Editor:

Dear Authors:
Your revised submission received some further comments and suggestions for improvements by two out of the three previous reviewers. Please, see if additional changes (as suggested by these reviewers) are feasible and upload a point-by-point reply. Should you disagree with some of these additional comments, please explain why clearly. If any changes are made, upload of course the new revised version of your paper.

Dear Nunzio Romano,

We revised the manuscript where necessary and feasible following the comments and suggestions of the two referees. In case we did not change the manuscript as suggested we explain our reasoning for this in the response to the referees. For some of the main comments mainly a clarification was needed (e.g. concerning the power law relationships, comment #1 of referee #2, the exfiltration into topographic depressions, comment #1 of referee #3, the reasonableness of daily steps, comment #2 of referee #3). In these cases we added a short clarification in the manuscript and replied in more detail to the referees.

We thank you for handling our manuscript and hope our revised manuscript is now ready for publication.

Kind regards,

Barbara Glaser
On behalf of all co-authors

Referee #2:

Dear Authors,
The paper manuscript presents physically base modeling results of saturation area dynamics and compares these results with field observations using Thermal Infrared Image collected for a large number of snapshots and derived maps of saturated areas of 7 stream sections in the Weierbach, Luxembourg. I appreciate that the authors have responded to each comment and incorporated/addressed the feedback. The paper manuscript has further improved and I only want to raise three points:

We thank the referee for the re-assessment of our manuscript and for the given suggestions for further improvement.

1) The authors talk about a power law relation between discharge and normalized saturation and I assume, this is dawn from Figure 5. However, I ask the authors to plot Fig 5 (not necessarily for the manuscript but for their interpretation) using linear x-axis for discharge. It is likely that the relation is not following a power law relation and the authors might want to reconsider their interpretation in the abstract, the results and conclusions.

We fitted the power law relationships to the original, non-logarithmic observation data (normalized surface saturation and catchment discharge, excluding data taken during rainfall or rising discharge). Therefore, our interpretations, and conclusions do not change with the scale of the axis.

Of course, by plotting the discharge in logarithmic scale the relationships appear different than by using a linear x-axis. The fact that the relationships look like following power law relations in both

cases is caused by the fitted power law exponents. The exponents ranged between 0.29 and 0.69 for the different areas, which results in concave shapes when plotting the data in linear x-y space. When plotting the data using a logarithmic x-axis, the relationship turn into convex shapes, appearing similar to power law relationships with exponents >1 plotted in linear space.

To clarify and avoid confusion, we added Figure S2 to the supplemental material that shows the data of Figure 6 plotted in the linear surface saturation – discharge space. Moreover, we added the information on the power law exponents in the manuscript and the caption of Figure 6. Finally, we carefully rechecked all our statements relating to the description of the relationship (especially section 4.4) with regard to their correctness for the linear scale.

2) In general, the paper manuscript has further improved during the first round of review. However, in the conclusion section some formulation is no longer clear the way the text has been edited (see my comments in the pdf).

We revised the conclusion following the edits and comments given in the commented PDF (see our responses there).

3) Besides that, I found some very minor things that the authors might want to change before final publishing (see comments in the pdf).

We changed the manuscript as suggested for most of the highlighted points. For the cases where we did not revise the manuscript as suggested, we responded to the respective comments in the PDF to explain our reasoning.

Referee #3:

The manuscript has been significantly improved. I only have five remaining comments, and include some suggestions to further improve the clarity of the manuscript and some minor comments and suggestions in the annotated pdf.

We thank the referee for the re-assessment of our manuscript and for the given comments and suggestions for further clarity

1. The wording regarding the exfiltration of groundwater in microtopographic depressions could be clearer. Is there really only exfiltration in these centimeter to meter scale depressions (and thus not the surrounding areas) or is there exfiltration in a larger area but only sufficient accumulation and ponding in these depressions for the surface to be considered saturated?

It is true that in theory water could exfiltrate from the subsurface over a large area and that the surface might not be considered saturated due to insufficient accumulation and ponding. However, there are two reasons why we are confident that this does not occur to a relevant extent in the Weierbach catchment based on our results:

1) 'We define the surface as saturated as soon as water is standing or flowing on the ground surface' (cf. P6, L2), i.e. even a thin film of water of more than 0.1 mm thickness was accounted for in the detection of surface saturation with the TIR imagery and simulation (cf. P10, L8-12). This excludes that accumulation and ponding of water on the surface in other areas than in the topographic depressions is insufficient to be detected with the TIR imagery and simulation.

2) We compared the simulated extent of surface saturation with the simulated extent of groundwater reaching the surface (cf. Fig. 4) in order to identify the 'importance of saturation excess and groundwater exfiltration in comparison to infiltration excess and overland flow' (cf. P10, L18). The comparison showed that the simulated extent of the surface saturation was always slightly larger than the simulated extent of groundwater reaching the surface for all observation dates (cf. Fig. 4), thus we exclude that exfiltration of groundwater occurred in much larger areas than the accumulation of water on the surface.

Nonetheless, it is true that the exfiltration of water and thus the generation of surface saturation was not exclusively observed and simulated within microtopographic depressions with extents of tens of centimeters to few meters, but also in topographic depressions of larger extent (e.g.in the area above source area R3, in the stream valley in general). We carefully rechecked our wording to be more precise on this circumstance (e.g. P18, L26-27, P22, L15-16).

2. Please comment on the reasonableness of the daily time step for precipitation (and ET) for analysing spatial patterns in surface saturation and how this may have affected your results and interpretations. I recall from other papers that the streamflow response in this catchment can be fairly quick. Does this reduce the maximum size of the saturated area and is this partly a reason for a lack of hysteresis?

Although we used daily precipitation and ET input data for forcing the model, the model runs with sub-daily time steps (minimum time step < 1 min) for periods with highly dynamic hydrologic conditions (e.g. rainfall-runoff events). We clarified this now in the methods section (P9, L17-19). Still, an even distribution of the daily input rates over sub-daily time steps results in different intensities than by using hourly input data. This can result in an underestimation of the rain intensities and consequently it might also result in an underestimation of the extent of saturated areas during an event. Furthermore, it is correct that streamflow in the Weierbach catchment is characterized by two peaks as response to precipitation: an immediate, flashy first peak and a broad, delayed second peak. Thus, we agree that using daily rates of precipitation and ET may not be sufficient to analyze the surface saturation during events.

However, the focus of our manuscript is on the seasonal and yearly dynamics of surface saturation and not on the event scale. We compared the effect of using hourly vs daily meteorological forcing data in a previous study and showed that the resulting difference in the simulated seasonal to yearly discharge dynamics in the Weierbach catchment is negligible (see Glaser et al. 2019). Thus, we are confident that the results and interpretations presented here would not change with hourly instead of daily resolved input data. Moreover, during our field work we observed that intra-event variations of surface saturation were rather small (TIR images with 15min resolution, data not used in this study) and highest extents of surface saturation coincide with periods of highest discharge (cf. Fig. 6), which are resolved with daily time steps.

Concerning hysteresis we agree that it is likely that an event scale analysis would show different results than the weekly to biweekly analysis (cf. our response to comment #3), still we think that our interpretation of the weekly to biweekly simulation data would not change significantly when using hourly instead of daily rates of precipitation and ET as input data. We added a sentence in the Introduction and one sentence in the discussion section relating to the hysteresis (cf. our response to comment #3) to better highlight the seasonal focus of our study.

3. There is a fair bit of discussion on the lack of hysteresis in the saturated area - discharge relationship. I agree that the authors show that there is no clear hysteresis on the seasonal scale.

However, there may be hysteresis on the event time scale. This is not clear for the observations because the photos weren't taken during events but could be analysed for the model results. I suggest that either the authors weaken their statements on a lack of hysteresis and mention more clearly that they focus on hysteresis at the seasonal (not event) time scale or that they actually look at the model results for individual events and check if there is hysteresis at the event time scale or not. It could be useful to add the simulation results for all time steps in a light shading in the background of figure 6 to show that the selected times are representative.

We fully agree that there might be hysteresis on event time scale. We mention that the lack of a clear hysteresis might result from the available resolution of the observation data (P19, L7-8 in the previous version of the manuscript). We also agree that the model results could be used to analyze the relationship at the event time scale. However, this would require a different model setup in terms of output and data management than what we did for this study. We focused on the comparison of the simulation results with the observation data and on the long-term seasonal to yearly time scale. Thus, we only processed and analyzed the simulation results for the dates where TIR images were taken. Analyzing individual events and adding the simulation results for all time steps in light shading in the background of Figure 6 is definitely of interest. Yet this would require quite some additional modelling and data processing work, which is far beyond the scope of the current study.

We now state explicitly in the discussion on the hysteresis in section 5.2 that our study focusses on the seasonal scale and that the model output might be used for more detailed analyses in the future. Moreover, we clarified why we see some indications for hysteresis in the observation data (i.e. scattering of the observation data) while the model does not show any indication for hysteresis on seasonal time scale (i.e. no scattering, more similar relationships between different areas than observed) and revised our wording to not to talk about 'hysteretic processes' anymore.

Please also see our responses in the commented PDF.

4. The abstract and conclusion could focus a bit more on what was learned from the modeling regarding the saturated area dynamics and its relation to discharge or the flow pathways that couldn't have been inferred from the field data alone.

The benefit of our approach to use the simulation as complement to and in combination with the field data is that we can 1) identify the match and mismatch between observations and simulations and 2) subsequently can infer from the model setup which features control the occurrence of surface saturation. With field data alone the identification of key controls on surface saturation would have remained speculative. With the help of the simulations, we could show clear indications for the importance of several controlling factors. For the aspects where the simulation reproduced the observations, we inferred the controlling factors from the features that were included in the model setup (i.e. microtopography, groundwater reaching exfiltration) and apparently were sufficient to reproduce the observed behavior. For the aspects where we observed a mismatch between observations and simulations, it became clear that some important controlling features were not included in the model setup and we discuss which features these might be (e.g. subsurface heterogeneity).

The benefit of analyzing the match and mismatch between observations and simulations is mentioned repeatedly throughout the manuscript. Also in the abstract and conclusion we highlight the identified matches and mismatches and what we inferred from them about the controlling factors (i.e. P1, L23-25, P23 L5-6, P23 L17-19 for matches; P1, L25-30, P23, L19-21 for mismatches). We replaced 'The match between simulation and observations suggested…' by 'The good

performance of the simulations suggested…' in the abstract and added an additional sentence in the conclusions to further highlight the relevant role of the simulations in our approach.

5. Overall, the discussion (particularly sections 5.1 and 5.4) could focus a bit more on the broader message and/or implications of this study for understanding runoff generation mechanisms in other catchments. Now these sections tend to be very specific for the modeling done in this study or interpretation for runoff mechanisms in this study catchment.

We agree that the discussion and especially section 5.1 and 5.4 are rather specific for the Weierbach catchment with regard to the discussed processes. This is due to the fact that we discuss which factors / processes we could infer from the simulation in combination with the observations. Since we only applied our approach for the Weierbach catchment, we can also only infer these factors / processes for the Weierbach catchment. Certainly, it is possible to discuss which implication the identified controls on surface saturation and the identified intra-catchment variability of surface saturation may have for investigations in other catchments. We do this in Section 5.2 and 5.3, where we compare our results with other studies on surface saturation. We added additional parts of discussion at the end of these two section following the suggestions given in the commented PDF (see also our reply there). Yet we think that it is rather difficult to discuss our results in a broader context for section 5.1 and 5.4. There is a lack of studies that have done similar analyses where we could relate to (some references are given in section 5.1). Thus we can only state that future studies might need to consider/analyze if our findings are also valid in other catchments, e.g. by using a similar approach to ours and combine observation and ISSHM simulations, and if we can generate transferable knowledge based on this. We now added such a statement at the end of section 5.1 and we rephrased the last sentence of section 5.4 to 'Future work should analyse potential water sources and surface saturation generation processes in the *Weierbach catchment and beyond*'.

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

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 (b.glaser@lmu.de)

Abstract

In this study, we explored the spatio-temporal variability of surface saturation within a forested headwater catchment ~~in~~using 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~~images 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~~ good performance of the simulation ~~and observations~~ suggested that surface saturation in the Weierbach catchment ~~was~~is 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 between observations and simulation 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~~persistence of surface saturation ~~and catchment discharge~~ due to ~~the~~ local ~~existence of~~morphological features like perennial springs.

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. In general, such surface saturated areas arise from 1) water ponding on the surface due to precipitation intensity exceeding 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 flowing into the floodplain (e.g. Megahan and King, 1985). When 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~~was 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 different water sources 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)~~(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 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 processes affects 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 is always ~~represents a simplification~~simplified representation of reality, the big asset of spatially distributed and dynamic hydrological models is that they ~~-~~allow detailed investigation of surface saturation at any desired location and time. This goes far beyond the information that can be gained by 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~~an integrated surface-subsurface hydrologic ~~models~~model ~~(ISSHMs, Sebben et al., 2013)~~(ISSHM, Sebben et al., 2013), which can simulate the interplay of different surface and subsurface processes ~~of~~leading to surface saturation ~~generation~~ (e.g. ponding of precipitation from the surface, exfiltration from the subsurface). However, 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 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 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 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 use 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 to identify 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)~~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)~~(2020) 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~~(2020) analysed the saturation dynamics in various riparian areas and related them to meteorological and hydrological conditions. They found similar seasonal extension and contraction dynamics of surface saturation ~~in~~between 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~~(2020) also identified some local differences of saturation dynamics related to 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 a combination of field observation and modelling. We perform the study in the Weierbach catchment~~,~~ (42 ha), where we can rely on existing TIR imagery data ~~(cf. Antonelli et al., 2019; Glaser et al., 2018)~~(cf. Antonelli et al., 2020; 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 headwater by accounting for ~~a~~ layering of the subsurface, while spatial heterogeneity was ~~only represented by microtopography~~simplified to topography, including microtopographical features with extents of tens of centimetres to few metres, and a different sequence of subsurface layers in the riparian zone compared to the hillslopes and plateau area. 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 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 us to study the potential occurrence of different hydrological processes within the different areas that was suggested by the ~~precedent~~previous analysis of the TIR imagery observations ~~(cf. Antonelli et al., 2019).~~(cf. Antonelli et al., 2020). The temporal focus of our study is on the seasonal dynamics of surface saturation. By contrasting the simulation results with observed surface saturation characteristics from TIR imagery, we ~~aim to~~ use the model as a learning tool and address the following research questions:

1) To what extent can we reproduce the observed intra-catchment variability of surface saturation characteristics with a rather homogenously set up ISSHM?
2) What key controls for the intra-catchment variability of surface saturation characteristics can we identify based on the match and mismatch 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~~data ~~(Antonelli et al., 2019; Glaser et al., 2018)~~(Antonelli et al., 2020; 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~~during 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~~based on the previous modelling work for the 6 ha headwater (Glaser et al., 2016, 2019). We first ~~perform an evaluation of~~evaluate 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~~to answer the two research questions.

2 Study site and data

2.1 Physiography, climate and hydrometry

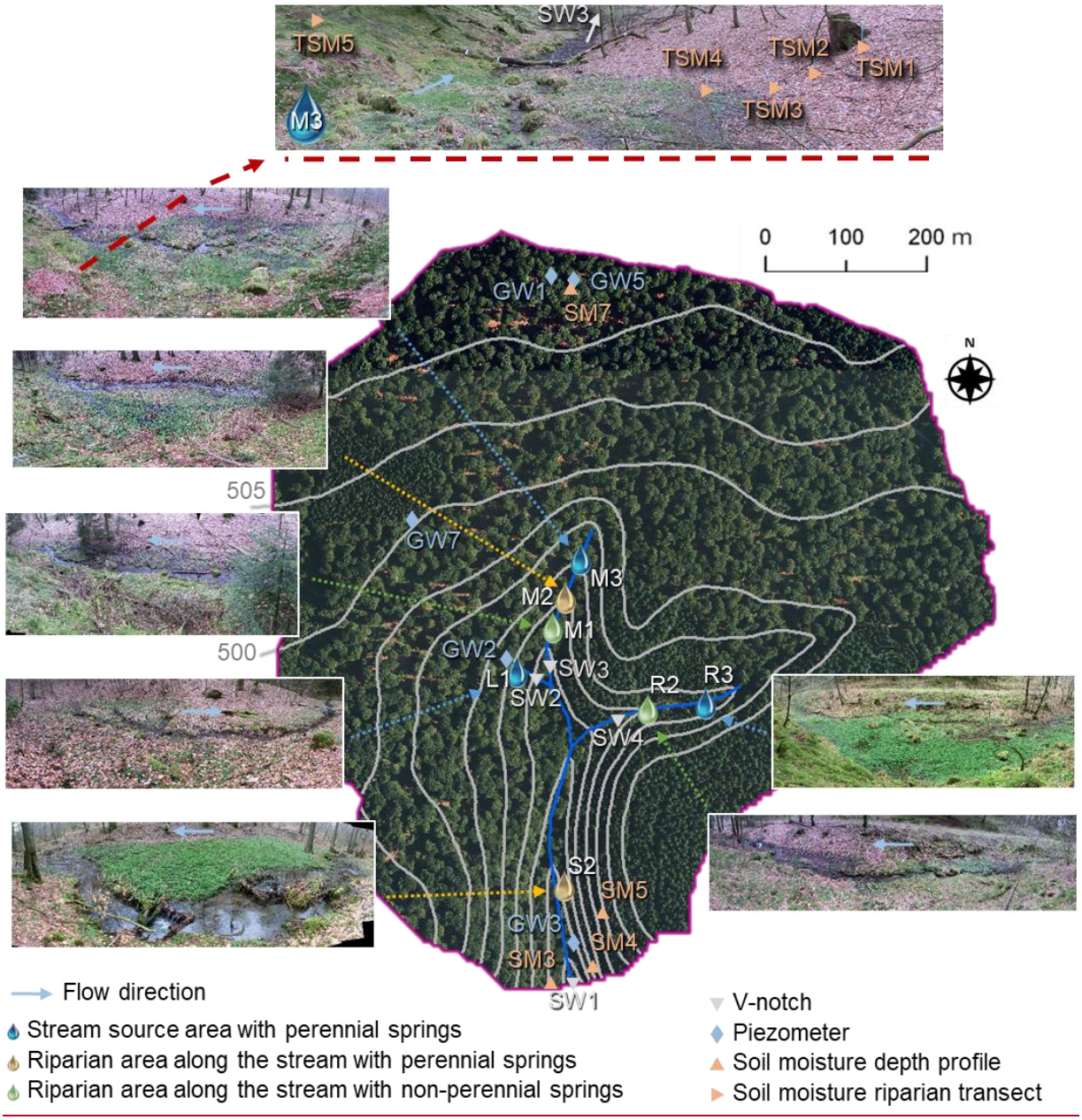
The Weierbach catchment is an intensively studied headwater catchment (42 ha) in western Luxembourg. About half of the catchment area is characterized by gentle slopes <5°, forming a plateau landscape unit (Martínez-Carreras et al., 2016). The rest of the catchment is characterized by hillslopes with 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 tributary is situated west of the central stream valley. ~~Riparian zones along the stream account for 1.2 % of the catchment area (Antonelli et al., 2019).~~Riparian zones along the stream account for 1.2 % of the catchment area (Antonelli et al., 2020). 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 trees and covered with ferns, moss, and herbaceous plants. Siltic, sceletic Cambisols, developed from Pleistocene Periglacial Slope Deposits, are shallow and highly-permeable 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 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 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 weirs, installed at the outlet of the catchment (SW1, Fig.1) and upstream of the confluences of the three 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 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 at 10 cm depth was monitored with water content reflectometers (CS616, Campbell Scientific, 30 min logging intervals) at five locations transecting the riparian zone of the stream source area of the middle tributary (Fig. 1, 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 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 at 3 m height and converted to wind speed ~~in~~at 2 m height (data gaps ~~closed~~were filled 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.


SW3
TSM5
TSM4
TSM2
TSM1
TSM3
M3
0
100
200 m
GW1
GW5
SM7
N
505
GW7
M3
M2
M1
500
GW2
L1
SW3
SW2
R2
R3
SW4
S2
SM5
GW3
SM4
SM3
SW1
Flow direction
Stream source area with perennial springs
Riparian area along the stream with perennial springs
Riparian area along the stream with non-perennial springs
V-notch
Piezometer
Soil moisture depth profile
Soil moisture riparian transect

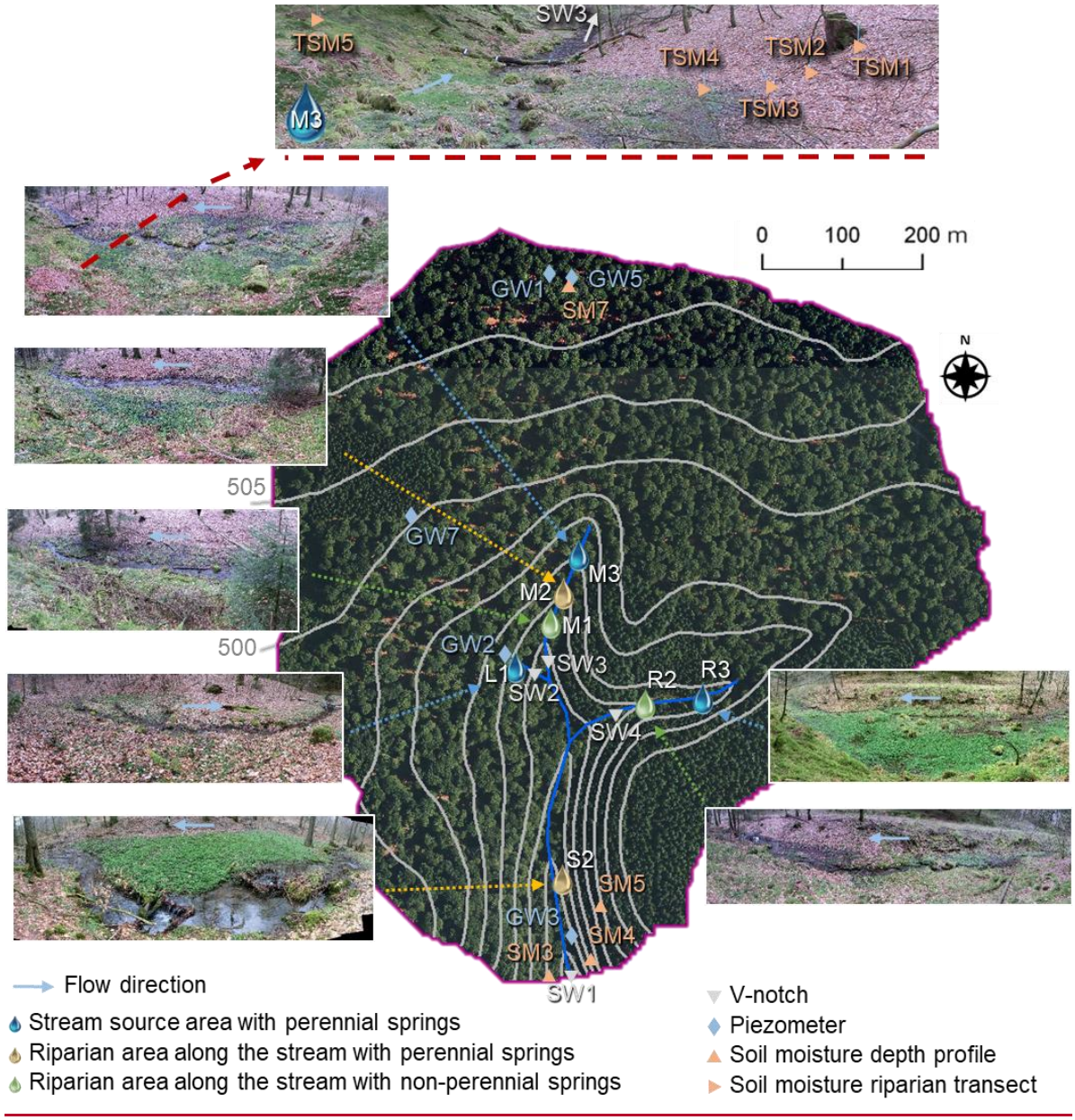


Figure 1: Weierbach catchment with the locations of the installed v-notch weirs for measuring discharge (SW1-SW4), piezometers for measuring groundwater levels (GW1-GW3, GW5, GW7), soil moisture sensors for measuring soil moisture in different depth profiles (SM3-SM5, SM7) and a riparian transect (TSM1-TSM5), 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 as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This includes water bodies such as 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 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). The seven investigated areas include about one quarter of the total stream network and ~~area sizes~~areas 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).~~(cf. Antonelli et al., 2020). According to their main hydro-morphological features, the seven areas can be classified into three different categories ~~(cf. Antonelli et al., 2019)~~(cf. Antonelli et al., 2020): i) stream source areas with

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

We mapped the surface saturation in the 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~~(2020). In brief, we took a panoramic photograph 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, 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, slight shifts in perspective between the panoramas of different dates were inevitable and the different features (e.g. streambed, stones, trees) in the images the images did not always overlap exactly pixel by pixel.

The~~-~~ locations of surface saturation (including the stream) were identified from the TIR panoramas based on the acquired temperature information. Each pixel in a TIR panorama was ~~assigned~~either considered to be saturated or unsaturated based on the temperature range of locations that were obviously saturated according to field observations and ~~visual~~visible light 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~~(cf. Antonelli et al., 2020). While this is a subjective and laborious approach, previous 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~~the same person for all the images ~~(cf. Antonelli et al., 2019; Glaser et al., 2018)~~(cf. Antonelli et al., 2020; 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, the images were analysed but flagged as less reliable. Altogether, we obtained from 63 monitoring dates a total of 291 binary panoramic images (i.e. 66 % of the 441 panoramas taken) showing the temporal dynamics of surface saturation patterns in the seven studied riparian areas. The total number of analysed binary images per site ranged 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 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. 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 riparian area how often the individual pixels of the panoramic TIR images were classified as saturated and normalised the resulting frequency numbers by the total number of TIR images analysed for that area. The resulting maps of normalised saturation frequency suggest that very few pixels were always saturated (i.e. reaching a normalised frequency of 1). Field experience and the analysis of individual TIR and visual images showed that in reality, surface saturation was permanent at more places than indicated by the frequency maps. The reason for this artefact is that the perspective and distortion 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 for instances where the extent of surface saturation was ~~narrow.~~ small. As a result, the generated saturation frequency maps are blurred. Nonetheless, the maps of normalised saturation frequencies are very useful to quickly detect where surface saturation occurs more ~~and~~or less frequent within an area and 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 address 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 already 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 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, 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).

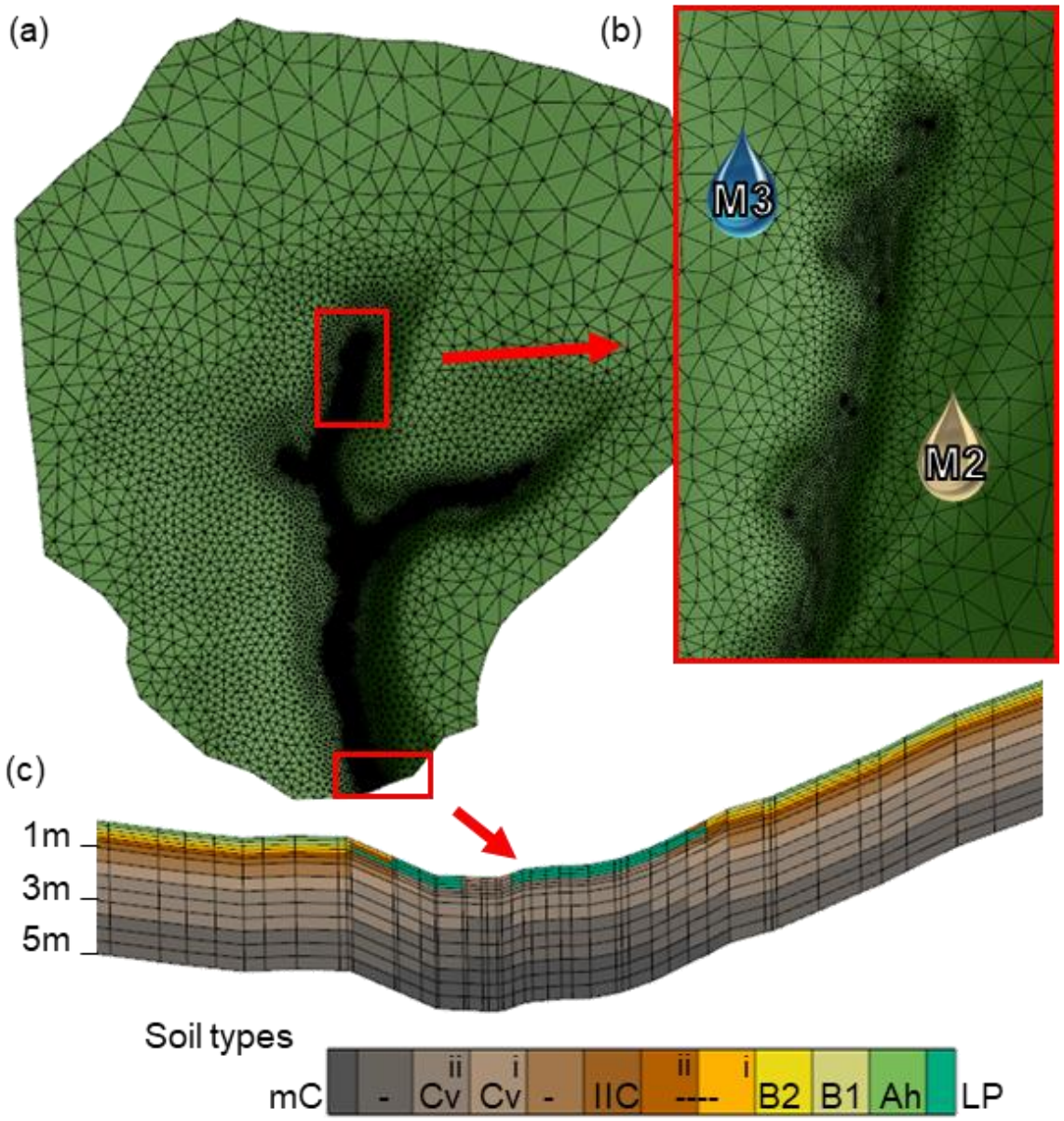


Figure 2: Setup of the model mesh (a) with a zoom on the fine horizontal resolution in the riparian areas and the streambed (b) and a vertical cross section through the stream valley and adjacent hillslopes (c) showing the vertical discretization and assignment of different soil properties (cf. Tab. 1). LP = riparian Leptosol, Ah = topsoil, B1 and B2 =subsoil, IIC = basal layer, Cv = fractured bedrock, 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, mC), and transition layers between subsoil, basal layer, and fractured bedrock (Fig. 2c). A differing subsurface structure was implemented in the stream valleys because in this area the soil and basal layers are eroded and the outcropping fractured bedrock is overlain with organic, stagnic Leptosol (LP) in the riparian zones (Fig. 2c). We used the Mualem-van Genuchten soil hydraulic functions to 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 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, we added new parameter 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 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 values of Glaser et al. (2016), who assigned the ~~values~~evapotranspiration parameters 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. The application of the adaptive time-stepping scheme of HGS ensured simulations with sub-daily resolution (minimum time step < 1 minute) for periods with highly dynamic hydrologic conditions and coarser temporal resolution (maximum time step 1 day) for hydrologically less active periods. A critical depth boundary was assigned to the outer edge of the surface domain, 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 to obtain realistic initial conditions. The actual simulation spanned 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 the model performance with discharge, groundwater level, and soil moisture ~~measured at~~data from 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. ~~Additionally~~Furthermore, 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 to ~~calculate~~determine 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 two depth profiles at each monitoring site.

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 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 and perspective. First, we transformed the simulated surface water depths for the days with TIR images into binary saturation maps of the entire catchment following the processing described above. Then, we converted the model output into jpeg images with the same perspective and extent ~~of~~as 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 normalised surface saturation extent, estimate the Kendall correlation for the relationship between normalised surface saturation extent and catchment discharge, and generate maps of 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~~percentage of saturated pixels and saturation frequency maps generated from the TIR and model images.

4 Results

4.1 Simulation of discharge, groundwater ~~level~~levels, 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 reasonably well with ~~KGEs~~KGE values 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 discrepancy was reflected in low ~~KGEs~~KGE values (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 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). Pearson correlation coefficients indicated overall a good agreement between simulated and observed soil moisture dynamics (Tab. 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, while the simulation showed a decrease in water content during dry periods (TSM4, Fig. 3). 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 and mismatch of the volumetric water content did not ~~clearly depend on~~show any spatial patterns and did not correspond to specific areas or landscape units. Moreover, we think that moisture dynamics and responsiveness are more ~~informative~~appropriate for model ~~performance~~evaluation than the absolute water content values, since also the measured values of volumetric water content differed within small distances (e.g. measurements of water content in 10 cm depth at profile SM7, Fig. 3).

4.2 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. 4a). During the wettest conditions of the study period (winter 2017/2018), surface saturation also ~~occurred as extension of the right tributary~~ extended into the hillslope above the source area R3 of the right tributary (cf. Fig. 1). This simulated occurrence of surface saturation across the catchment is consistent with field evidence, 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. 4b) were very similar to the simulated surface saturation frequency map of the catchment (Fig. 4a). The only obvious difference occurred in the area above the source area of the right tributary (R3), with a smaller simulated extent of groundwater reaching the surface than the simulated surface saturation extent.

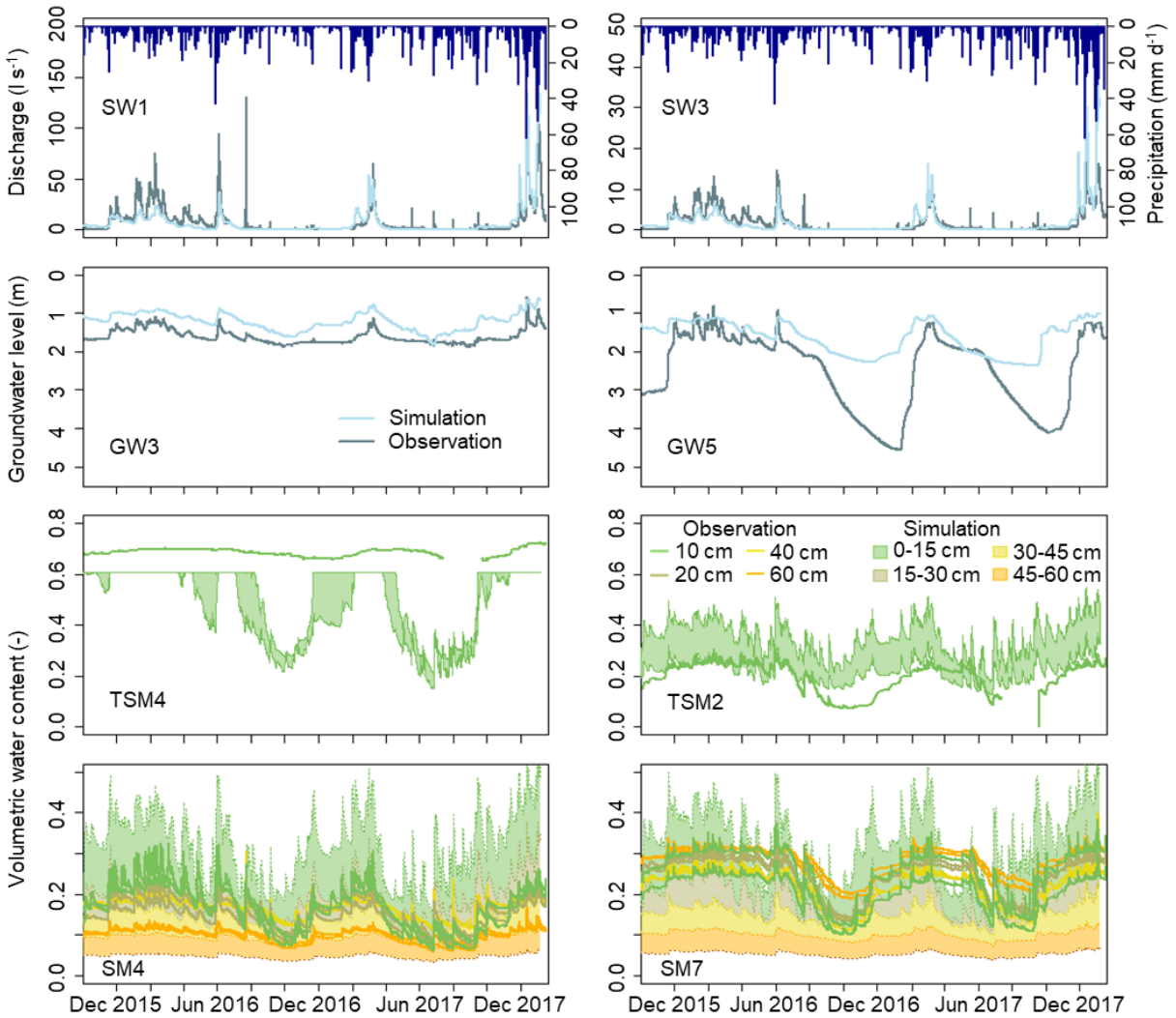
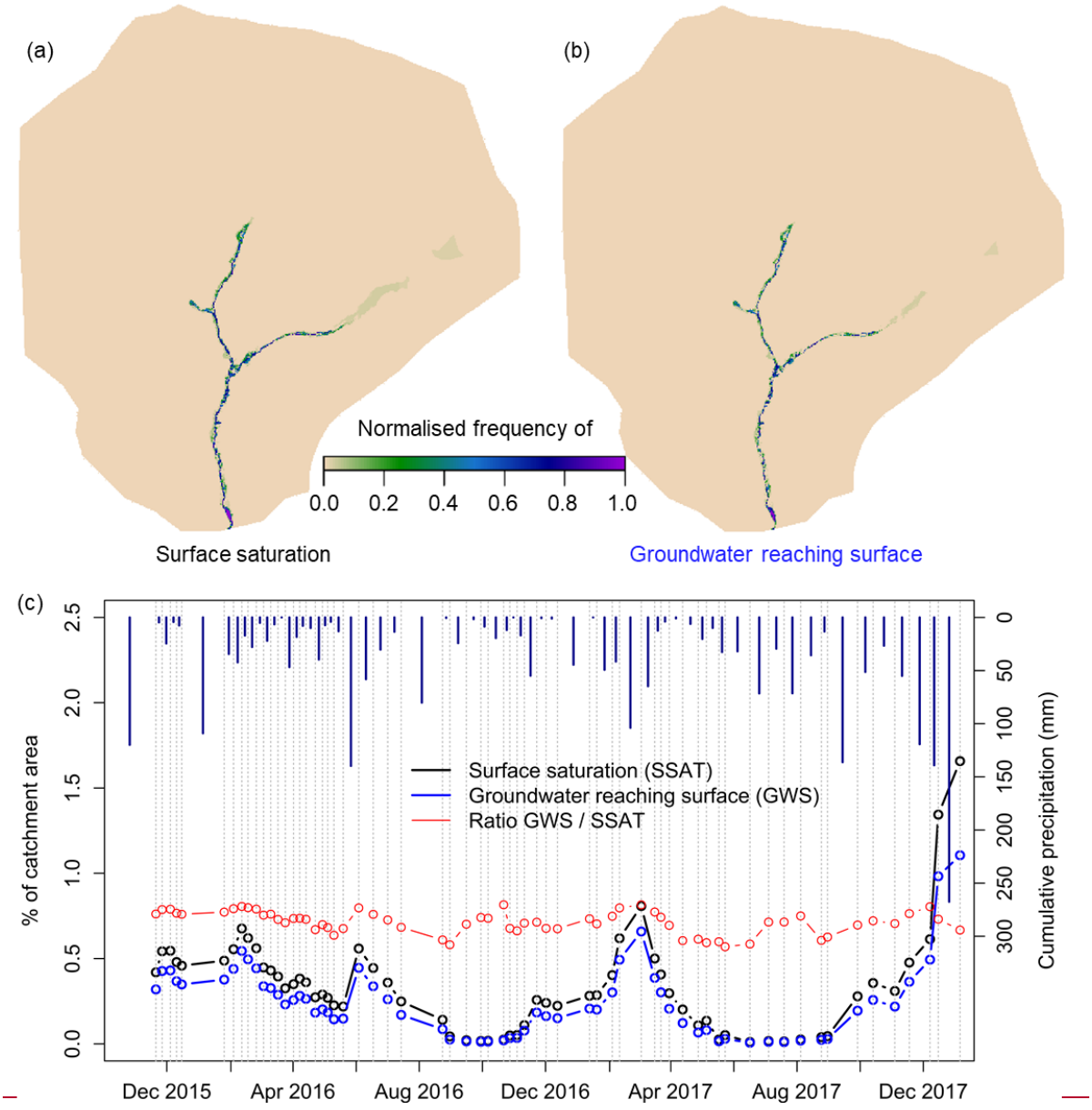


Figure 3: Simulated and observed time series of discharge, groundwater ~~level~~levels below the surface, and volumetric water content. Colour bands indicate the possible span of simulated volumetric water contents in ~~the depths~~ between two model nodes. The time series ~~of~~for the observation locations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

Table 2: 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					


(a)
(b)
Normalised frequency of
0.0 0.2 0.4 0.6 0.8 1.0
Surface saturation
Groundwater reaching surface
(c)
% of catchment area
Cumulative precipitation (mm)
Surface saturation (SSAT)
Groundwater reaching surface (GWS)
Ratio GWS / SSAT
Dec 2015
Apr 2016
Aug 2016
Dec 2016
Apr 2017
Aug 2017
Dec 2017

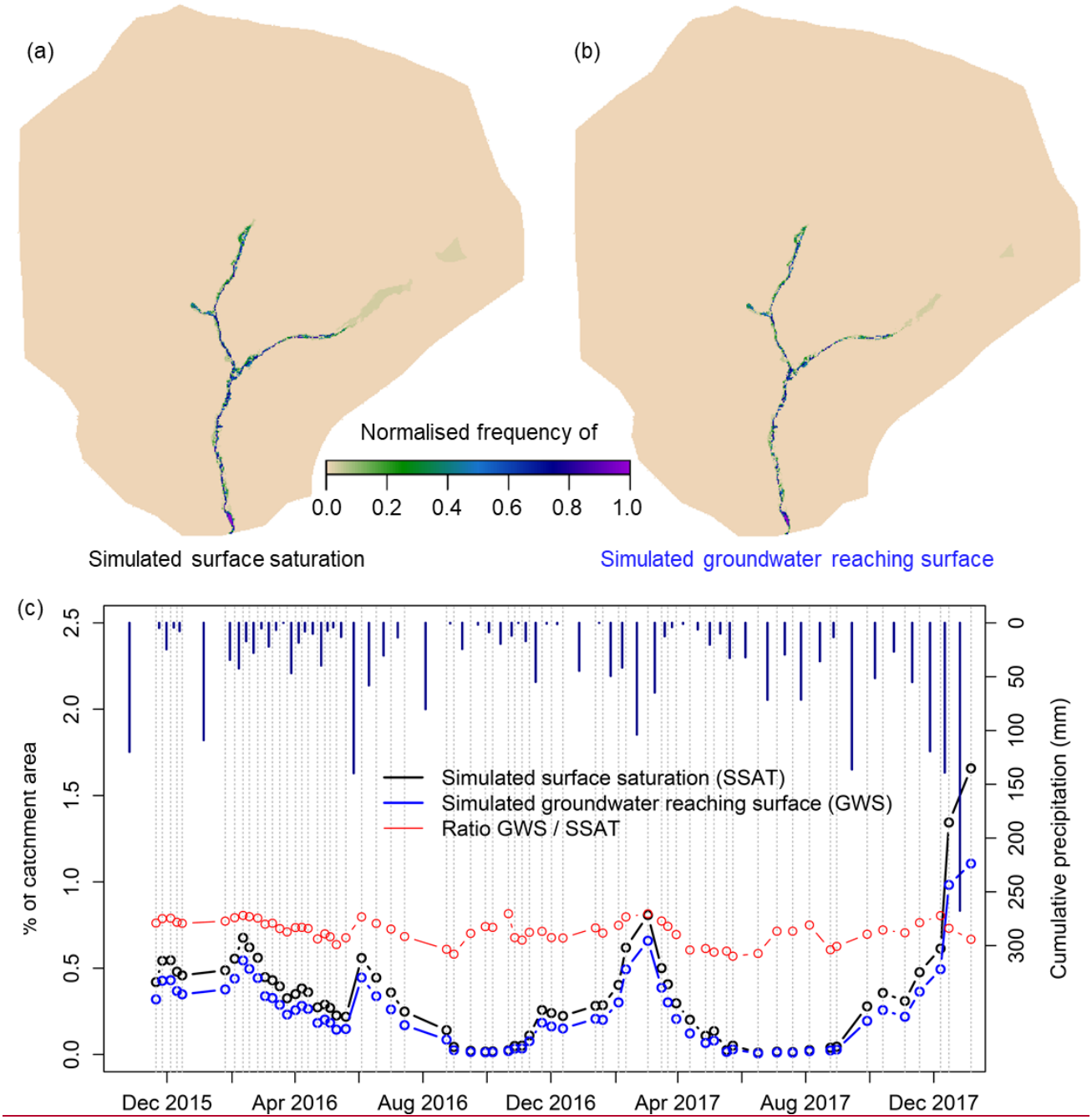


Figure 4: ~~Simulated frequency~~Frequency maps (a, b) and time series of areal percentage (c) of simulated surface saturation and simulated groundwater reaching the surface in the Weierbach catchment. Surface saturation (a, black line in c) corresponds to a simulated water depth >10^{-4} m in the surface 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 levels 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. 4c). The biggest absolute difference between the areal extent of surface saturation and groundwater reaching the surface was simulated during winter 2017/2018 (1.~~66~~7 % vs 1.1 % of catchment area), ~~where~~when the conditions were very wet with high discharge and high cumulative precipitation ~~and where~~. At this time, the difference in areal extent was also visible in the frequency maps (Fig. 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 varied 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.

4.3 Temporal dynamics of surface saturation extent within distinct riparian areas

The time series of observed normalised surface saturation extent (Fig. 5, coloured lines) were similar for all seven investigated riparian areas and followed the same seasonal trend as 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. 5) 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. 5). Similar to the observations, the simulated dynamics of normalised surface saturation (Fig 5, black lines) followed the trend of the simulated discharge dynamic. The simulation showed a faster decrease and increase of the normalised saturation during dry periods compared to what was observed in most areas. However, simulated discharge also seemed to decrease and increase earlier than ~~–~~the observed discharge (c.f. Section 4.4). The simulated saturation dynamics did not clearly differ between the different locations and thus ~~behaved~~were 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 normalised saturation was better for some areas (e.g. M1, R2, Fig. 5) than for others (e.g. S2, L1, Fig. 5).

The ~~dynamic~~temporal changes of simulated normalised ~~simulated~~ saturation matched the ~~normalised~~ observations generally well, despite under- and over-estimation of the minimum and maximum absolute saturation for all areas. The minimum 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. 5). In addition, simulated normalised saturation ~~stayed longer~~ remained close to the minimum longer 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. 5). The overestimation was not ~~equally strong at~~the same for the seven investigated areas and as a result, the distinction between areas with higher or lower maximum saturation was not the same for the observations and simulations (e.g. for R3 ~~showing one of~~ the second highest maximum saturation ~~in the observation~~was observed compared to the other areas, but ~~one of~~ the second lowest maximum saturation ~~in the simulation compared to the other areas~~was simulated).

4.4 Relationship between surface saturation extent and catchment discharge for distinct riparian areas

Kendall correlation between normalised surface saturation and discharge at the outlet (SW1) was > 0.60 for both the ~~simulation~~observed and ~~the observation in~~simulated for all riparian areas, indicating a monotonic relationship between the weekly to biweekly variation of surface saturation and discharge in all areas (Fig. 6). The simulated relationships between normalised surface saturation and catchment discharge resembled the observed relationships in terms of ~~value~~ range and shape (Fig. 6, Fig. S2), although the observation data were distinctly more scattered than the simulation data. A power law relationship approximated the observed relationship between discharge and saturation for all seven areas~~;~~ (exponents ranging between 0.29 and 0.69), when data that were taken during rainfall or rising discharge were excluded ~~(cf. Antonelli et al., 2019).~~(cf. Antonelli et al., 2020). 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~~observations 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~~indicated 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 normalised saturation and its change relative to discharge was smallest in the source areas (L1, M3, R3). In the simulated relationships, the increase in saturation for high discharge (> 5 l s^{-1}) was ~~strongest~~largest 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~~extent of normalised

saturation, being highest for areas in the right tributary (R2, R3), followed by the areas in the middle tributary (M1, M2, M3), and L1 and S2.

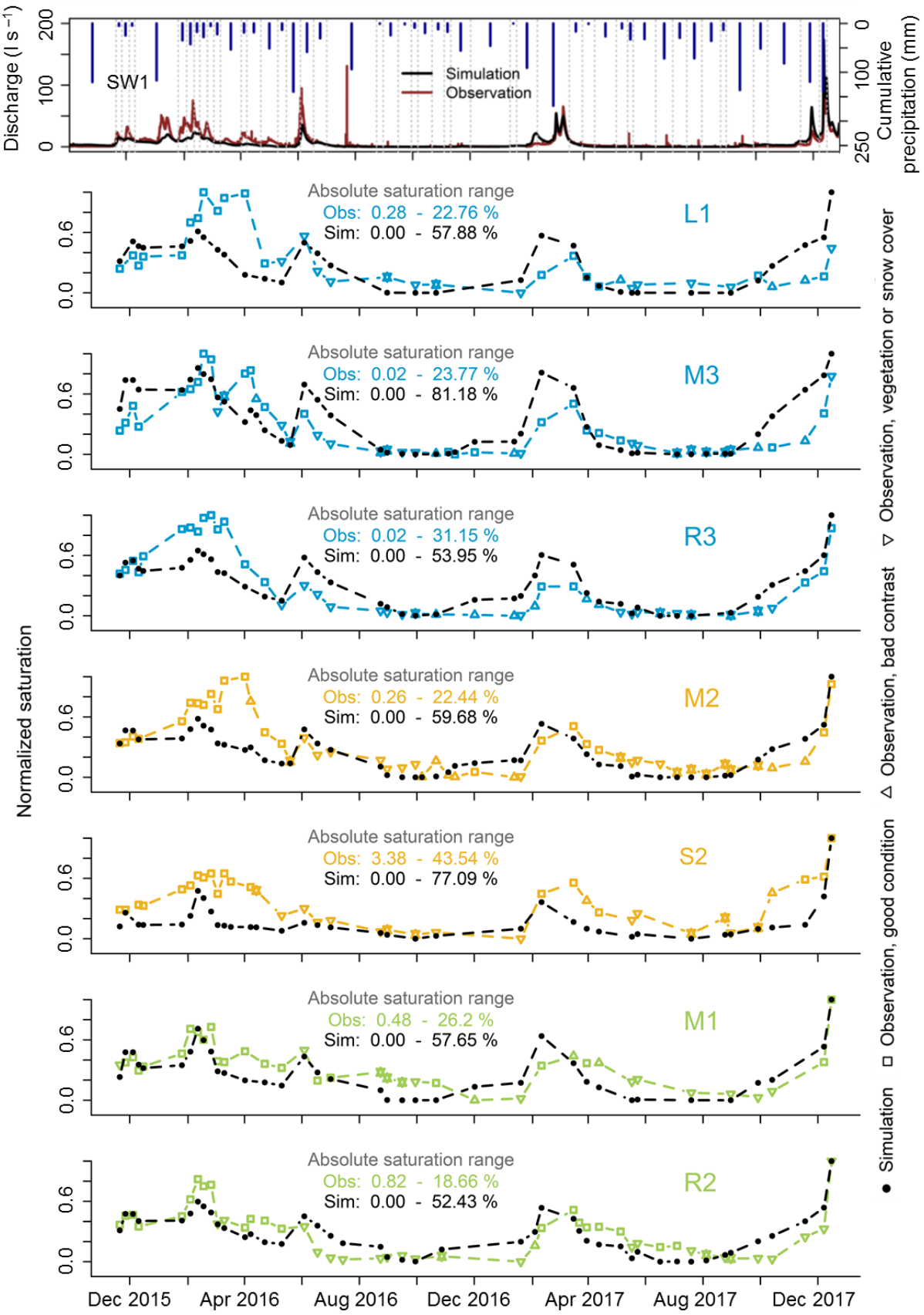
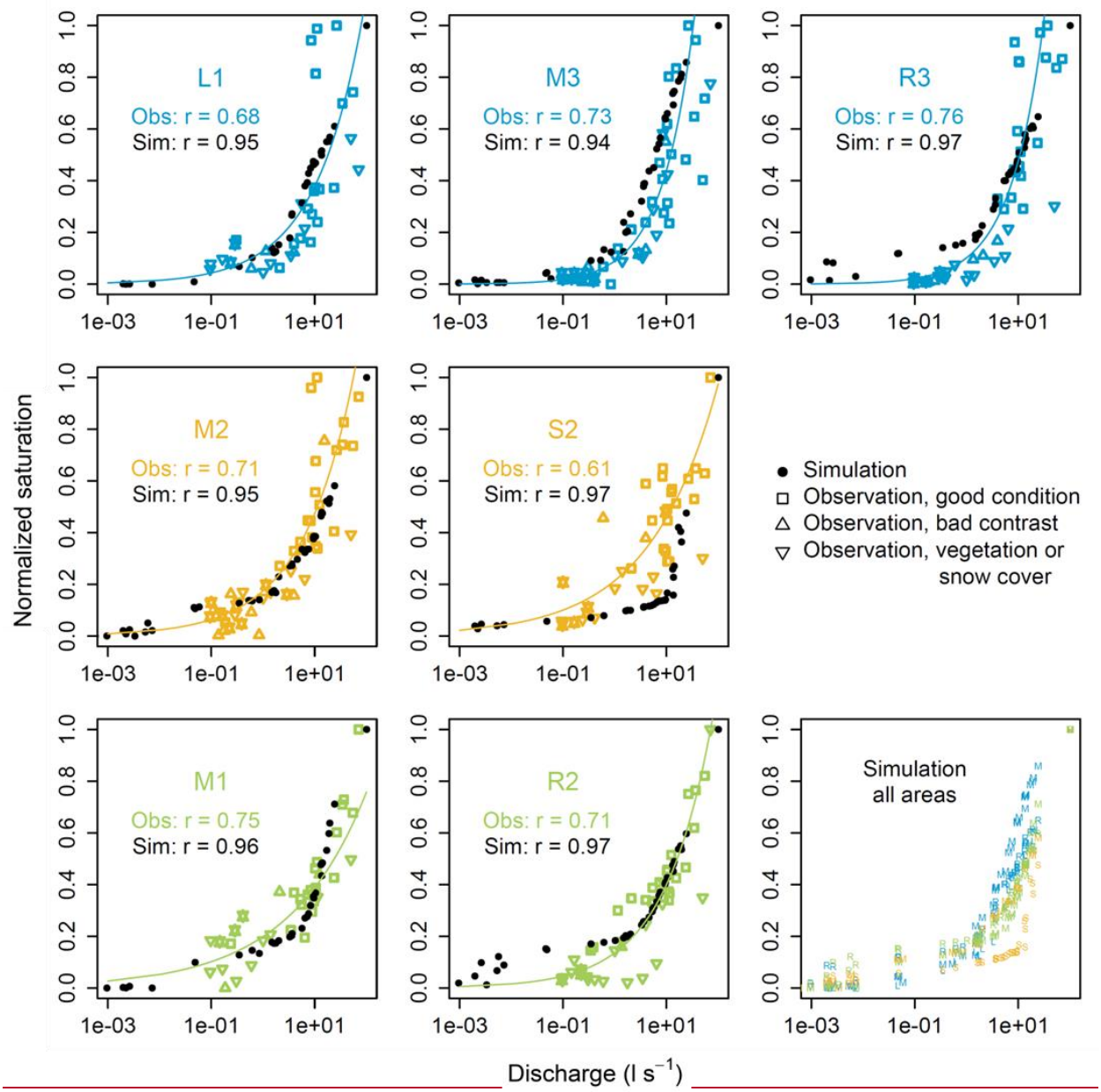


Figure 5: Time series of observed and simulated surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond to the colours of the icons in Fig. 1 and represent the different categories of 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~~indicated with ~~triangular symbols~~triangles (see symbol legend displayed on the right). Cumulative precipitation between the measurement dates (grey dashed lines) and discharge at the catchment outlet SW1 are shown in the top panel to facilitate the comparison to precipitation and flow conditions.


L1
Obs: r = 0.68
Sim: r = 0.95
M3
Obs: r = 0.73
Sim: r = 0.94
R3
Obs: r = 0.76
Sim: r = 0.97
M2
Obs: r = 0.71
Sim: r = 0.95
S2
Obs: r = 0.61
Sim: r = 0.97
Simulation
Observation, good condition
Observation, bad contrast
Observation, vegetation or snow cover
M1
Obs: r = 0.75
Sim: r = 0.96
R2
Obs: r = 0.71
Sim: r = 0.97
Simulation all areas
Normalized saturation
Discharge (l s^{-1})

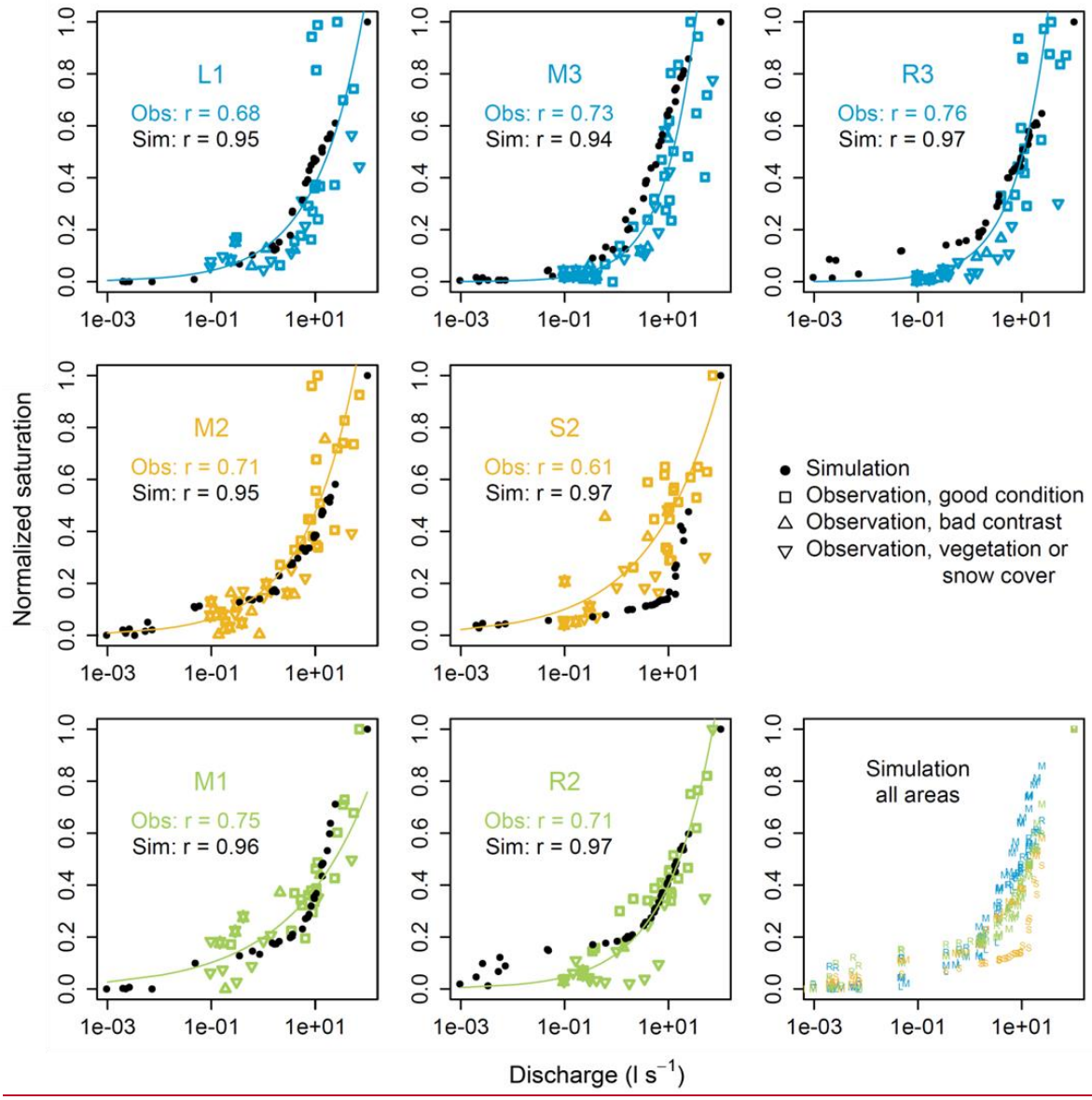


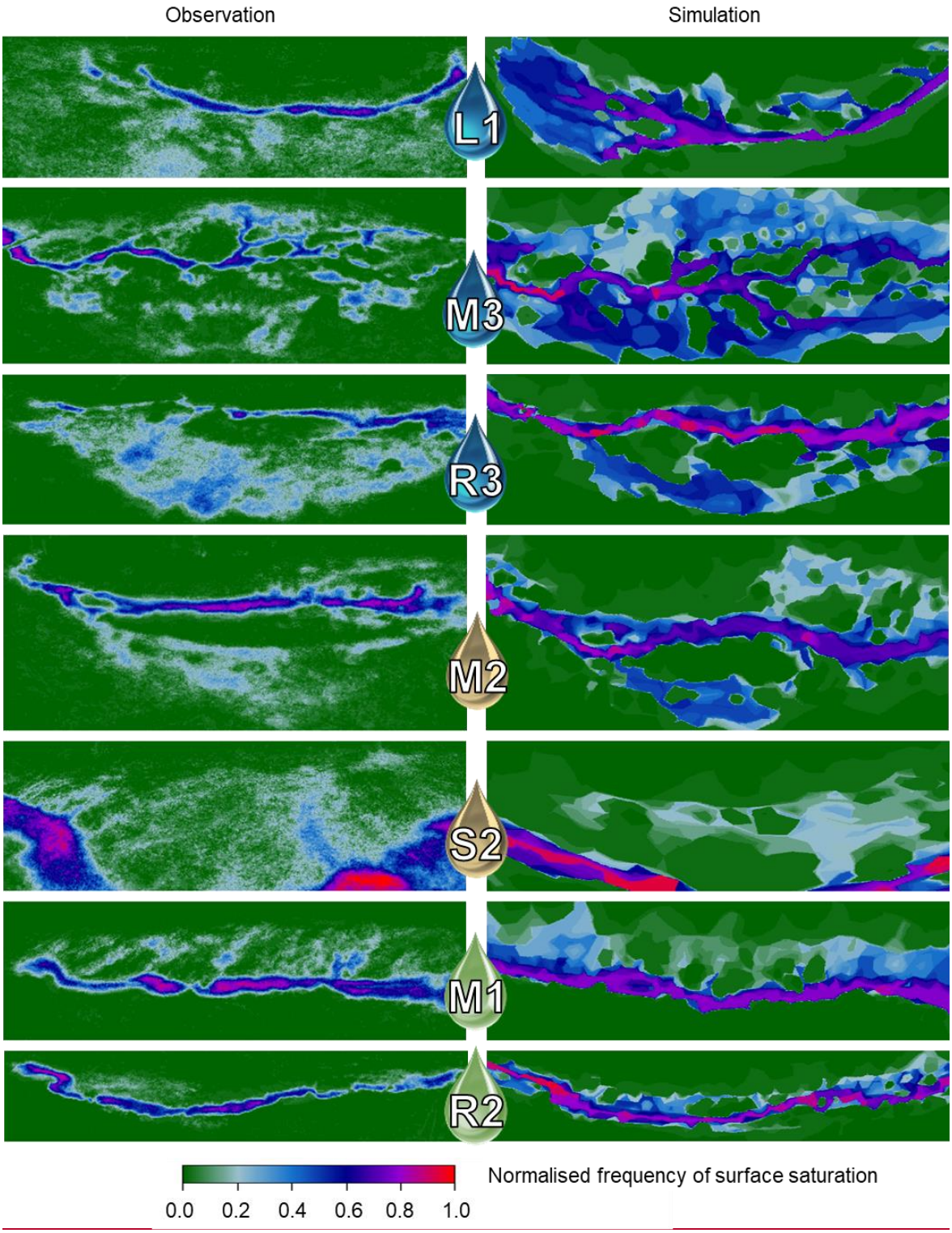
Figure 6: Observed and simulated relationships and Kendall correlations between normalised surface saturation and discharge at the catchment outlet SW1 for the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond 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.~~indicated with triangles. Simulated data points correspond to the times of observations (cf. Fig. 5); the plotted amount of data points with low and high discharge values differs between the observations and simulations due to mismatches between observed and simulated discharge at these times (cf. Fig. 5). Solid lines are power law curves fitted to the observation data, excluding data taken during rainfall or rising discharge. Note that discharge is plotted in logarithmic scale, showing relationships of convex shapes although power law exponents are < 1 for all areas. See Fig. S2 for plots with non-logarithmic abscissa and ordinate. To facilitate the comparison between the seven areas, the simulated data points from all seven areas are shown with the different colours and letters of the respective areas in the panel on the bottom right.

4.5 Spatial patterns of surface saturation: Occurrence and ~~frequencies~~frequency within distinct riparian areas

The realism of the simulated patterns of surface saturation was evaluated for each riparian area by visually comparing the surface saturation frequency maps ~~obtained from~~for the simulations and observations (Fig. 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~~observations and ~~simulation~~simulations 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. 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 patterns 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. 7) than in others (e.g. R3, Fig. 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.



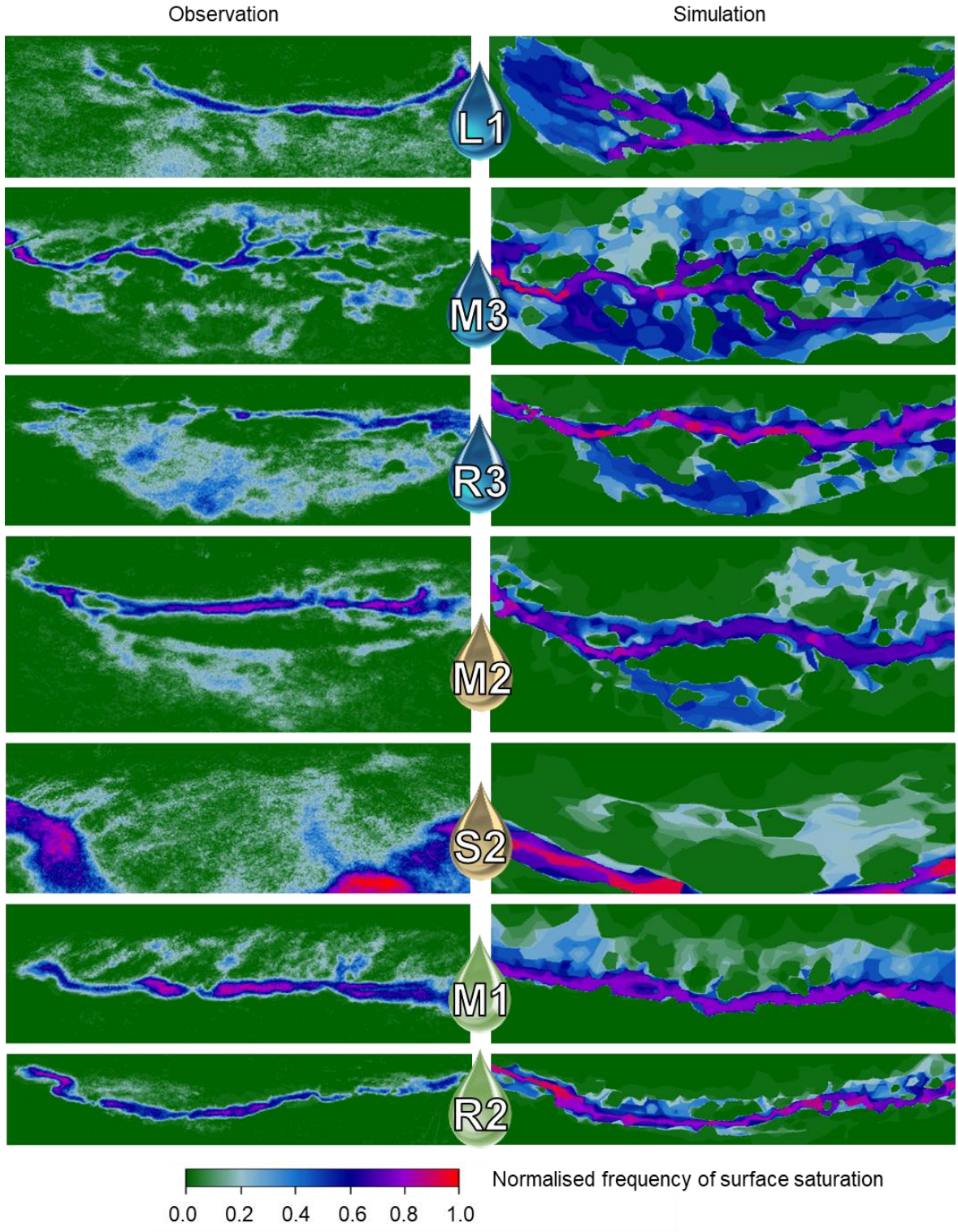


Figure 7: Observed (left) and simulated (right) frequencies of surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary, and the central stream (S2) (cf. Fig. 1). The maps were created by first counting how often the individual pixels 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.

5 Discussion

The aim of this study was to use an ISSHM as a complementary tool to field observations to analyse the spatio-temporal variability of surface saturation ~~within~~in the Weierbach catchment, with a focus on the stream valleys and riparian zones. We found some discrepancies between observed and simulated discharge, groundwater levels, and soil moisture -in terms of absolute values. Particularly, the model had some problems to reproduce soil moisture and groundwater levels during the dry

conditions at the hillslopes and plateau sites. We nonetheless argue that the match between the 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. 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~~the intra-catchment variability in surface saturation.

5.1 Temporal dynamics of surface saturation extent

The model reproduced ~~well~~ the observed dynamics of surface saturation in the seven investigated riparian areas over different seasons and wetness conditions well. Our work goes beyond previous studies 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 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 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~~saturated areas observed in a catchment in the Scottish Highlands, our model reproduced the distributed dynamics of surface saturation well, without clear differences in performance for different wetness conditions.

Simulations and observations both showed ~~both~~ that the temporal dynamics of the extent of surface saturation were mostly consistent across the seven different riparian sites of 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. 4). This suggests that ~~the generation of~~ surface saturation in the Weierbach catchment is largely driven by ~~the synchronous~~ exfiltration of groundwater ~~in~~into the stream valley and topographic depressions. ~~The~~Note that 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~~consistent 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 also 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~~(2020) drew conclusions that are consistent 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. This demonstrates that the simulation could confirm the assumed important role of groundwater exfiltration for surface saturation generation in the Weierbach catchment. We think that future applications of a combined observation and ISSHM simulation approach in other catchments will help to prove if the identified dominance of groundwater level dynamics in generating saturated soils and subsequent saturation excess surface saturation in the Weierbach is valid for catchments with similar and differing environmental conditions and characteristics.

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 (cf. Fig. 6). This is consistent with earlier studies that showed power law relationships between contiguous connected surface saturated areas and discharge for events of daily to seasonal time scales (Mengistu and Spence, 2016; Weill et al., 2013). In contrast to these studies that analysed hourly to daily simulation data, we did not observe clear hysteretic loops in the seasonal relationship between saturation and streamflow~~. Nonetheless, the~~ with our weekly to biweekly data base. 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~~possible 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).~~(cf. Antonelli et al., 2020; Dunne et al., 1975). The HGS model set up for this study might be used in the future to further analyse a possible hysteretic behaviour at the time-scale of events. However, there are indications from the weekly to biweekly data that the simulation is currently not able to reproduce possible hysteretic behaviour: i) the scatter in the simulated discharge – surface saturation relationships identified at seasonal time scale was distinctly less than in the observed relationships and ii) the model had a tendency for less persistent and faster contracting surface saturation than observed.

~~The lack of such a~~ Assuming that the scatter in the observed discharge – surface saturation relationships results from hysteretic ~~process~~extension and contraction of surface saturation, a lack of representation of this phenomenon 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 generation (and thus less hysteretic~~) generation~~ behaviour) 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. 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 of the ~~observation data~~observations 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 tributaries, cf. Section 4.4). Nonetheless, our findings are in line with experimental studies that ~~discussed~~showed 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).

The ~~degree of~~ variability of shapes of the discharge – surface saturation relationships for the different ~~areas~~ studied riparian areas within the Weierbach catchment is comparable to the variability of shapes of the discharge – surface saturation relationships for different catchments ~~presented~~summarized by Latron and Gallart (2007) (Figure 8). We cannot compare our

results directly with the results shown in Latron and Gallart (2007), since we evaluated absolute discharge and normalised saturation, while they evaluated connected saturated areas in percentage of catchment area, but ~~discharge~~ normalised ~~to the catchment area~~discharge. 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 normalised discharge and absolute saturation (Fig. 8a), and once with absolute discharge and normalised saturation (Fig. 8b). The shape of the relationship for the entire Weierbach catchment was nearly linear in double-logarithmic space, 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 in the Weierbach differed from the catchment relationship and between each other (Fig. 8b). For example, for area S2 and M1 ~~showed a~~the shape of the relation in double-logarithmic space was convex ~~shape~~and similar to the observations in the Vermont W2 catchment made by Dunne et al. (1975), for area M3 ~~showed a~~it was rather concave ~~shape,~~ similar to the relationships found for a sub-catchment of the Asker basin (Myrabø, 1986) and the Strengbach catchment (Latron, 1990), but for area M2 ~~showed a~~it was 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 intra-catchment variability. Different areas within a catchment can behave hydrologically differently due to their location along the stream and differing morphological features. For example, the stream source areas showed a faster increase of surface saturation extent with increasing catchment discharge at low flow but smaller additional increase with increasing catchment discharge at high flow than the riparian areas further downstream (cf. Fig. 8b). This demonstrates that one relationship for the entire catchment is not sufficient to fully resolve the spatially variable relation between discharge and surface saturation since the catchment is internally structured. The impact of this intra-catchment variability on the inter-catchment comparison of landscape controls on the discharge – surface saturation relationships may be a topic for future research.

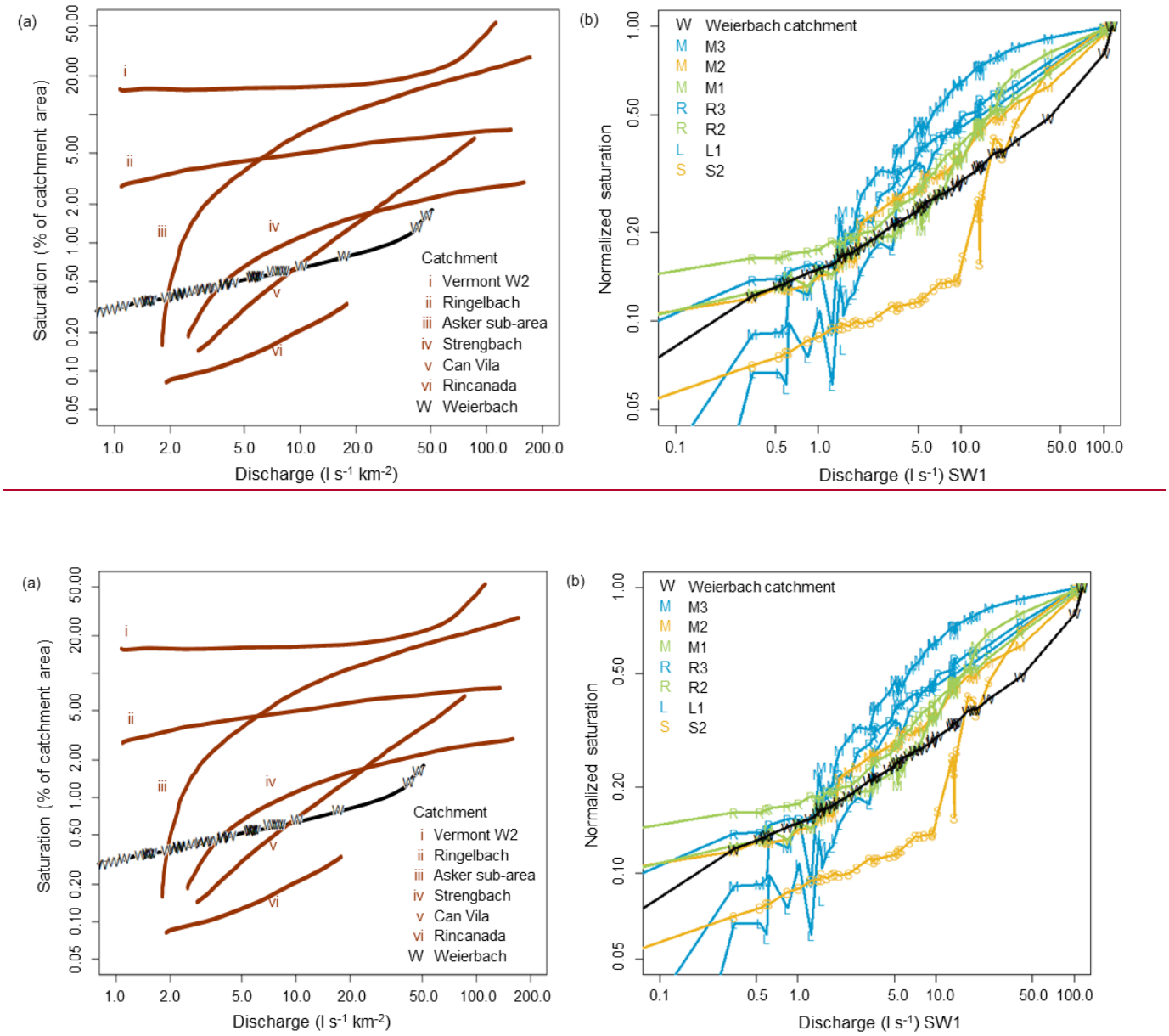


Figure 8: Simulated relationship between discharge and surface saturation of the entire Weierbach catchment (marked with W) in comparison to (a) the relationships observed in other catchments (Figure modified from Latron and Gallart (2007) and (b) the relationships simulated for the seven investigated riparian areas within the catchment~~.~~ (cf. Fig. 6). The ~~presented~~data for the relationships ~~of~~from the other catchments ~~were investigated by~~are from 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.5-7).~~

5.3 Spatial patterns of surface saturation occurrence

The observed spatial patterns of surface saturation occurrence were reproduced well by the simulations for most of the investigated areas. We attribute the successful simulation of the spatial patterns to microtopography (local topographical features with extents of tens 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~~S3, 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~~to identify 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, but still microtopography is not often considered in the simulation of surface saturation patterns.

When microtopography is not resolved in sufficient detail, it is more likely that the simulated surface water extends over a large area instead of being confined to occurring in topographic depressions and thus the model 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 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 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 infor small catchments (< 0.1 km^2), whereas studies that analysed the extent of surface saturation without considering microtopography in larger catchments (1 - 100 km^2) often simulated similar or smaller maximum extents of surface saturation than observed (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 extent of surface saturation reached a maximum of 1.6 % of the catchment area during the very wet conditions in winter 2017/2018 (cf. Fig. 4c). This simulated maximum extent of surface saturation is small compared to the estimationestimates in other simulation studies, but 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 the maximum saturation within the individual areas was overestimated compared to the observations (cf. Fig. 5). 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. 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 number of saturated pixels in the TIR images was thus lower than in the model images, while the minimum possible extent of saturation was not affected. This could also explain why overestimation of total 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. 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 left tributary (L1), where saturation was simulated at the clearly wrong side along the stream. Further field investigations and simulations will be needed to clarify if a varying subsurface structure is sufficient to correct the current mismatch between observed and simulated patterns of surface saturation or if further aspects, such as

differing evapotranspiration rates due to varying vegetation cover or exposure to the sun, have a relevant control on the observed saturation patterns as well.

5.4 Spatial patterns of surface saturation frequencies

The frequency maps of surface saturation (cf. Fig. 7) combine information on ~~when~~how often and where surface saturation occurs. We do not think that the exfiltration of subsurface water into the stream valley and local topographic depressions with extents of tens of centimetres to several metres (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. Instead, we assume that the differences in saturation frequency were controlled by additional water sources than exfiltrating groundwater, such as stream water from upstream 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 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 surface saturation generation processes ~~of surface saturation~~in the Weierbach catchment and beyond 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~~identify possible 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~~based on observations and physically-based simulations. We showed that the model ~~could reproduce~~reproduced 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 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. The use of the model as learning tool complementary to the field observations allowed a better understanding of the controls on surface saturation. Based on the identified match and mismatch between the simulation results and observations, we ~~could identify~~identified groundwater exfiltration and microtopography as key factors controlling the ~~occurrence~~dynamics and patterns of surface saturation. Yet ~~these two~~the model results showed that additional factors ~~alone were not sufficient~~are needed to explain the full variability of the different characteristics of surface saturation that were observed between the different areas.

The observed and simulated temporal dynamics of surface saturation extent were ~~observed and simulated to be~~ similar across the catchment, which we ~~related~~inferred – based on the simulation results – to ~~a large influence of~~be largely driven by synchronous groundwater level fluctuations across the catchment. Still, we 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 ~~–~~relate to differing 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.~~. The model could not fully reproduce the observed ~~varying~~ variability in persistence of surface saturation~~ and hysteretic relationships between surface saturation extent and catchment discharge.~~. Nonetheless, ~~the~~our simulation results demonstrated that~~ the shape of~~ the relationship between surface saturation and discharge for different riparian areas within a catchment can be as variable in shape as it has been observed between different catchments around the globe with different topographical and morphological conditions. Future research should synthesise inter- and intra-catchment landscape controls on dynamic surface saturation extension and contraction and its relation to runoff generation between and within catchments.

The observed spatial occurrence of surface saturation ~~differed between~~across and within the seven investigated riparian areas was captured well by the simulations, which we ~~could mainly relate~~related to ~~the~~a large influence of microtopography~~.~~ ~~Nonetheless~~ on the spatial patterns of surface saturation occurrence. However, 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~~Still, the model could satisfactorily reproduce the observed patterns of surface saturation frequencies for most parts of the different riparian areas. We suggest that the spatially varying frequencies of surface saturation might reflect a locally varying ~~relevance~~importance of different water sources. ~~Since~~For locations where the model could reproduce the observed frequencies, the model may be used in a future study to analyse such a potential mixing of different water sources and their variation in space and time.

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~~A. J., Klaus, J. and Pfister, L.: Saturated areas through the lens: 1. Spatio-temporal variability of surface saturation documented through ~~Thermal Infrared~~thermal infrared imagery, Hydrol. Process., ~~under review (HYP-19-0453), 2019~~1–23, doi:10.1002/hyp.13698, 2020.

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.

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

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 (b.glaser@lmu.de)

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.

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. In general, such surface saturated areas arise from 1) water ponding on the surface due to precipitation intensity exceeding 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 flowing into the floodplain (e.g. Megahan and King, 1985). When 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,

Summary of comments: glaser-et-al_2019_hess-2019-203_review2_commented.pdf

Page:1

Number: 1 Author: reviewer Subject: Note Date: 2020-01-05 18:03:52

Please plot Fig 5 using linear x-axis (not necessarily for the manuscript but for your interpretation). It is likely that the relation between discharge and normalized saturation is not a power law. Good news: The Kendall correlation will not change as it is non-parametic. If (but only if) the data fulfill 1) normality, 2) homoscedacticity and 3) linearity, your previous attempt to calculate Pearson correlation would work! Typically empirical data do not fulfill all three why it is wise to use non-parametric measures.

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 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 processes affects 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 allow detailed investigation of surface saturation at any desired location and time. This goes far beyond the information that can be gained by 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 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 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 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 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 use 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

No Comments.

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 to identify 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 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 headwater by accounting for a layering of the subsurface, while spatial heterogeneity was only represented by microtopography and 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 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 results with observed surface saturation characteristics from TIR imagery, we aim to use the model as learning tool and address the following research questions:

1) To what extent can we reproduce the observed intra-catchment variability of surface saturation characteristics with a rather homogenously set up ISSHM?
2) What key controls for the intra-catchment variability of surface saturation characteristics can we identify based on the match and mismatch 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

Page:3

Number: 1 Author: reviewer Subject: Note Date: 2020-01-05 12:06:19

include (42 ha) to be just even more clear that previous work was done in only a part if the Weierbach and this study goes beyond previous work!

Number: 2 Author: reviewer Subject: Insert Text Date: 2020-01-05 11:53:19

(42 ha)

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 tributary is situated 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 trees and covered with ferns, moss, and herbaceous plants. Siltic, sceletic Cambisols developed from Pleistocene Periglacial Slope Deposits are shallow and highly-permeable 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 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 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 weirs, installed at the outlet of the catchment (SW1, Fig.1) and upstream of the confluences of the three 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 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 at 10 cm depth was monitored with water content reflectometers (CS616, Campbell Scientific, 30 min logging intervals) at five locations transecting the riparian zone of the stream source area of the middle tributary (Fig. 1, 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 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

Page:4

Number: 1 Author: reviewer Subject: Insert Text Date: 2020-01-05 12:08:41

,

Number: 2 Author: reviewer Subject: Insert Text Date: 2020-01-05 12:08:55

,

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 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.

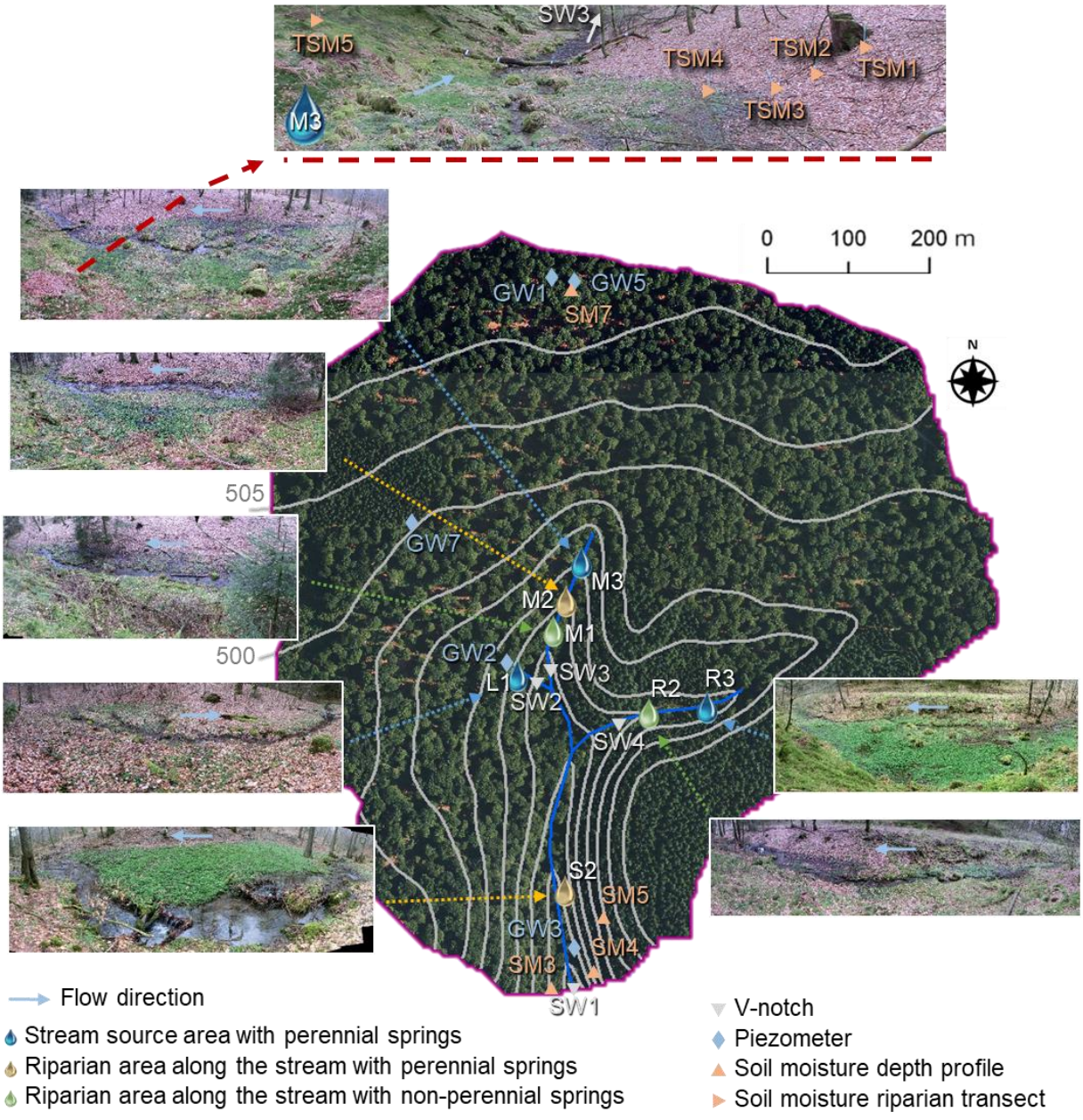


Figure 1: Weierbach catchment with the locations of the installed v-notch weirs for measuring discharge (SW1-SW4), piezometers for measuring groundwater levels (GW1-GW3, GW5, GW7), soil moisture sensors for measuring soil moisture in different depth profiles (SM3-SM5, SM7) and a riparian transect (TSM1-TSM5), 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).

No Comments.

2.2 Surface saturation

Here, we define the surface as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This includes water bodies such as 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 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). 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 can be classified into three different categories (cf. Antonelli et al., 2019): i) stream source areas with perennial springs (L1, M3, R3, blue icons Fig. 1), ii) areas along the stream with perennial springs (M2, S2, yellow icons Fig. 1), and iii) areas along the stream with non-perennial springs (M1, R2, green icons Fig.1).

We mapped the surface saturation in the 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, 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, slight shifts in perspective between the panoramas of different dates were inevitable and the different features (e.g. streambed, stones, trees) in the images the images did not always overlap exactly pixel by pixel.

The locations of surface saturation (including the stream) were identified from the TIR panoramas based on the acquired temperature information. Each pixel in a TIR panorama was assigned to be saturated or unsaturated based on the temperature range of locations that were obviously saturated according to 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, 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 total number of analysed images per site ranged 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 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. 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 normalised the resulting frequency numbers by the total number of TIR images analysed for that area. The resulting maps of normalised saturation frequency suggest that very few pixels were always saturated (i.e. reaching a

No Comments.

normalised frequency of 1). Field experience and the analysis of individual TIR and visual images showed that in reality, surface saturation was permanent at more places than indicated by the frequency maps. The reason for this artefact is that the perspective and distortion 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 for instances where the extent of surface saturation was narrow. As a result, the generated saturation frequency maps are blurred. Nonetheless, the maps of normalised saturation frequencies are very useful to quickly detect where surface saturation occurs more and less frequent within an area and 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 address 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 already 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 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, 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).

No Comments.

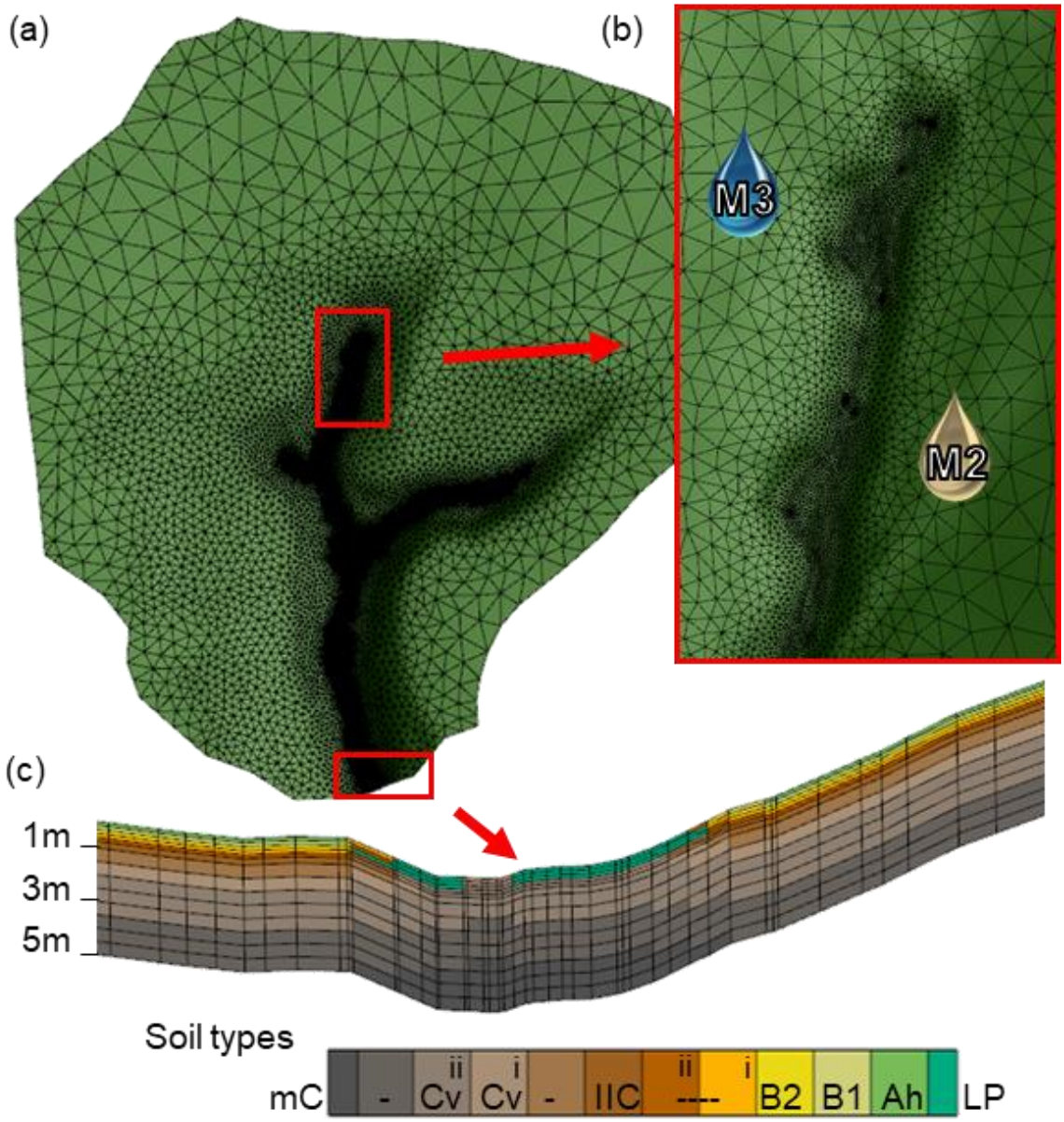


Figure 2: Setup of the model mesh (a) with a zoom on the fine horizontal resolution in the riparian areas and the streambed (b) and a vertical cross section through the stream valley and adjacent hillslopes (c) showing the vertical discretization and assignment of different soil properties (cf. Tab. 1). LP = riparian Leptosol, Ah = topsoil, B1 and B2 =subsoil, IIC = basal layer, Cv = fractured bedrock, 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, mC), and transition layers between subsoil, basal layer, and fractured bedrock (Fig. 2c). A differing subsurface structure was implemented in the stream valleys because in this area the soil and basal layers are eroded and the outcropping fractured bedrock is overlain with organic, stagnic Leptosol (LP) in the riparian zones (Fig. 2c). We used the Mualem-van Genuchten soil hydraulic functions to 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 catchment as well as from the shallow soil (5 cm and 35 cm) of nine locations

No Comments.

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, we added new parameter 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 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 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 outer edge of the surface domain, 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 to obtain realistic initial conditions. The actual simulation spanned 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 the model performance with discharge, groundwater level, and soil moisture measured 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. Additionally, we particularly evaluated the simulated groundwater level dynamics rather than absolute values based on Pearson correlation coefficients. Soil moisture was also

No Comments.

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 to 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 two depth profiles at each monitoring site.

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 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 and perspective. First, we transformed the simulated surface water depths for the days with TIR images into binary saturation maps of the entire catchment following the processing described above. Then, we converted 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 normalised surface saturation extent, estimate the Kendall correlation for the relationship between normalised surface saturation extent and catchment discharge, and generate maps of 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.

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 reasonably well with KGEs of

No Comments.

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 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 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). Pearson correlation coefficients indicated overall a good agreement between simulated and observed soil moisture dynamics (Tab. 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, while the simulation showed a decrease in water content during dry periods (TSM4, Fig. 3). 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 and mismatch 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 within small distances (e.g. measurements of water content in 10 cm depth at profile SM7, Fig. 3).

4.2 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. 4a). During the wettest conditions of the study period (winter 2017/2018), surface saturation also occurred as extension of the right tributary into the hillslope above the source area R3 (cf. Fig. 1). This simulated occurrence of surface saturation across the catchment is consistent with field evidence, 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. 4b) were very similar to the surface saturation frequency map of the catchment (Fig. 4a). The only obvious difference occurred in the area above the source area of the right tributary (R3), with a smaller extent of groundwater reaching the surface than the surface saturation extent.

Page:11

Number: 1 Author: reviewer Subject: Insert Text Date: 2020-01-05 12:59:23

simulated

Number: 2 Author: reviewer Subject: Note Date: 2020-01-05 18:05:23

I know that this section is on simulated patterns but to avoid any misunderstanding you might want to include "simulated" here again.

Number: 3 Author: reviewer Subject: Insert Text Date: 2020-01-05 12:59:46

simulated

Number: 4 Author: reviewer Subject: Insert Text Date: 2020-01-05 13:01:48

simulated

Number: 5 Author: reviewer Subject: Insert Text Date: 2020-01-05 13:01:56

simulated

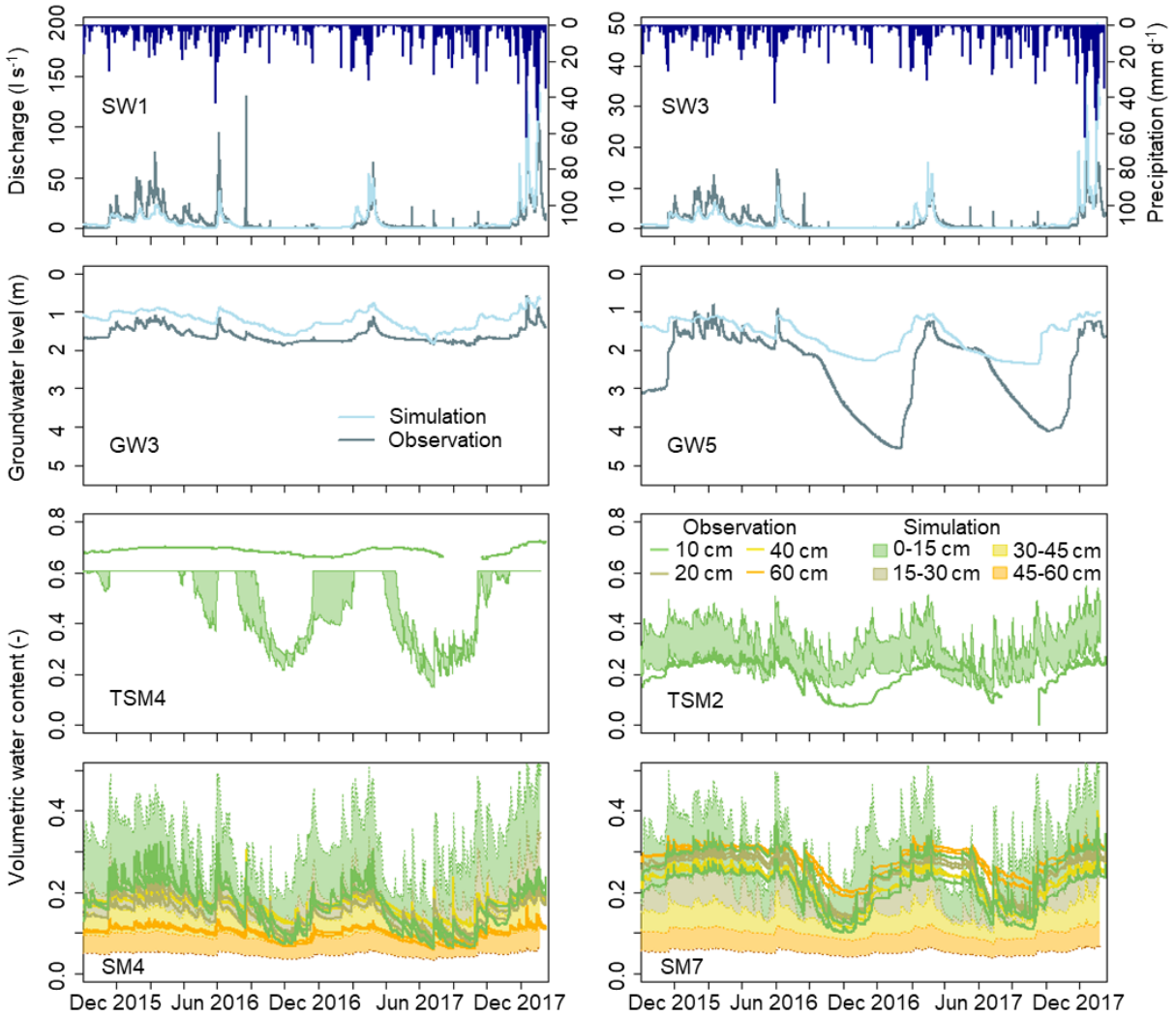


Figure 3: Simulated and observed time series of discharge, groundwater level below the surface, and volumetric water content. Colour bands indicate the possible span of simulated volumetric water contents in the depths between two model nodes. The time series of the observation locations (cf. Fig. 1)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

5

Table 2: 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					

No Comments.

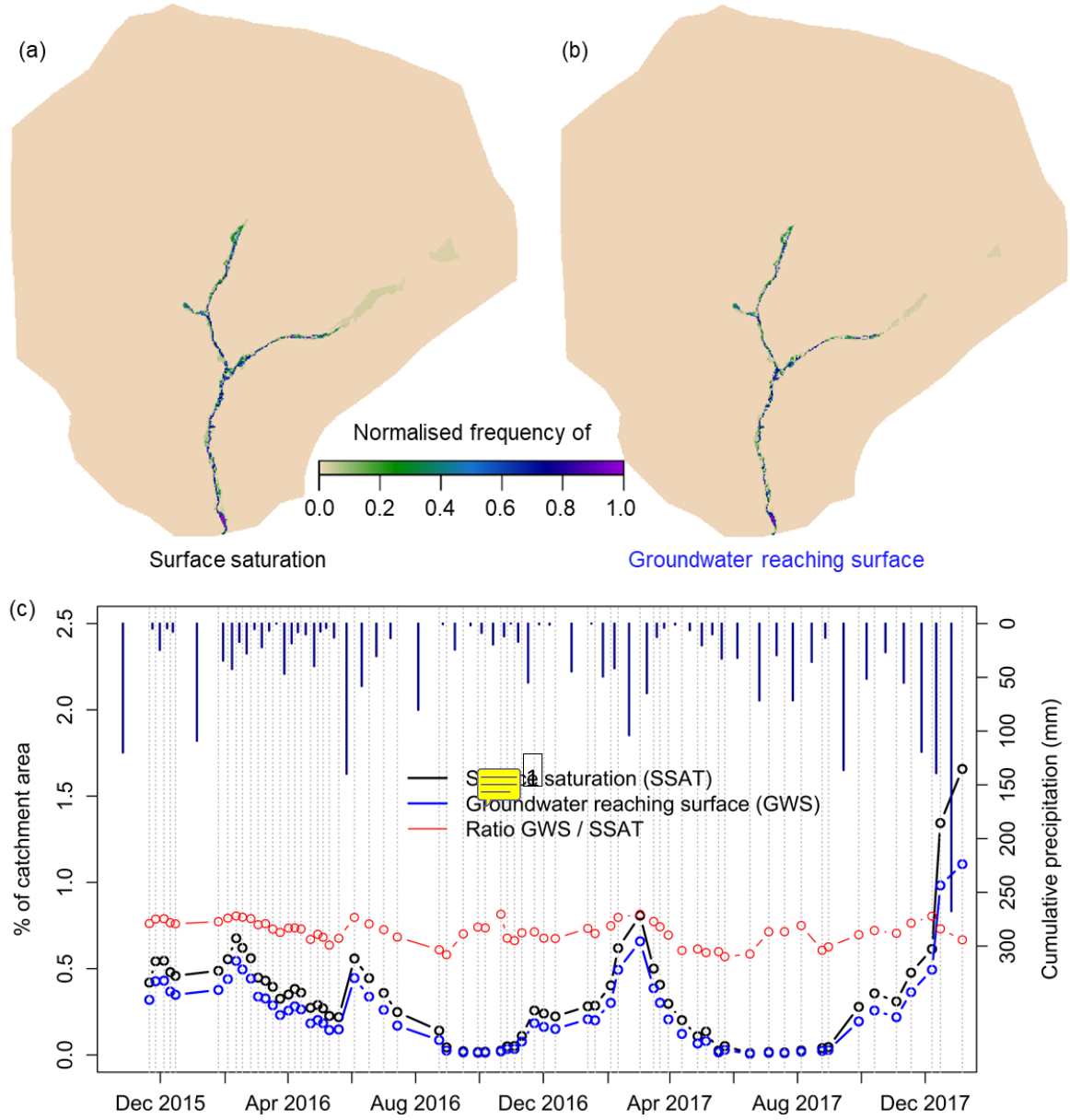


Figure 4: Simulated frequency maps (a, b) and time series of percentage (c) of surface saturation and groundwater reaching the surface in the Weierbach catchment. Surface saturation (a, black line in c) corresponds to a simulated water depth >10^{-4} m in the surface domain of the model. Groundwater 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. 4c). 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. 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 varied 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.

Page:13

Number: 1 Author: reviewer Subject: Note Date: 2020-01-05 13:04:28

maybe consider to also add "simulated" to the legend text of the black and blue line.

Number: 2 Author: reviewer Subject: Insert Text Date: 2020-01-05 13:02:39

simulated

Number: 3 Author: reviewer Subject: Insert Text Date: 2020-01-05 13:02:31

simulated

4.3 Temporal dynamics of surface saturation extent within distinct riparian areas

The time series of observed normalised surface saturation extent (Fig. 5, coloured lines) were similar for all seven investigated riparian areas and followed the same seasonal trend as 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. 5) 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. 5). Similar to the observations, the simulated dynamics of normalised surface saturation (Fig 5, black lines) followed the trend of the simulated discharge dynamic. The simulation showed a faster decrease and increase of the normalised saturation during dry periods compared to what was observed in most areas. However, simulated discharge also seemed to decrease and increase earlier than the observed discharge (c.f. Section 4.4). 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 normalised saturation was better for some areas (e.g. M1, R2, Fig. 5) than for others (e.g. S2, L1, Fig. 5).

The dynamic changes of normalised simulated saturation matched the normalised observations generally well, despite under- and over-estimation of the minimum and maximum absolute saturation for all areas. The minimum 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. 5). In addition, simulated normalised 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. 5). The overestimation was not equally strong at the seven investigated areas and as a result, the distinction between areas 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).

4.4 Relationship between surface saturation extent and catchment discharge for distinct riparian areas

Kendall correlation between normalised surface saturation and discharge at the outlet SW1 was > 0.60 for both the simulation and the observation in all riparian areas, indicating a monotonic relationship between surface saturation and discharge in all areas (Fig. 6). The simulated relationships between normalised surface saturation and catchment discharge resembled the observed relationships in terms of value range and shape (Fig. 6), although the observation data were distinctly more scattered 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 normalised saturation and its change relative to discharge was smallest in the source areas (L1, M3, R3). 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 normalised saturation, being highest for areas in the right tributary (R2, R3), followed by the areas in the middle tributary (M1, M2, M3), and L1 and S2.

No Comments.

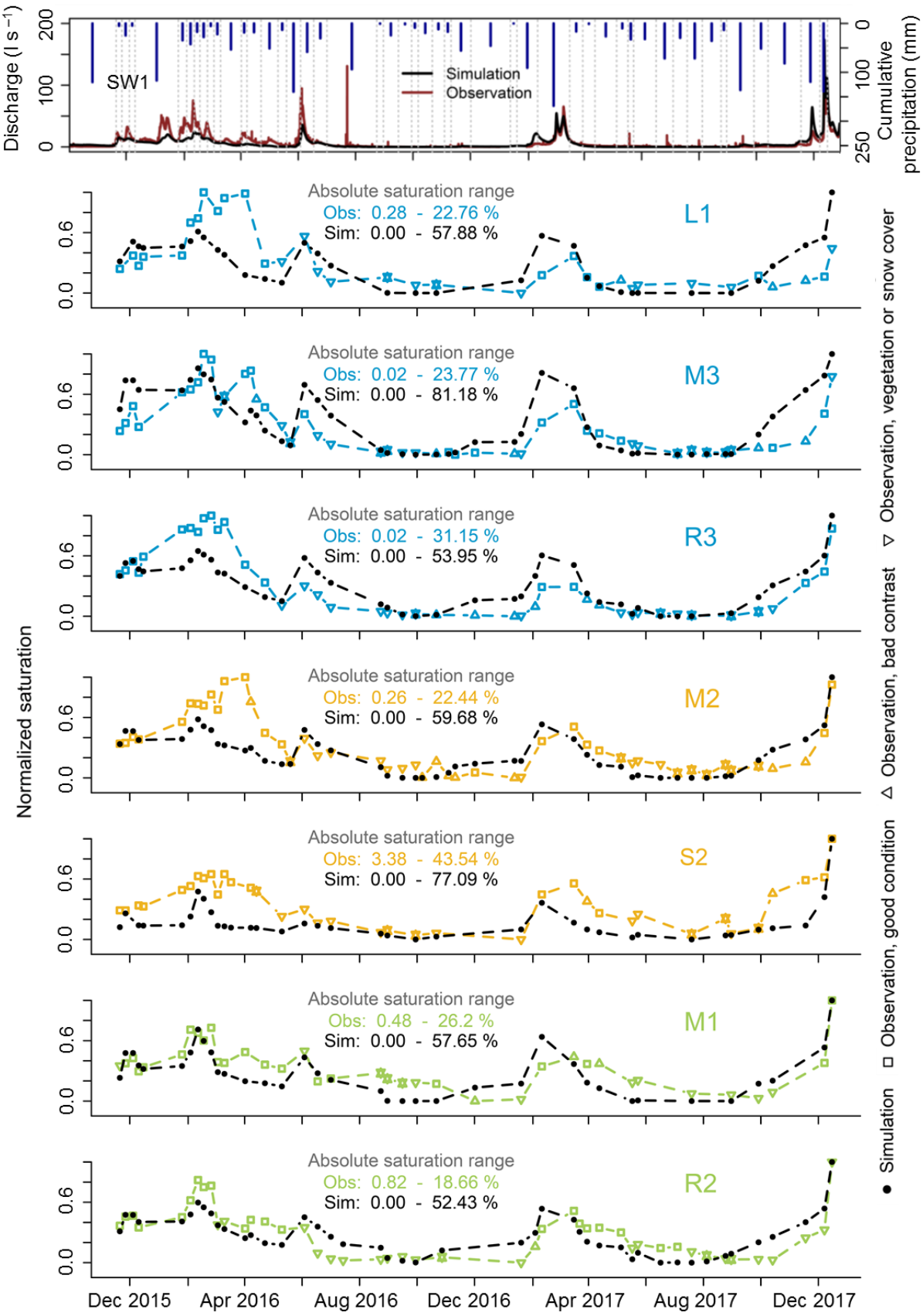


Figure 5: Time series of observed and simulated surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond to the colours of the icons in Fig. 1 and represent the different categories of 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 to facilitate the comparison to precipitation and flow conditions.

No Comments.

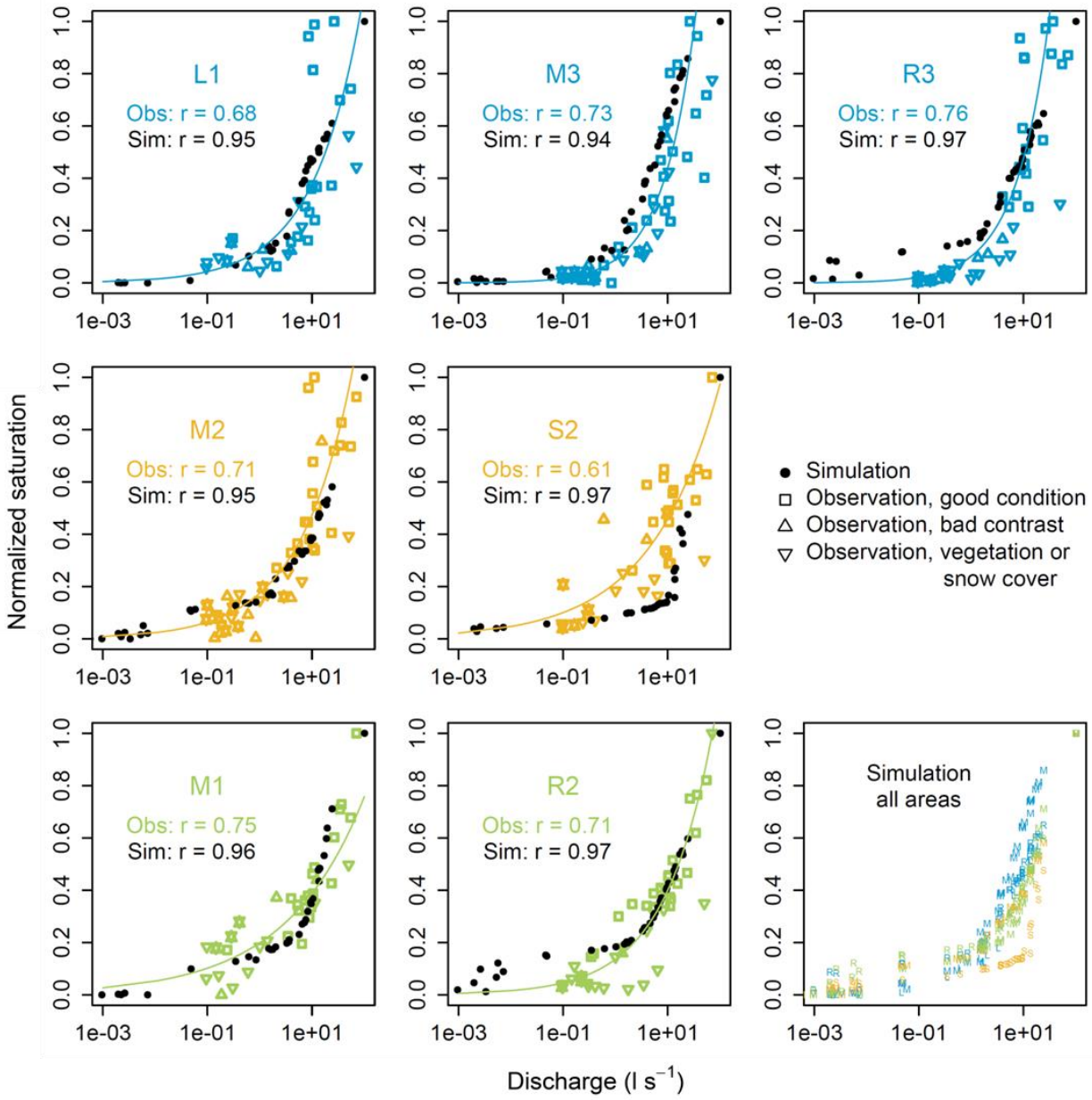


Figure 6: Observed and simulated relationships and Kendall correlations between normalised surface saturation and discharge at the catchment outlet SW1 for the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond 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. To facilitate the comparison between the seven areas, the simulated data points from all seven areas are shown with the different colours and letters of the respective areas in the panel on the bottom right.

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. 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 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. 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).

Number: 1 Author: reviewer Subject: Note Date: 2020-01-05 18:06:43

Please plot Fig 5 using linear x-axis (not necessarily for the manuscript but for your interpretation). It is likely that the relation between discharge and normalized saturation is not a power law. Good news: The Kendall correlation will not change as it is non-parametic. If (but only if) the data fulfill 1) normality, 2) homoscedacticity and 3) linearity, your previous attempt to calculate Pearson correlation would work! Typically empirical data do not fulfill all three why it is wise to use non-parametric measures.

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. 7) than in others (e.g. R3, Fig. 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.

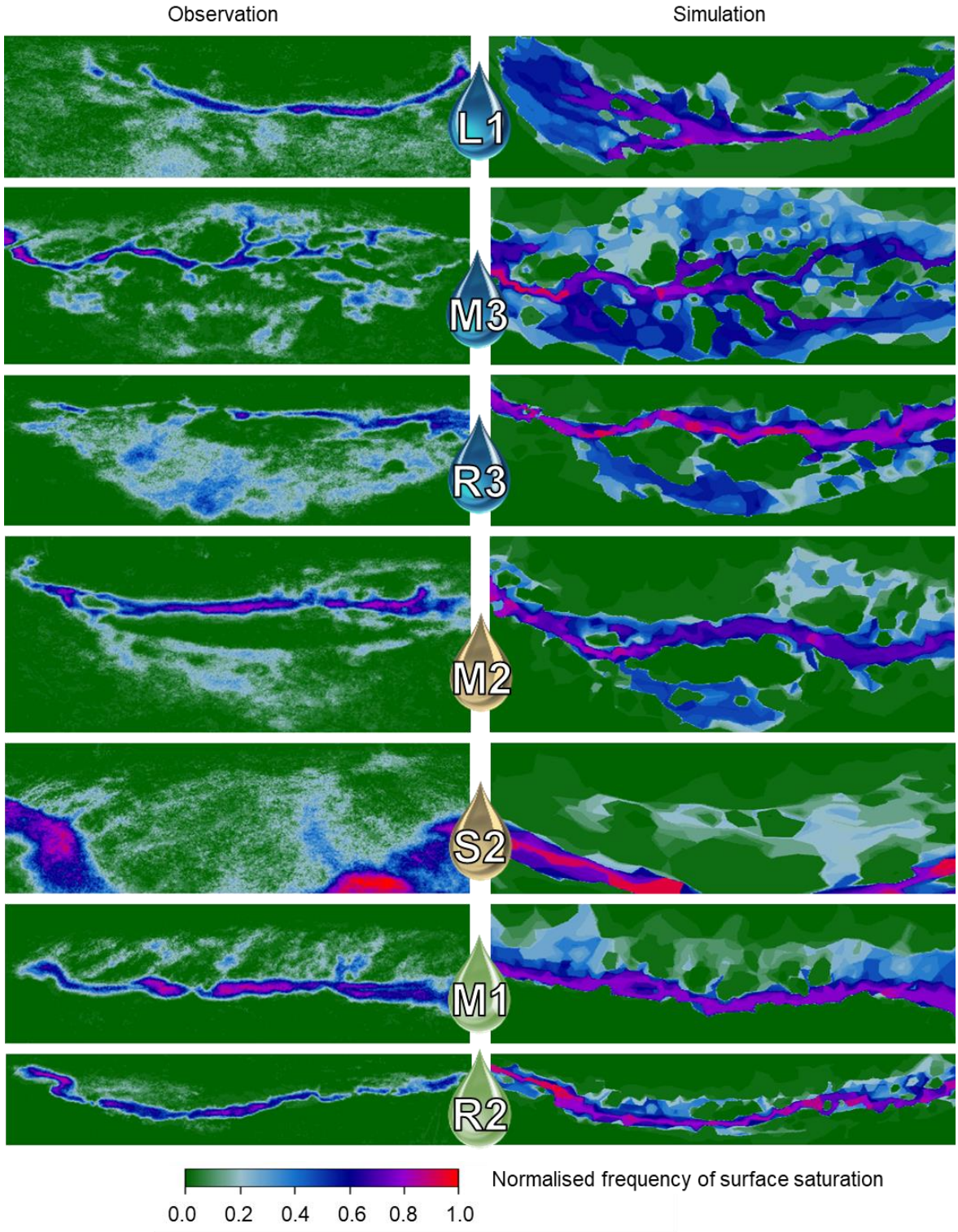


Figure 7: Observed (left) and simulated (right) frequencies of surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary, and the central stream (S2) (cf. Fig. 1). The maps were created by first counting how often the individual pixels 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.

No Comments.

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. We found some discrepancies between observed and simulated discharge, groundwater levels, and soil moisture 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 nonetheless argue that the match between the 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. 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 dynamics of surface saturation in the seven investigated riparian areas over different seasons and wetness conditions. Our work goes beyond previous studies 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 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 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 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. 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.

Page:18

Number: 1 Author: reviewer Subject: Insert Text Date: 2020-01-05 15:16:08

seven different riparian sites of the

Number: 2 Author: reviewer Subject: Note Date: 2020-01-05 15:51:18

I hope that Antonelli et al., 2019 is published by the time this paper is also accepted. Unlike with HESSD, papers under review at Hydrological Processes need to be cited as (in review) unless they are accepted. Please adapt throughout the manuscript.

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 (cf. Fig. 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 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. 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 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 tributaries, cf. Section 4.4). 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).

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 cannot compare our results directly with the results shown in Latron and Gallart (2007), since we evaluated absolute discharge and normalised saturation, while they evaluated connected saturated areas in percentage of catchment area, but discharge normalised 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 normalised discharge and absolute saturation (Fig. 8a), and once with absolute discharge and normalised 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

Page:19

Number: 1 Author: reviewer Subject: Note Date: 2020-01-05 18:08:34

I assume your conclusions are drawn from Figure 5 where you plot normalized saturation as a function of dischage. However - and I apologize that I had overseen this during the first round of reviews - the x-axis in Fig 5 are log. While this is useful to plot in the manuscript it is misleading in terms of the interpretation that normalized saturation would follow a power law relation with dischage. I ask you to plot discharge on a linear axis so you can better interpret the functional relation. I would assume it is not a power law. Please adapt your interpretation throughout the text and abstract accordingly.

Number: 2 Author: reviewer Subject: Note Date: 2020-01-05 15:52:02

As Antonelli et al., 2019 is not yet accepted your argument would be stronger if you could cite other published papers that suggest the same.

Number: 3 Author: reviewer Subject: Note Date: 2020-01-05 16:01:46

quantitatively the variability between the normalized saturation - dischage relations are likely smaller that the variability of saturation-dischage rations between different catchments (check difference in the range of y-axis). I liked your previous focus on the similarity of shape better.

Number: 4 Author: reviewer Subject: Note Date: 2020-01-05 16:11:22

linear - correct! But only because you are plotting your data in log-log space!

Number: 5 Author: reviewer Subject: Insert Text Date: 2020-01-05 16:12:33

in the Weierbach

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 intra-catchment variability.

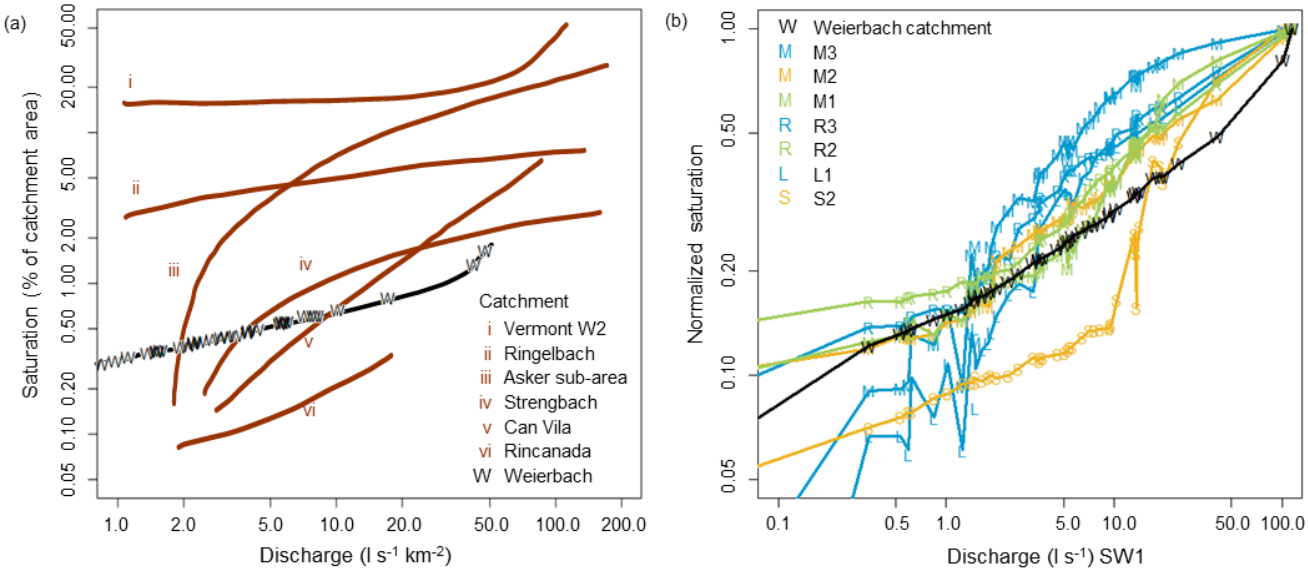


Figure 8: Simulated relationship between discharge and surface saturation of the entire Weierbach catchment (marked with W) in comparison to (a) the relationships observed in other catchments (Figure modified from Latron and Gallart (2007) and (b) the relationships simulated for the seven investigated 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.5-7).

5.3 Spatial patterns of surface saturation occurrence

The observed spatial patterns of surface saturation occurrence were reproduced well by the simulations 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 sufficient detail, it is more likely that the simulated surface water extends over a large area instead of being confined to occurring in topographic depressions and thus the model 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 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 suggest maximum extents around 10 % (Ambroise, 2016; Grabs et al., 2009; Güntner et al.,

No Comments.

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 studies that analysed the extent of surface saturation without considering microtopography in larger catchments (1 - 100 km^2) often simulated similar or smaller maximum extents of surface saturation than observed (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 extent of surface saturation reached a maximum of 1.6 % of catchment area during the very wet conditions in winter 2017/2018 (cf. Fig. 4c). This simulated maximum extent of surface saturation is small compared to the estimation in other simulation studies, but 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 the maximum saturation within the individual areas was overestimated compared to the observations (cf. Fig. 5). 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. 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 number of saturated pixels in the TIR images was thus lower than in the model images, while the minimum possible extent of saturation was not affected. This could also explain why overestimation of total 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. 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 left tributary (L1), where saturation was simulated at the clearly wrong side along the stream.

5.4 Spatial patterns 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. Instead, we assume 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 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

Page:21

Number: 1 Author: reviewer Subject: Note Date: 2020-01-05 16:42:24

but 0.1 km2 = 100 ha which is more than twice the Weierbach

Number: 2 Author: reviewer Subject: Insert Text Date: 2020-01-05 16:40:45

in our study (?)

Number: 3 Author: reviewer Subject: Replace Date: 2020-01-05 17:03:44

how often

Number: 4 Author: reviewer Subject: Insert Text Date: 2020-01-05 18:01:34

from upstream

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 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 and mismatch between the simulation results and observations, we could identify groundwater exfiltration and microtopography as 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 of surface saturation extent 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 the catchment. Still, we 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 relate to differing 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~~. The model could not fully reproduce the observed varying persistence of surface saturation and hysteretic relationships between surface saturation extent and catchment discharge. Nonetheless, the simulation results demonstrated that the shape of the relationship between surface saturation and discharge for different riparian areas within a catchment can be as variable as it has been observed between different catchments with different topographical and morphological conditions.

The spatial occurrence of surface saturation differed between and within the seven investigated riparian areas, which we could mainly 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 ~~a~~ locally varying frequencies of surface saturation might reflect a locally varying relevance of different water sources. Since the model could reproduce the observed frequencies, the model may be used in a future study to analyse such a potential mixing of different water sources and their variation in space and time.

Page:22

Number: 1 Author: reviewer Subject: Insert Text Date: 2020-01-05 17:03:15

model

Number: 2 Author: reviewer Subject: Note Date: 2020-01-05 17:15:19

Previously you stated that you did not see hysteretic behaviour in your data(and you give plausible reasons why e.g. the temporal frequency of your campaigns) . So I would write a bit differently ....

Number: 3 Author: reviewer Subject: Replace Date: 2020-01-05 17:17:56

. H

Number: 4 Author: reviewer Subject: Replace Date: 2020-01-05 17:19:24

like reported from other experimental catchments were not resolved in our data. Also the model ...

Number: 5 Author: reviewer Subject: Note Date: 2020-01-05 17:20:59

see my comment of the variability between other catchments and within the Weilerbach and my suggestion to more focus on the shapes of these relations.

Number: 6 Author: reviewer Subject: Note Date: 2020-01-05 17:29:36

After your changes to this paragraph this sentence is no longer clear to me. Do you refer to observations only or to the differences between observations and modeling results?

Number: 7 Author: reviewer Subject: Strikeout Date: 2020-01-05 17:30:10

Number: 8 Author: rinderer Subject: Replace Date: 2020-01-05 17:30:24

T

Number: 9 Author: rinderer Subject: Replace Date: 2020-01-05 17:31:15

Overall

Number: 10 Author: reviewer Subject: Note Date: 2020-01-05 18:11:02

After you edited this last paragraph (in this round of review) your argumentation is contradicting. You say that the model was not capable to reproduce the spatial patterns of surface saturation occurrence equal well in all investigated areas and that this suggests that the model is missing some aspects. But then you say you want to used this model to learn more about something even more complex i.e., the sources of different water origines.

I would rather use this thought for an outlook and say something like:" If we get the models better in reproducing the spatial extent and frequency of saturation occurrence we might even be able to learn something about the origin of water (groundwater or stream water from upsteam) that causes saturation.

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-

No Comments.

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,

No Comments.

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.

No Comments.

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

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 (b.glaser@lmu.de)

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.

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. In general, such surface saturated areas arise from 1) water ponding on the surface due to precipitation intensity exceeding 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 flowing into the floodplain (e.g. Megahan and King, 1985). When 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 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 processes affects 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 allow detailed investigation of surface saturation at any desired location and time. This goes far beyond the information that can be gained by 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 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 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 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 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 use 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 to identify 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 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 headwater by accounting for ~~a~~ layering of the subsurface, while spatial heterogeneity was only represented by microtopography and 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 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 results with observed surface saturation characteristics from TIR imagery, we ~~aim to~~ use the model as learning tool and address the following research questions:

1) To what extent can we reproduce the observed intra-catchment variability of surface saturation characteristics with a rather homogenously set up ISSHM?
2) What key controls for the intra-catchment variability of surface saturation characteristics can we identify based on the match and mismatch 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 pattern of surface saturation occurrence,
iv) the spatial pattern 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 tributary is situated 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 trees and covered with ferns, moss, and herbaceous plants. Siltic, sceletic Cambisols developed from Pleistocene Periglacial Slope Deposits are shallow and highly-permeable 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 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 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 weirs, installed at the outlet of the catchment (SW1, Fig.1) and upstream of the confluences of the three 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 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 at 10 cm depth was monitored with water content reflectometers (CS616, Campbell Scientific, 30 min logging intervals) at five locations transecting the riparian zone of the stream source area of the middle tributary (Fig. 1, 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 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 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 0.9[illegible] = 0.81) with the measurements close to the study site in the years before.

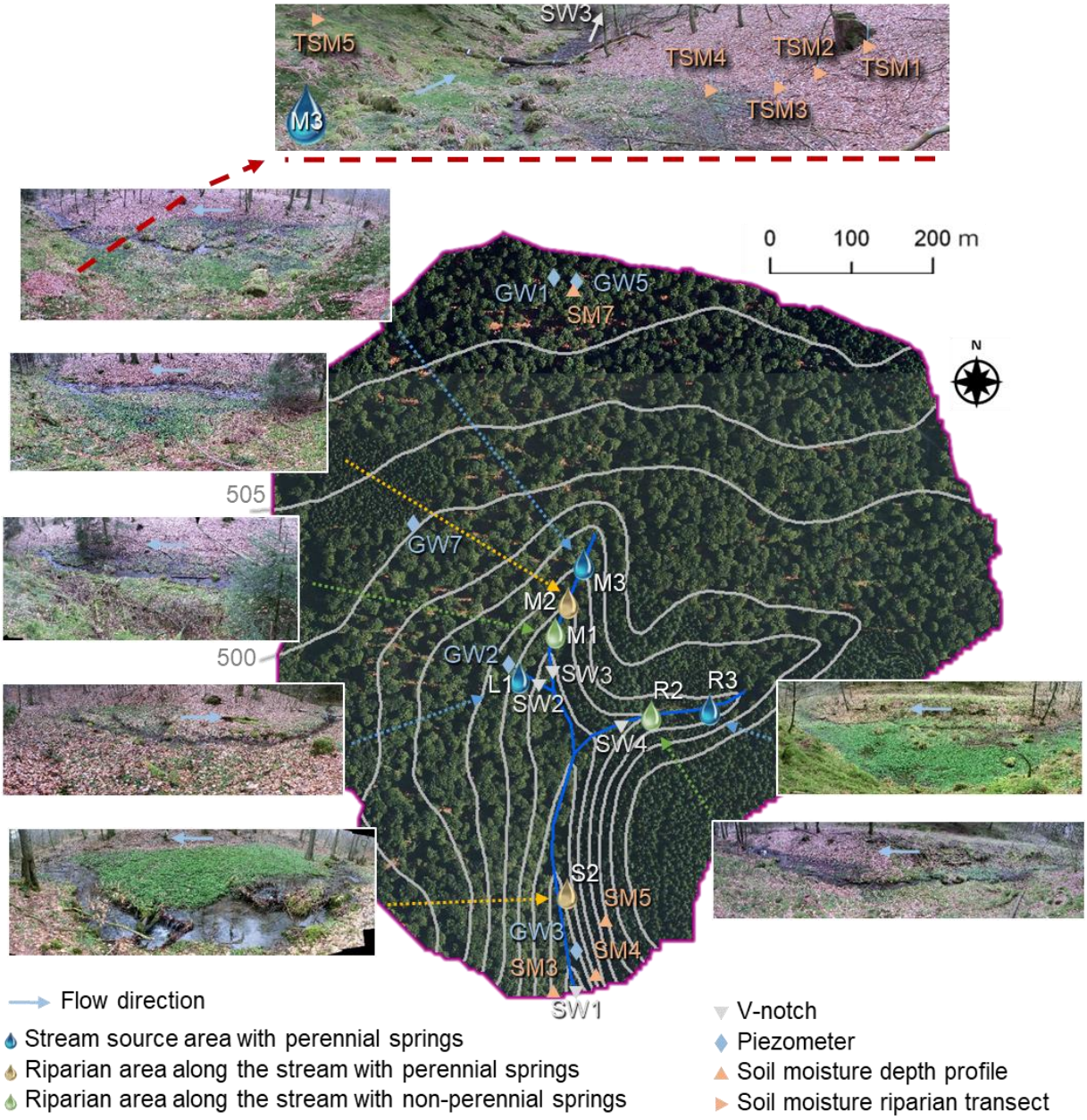


Figure 1: Weierbach catchment with the locations of the installed v-notch weirs for measuring discharge (SW1-SW4), piezometers for measuring groundwater levels (GW1-GW3, GW5, GW7), soil moisture sensors for measuring soil moisture in different depth profiles (SM3-SM5, SM7) and a riparian transect (TSM1-TSM5), 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 (geoporta[illegible]).

2.2 Surface saturation

Here, we define the surface as saturated as soon as water is standing or flowing on the ground surface (Glaser et al., 2018). This includes water bodies such as 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 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). 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 can be classified into three different categories (cf. Antonelli et al., 2019): i) stream source areas with perennial springs (L1, M3, R3, blue icons Fig. 1), ii) areas along the stream with perennial springs (M2, S2, yellow icons Fig. 1), and iii) areas along the stream with ~~non-perennial~~ springs (M1, R2, green icons Fig.1).

We mapped the surface saturation in the 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, 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, slight shifts in perspective between the panoramas of different dates were inevitable and the different features (e.g. streambed, stones, trees) in the images the images did not always overlap exactly pixel by pixel.

The locations of surface saturation (including the stream) were identified from the TIR panoramas based on the acquired temperature information. Each pixel in a TIR panorama was ~~assigned~~ to be saturated or unsaturated based on the temperature range of locations that were obviously saturated according to field observations and visible 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, 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 total number of analysed images per site ranged 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 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. 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 normalised the resulting frequency numbers by the total number of TIR images analysed for that area. The resulting maps of normalised saturation frequency suggest that very few pixels were always saturated (i.e. reaching a

normalised frequency of 1). Field experience and the analysis of individual TIR and visual images showed that in reality, surface saturation was permanent at more places than indicated by the frequency maps. The reason for this artefact is that the perspective and distortion 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 for instances where the extent of surface saturation was ~~narrow~~. As a result, the generated saturation frequency maps are blurred. Nonetheless, the maps of normalised saturation frequencies are very useful to quickly detect where surface saturation occurs more ~~and~~ less frequent within an area and 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 address 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 already 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 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, 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).

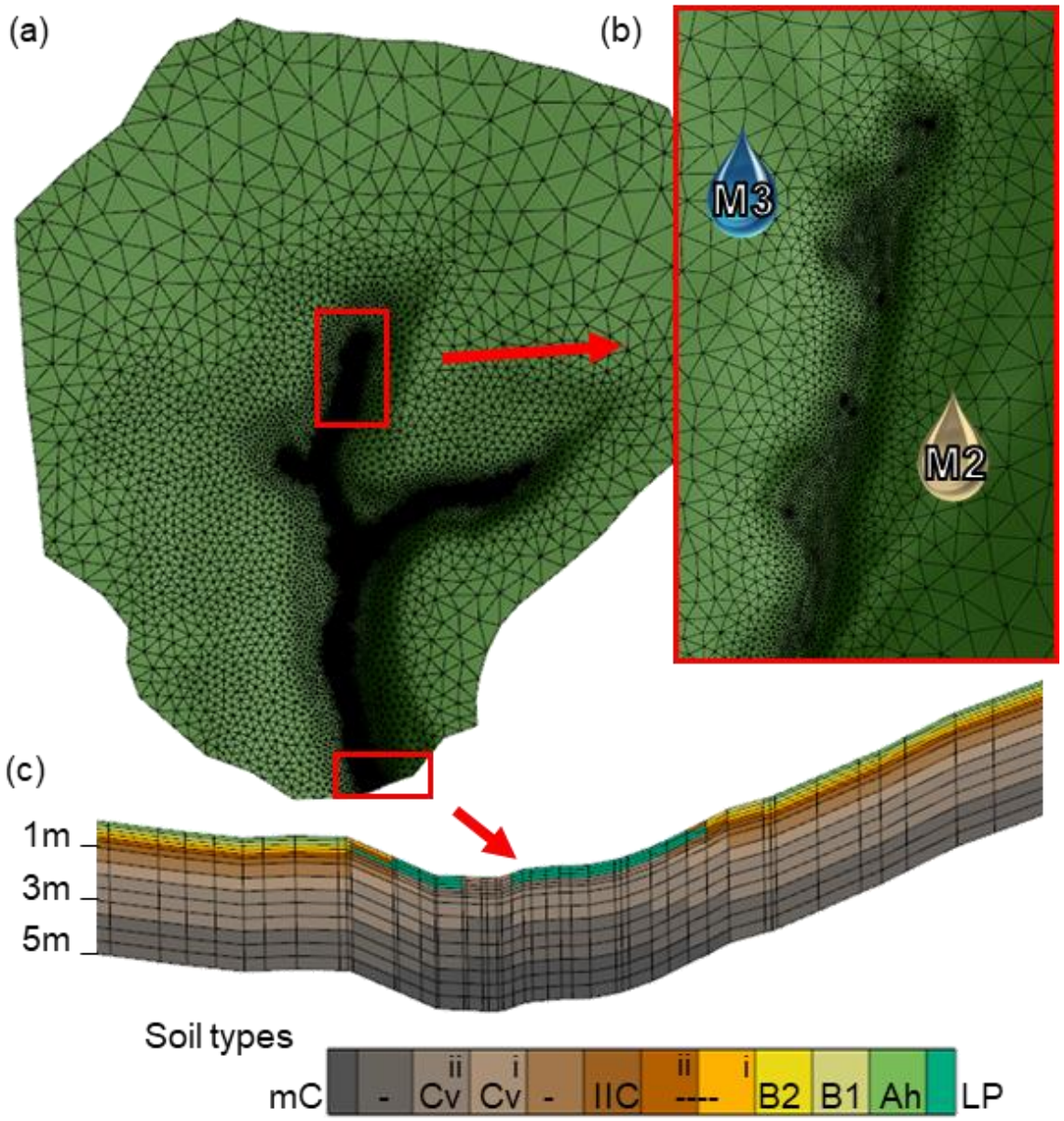


Figure 2: Setup of the model mesh (a) with a zoom on the fine horizontal resolution in the riparian areas and the streambed (b) and a vertical cross section through the stream valley and adjacent hillslopes (c) showing the vertical discretization and assignment of different soil properties (cf. Tab. 1). LP = riparian Leptosol, Ah = topsoil, B1 and B2 =subsoil, IIC = basal layer, Cv = fractured bedrock, 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, mC), and transition layers between subsoil, basal layer, and fractured bedrock (Fig. 2c). A differing subsurface structure was implemented in the stream valleys because in this area the soil and basal layers are eroded and the outcropping fractured bedrock is overlain with organic, stagnic Leptosol (LP) ~~in the riparian zones~~ (Fig. 2c). We used the Mualem-van Genuchten soil hydraulic functions to 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 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, we added new parameter 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 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 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 outer edge of the surface domain, 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 to obtain realistic initial conditions. The actual simulation spanned 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 the model performance with discharge, groundwater level, and soil moisture ~~measured 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. ~~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 to ~~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 two depth profiles at each monitoring site.

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 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 and perspective. First, we transformed the simulated surface water depths for the days with TIR images into binary saturation maps of the entire catchment following the processing described above. Then, we converted 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 normalised surface saturation extent, estimate the Kendall correlation for the relationship between normalised surface saturation extent and catchment discharge, and generate maps of 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.

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 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 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 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). Pearson correlation coefficients indicated overall a good agreement between simulated and observed soil moisture dynamics (Tab. 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, while the simulation showed a decrease in water content during dry periods (TSM4, Fig. 3). 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 and mismatch 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 within small distances (e.g. measurements of water content in 10 cm depth at profile SM7, Fig. 3).

4.2 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. 4a). During the wettest conditions of the study period (winter 2017/2018), surface saturation also ~~occurred as~~ extension ~~of~~ the right tributary into the hillslope above the source area R3 (cf. Fig. 1). This simulated occurrence of surface saturation across the catchment is consistent with field evidence, 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. 4b) were very similar to the surface saturation frequency map of the catchment (Fig. 4a). The only obvious difference occurred in the area above the source area of the right tributary (R3), with a smaller extent of groundwater reaching the surface than the surface saturation extent.

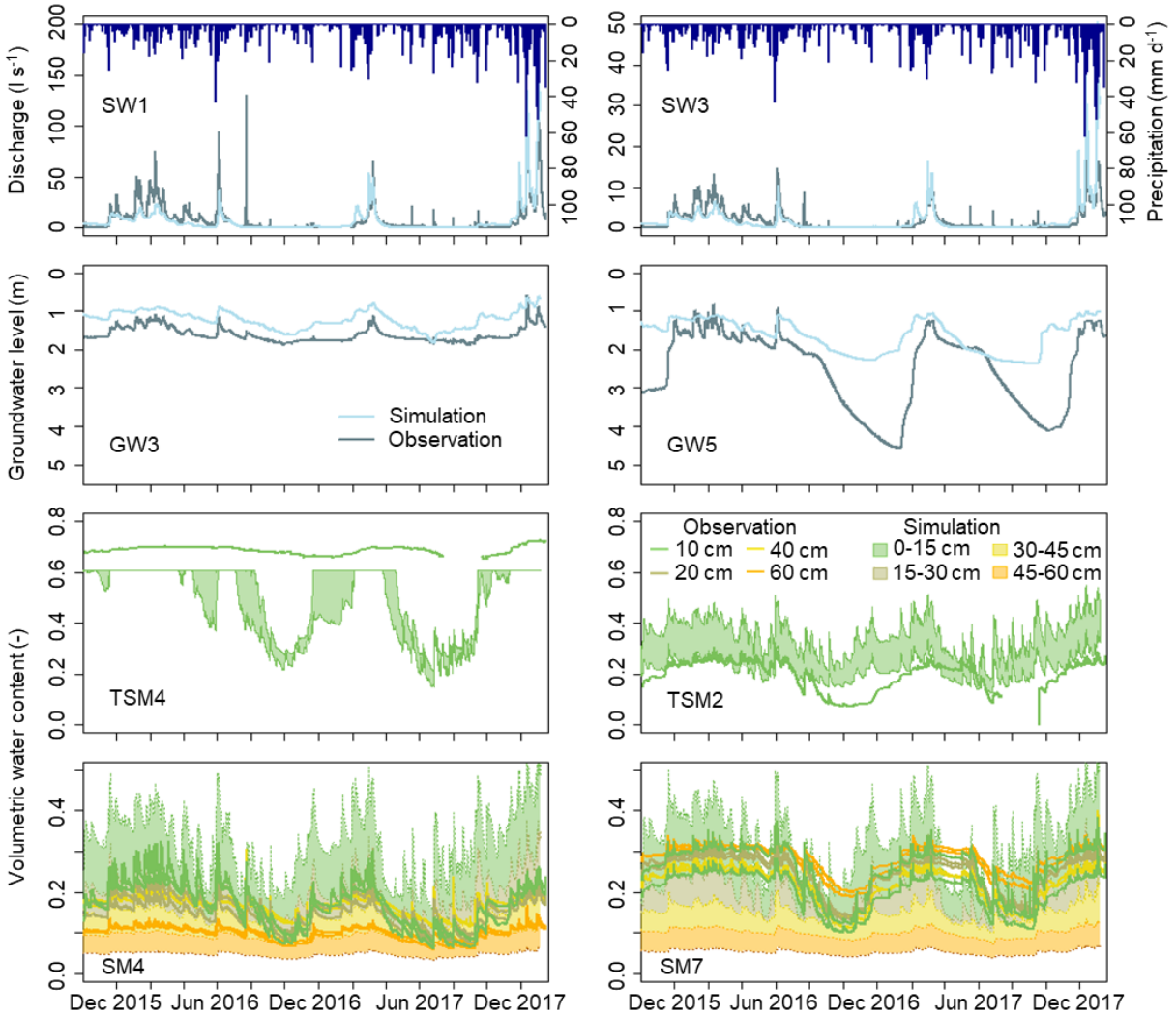


Figure 3: Simulated and observed time series of discharge, groundwater level below the surface, and volumetric water content. Colour bands indicate the possible span of simulated volumetric water contents in ~~the depths~~ between two model nodes. The time series ~~of~~ the observation locations (cf.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

Table 2: 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					

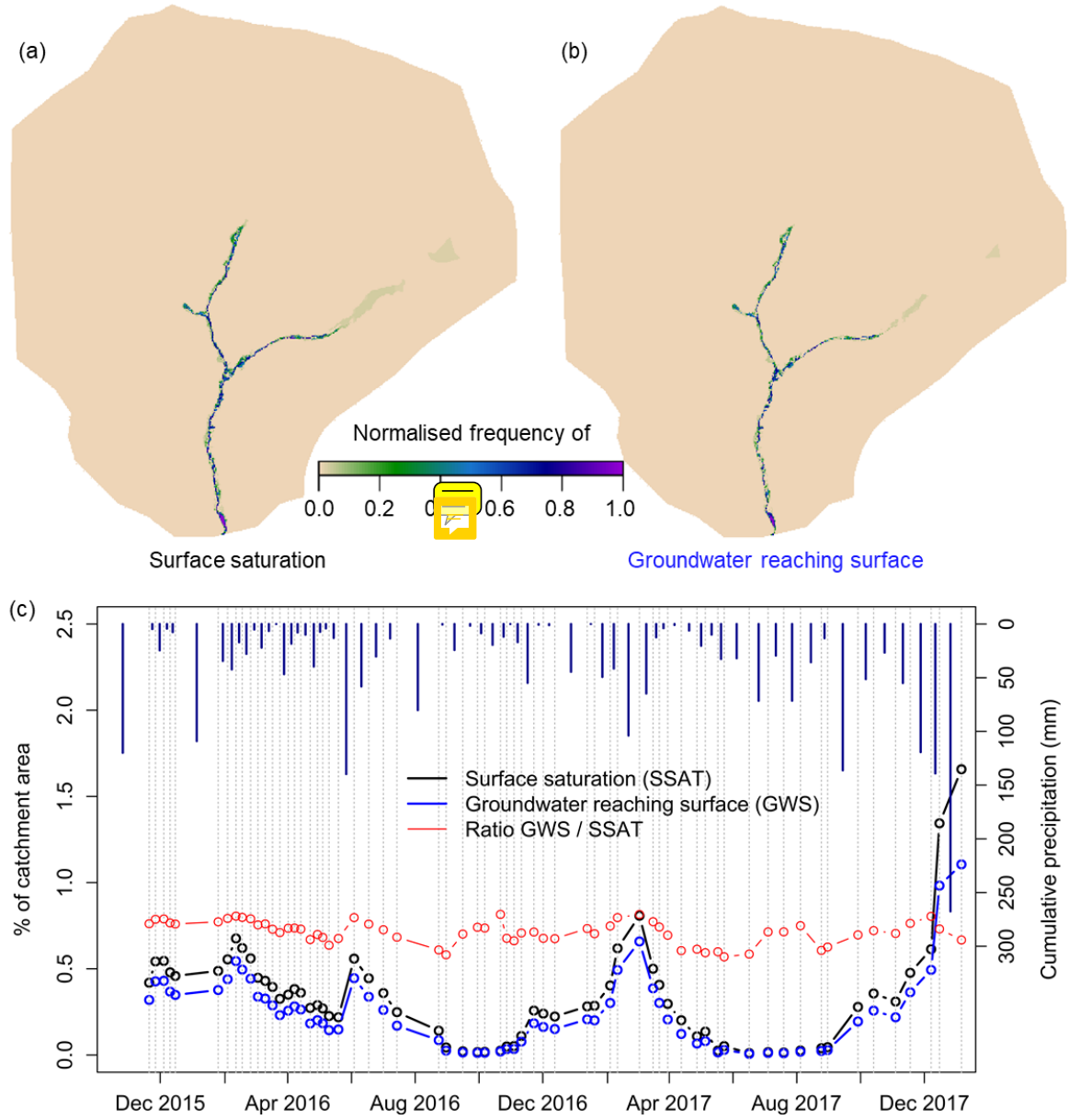


Figure 4: Simulated frequency maps (a, b) and time series of percentage (c) of surface saturation and groundwater reaching the surface in the Weierbach catchment. Surface saturation (a, black line in c) corresponds to a simulated water depth >10^{-4} m in the surface domain of the model. Groundwater 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. 4c). 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. 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 varied 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.

4.3 Temporal dynamics of surface saturation extent within distinct riparian areas

The time series of observed normalised surface saturation extent (Fig. 5, coloured lines) were similar for all seven investigated riparian areas and followed the same seasonal trend as 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. 5) 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. 5). Similar to the observations, the simulated dynamics of normalised surface saturation (Fig 5, black lines) followed the trend of the simulated discharge dynamic. The simulation showed a faster decrease and increase of the normalised saturation during dry periods compared to what was observed in most areas. However, simulated discharge also seemed to decrease and increase earlier than the observed discharge (c.f. Section 4.4). 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 normalised saturation was better for some areas (e.g. M1, R2, Fig. 5) than for others (e.g. S2, L1, Fig. 5).

The ~~dynamic~~ changes of ~~normalised~~ simulated saturation matched the ~~normalised~~ observations generally well, despite under- and over-estimation of the minimum and maximum absolute saturation for all areas. The minimum 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. 5). In addition, simulated normalised 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. 5). The overestimation was not ~~equally strong at~~ the seven investigated areas and as a result, the distinction between areas 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~~).

4.4 Relationship between surface saturation extent and catchment discharge for distinct riparian areas

Kendall correlation between normalised surface saturation and discharge at the outlet SW1 was > 0.60 for both the simulation and the observation ~~in~~ all riparian areas, indicating a monotonic relationship between surface saturation and discharge in all areas (Fig. 6). The simulated relationships between normalised surface saturation and catchment discharge resembled the observed relationships in terms of ~~value~~ range and shape (Fig. 6), although the observation data were distinctly more scattered 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 normalised saturation and its change relative to discharge was smallest in the source areas (L1, M3, R3). 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~~ normalised saturation, being highest for areas in the right tributary (R2, R3), followed by the areas in the middle tributary (M1, M2, M3), and L1 and S2.

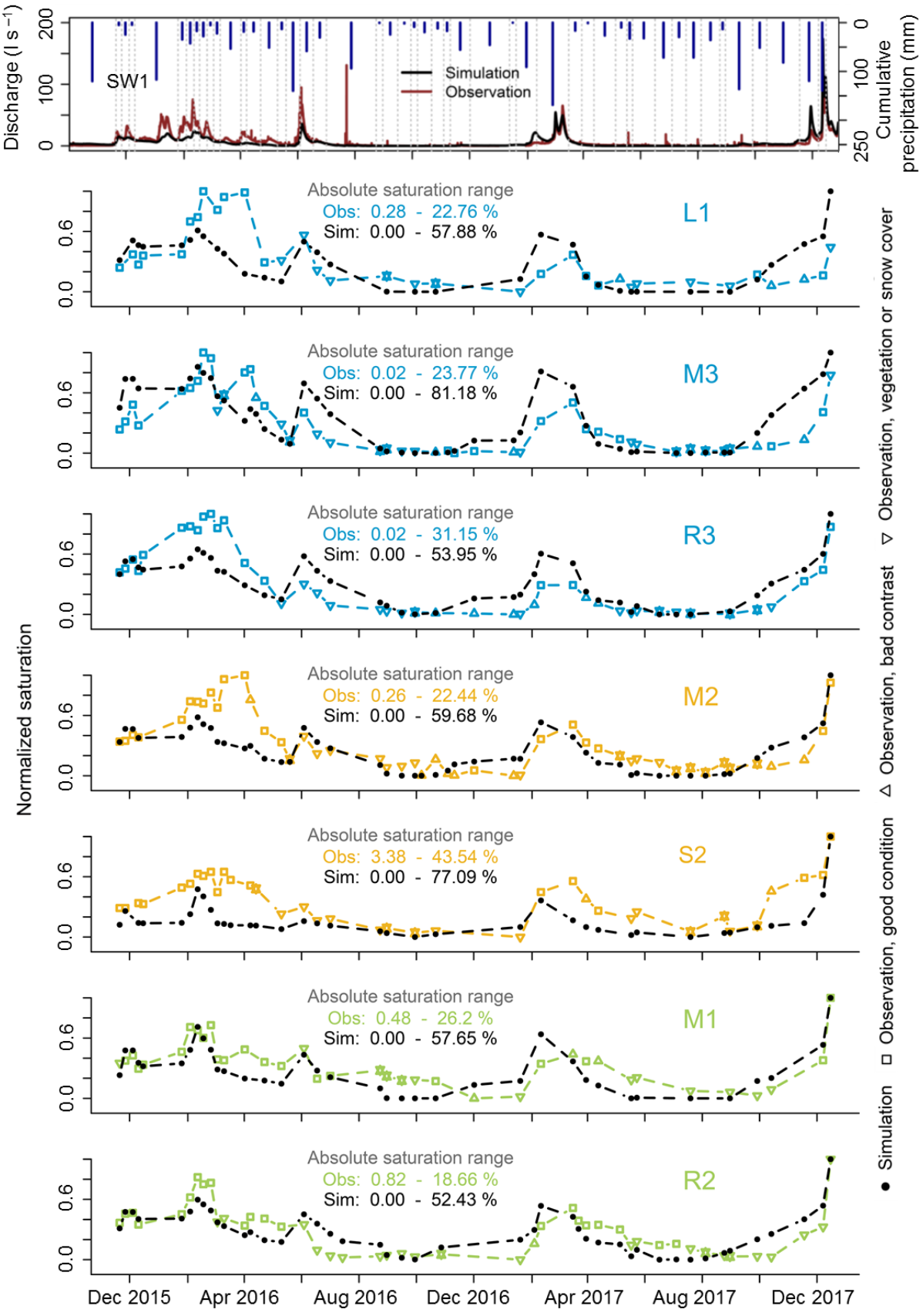


Figure 5: Time series of observed and simulated surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond to the colours of the icons in Fig. 1 and represent the different categories of 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 to facilitate the comparison to precipitation and flow conditions.

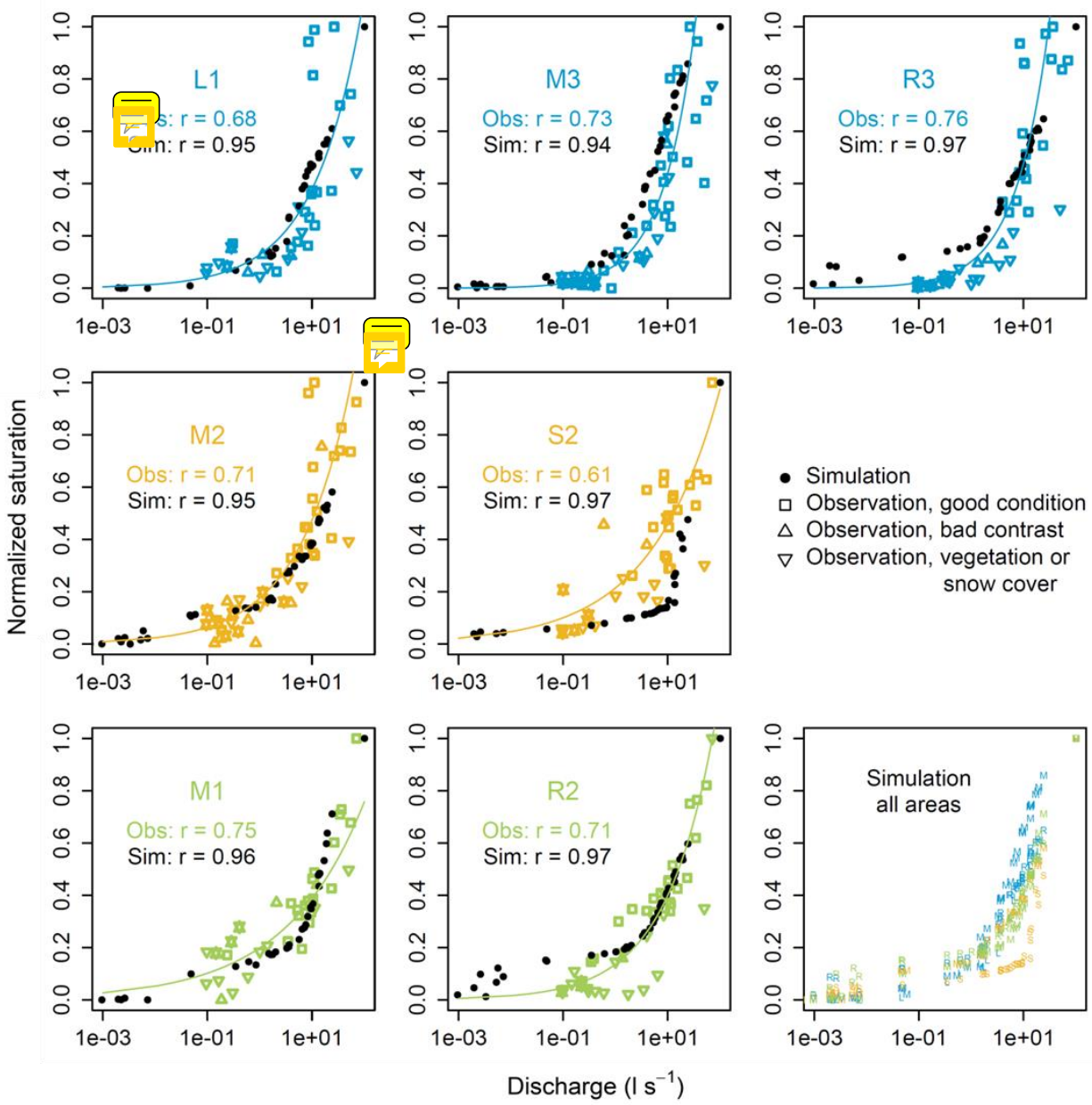


Figure 6: Observed and simulated relationships and Kendall correlations between normalised surface saturation and discharge at the catchment outlet SW1 for the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary and the central stream (S2). Colours correspond 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. To facilitate the comparison between the seven areas, the simulated data points from all seven areas are shown with the different colours and letters of the respective areas in the panel on the bottom right.

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. 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 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. 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. 7) than in others (e.g. R3, Fig. 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.

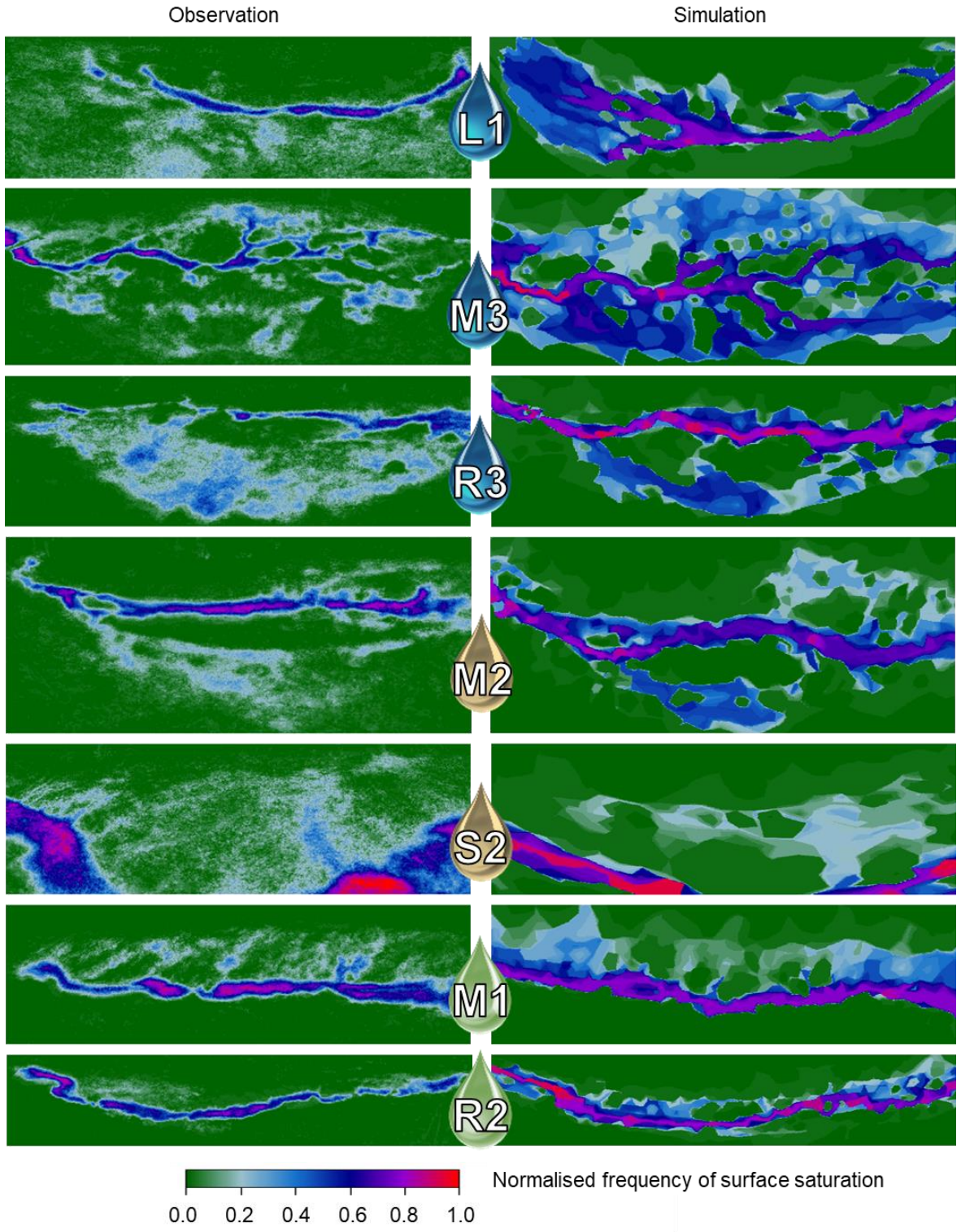


Figure 7: Observed (left) and simulated (right) frequencies of surface saturation in the seven investigated riparian areas along the left (L1), middle (M1-M3), and right (R2, R3) tributary, and the central stream (S2) (cf. Fig. 1). The maps were created by first counting how often the individual pixels 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.

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. We found some discrepancies between observed and simulated discharge, groundwater levels, and soil moisture 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 nonetheless argue that the match between the 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. 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 dynamics of surface saturation in the seven investigated riparian areas over different seasons and wetness conditions. Our work goes beyond previous studies 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 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 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 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. 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 (cf. Fig. 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 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~~ du[illegible] 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 [illegible]1 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. 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 of the observation ~~data~~ did not result from hy[illegible]tic behaviour, but from uncertainties in the TIR methodology. A good argument for a correct simulation of the dischar[illegible] 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~~l~~ 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 tributaries, cf. Section 4.4). 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).

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 cannot compare our results directly with the results shown in Latron and Gallart (2007), since we evaluated absolute discharge and normalised saturation, while they evaluated connected saturated areas in percentage of catchment area, but ~~discharge~~ normalised ~~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 normalised discharge and absolute saturation (Fig. 8a), and once with absolute discharge and normalised 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 intra-catchment variability.

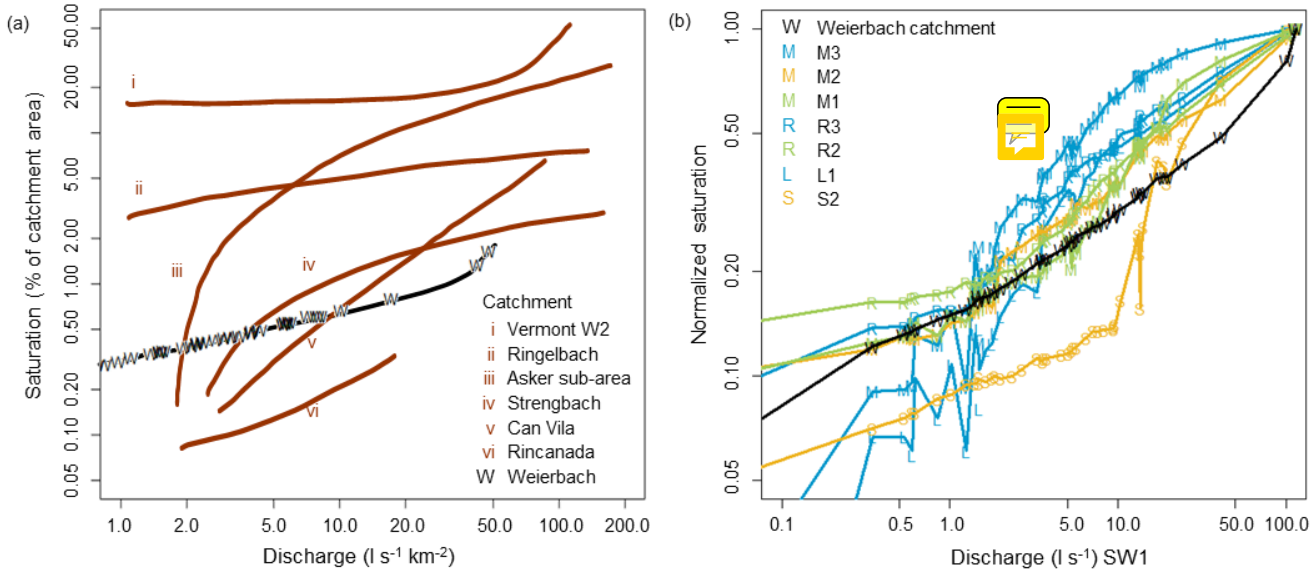


Figure 8: Simulated relationship between discharge and surface saturation of the entire Weierbach catchment (marked with W) in comparison to (a) the relationships observed in other catchments (Figure modified from Latron and Gallart (2007) and (b) the relationships simulated for the seven investigated 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.5-7).~~

5.3 Spatial patterns of surface saturation occurrence

The observed spatial patterns of surface saturation occurrence were reproduced well by the simulations 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~~ identify~~ing~~ 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). ~~S~~still, microtopography is not often considered in the simulation of surface saturation patterns.

When microtopography is not resolved in sufficient detail, it is more likely that the simulated surface water extends over a large area instead of being confined to ~~occurring in~~ topographic depressions and thus the model 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 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 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 studies that analysed the extent of surface saturation without considering microtopography in larger catchments (1 - 100 km^2) often simulated similar or smaller maximum extents of surface saturation than observed (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 extent of surface saturation reached a maximum of 1.6 % of catchment area during the very wet conditions in winter 2017/2018 (cf. Fig. 4c). This simulated maximum extent of surface saturation is small compared to the estimation in other simulation studies, but 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 the maximum saturation within the individual areas was overestimated compared to the observations (cf. Fig. 5). 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. 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 number of saturated pixels in the TIR images was thus lower than in the model images, while the minimum possible extent of saturation was not affected. This could also explain why overestimation of total 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. 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 left tributary (L1), where saturation was simulated at the clearly wrong side along the stream.

5.4 Spatial patterns 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 in 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. Instead, we assume 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 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 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 and mismatch between the simulation results and observations, we ~~could~~ identify groundwater exfiltration and microtopography as 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 of surface saturation extent 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 the catchment. Still, we 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 relate to differing 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. The model could not fully reproduce the observed ~~varying~~ persistence of surface saturation and hysteretic relationships between surface saturation extent and catchment discharge. Nonetheless, the simulation results demonstrated that the shape of the relationship between surface saturation and discharge for different riparian areas within a catchment can be as variable as ~~it~~ has been observed between different catchments with different topographical and morphological conditions.

~~The spatial occurrence of surface saturation differed between and within the seven investigated riparian areas, which we could mainly 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. Since the model could reproduce the observed frequencies, the model may be used in a future study to analyse such a potential mixing of different water sources and their variation in space and time.

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.