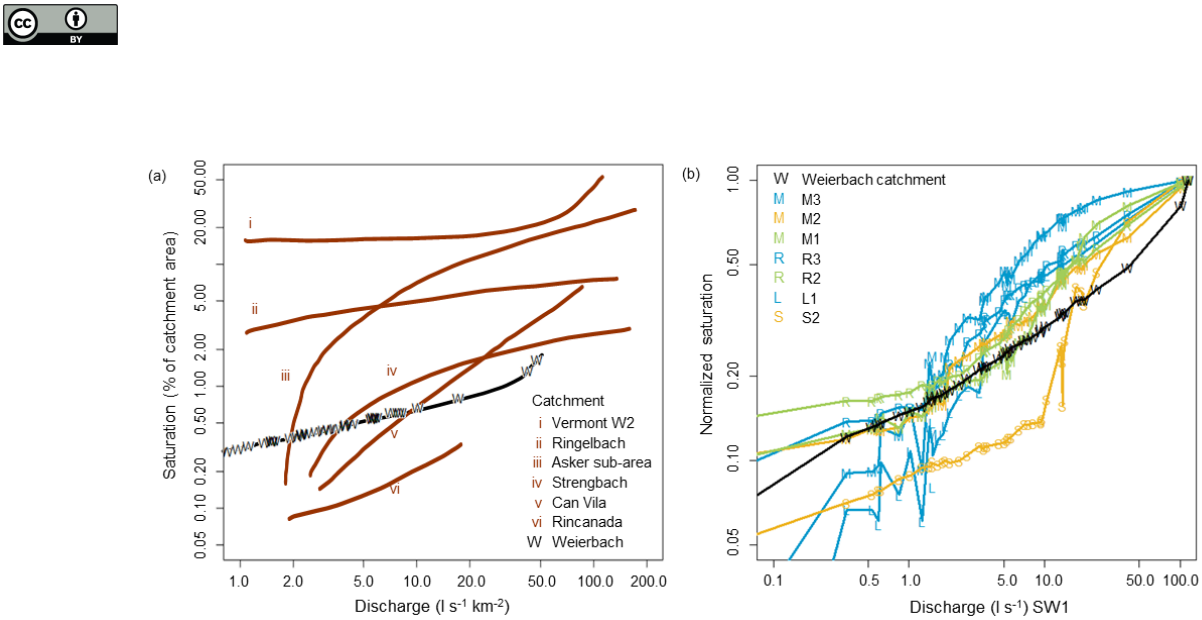

Figure 8: Simulated relationship between discharge and surface saturation of the entire Weierbach catchment (marked with W) in comparison to (a) the relationships observed in other catchments (Figure modified from Latron and Gallart (2007) and (b) the relationships simulated for the seven investigated riparian areas within the catchment. The presented relationships of the other catchments were investigated by i) Dunne et al. (1975), ii) Ambroise (1986), iii) Myrabø (Myrabø, 1986), iv) Latron (Latron, 1990), v) Latron and Gallart (2007), and vi) Martinez-Fernandez et al. (2005). Area affiliation for the investigated riparian areas of the Weierbach catchment is indicated with the respective colour and letter (cf. Fig. 4-6).

5.3 Spatial patterns of surface saturation

The observed spatial patterns of surface saturation were reproduced with the simulations in great detail for most of the investigated areas. We attribute the successful simulation of the spatial patterns to microtopography (local topographical features with extents of centimetres to few metres) since i) microtopography described the main spatial variability between the seven investigated areas in the model setup and ii) we observed that small changes in the setup and resolution of the model mesh in the riparian zones changed some details of the simulated surface saturation patterns (Fig. S2, especially area M2, S2). Therefore, we would like to stress that not only major topographic features of the catchment (e.g. hillslope shape, slope angle, valley width) but also its microtopography needs to be considered for identifying locations where surface saturation may occur. This may sound trivial and several studies have already pointed out the importance of microtopography for the simulation of different hydrological aspects such as hydraulic heads, hyporheic surface-subsurface water exchange, bank storage and overbank flooding, water quality of shallow groundwater systems and runoff generation (e.g. Aleina et al., 2015; Frei et al., 2010; Käser et al., 2014; Van der Ploeg et al., 2012; Tang et al., 2018). Still, microtopography is not often considered in the simulation of surface saturation patterns.

When microtopography is not resolved detailed enough, it is more likely that the simulated surface water extends over a large area instead of accumulating in topographic depressions and thus overrates the extent of surface saturation. In this context it is interesting to note that there are studies that simulated maximum extents of surface saturation up to 80 % of the study area (Qu and Duffy, 2007; Weill et al., 2013), while field observations have only reached maximum extents up to 25 % - 50 % of catchment area (Ali et al., 2014; Birkel et al., 2010; Dunne et al., 1975; Mengistu and Spence, 2016) and often show maximum extents around 10 % (Ambroise, 2016; Grabs et al., 2009; Güntner et al., 2004; Latron and Gallart, 2007; Tanaka et al., 1988). Microtopography might partly explain this discrepancy, even though the maximum extent of surface saturation certainly also depends on the climatic and physiographic conditions of the catchment and on the timing of the observations (e.g. baseflow conditions vs storm events) and there are some studies that simulated similar or less maximum extent of surface saturation

No Comments.

© Author(s) 2019. CC BY 4.0 License.

than observed without considering the microtopography (e.g. Ali et al., 2014; Birkel et al., 2010; Grabs et al., 2009; Güntner et al., 2004; Mengistu and Spence, 2016).

In our study, the simulated maximum extent of surface saturation was 1.6 % of catchment area, which is small compared to other simulation studies, but matches the observation that surface saturation commonly only occurs within the riparian zone and streambed (extent of 1.2 %). Nonetheless, also our maximum saturation within the individual areas was overestimated compared to the observations (cf. Fig. 4). Besides the effect of microtopography, there are two other possible explanations for this. First, the largest simulated saturation occurred during winter 2017/2018, which is the same period where the model clearly overestimated discharge. This mismatch could partly explain the overestimation of saturation, assuming that the relationship between discharge and saturation was correctly captured with the model (cf. Section 5.2). Second, the overestimation of absolute saturation could result from different perspectives and extensions of model output and TIR images (cf. section 3.2, Fig. 6). The TIR images included parts of the hillslopes around the riparian zones, which were not included to the same extent in the extracted model images. Since the hillslopes normally remained unsaturated, the maximum possible amount of saturated pixels in the TIR images was thus lower than in the model images, while the minimum possible amount of saturation was not affected. This could also explain why overestimation of total amounts of saturation was different between the different areas.

Despite the importance of microtopography, the model results showed that microtopography alone was not sufficient to capture the spatial patterns of surface saturation correctly. The simulated patterns of surface saturation clearly did not match the observed patterns equally well in all seven investigated areas (cf. Fig. 6), although the topographical information source and mesh resolution was consistent for the simulated riparian areas. This means that there are additional factors that control the spatial patterns of surface saturation that were not accounted for in the simulations. Such a factor could for example be the structure of the subsurface, which was treated as being homogeneous between all investigated riparian areas in the simulations. In reality, the subsurface structure may locally differ to some degree, for example in the riparian area of the western stream branch (L1), where saturation was simulated at a clearly wrong side along the stream.

5.4 Frequency maps of surface saturation

The frequency maps of surface saturation combine information on when and where surface saturation occurs. We do not think that the exfiltration of subsurface water into local depressions (cf. Section 5.1 and 5.2) can fully explain the spatial variability of saturation frequencies that was observed and simulated satisfactorily within the different riparian areas (Fig. 6). Instead, we suppose that the differences in saturation frequency were controlled by additional water sources than exfiltrating groundwater, such as stream water or direct precipitation, and that the contribution of these additional water sources to surface saturation varied in space and time. For example, the lower frequencies of surface saturation observations at the streamsides compared to the streambed and the lower frequencies in the streambed of the source areas (L1, M3, R3) compared to the mid- and downstream areas (M2, S2, M1, R2) might reflect a lower and less frequent contribution of upstream water in these areas. The overestimation of simulated saturation frequencies in the streambed of R3 could thus indicate an overestimated upstream contribution due to simulating the stream extent too far upstream from the source area. Future work should analyse potential water sources and generation processes of surface saturation with a suitable model framework (cf. Partington et al., 2013; Weill et al., 2013) in order to complement the interpretation of the observation data and to identify the mixture of different water sources of surface saturation (e.g. stream water, exfiltrating subsurface water, ponding precipitation), how the sources might vary in space and time, and how this might reflect in the surface saturation frequencies.

6 Summary and conclusions

We explored the intra-catchment variability of surface saturation in the Weierbach catchment with joint observations and simulations. We showed that the model could reproduce the observed variability of the surface saturation characteristics

No Comments.

© Author(s) 2019. CC BY 4.0 License.

(dynamics, frequencies, patterns) with great detail, although the model setup was rather homogeneous and parameters were not calibrated at catchment scale. Our results demonstrated that a spatially distributed, physically-based, integrated hydrologic model such as HGS is well-suited for reproducing and analysing the generation and development of surface saturation in space and time.

Based on the matches and mismatches between the simulation results and observations, we could identify some key factors controlling the surface saturation generation. The temporal occurrence of surface saturation was observed and simulated to be similar across the catchment, which we related – based on the simulation results – to a large influence of groundwater that reacts synchronous across the catchment. The spatial occurrence of the surface saturation differed between and within the seven investigated riparian areas, which we mainly could relate to the influence of microtopography. Furthermore, we discussed that the full variability between the different areas and the mismatches between observations and simulation can only be explained with additional factors besides groundwater exfiltration and microtopography.

The spatially varying frequencies of surface saturation within the riparian areas indicated that there might be additional water sources than subsurface water that contribute to the generation of surface saturation. Since the model could reproduce the observed frequencies, the model can be used in a future study to analyse such a potential mixing of different water sources and their variation in space and time. The observed differences between the investigated riparian areas with regard to the seasonal dynamics of saturation extension and contraction and the surface saturation – discharge relationship likely resulted from different morphological characteristics (width, existence of perennial springs) of the riparian areas. Although the model could not reproduce a varying hysteretic occurrence and persistence of surface saturation in the different investigated areas, also the simulation results demonstrated that the relationship between surface saturation and discharge can differ within a catchment in the same manner as between catchments with different topographical and morphological conditions.

Data availability. Data underlying the study are property of the Luxembourg Institute of Science and Technology. They are available on request from the authors.

Author contributions. BG, LH and JK designed and directed the study. BG and MA planned and carried out the field work and processed the TIR images. BG set up the simulation and processed the model output. BG, MA, LH and JK discussed and interpreted the results. BG prepared the manuscript with contributions from JK and LH.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. We wish to thank Jean-Francois Iffly, Jérôme Juilleret, the Observatory for Climate and Environment of LIST, and the Administration des Services Techniques de l'Agriculture (ASTA) for the collection and provision of the hydrometrical and meteorological data. We acknowledge deployment of a trial version of AlgoMesh by HydroAlgorithmics Pty Ltd. Barbara Glaser thanks the Luxembourg National Research Fund (FNR) for funding within the framework of the FNR-AFR Pathfinder project (ID 10189601). Marta Antonelli was funded by the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no. 607150 (FP7-PEOPLE-2013-ITN – INTERFACES – Ecohydrological interfaces as critical hotspots for transformation of ecosystem exchange fluxes and bio-geochemical cycling).

References

Ala-aho, P., Rossi, P. M., Isokangas, E. and Kløve, B.: Fully integrated surface–subsurface flow modelling of groundwater–lake interaction in an esker aquifer: Model verification with stable isotopes and airborne thermal imaging, J. Hydrol., 522, 391–406, doi:10.1016/j.jhydrol.2014.12.054, 2015.

Page:19

Number: 1 Author: reviewer Subject: Note Date: 2019-07-12 12:10:55

not re-calibrated but taken fro Glaser et al., (2016).

Number: 2 Author: reviewer Subject: Note Date: 2019-07-12 12:12:40

I suggest to limit this statement to environmental conditions similar to the Weierbach Catchment?

Number: 3 Author: reviewer Subject: Insert Text Date: 2019-07-12 12:15:46

of surface saturation

© Author(s) 2019. CC BY 4.0 License.

Aleina, F. C., Runkle, B. R. K., Kleinen, T., Kutzbach, L., Schneider, J. and Brovkin, V.: Modeling micro-topographic controls on boreal peatland hydrology and methane fluxes, Bio, 12, 5689–5704, doi:10.5194/bg-12-5689-2015, 2015.

Ali, G., Birkel, C., Tetzlaff, D., Soulsby, C., Mcdonnell, J. J. and Tarolli, P.: A comparison of wetness indices for the prediction of observed connected saturated areas under contrasting conditions, Earth Surf. Process. Landforms, 39(3), 399–413, doi:10.1002/esp.3506, 2014.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration (guidelines for computing crop water requirements), FAO Irrig. Drain. Pap., 56, 1998.

Ambroise, B.: Rôle hydrologique des surfaces saturées en eau dans le basin du Ringelbach à Soultzeren (Hautes-Vosges), France, in Recherches sur l'Environnement dans la Région, Actes du 1er Colloque Scientifique des Universités du Rhîn Supérieur, edited by O. Rentz, J. Streith, and L. Ziliox, pp. 620–630, Université Louis Pasteur - Conseil de l'Europe, Strasbourg., 1986.

Ambroise, B.: Variable water-saturated areas and stream fl ow generation in the small Ringelbach catchment (Vosges Mountains , France): the master recession curve as an equilibrium curve for interactions between atmosphere , surface and ground waters, Hydrol. Process., 30, 3560–3577, doi:10.1002/hyp.10947, 2016.

Antonelli, M., Glaser, B., Teuling, R., Klaus, J. and Pfister, L.: Saturated areas through the lens: 1. Spatio-temporal variability of surface saturation documented through Thermal Infrared imagery, In preparation, 2019.

Birkel, C., Tetzlaff, D., Dunn, S. M. and Soulsby, C.: Towards a simple dynamic process conceptualization in rainfall – runoff models using multi-criteria calibration and tracers in temperate, upland catchments, Hydrol. Process., 24, 260–275, doi:10.1002/hyp.7478, 2010.

Carrer, G. E., Klaus, J. and Pfister, L.: Assessing the Catchment Storage Function Through a Dual-Storage Concept, Water Resour. Res., 55, 476–494, doi:10.1029/2018WR022856, 2019.

Davison, J. H., Hwang, H.-T., Sudicky, E. A., Mallia, D. V. and Lin, J. C.: Full Coupling Between the Atmosphere, Surface, and Subsurface for Integrated Hydrologic Simulation, J. Adv. Model. Earth Syst., 10, 43–53, doi:10.1002/2017MS001052, 2018.

Dunne, T., Moore, T. R. and Taylor, C. H.: Recognition and prediction of runoff-producing zones in humid regions, Hydrol. Sci. Bull., 20, 305–327, 1975.

Erler, A. R., Frey, S. K., Khader, O., Orgeville, M., Park, Y.-J., Hwang, H.-T., Lapen, D. R., Peltier, W. R. and Sudicky, E. A.: Simulating Climate Change Impacts on Surface Water Resources Within a Lake-Affected Region Using Regional Climate Projections, Water Resour. Res., 55, 130–155, doi:10.1029/2018WR024381, 2019.

Frei, S., Lischeid, G. and Fleckenstein, J. H.: Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland --- A modeling study, Adv. Water Resour., 33(11), 1388–1401, doi:10.1016/j.advwatres.2010.07.006, 2010.

Gburek, W. J. and Sharpley, A. N.: Hydrologic Controls on Phosphorus Loss from Upland Agricultural Watersheds, J. Environ. Qual., 27, 267–277, doi:10.2134/jeq1998.00472425002700020005x, 1998.

Glaser, B., Klaus, J., Frei, S., Frentress, J., Pfister, L. and Hopp, L.: On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum, Water Resour. Res., 52, 8317–8342, doi:10.1002/2015WR018414, 2016.

Glaser, B., Antonelli, M., Chini, M., Pfister, L. and Klaus, J.: Technical note: Mapping surface-saturation dynamics with thermal infrared imagery, Hydrol. Earth Syst. Sci., 22(11), 5987–6003, doi:10.5194/hess-22-5987-2018, 2018.

Glaser, B., Jackisch, C., Hopp, L. and Klaus, J.: How meaningful are plot-scale observations and simulations of preferential flow for catchment models ?, Vadose Zo. J., doi:10.2136/vzj2018.08.0146, 2019.

Gourdol, L., Clément, R., Juilleret, J., Pfister, L. and Hissler, C.: Large-scale ERT surveys for investigating shallow regolith properties and architecture, Hydrol. Earth Syst. Sci. Discuss., 1–39, doi:10.5194/hess-2018-519, 2018.

No Comments.

© Author(s) 2019. CC BY 4.0 License.

Grabs, T., Seibert, J., Bishop, K. and Laudon, H.: Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model, J. Hydrol., 373(1-2), 15–23, doi:10.1016/j.jhydrol.2009.03.031, 2009.

Güntner, A., Seibert, J. and Uhlenbrook, S.: Modeling spatial patterns of saturated areas: An evaluation of different terrain indices, Water Resour. Res., 40, W05114, doi:10.1029/2003WR002864, 2004.

Gupta, H. V, Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Juilleret, J., Iffly, J. F., Pfister, L. and Hissler, C.: Remarkable Pleistocene periglacial slope deposits in Luxembourg (Oesling): pedological implication and geosite potential, Bull. la Société des Nat. Luxemb., 112(1), 125–130 [online] Available from: http://www.snl.lu/publications/bulletin/SNL_2011_112_125_130.pdf, 2011.

Käser, D., Graf, T., Cochand, F., McLaren, R., Therrien, R. and Brunner, P.: Channel Representation in Physically Based Models Coupling Groundwater and Surface Water : Pitfalls and How to Avoid Them, Groundwater, 52(6), 827–836, doi:10.1111/gwat.12143, 2014.

Kristensen, K. J. and Jensen, S. E.: A model for estimating actual evapotranspiration from potential evapotranspiration, Nord. Hydrol., 6(3), 170–188, doi:10.2166/nh.1975.012, 1975.

Latron, J.: Caractérisation géomorphologique et hydrologique du bassin versant versant du Strengbach (Aubure), Université Louis Pasteur, Strasbourg I., 1990.

Latron, J. and Gallart, F.: Seasonal dynamics of runoff-contributing areas in a small mediterranean research catchment (Vallcebre, Eastern Pyrenees), J. Hydrol., 335(1-2), 194–206, doi:10.1016/j.jhydrol.2006.11.012, 2007.

Martínez Fernández, J., Ceballos Barbancho, A., Morán Tejeda, C., Casado Ledesma, S. and Hernández Santana, V.: Procesos hidrológicos en una cuenca forestal del Sistema Central: Cuenca experimental de Rinconada, Cuad. Investig. Geográfica, 31, 7–25 [online] Available from: https://dialnet.unirioja.es/servlet/articulo?codigo=1975892, 2005.

Martínez-Carreras, N., Hissler, C., Gourdol, L., Klaus, J., Juilleret, J., Iffly, J. F. and Pfister, L.: Storage controls on the generation of double peak hydrographs in a forested headwater catchment, J. Hydrol., 543, 255–269, doi:10.1016/j.jhydrol.2016.10.004, 2016.

Megahan, W. F. and King, P. N.: Identification of critical areas on forest lands for control of nonpoint sources of pollution, Environ. Manage., 9(1), 7–17, doi:10.1007/BF01871440, 1985.

Mengistu, S. G. and Spence, C.: Testing the ability of a semidistributed hydrological model to simulate contributing area, Water Resour. Res., 52, 4399–4415, doi:10.1002/2016WR018760, 2016.

Moragues-Quiroga, C., Juilleret, J., Gourdol, L., Pelt, E., Perrone, T., Aubert, a., Morvan, G., Chabaux, F., Legout, a., Stille, P. and Hissler, C.: Genesis and evolution of regoliths: Evidence from trace and major elements and Sr-Nd-Pb-U isotopes, Catena, 149, 185–198, doi:10.1016/j.catena.2016.09.015, 2017.

Munz, M., Oswald, S. E. and Schmidt, C.: Coupled Long-Term Simulation of Reach-Scale Water and Heat Fluxes Across the River-Groundwater Interface for Retrieving Hyporheic Residence Times and Temperature Dynamics, Water Resour. Res., (53), 8900–8924, doi:10.1002/2017WR020667, 2017.

Myrabø, S.: Runoff Studies in a Small Catchment, Nord. Hydrol., 17, 335–346, doi:10.2166/nh.1986.0025, 1986.

Nasta, P., Boaga, J., Deiana, R., Cassiani, G. and Romano, N.: Comparing ERT- and scaling-based approaches to parameterize soil hydraulic properties for spatially distributed model applications, Adv. Water Resour., 126, 155–167, doi:10.1016/j.advwatres.2019.02.014, 2019.

Nippgen, F., McGlynn, B. L. and Emanuel, R. E.: Water Resources Research, Water Resour. Res., 51, 4550–4573, doi:10.1002/2014WR016719.Received, 2015.

Ogden, F. L. and Watts, B. a.: Saturated area formation on nonconvergent hillslope topography with shallow soils: A numerical investigation, Water Resour. Res., 36(7), 1795, doi:10.1029/2000WR900091, 2000.

No Comments.

© Author(s) 2019. CC BY 4.0 License.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., Dandy, G. C. and Fleckenstein, J. H.: Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, Water Resour. Res., 49(9), 5501–5519, doi:10.1002/wrcr.20405, 2013.

Pfister, L., McDonnell, J. J., Hissler, C. and Hoffmann, L.: Ground-based thermal imagery as a simple, practical tool for mapping saturated area connectivity and dynamics, Hydrol. Process., 24(21), 3123–3132, doi:10.1002/hyp.7840, 2010.

Van der Ploeg, M. J., Appels, W. M., Cirkel, D. G., Oosterwoud, M. R., Witte, J.-P. M. and Van der Zee, S. E. A. T. M.: Microtopography as a Driving Mechanism for Ecohydrological Processes in Shallow Groundwater Systems, Vadose Zo. J., 11(3), doi:10.2136/vzj2011.0098, 2012.

Qu, Y. and Duffy, C. J.: A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43(8), 1–18, doi:10.1029/2006WR005752, 2007.

Reaney, S. M., Bracken, L. J. and Kirkby, M. J.: The importance of surface controls on overland flow connectivity in semi-arid environments: Results from a numerical experimental approach, Hydrol. Process., 28(4), 2116–2128, doi:10.1002/hyp.9769, 2014.

Scaini, A., Audebert, M., Hissler, C., Fenicia, F., Gourdol, L., Pfister, L. and Beven, K. J.: Velocity and celerity dynamics at plot scale inferred from artificial tracing experiments and time-lapse ERT, J. Hydrol., 546, 28–43, doi:10.1016/j.jhydrol.2016.12.035, 2017.

Schilling, O. S., Gerber, C., Partington, D. J., Purtschert, R., Brennwald, M. S., Kipfer, R., Hunkeler, D. and Brunner, P.: Advancing Physically-Based Flow Simulations of Alluvial Systems Through Atmospheric Noble Gases and the Novel 37Ar Tracer Method, Water Resour. Res., 53, 10,465–10,490, doi:10.1002/2017WR020754, 2017.

Sebben, M. L., Werner, A. D., Liggett, J. E., Partington, D. and Simmons, C. T.: On the testing of fully integrated surface – subsurface hydrological models, Hydrol. Process., 27, 1276–1285, doi:10.1002/hyp.9630, 2013.

Silasari, R., Parajka, J., Ressl, C., Strauss, P. and Blöschl, G.: Potential of time-lapse photography for identifying saturation area dynamics on agricultural hillslopes, Hydrol. Process., 1–18, doi:10.1002/hyp.11272, 2017.

Tanaka, T., Yasuhara, M., Sakai, H. and Marui, A.: The Hachioji experimental basin study -- storm runoff processes and the mechanism of its generation, J. Hydrol., 102, 139–164, 1988.

Tang, Q., Schilling, O. S., Kurtz, W., Brunner, P., Vereecken, H. and Hendricks Franssen, H.-J.: Simulating Flood-Induced Riverbed Transience Using Unmanned Aerial Vehicles , Physically Based Hydrological Modeling, and the Ensemble Kalman Filter, Water Resour. Res., (54), 9342–9363, doi:10.1029/2018WR023067, 2018.

Weill, S., Altissimo, M., Cassiani, G., Deiana, R., Marani, M. and Putti, M.: Saturated area dynamics and streamflow generation from coupled surface-subsurface simulations and field observations, Adv. Water Resour., 59, 196–208, doi:10.1016/j.advwatres.2013.06.007, 2013.

No Comments.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Intra-catchment variability of surface saturation – insights from long-term observations and simulations

Barbara Glaser[1,2], Marta Antonelli[3,1], Luisa Hopp[2], Julian Klaus[1]

[1]Catchment and Eco-Hydrology Research Group, Luxembourg Institute of Science and Technology, Esch/Alzette, 4362, Luxembourg
[2]Department of Hydrology, University of Bayreuth, Bayreuth, 95447, Germany
[3]Hydrology and Quantitative Water Management Group, Wageningen University & Research, Wageningen, 6700, The Netherlands

Correspondence to: Barbara Glaser (barbara.glaser@list.lu)

Abstract

~~The inundation~ of flood-prone areas varies in space and time and can have crucial impacts on runoff generation and water quality when the ~~surface~ saturated areas become connected to the stream. In this study, we aimed to investigate and explain the variability ~~of~ surface saturation patterns ~~and dynamics~ within a forested headwater catchment. On the one hand, we mapped surface saturation in seven distinct riparian areas of the Weierbach catchment (Luxembourg) ~~with~ thermal infrared images, taken weekly to bi-weekly over a period of two years. On the other hand, we simulated the surface saturation generation in the catchment with the integrated surface subsurface hydrologic model HydroGeoSphere ~~over~ the same period. Both the observations and simulations showed that the saturation dynamics were similar across the catchment, but that small differences between the dynamics at different areas ~~occurred.~ Moreover, the model reproduced the observed saturation patterns well for all seasonal and hydrologic conditions and at all investigated locations. Based on the observations and simulation results and the match~~es~ and mismatch~~es~ between them, we conclude~~d~ that ~~the generation of~ surface saturation in the Weierbach catchment was largely controlled by exfiltration of groundwater into local depressions. However, we also illustrate that the ~~entire~ variability of the patterns, dynamics and frequencies of surface saturation within the different riparian areas of the catchment ~~can only~ result from ~~additional controlling~ factors ~~to~ microtopography and groundwater exfiltration, such as differing hysteretic behaviour, differing subsurface structures, or additional water sources.

1 Introduction

It is critical for flood risk assessment to understand where and when water is standing or flowing on the ground surface outside of perennial surface water bodies. When such surface saturated areas connect to the stream via overland flow, they also become crucial for runoff ~~gen~ration and water quality. In general, surface saturated areas arise from 1) water ponding on the surface due to exceedance of the infiltration capacity of unsaturated soil, 2) water ponding on impermeable surfaces or saturated soil, 3) water exfiltrating from the subsurface or, 4) stream water ~~extending~ into the floodplain (e.g. Megahan and King, 1985). Over the past years and decades, various field studies mapped and analysed the spatial and temporal occurrence of surface saturation within different landscapes (e.g. Ambroise, 1986, 2016; Dunne et al., 1975; Gburek and Sharpley, 1998; Latron and Gallart, 2007; Silasari et al., 2017; Tanaka et al., 1988). From the field studies it is well recognized that surface saturation varies in space and time and ~~that its appearance~ is affected by structural (e.g. topography) and dynamic factors (e.g. precipitation intensity, antecedent moisture). Yet there is limited understanding on how surface saturation evolves spatially and temporally between and within landscapes and how the interplay of different controlling factors and ~~generation~ processes controls the spatio-temporal variability of surface saturation.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Spatially distributed and dynamic hydrological models are potential tools for analysing the generation and development of surface saturation in space and time. Such models allow a detailed investigation of surface saturation at any desired location and time that goes far beyond the information that can be gained by any field observation. Several simulation studies systematically assessed the influence of static and dynamic factors on the temporal evolution, connectivity, and spatial distribution of surface saturation by performing virtual experiments with hillslope models (Ogden and Watts, 2000; Reaney et al., 2014) or by testing a range of terrain indices for predicting time-integrated saturation patterns (Güntner et al., 2004). Other studies relied on dynamic distributed and semi-distributed simulations for analysing connectivity of surface saturation in relation to wetness conditions and catchment runoff (Mengistu and Spence, 2016; Qu and Duffy, 2007; Weill et al., 2013). Weill et al. (2013) and Partington et al. (2013) analysed the processes and water sources that generate surface saturation in a wetland and a pre-alpine grassland headwater, respectively. Both studies applied a model belonging to the group of integrated surface-subsurface hydrologic models (ISSHMs, Sebben et al., 2013), which can simulate the interplay of different surface and subsurface processes of surface saturation generation (e.g. ponding of precipitation from the surface, exfiltration from the subsurface). Modelling studies that focus on a comprehensive spatio-temporal analysis of surface saturation dynamics within a landscape by evaluating the spatially distributed model outputs rather than aggregating the outputs are scarce (e.g. Nippgen et al. (2015) for subsurface saturated areas)

When complementing field observations with simulations to analyse the generation and development of surface saturation in space and time, it is important to ensure that the model yields realistic results. Glaser et al. (2016) demonstrated for a small riparian area that a good match between modelled and observed discharge or soil moisture does not automatically imply a realistic simulation of saturation patterns. They concluded that a spatial validation of the dynamic saturation patterns itself is crucial. However, only few of the existing modelling studies explicitly checked the realism of their simulated surface saturation with field observations before using them for further analyses. These studies focussed either on temporally integrated spatial patterns (Grabs et al., 2009; Güntner et al., 2004) or on temporal dynamics of overall catchment saturation (Birkel et al., 2010; Mengistu and Spence, 2016), but barely any study combined the observation and simulation of both surface saturation patterns and dynamics (Ali et al., 2014; Glaser et al., 2016). The lack of such studies is certainly explainable by the resources that are necessary for obtaining appropriate field data. Today, we still lack a standard method to map surface saturation and the different existing methods such as the 'squishy boot' method, the usage of 'on-off' surface saturation sensors, the mapping of soil morphology or vegetation as surrogates, or the usage of remote sensing techniques (e.g. Dunne et al., 1975; Gburek and Sharpley, 1998; Güntner et al., 2004; Latron and Gallart, 2007; Mengistu and Spence, 2016; Silasari et al., 2017) all have their own advantages and disadvantages.

A relatively new and powerful method for mapping surface saturation is thermal infrared (TIR) imagery. TIR mapping relies on the difference between the surface temperature of water and other materials for identifying surface saturation. Previous work showed that recurrent mapping of surface saturation with high spatial resolution is possible with TIR imagery (Glaser et al., 2016; Pfister et al., 2010). Glaser et al. (2018) and Antonelli et al. (2019) applied TIR imagery mapping in the 42 ha forested Weierbach catchment in western Luxembourg and monitored the dynamics of surface saturation within several distinct riparian areas along the Weierbach stream with a weekly to biweekly mapping frequency over several seasons.

In this study, we explore the intra-catchment variability of temporal and spatial characteristics of surface saturation (dynamics, frequencies, patterns) with a combination of field observation and modelling. We perform the study in the Weierbach catchment, where we can rely on existing TIR imagery data (Antonelli et al., 2019; cf. Glaser et al., 2018) and on previous modelling work for a 6 ha headwater of the catchment (Glaser et al., 2016, 2019) with the ISSHM HydroGeoSphere. Glaser et al. (2016, 2019) simulated the 6 ha area of the catchment by accounting for a layering of the subsurface, while spatial heterogeneity was only represented by microtopography and by a different sequence of subsurface layers in the riparian zone compared to the hillslopes and plateau. Here, we extend the model setup to the entire 42 ha catchment without introducing additional heterogeneity and without performing a re-calibration. We simulate surface saturation in the catchment and contrast

© Author(s) 2019. CC BY 4.0 License.

the results with the observed saturation patterns from the TIR imagery, focussing on the long-term saturation dynamics over different seasons and wetness conditions (25 months with weekly to biweekly mapping resolution) and the spatial patterns of surface saturation occurrence and frequency at seven different riparian areas across the catchment. The two research ~~objectives~~ that we aim to address with this approach are:

1) How variable are surface saturation dynamics and patterns within a catchment and to what extent can we reproduce the variability of the saturation characteristics (dynamics, frequencies, patterns) with a rather homogenously set-up ISSHM?
2) What do we learn about the reasons for the intra-catchment variability of surface saturation characteristics from the match~~es~~ and mismatch~~es~~ between simulation results and observations?

2 Study site and data

2.1 Physiography, climate and hydrometry

The Weierbach catchment is an intensively studied headwater catchment (42 ha) in western Luxembourg. About half of the catchment area is characterized by gentle slopes <5°, forming a plateau landscape unit (Martínez-Carreras et al., 2016). The rest of the catchment is characterized by hillslopes with slopes > 5°, forming a central V-shaped stream valley from north to south and a V-shaped tributary valley in the east. A third, few metres long stream branch is situated ~~in the~~ west of the central stream valley. Riparian zones along the stream account for 1.2 % of the catchment area (Antonelli et al., 2019). Large parts of the catchment are forested with deciduous trees (mainly European beech and Sessile oaks), the south-east and some other small parts are forested with conifers (mainly Norway spruce and Douglas spruce). The riparian zones are free of tree ~~canopy~~ and covered with ferns, moss, and herbaceous plants. ~~Soil~~ developed from Pleistocene Periglacial Slope Deposits ~~as~~ shallow and highly-permeable ~~siltic, sceletic Cambisol~~ with a depth ranging between 0.4 and 0.9 m (Gourdol et al., 2018; Juilleret et al., 2011; Moragues-Quiroga et al., 2017). Beneath the solum, a 0.5 – 1 m thick basal layer with bedrock clasts oriented parallel to the slope overlies fractured Devonian slate and phyllites (Gourdol et al., 2018; Juilleret et al., 2011; Moragues-Quiroga et al., 2017; Scaini et al., 2017). In the riparian zones, soil and basal layer have been eroded and the fractured bedrock is overlain by shallow organic Leptosols (Glaser et al., 2016).

The climate is oceanic-continental without apparent seasonality in precipitation and w[illegible]egligible amounts of snow (Carrer et al., 2019). Mean annual precipitation during the period from October 2013 to September 2017 was 955±53 mm. Mean annual discharge was 546±253 mm, with exceptionally dry conditions in the hydrological year 2017. During wet periods, discharge is characterized by double peak hydrographs with first peaks appearing ~~as~~ immediate response to precipitation and second pronounced peaks appearing 48h to 72h later (cf. Martínez-Carreras et al., 2016). During dry periods, only first hydrograph peaks occur a[illegible] stream dries out intermittently starting from the source areas downstream.

Hydrological and meteorological data that were used in this study were measured from October 2013 to January 2018. Data from the period from October 2013 to September 2015 were used for spin-up simulations, data from the period from October 2015 to January 2018 were used to drive and validate the actual simulation (cf. Section 3). Discharge was measured with water pressure transducers (ISCO 4120 Flow Logger, 15 min logging intervals) at four v-notch~~es~~, installed at the outlet of the catchment (SW1, Fig.1) and upstream of the confluences of the three ~~branches~~ (SW2-SW4). Groundwater levels were continuously recorded every 15 minutes with pressure sensors (OTT CTD) in five piezometers installed in different landscape units (riparian zone, hillslope, plateau) of the catchment (Fig. 1, GW1-3, GW5, GW7). Soil moisture was continuously monitored (30 min logging intervals) with water content reflectometers (CS650, Campbell Scientific) installed horizontally ~~in~~ 10, 20, 40 and 60 cm depth at four different sites (Fig.1, SM3-5, SM7). At each site, two depth profiles were monitored. In addition, soil moisture ~~in~~ 10 cm depth was monitored with water content reflectometers (CS616, Campbell Scientific, 30 min logging intervals) at five locations crosscutting the riparian zone of the stream source area of the middle stream ~~branch~~ (Fig. 1, SSM1-5).

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

Cumulative precipitation was recorded every 5 minutes with a tipping bucket raingauge (Young 52203, unheated, 1 m height) at an open area within the catchment (data gaps were filled ~~by estimating~~ a linear regression to data from a station approximately 4.5 km southward). Potential reference evapotranspiration was estimated based on measured air temperature, relative humidity, wind speed, and net radiation according to the FAO Penman-Monteith formulation (Allen et al., 1998). Air temperature and relative humidity data were recorded next to the soil moisture profile SM5 (Fig. 1, HMP45C-LC, Campbell Scientific, 15 min logging intervals, 2 m height). Wind speed and radiation data were recorded approximately 4.5 km southward of the study site. Wind speed (Young Wind Monitor 05103, Vector A100R Anemometer) was recorded every 15 minutes ~~in~~ 3 m height and converted to wind speed in 2 m height (data gaps closed with data from a station approximately 11.5 km north-eastward) following the FAO guidelines (Allen et al., 1998). Net radiation was recorded every 15 minutes (Kipp & Zonen NR Lite net radiometer) until May 2017. From June 2017 on (and for closing other data gaps), we used net radiation data recorded every 5 minutes close to Luxembourg Airport (~40 km southeast of the study site), as these measurements were highly correlated (linear regression with an intercept of 7.6 W m^{-2}, a slope of 0.92, $R^2 = 0.81$) with the measurements close to the study site in the years before.

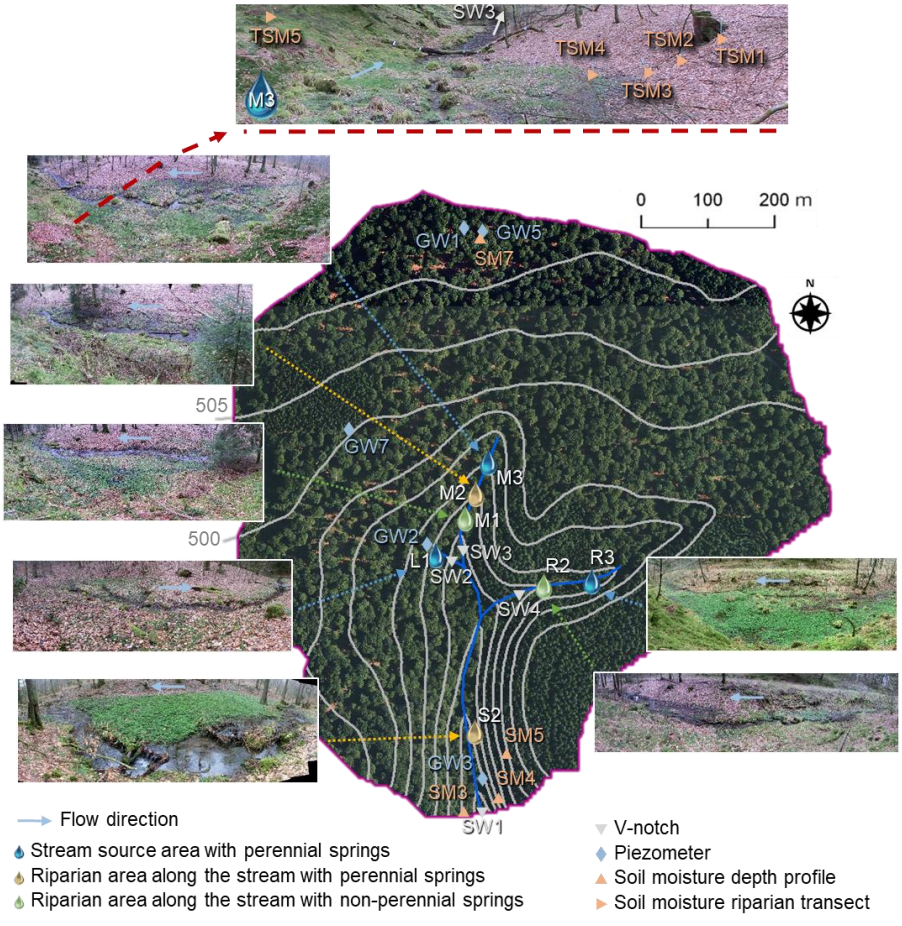

Figure 1: The Weierbach catchment with the locations of the installed v-notches, piezometers, soil moisture sensors and the seven investigated riparian areas.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

2.2 Surface saturation

Here, we define surfaces as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This involves water bodies such as lakes and streams, but excludes mere saturation in the topsoil. According to this definition, surface saturation in the Weierbach catchment generally only occurs in the streambed and the adjacent riparian zones. Other areas that were occasionally observed to be surface saturated during very wet conditions or 'rain on snow' events are forest roads and the prolongation of the streambed above the source regions into the hillslopes. We focus in this study on seven distinct riparian areas in the catchment, which can be classified into three different categories (cf. Antonelli et al., 2019): i) stream source areas with perennial springs (L1, M3, R3, blue areas Fig. 1), ii) areas along the stream with perennial springs (M2, S2, yellow areas Fig. 1), and iii) areas along the stream with non-perennial springs (M1, R2, green areas Fig.1).

We mapped the surface saturation in these seven riparian areas weekly to biweekly from November 2015 to December 2017 with thermal infrared imagery (TIR). Details on the identification of surface saturation with TIR imagery and on the collected surface saturation dataset are presented and discussed in Glaser et al. (2018) and Antonelli et al. (2019). In brief, we created panoramic TIR images of the distinct areas and identified the locations of surface saturation (including the stream) within the images. To do this, each pixel in an image was assigned to be saturated or unsaturated based on the temperature range of locations that were obvious to be derived from field observations and visual images. In case the contrast between water temperature and temperature of surrounding materials was not sufficient for a reliable pixel classification, the images were excluded from the analysis. In case the pixel classification was affected by a poor temperature contrast or by pixels representing vegetation or snow cover in the images, the images were analysed but flagged as less reliable. Altogether, we obtained 291 binary panoramic images showing the temporal dynamics of surface saturation patterns in the seven studied riparian areas with total numbers of images per site ranging between 34 (L1) and 48 (M2).

Time series of saturation were created for each area by accounting for the percentage of saturated pixels within the individual panoramic images. We normalized the saturation percentages to the maximum observed percentage of saturation in the distinct areas in order to allow a comparison of the saturation dynamics between the different riparian areas. For picturing the spatial surface saturation dynamics within a distinct riparian area, we created maps of saturation frequency. We counted for each area how often the individual pixels of the panoramic TIR images were classified as saturated and normalized the resulting frequency numbers by the total number of TIR images analysed for that area.

The resulting maps of normalized saturation frequency rarely showed pixels that were always saturated (i.e. reaching a normalized frequency of 1). In reality, surface saturation was more persistent than indicated by the frequency maps. The reason for this artefact is that the perspective of the individual TIR panoramas was not 100% identical for all mapping instances and that vegetation sometimes covered parts of the saturated surface, especially during near-by conditions. We co-registered the individual panoramas against a reference panorama for each area, but slight positional shifts were inevitable. As a result, the images that were placed on top of each other did not always overlap exactly and the generated saturation frequency maps are blurred. Nonetheless, the maps of normalized saturation frequencies are very useful to understand at a glance where surface saturation occurs more and less frequent within an area and to be used for model validation.

3 Catchment model

3.1 Model setup and parameterisation

We simulated the spatio-temporal dynamics of surface saturation across the Weierbach catchment with HydroGeoSphere (HGS, Aquanty Inc.). HGS is an integrated surface subsurface hydrological model and allows simultaneous simulation of transient surface and subsurface flow. Subsurface flow is simulated based on the 3D Richards equation. Surface flow is simulated based on the diffusive-wave approximation of the 2D Saint Venant equation. Evapotranspiration is simulated with a comparatively simple approach, following the mechanistic concept of Kristensen and Jensen (1975). The equations are

© Author(s) 2019. CC BY 4.0 License.

linearized implicitly using the Newton-Raphson approach and solved in an unstructured finite element grid. HGS has been used in the past ~~for~~ addressing diverse questions at various temporal and spatial scales (e.g. Ala-aho et al., 2015; Davison et al., 2018; Erler et al., 2019; Frei et al., 2010; Munz et al., 2017; Nasta et al., 2019; Partington et al., 2013; Schilling et al., 2017; Tang et al., 2018). It also ~~has~~ already ~~been~~ applied for a 6 ha headwater region of the Weierbach catchment (Glaser et al., 2016, 2019). In this study, we applied the parameterization of Glaser et al. (2016) to the entire 42 ha catchment without ~~performing~~ an additional parameter calibration.

The catchment was spatially discretized into 42,274 triangular elements, using the mesh generator AlgoMesh (HydroAlgorithmics Pty Ltd). Edge lengths of the mesh elements ranged from > 30 m at the plateau to < 0.4 m for the seven analysed riparian zones and the streambed (Fig. 2). It was crucial to use such a fine mesh resolution in the riparian zone in order to enable a comparable spatial detail as obtained with the TIR imagery for the surface saturation patterns. Vertically, the model grid comprised 5 m, which ~~were~~ divided into 14 layers with element depths ranging from 0.15 m for the top layers to 0.5 m for the bottom layers (Fig. 2).

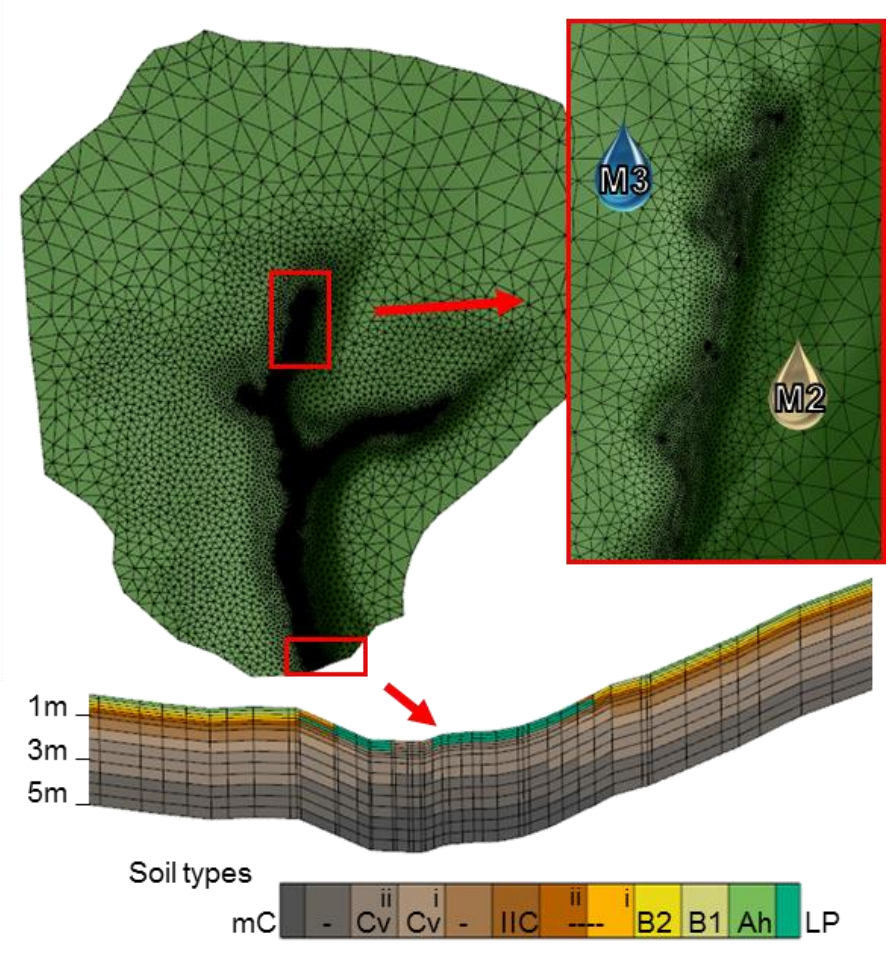

Figure 2: Setup of the model mesh with a zoom on the fine horizontal resolution in the riparian areas and the streambed (inset on the right) and a vertical cross section through the stream valley and adjacent hillslopes (bottom) showing the vertical discretization and assignment of different soil properties (cf. Tab. 1). Ah = topsoil, B1 and B2 =subsoil, IIC = basal layer, Cv = fractured bedrock, Cm = fresh bedrock.

© Author(s) 2019. CC BY 4.0 License.

The subsurface was parameterized homogeneously with 10 different property layers, representing top- and subsoil (Ah, B1, B2), the basal layer (IIC), fractured and fresh bedrock (Cv, Cm), and layers ~~of transition~~ between subsoil, basal layer, and fractured bedrock (Fig. 2). We implemented spatial heterogeneity in the stream valleys, ~~where~~ soil and basal layer were eroded and the outcropping fractured bedrock ~~was~~ overlain with organic, stagnic Leptosol ~~in the ripar[illegible]ones~~ (Fig. 2). We used the Mualem-van Genuchten soil hydraulic functions ~~for~~ describing the saturation-pressure relation. The necessary soil hydraulic parameter values for the different property layers (porosity, residual saturation, van Genuchten α, van Genuchten β, saturated hydraulic conductivity, Tab. 1) were assigned according to Glaser et al. (2016). We only ~~parameterised~~ one additional layer ~~for~~ the fractured bedrock (Cv (ii)) in order to account for the adapted depth of 5 m in the catchment model compared to the depth of 3 m in the headwater model.

Table 1: Soil hydraulic parameters of the different soil property zones. Table adapted from Glaser et al. (2016)

Soil property zone	Residual saturation	van Genuchten parameter α [m^{-1}]	van Genuchten parameter β	Porosity	Saturated hydraulic conductivity [m d^{-1}]
Ah	0.12	6.6	1.46	0.74	1.71E+01
B1	0.10	22.1	1.42	0.61	1.71E+01
B2	0.10	22.1	1.42	0.45	4.59E+01
B2-IIC (i)	0.10	22.1	1.42	0.3	9.30E+02
B2-IIC (ii)	0.10	22.1	1.42	0.15	2.04E+03
IIC	0.02	6.0	1.50	0.20	8.40E+02
IIC-Cv	0.02	6.0	1.50	0.15	3.00E+00
Cv (i)	0.02	6.0	1.50	0.10	1.20E-02
Cv (ii)	0.02	6.0	1.50	0.07	1.20E-02
Cv-mC	0.02	6.0	1.50	0.05	9.00E-04
mC	0.02	6.0	1.50	0.01	2.40E-05
LP	0.10	22.1	1.42	0.61	7.80E+00

Surface and subsurface flow were coupled via a Darcy flux exchange through a thin coupling layer (10^{-4} m). We assumed different Manning's surface roughness values for the forested area ($1.24*10^{-6}$ d $m^{-1/3}$), the riparian zone ($9.41*10^{-7}$ d $m^{-1/3}$), and the stream bed ($4.4*10^{-7}$ d $m^{-1/3}$) (cf. Glaser et al., 2016). Evapotranspiration ~~properties~~ (Tab. S1) were assigned individually for the deciduous forest, the coniferous forest in the southeast of the catchment, and the riparian zones including the streambed and values were based on the calibrated values of Glaser et al. (2016). The simulation was driven with daily sums of precipitation and reference evapotranspiration, which were treated as being spatially uniform. The outer edge of the surface domain was assigned ~~as~~ critical depth boundary, allowing water to leave the model domain via surface flow. Side and bottom boundaries of the subsurface domain were no flow boundaries. A spin-up simulation drained the catchment from full saturation to steady state conditions (for 1 mm d^{-1} of precipitation, no evapotranspiration) and subsequently repeated the period from October 2013 to October 2015 three times ~~for~~ obtaining realistic initial conditions. The actual simulation spanned ~~over~~ the period from October 2015 to January 2018, the period where we mapped surface saturation with TIR imagery.

3.2 Assessment of model performance

We benchmarked the model against measured discharge, groundwater level, soil moisture, and surface saturation patterns and dynamics at various locations (Fig. 1). We calculated the Kling Gupta Efficiency (KGE) as a combined measure for correlation, bias, and relative variability (Gupta et al., 2009) between simulated and observed discharge. We also calculated KGEs for the simulated groundwater levels, but particularly evaluated the groundwater level dynamics rather than absolute values based on

© Author(s) 2019. CC BY 4.0 License.

Pearson correlation coefficients. Soil moisture was also evaluated based on its dynamics with Pearson correlation coefficients, while absolute values were only compared visually. Since simulated soil moisture was extracted from model nodes whose depths did not exactly correspond with the measurement depths, we interpolated depth-weighted average values from the model output ~~for~~ calculat~~ing~~ the correlation with the observations in the respective depths. The interpolated model values of volumetric water content were then correlated with the observations of water content, averaging the measurements of the ~~twin~~ depth profiles at ~~the~~ monitoring sites.

For compar~~ing~~ the simulation output with the surface saturation information obtained with the TIR images, it was necessary to convert the model output into a comparable format via several processing steps: First, we extracted the surface water depths in the surface domain of the model ~~for~~ noon of the days where TIR images were taken ~~and analysed~~. Next, we transformed the surface water depths into a binary saturation map of the entire catchment by classifying the surface domain cells as saturated if water depths were $>10^{-4}$ m. The depth of 10^{-4} m corresponds to the penetration depth of the used TIR camera for water columns and thus is the minimum depth that could be detected ~~as~~ pure water temperature signal with the camera. Finally, we ~~proj~~ the model output into jpeg images with the same perspective and extent of the TIR panoramic images by turning, bending, and cutting the modelled saturation maps according to each of the seven riparian areas individually. This model output processing allowed us to perform the same calculations for the model output as for the TIR images, i.e. to create time series of normalized saturation and maps of normalized saturation frequencies for the seven riparian areas with comparable perspectives and extents. Since it was not possible to project the model output identically to the perspectives of the TIR images, we compared the saturation dynamics and patterns of the model images with the observations qualitatively (visually) only. A quantitative comparison would have been biased by differences in image distortions and total area extent.

Furthermore, we compared the simulated frequency of surface saturation with the simulated frequency of groundwater reaching the surface. To do this, we marked the surface cells below which the subsurface domain was fully saturated from the bottom to the surface as cells where groundwater reached the surface. This binary information was transformed into a frequency map analogous to the procedure for creating the surface saturation frequency maps, using the same output times.

4 Results

4.1 Simulation of discharge, groundwater level and soil moisture

The model reproduced the seasonal dynamics of measured discharge very well (Fig. 3, Fig. S1). The best fit was obtained at the outlet (SW1) with a KGE of 0.74. Discharge at SW2, SW3, and SW4 was reproduced ~~equally~~ well with KGEs of 0.49, 0.48, and 0.47. Groundwater levels were captured well with the model at the locations close to the riparian zone (KGE=0.57, r=0.78 for GW2; KGE=0.64, r=0.84 for GW3). At hillslopes and plateau, simulated groundwater levels were similar to the observed levels during the wet season, but during dry conditions the groundwater levels did not fall deep enough (Fig. 3, Fig. S1). This ~~level~~ discrepancy was reflected in low KGEs (0.30 for GW1, 0.21 for GW5, 0.02 for GW7). However, general dynamics of level increasing and decreasing were also captured at hillslopes and plateau (r = 0.66 for GW5, r = 0.62 for GW7, and r = 0.76 for GW1; note that the value for GW1 only includes data for wet periods, since the piezometer fell dry during summer months).

Simulated soil moisture generally showed a transition from higher to lower responsiveness from topsoil to subsoil layers consistent with the monitored soil moisture (Fig. 3, Fig. S1) ~~and~~ Pearson correlation coefficients indicated overall a good agreement between simulated and observed soil moisture dynamics (Tab. 1). As for the groundwater levels at the hillslopes and plateau, soil moisture observations showed a distinct decrease in water content during dry periods, which the simulation could not reproduce to the same extent. The observed water content in the riparian zone was always close to saturation (TSM4, Fig. 3), while the simulation showed a decrease in water content during dry periods ~~in the riparian zone~~. Yet the simulation also showed a spatial trend for more permanent soil saturation in the riparian zone (TSM4) and its vicinity (TSM3, Fig. S1)

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

than at the hillslopes and plateau. The simulated values of water content were similar to the observed values at some locations (e.g. TSM2, SM4, Fig. 3) and clearly differed at other locations (e.g. SM7, Fig. 3), but the matches and mismatches of the volumetric water content did not clearly depend on specific areas or landscape units. Moreover, we think that moisture dynamics and responsiveness are more informative for model performance than the absolute water content values, since also the measured values of volumetric water content differed from each other within small distances (e.g. measurements of water content in 10 cm depth at profile SM7, Fig. 3).

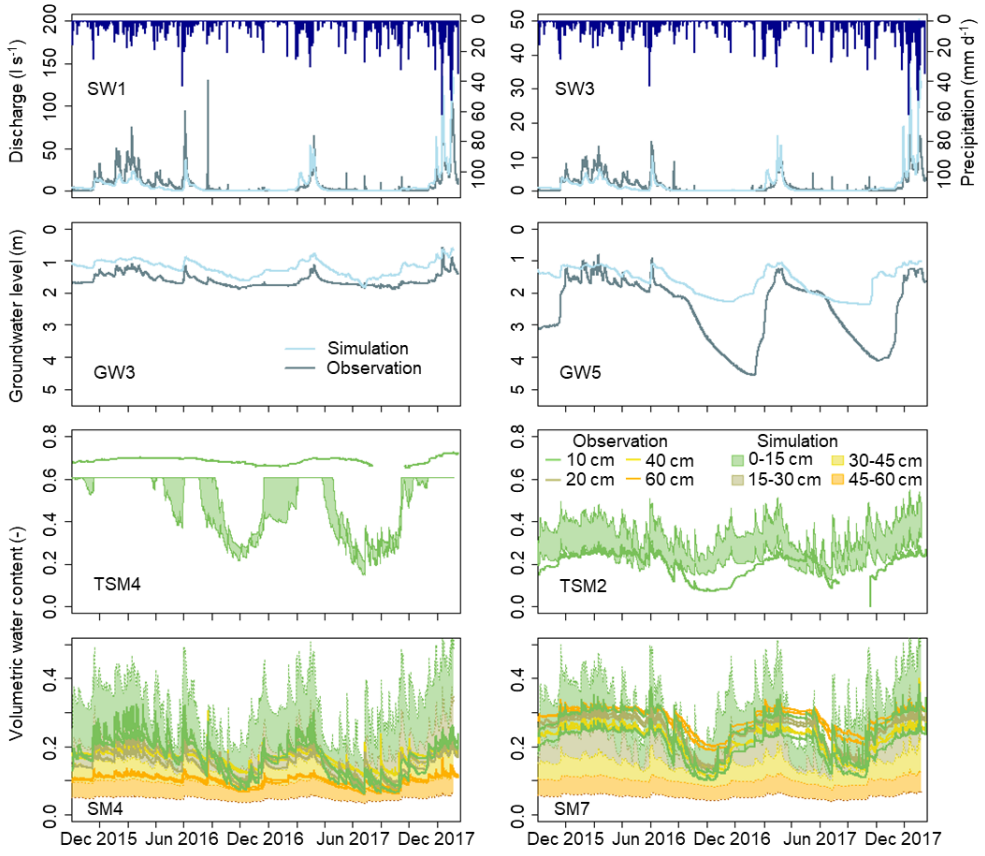

Figure 3: Simulated and observed time series of discharge, groundwater level below the surface, and volumetric water content. Colour bands indicate the possible span of simulated volumetric water contents in the depths between two model nodes. The time series of the observation locations (cf. Figure 1) that are not shown here, are shown in the supplemental material (Figure S1).

Table 2: Coefficients of Pearson correlation between simulated and observed volumetric water content of the soil for the different measurement locations and depths (cf. Fig. 1).

	SM3	SM4	SM5	SM7	TSM1	TSM2	TSM3	TSM4	TSM5
10cm	0.54	0.75	0.70	0.59	0.60	0.62	0.67	0.30	0.85
20cm	0.67	0.82	0.76	0.62					
40cm	0.82	0.89	0.88	0.79					
60cm	0.85	0.92	0.91	0.82					

© Author(s) 2019. CC BY 4.0 License.

4.2. Dynamics of surface saturation

The observed dynamics of normalized surface saturation (Fig. 4, coloured lines) were similar for all seven investigated riparian areas and followed the seasonal trend of the ~~catchment~~ discharge. Yet some differences between the studied areas were discernible. For example, saturation was less persistent between February and April 2016 in the two areas without perennial springs (M1, R2, Fig. 4) than in the other areas. Maximum saturation was reached in December 2017 at M1, R2 and S2, but between February and April 2016 at the other locations (Fig. 4). ~~As for~~ the observations, the simulated dynamics of normalized surface saturation (Fig 4, black lines) followed the general trend of the simulated discharge dynamic. The simulation showed a faster decrease and increase of the normalized saturation during dry periods than ~~it~~ was observed in most areas. However, simulated discharge also seemed to decrease and increase earlier than it was observed (c.f. Section 4.3). The simulated saturation dynamics did not clearly differ between the different locations and thus behaved more synchronous than the observations (e.g. maximum simulated saturation in December 2017 in all areas). As a result, the match between simulated and observed dynamics of normalized saturation was better for some areas (e.g. M1, R2, Fig. 4) than for others (e.g. S2, L1, Fig. 4).

The dynamic changes of normalized simulated saturation matched the normalized observations generally well, despite ~~of~~ under- and over-estimated ~~amounts~~ of minimum and maximum absolute saturation for all areas. The minimum ~~amount~~ of saturated pixels in the TIR panoramas ranged between 0.02 % at M3 and R3 and 3.38 % at S2, while the model did not simulate any surface saturation during the driest period (Fig. 4). In addition, simulated normalized saturation stayed longer close to the minimum than the observed saturation for several areas (L1, S2, M1). These results show that the model simulated a stronger dry-out than observed in the Weierbach. At the same time, the simulation overestimated maximum saturation in the riparian zone (Fig. 4). The overestimation was not equally strong at the seven investigated areas and as a result, the distinction between areas ~~showing~~ higher or lower maximum saturation was not the same for observation and simulation (e.g. R3 showing one of the highest maximum saturation in the observation, but one of the lowest maximum saturation in the simulation compared to the other areas).

4.3 Discharge – surface saturation relationship

The Pearson correlation between normalized saturation and discharge at the outlet SW1 was > 0.65 for both the simulation and the observation in almost all riparian areas. L1 was the only exception with r_{obs} = 0.54 (Fig. 5). The simulated relationships between normalized saturation and discharge resembled the observed relationships in terms of value range and shape (Fig. 5), although the observation data ~~scattered~~ distinctly more than the simulation data. A power law relationship approximated the observed relationship between discharge and saturation for all seven areas, when data that were taken during rainfall or rising discharge were excluded (cf. Antonelli et al., 2019). For some areas, the simulation matched the trend lines of the observation data closely (e.g. L1, M2). For other areas, the visual fit of the model output to the observation data was less good (e.g. S2, R3), but still described a similar trend.

Despite the common shape of a power law function, the saturation – discharge relationships were slightly different between the different areas, both for observation and simulation data. For example, the power law functions fitted to the observations showed that saturation during high flow conditions (> 5 l s^{-1}) increased most strongly with discharge in the sources areas (especially M3 and R3). During low flow conditions (< 1 l s^{-1}), the source areas (L1, M3, R3) showed the lowest amount of normalized saturation and the least change relative to discharge compared to the other areas. In the simulated relationships, the increase in saturation for high discharge (> 5 l s^{-1}) was strongest for M3 and S2. The simulated relationship between discharge and surface saturation during low flow (< 1 l s^{-1}) was similar for all areas in terms of slope, but differed in the amount of normalized saturation, being highest for areas in the east stream ~~branch~~ (R2, R3), followed by the middle upstream ~~branch~~ (M1, M2, M3), and L1 and S2.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

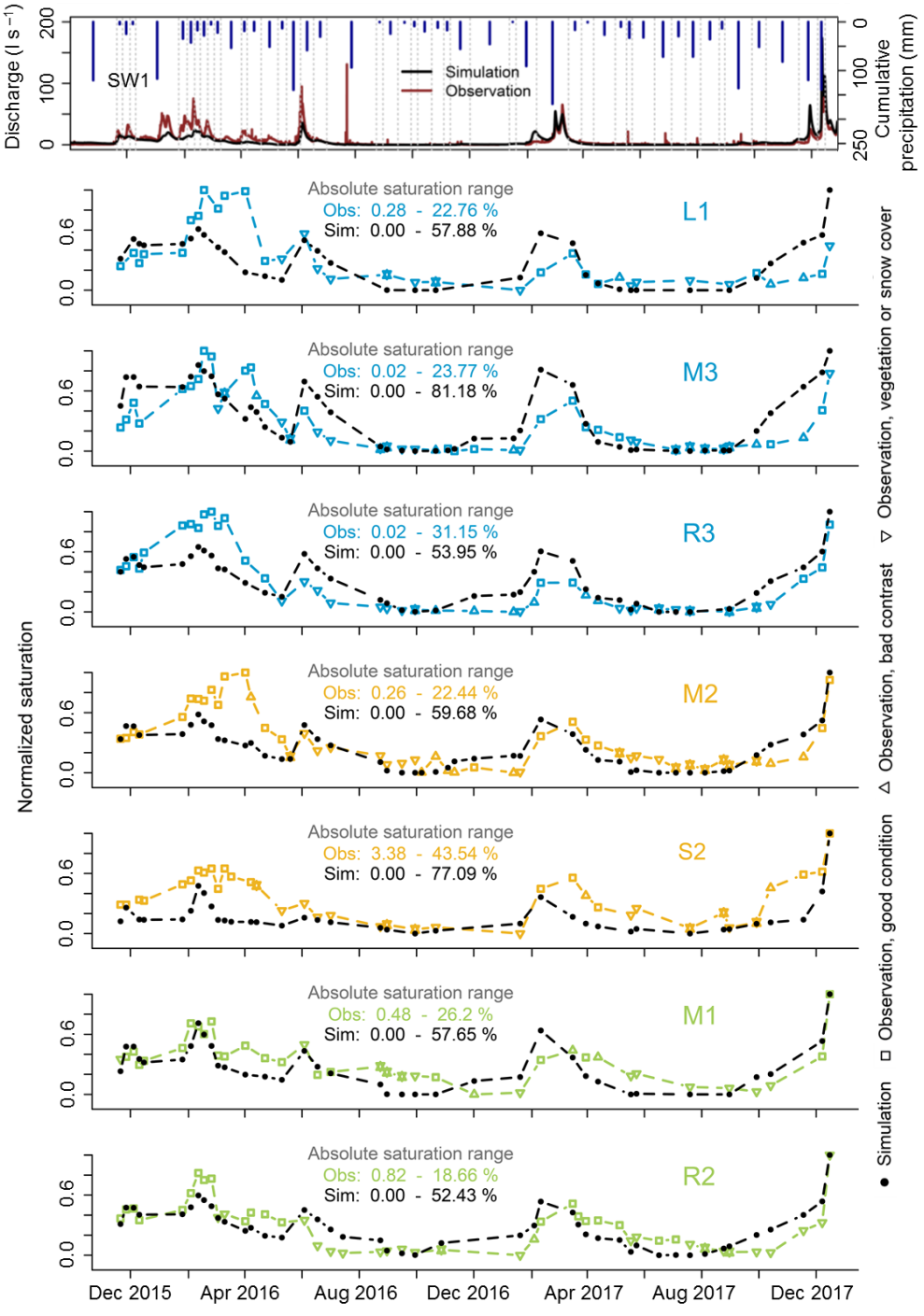

Figure 4: Time series of observed and simulated surface saturation in the seven investigated riparian areas. Surface saturation is normalized to the minimum and maximum amount of saturation that was observed and simulated in the individual areas, respectively. Observations that were derived from TIR images with a poor temperature contrast or with influences of vegetation and snow cover are deemed less reliable. Cumulative precipitation between the measurement dates (grey dashed lines) and discharge at catchment outlet SW1 are shown in the top panel for facilitating the comparison to precipitation and flow conditions.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

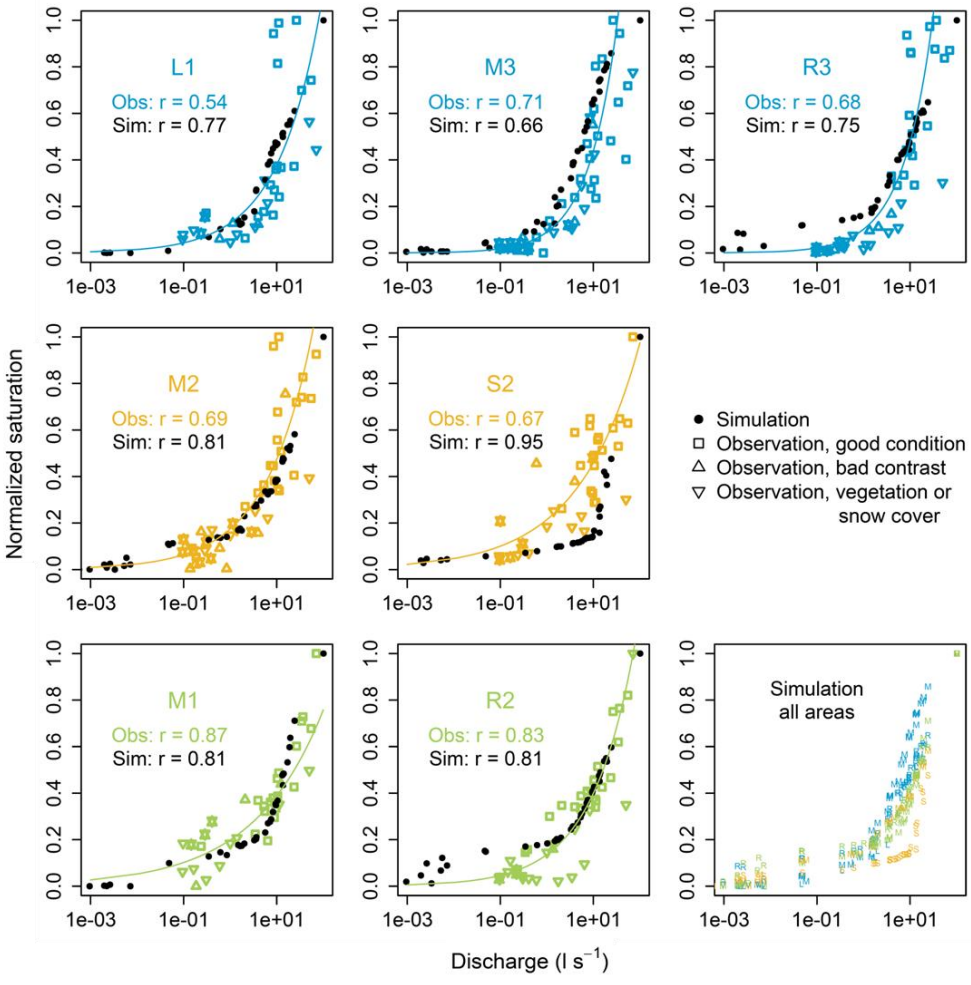

Figure 5: Observed and simulated relationships and Pearson correlations between normalized surface saturation and discharge at the catchment outlet SW1 for the seven investigated riparian areas. Observations that were derived from TIR images with a poor temperature contrast or with influences of vegetation and snow cover are deemed less reliable. Solid lines are power law curves fitted to the observation data, excluding data taken during rainfall or rising discharge. ~~For~~ facilitating the comparison between the seven areas, the panel on the bottom r[illegible]ontains the simulated data points from all seven areas ~~and the area affiliation is indicated with the respective colour and letter.~~

4.4 Spatial patterns of surface saturation

The realism of simulated patterns of surface saturation was evaluated for each riparian area by visually comparing the surface saturation frequency maps obtained from the simulations and observations (Fig. 6). The model captured the location of the stream and the locations that intermittently became surface saturated well for most of the seven investigated areas. For example, both observation and simulation showed that only the right side of the stream became saturated in M1, that the riparian zone of the right streamside in M2 became saturated only in the upstream part, and that saturation mainly developed on the left streamside in R3, surrounding some permanently dry areas next to the stream (Fig. 6). The only area with a clear mismatch between observed and simulated patterns of surface saturation was area L1, where surface saturation was simulated on the opposite streamside and at a clearly wrong position along the stream (upstream vs downstream).

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

The simulated surface saturation also reflected the observed saturation frequencies well. The simulation reproduced the general picture of more frequent surface saturation in the streambed than at the streamsides, but - as for the saturation patterns - simulated and observed frequencies corresponded better in some areas (e.g. S2, Fig. 6) than in others (e.g. R3, Fig. 6). For example, the observed frequency of surface saturation in the streambed was generally lower in the source areas (L1, M3, R3) than in the mid- and downstream areas (M2, S2, M1, R2), while the simulated frequency of surface saturation in the streambed was more similar between the areas and particularly overestimated in L1 and R3.

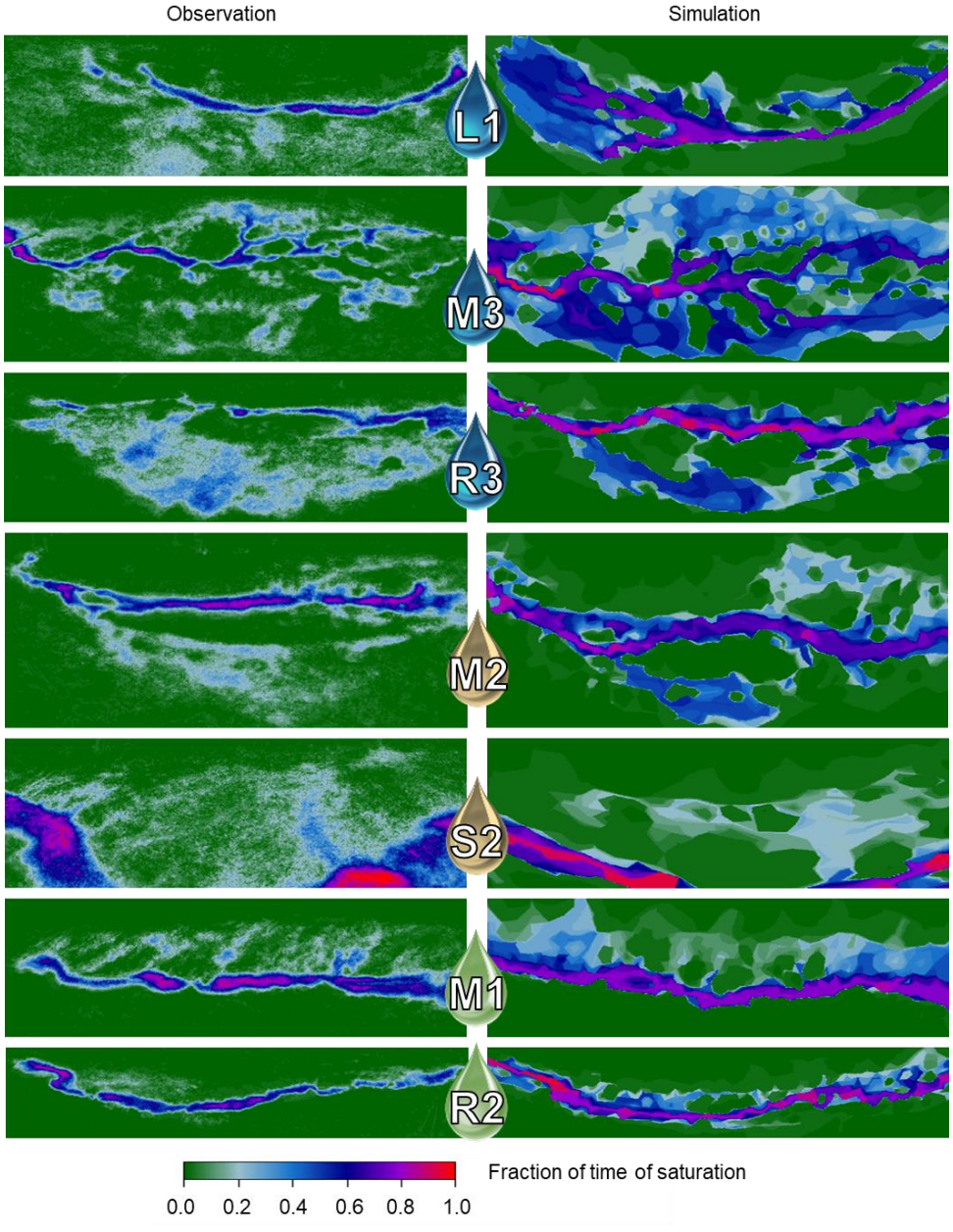

Figure 6: Observed (left) and simulated (right) frequencies of surface saturation in the seven investigated riparian areas. The maps were created by first counting how often the individual pixels were classified as saturated in the individual panoramic images and second normalizing the resulting frequency numbers by the total number of images analysed for the respective area.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

4.5 Simulated patterns and dynamics of surface saturation versus groundwater reaching the surface at catchment scale

Simulated surface saturation generally occurred only in the streambed and adjacent riparian zones (Fig. 7a). During the wettest conditions of the study period (winter 2017/2018), surface saturation also occurred as prolongation of the eastern stream branch into the hillslope above the source area R3. This simulated occurrence behaviour of surface saturation across the catchment is in accordance with field evidence, where we observed surface saturation outside of the valley bottom only during very wet conditions or rain on snow events (cf. Section 2.2). The simulated patterns of where and how frequently groundwater reached the ground surface (Fig. 7b) proved to be very similar to the surface saturation frequency map of the catchment (Fig. 7a). The only obvious difference occurred in the area above the source area of the eastern stream branch (R3), with a smaller extent of groundwater reaching the surface than ~~extent of~~ surface saturation.

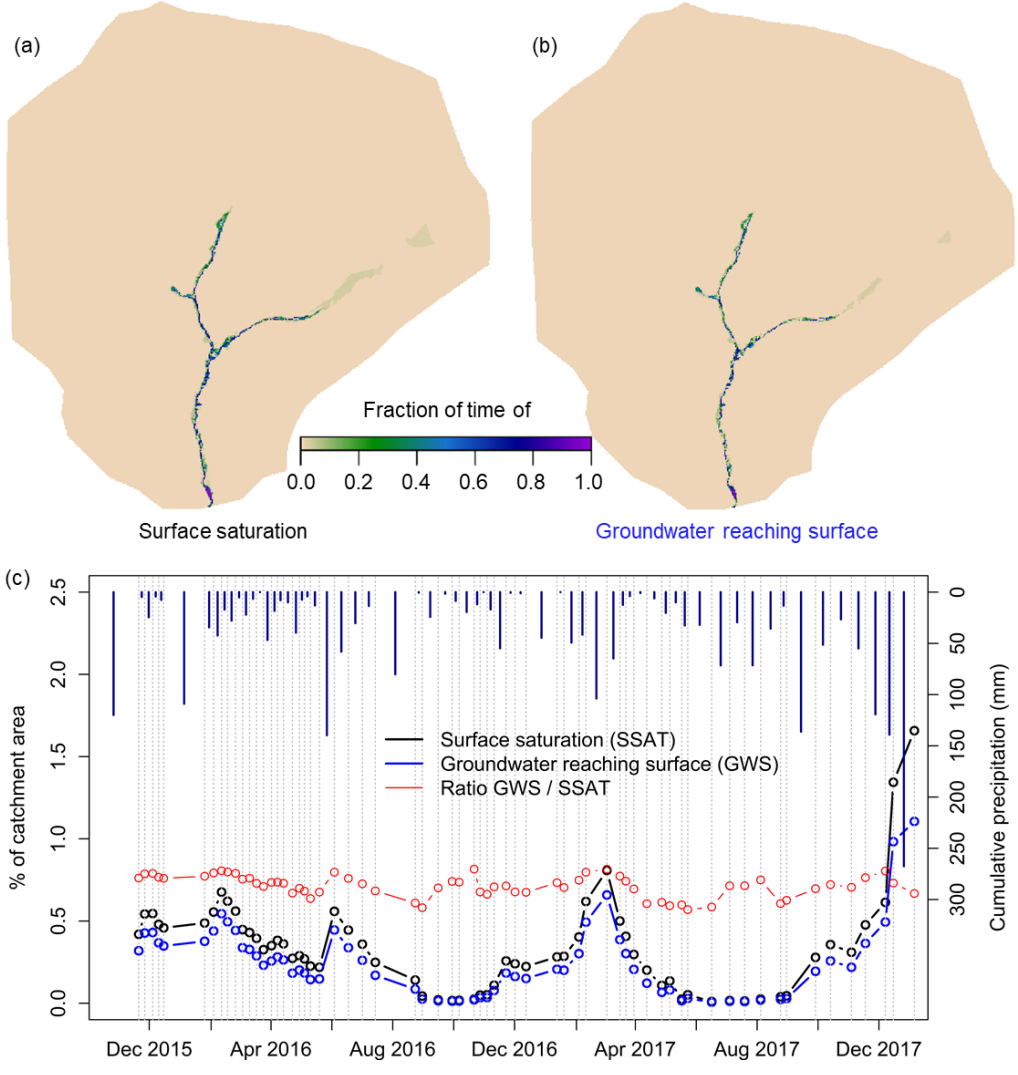

Figure 7: Simulated frequency maps (a, b) and time series of percentage (c) of surface saturation and groundwater reaching the surface in the Weierbach catchment. Precipitation is given as cumulative amounts between the observation dates (grey dashed lines).

© Author(s) 2019. CC BY 4.0 License.

The time series of simulated percentage of catchment area with surface saturation and groundwater reaching the surface revealed that the area where groundwater reached the surface was always smaller in extent than the surface saturated area, even after dry condition (Fig. 7c). The biggest absolute difference between the areal extent of surface saturation and groundwater reaching the surface was simulated during winter 2017/2018 (1.66 % vs 1.1 % of catchment area), where the conditions were very wet with high discharge and high cumulative precipitation and where the difference in areal extent was also visible in the frequency maps (Fig. 7a and b). However, the ratio between the extent of groundwater reaching the surface and the extent of surface saturation was not exceptionally high during winter 2017/2018. Instead, the ratio ~~scattered~~ without a clear trend between 0.57 and 0.82 during the entire simulation period, apparently independent from the cumulative amount of precipitation or surface saturation.

5 Discussion

The aim of this study was to analyse the spatio-temporal variability of surface saturation within the Weierbach catchment, with a focus on the stream valleys and riparian zones. Even though simulated discharge, groundwater levels and soil moisture ~~showed some discrepancies to~~ observations in terms of absolute values, we would argue that the performance of the different time series at different locations was quite good for a model that was not calibrated and set up ~~rather homogenously~~ across the catchment. While the model had some problems to reproduce soil moisture and groundwater levels during the dry conditions at hillslopes and plateau, the simulated time series matched the observations especially well in the riparian zone and vicinity. This gives us confidence that the model setup was valid for evaluating and analysing the spatio-temporal dynamics of surface saturation and its intra-catchment variability.

5.1 Temporal dynamics of surface saturation

The model reproduced the observed long-term dynamics of surface saturation over different seasons and wetness conditions well. Our study goes beyond previous works that compared the simulation of surface saturation dynamics with observations (e.g. Ali et al., 2014; Birkel et al., 2010; Glaser et al., 2016; Mengistu and Spence, 2016) by relying on a longer study period and a ~~higher~~ number of observations in time. This allowed us to analyse and compare various hydrological conditions and the dynamic transition between them over all seasons with a ~~frequent~~ number of observations. Moreover, we accounted for spatial variability of saturated area dynamics within the catchment. Unlike the various quasi dynamic wetness indices presented in Ali et al. (2014), which could not satisfyingly reproduce the spatio-temporal variability of connected surface saturation observed in a catchment in the Scottish Highlands, our model reproduced the distributed dynamics of surface saturation well, without clear ~~performance~~ differences for different wetness conditions.

Simulations and observations showed both that the temporal dynamics of surface saturation were mostly consistent across the catchment. Moreover, our simulations showed that the spatio-temporal development of surface saturation was very similar to the spatio-temporal dynamics of groundwater reaching the surface (cf. Fig. 7). This suggests that the generation of surface saturation in the Weierbach catchment is largely driven by the synchronous exfiltration of groundwater in topographic depressions. Antonelli et al. (2019) drew consistent conclusions based on a statistical analysis of the observation data.

5.2 Relation between surface saturation and discharge

We found that the observed and simulated relationships between surface saturation and discharge resembled power law relationships (cf. Fig. 5). This is consistent with earlier studies that showed power law relationships between contiguous connected surface saturated areas and discharge (Mengistu and Spence, 2016; Weill et al., 2013). In contrast to these studies, we did not observe hysteretic loops in the relationship between saturation and ~~high~~ streamflow. Nonetheless, the scatter in the observed discharge – surface saturation relationships might indicate that the development of surface saturation in the

© Author(s) 2019. CC BY 4.0 License.

Weierbach catchment follows hysteretic loops, but that the hysteresis was not resolved with the available temporal resolution of the observations. For example, it is likely that surface saturation evolved in the riparian areas during high flow conditions and persisted on the ground surface during decreasing streamflow due to restricted infiltration capacities of the riparian soil (cf. Antonelli et al., 2019).

The lack of such a hysteretic process in the simulation could explain why the model showed the tendency for less persistent and faster contracting surface saturation. It may also explain why the simulated saturation dynamics differed less between the different investigated areas than the observed dynamics. It is likely that the observed saturation dynamics were not synchronous between the different areas due to a less persistent (and thus hysteretic) generation of surface saturation in the relatively narrow riparian areas without perennial springs (M1 and R2) compared to the wider riparian areas with perennial springs (cf. observation of less persistent saturation in M1 and R2 during February and April 2016, Fig. 4). The model, instead, simulated a non-hysteretic saturation behaviour for all investigated riparian areas, which resulted in a better fit between simulated and observed dynamics in the areas M1 and R2 compared to the other areas.

At the same time, it might also be that the simulated relationship between saturation and discharge was correct in all riparian areas and that the scattering of the observation data did not result from hysteretic behaviour, but from uncertainties in the TIR methodology. A good argument for a correct simulation of the discharge – surface saturation relationship is that not only simulated saturation but also simulated discharge seemed to be less persistent and to decrease and increase earlier than it was observed. In reality, the scatter of the observation data is likely related to both measurement uncertainties and hysteretic aspects and a future study with higher temporal resolution of field observations and corresponding simulation output could further analyse this.

Independently from the question on hysteretic loops, we found that the discharge – surface saturation relationships somewhat differed between the different areas. We could connect the main differences to different topographical and morphological features, yet we cannot decipher why the main controlling feature for the discharge – surface saturation relationship was different between observations (source areas vs non-source areas) and simulations (different stream branches, cf. Section 4.3). Nonetheless, our findings are in line with experimental studies that discussed that the relationships between baseflow discharge and total extent of contributing saturated areas differ between catchments with different physiographic characteristics (e.g. Dunne et al., 1975; Latron and Gallart, 2007).

By comparing our model results to the double logarithmic plot presented by Latron and Gallart (2007) (Figure 8), we could identify similar shape varieties of the discharge – surface saturation relationship for the different areas studied within the Weierbach catchment as observed for the different catchments presented in Latron and Gallart (2007). We cannot compare our results directly with the results shown in Latron and Gallart (2007), since we evaluated absolute discharge and normalized saturation, while they evaluated connected saturated areas in percentage of catchment area, but discharge normalized to the catchment area. In order to facilitate the comparison and to connect the two plots (Fig. 8a, 8b), we show the simulated relationship between discharge and surface saturation of the entire Weierbach catchment in both plots, once with normalized discharge and absolute saturation (Fig. 8a), and once with absolute discharge and normalized saturation (Fig. 8b). The shape of the relationship for the entire Weierbach catchment was nearly linear, similar to the relationship observed in the Can Vila catchment investigated by Latron and Gallart (2007) (Fig. 8a). The relationships of the seven studied riparian areas differed from the catchment relationship and between each other (Fig. 8b). For example, area S2 and M1 showed a convex shape similar to the observations in the Vermont W2 catchment made by Dunne et al. (1975), area M3 showed a rather concave shape similar to the relationships found for a sub-catchment of the Asker basin (Myrabø, 1986) and the Strengbach catchment (Latron, 1990), area M2 showed a rather linear shape similar to the Can Vila catchment studied by Latron and Gallart (2007). This clearly shows that differences in the relationship between surface saturation and discharge do not only occur between different catchments, but that they also occur as intra-catchment variability.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-203
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 17 May 2019
© Author(s) 2019. CC BY 4.0 License.

Hydrology and Earth System Sciences
Discussions

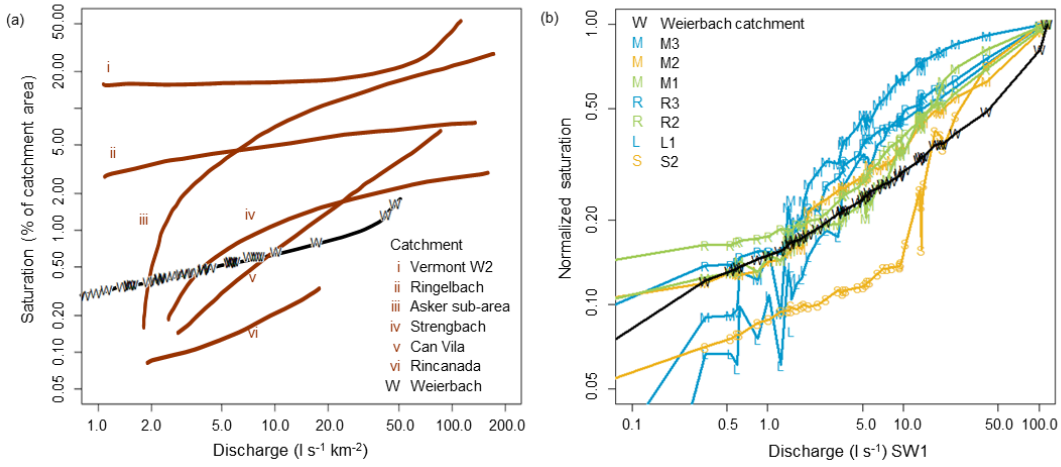

Figure 8: Simulated relationship between discharge and surface saturation of the entire Weierbach catchment (marked with W) in comparison to (a) the relationships observed in other catchments (Figure modified from Latron and Gallart (2007) and (b) the relationships simulated for the seven investigated riparian areas within the catchment. The presented relationships of the other catchments were investigated by i) Dunne et al. (1975), ii) Ambroise (1986), iii) Myrabø (Myrabø, 1986), iv) Latron (Latron, 1990), v) Latron and Gallart (2007), and vi) Martinez-Fernandez et al. (2005). Area affiliation for the investigated riparian areas of the Weierbach catchment is indicated with the respective colour and letter (cf. Fig. 4-6).

5.3 Spatial patterns of surface saturation

The observed spatial patterns of surface saturation were reproduced ~~with~~ the simulations ~~in great detail~~ for most of the investigated areas. We attribute the successful simulation of the spatial patterns to microtopography (local topographical features with extents of centimetres to few metres) since i) microtopography described the main spatial variability between the seven investigated areas in the model setup and ii) we observed that small changes in the setup and resolution of the model mesh in the riparian zones changed some details of the simulated surface saturation patterns (Fig. S2, especially area M2, S2). Therefore, we would like to stress that not only major topographic features of the catchment (e.g. hillslope shape, slope angle, valley width) but also its microtopography needs to be considered for identifying locations where surface saturation may occur. This may sound trivial and several studies have already pointed out the importance of microtopography for the simulation of different hydrological aspects such as hydraulic heads, hyporheic surface-subsurface water exchange, bank storage and overbank flooding, water quality of shallow groundwater systems and runoff generation (e.g. Aleina et al., 2015; Frei et al., 2010; Käser et al., 2014; Van der Ploeg et al., 2012; Tang et al., 2018). Still, microtopography is not often considered in the simulation of surface saturation patterns.

When microtopography is not resolved ~~detailed~~ enough, it is more likely that the simulated surface water extends over a large area instead of ~~accumulating in~~ topographic depressions and thus ~~overrates~~ the extent of surface saturation. In this context it is interesting to note that there are studies that simulated maximum extents of surface saturation up to 80 % of the study area (Qu and Duffy, 2007; Weill et al., 2013), while field observations have only ~~reached~~ maximum extents up to 25 % - 50 % of catchment area (Ali et al., 2014; Birkel et al., 2010; Dunne et al., 1975; Mengistu and Spence, 2016) and often ~~show~~ maximum extents around 10 % (Ambroise, 2016; Grabs et al., 2009; Güntner et al., 2004; Latron and Gallart, 2007; Tanaka et al., 1988). Microtopography might partly explain this discrepancy, even though the maximum extent of surface saturation certainly also depends on the climatic and physiographic conditions of the catchment and on the timing of the observations (e.g. baseflow conditions vs storm events) and there are some studies that simulated similar or ~~less~~ maximum extent of surface saturation

© Author(s) 2019. CC BY 4.0 License.

than observed without considering the microtopography (e.g. Ali et al., 2014; Birkel et al., 2010; Grabs et al., 2009; Güntner et al., 2004; Mengistu and Spence, 2016).

In our study, the simulated maximum extent of surface saturation was 1.6 % of catchment area, which is small compared to other simulation studies, but matches the observation that surface saturation commonly only occurs within the riparian zone and streambed (extent of 1.2 %). Nonetheless, also our maximum saturation within the individual areas was overestimated compared to the observations (cf. Fig. 4). Besides the effect of microtopography, there are two other possible explanations for this. First, the largest simulated saturation ~~occurred~~ during winter 2017/2018, which is the same period where the model clearly overestimated discharge. This mismatch could partly explain the overestimation of saturation, assuming that the relationship between discharge and saturation was correctly captured with the model (cf. Section 5.2). Second, the overestimation of absolute saturation could result from different perspectives and extensions of model output and TIR images (cf. section 3.2, Fig. 6). The TIR images included parts of the hillslopes around the riparian zones, which were not included to the same extent in the extracted model images. Since the hillslopes normally remained unsaturated, the maximum possible ~~amount~~ of saturated pixels in the TIR images was thus lower than in the model images, while the minimum possible ~~amount~~ of saturation was not affected. This could also explain why overestimation of total amounts of saturation was different between the different areas.

Despite the importance of microtopography, the model results showed that microtopography alone was not sufficient to capture the spatial patterns of surface saturation correctly. The simulated patterns of surface saturation clearly did not match the observed patterns equally well in all seven investigated areas (cf. Fig. 6), although the topographical information source and mesh resolution was consistent for the simulated riparian areas. This means that there are additional factors that control the spatial patterns of surface saturation that were not accounted for in the simulations. Such a factor could for example be the structure of the subsurface, which was treated as being homogeneous between all investigated riparian areas in the simulations. In reality, the subsurface structure may locally differ to some degree, for example in the riparian area of the western stream ~~branch~~ (L1), where saturation was simulated at ~~a~~ clearly wrong side along the stream.

5.4 Frequency maps of surface saturation

The frequency maps of surface saturation combine information on when and where surface saturation occurs. We do not think that the exfiltration of subsurface water into local depressions (cf. Section 5.1 and 5.2) can fully explain the spatial variability of saturation frequencies that was observed and simulated satisfactorily within the different riparian areas (Fig. 6). Instead, we ~~suppose~~ that the differences in saturation frequency were controlled by additional water sources than exfiltrating groundwater, such as stream water or direct precipitation, and that the contribution of these additional water sources to surface saturation varied in space and time. For example, the lower frequencies of surface saturation ~~observations~~ at the streamsides compared to the streambed and the lower frequencies in the streambed of the source areas (L1, M3, R3) compared to the mid- and downstream areas (M2, S2, M1, R2) might reflect a lower and less frequent contribution of upstream water in these areas. The overestimation of simulated saturation frequencies in the streambed of R3 could thus indicate an overestimated upstream contribution due to simulating the stream extent too far upstream from the source area. Future work should analyse potential water sources and generation processes of surface saturation with a suitable model framework (cf. Partington et al., 2013; Weill et al., 2013) in order to complement the interpretation of the observation data and to identify the mixture of different water sources of surface saturation (e.g. stream water, exfiltrating subsurface water, ponding precipitation), how the sources might vary in space and time, and how this might reflect in the surface saturation frequencies.

6 Summary and conclusions

We explored the intra-catchment variability of surface saturation in the Weierbach catchment with joint observations and simulations. We showed that the model could reproduce the observed variability of the surface saturation characteristics

© Author(s) 2019. CC BY 4.0 License.

(dynamics, frequencies, patterns) with great detail, although the model setup was rather homogeneous and parameters were not calibrated at catchment scale. Our results demonstrated that a spatially distributed, physically-based, integrated hydrological model such as HGS is well-suited for reproducing and analysing the generation and development of surface saturation in space and time.

Based on the matches and mismatches between the simulation results and observations, we could identify some key factors controlling the surface saturation generation. The temporal occurrence of surface saturation was observed and simulated to be similar across the catchment, which we related – based on the simulation results – to a large influence of groundwater that reacts synchronous across the catchment. The spatial occurrence of the surface saturation differed between and within the seven investigated riparian areas, which we mainly could relate to the influence of microtopography. Furthermore, we discussed that the full variability between the different areas and the mismatches between observations and simulation can only be explained with additional factors besides groundwater exfiltration and microtopography.

The spatially varying frequencies of surface saturation within the riparian areas indicated that there might be additional water sources than subsurface water that contribute to the generation of surface saturation. Since the model could reproduce the observed frequencies, the model can be used in a future study to analyse such a potential mixing of different water sources and their variation in space and time. The observed differences between the investigated riparian areas with regard to the seasonal dynamics of saturation extension and contraction and the surface saturation – discharge relationship likely resulted from different morphological characteristics (width, existence of perennial springs) of the riparian areas. Although the model could not reproduce a varying hysteretic occurrence and persistence of surface saturation in the different investigated areas, also the simulation results demonstrated that the relationship between surface saturation and discharge can differ within a catchment in the same manner as between catchments with different topographical and morphological conditions.

Data availability. Data underlying the study are property of the Luxembourg Institute of Science and Technology. They are available on request from the authors.

Author contributions. BG, LH and JK designed and directed the study. BG and MA planned and carried out the field work and processed the TIR images. BG set up the simulation and processed the model output. BG, MA, LH and JK discussed and interpreted the results. BG prepared the manuscript with contributions from JK and LH.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. We wish to thank Jean-Francois Iffly, Jérôme Juilleret, the Observatory for Climate and Environment of LIST, and the Administration des Services Techniques de l'Agriculture (ASTA) for the collection and provision of the hydrometrical and meteorological data. We acknowledge deployment of a trial version of AlgoMesh by HydroAlgorithmics Pty Ltd. Barbara Glaser thanks the Luxembourg National Research Fund (FNR) for funding within the framework of the FNR-AFR Pathfinder project (ID 10189601). Marta Antonelli was funded by the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no. 607150 (FP7-PEOPLE-2013-ITN – INTERFACES – Ecohydrological interfaces as critical hotspots for transformation of ecosystem exchange fluxes and bio-geochemical cycling).

References

Ala-aho, P., Rossi, P. M., Isokangas, E. and Kløve, B.: Fully integrated surface–subsurface flow modelling of groundwater–lake interaction in an esker aquifer: Model verification with stable isotopes and airborne thermal imaging, J. Hydrol., 522, 391–406, doi:10.1016/j.jhydrol.2014.12.054, 2015.

© Author(s) 2019. CC BY 4.0 License.

Aleina, F. C., Runkle, B. R. K., Kleinen, T., Kutzbach, L., Schneider, J. and Brovkin, V.: Modeling micro-topographic controls on boreal peatland hydrology and methane fluxes, Bio, 12, 5689–5704, doi:10.5194/bg-12-5689-2015, 2015.

Ali, G., Birkel, C., Tetzlaff, D., Soulsby, C., Mcdonnell, J. J. and Tarolli, P.: A comparison of wetness indices for the prediction of observed connected saturated areas under contrasting conditions, Earth Surf. Process. Landforms, 39(3), 399–413, doi:10.1002/esp.3506, 2014.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration (guidelines for computing crop water requirements), FAO Irrig. Drain. Pap., 56, 1998.

Ambroise, B.: Rôle hydrologique des surfaces saturées en eau dans le basin du Ringelbach à Soultzeren (Hautes-Vosges), France, in Recherches sur l'Environnement dans la Région, Actes du 1er Colloque Scientifique des Universités du Rhîn Supérieur, edited by O. Rentz, J. Streith, and L. Ziliox, pp. 620–630, Université Louis Pasteur - Conseil de l'Europe, Strasbourg., 1986.

Ambroise, B.: Variable water-saturated areas and stream fl ow generation in the small Ringelbach catchment (Vosges Mountains , France): the master recession curve as an equilibrium curve for interactions between atmosphere , surface and ground waters, Hydrol. Process., 30, 3560–3577, doi:10.1002/hyp.10947, 2016.

Antonelli, M., Glaser, B., Teuling, R., Klaus, J. and Pfister, L.: Saturated areas through the lens: 1. Spatio-temporal variability of surface saturation documented through Thermal Infrared imagery, In preparation, 2019.

Birkel, C., Tetzlaff, D., Dunn, S. M. and Soulsby, C.: Towards a simple dynamic process conceptualization in rainfall – runoff models using multi-criteria calibration and tracers in temperate, upland catchments, Hydrol. Process., 24, 260–275, doi:10.1002/hyp.7478, 2010.

Carrer, G. E., Klaus, J. and Pfister, L.: Assessing the Catchment Storage Function Through a Dual-Storage Concept, Water Resour. Res., 55, 476–494, doi:10.1029/2018WR022856, 2019.

Davison, J. H., Hwang, H.-T., Sudicky, E. A., Mallia, D. V. and Lin, J. C.: Full Coupling Between the Atmosphere, Surface, and Subsurface for Integrated Hydrologic Simulation, J. Adv. Model. Earth Syst., 10, 43–53, doi:10.1002/2017MS001052, 2018.

Dunne, T., Moore, T. R. and Taylor, C. H.: Recognition and prediction of runoff-producing zones in humid regions, Hydrol. Sci. Bull., 20, 305–327, 1975.

Erler, A. R., Frey, S. K., Khader, O., Orgeville, M., Park, Y.-J., Hwang, H.-T., Lapen, D. R., Peltier, W. R. and Sudicky, E. A.: Simulating Climate Change Impacts on Surface Water Resources Within a Lake-Affected Region Using Regional Climate Projections, Water Resour. Res., 55, 130–155, doi:10.1029/2018WR024381, 2019.

Frei, S., Lischeid, G. and Fleckenstein, J. H.: Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland --- A modeling study, Adv. Water Resour., 33(11), 1388–1401, doi:10.1016/j.advwatres.2010.07.006, 2010.

Gburek, W. J. and Sharpley, A. N.: Hydrologic Controls on Phosphorus Loss from Upland Agricultural Watersheds, J. Environ. Qual., 27, 267–277, doi:10.2134/jeq1998.00472425002700020005x, 1998.

Glaser, B., Klaus, J., Frei, S., Frentress, J., Pfister, L. and Hopp, L.: On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum, Water Resour. Res., 52, 8317–8342, doi:10.1002/2015WR018414, 2016.

Glaser, B., Antonelli, M., Chini, M., Pfister, L. and Klaus, J.: Technical note: Mapping surface-saturation dynamics with thermal infrared imagery, Hydrol. Earth Syst. Sci., 22(11), 5987–6003, doi:10.5194/hess-22-5987-2018, 2018.

Glaser, B., Jackisch, C., Hopp, L. and Klaus, J.: How meaningful are plot-scale observations and simulations of preferential flow for catchment models ?, Vadose Zo. J., doi:10.2136/vzj2018.08.0146, 2019.

Gourdol, L., Clément, R., Juilleret, J., Pfister, L. and Hissler, C.: Large-scale ERT surveys for investigating shallow regolith properties and architecture, Hydrol. Earth Syst. Sci. Discuss., 1–39, doi:10.5194/hess-2018-519, 2018.

© Author(s) 2019. CC BY 4.0 License.

Grabs, T., Seibert, J., Bishop, K. and Laudon, H.: Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model, J. Hydrol., 373(1-2), 15–23, doi:10.1016/j.jhydrol.2009.03.031, 2009.

Güntner, A., Seibert, J. and Uhlenbrook, S.: Modeling spatial patterns of saturated areas: An evaluation of different terrain indices, Water Resour. Res., 40, W05114, doi:10.1029/2003WR002864, 2004.

Gupta, H. V, Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Juilleret, J., Iffly, J. F., Pfister, L. and Hissler, C.: Remarkable Pleistocene periglacial slope deposits in Luxembourg (Oesling): pedological implication and geosite potential, Bull. la Société des Nat. Luxemb., 112(1), 125–130 [online] Available from: http://www.snl.lu/publications/bulletin/SNL_2011_112_125_130.pdf, 2011.

Käser, D., Graf, T., Cochand, F., McLaren, R., Therrien, R. and Brunner, P.: Channel Representation in Physically Based Models Coupling Groundwater and Surface Water : Pitfalls and How to Avoid Them, Groundwater, 52(6), 827–836, doi:10.1111/gwat.12143, 2014.

Kristensen, K. J. and Jensen, S. E.: A model for estimating actual evapotranspiration from potential evapotranspiration, Nord. Hydrol., 6(3), 170–188, doi:10.2166/nh.1975.012, 1975.

Latron, J.: Caractérisation géomorphologique et hydrologique du bassin versant versant du Strengbach (Aubure), Université Louis Pasteur, Strasbourg I., 1990.

Latron, J. and Gallart, F.: Seasonal dynamics of runoff-contributing areas in a small mediterranean research catchment (Vallcebre, Eastern Pyrenees), J. Hydrol., 335(1-2), 194–206, doi:10.1016/j.jhydrol.2006.11.012, 2007.

Martínez Fernández, J., Ceballos Barbancho, A., Morán Tejeda, C., Casado Ledesma, S. and Hernández Santana, V.: Procesos hidrológicos en una cuenca forestal del Sistema Central: Cuenca experimental de Rinconada, Cuad. Investig. Geográfica, 31, 7–25 [online] Available from: https://dialnet.unirioja.es/servlet/articulo?codigo=1975892, 2005.

Martínez-Carreras, N., Hissler, C., Gourdol, L., Klaus, J., Juilleret, J., Iffly, J. F. and Pfister, L.: Storage controls on the generation of double peak hydrographs in a forested headwater catchment, J. Hydrol., 543, 255–269, doi:10.1016/j.jhydrol.2016.10.004, 2016.

Megahan, W. F. and King, P. N.: Identification of critical areas on forest lands for control of nonpoint sources of pollution, Environ. Manage., 9(1), 7–17, doi:10.1007/BF01871440, 1985.

Mengistu, S. G. and Spence, C.: Testing the ability of a semidistributed hydrological model to simulate contributing area, Water Resour. Res., 52, 4399–4415, doi:10.1002/2016WR018760, 2016.

Moragues-Quiroga, C., Juilleret, J., Gourdol, L., Pelt, E., Perrone, T., Aubert, a., Morvan, G., Chabaux, F., Legout, a., Stille, P. and Hissler, C.: Genesis and evolution of regoliths: Evidence from trace and major elements and Sr-Nd-Pb-U isotopes, Catena, 149, 185–198, doi:10.1016/j.catena.2016.09.015, 2017.

Munz, M., Oswald, S. E. and Schmidt, C.: Coupled Long-Term Simulation of Reach-Scale Water and Heat Fluxes Across the River-Groundwater Interface for Retrieving Hyporheic Residence Times and Temperature Dynamics, Water Resour. Res., (53), 8900–8924, doi:10.1002/2017WR020667, 2017.

Myrabø, S.: Runoff Studies in a Small Catchment, Nord. Hydrol., 17, 335–346, doi:10.2166/nh.1986.0025, 1986.

Nasta, P., Boaga, J., Deiana, R., Cassiani, G. and Romano, N.: Comparing ERT- and scaling-based approaches to parameterize soil hydraulic properties for spatially distributed model applications, Adv. Water Resour., 126, 155–167, doi:10.1016/j.advwatres.2019.02.014, 2019.

Nippgen, F., McGlynn, B. L. and Emanuel, R. E.: Water Resources Research, Water Resour. Res., 51, 4550–4573, doi:10.1002/2014WR016719.Received, 2015.

Ogden, F. L. and Watts, B. a.: Saturated area formation on nonconvergent hillslope topography with shallow soils: A numerical investigation, Water Resour. Res., 36(7), 1795, doi:10.1029/2000WR900091, 2000.

© Author(s) 2019. CC BY 4.0 License.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., Dandy, G. C. and Fleckenstein, J. H.: Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, Water Resour. Res., 49(9), 5501–5519, doi:10.1002/wrcr.20405, 2013.

Pfister, L., McDonnell, J. J., Hissler, C. and Hoffmann, L.: Ground-based thermal imagery as a simple, practical tool for mapping saturated area connectivity and dynamics, Hydrol. Process., 24(21), 3123–3132, doi:10.1002/hyp.7840, 2010.

Van der Ploeg, M. J., Appels, W. M., Cirkel, D. G., Oosterwoud, M. R., Witte, J.-P. M. and Van der Zee, S. E. A. T. M.: Microtopography as a Driving Mechanism for Ecohydrological Processes in Shallow Groundwater Systems, Vadose Zo. J., 11(3), doi:10.2136/vzj2011.0098, 2012.

Qu, Y. and Duffy, C. J.: A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43(8), 1–18, doi:10.1029/2006WR005752, 2007.

Reaney, S. M., Bracken, L. J. and Kirkby, M. J.: The importance of surface controls on overland flow connectivity in semi-arid environments: Results from a numerical experimental approach, Hydrol. Process., 28(4), 2116–2128, doi:10.1002/hyp.9769, 2014.

Scaini, A., Audebert, M., Hissler, C., Fenicia, F., Gourdol, L., Pfister, L. and Beven, K. J.: Velocity and celerity dynamics at plot scale inferred from artificial tracing experiments and time-lapse ERT, J. Hydrol., 546, 28–43, doi:10.1016/j.jhydrol.2016.12.035, 2017.

Schilling, O. S., Gerber, C., Partington, D. J., Purtschert, R., Brennwald, M. S., Kipfer, R., Hunkeler, D. and Brunner, P.: Advancing Physically-Based Flow Simulations of Alluvial Systems Through Atmospheric Noble Gases and the Novel 37Ar Tracer Method, Water Resour. Res., 53, 10,465–10,490, doi:10.1002/2017WR020754, 2017.

Sebben, M. L., Werner, A. D., Liggett, J. E., Partington, D. and Simmons, C. T.: On the testing of fully integrated surface – subsurface hydrological models, Hydrol. Process., 27, 1276–1285, doi:10.1002/hyp.9630, 2013.

Silasari, R., Parajka, J., Ressl, C., Strauss, P. and Blöschl, G.: Potential of time-lapse photography for identifying saturation area dynamics on agricultural hillslopes, Hydrol. Process., 1–18, doi:10.1002/hyp.11272, 2017.

Tanaka, T., Yasuhara, M., Sakai, H. and Marui, A.: The Hachioji experimental basin study -- storm runoff processes and the mechanism of its generation, J. Hydrol., 102, 139–164, 1988.

Tang, Q., Schilling, O. S., Kurtz, W., Brunner, P., Vereecken, H. and Hendricks Franssen, H.-J.: Simulating Flood-Induced Riverbed Transience Using Unmanned Aerial Vehicles , Physically Based Hydrological Modeling, and the Ensemble Kalman Filter, Water Resour. Res., (54), 9342–9363, doi:10.1029/2018WR023067, 2018.

Weill, S., Altissimo, M., Cassiani, G., Deiana, R., Marani, M. and Putti, M.: Saturated area dynamics and streamflow generation from coupled surface-subsurface simulations and field observations, Adv. Water Resour., 59, 196–208, doi:10.1016/j.advwatres.2013.06.007, 2013.