Response to Anonymous Referee #1

The study assesses the practicability of applying thermal infrared (TIR) imagery for mapping surface saturation dynamics. The experimental work was based on an 18-month field campaign, where the authors tried to try to outline under which conditions the method works best and what problems may occur.

General comments

I found the topic very interesting and I really enjoyed reading the paper. Although I do not feel qualified to judge the entire process of acquisition and post-processing of the images since it does not fall within my field of expertise (thus I remind the editor decision to the comments of another more expert reviewer) I see a very high potential for this technique especially its application to larger scales.

We wish to thank the reviewer for the positive feedback and we are pleased that the reviewer is interested in the technique and enjoyed reading the manuscript.

For this reason, my main comment/suggestion is to put an effort to expand the description of the potential transfer of this technique to larger spatial scales (for instance by using air-born space-born instruments) and the associated difficulties/simplifications/implications this transfer entails.

We thank the reviewer for pointing out this. We emphasised the potential scale transfer and related challenges by adding a paragraph on this in the discussion section (P13, L4-14).

Technical comment

The paper is really well written and clear. It is also well structure and organized so I do not have specific comments except the suggestions of inclusion of scale bars in the pictures that help the reader to understand the spatial dimension of the figures otherwise too difficult.

The images are taken with a low angle and the images are thus spatially distorted. This means that it is not possible to include scale bars that are valid for all of the image. In order to facilitate the perception for the spatial dimension we indicated the length of a distinct stream section within the images instead (with the exception of Fig. 2, since this would have overloaded the images)

Based on that my suggestion is the acceptance of the paper after minor revisions.

Response to Anonymous Referee #2

This technical not describes the opportunities, methodological considerations and challenges of applying thermal infrared imagery to map surface saturation. This technique shows great promise to understand the spatiotemporal dynamics of surface saturation and the hydrological processes that induce or are a result of these dynamics. Overall the manuscript is well written and articulates the challenges and opportunities well and is appropriate as a technical note for HESS. I would recommend publication with minor revisions with the main comments and technical edits provided below.

We wish to thank the reviewer for the assessment of the manuscript and for the valuable comments and suggested edits.
Main comments:

1. Overall the manuscript conveys a lot of information, but I struggle with the overall organization. Technical notes obviously are not full research articles, but I would still expect a similar format. Intro/Methodology/Results/Discussion/Conclusions. In this work the methodology, results, and discussion seem to overlap in some cases. There is no specific results section, so the findings are not clear before a discussion section begins rather abruptly. I would suggest the following organization:

1. Introduction 2. Methodology
-in fundamental principles it would be good to see the full equation for how to relate what is seen with TIR to absolute temperature. Will help in the communication of the challenges of this method and why for example emissivity and environmental conditions are important.

We included the equation for the Stefan Boltzmann law and for the radiometric corrections (P3, L30 - P4, L9).

-in image acquisition if the various challenges could have their own headings. Ie.Weather conditions, view obstruction, view angle… etc..

We added three sub-headings (Impact of weather conditions, Camera position, and Measurement artefacts during image acquisition)

-"4. Building saturation maps" is still a lot of methodology. Could it be incorporated in this section?

3. Results/Application examples

4. Discussion

5. Conclusions

As is the combination of methodology, results and discussion throughout makes the article feel muddled and at time repetitive even though the information is all very relevant.

We thank the reviewer for the suggestions on the structure. It is very important to us that the structure of the manuscript is easy to follow and comprehensible.
Yet, as the manuscript combines both review and own experimental work on how to use the TIR methodology (see objectives in the introduction) the organization in the 'classical' way is in our opinion not the best option.
We agree with the reviewer that the structure was a bit muddled and iterated the structure of the manuscript and improved various sections:

1. We integrated section 3 'Application examples' of the previous manuscript into section 2 (now 2.4).
2. We disentangled section 3 'Quantification of saturation through pixel classification' (former section 4 'Building saturation maps') and subdivided it into two parts, 'Methods for generating binary saturation maps' and 'Comparison of methods for generating binary saturation maps for TIR images'.

The reason behind this revision of the structure is that the manuscript contains two parts that evaluate different aspects of the TIR technology. The first part (now section two) deals with the identification of saturation. The second part (section 3 in the new manuscript) assesses approaches on how to generate binary saturation maps from the images. Each of the two sections consists of the methodological aspects but also on a review of current literature. We round the paper with a general discussion on both key sections and the conclusions. This structure is also explicitly outlined at the end of the Introduction of the new version of the manuscript (P3, L21-25).

2. Generalize. Portions of the article are very specific to the software and camera that were selected for this study. To make this more relevant to a wider audience certain sections could be removed or be made to be more generalized. Ie(Page 5 line33-34, Page 6 line 15-24, and page 7 line 10-19).

We agree that some parts are very specific and we generalized especially the descriptions in section 2.3 (see also our reply to the comment of Reviewer #3 on P6, L 13-20). Yet, we did not remove all parts that could not be completely generalized (i.e. P5, L33-34, now P7, L3-5), since we think that the information contained in these parts is also of interest for users employing different camera types / software.

3. The influence of difference of surface emissivity's were only very briefly mentioned. In TIR, depending on what is in the scene, the differences in emissivity's can be important to the reported temperatures- will have implications for absolute temperature values and gradients across the image. I would expect more of a discussion at least so, even if this article doesn't do it, others can incorporate these important corrections in their own work. An example for this in another paper can be found in: Aubryâ˘A˘RWake,Caroline, et al. "Measuring glacier surface temperatures with groundâ˘AR˘ based thermalinfrared imaging." Geophysical Research Letters 42.20 (2015): 8489-8497.

We put more emphasis on it in the revised manuscript (P6, L34 - P7, L3).

Specific comments:

Throughout: please use an oxford comma We now consistently use the oxford comma
Page 2. Line 6-10. Unclear sentence structure We reformulated the sentences.
Page 2 Line 16. "up to now" -> to date Done
Page 2 Line 28-35. Paragraph is muddled. Please improve structure We rearranged the paragraph
Page 3 Line 1-2: sentence is awkward We rewrote the sentence.
Page 3 Line 13: "Yet," remove Done
Page 3 Line 18: "allow to obtain an areal picture"- please rewrite as this is awkward Done
Page 3 Line 29-30: please define surface saturation more clearly on its own as this is critical to the entire paper. We removed the first half of the sentence to make the definition more stand alone.
Page 4 Line 3-6. Long complex sentence. Breakup. "or" -> "of" Done
Page 4 Line 7: "as expressed from" -> relative to Thanks for that suggestion
Page 5 Line 14: clarify how one can still observe ground temperatures even if there is vegetation. We clarified that this depends on the density of the vegetation cover.
Page 6 Line 5. Define what image vignetting is in this situation and why it is a challenge We added two sentences to address this.
Page 6 Line 29-34: discussion of consistent temperature scale is redundant We removed it

Page 7 Line 21-25. Georectification of terrestrial photos has also been extensively worked on by Corrpio and Harer and should be cited here as well as examples:

Härer, S., M. Bernhardt, J. G. Corripio, and K. Schulz. 2013. "PRACTISE – Photo Rectification And ClassificaTIon SoftwarE (V.1.0)." Geoscientific Model Development 6 (3): 837–848. doi:10.5194/gmd-6-837-2013.

Corripio, J. G. 2004. “Snow Surface Albedo Estimation Using Terrestrial Photography.”International Journal of Remote Sensing 25 (24): 5705–5729.doi:10.1080/01431160410001709002.

We added the two references in the manuscript.

Page 8 Line 2-4. Can the usability of an image be related to any metrics that would be helpful for fieldwork planning? This would be very useful information from a practical fieldwork perspective- help improve fieldwork efficiency.

Sadly, we cannot relate this to any metric. The applicability depends on the combination of various conditions affecting the image quality as mentioned and explained in section 2.2. In addition, it depends on the intended purpose up to which point an image can be used.
Fieldwork efficiency can be improved by avoiding such unfavourable conditions, but it is not possible to predict with certainty if the temperature contrast will be good enough.

Page 9 Line 28-29: please clarify as I’m unclear what the 90% means. We clarified the formulation (P11, L11-12).

Page 12 Line 21. What does the “(non)-“ add to this statement. Confusing as is.
We deleted “(non-)permanent”

Figure 4. The b scene with snow makes me wonder about how much the snow on the ground combined with the low camera angle is obscuring saturated area. Perhaps discuss this as a challenge in the article.
Snow can indeed obscure the saturated area. We mentioned snow as a possible view obstruction and added two sentences to further discuss the effect of snow in the revised manuscript (P6, L12-14). If the amount of snow is low, the saturated areas mainly stay uncovered (due to a warmer temperature of the water and thus a fast melting). This is the case for Figure 4b. If the snow cover is thicker, saturated areas will be covered and the snow surface has to be interpreted as the new ground surface.

Figure 7. Please use upper case on first characters of axis labels and put percentage into a unit. “percentage of saturated pixel” is unclear to me. Is this some sort of a cumulative distribution?
We changed the axis labels as demanded. We also clarified in the labels that ‘% of saturated pixels’ is equivalent to all pixels having a higher/lower temperature than a temperature range threshold. So yes, the plot shows the cumulative saturation distribution.

Response to Anonymous Referee #3

HESS manuscript HESS-2018-334 presents a technical note describing the use of thermal infrared imagery to map surface hydrological saturation. While similar techniques have been previously applied in a limited sense for mapping surface saturation, the manuscript is novel in that it attempts to review and develop best-practice processes for TIR-based saturation mapping. The manuscript is general well-written and thorough and contains a range of good advice. It will therefore be of interest to a broad cross-section of HESS's readership. I have nonetheless included a range of relatively minor comments/suggestions that should be addressed or clarified. Provided the authors are able to make these revisions, I would support its publication in HESS.

We wish to thank the reviewer for the assessment of the manuscript and for the helpful and detailed comments and suggestions.

—

General comments

- Although the article is generally well written and flows very nicely, some sentences and uses of grammar might appear a bit 'clunky' to native English speakers. This is not to say that the quality of the English isn't already very good, but it might nonetheless be worth passing it to a native English speaker for a quick check.

We used our internal institutional language editing.

- The 'Building saturation maps' section is quite long and brings up image processing techniques that are already well-established in the remote sensing literature. It can therefore be shortened. It would also benefit from more consistent referencing to existing image processing/classification/ segmentation literature.

We agree that the section is rather long and includes techniques that are well-established in remote sensing literature. We should of course not present the obvious or repeat findings from other fields. Yet, we think that the given information is necessary to understand the approaches, especially for readers with a different background (e.g. field hydrologists). These two aspects need to be balanced as one purpose of the work is to establish these methods more in the field of experimental hydrology. Eventually, we did not considerably shorten the section. Instead, we improved the structure of this section by dividing it into two subsections: '3.1 Methods for generating binary saturation maps' and '3.2 Comparison of methods for generating binary saturation maps for TIR images' (see also our response to Referee #2 concerning the structure). We think that this will allow for a remote sensing expert to decide if especially the first part of the section can be skipped for reading. In addition, we followed the reviewer's advice and included more references from the remote sensing literature in the first part of the section to demonstrate the common use of these approaches at other scales.

- Although I appreciate that it was not the purpose of the MS, it would still be nice to see some validation of the saturation maps. Do you have any 'squishy boot' data that you can present to validate this data? Or soil moisture data or similar? If not, it would nonetheless be nice to add a section (maybe a paragraph of text) detailing a) the importance of thoroughly validating the TIR data and b) potential validation methods.

We basically used our observations from the field and the respective VIS images as ground truth / validation. We included a separate paragraph to explicitly state this in the 'Fundamental principles' section. In this paragraph we also discuss why other data than visual observation are difficult to be used as ground truth (P5, L7-14). We also added a sentence on the ground truth data used for comparing the different pixel classification approaches (P10, L26-27).

—

Abstract

- Consider adding 1-2 sentences at the start of abstract explaining surface saturation dynamics and their importance Done
P1 L18: Unclear what you mean by 'intuitive character'. Can you rephrase? Done.

P1 L19: No need for comma after ‘methodological principles’. We removed it

—

Intro

P2 L1: Change ‘albeit’ to ‘despite’. Done

P2 L9: Change location of ‘only’: ‘Hydrometric measurements ONLY have the potential to monitor…’

We rephrased the sentence to make it shorter and clearer (cf. comment of Referee #2) and don’t use ‘only’ anymore.

P2 L18-19: This sentence is factually incorrect - there are numerous examples of research using drones equipped with near infrared cameras to do things like NDVIs. Please delete/re-think this statement.

Indeed, this was not well formulated and we changed it. Our intention was to refer only to research using NDVI/NDWI for mapping surface saturation. We are not aware of any study that does this with a drone or ground-based and we rephrased the sentence accordingly. If the reviewer is aware of a study using drones to map surface saturation with NDVI/NDWI, we would be happy if (s)he could provide us a reference for this.

—

Acquisition of TIR images for mapping surface saturation patterns

P4, L20-22: In my experience, it’s less the similarity between air and water temperature that makes identification of waterbodies difficult to detect; it’s more when the ground surface temperature and water surface temperature are the same. This means that the ‘detectability’ of surface water using TIR is often a function of time of day (ie. During summer, water will be cooler than land in the morning, but a similar temperature in the late afternoon/early evening). I know that you’ve alluded to this in lines 22-26, but it might be nice to clarify this point.

We agree and rephrased the sentence to clarify the point on the temperature similarities. The link to the time of day is mentioned at the end of the section (P5, L32-P6, L2).

P4 L30: Is ‘sun memory effects’ an accepted term? Consider rephrasing. We replaced it.

P5 L2: Can you elaborate on why cloudy conditions are better? In my experience, TIR seems to produce better data when there are moderate amounts of cloud (not too clear, not too cloudy, eg. high level cirrus clouds), but persistent low level stratus-type cloud cover can cause reflections that are pernicious in TIR imagery.

We agree with your experience. However, we think that this may be overly complex to predict. Generally, we advise cloudy conditions in order to avoid the effect of sun. This is specified in the revised manuscript (P6, L2-3).

P6 L1-3: I believe that the explanation for these false ‘negative’ values is that during clear sky conditions (if the water is sufficiently still), the water acts like a mirror and reflects the clear sky. However, because the emissivity value at the camera will be set to that of water (0.97) and not of air/sky, the resulting temperature value reported by the camera is incorrect. You get similar effects when filming very reflective surface such as aluminium using TIR cameras.

We added one sentence about the possible explanation (P7, L6-8).

P6 L10: Change ‘shades‘ to ‘shadows’. Done

P6 L29 – P7 L9 and Figure 3, stage 1: I do not like the proposed technique for colour transformation prior to panorama creation. Personally, I feel that these ‘colour transformation’ type techniques are sub-optimal, and by converting the images to a simple 8 bit raster, you may lose contrast in the dataset that could be important.

A better way to do it would be to convert the radiometric TIR images to floating-point TIFF files (which preserve 'raw' temperature data), and then create the panorama using these files. This can be done with both of the cameras used in this study (FLIR b425 and T640). Furthermore, if working with videos, it is better to acquire them as .SEQ or .FCF files (essentially sequences of radiometric TIR images) rather than 'conventional' video formats (eg. MP4 files), as these a) preserve temperature data and b) are uncompressed. In my experience, the resulting panoramas are of better quality and have the added benefit of preserving 'real' temperature data.
I do not have a problem with the method you have used, per-se, for the purposes of this manuscript. However, you should not promote this method as the 'best practice' technique for converting TIR images. I would therefore simplify this section of text and Figure 3 to simply say something like 'convert TIR images to a file format that can be used by panorama software'.

We carefully checked the formulations and given information of this section to ensure that it does not read as a 'best practice' but rather as one possible approach. We acquired the images and the videos as radiometric TIR images (thus as .SEQ files for the videos). However, we do not know a standard software that can read these radiometric files for creating a panorama. Therefore, we exported the images and videos in a more standard file format (JPEG, .wmv). We exported the images as grey-scale images / video frames, knowing which colour value corresponds to which temperature value, in order to retain the temperature information. Relying on floating-point TIFFs instead did not appear to us to be more advantageous in terms of contrast/detail and practicability than the grey-converted images. This might depend on specific image characteristics and we are open for further discussions and exchange on this.

P6 L13-20: Similar to the above, the findings here are specific to the panorama software used, and are not necessarily 'best practice'. I would therefore simplify this text, and just talk about the mosaicking process in general terms, rather than talking about the pros/cons of different software packages and using video vs. using still images.

Referee #2 pointed out as well that some parts of the manuscript (including this one) should be generalized. We simplified and shortened this section to account for this.

—
Application examples

P8 L15-16: How can you be sure that this is groundwater exfiltration? Could it not just be runoff occurring from a terrace above the stream?

We think the reviewer's suggestion of adding the VIS images to Figure 5 will help to clarify this. There is no 'terrace' above the stream.

—
Building saturation maps

- This section is quite long, and a lot of the advantages/disadvantages of the image classification/segmentation approaches are common to all types of single-band imagery (not just TIR data). These approaches are thus well established in the remote

sensing literature (see histogram thresholding, segmentation using k-means classification, etc etc), and you probably don't need to go into such detail. You could therefore probably shorten this section by around 50% and just preserve the key findings. It might also be nice to include more references to image processing/segmentation/ classification from the remote sensing literature.

See our reply to the second general comment.

—

Discussion

- It would be nice to see some discussion of future work. For example, it would be fairly simple to combine TIR, VIS and NIR data to create multispectral images, thus allowing for advanced image classification procedures to better map connectivity patterns. This would help further improve the quality saturation maps by combining the advantages of these three approaches.

- Furthermore, you could also talk about using new computer vision techniques (eg. deep learning) to improve classification and thus generation of saturation maps from TIR imagery.

Both are good and interesting points and we took them up at the end of the discussion. (P14, L2-6).

—

Figures

Figure 3: See my comments about 'best practice' techniques for converting TIR images to formats that can be used with panorama software. Steps 1 and 2 should be revised to reflect these comments.

We revised step 1. Step 2 is identical to the original version, since we think that this step is valid for all panorama software, independently of how the images were converted.

Figure 4(e): There is something interesting going on in this image, where the surface water temperature appears to change on the left hand side (ie. much warmer than on the right, as evidences by the light/white colour). Do you know what might be causing this?

This change in temperature is very likely caused by the fact that the temperature of warm, exfiltrating groundwater (starting point of the stream on the right) is influenced and changed by the temperature of the surroundings. Since the water volume/depth is low, this can happen within a very short distance

Figure 5. It would be nice to see the visible images that accompany these TIR photos to aid interpretation

We added the visual images for both areas.

Technical note: Mapping surface saturation dynamics with thermal infrared imagery

Barbara Glaser[1,2], Marta Antonelli[1,3], Marco Chini[4], Laurent Pfister[1,5], 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
[4]Remote Sensing and Ecohydrological ~~modelling~~Modelling Research Group, Luxembourg Institute of Science and Technology, Esch/Alzette, 4362, Luxembourg
[5]Faculty of Science, Communication and Technology, University of Luxembourg, Maison du Savoir, 2 Avenue de l'Université, L-4365 Esch-sur-Alzette

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

Abstract~~.~~

~~In this study~~Surface saturation can have a critical impact on runoff generation and water quality. Saturation patterns are dynamic and thus their potential control on discharge and water quality is also variable in time. In this study, we assess the practicability of applying thermal infrared (TIR) imagery for mapping surface saturation dynamics. The ~~advantage~~advantages of TIR imagery compared to other surface saturation mapping methods ~~is~~are its large spatial and temporal flexibility ~~combined with a~~, its non-invasive character, and the fact that it allows a rapid and intuitive ~~character.~~visualization of surface-saturated areas. Based on an 18-month field campaign, we review and discuss the methodological principles, ~~under~~in which conditions the method works best, and ~~what~~the problems that may occur. These considerations enable potential users to plan efficient TIR imagery mapping campaigns and ~~to~~ benefit from the full potential offered by TIR imagery, which we demonstrate with several application examples. In addition, we elaborate on image post-processing and test different methods for the generation of binary saturation maps from the TIR images. ~~The method testing is performed~~We test the methods on various images with different image characteristics. Results show that the best method, in addition to a manual image classification, is a statistical ~~based~~ approach that combines ~~distribution~~the fitting of two pixel ~~classes~~class distributions, adaptive thresholding, and region growing.

1 Introduction

~~Patterns~~The patterns and dynamics of surface saturation areas have ~~remained~~been on ~~hydrologic~~hydrological research agendas ever since the formulation of the variable source area (VSA) concept by Hewlett and Hibbert (1967). Surface saturation is relevant for runoff generation and for water quality, due to variable active and contributing areas (Ambroise, 2004), as well as critical source areas (e.g. Doppler et al., 2014; Frey et al., 2009; Heathwaite et al., 2005). Likewise, surface saturation patterns

and their dynamics are closely linked to groundwater-surface water interactions (e.g. Frei et al., 2010; Latron and Gallart, 2007) and catchment storage characteristics and dynamics (e.g. Soulsby et al., 2016; Whiting and Godsey, 2016).

~~Albeit~~Despite the prominent role of saturated areas in hydrological processes research, ~~their~~ mapping them remains a challenging exercise. The most straightforward mapping method consists ~~in~~of locating saturated areas by walking through the catchment. However, this simple but labour~~-~~-intensive 'squishy-boot' method (e.g. Blazkova et al., 2002; Creed et al., 2003; Latron and Gallart, 2007; Rinderer et al., 2012) is suitable neither ~~applicable to~~ for large areas nor ~~is it suited~~ for fine-scale spatial resolutions. Dunne et al. (1975) introduced topography, soil morphology, hydrometric measurements (soil moisture, water table level, baseflow), and vegetation as useful indicators for delineating saturated areas. ~~Even if it~~It is still a ~~matter of~~valid research question today of how to best make use of these catchment characteristics to delineate saturated areas (e.g. Ali et al., 2014; Doppler et al., 2014; Grabs et al., 2009; Kulasova et al., 2014a, 2014b)~~, only hydrometric~~. Hydrometric measurements ~~have~~offer the potential ~~to monitor~~for monitoring the local temporal evolution (in increments ranging from minutes to months) of dynamic surface saturation. ~~Their major disadvantage is that area representative hydrometric measurements are difficult to carry out for larger~~The analysis of topography, soil morphology, or vegetation allows lasting saturation patterns to be identified for large contiguous areas.

Remote sensing ~~methods have~~has proven to be well~~-~~-suited for mapping temporal dynamic patterns of surface saturation over large areas. It is possible to extract flooded areas in the order of metres to kilometres ~~to metres~~ from data acquired with satellite and airborne platforms, such as synthetic aperture radar (SAR) images (e.g. Matgen et al., 2006; Verhoest et al., 1998), or the ~~normalised~~normalized difference water index (NDWI) and the ~~normalised~~normalized difference vegetation index (NDVI) (de Alwis et al., 2007; Mengistu and Spence, 2016). Observations at higher spatial resolution (order of centimetres) require unmanned aerial vehicles (UAVs) or ground-based instruments. Due to various technical constraints, ~~up~~ to ~~now~~date, SAR image acquisitions are ~~scarcely~~rarely used for UAV-based applications or for ground-based applications that are not restricted to a fixed location (e.g. Li and Ling, 2015; Luzi, 2010). ~~NDWI and NDVI are theoretically applicable at these scales, however, to the best of our knowledge the necessary simultaneous acquisition of short wave infrared and visible light (VIS) images has not yet been performed via UAVs or on the ground~~NDWI and NDVI are applicable at these scales (e.g. Orillo et al., 2017; Wahab et al., 2018), however, to the best of our knowledge, the necessary simultaneous acquisition of short-wave infrared and visible light (VIS) images has not yet been performed by UAVs or on the ground for mapping surface saturation.

Ishaq and Huff (1974) and Dunne et al. (1975) suggested the use of VIS or infrared photographs for mapping surface saturation. ~~However, even though VIS cameras have been deployed on the ground and mounted on UAVs, airborne or satellite platforms for a long time, this suggestions was rarely~~However, this suggestion has rarely been followed in the last 40 years ~~(exception e.g. Portmann, 1997).~~(with Portmann, 1997, being a notable exception) despite VIS cameras having been deployed on the ground and mounted on UAVs, airborne or satellite platforms for a long time. Recently, Chabot and Bird (2013) and Spence and Mengistu (2016) successfully used VIS cameras mounted on UAVs for mapping surface water (a wetland of 128 ha and an intermittent stream surveyed via three transects of 2 km each). Silasari et al. ~~(2017)~~(2017) mapped surface~~-~~-saturated areas

on an agricultural field (100 m x 15 m) ~~by~~ using a VIS camera mounted on a weather station for high-frequency image acquisition.

~~Today, thermal infrared (TIR) imagery features the same temporal and spatial flexibility as VIS imagery. Pfister et al.~~Since the advent of affordable, handheld thermal infrared (TIR) cameras, TIR imagery features the same temporal and spatial flexibility as VIS imagery. In the context of this technical advancement, TIR imagery started to be used for analysing hydrological processes such as groundwater – surface (water) interactions (e.g. Ala-aho et al., 2015; Briggs et al., 2016; Pfister et al., 2010; Schuetz and Weiler, 2011) or water flow paths, velocities, and mixing (e.g. Antonelli et al., 2017; Deitchman and Loheide, 2009; Schuetz et al., 2012). However, applications of TIR imagery for mapping surface saturation are rare. Two examples are from Pfister et al. (2010) and Glaser et al. (2016) ~~demonstrated the potential of TIR imagery for mapping surface saturation by carrying out repeated infrared image acquisitions at small spatial scales (centimetres to metres) with handheld cameras. So far, these two studies represent rare examples of using TIR imagery for mapping surface saturation, while the usage of TIR imagery for analysing groundwater – surface (water) interactions (e.g. Ala-aho et al., 2015; Briggs et al., 2016; Pfister et al., 2010; Schuetz and Weiler, 2011) or water flow paths, velocities, and mixings (e.g. Antonelli et al., 2017; Deitchman and Loheide, 2009; Schuetz et al., 2012) became rather common with the advent of affordable, handheld TIR~~, who demonstrated the potential for TIR imagery to map surface saturation by carrying out repeated TIR image acquisitions at small spatial scales (centimetres to metres) with handheld cameras.

~~A~~One reason ~~why~~ for the scarce number of studies ~~using~~that use TIR imagery for mapping surface saturation ~~are still scarce~~ is certainly that ~~existing~~few descriptions of the methodological advantages and challenges exist. However, there are ~~sparse. Several~~several general guidelines and methodological descriptions for TIR imagery applications ~~exist. For example, they~~. These studies focus on one specific aspect of TIR imagery, such as co-registration (Turner et al., 2014; Weber et al., 2015) or on how to acquire correct surface water temperatures, which is the ~~main~~most common application of TIR imagery in hydrology (e.g. Dugdale, 2016; Handcock et al., 2006, 2012; Torgersen et al., 2001). Many of these recommendations can be directly applied for mapping surface saturation via TIR imagery (e.g. choice of sensor type). However, some recommendations are redundant (e.g. temperature corrections) or different (e.g. optimal time scheduling) for the application of TIR imagery for surface saturation mapping.

Here, we go beyond the mere demonstration of the potential for TIR imagery to map saturated surface areas and address the related application-specific technical and methodological challenges. ~~We~~The novelty of this work is that we assimilate, within one study, fundamental principles, technical aspects, and methodological possibilities and challenges with an exclusive focus on the mapping of surface saturation. This includes all steps, from image acquisition to the generation of binary saturation maps. To do this, we (1) review relevant technical and methodological aspects from existing TIR imagery literature and (2) complement them with our expertise and results from an 18-~~months~~month field campaign. The field campaign ~~was~~ focused on the recurrent acquisition of panoramic images ~~of~~with a portable TIR camera in seven distinct riparian areas ~~with a portable TIR camera. Yet, the~~. The precautions and considerations that we describe in this technical note are also valid for surface

saturation mapping campaigns with permanently installed ground-based TIR cameras and TIR cameras mounted on UAVs, and airborne or satellite platforms.

~~**2 Acquisition of TIR images for mapping surface saturation patterns**~~

The manuscript is structured in two main parts. The first part (Section 2) focusses on the mapping approach itself and combines a literature review with examples of our own experience. The second part (Section 3) demonstrates the application of different pixel classification techniques for generating binary saturation maps from TIR images by applying and comparing them for different example images. A discussion and a conclusion section evaluate the key features of the manuscript and outline perspectives for future research and applications for TIR imagery in hydrological sciences.

2 Mapping surface saturation with TIR imagery: state of the art and examples

2.1 Fundamental principles

TIR cameras ~~allow to obtain an areal picture of~~ are used for measuring surface temperatures remotely (e.g. 100 µm penetration depth for water columns~~).~~) within an area of interest. The cameras sense the intensity of thermal infrared radiation emitted by the objects the camera is pointed at. The surface temperature T (K) of the objects is then calculated from the sensed ~~radiation~~radiant intensity~~,~~ W (Wm^{-2}), based on Stefan-Boltzmann's law ~~and considering some~~with the Stefan-Boltzmann constant $\sigma = 5.67*10^{-8}\ Wm^{-2}K^{-4}$

$$T = \sqrt[4]{(W/\sigma)} \quad (1)$$

Considering radiometric corrections~~, such as~~ for material~~-~~-specific emissivity~~, reflected~~ ε, for reflections of radiation from the surroundings, and for atmospheric induced ~~radiation.~~ and attenuated radiation, the radiant intensity W is split into the emission from the object (W_{obj}), from the ambient sources (W_{refl}), and the atmosphere (W_{atm})

$$W = \varepsilon\tau W_{obj} + (1-\varepsilon)\tau W_{refl} + (1-\tau)W_{atm} \quad (2)$$

with τ being the transmittance of the atmosphere, which depends on the distance between the object and the camera sensor, as well as on relative air humidity. Ultimately, values for the temperature of the ambient sources and the atmosphere, the targeted object's emissivity, the distance between object and camera, and the relative humidity are required for accurately estimating an object's surface temperature T.

Details on the principles of TIR imagery, TIR sensor types (i.e. wave length, sensitivity), and considerations for choosing the most appropriate camera and remote sensing platform for the desired acquisition (i.e. accuracy, resolution) are provided in the literature (cf. Dugdale, 2016; Handcock et al., 2012). For this study, we ~~used~~relied on two different handheld TIR camera models: a FLIR B425 with a resolution of 320 x 240 pixels and an angle of view of 25° and a FLIR T640 with a resolution of 640 x 480 pixels and an angle of view of 45° (FLIR Systems, Wilsonville. USA). The wider angle of view of the FLIR T640

clearly ~~eased~~facilitated the image acquisition in this study, while a pixel resolution lower than the resolutions of the two cameras would still have been sufficient for the identification of surface saturation patterns.

~~Determining~~We define surface saturation ~~with TIR imagery implies that surface saturation is defined~~ as water ponding or flowing on the ground surface (even if only present as a very thin layer). Mapping surface saturation with TIR imagery requires (1) a sufficient temperature contrast between surface water and the surrounding environment (e.g. dry soil, rock, vegetation) and (2) at least one pixel of the TIR image being known to correspond to surface water. When these two requirements are met, it is possible to visually identify the surface saturation patterns in a TIR image. This is exemplified with a TIR image of a riparian-stream zone (Fig. 1). The substantial temperature contrast (requirement 1) allows us to differentiate between two TIR pixel groups, i.e. surface water pixels and surrounding environment pixels. With ground truth data at hand (here: VIS image, alternatives ~~are e.g.~~ include stream water temperature or knowing the ~~stream course~~location of the creek) for point 1 of Fig. 1 (requirement 2), the group of pixels with higher temperatures can be identified as surface water~~, whereas the~~. The group of pixels with lower temperatures can be ~~assigned to the~~regarded as non-saturated surrounding environment (cf. Fig. 1, point 2). With this classification in mind, the TIR image significantly amplifies the appearance of surface ~~-~~saturated areas ~~as expressed from~~relative to a VIS image. Moreover, the TIR image reveals additional surface ~~-~~saturated areas that are not clearly identifiable (cf. points 3 in Fig.1) or not visible (cf. area above point 6, Fig.1) within a VIS image.

The example shows that the identification of surface saturation relies on temperature contrasts between surface water and the surrounding environment. Radiometric corrections of TIR images for obtaining correct temperature values are thus not necessary. However, interferences that affect temperature, such as shadow casts or reflections (cf. Dugdale, 2016; Handcock et al., 2012), cannot be disregarded as they can influence the temperature contrast (see ~~section~~Section 2.2). In cases where the water temperature is too similar to the surrounding materials, saturated areas might be falsely identified as dry, whereas surrounding materials might be falsely identified as wet. In cases where non-uniform water temperatures occur, different water sources may be distinguished (cf. Fig.1, where point 4 likely represents stream water, ~~point~~points 5 and 7 likely represent the exfiltration of warmer groundwater). However, a bimodal distribution of water temperatures (e.g. cold stream and warm exfiltrating groundwater or warm ponding water) can also lead to a misinterpretation of temperature contrasts to the surrounding environment (e.g. a surrounding material with a temperature that is in-between the water temperatures might be identified as water).

For the above-mentioned reasons, it is important to evaluate the applicability of the TIR images for identifying the surface-saturated areas with some ground truth / validation data. For the validation, we relied on immediate visual verification during image acquisition, as well as on VIS images. Another option is to install sensors that can verify the presence or absence of water on the ground surface locally, yet this is an experimental effort and only results in validation data for selective points. Validating the TIR images with other saturation mapping techniques is difficult, since most of these techniques implicitly include saturation in the upper soil layer, while the current use of TIR imagery excludes the soil. For example, saturated areas inferred via the squishy boot method account for areas where water is squeezed out of the soil when stepping on it, whereas such areas are not detected as saturated areas by the non-invasive TIR imagery.

2.2 Image acquisition interferences

Impact of weather conditions

Weather conditions can interfere with TIR image acquisition (e.g. Dugdale, 2016; Handcock et al., 2012). ~~We identified similarity between air and water temperatures as the main reason for insufficient temperature contrasts between~~The main problem stems from similar temperatures of water and the surrounding environment, compromising an identification of surface saturation with TIR images (Fig. 2a). Water has a higher thermal capacity than most environmental materials and therefore the water surface temperature generally aligns more slowly with the air temperature than the surface temperatures of surrounding materials ~~do~~. During our field campaign, it became clear that particularly during day-night-day or seasonal transitions, this difference in thermal capacities induced a convergence of the surrounding environment's temperatures (which ~~are aligning~~align to the air temperature) to the water temperature. Furthermore, direct exposure of the study site to sunlight, combined with shadow casts, commonly distorted the temperature contrasts. ~~Shadowed surrounding~~Surrounding materials in the shade with ~~different~~ temperatures ~~than~~different to the same~~, sunlit~~ surrounding materials in sunlight led to reduced temperature contrasts between these materials and the surface water (Fig. 2b). Once the direct sun exposure ceased, the different thermal capacities of different materials heated by the sun could still cause ~~sun memory effects with~~ patches of warmer and colder temperatures. Rain and fog may also influence image quality due to water droplets falling between the TIR sensor and the ground, eventually blurring the images and causing uniform temperature signatures (Fig. 2c).

To avoid the acquisition of non-useable TIR images, we advise ~~to plan~~planning field campaigns adapted to the weather forecasts. The ideal situation is to work during dry ~~and cloudy~~ weather with warm or cold air temperatures in order to ensure a clear difference between the temperature of the surrounding materials and the more temperate water surface temperatures. Dugdale (2016) reported the time period from mid-afternoon to night ~~–~~time as an optimal TIR image acquisition ~~time~~period for monitoring water surface temperatures. Based on our 18-~~months~~month field campaign, we suggest that the optimal TIR image acquisition time for identifying surface saturation patterns is early ~~daytime~~morning. At this time, there are no undesirable effects due to sunlight (shadows, ~~memory effect) are non existent and~~warming-up) and there are generally high temperature contrasts between water surfaces and the surrounding environment ~~are commonly large.~~. Cloudy conditions can also help to avoid the effect of direct sunlight. A site ~~-~~specific analysis of the sun exposure ~~over~~throughout the day can help ~~to understand~~pinpoint at what other times images can be taken ~~under~~in favourable conditions for a specific study site.

~~View obstructions~~

Camera position

Obstructions in the TIR camera's field of view are obviously problematic. Yet, permanent view obstructions on the ground (e.g. tree trunks, Fig. 2d, point 6) ~~showed~~proved to be useful ground reference points ~~in~~during our field campaign. Temporary view obstructions, such as growing vegetation (Fig. 2d), recent litter, and snow cover ~~proved to be problematic~~are a problem for repeated imaging campaigns. Cutting the vegetation during the growing season ~~can be~~is an option for small study sites. Our experience is that the coverage of grasses and herbaceous plants with small leaves ~~still~~ is normally low enough to permit

the recording of the ground surface temperature, while the coverage of ferns or tree leaves ~~are~~is normally completely opaque. Snow cover usually hides surface saturation. Yet, periods where the amount of snow is low are commonly not problematic, since the saturated areas mainly stay uncovered due to a warmer water temperature and thus fast melting of the snow.

Ideally, images are taken from above and at nadir to the study site. Oblique angles of view (>30° of nadir) reduce the object's emissivity and thus distort the detected temperatures in the TIR images (Dugdale, 2016). The incorrect temperature values ~~as such~~ are not critical as such for mapping surface saturation patterns, but we observed that wide ranges of angles can result in distinct temperature distortions and thus reduced temperature contrasts within the images. In a similar way, varying distances between camera and ground surface for different positions within one image (e.g. top / bottom, left / right) do not only provoke pixels with varying area equivalents, but can also distort the temperature detection and thus temperature contrast. Therefore, ground-based cameras should be positioned at locations that minimize the range of angles of view and the distances between camera and ground surface. In ~~case~~the event of repeated image acquisitions of a given area of interest, we took the pictures ~~each time~~ from the same position each time in order to facilitate subsequent image comparisons. For repeated image campaigns, it could be useful to install a structure that allows ~~to acquire~~ several images to be acquired by moving the camera to specific positions with fixed heights above the ground and fixed angles of view. This could simplify the post-processing and assemblage of the images into panoramic images (cf. ~~section~~Section 2.3).

Measurement artefacts during image acquisition

For determining surface saturation, the TIR images should cover ~~a part of the stream or another~~an area known to be surface~~-~~saturated (e.g. stream, visually obvious wet spots) in order to have a reference for water temperature (cf. ~~section~~Section 2.1). In addition, a VIS image should be acquired simultaneously ~~with~~to the TIR image for comparison. The TIR imagery parameters ~~that are~~ necessary for correcting and converting the radiation signal to temperature values (e.g. air temperature, humidity) do not need to correspond to the actual conditions, since only the temperature contrast and not the correct temperature value is required for defining saturated areas. ~~However, setting realistic values for the distance between camera and ground surface helped the auto-focus process of the camera. Working with proper values for the emissivity (0.96–0.98 for water) and for the actual reflection temperature helped to prevent the observation of unrealistic surface temperatures. Nonetheless, in case of clear sky or cold winter days~~Certainly, 'wrong' temperatures influence the temperature contrast between the surroundings and the water, but this effect on the contrast can be negative or positive. If correct temperatures are targeted, radiometric corrections need to be applied during the image post-processing procedure. This allows for example to consider different emissivities for different surface materials by using appropriate values for each individual image pixel (Aubry-Wake et al., 2015). However, in our experience, setting realistic parameter values during the image acquisition helped the auto-focus process of the camera and prevented the observation of unrealistic surface temperatures. Nonetheless, in the event of clear skies or on cold winter days, we occasionally observed negative temperatures for flowing water. The explanations for these observations ~~stay purely speculative (Antonelli et al., 2017). However, for the identification of surface saturation patterns such unrealistic~~remain speculative. Potentially, a particularly strong reflection of the radiation from the surroundings and the sky in the water

influenced the temperature detection. However, for the identification of surface saturation patterns, such unrealistic negative temperatures do not pose a problem since the temperature range stays correct (Antonelli et al., 2017).

Reflections of surrounding objects on the water surface (Fig. ~~2e) and image vignetting (cf. aura effect in Antonelli et al., 2017) can occur during the image acquisition and can compromise a further use of the TIR images. In this study, the image vignetting was unproblematic, especially in case a panorama was built from several images (cf. section 2.3). This is due to the fact that image vignetting only occurs at the edges of the pictures and that it is of minor relevance in images with high temperature contrasts. Reflections of surrounding objects on the water surface limit the value of the images for saturation identifications in a similar way as shades (cf. Fig. 2d and 2e). The difference to shades~~2e) and image vignetting can occur during image acquisition and can compromise a further use of the TIR images. Vignetting is the falloff of radiation intensity towards the edges of the image, which is mainly generated by the geometry of the sensor optics (esp. wide angle lenses) (cf. Kelcey and Lucieer, 2012). As a consequence, the monitored temperature can change towards the edges of the picture (cf. aura effect in Antonelli et al., 2017). In this study, the image vignetting was unproblematic, especially where a panorama was built from several images (cf. Section 2.3). This is due to the fact that the effect of image vignetting only occurs at the edges of the pictures and it is of minor relevance in images with high temperature contrasts. Reflections of surrounding objects on the water surface limit the value of the images for saturation identifications in a similar way to shadows (cf. Fig. 2d and 2e). The difference with shadows is that reflections also occur with diffuse light, which makes it difficult to predict their occurrence and thus to avoid them.

2.3 Generation of TIR panorama images

We acquired the images ~~that were needed~~used for the assemblage of a panoramic view in two different ways: (1) by taking single, overlapping images, ~~or~~and by (2) taking a video of the area of interest. While both approaches deliver similar final results, videos are ~~taken~~recorded faster than sequences of individual images. Independently from the chosen data format, we ~~kept the camera's parameter settings (cf. section 2.2) constant during image / video acquisition for the area of interest and~~ ensured that the saving format retained the temperature information as radiometric data~~. These two precautions were necessary~~ for further image processing (see below and Fig. 3). Sun (dis)appearance and automatic noise corrections by the camera (non-uniformity corrections, cf. Dugdale, 2016) can lead to ~~clear~~considerable shifts in recorded temperatures from one image / video frame to another. ~~In order to control if such a temperature shift occurred, we fixed the temperature – colour scale while taking the video/single images of an observation area. In case the colour (and thus temperature) of overlapping image parts shifted, we restarted the image acquisition, since a correction of~~Since correcting such temperature shifts is difficult (cf. Dugdale, 2016)~~.~~, we opted to control them by fixing the temperature – colour scale and restarting image acquisition if the colour (and thus temperature) of overlapping image parts changed.

We acquired the images / video frames in such a way that the area of interest ~~constituted~~formed the central part of a panorama. This allowed us to avoid image gaps and distortion effects at the borders of the area of interest. When possible, we ensured that the single pictures / video frames included overlapping parts with identifiable structures such as the stream bank, tree

stems, or stones as natural reference points. For videos, it was essential to move the camera slowly enough to obtain sharp images and to use a low frame rate (e.g. 2 Hz) to keep the number of video frames reasonable (enough frames for obtaining area overlaps, but not too many frames showing the same area).

The generation of a panorama from overlapping TIR images / video frames acquired with a ground-based camera involves some challenges that specifically relate to TIR and / or ground-based images. This needs to be addressed in TIR-specific panorama generation and image processing steps, as presented ~~in a nutshell~~briefly by Cardenas et al. (2014). ~~Before blending the images / video~~Our approach consisted of transforming the acquired images / video frames containing the radiometric information (see above) into grey-scaled, standard format images / videos (Fig.3, step 1) in order to ~~a~~allow the use of ordinary panorama ~~image, one needs to ensure that all single images / video frames have the identical temperature~~ assemblage software. We relied on grey colour scale~~.~~

~~Ideally,~~ images, linearly splitting the ~~temperature~~ colour ~~scale is a linear greyscale and ranges from~~shades over the global ~~minimum to the global maximum~~ temperature ~~value~~range of the acquired images / video frames~~. This~~, since this prevents artefactual colour mixing effects and ~~it allows~~allowed us to embed the temperature information in the generated panoramas. ~~The camera types that we used for this study did not allow to fix such a linear, grey temperature colour scale prior to image acquisition. Therefore, we transformed the acquired images / video in a colour conversion step to grey scaled images (Fig.3, step 1) based on the temperature information retained in the acquired images / video (see above). In cases where~~When the extreme temperature values of an image were not relevant for the identification of saturated areas, we truncated the global temperature range ~~retained for the colour conversion~~ in favour of a better colour contrast and a finer temperature class width retained in the grey values (e.g. the retained temperature class width is 0.1 °C in case of a temperature range of 25.5 °C and an image with 255 grey values~~, the retained temperature class width is 0.1 °C~~).

We employed Microsoft's Image Composite Editor (ICE) and the PTGui panorama software (New House Internet Services) ~~for creating~~to create panorama images (Fig.3, step 2). ICE and PTGui allow ~~to create~~ the creation of panoramas from single images (and from video frames for ICE) with an automatic mosaicking function (i.e. a function that geometrically transforms, aligns, and overlaps the single images). ~~Unlike PTGUI, the ICE software additionally allows to automatically mosaic video frames. However, in cases where the automatic mosaicking fails, only PTGUI offers adequate possibilities to manually interact with the image alignment (i.e. defining control points for matching distinct points in overlapping images).~~ TIR images generally show less identifiable features and lower contrasts than VIS images (cf. Weber et al., 2015). Therefore, a (partial) failure of automatic mosaicking is not uncommon and manual interactions ~~became frequently necessary for the images of our 18-months field campaign. Nevertheless, we always first attempted to create a panorama from a video with ICE, since the video acquisition showed to be more efficient for the image acquisition and for the grey-scale conversion. In cases where the image mosaicking of a video (partly) failed, we extracted single images from the video and processed them as individual images, eventually resorting to PTGui.~~with image alignment (i.e. defining control points for matching distinct points in overlapping images in PTGui) were frequently necessary for the TIR images taken during our 18-month field campaign.

In order to compare several panorama images of the same area, one needs to co-register the panoramas ~~need to be co-registered~~ (Fig. 3, step 3). In principle, it is possible to ~~georectify~~geo-rectify the TIR images by allocating ~~in the images~~ geographical coordinates ~~that~~to the images, which are derived from ground control points ~~(Keys et al., 2016; Silasari et al., 2017)~~(cf. Keys et al., 2016; Silasari et al., 2017) or from a virtually projected elevation model ~~(cf. DSM-derived virtual perspectives in Cardenas et al., 2014). However, this can result in strong interpolations or distortions, due to view obstruction in the picture. Instead~~(cf. Cardenas et al., 2014; Corripio, 2004; Härer et al., 2013). However, this can result in large gaps or strong interpolations and distortions in the images, due to view obstructions in the picture. Instead of this, therefore, we co-registered TIR panoramas of the same area against each other (cf. Cardenas et al., 2014; Glaser et al., 2016). More specifically, we registered and ~~cut~~cropped them to the dimensions of a reference TIR panorama of the area of interest (Fig. 3, step 3).

~~3~~2.4 Application examples

~~In this section we present three examples from our 18-months field campaign that demonstrate the capability of TIR imagery for analysing surface saturation patterns and their dynamics. All images were taken in the Weierbach catchment – a forested, 42 ha headwater research catchment in western Luxembourg (Glaser et al., 2016; Klaus et al., 2015; Martínez-Carreras et al., 2016; Schwab et al., 2018). We avoided unfavourable environmental conditions for the image acquisitions (cf. 2.2, Fig. 2) by allowing a few days tolerance to the targeted (bi-)weekly recurrence frequency. Additionally, we cut ferns that obstructed the camera view during the summer months. The 364 acquired panorama were divided into three groups classified as usable without restrictions (32.4 %), usable with some restrictions (small negative effects of low temperature contrasts or covering vegetation visible, 31.1 %) and unusable (36.5 %).~~

In this section, we present three examples from our 18-month field campaign that demonstrate the potential for TIR imagery to analyse surface saturation patterns and their dynamics. All images were taken in the Weierbach catchment - a forested, 42 ha headwater research catchment in western Luxembourg (Glaser et al., 2016; Klaus et al., 2015; Martínez-Carreras et al., 2016; Schwab et al., 2018). We avoided unfavourable environmental conditions for the image acquisitions (cf. 2.2, Fig. 2) by allowing a few days tolerance around the targeted (bi-)weekly recurrence frequency. Additionally, we cut ferns that obstructed the camera view during the summer months. The 364 acquired panorama images were divided into three groups classified as usable without restrictions (32.4 %), usable with some restrictions (small negative effects of low temperature contrasts or covering vegetation visible, 31.1 %), and unusable (36.5 %).

The usable panoramas captured the temporal evolution of surface saturation over the 18-~~months~~month field campaign. This demonstrates the robustness of TIR imagery through the complete range of seasonal conditions (Fig. 4), including snow and growing vegetation, as well as warm and cold water. The full extent of added value provided by TIR imagery compared to VIS imagery was documented for cases with different seasonal conditions (Fig 4 a/e vs. b vs. c/d), particularly for situations with less pronounced differences in discharge levels (Fig. 4 a vs. b vs. c, d vs. e). For example, the comparison of the VIS images of December 2015 and June 2016 (Fig. 4 a vs. c) suggests wetter conditions for December 2015, while the two TIR images show similar saturation patterns for the two dates.

In addition to surface saturation dynamics, the TIR images can also reveal distinct types of saturation patterns. For example, the orientation of saturated areas may ~~be oriented within~~change over a few metres from perpendicular (Fig. 5 top) to parallel (Fig. 5 bottom) to the adjacent stream. ~~The parallel direction on~~The extension of saturated areas along the left bank (Fig. 5 bottom) appears to be created by a parallel ~~flow~~extension of the stream in a flat riparian zone that becomes an extended stream bed~~, while the perpendicular direction~~. The surface saturation oriented perpendicularly to the stream at the right bank (Fig. 5) appears to be generated from exfiltrating groundwater that flows downhill to the stream at the soil surface. Thus, the different directional extents of the saturated areas can indicate different processes underlying the surface saturation formation.

Finally, the images allow us to identify the spatial heterogeneity of temporal saturation dynamics across different study sites. Figure 6 shows TIR images of the riparian zone of two different source areas with different degrees and dynamics of surface saturation. In area 1 (Fig. 6, left panels~~)~~), the pattern of saturation areas ~~had~~ barely changed from February to April, while in area 2 (Fig. 6, right panels) some locations had dried out (red circles). In December 2016, the riparian zones of both source areas were completely dry and the stream started further downstream in comparison to the other observation dates (red arrows). This suggests that both source areas evolve from very wet to very dry conditions (during which surface saturation is mainly represented by spots with stable groundwater exfiltration) with distinctly different transition dynamics.

~~4 Building~~3 Quantification of saturation through pixel classification

3.1 Methods for generating binary saturation maps

~~The application examples described in section 3 demonstrate the potential of TIR images for a rapid and intuitive visualisation of surface saturated areas. The 'raw data' images can be used without any additional processing to study surface saturated areas, their evolution over time, and how and where they occur – ultimately contributing to a better mechanistic understanding of the hydrological processes prevailing in the studied area. This is even more valid, since the images do not only provide surface saturation extensions, but potentially also information on the different mechanisms that generate saturated patches (cf. Fig.1 and Fig.5, groundwater inflows vs.~~ The application examples described in Section 2.4 demonstrate the potential for TIR images to rapidly and intuitively visualize surface-saturated areas. However, the 'raw data'~~stream water). The pure visual information provided *per se* by the images is also usable as soft data, e.g. for model validation (e.g. different types of extent compared to stream, Fig.5, more and less stable saturation patterns, Fig.6). However, the~~ images need to be transformed into binary saturation maps for further analyses based on quantitative values (as e.g. saturation percentages~~) or for applying map comparison methods (e.g. confusion matrices, kappa coefficients).~~

~~One possibility to transform a TIR image into a binary saturation map is to take the temperature range of pixels that are known to be saturated (i.e. stream pixels) and to define all pixels of the~~). A common approach to binarizing an image is histogram thresholding (e.g. Rosin, 2002). This allows a TIR image to be transformed into a binary saturation map by taking the temperature range of pixels that are known to be saturated (i.e. stream pixels) and defining all pixels in that image that fall into that temperature range as saturated (cf. Glaser et al., 2016; Pfister et al., 2010). ~~This approach requires a sufficient temperature~~

~~contrast. Furthermore, artefacts (such as pixels corresponding to vegetation covering the stream) may induce some uncertainty in pixel classification, eventually leading to discrepancies to visual saturation pattern identifications. The selection of the temperature range for surface saturation can be done manually by adapting the chosen range until the resulting saturation map matches best the visual assessment of the original TIR and – if possible - VIS image. Up to a certain level it is straightforward to visually reject a temperature range by qualifying it as being too wide or too small (cf. Fig.~~Several thresholding algorithms can be found in the literature, each of which has its characteristic assumptions with respect to image content (Patra et al., 2011). Unsupervised approaches other than thresholding are also used for image binarization, e.g. clustering (Li et al., 2015). Yet, thresholding is the most rapid technique for achieving a binary classification of an image, even though the selection of an adequate threshold value represents a critical step and its choice strongly influences the classification outcome.~~7). However, finding an unequivocal temperature range is not feasible and the selection of the most plausible temperature range (Fig. 7, dark green asterisk) remains somewhat subjective. Consequently, a pixel classification based on this procedure remains tarnished by uncertainties and the definition of an uncertainty range within which the temperature range is considered plausible (Fig. 7, dark green, dashed lines) is a subjective exercise as well. In our experience from the 18 months field campaign, the uncertainty range was generally small for images with low saturation and gradually increased with higher saturation (compare Fig. 7 d-b). Accordingly, images with a large difference in percentages of saturated pixels (e.g. Fig. 7b vs 7d) did not encounter an overlap of the uncertainty ranges. For some images the uncertainty range was very high (Fig. 7a) and a comparison with other images with percentages of saturated pixels in the same range can thus be problematic. In such cases it is good if only one person defines the optimal temperature ranges and thus saturation patterns for all images that are intended to be compared in order to ensure consistency in the image interpretation (e.g. we realised that some persons consistently favoured higher and others consistently favoured lower saturations within the uncertainty range of a set of images).~~

One possibility to select a threshold value for classifying surface saturation is to manually adapt the temperature range until the resulting saturation map matches best the visual assessment of the original TIR and – if possible – VIS image. A more objective and, for time-lapsed images, faster ~~option for constraining the temperature range of saturated pixels~~method consists ~~in using~~of relying on the temperature of preselected pixels or a predefined mask for saturated and unsaturated parts in all images. Such pixels / masks can be selected based on a visual interpretation of the images or on information obtained from reference sensors in the field, indicating whether a location was wet or dry at the surface at the time of ~~the~~ image acquisition. ~~As an example we chose a mask of 2000 pixels falling into an area that always stayed dry and 2000 pixels falling into an area where the stream was flowing all year (red rectangles, Fig.7). Based on this mask we defined a minimum and maximum temperature range for each image in such a manner that 90% of the pixels falling below the mask are defined as saturated and dry, respectively. The resulting uncertainty ranges of saturation percentages (Fig. 7, blue points) are higher than for the manual temperature range selection (Fig. 7, dark green, dashed lines). The saturation identifications based on the dry part of the mask were clearly not constrained enough. The saturation identifications based on the wet part of the mask sometimes approached the manual saturation identifications (Fig. 7 a, c) but in other cases they even exceeded it (Fig. 7d). The mask defined uncertainty range of saturation could be narrowed by selecting a value higher than 90% of the pixels for the temperature range~~

~~definition. However, this increased the risk to obtain a clearly wrong value (cf. Fig. 7d), since the wet / dry mask can cover pixels of the 'wrong' category (due to artefacts like vegetation covering the stream or due to distorted co-registered images, resulting in a shifted mask). A reduced mask size prevents such 'wrong' pixels, but reduces the captured variability in temperature (in an extreme case down to one temperature value), which increases the risk to miss the warmest or coldest temperature of the water or dry areas.~~

~~Another approach is to derive a threshold value for the temperature range from image statistics as e.g. the probability density function (Pfister et al., 2010). However, this is only straightforward in cases where the temperature distribution is clearly bimodal between water and the surrounding environment.~~ Silasari et al. ~~(2017)~~(2017) applied an automatic image classification for unimodal distributions based on a threshold parameter that needs to be calibrated to specific image conditions (in ~~their case brightness of VIS images). In our case, the image conditions vary to such an extent (e.g. very wet/dry conditions, water warmest/coldest material, cf. location of dark green asterisk on cumulative saturation curves, Fig. 7) that each image would require its own calibration to manual selected temperature ranges / saturation patterns.~~this case, the brightness of VIS images).This is only straightforward in cases where the temperature distribution between water and the surrounding environment is clearly bimodal. Chini et al. (2017) presented a parametric adaptive thresholding algorithm especially suited for images that do not show a clear bimodal ~~pixel values histogram distinguishing two different classes (e.g. water and surrounding environment).~~distribution. The algorithm makes use of an automatic selection of image subsections with clear bimodal distributions, a hierarchical split-based approach (HSBA), and a subsequent ~~parameterisation~~parameterization of the distributions of the two pixel classes. Since the two decomposed distributions might still overlap to a certain extent, Chini et al. (2017) advise ~~to complement~~complementing the decomposed distribution information with contextual information of the image for the final generation of a binary image instead of selecting a single threshold value between the two decomposed distributions. ~~For this, they apply a region growing algorithm where the seeds and the stopping criteria are constrained to the identified distribution of the class of interest (i.e.~~Several approaches are available in the literature for including contextual information in the classification of a single spectral image, such as mathematical morphology (Chini et al., 2009) or second order textural parameters (Pacifici et al., 2009). Chini et al. (2017) suggested a region growing algorithm where the seeds and the stopping criteria are constrained by the identified distribution of the class of interest (i.e. here saturation).

3.2 Comparison of methods for generating binary saturation maps for TIR images

We applied three of the approaches described above to generate binary saturation maps on our TIR image data set. Here, we present the results for four example images with differing conditions during image acquisition (e.g. very wet/dry conditions, water being the warmest / coldest material, Fig. 7). We evaluated the results of the three different approaches based on our observations from the field and the corresponding VIS image as ground truth.

First, we manually chose a temperature range of saturation for each image. By its nature, this pixel classification approach creates results that are very close to ground truth. However, finding an unequivocal temperature range was not feasible and the selection of the most plausible temperature range (Fig. 7, dark-green asterisk) remained somewhat subjective. Furthermore,

artefacts (such as pixels corresponding to vegetation covering the stream) induced some uncertainty in the pixel classification, eventually leading to discrepancies compared to visually identified saturation patterns. Consequently, a pixel classification based on this manual procedure remained tarnished by some uncertainties. The definition of an uncertainty range within which the temperature range can be considered plausible (Fig. 7, dark-green, dashed lines) was also subjective. Generally, the uncertainty range was small for images with low saturation and gradually increased with higher saturation (compare Fig. 7 d-b). Accordingly, images with a large difference in percentages of saturated pixels (e.g. Fig. 7b vs. 7d) did not encounter an overlap of the uncertainty ranges. For some images, the uncertainty range was rather high (Fig. 7a) and a comparison with other images with percentages of saturated pixels in the same range was thus problematic. In such cases, it is preferable that only one person defines the optimal temperature ranges and thus saturation patterns for all images that are intended to be compared in order to ensure consistency in the image interpretation.

Secondly, we performed an objective selection of the temperature range of saturation based on masks with known pixel classes. For this, we used two masks, one with 2000 pixels falling into an area that always stayed dry and one with 2000 pixels falling into an area where the stream was flowing all year (red rectangles, Fig.7). Based on the mask, we selected the threshold for the temperature range as the 90th percentile and 10th percentile of the temperature of the stream mask pixels and dry mask pixels, respectively (i.e. 90% of the pixels falling below the mask were defined as saturated and dry, respectively). By using the two different masks, we obtained two temperature ranges, resulting in two different saturation percentages for each image (Fig. 7, blue points). The identification of saturated areas based on the dry mask was clearly not constrained enough. The identification of saturated areas based on the stream mask sometimes approached the manual identification of saturation (Fig. 7 a, c) but in other cases even exceeded it (Fig. 7d). The uncertainty range of saturation obtained with the two masks could be reduced by selecting a more extreme percentile for the temperature threshold definition. However, this increased the risk of obtaining a clearly incorrect value (cf. Fig. ~~We tested the usability of this approach for our TIR images by constraining the~~7d), since the stream / dry mask can cover pixels of the 'wrong' category (due to artefacts like vegetation covering the stream or due to distorted co-registered images, resulting in a shifted mask). A reduced mask size prevents such 'wrong' pixels, but also reduces the captured variability in temperature (in an extreme case down to one temperature value), which in turn increases the risk of missing the warmest or coldest temperature of the water or dry areas.

Finally, we tested the usability of the approach proposed by Chini et al. (2017), constraining a region growing algorithm to a) a bimodal distribution derived from the HSBA applied to the entire image, b) a bimodal distribution derived from the HSBA where the selection of bimodal image subsections was constrained to image-specific manual predefinitions of temperature ranges of saturation, and c) a bimodal distribution derived from pre-selected parts of the image that include clearly wet and dry areas. While in some cases the fully automatic image classification (a) worked very well in comparison ~~with~~to the manual selection of a temperature range (cf. Fig. 8 04/12/15, 30/08/16), for the other cases ~~the~~, saturation was mostly underestimated (cf. Fig. 8 25/02/16, 03/06/16). The additional constraint with image-~~-~~specific temperature ranges (b) overall improved the matches ~~to~~with the manually defined saturation patterns, but the result was strongly influenced by the match of the given constraint range to the range that was defined as the optimum for the image. A constraint with a rough estimated temperature

for saturation worked ~~poorer~~less well than a constraint with the temperature range as selected in the detailed manual assessment described earlier in the ~~section~~Section (cf. Fig. 7 green asterisks and lines). The classification based on pre-selected parts of the image (c) tended to result in higher saturation amounts. This improved the match for the cases that were underestimated with the fully automatic classification (a) (cf. Fig. 8 25/02/15, 03/06/16), but ~~it~~ overestimated saturation for the cases where the fully automatic classification (a) showed good results (cf. Fig. 8 04/12/15, 30/08/16).

~~5~~4 Discussion

4.1 Mapping surface saturation with TIR imagery

The main advantages of TIR imagery in comparison to other surface saturation mapping methods are its non-invasive character and its large temporal and spatial flexibility (centimetres to kilometres, minutes to months). Another advantage is that TIR images allow a rapid and intuitive identification and analysis of the dynamics of surface saturation patterns. The 'raw data' images can be used without any additional processing to study surface-saturated areas, their evolution over time, and how and where they occur – ultimately contributing to a better mechanistic understanding of the hydrological processes prevailing in the studied area. The pure visual information provided *per se* by the images is also usable as soft data, e.g. for model validation (e.g. different types of extent compared to stream, Fig.5, more and less stable saturation patterns, Fig.6). VIS imagery offers similar advantages ~~(Silasari et al., 2017)~~(Silasari et al., 2017), but commonly the saturated areas are not as clearly visible as with TIR imagery (cf. Fig. 1, Fig. 4). Moreover, VIS imagery is not usable during the night ~~time~~ and ~~it~~ cannot provide additional information about water sources and processes underlying the surface saturation formation (cf. Fig. 1, Fig. 5, groundwater inflow vs. stream water).~~).~~ Nevertheless, VIS imagery provides good ~~complimentary~~complementary information to the TIR imagery and should always be considered as a ground truth information source.

In our study, unfavourable image acquisition conditions (cf. ~~section~~Section 2.2) caused 36.5 % of the acquired images to be ~~not usable~~unusable for further processing. High amounts of unusable images are a common problem in environmental imagery (cf. e.g. cloud cover for satellite images, night-time for VIS images ~~(DeAlwis et al., 2007; Silasari et al., 2017)~~(DeAlwis et al., 2007; Silasari et al., 2017)). Flexibility in the scheduling of a field campaign is thus necessary to reduce the number of acquisitions during unfavourable conditions. A concern for the use of TIR imagery for mapping saturation patterns is that some saturated areas (e.g. warmed~~-~~-up ponding water) might not be identified as saturated due to a temperature that is very different from stream temperature. This relates back to the fact that temperature is only used as an indicator for saturation. Compared to other saturation indicators, such as vegetation mapping or hydrometric measurements (cf. Dunne et al., 1975), we ~~deem the~~consider TIR imagery with the above-mentioned advantages as the better indirect mapping method. However, the only way to directly map surface saturation consists ~~in~~of walking through the area of interest (e.g. squishy boot method), which remains restricted to small areas and / or low mapping frequencies.

The amount of ~~field work~~fieldwork for imagery mapping is generally reduced compared to other methods for mapping surface saturation (e.g. vegetation/soil mapping), allowing more frequent campaigns with higher spatial precision. Yet, ~~consistent~~in

consistency with other imagery mapping studies (e.g. Spence and Mengistu, 2016), the image post-processing in this study was time-consuming. Mosaicking and the co-registering of images is often considered particularly difficult for TIR images, since ground control points with a thermal signature are needed (Dugdale, 2016; Weber et al., 2015). Our experience showed that the images normally offered enough natural thermal ground control points (e.g. the stream bank) in cases where the temperature contrast between water and ambient materials was good enough for image usability. In combination with the presented post-processing workflow presented, the post-processing effort was reasonable. More automatized workflows like the one proposed by Turner et al. (2014) for mosaicking UAV acquired TIR images could also be adapted and applied.

The image acquisition considerations, post-processing steps, and application examples described focused on (bi-) weekly panoramic images of small areas, acquired with a portable TIR camera. A transfer of the TIR imagery technique to different temporal or spatial scales does not change the principles and possibilities of the technique, but will require some additional scale- and platform-dependent considerations. For example, using permanently installed ground-based cameras for image acquisitions with high temporal frequencies might challenge technical aspects such as protection of the camera against environmental influences, an automatic triggering of image acquisition, and power supply. These aspects might also be relevant for TIR imagery acquisition at larger spatial scales, especially when using UAVs. Besides, image acquisitions based on UAV or aeroplane overflights might for example require considerations of overflight regulations, saturation patterns within a forest could only – if at all – be mapped during the dormant season, and ground control points and ground truth data might be more difficult to obtain. Partly, such challenges are addressed in existing literature (e.g. Vivoni et al., 2014; Weber et al., 2015), others will need to be figured out by applying the TIR technique at such different scales.

4.2 Pixel classification methods

More challenging than TIR image mosaicking and co-registering was the generation of saturation maps from the TIR images. The different tested processing pixel classification methods tested all yieldyielded somewhat different results compared to pixel classification based on manual, visual assessment. Nevertheless, realisingrealizing an objective, automatic classification of saturated areas is not more challenging than for other surface saturation mapping methods. Saturation maps created frombased on the squishy boot method or vegetation/soil mapping are subjective due to decisions takenmade during field workthe fieldwork. The (un)supervised classification methods that are commonly used for creating saturation maps from remote sensing data (e.g. VIS images / NDVI/NDWI) also contain some uncertainty (Chabot and Bird, 2013; DeAlwis et al., 2007; Mengistu and Spence, 2016; Spence and Mengistu, 2016).

Moreover, the main problem for all of the tested saturation map generation methods (cf. sectionSection 4) is that they are not applicable without adapting them to individual image conditions (very wet, very dry, water being the warmest / coldest material, slightly different field of views). Other image processing methods for deriving saturation maps also do not fulfil this requirement and it is necessary to adapt the parameters (e.g. Silasari et al., 2017) or to(e.g. Silasari et al., 2017) or do a new supervision (with new classification pixels/masks) for the classification of images with different conditions (e.g. Chabot and Bird, 2013; Keys et al., 2016). By nowAt this stage, we consider a manual choice of the temperature range for saturated pixels

~~as~~ the best approach for time-lapsed images with very variable conditions and slight perspective shifts, even though it is labour ~~-~~intensive and somewhat subjective. For time-lapsed images with a fixed vantage point and for time spans with similar conditions (e.g. storm events), the ~~presented~~ automatable methods presented represent valuable options. ~~Especially~~In particular, the combination of an automatic decomposition of two pixel class distributions with a region growing algorithm yielded objective saturation maps close to the manual saturation classification and visual assessment of the TIR images (Fig. 8). Small adaptations of the constraint for the decomposition of two pixel class distributions were sufficient to obtain good results for the different image conditions (cf. Fig. 8 a – c) and further developments of the method might even allow ~~to perform such adaptations in a (semi)automatic way.~~such adaptations to be performed in a (semi)automatic way. More work on pixel classification might also include the application of machine learning techniques or, especially for time-lapsed images, the analysis of the temperature signals of individual pixels over time. Another interesting option may consist of combining the TIR images with additional data (e.g. VIS images or NIR images), which will allow multi-spectral classification methods to be applied (Chini et al., 2008) and at the same time, contextual information to be integrated (Chini et al., 2014).

~~6~~5 Summary and conclusions

This technical note presents recent work carried out in the Weierbach catchment, where we tested the ~~capabilities of~~ potential for TIR imagery ~~for mapping~~to map surface saturation dynamics. ~~We reviewed~~For the best of our knowledge, this is the first comprehensive review and ~~summarized the~~summary of the TIR imagery-related methodological principles and the required precautions and considerations for a successful application of TIR imagery~~.~~ for mapping surface saturation. We give advice for all steps, from image acquisition to processed saturation maps. The main requirement is a clear temperature contrast between water and the surrounding environments. Image acquisition during an 18-~~months~~month campaign showed that the method works best during dry ~~night time~~nights or dry early ~~daytime~~mornings and that images should be taken from well-chosen positions without ~~(non-)permanent view~~ obstructions ~~to~~ in view towards the ground. The workflow presented ~~workflow~~ for acquiring panoramic images is particularly suitable for small areas of interest (centimetres to metres) that are monitored with ~~intended~~ intermediate to low mapping frequencies (days to months). Moreover, the information contained in this technical note is also beneficial for applications at different temporal and spatial scales (fixed cameras for high frequency images, drone/satellite images for larger spatial scales), considering that some adaptions and further developments of the methodology might be necessary.

We demonstrated with three examples that TIR imagery is applicable throughout the year and can reveal spatially heterogeneous surface saturation dynamics and distinct types of saturation patterns. The saturation patterns can also be used to identify different processes underlying the surface saturation formation, such as groundwater exfiltration or stream expansion. The surface saturation information ~~visualised~~visualized in the images can be used directly ~~used~~ as soft data for ~~characterising~~characterizing field conditions, for analysing ongoing hydrologic processes~~,~~ and for model validation.

The methods presented ~~methods~~ for obtaining binary, objective saturation maps from TIR images contain some uncertainties and are not automatable for datasets containing many images with varying characteristics (e.g. very wet / dry, water warmest / coldest material, slightly different field of views). In such cases, a manual choice of the temperature range for saturated pixels is the most reliable approach. Yet, for image subsets with similar conditions, the ~~tested~~ pixel classifications tested work well and we think that the combination of an automatic decomposition of the image distribution in two pixel classes and a region growing algorithm is a very promising option for obtaining objective, comparable saturation maps. In conclusion, we consider the TIR imagery a very powerful method for mapping surface saturation in terms of practicability and spatial and temporal flexibility and we believe it can provide new insights ~~in~~into the role of saturated areas and subsequent spatial and temporal dynamics in rainfall runoff transformation.

Acknowledgments

We would like to thank the three anonymous reviewers whose suggestions helped to improve the manuscript. Barbara Glaser ~~was funded~~thanks the Luxembourg National Research Fund (FNR) for funding within the framework of the FNR-AFR Pathfinder project (ID 10189601~~) by the Luxembourg National Research Fund.~~). Marta Antonelli was funded by the European Union's Seventh Framework ~~Program~~Programme for research, technological development and demonstration under grant agreement no. 607150 (FP7-PEOPLE-2013-ITN – INTERFACES – Ecohydrological interfaces as critical hotspots for transformations of ecosystem exchange fluxes and bio-geochemical cycling).

References

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

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

De Alwis, D. A., Easton, Z. M., Dahlke, H. E., Philpot, W. D. and Steenhuis, T. S.: Unsupervised classification of saturated areas using a time series of remotely sensed images, Hydrol. Earth Syst. Sci., 11(5), 1609–1620, doi:10.5194/hess-11-1609-2007, 2007.

Ambroise, B.: Variable "active" versus "contributing" areas or periods: a necessary distinction, Hydrol. Process., 18(6), 1149–1155, doi:10.1002/hyp.5536, 2004.

Antonelli, M., Klaus, J., Smettem, K., Teuling, A. J. and Pfister, L.: Exploring streamwater mixing dynamics via handheld thermal infrared imagery, Water, 9, 1–16, doi:10.3390/w9050358, 2017.

Blazkova, S., Beven, K. J. and Kulasova, A.: On constraining TOPMODEL hydrograph simulations using partial saturated area information, Hydrol. Process., 16(2), 441–458, doi:10.1002/hyp.331, 2002.

Briggs, M. A., Hare, D. K., Boutt, D. F., Davenport, G. and Lane, J. W.: Thermal infrared video details multiscale groundwater discharge to surface water through macropores and peat pipes, Hydrol. Process., n/a–n/a, doi:10.1002/hyp.10722, 2016.

Cardenas, M. B., Doering, M., Rivas, D. S., Galdeano, C., Neilson, B. T. and Robinson, C. T.: Analysis of the temperature dynamics of a proglacial river using time-lapse thermal imaging and energy balance modeling, J. Hydrol., 519, 1963–1973, doi:10.1016/j.jhydrol.2014.09.079, 2014.

Chabot, D. and Bird, D. M.: Small unmanned aircraft: precise and convenient new tools for surveying wetlands, J. Unmanned Veh. Syst., 01, 15–24, doi:10.1139/juvs-2013-0014, 2013.

Chini, M., Hostache, R., Giustarini, L. and Matgen, P.: A hierarchical split-based approach for parametric thresholding of SAR images: Flood inundation as a test case, IEEE Trans. Geosci. Remote Sens., 55(12), 6975–6988, doi:10.1109/TGRS.2017.2737664, 2017.

Creed, I. F., Sanford, S. E., Beall, F. D., Molot, L. A. and Dillon, P. J.: Cryptic wetlands: integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes, Hydrol. Process., 17(18), 3629–3648, doi:10.1002/hyp.1357, 2003.

DeAlwis, D. a., Easton, Z. M., Dahlke, H. E., Philpot, W. D. and Steenhuis, T. S.: Unsupervised classification of saturated areas using a time series of remotely sensed images, Hydrol. Earth Syst. Sci. Discuss., 4(3), 1663–1696, doi:10.5194/hessd-4-1663-2007, 2007.

Deitchman, R. S. and Loheide, S. P.: Ground-based thermal imaging of groundwater flow processes at the seepage face, Geophys. Res. Lett., 36(14), 1–6, doi:10.1029/2009GL038103, 2009.

Doppler, T., Honti, M., Zihlmann, U., Weisskopf, P. and Stamm, C.: Validating a spatially distributed hydrological model with soil morphology data, Hydrol. Earth Syst. Sci., 18(9), 3481–3498, doi:10.5194/hess-18-3481-2014, 2014.

Dugdale, S. J.: A practitioner's guide to thermal infrared remote sensing of rivers and streams: recent advances, precautions and considerations, Wiley Interdiscip. Rev. Water, doi:10.1002/wat2.1135, 2016.

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.

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.

Frey, M. P., Schneider, M. K., Dietzel, A., Reichert, P. and Stamm, C.: Predicting critical source areas for diffuse herbicide losses to surface waters: Role of connectivity and boundary conditions, J. Hydrol., 365(1-2), 23–36, doi:10.1016/j.jhydrol.2008.11.015, 2009.

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, , 1–26, doi:10.1002/2015WR018414, 2016.

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

Handcock, R. N., Gillespie, a. R., Cherkauer, K. a., Kay, J. E., Burges, S. J. and Kampf, S. K.: Accuracy and uncertainty of thermal-infrared remote sensing of stream temperatures at multiple spatial scales, Remote Sens. Environ., 100, 427–440, doi:10.1016/j.rse.2005.07.007, 2006.

Handcock, R. N., Torgersen, C. E., Cherkauer, K. A., Gillespie, A. R., Tockner, K., Faux, R. N. and Tan, J.: Thermal Infrared Remote Sensing of Water Temperature in Riverine Landscapes, in Fluvial Remote Sensing for Science and Management, edited by Carbonneau, Patrice E. and P. Hervé, pp. 85–113, Wiley-Blackwell., 2012.

Heathwaite, A. L., Quinn, P. F. and Hewett, C. J. M.: Modelling and managing critical source areas of diffuse pollution from agricultural land using flow connectivity simulation, J. Hydrol., 304(1-4), 446–461, doi:10.1016/j.jhydrol.2004.07.043, 2005.

Hewlett, J. D. and Hibbert, A. R.: Factors affecting the response of small watershed to precipitation in humid areas, For. Hydrol., 275–279 [online] Available from: http://coweeta.ecology.uga.edu/publications/851.pdf, 1967.

Ishaq, R. M. and Huff, D. D.: Application of remote sensing to the location of hydrodynamically active (source) areas, in Proceedings of 9th international symposium on remote sensing of the environment, pp. 653–666., 1974.

Keys, T. a., Jones, C. N., Scott, D. T. and Chuquin, D.: A cost-effective image processing approach for analyzing the ecohydrology of river corridors, Limnol. Oceanogr. Methods, 14, 359–369, doi:10.1002/lom3.10095, 2016.

Klaus, J., Wetzel, C. E., Martínez-Carreras, N., Ector, L. and Pfister, L.: A tracer to bridge the scales: on the value of diatoms for tracing fast flow path connectivity from headwaters to meso-scale catchments, Hydrol. Process., 29(25), 5275–5289, doi:10.1002/hyp.10628, 2015.

Kulasova, A., Beven, K. J., Blazkova, S. D., Rezacova, D. and Cajthaml, J.: Comparison of saturated areas mapping methods in the Jizera Mountains, Czech Republic, J. Hydrologuy Hydromechanics, 62(2), 160–168, doi:10.2478/johh-2014-0002, 2014a.

Kulasova, A., Blazkova, S., Beven, K., Rezacova, D. and Cajthaml, J.: Vegetation pattern as an indicator of saturated areas in a Czech headwater catchment, Hydrol. Process., 28(20), 5297–5308, doi:10.1002/hyp.10239, 2014b.

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

Li, C. J. and Ling, H.: Synthetic Aperture Radar Imaging Using a Small Consumer Drone, IEEE Int. Symp. Antennas Propagation, Vancouver, Canada, 685–686, doi:10.1109/APS.2015.7304729, 2015.

Luzi, G.: Ground based SAR interferometry: a novel tool for Geoscience, in Geoscience and Remote Sensing New Achievements, edited by P. Imperatore and D. Riccio, pp. 1–26, InTech., 2010.

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.

Matgen, P., El Idrissi, a., Henry, J. B., Tholey, N., Hoffmann, L., de Fraipont, P. and Pfister, L.: Patterns of remotely sensed floodplain saturation and its use in runoff predictions, Hydrol. Process., 20(8), 1805–1825, doi:10.1002/hyp.5963, 2006.

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.

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

Portmann, F. T.: Hydrological runoff modelling by the use of remote sensing data with reference to the 1993-1994 and 1995 floods in the river Rhine catchment, Hydrol. Process., 11(10), 1377–1392, doi:10.1002/(sici)1099-1085(199708)11:10<1377::aid-hyp533>3.0.co;2-r, 1997.

Rinderer, M., Kollegger, A., Fischer, B. M. C., Stähli, M. and Seibert, J.: Sensing with boots and trousers - qualitative field observations of shallow soil moisture patterns, Hydrol. Process., 26(26), 4112–4120, doi:10.1002/hyp.9531, 2012.

Schuetz, T. and Weiler, M.: Quantification of localized groundwater inflow into streams using ground-based infrared thermography, Geophys. Res. Lett., 38(3), L03401, doi:10.1029/2010gl046198, 2011.

Schuetz, T., Weiler, M., Lange, J. and Stoelzle, M.: Two-dimensional assessment of solute transport in shallow waters with thermal imaging and heated water, Adv. Water Resour., 43, 67–75, doi:10.1016/j.advwatres.2012.03.013, 2012.

Schwab, M. P., Klaus, J., Pfister, L. and Weiler, M.: Diel fluctuations of viscosity-driven riparian inflow affect streamflow DOC concentration, , 2177–2188, 2018.

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.

Soulsby, C., Bradford, J., Dick, J., McNamara, J. P., Geris, J., Lessels, J., Blumstock, M. and Tetzlaff, D.: Using geophysical surveys to test tracer-based storage estimates in headwater catchments, Hydrol. Process., 30, 4434–4445, doi:10.1002/hyp.10889, 2016.

Spence, C. and Mengistu, S. G.: Deployment of an unmanned aerial system to assist in mapping an intermittent stream, Hydrol. Process., 30, 493–500, doi:10.1002/hyp.10597, 2016.

Torgersen, C. E., Faux, R. N., McIntosh, B. a., Poage, N. J. and Norton, D. J.: Airborne thermal remote sensing for water temperature assessment in rivers and streams, Remote Sens. Environ., 76(3), 386–398, doi:10.1016/S0034-4257(01)00186-9, 2001.

Turner, D., Lucieer, A., Malenovský, Z., King, D. H. and Robinson, S. a.: Spatial co-registration of ultra-high resolution visible, multispectral and thermal images acquired with a micro-UAV over antarctic moss beds, Remote Sens., 6, 4003–4024, doi:10.3390/rs6054003, 2014.

Verhoest, N. E. C., Troch, P. A., Paniconi, C. and De Troch, F. P.: Mapping basin scale variable source areas from multitemporal remotely sensed observations of soil moisture behavior, Water Resour. Res., 34(12), 3235–3244, doi:10.1029/98wr02046, 1998.

Weber, I., Jenal, A., Kneer, C. and Bongartz, J.: PANTIR-a dual camera setup for precise georeferencing and mosaicing of thermal aerial images, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch., XL-3/W2, 269–272, doi:10.5194/isprsarchives-XL-3-W2-269-2015, 2015.

Whiting, J. A. and Godsey, S. E.: Discontinuous headwater stream networks with stable flowheads, Salmon River basin, Idaho, Hydrol. Process., 30, 2305–2316, doi:10.1002/hyp.10790, 2016.

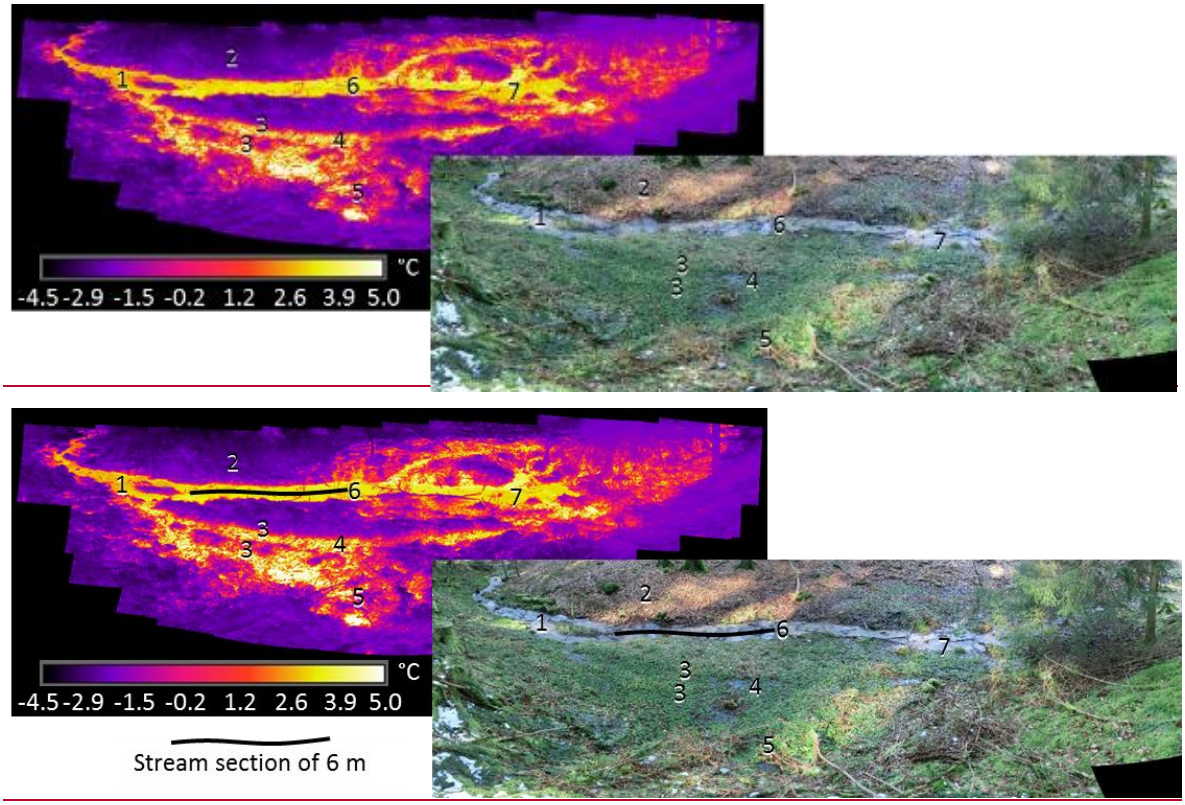
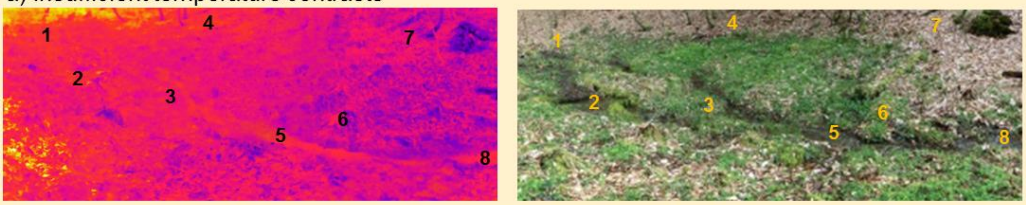


Figure 1: TIR image and VIS image of a riparian-stream zone. The temperature contrast between the water and the surrounding environment allows us to clearly differentiate between surface--saturated and dry areas in the TIR image. The numbers indicate identical locations in the TIR and VIS ~~image~~images and relate to dry areas (2), stream water (1,4,6), points of supposed groundwater exfiltration (5,7, warmer water temperatures~~))~~), and locations in which surface saturation is clearly visible in the TIR image but not in the VIS image (3, area above 6).

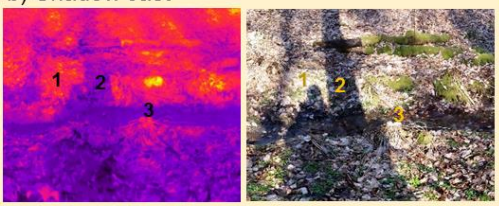
Unfavourable acquisition circumstances

a) Insufficient temperature contrasts



1, 2: stream, very difficult to identify by temperature contrast 3: dry soil with slightly different temperature than stream 4,7: hillslope with very similar temperature as stream 5,8: stream, clearly identifiable by temperature contrast 6: hillock with two contrasting temperatures, left side similar to stream

b) Shadow cast



1: sunlit stream bank 2: shadowed stream bank 3: sunlit stream

c) Rain / fog

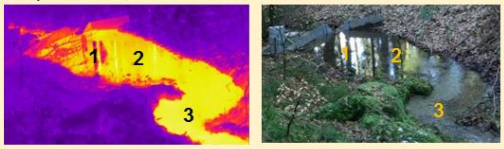
First image, no rain Next image, 8 sec later: Rain

d) (Temporary) view obstructions

1: uncovered stream the growing season 2-5: fern covering the stream temporary during 6: tree trunk permanently covering soil behind

e) Additional effects to consider

such as reflection, image vignetting, unrealistic temperatures, automatic recalibration of the camera



1: reflections visible in TIR strong reflections 2: reflections not visible in TIR 3: no

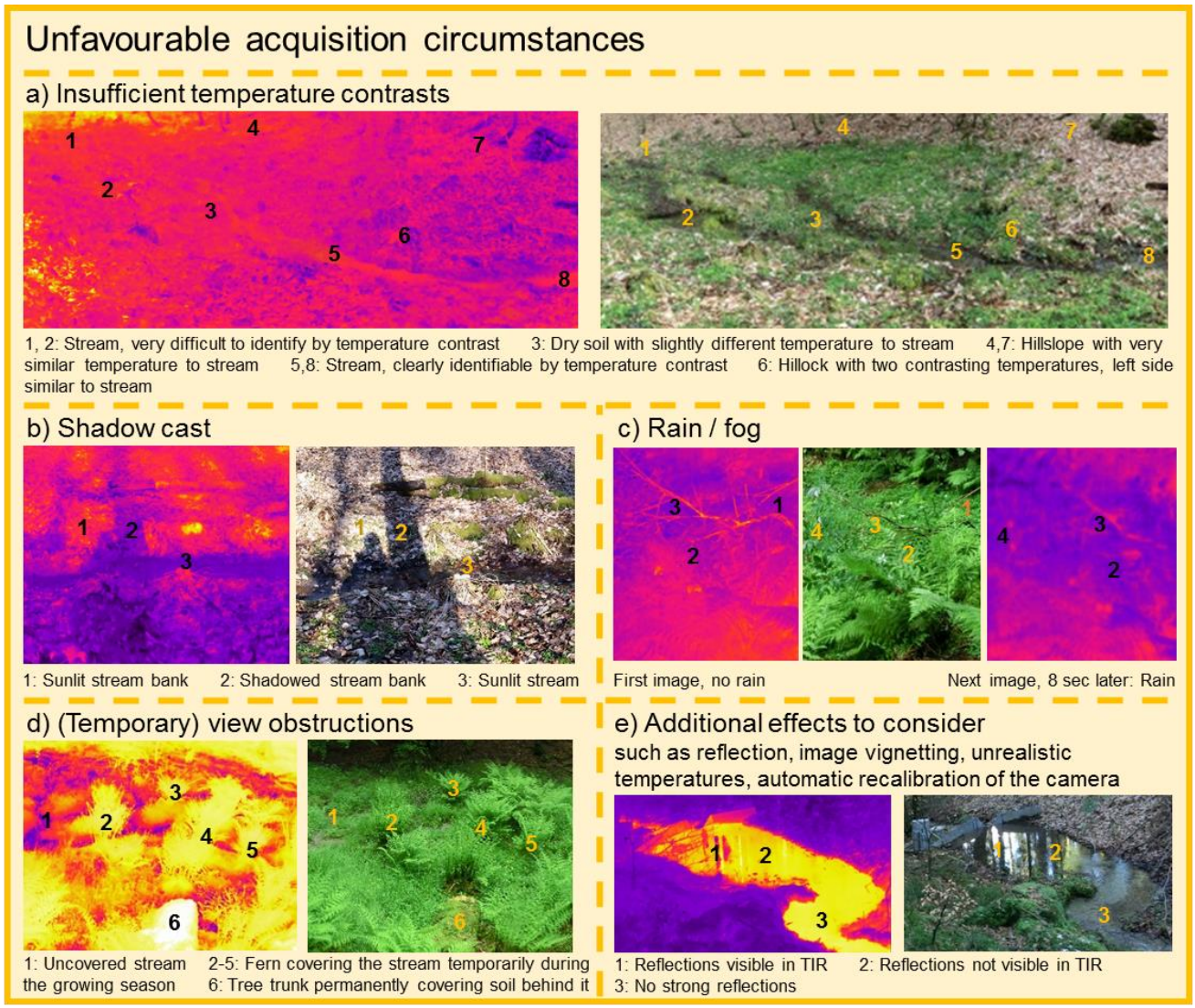
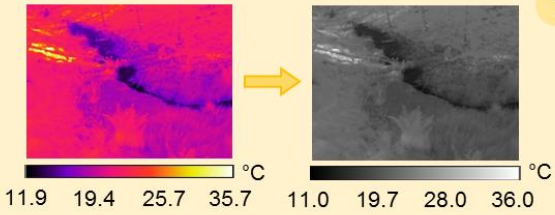


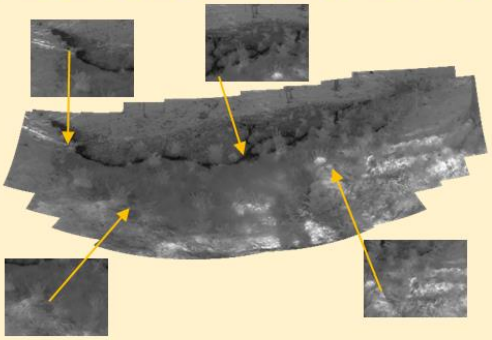
Figure 2: Example images showing how unfavourable image acquisition circumstances influence the usability of TIR imagery for the identification of surface saturation. The numbers indicate identical locations in the TIR and VIS images.



1 Colour conversion

- Transform the colour scale to a linear, grey scale, identical for all single images / video frames

Note:Save images / videos without embedded colour bar or any other disturbing element, but note down chosen temperatrue range to keep the tempereature information

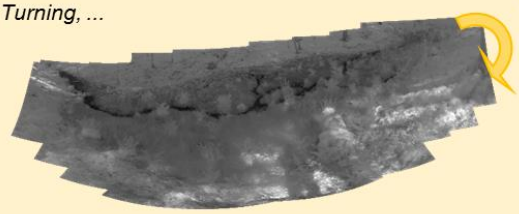


2 Generation of panorama images (PI)

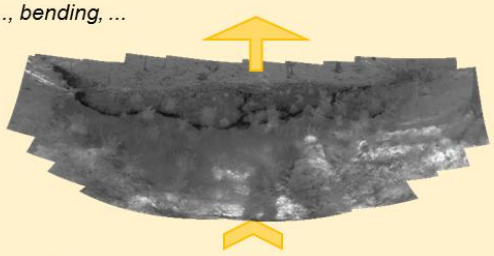
- Mosaick single images / video frames with existing open-source or fee-based software (e.g. PTGui, ICE)
- Manually adjust image alignment and overlaps if needed

Note: Save PI in a format that allows to denote blank values originating from the non-rectangular shape of the created panoram (e.g. TIFF including transparency values)

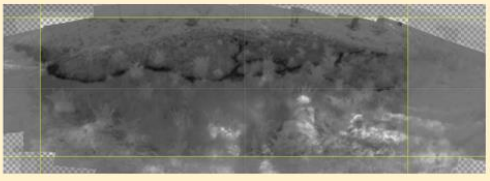
Turning, ...



..., bending, ...



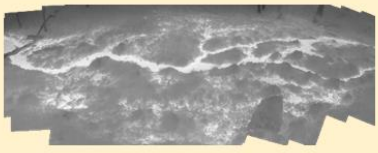
..., cutting ....



3 Co-registration of panoramas

- Select a reference panorama (RPI) with identifiable, permanent existing features (e.g. stones, course of streambed) and without extreme wet or dry conditions
- Identify corresponding locations between PI and RPI, spread over the entire image area
- Apply a co-registration algorithm (e.g. implemented in ArcGIS, Matlab, PTGui) that adapts the projection and field of view of the PI to the RPI without modifying the RPI
- Cut PI to the same section with the same pixel extent as the RPI

Note: Co-registration is eased and distortion reduced if the projection and field of view of PI and RPI are similar before applying a co-registration algorithm (→ adjust projection to RPI before saving created PI, step 2)



... to same perspective and size as reference panorama RPI

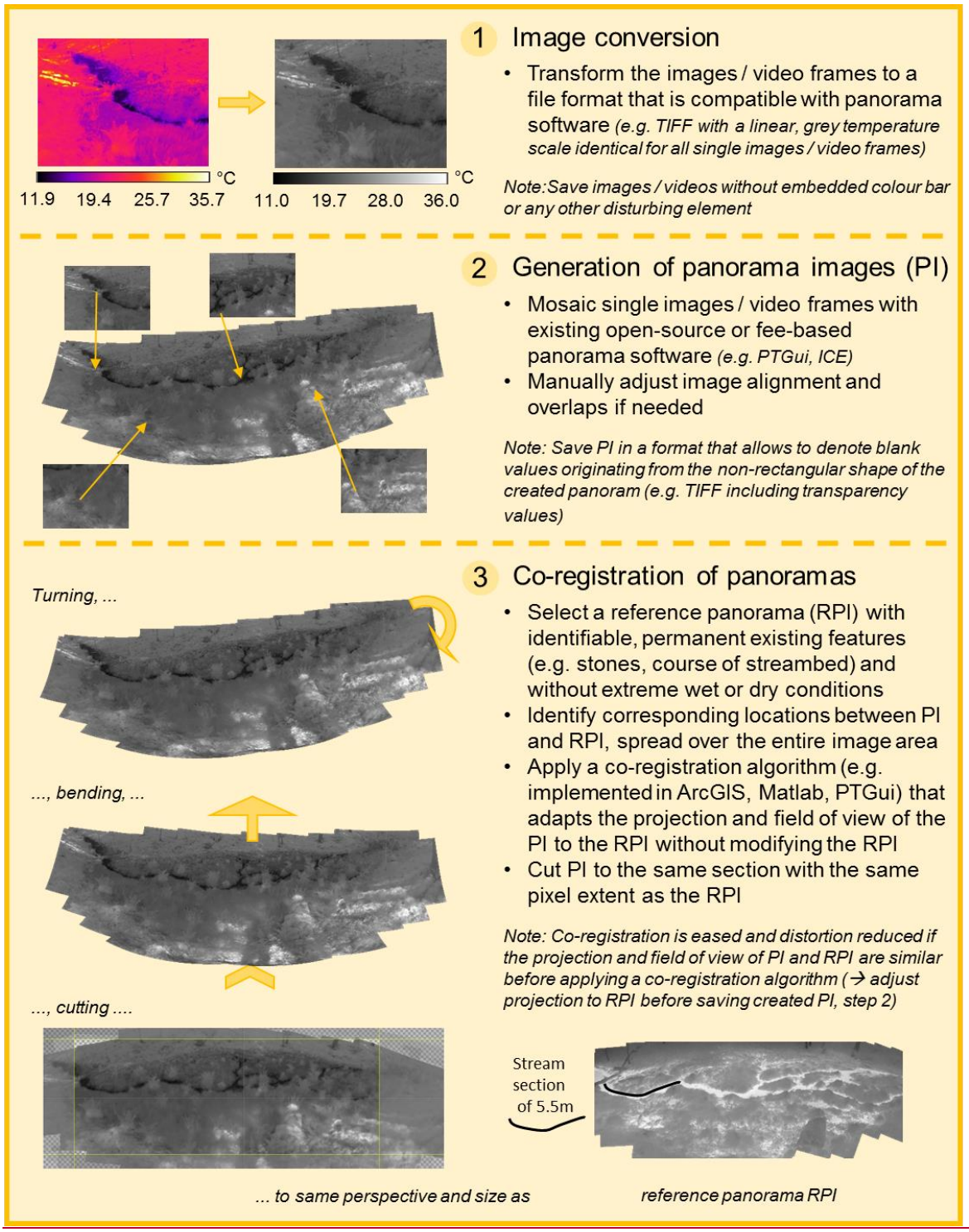
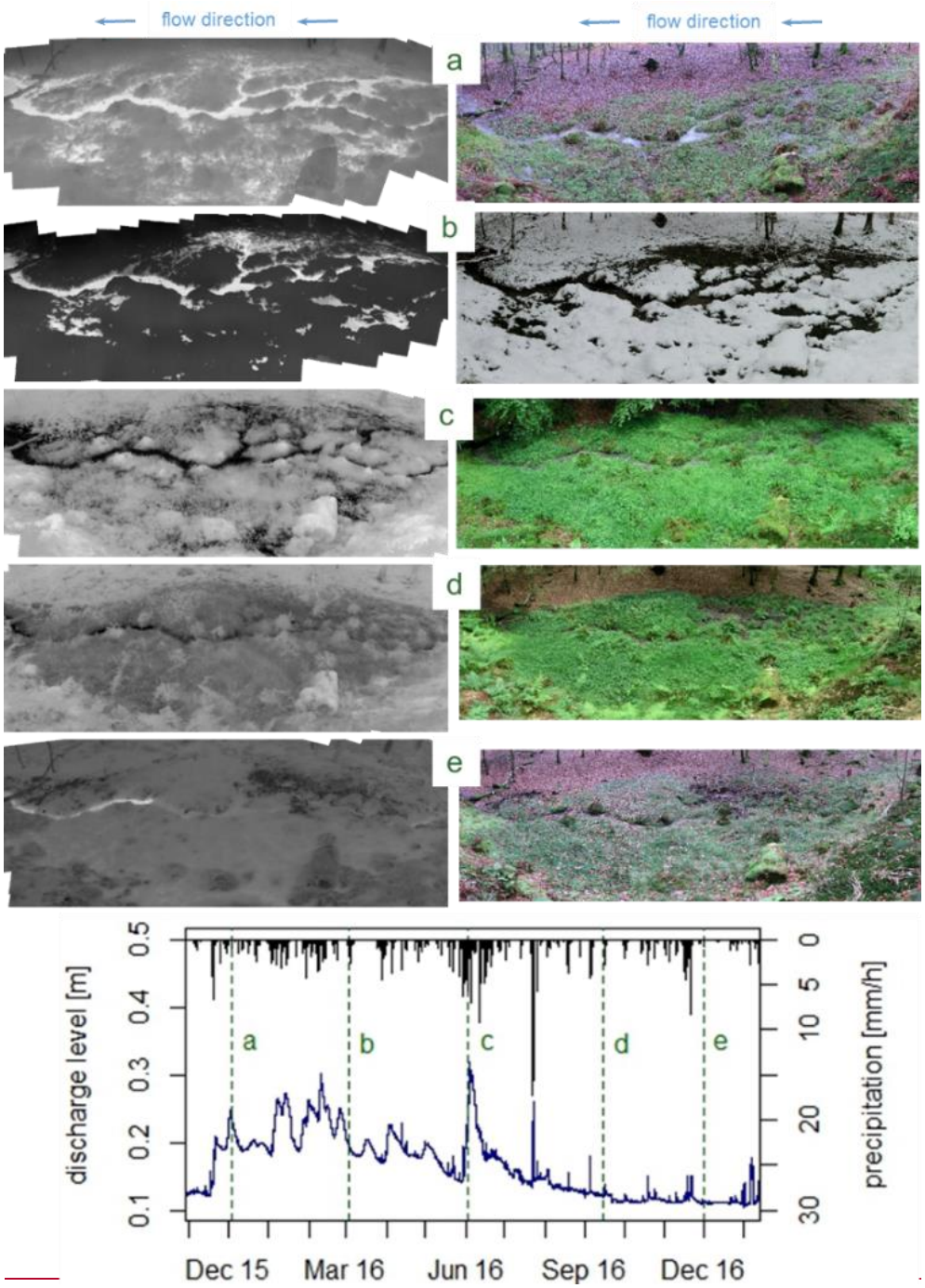


Figure 3: Workflow for processing single TIR images / video frames to co-registered panoramic images.


flow direction
flow direction
a
b
c
d
e
discharge level [m]
0.1
0.2
0.3
0.4
0.5
precipitation [mm/h]
0
5
10
20
30
a
b
c
d
e
Dec 15
Mar 16
Jun 16
Sep 16
Dec 16

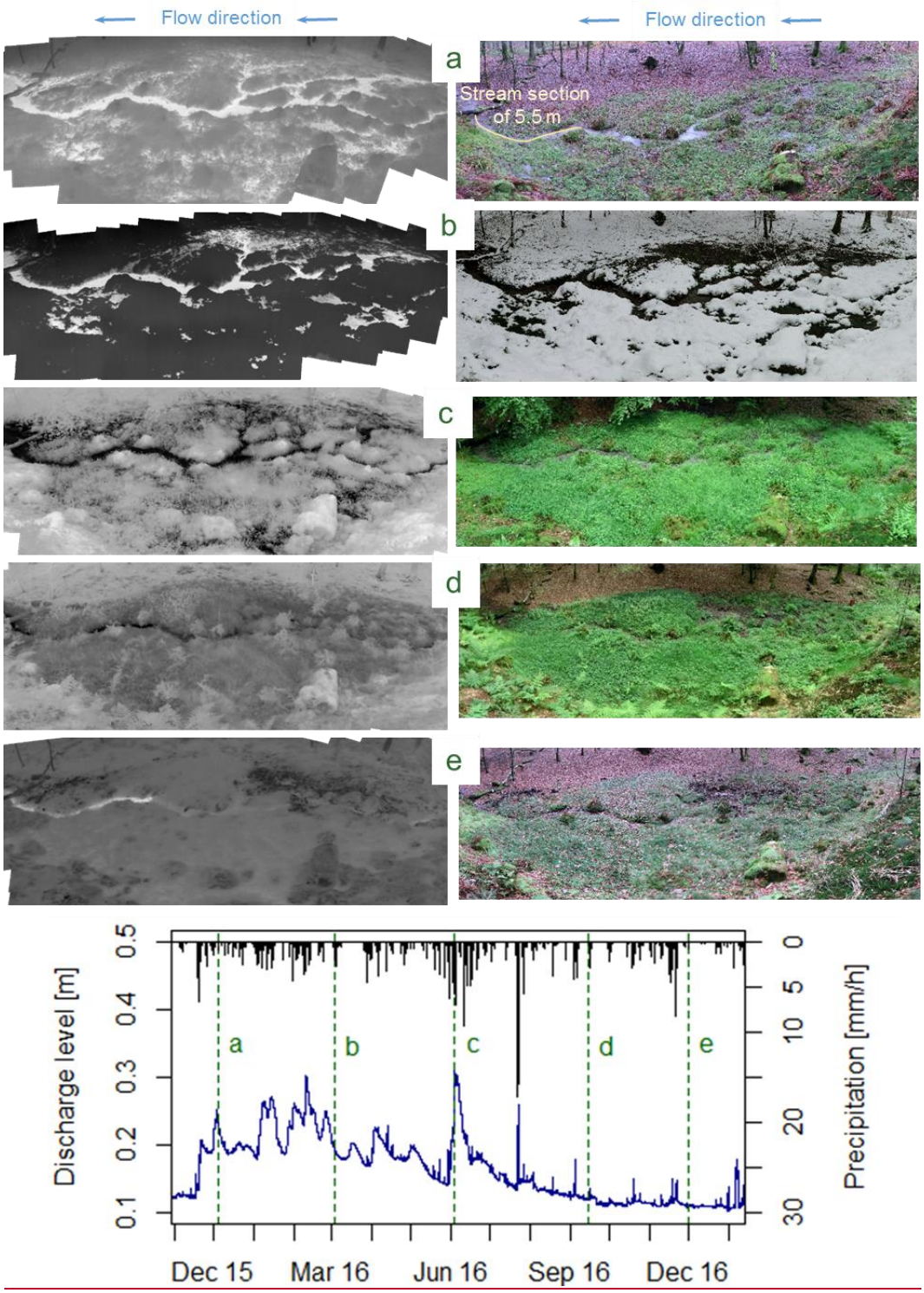
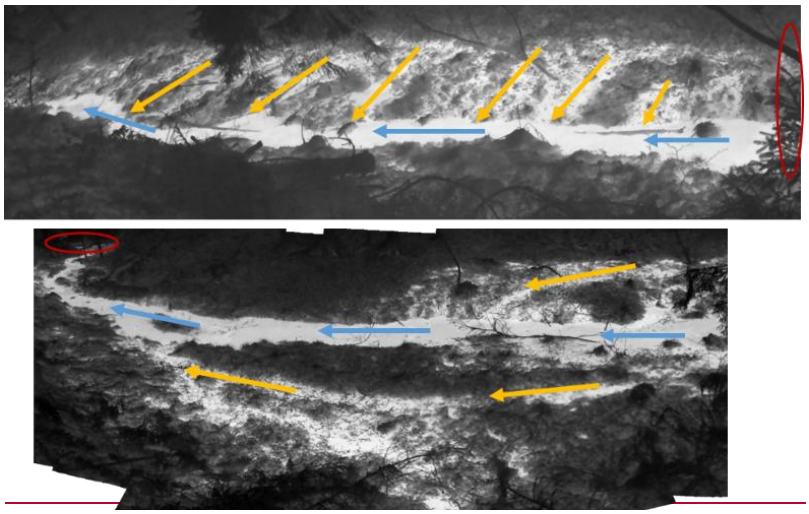


Figure 4: Time lapse TIR and VIS panoramas showing the variation of surface saturation patterns with varying discharge levels under diverse seasonal conditions.



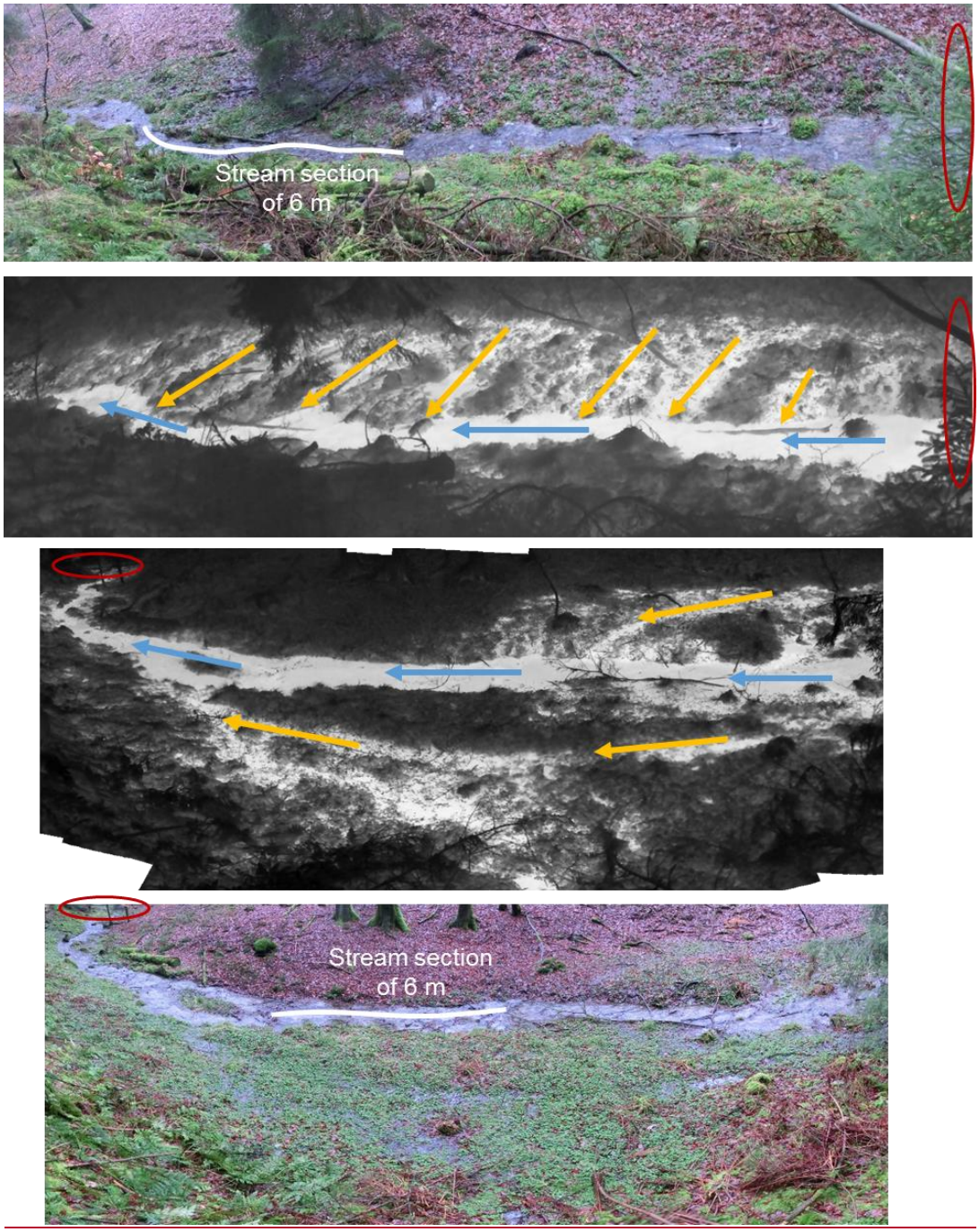


Figure 5: Comparison of different types of surface saturation patterns. The yellow arrows indicate the orientation of the saturated areas towards the stream (blue arrows = flow direction). The perpendicular direction (top) is likely caused by exfiltrating groundwater connecting to the stream, the parallel direction (bottom) by a parallel flow of the stream expanding into the riparian zone. The red ovals indicate where the two panorama images connect.

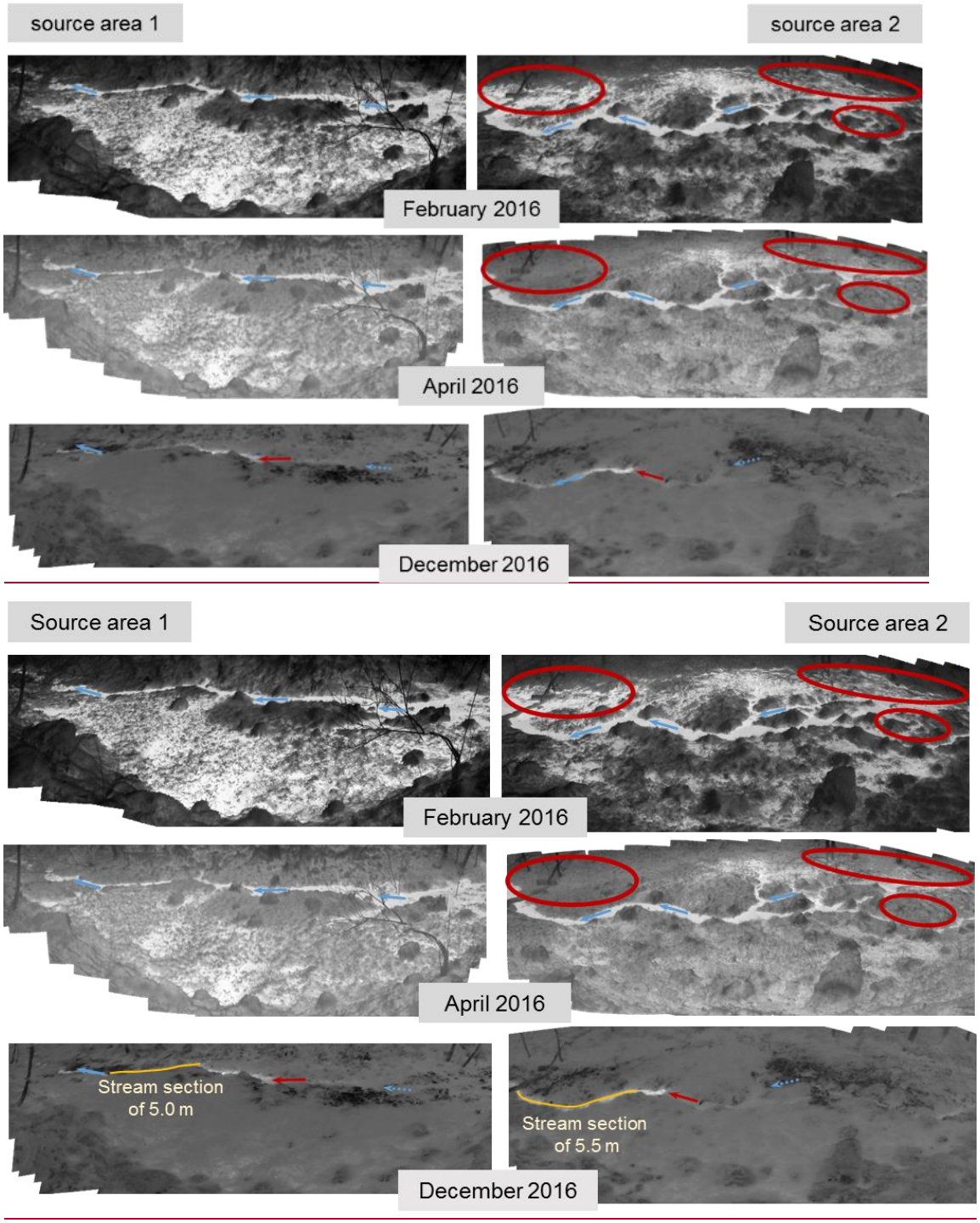
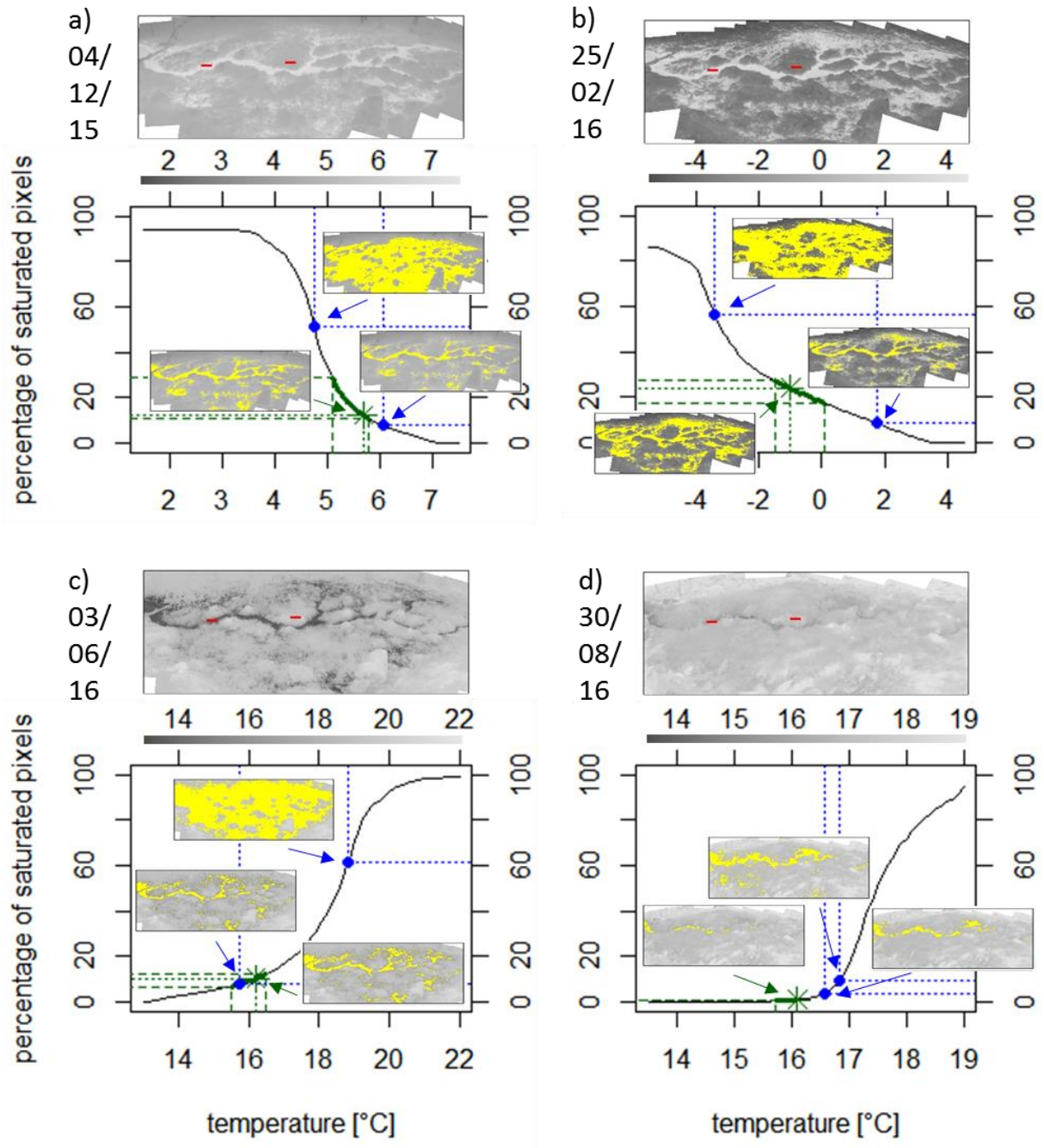


Figure 6: Transition of two source areas (left vs. right) from very wet (top) to very dry conditions (bottom). Surface saturation in source area 1 (left) barely changed between February and April 2016, whereas source area 2 is clearly drier at some locations (red ovals) in April 2016. In December 2016, both source areas were completely dry ~~aside~~on each side of the stream (blue arrows = flow direction) and the stream started further downstream (red arrows).


a)
04/
12/
15
b)
25/
02/
16
c)
03/
06/
16
d)
30/
08/
16
percentage of saturated pixels
temperature [°C]

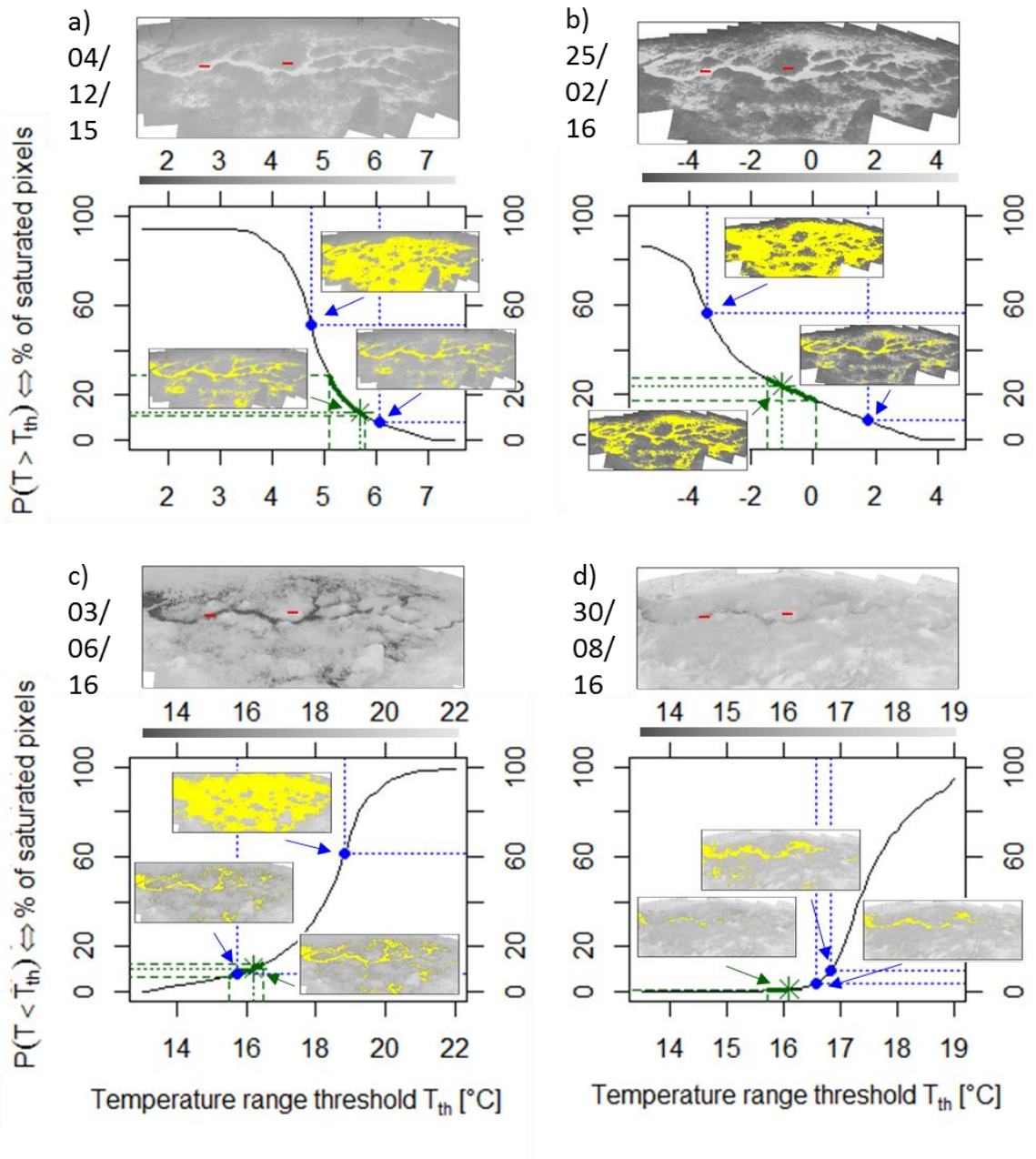


Figure 7: Example TIR images with their cumulative saturation curves showing the percentage of pixels that have a higher (a, b) or lower (c, d) temperature than the temperature range threshold T_{th} and are thus defined as saturated (marked as yellow pixels in the inset TIR images). The green asterisks mark the temperature ranges that were manually chosen as optimum ~~from~~following a visual assessment of the images. Green dashed lines define the uncertainty of the optimum temperature ranges. ~~Red~~The red rectangles in the TIR images ~~are~~depict the ~~mask~~masks used for the identification of temperature ranges from a constantly wet (left) and constantly dry (right) area. The respective temperature ranges and saturation percentages are marked in blue. For reference for the spatial dimension of the images, we refer the reader to the indicated stream section in Fig. 3 or Fig. 4.

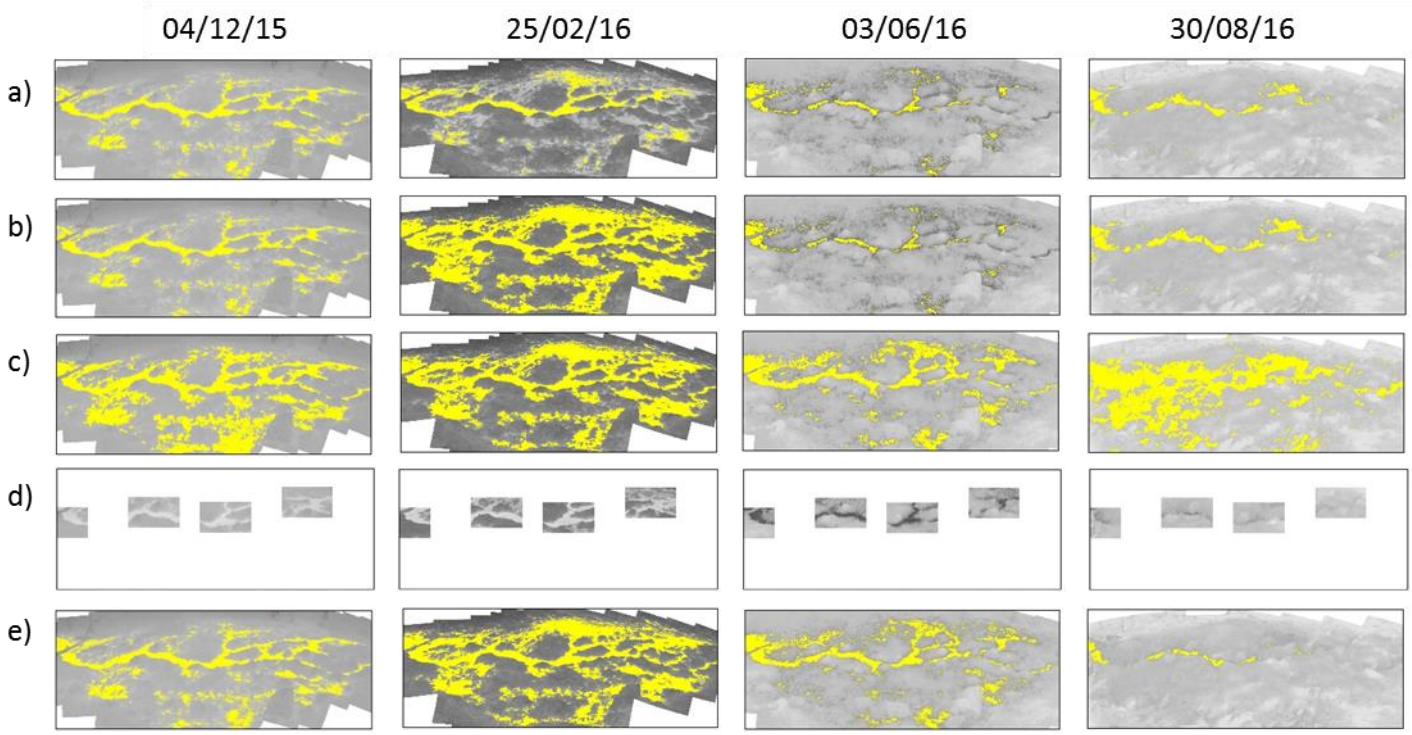


Figure 8: Comparison of saturation maps (yellow = saturation) generated with a region growing process whose seeds and stopping criteria were automatically constrained to a) bimodal distributions derived from the HBSA applied to the entire image, b) bimodal distributions derived from the HBSA where the selection of bimodal image subsections was constrained to image-specific manual predefinitions of temperature ranges of saturation, c) bimodal distributions derived from pre-selected parts of the image (shown in d) that include clearly wet and dry areas. The saturation maps generated with ~~manual~~manually selected temperature ranges based on visual assessment (cf. Fig. 7, green asterisk) are shown for comparison (e). For reference for the spatial dimension of the images, we refer the reader to the indicated stream section in Fig. 3 or Fig. 4.