Cumulative Dissertation

Hillslope-Stream Connectivity across Scales

For the degree of Doctor of Natural Sciences (Dr. rer. nat.) in Geoecology

by Lisa Angermann

Submitted to the
Faculty of Mathematics and Natural Sciences, Institute of Earth and Environmental Sciences at the University of Potsdam, Germany

Prepared at the
Section of Hydrology of the German Research Centre for Geosciences (GFZ)

Submitted: June 20, 2018
Defended: December 20, 2018

Referees:
Bruno Merz — University of Potsdam, Germany
Theresa Blume — GFZ, Germany
Kevin J. McGuire — Virginia Tech, USA

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-42454
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-424542

Hillslope-Stream Connectivity across Scales

Lisa Angermann

June 2018

Contents

List of Figures

List of Tables

Summary

The concept of hydrologic connectivity summarizes all flow processes that link separate regions of a landscape. As such, it is a central theme in the field of catchment hydrology, with influence on neighboring disciplines such as ecology and geomorphology. It is widely acknowledged to be an important key in understanding the response behavior of a catchment and has at the same time inspired research on internal processes over a broad range of scales. From this process-hydrological point of view, hydrological connectivity is the conceptual framework to link local observations across space and scales.

This is the context in which the four studies this thesis comprises of were conducted. The focus was on structures and their spatial organization as important control on preferential subsurface flow. Each experiment covered a part of the conceptualized flow path from hillslopes to the stream: soil profile, hillslope, riparian zone, and stream.

For each study site, the most characteristic structures of the investigated domain and scale, such as slope deposits and peat layers were identified based on preliminary or previous investigations or literature reviews. Additionally, further structural data was collected and topographical analyses were carried out. Flow processes were observed either based on response observations (soil moisture changes or discharge patterns) or direct measurement (advective heat transport). Based on these data, the flow-relevance of the characteristic structures was evaluated, especially with regard to hillslope to stream connectivity.

Results of the four studies revealed a clear relationship between characteristic spatial structures and the hydrological behavior of the catchment. Especially the spatial distribution of structures throughout the study domain and their interconnectedness were crucial for the establishment of preferential flow paths and their relevance for large-scale processes. Plot and hillslope-scale irrigation experiments showed that the macropores of a heterogeneous, skeletal soil enabled preferential flow paths at the scale of centimeters through the otherwise unsaturated soil. These flow paths connected throughout the soil column and across the hillslope and facilitated substantial amounts of vertical and lateral flow through periglacial slope deposits.

In the riparian zone of the same headwater catchment, the connectivity between hillslopes and stream was controlled by topography and the dualism between characteristic subsurface structures and the geomorphological heterogeneity of the stream

channel. At the small scale (1 m to 10 m) highest gains always occurred at steps along the longitudinal streambed profile, which also controlled discharge patterns at the large scale (100 m) during base flow conditions (number of steps per section). During medium and high flow conditions, however, the impact of topography and parafluvial flow through riparian zone structures prevailed and dominated the large-scale response patterns.

In the streambed of a lowland river, low permeability peat layers affected the connectivity between surface water and groundwater, but also between surface water and the hyporheic zone. The crucial factor was not the permeability of the streambed itself, but rather the spatial arrangement of flow-impeding peat layers, causing increased vertical flow through narrow "windows" in contrast to predominantly lateral flow in extended areas of high hydraulic conductivity sediments.

These results show that the spatial organization of structures was an important control for hydrological processes at all scales and study areas. In a final step, the observations from different scales and catchment elements were put in relation and compared. The main focus was on the theoretical analysis of the scale hierarchies of structures and processes and the direction of causal dependencies in this context. Based on the resulting hierarchical structure, a conceptual framework was developed which is capable of representing the system's complexity while allowing for adequate simplifications.

The resulting concept of the parabolic scale series is based on the insight that flow processes in the terrestrial part of the catchment (soil and hillslopes) converge. This means that small-scale processes assemble and form large-scale processes and responses. Processes in the riparian zone and the streambed, however, are not well represented by the idea of convergence. Here, the large-scale catchment signal arrives and is modified by structures in the riparian zone, stream morphology, and the small-scale interactions between surface water and groundwater. Flow paths diverge and processes can better be represented by proceeding from large scales to smaller ones. The catchment-scale representation of processes and structures is thus the conceptual link between terrestrial hillslope processes and processes in the riparian corridor.

Zusammenfassung

Das Konzept der hydrologischen Konnektivität umfasst alle Fließprozesse, welche verschiedene Bereiche einer Landschaft verbinden. Als solches ist es ein zentrales Thema in dem Forschungsbereich der Einzugsgebietshydrologie und beeinflusst auch benachbarte Disziplinen wie die Ökologie oder die Geomorphologie. Es ist allgemein akzeptiert, dass das Konzept der Konnektivität ein wichtiger Schlüssel zum Verständnis von Einzugsgebietsdynamiken ist, gleichzeitig inspiriert es die Erforschung interner Prozesse auf verschiedenen Skalen. Von dieser prozesshydrologischen Perspektive gesehen, bietet Konnektivität einen konzeptionellen Rahmen, um lokale Beobachtungen über Raum und Skalen miteinander in Verbindung zu setzen.

In diesem Kontext stehen die vier Studien dieser Doktorarbeit. Der Fokus lag dabei auf räumlichen Strukturen als wichtigem Kontrollfaktor für präferentielle Fließpfade als spezieller Form unterirdischer Fließprozesse. Die Experimente deckten dabei je einen Abschnitt des konzeptionellen Fließweges vom Hang zum Bach exemplarisch ab: Bodenprofil und Hang, Hang und Auenbreich, und Bachbett.

Für alle vier Studien wurden zunächst charakteristische Strukturen des Untersuchungsgebietes wie Schuttablagerungen am Hang oder Torfschichten im Flussbett auf Basis vorausgehender Untersuchungen und Literaturrecherchen identifiziert. Zusätzlich wurden weitere strukturelle Daten erfasst und digitale Geländemodelle ausgewertet. Anschließend wurde die Prozessrelevanz dieser Strukturen, vor allem im Hinblick auf die Hang-Bach-Konnektivität, untersucht.

Die Ergebnisse der einzelnen Studien zeigten eine deutliche Verbindung zwischen den charakteristischen räumlichen Strukturen und dem hydrologischen Verhalten des untersuchten Gebietes. Insbesondere die räumliche Anordnung von Strukturen, d.h. die räumliche Verteilung und der Grad der Konnektivität der Strukturen, war ausschlaggebend für die Ausbildung präferenzieller Fließpfade und deren Relevanz für größerskalige Prozesse. Die räumliche Organisation von Strukturen war in allen Untersuchungsgebieten ein wichtiger Kontrollfaktor für hydrologische Prozesse.

Die Beobachtungen auf verschiedenen Skalen und verschiedener Fließpfadabschnitte wurden miteinander in Verbindung gesetzt und verglichen. Besonderes Augenmerk lag dabei auf der theoretischen Analyse der Skalenhierarchie von Strukturen und Prozessen und der Richtung der Kausalität in diesem Zusammenhang. Auf dieser Grundlage wurde als Synthese der einzelnen Studien ein Konzept entwickelt, welches in der Lage ist, die Komplexität eines Einzugsgebietes abzubilden und gleichzeitig adequate Vereinfachungen zuzulassen.

Diese Konzept der parabelförmigen Skalenabfolge beruht auf der Erkenntnis, dass Fließprozesse im terrestrischen Bereich eines Einzugsgebietes, also im Boden und den Hängen, vorwiegend konvergieren und sich von der kleinen Skala zur größeren hin zusammenfügen. Die Prozesse in der Aue und dem Bachbett werden von diesem Prinzip der Konvergenz allerding nicht abgebildet. Die in den Böden und an den Hängen erzeugten Fließmuster des Einzugsgebiets werden von den Strukturen in der Aue, der Morphologie des Baches und den kleinskaligen Wechselwirkungen zwischen Fließgewässer und Sediment überprägt. Die Fließprozesse divergieren, und eine Beschreibung von der großen Skala hin zur kleineren ist hier besser geeignet. Die räumlich diskrete oder konzeptionelle Darstellung von Prozessen auf der Einzugsgebietsskala bietet so die Verbindung zwischen terrestrischer Hanghydrologie und der bachseitigen Auenhydrologie.

List of Publications

L. Angermann, C. Jackisch, N. Allroggen, M. Sprenger, E. Zehe, J. Tronicke, M. Weiler, T. Blume. Form and function in hillslope hydrology: characterization of subsurface flow based on response observations. Hydrology and Earth System Sciences, 21(7):3727 – 3748, 2017.

C. Jackisch, L. Angermann, N. Allroggen, M. Sprenger, T. Blume, J. Tronicke, E. Zehe. Form and function in hillslope hydrology: in situ imaging and characterization of flow-relevant structures. Hydrology and Earth System Sciences, 21(7):3749 – 3775, 2017.

L. Angermann, M. Weiler, T. Blume. The competing impact of stream morphology and catchment topography on hillslope-stream connectivity. *In preparation.*

L. Angermann, S. Krause, J. Lewandowski. Application of heat pulse injections for investigating shallow hyporheic flow in a lowland river. Water Resources Research, 48(12), 2012.

Chapter 1

Introduction

1.1 Hydrologic connectivity

The term connectivity in its most general linguistic use describes the quality, state, or capability of being connective or connected [Merriam-Webster]. As such, it is at the heart of hydrology, the study of the water cycle, connecting land, oceans and atmosphere. Understanding a hydrological system means to understand when and how it is connected, within itself or to other neighboring systems. Accordingly, the term is used extensively in hydrological publications, as technical term or in its general linguistic sense, to describe behavior, processes and structures of the investigated system.

As a technical term, connectivity of hydrological systems is a very broad concept used by many researchers and disciplines [e.g., Western et al., 2001, Lane et al., 2004, Michaelides and Chappell, 2009, Wainwright et al., 2011], resulting in a wide variety of approaches and definitions, which are not necessarily compatible to each other. The applicability of different definitions is mainly based on the research scale and perspective. Especially the latter one defines which structural features and/or transport processes are considered relevant. Depending on whether the research focus is on the transport of water, nutrients [e.g., Stieglitz et al., 2003], energy, [e.g., Ward, 1997], sediment [e.g., Fryirs et al., 2007], or organisms [e.g., Amoros and Bornette, 2002, Pringle, 2003], different concepts and definitions of connectivity may be applicable or suitable.

Bracken et al. [2013] compiled a list of available definitions and concluded that even among hydrologists there is not one single overarching definition which satisfies

all applications. However, in most generic terms hydrologic connectivity describes all kinds of water-mediated transport of mass, matter, energy or organisms within or between elements of the hydrologic cycle [Pringle, 2001]. A slightly stricter definition that resonates especially well with the processes-oriented hydrological community, and which will be basis of the following elaborations, describes hydrologic connectivity as the linkage of separate regions of a catchment via water flow [Blume and van Meerveld, 2015]. This definition emphasizes the actual movement of water through the system (as opposed to the function of water as a medium facilitating movement, e.g., migration of organisms) and defines a catchment as the domain in which connectivity takes effect.

Hydrologic connectivity is widely acknowledged to be an important factor in understanding processes and system behavior [Michaelides and Chappell, 2009]. Studies that explicitly address connectivity at the catchment scale do so by relating temporal dynamics of water tables or tracer signals in the stream with those from the catchment [functional connectivity Tetzlaff et al., 2007]. Another way to explore connectivity is by analyzing spatial structures such as topography or subsurface structures [structural connectivity Kirkby et al., 2002, Lane et al., 2009]. Besides these catchment-scale approaches, many studies focus on detailed and smaller-scale process understanding in different components of a catchment. These studies are not connectivity research *per se*, but provide the detailed nested-scale knowledge that is required for understanding the nature and dynamics of connectivity [e.g., Lexartza-Artza and Wainwright, 2009].

A field of research that creates process knowledge supporting our connectivity understanding is the field of hillslope hydrology. Here, the focus is on investigating the spatial and temporal characteristics of surface and subsurface flow paths across hillslopes as processes that establish connectivity, coining the term hillslope-stream connectivity as a sub-category of hydrologic connectivity [see Blume and van Meerveld, 2015, for a review of methods and approaches in this field].

Another field of hydrologic research that falls into the scope of catchment hydrologic connectivity is the field of surface water-groundwater (SW-GW) interaction [1]. Research on SW-GW interaction has gained attention during the last decades, also kindled by the call for perceiving and managing water resources coupled surface water-groundwater systems rather than as discrete entities, formulated by re-

[1]While the term "surface water" and related concepts are generally applicable to all surface water bodies, this thesis focuses on streams and rivers and uses the term surface water in this context

searchers and water management authorities [European Comission, 2000, Fleckenstein et al., 2010].

Both fields, hillslope hydrology and SW-GW interaction, investigate spatial characteristics, temporal dynamics and the relevance of connectivity establishing flow paths between catchment and stream. But despite the spatial and conceptual proximity of both fields, they remain separate research communities. This might be due to different perspectives which support or require different research foci, approaches, and scales, which will be summarized very briefly in the following.

1.1.1 Hillslope hydrology and connectivity

Hillslope hydrology is a research field of its own, and most studies do not explicitly focus on hydrological connectivity. Nevertheless, hillslope hydrological observations often yield important knowledge on how and when hillslopes connect to the stream [Blume and van Meerveld, 2015].

Many hillslope hydrological studies directly investigate the establishment and spatial characteristics of flow paths. Methods applied in this context include spatially distributed piezometers and groundwater wells, based on the assumption of connectivity through connected (perched) water tables [e.g., Tromp-Van Meerveld and McDonnell, 2006, Jencso et al., 2009], as well as observation trenches that provide information on surface and subsurface lateral flow [e.g., Bachmair and Weiler, 2012] or dye staining and excavation [e.g., Graham et al., 2010], which are capable of targeting preferential flow through discrete flow paths. Futher, geophysical methods for subsurface explorations, sometimes in combination with tracer applications, have been used to identify flow paths and residence time distributions [e.g., Holden, 2004a, Cassiani et al., 2006, Koch et al., 2009]. The investigated processes may finally lead to hillslope-stream connectivity, and detailed knowledge about the internal behavior of a hillslope significantly improves our mechanistic system understanding, also with regard to connectivity.

1.1.2 Surface water-groundwater interaction and connectivity

Research on SW-GW interaction usually has a surface water perspective. SW-GW interaction and hyporheic exchange is often treated as a property of the investigated

surface water body. Concepts like transient storage and hyporheic residence time distributions [Haggerty et al., 2002, Gooseff et al., 2003, Bencala et al., 2011] and other parameters describing hyporheic turnover are examples that are characteristic for this perception. As such, the interaction of the surface water body with the shallow groundwater is frequently used as a means to explain or target issues like water quality [Kalbus et al., 2006], habitat properties [e.g., Montgomery, 1999], and biogeochemical turnover [Brunke and Gonser, 1997] of a surface water system interacting with its surroundings.

In this context the hyporheic zone takes on a special role. It was named in recognition of a distinct subsurface biotop or ecotone [Schwoerbel, 1961] beneath the stream, which is characterized by the interplay of the adjacent water bodies. The term was quickly adopted by hydrologists investigating SW-GW interactions at multiple scales [Bencala, 1993]. In order to emphasize the difference between small-scale (1 m to 10 m), bi-directional exchange between surface water and sediment that conceptually happens along the entire stream, and larger-scale exchange flow, the term "parafluvial flow" has been used increasingly for the somewhat larger scale (10 m to 100 m). Depending on the definition, parafluvial flow includes longer and deeper flow paths more or less parallel to the riparian corridor that connect loosing and gaining reaches of a stream [Cook, 2013, Cartwright and Hofmann, 2016], but also flow through point bars and gravel banks [Cartwright et al., 2014] and the lateral riparian zone [Edwardson et al., 2003].

The process-oriented investigation of SW-GW interaction usually focuses on near-stream processes. The considered part of the groundwater body barely exceeds the riparian zone, and hillslope processes are rarely considered directly. Scales typically associated with SW-GW interactions and hyporheic exchange are in the range of 1 m to 10 m [Bencala et al., 2011].

1.1.3 Benefits from combining perspectives of connectivity

The two research fields summarized above represent the fluvial and terrestrial "ends" of the catchment domain in which connectivity takes effect. Or, transferring it to a very simplified conceptualization of the flow path through the catchment, the beginning and the end of that flow path. These two are linked via manifold processes across multiple scales. The investigation of these linking processes, their spatial extent and temporal dynamics is the goal behind studies on hydrological connectivity in a strict sense.

Approaches to catchment-scale connectivity can be categorized into two main classes: process-oriented and structure-oriented approaches. Indicators for structural connectivity are topography [e.g., Mayor et al., 2008], various wetness or flow indices [e.g., Lane et al., 2009, Ali et al., 2014], or subsurface heterogeneity [e.g., Western et al., 2001], or in general geostatistical measures as universal measure for hydrological connectivity [Michaelides and Chappell, 2009], *potentially* facilitating the establishment of flow or travel paths. This approach is particularly common among ecologists, who are interested in the connectivity of aquatic habitats [e.g., Taylor et al., 1993], and geomorphologists [e.g., Bracken et al., 2015] but is also widely applied by hydrologists. Especially at the large scale, GIS based geospatial analyses are a convenient tool to classify and categorize entire catchments based on remote sensing data [e.g., Lang et al., 2012].

Regarding the process-oriented approach, Blume and van Meerveld [2015] distinguish stream-centered and hillslope-centered methods, but remark that studies following hillslope-centered approaches are often not explicitly focused on connectivity. Stream-centered methods are based on the interpretation of the response signal in the stream as fingerprint of the internal processes in the catchment. Methods include tracers such as stable isotopes [e.g., Ala-Aho et al., 2018], radon, temperature [Selker et al., 2006], incremental streamflow gauging [Briggs et al., 2012, Payn et al., 2012], and hydrograph separation [Buttle, 1994, Klaus and McDonnell, 2013]. The focus often is on the distinction of different runoff generating source areas, based on catchment elements (riparian zone, hillslopes, plateaus, e.g., McGlynn and McDonnell [2003]), landscape properties (geology, pedology, vegetation, etc., e.g., Jencso and McGlynn [2011]), or structures (topography, surface and subsurface structures, e.g., Hopp and McDonnell [2009]).

Many studies on hydrological connectivity already combine connectivity measures and process observations in the stream, the riparian zone or at hillslopes. For a few well studied experimental hillslopes, a broad data base was accumulated which allows for a more holistic interpretation. Examples for such well studied catchments are the Maimai catchment [Brammer and McDonnell, 1996, McGlynn et al., 2002, Gooseff and McGlynn, 2005] or the Panola Mountain Research Watershed [Burns et al., 2001, Peters et al., 2003, Tromp-van Meerveld and McDonnell, 2006]). However, such vast data sets are the exception to the rule. Catchments are too complex to investigate in their entirety with adequate effort in the frame of one or few studies, and data sets covering different elements of a catchment are scarce.

This is a dilemma, not only for the sake of scientific progress. Bencala et al. [2011] identified the need to study hyporheic flow and transient storage at the intermediate scale (10 m to 1000 m), as this scale is relevant for many policy and management issues such as river restoration [see also Kondolf et al., 2006, Hester and Gooseff, 2010], aquatic ecosystem health [Tetzlaff et al., 2007], and even land use and watershed management [Wright et al., 2005]. At this intermediate scale, SW-GW interactions are "inherently messy" [Bencala et al., 2011], as the complexity of the stream-catchment system becomes apparent, and the influence of the variability of surface und subsurface transport processes in the catchment becomes more and more important [Woessner, 2000]. Understanding the multi-scale dependencies between local processes and large-scale catchment behavior are an important factor in our hydrological understanding of catchments, which also has implications for related fields:

Water quality: The fate of nutrients and pollutants in a stream is controlled by their transport through and out of the terrestrial part of the catchment [e.g., Musolff et al., 2015], but also their turnover in the stream and the hyporheic zone [e.g., Mulholland et al., 2008]. While processes can be studied separately or at least partially isolated to a certain extent, management decisions need to be based on a more holistic basis, considering catchment and in-stream processes.

Hydrogeology: From a hydrogeological perspective, the management of riparian groundwater but also stream health and flood plain ecological functioning would benefit from including riparian processes into groundwater research and management approaches [Woessner, 2000]. This requires the adaption of traditional hydrogeological approaches to scales that fit processes in the riparian zone and the stream channel.

Ecology: Hydrogeomorphology and related hydrological and hydrogeological conditions were suggested as basis for studying stream ecology [Poole, 2010], as these factors have significant impact on habitat properties and fragmentation. To organize the investigation of geomorphic influences on stream ecology, Montgomery [1999] suggests the use of lithotopo units, based on the classification of the landscape according to lithology and topography, as a framework for in-stream patch dynamics but also as context for addressing watershed processes and their influence on stream integrity.

Fausch et al. [2002] criticized the commonly local view on habitat quality and called for the perception or stream and riparian ecosystems as parts of riverscapes. Applied to hydrologic controls of habitat properties, this means that hyporheic hydrology influences local habitat quality while the connectivity between catchment and stream controlled the habitat connectivity of the fluvial habitat as a whole [Tetzlaff et al., 2007].

All of the dependencies and processes described above are relevant for management decisions and can not easily be assessed and evaluated at single scales. Synthesis of observations and process understanding at different scales and from different catchments and catchment elements calls for a framework to combine, structure and synthesize observations.

1.1.4 How to get there?

A prerequisite for the synthesis of knowledge about processes in all elements of a catchment requires a schematic to link observations across scales and between locations. A first step in that direction is to relate observations and identify possible patterns and dependencies across scales and space. Based on these patterns and dependencies a unifying concept may be developed.

To do so, four studies focusing on different landscape elements of catchments have been carried out. These elements are: the soil, hillslopes, riparian zone, and streambed (Chapter 2 to 5). The focus of each study is on connectivity in the scope of its observations. In the synthesis of this thesis (Chapter 6), the results of the four studies will be discussed with regard to inter-scale dependencies and a unifying framework will be developed, which allows for a representations and relation of the observations of all four studies.

In the following, several concepts will be introduced, which will be taken up on in the four publications composing this thesis, and which will provide the basis for the synthesis at the end. These concepts include: preferential flow as a theme emerging in all of the four field studies, the problem of scaling in catchment hydrology, and the concept of spatial organization of structures and processes. The approaches of the studies will be summarized at the end of this introduction (Sec. 1.3).

1.2 Conceptual background

1.2.1 Preferential flow

In the four experimental studies included in this thesis, preferential flow – as a special form of subsurface flow – and its importance for hydrological connectivity from hillslopes to streams is a central theme. Preferential flow is a frequently addressed topic in the context of connectivity: geostatistical methods were used to model or characterize preferential flow through heterogeneous or fractured media to represent flow through aquifers [e.g., Delorme et al., 2008] or soil profiles in vertical or lateral direction [e.g., Sidle et al., 2001]. At the hillslope scale, lateral flow through macropores or at the bedrock interface were shown to create hillslope connectivity and influence hillslope runoff generation [e.g., Uchida et al., 2005, Tromp-Van Meerveld and McDonnell, 2006] Yet, Beven and Germann [2013] criticize that despite the acknowledged importance of preferential flow, hydrologists stick to the established concepts of Darcy [Darcy, 1856] and Richards [Richards, 1931] rather than working on process understanding and proper model representations of preferential flow.

While the studies cited above concentrate on flow through macropores that connect to larger, rather discrete flow path networks, the term can be defined somewhat broader: Generally speaking, preferential flow is the manifestation of increased flow rates though cavities or regions of higher hydraulic conductivity in comparison to the surrounding area, or the concentration of flow lines in certain areas, such as funneling due to surface or subsurface topography. This definition covers the discrete flow paths, but also allows us to transfer the concept of preferential flow to the catchment scale.

From a practical perspective, the characterization of flow processes as preferential flow based on this definition strongly depends on the observation scale: Flow through fissures in otherwise impermeable bedrock is obviously perceived as preferential flow when observed at the small scale. The same flow processes, however, can be summarized as hydraulic conductivity, when observed at a sufficiently large scale. In contrast, slightly different flow velocities due to differences in hydraulic conductivity of the sediment at the scale of 10 m will not be identified as preferential flow when observed at the scale of meters. At an observation scale of 100 m, the resulting flow patterns can be categorized as preferential flow, while the flow processes at the scale of single pores may become irrelevant.

Strictly homogeneous flow fields are practically non-existent in natural system. Thus, "all processes are preferential" [Uhlenbrook, 2006], and the characterization and evaluation of processes as preferential flow is just a question of scale: the observation scale is decisive for the identification of preferential flow as such, and the scale of effect for its relevance for the investigated system.

1.2.2 Scale issues

The spatial and temporal observation scale plays an important role in how we perceive things and how detailed we are able to describe them. Sometimes processes can not be observed at the scale of the phenomenon due to practical or methodological reasons, which requires a transfer of information across scales [Hooper, 2001]. Sometimes a hierarchical perspective is useful to organize examinations and to identify dependencies [Montgomery, 1999]. This is why scaling is a recurring theme in many environmental disciplines. The "scaling issue" [Blöschl, 1999] can be broken down into two main challenges:

The identification of relevant scales

The scale of interest, scale of effect, or model scale [Blöschl, 2001] is preset by the research question. However, to understand and explain behavior at this scale, knowledge of processes at other scales might be necessary (relevant scales). In the example of preferential flow at the hillslope, the knowledge of the morphology of macropores in the soil and the processes that lead to connecting them was necessary to understand internal processes as well as the resulting runoff generation behavior at the large scale. The same holds true for knowledge of larger-scale processes, when regional groundwater flow systems impact local processes.

The identification of relevant scales is a challenge for environmental modelers, who strive to reduce complexity to avoid over-parameterization [Blöschl, 2001]. It is also a challenge for experimental hydrologists, who are constrained by practical effort but also by the observation itself: observations are limited by the hypotheses with regard to choice of observation scales, and by the applied methods with regard to resolution, integration volumes and time spans.

The transfer of information across scales

Once the inter-scale dependencies are identified, the next challenge is the transfer of information across scales [Vogel and Roth, 2003]. This problem is – again – immanent in the context of model development, but also applies to practical observations, be it with regard to experimental design, data analysis, or interpretation.

Roth et al. [1999] and [Vogel and Roth, 2003] suggested the scaleway concept to handle multi-scale heterogeneity and applied this concept to model flow processes within the soil column. The concept is built on the idea that knowledge about material properties, structure, and an adequate process model at a smaller scale are sufficient to predict the effective material properties of the next larger scale.

A similar idea was pursued by Wu [1999], who used a "scaling ladder" to represent the hierarchical architecture of a landscape and transfer information across scales. The scaling ladder is based on a concept borrowed from hierarchy theory, which states that complex systems often consist of multiple, interrelated subsystems, which themselves are composed of another layer of subsystems [Simon, 1962]. The subsystems contained in one layer are called "holons" (Greek: holos = whole, suffix on = part, [Koestler et al., 1967]), a term that, as Wu [1999] elaborates, "conveys the idea that subsystems at each level within a hierarchy [...] act as 'wholes' when facing downwards and as 'parts' when facing upwards". This hierarchical structure of a complex system can be used to decompose it, vertically into levels and each level horizontally into holons, to receive less complex entities without significant loss of information [Wu, 1999]. The decomposability or near-decomposability of a system is the prerequisite for its observability and describability.

With regard to field observations, decomposability means that processes can be observed at a certain scale, with only basic knowledge of the scale context, and that the results can be merged to assemble a picture of the whole system while properly representing the internal processes. Sidle [2006] argues that field observations and practical process understanding are essential components in scaling and criticizes the increasing exclusive reliance on models and "surrogate" information extracted from digital elevation models and generated by remote sensing, especially in the context of changing climate and land use. However, field research on hydrological processes at multiple scales can be labor and cost-intensive, which brings us back to the first point in this chapter: the (*a priori*) identification of the relevant scales.

1.2.3 Structural and functional organization

The structural organization is the spatial arrangement of structural entities and features that follows a certain logic. This logic partly stems from scale-based, intrinsic conditions: Regardless of the soil type, a pedon can structurally always be decomposed into horizons, which themselves can always be characterized by their porosities, and so forth, resulting in a hierarchy of nested material properties that can be identified for every landscape element. The logic is furthermore based on the morphogenesis of the catchment, caused by past hydro-climatic and management regimes [Savenije, 2010, Zehe et al., 2014a]. Geologic and pedogenetic processes, erosion, sedimentation etc. are causing characteristic topologies, catenas or soil horizons. These structures are relatively stable at small temporal scales, and similar between landscapes that formed under similar conditions.

The spatially distributed structural features or entities within a landscape interact with each other in many different ways. Zehe et al. [2014b] coined the term "elementary functional units" for the smallest entities of consistent internal behavior. These units can connect vertically or laterally to other units and thus assemble larger topological units, depending on external drivers such as solar radiation and precipitation. The changing connectedness of these elementary functional units is described by the term functional organization.

Properly identifying and representing the structural and functional organization of a catchment can be critically important for hydrologic prediction [Western et al., 1999], especially under changing conditions. Applying the concept of scale hierarchy elaborated above (Sec. 1.2.2), the identification of scale levels and holons is a way of analysing the spatial organization of a catchment. The changing interactions between holons and scale levels then represent the dynamic functional organization of the catchment.

1.3 Research questions and approach

1.3.1 General research goal

This thesis consists of four parts, all of which are aiming at identifying and characterizing different structural features for the hydrological connectivity between hillslopes, riparian zones and streams. These studies focused on spatial structures that are deemed characteristic for systems that formed under similar conditions. In a

first assessment, characteristic structures were identified, based on previous investigations in those areas and literature reviews. The main goal was then to investigate the flow-relevance of these structures at the local scale, as well as their impact on the establishment of connectivity between hillslopes and streams.

The studies focused on different sections of the flow paths and different scales and aimed at identifying the (dis-)connecting structural elements of the investigated system. The observation scales were preset by the hypotheses and applied methods, but the development of comprehensive approaches allowed us to combine point observations with spatial patterns and integrated signals. To accomplish this goal, novel methodological techniques were developed or adapted accordingly.

The process-hydrological findings of the different parts of this thesis will be used to discuss and develop a framework to relate observations at different scales. This framework should allow for the required complexity to represent the systems spatial organization, while providing a clearly structured scheme that allows for an evaluation even if the picture is not complete.

1.3.2 Approach

The experimental approach of the four studies conceptually followed the path of water through the hillslope, towards the stream and through the streambed. Different stages of this path were investigated separately, focusing on structures that were prevalent at each stage. The first two parts investigated the establishment of vertical and lateral preferential flow and connectivity in slope deposits at the plot scale and the hillslope scale (Chapter 2 and 3), the third part dealt with the role of the riparian zone and its morphology as link or interface between hillslopes and stream (Chapter 4), and the fourth part examined the impact of streambed heterogeneity on the interaction between groundwater, surface water and the hyporheic zone (Chapter 5).

Part I and II of the thesis feature a hillslope-scale irrigation experiment at a hillslope of the Holtz headwater catchment. The experiment was designed to observe the establishment of subsurface preferential flow and to identify the scales and structures that were relevant for the observed processes. The observations under irrigation-driven conditions were compared to observations made during a rainfall event to evaluate their importance under natural conditions and at the larger scale (Chapter 2). Furthermore, the hillslope-scale observations were compared to ad-

ditional plot-scale irrigation experiments and the results of a structural survey of the subsurface (Chapter 3). Besides the investigation of hydrological processes, the application of GPR (ground penetrating radar) for hydrolological monitoring was tested and evaluated. The experiment furthermore introduced the concept of *form and function* to categorize and evaluate different methods and results in hydrological field research and analyzed the informational value of observations in the two categories. The methodological approach of these studies combined point observations of soil moisture, subsurface 2D monitoring via GPR, soil water and surface water stable isotope dynamics and integrated response observations.

Part III investigates the spatial patterns of gains along the 650 m long the Holtz headwater (Chapter 4). The gains were measured sequentially along the entire stream with a nested spatial resolution of a few meters up to 50 m, and under different hydrometeorological conditions. The variability of gain patterns was compared with geomorphological features of the stream channel, as well as the topography and known subsurface structures of the headwater catchment. The methodological challenge of this study was the application of Radon (Rn^{222}) as a naturally occurring tracer of groundwater inflow in a small, turbulently flowing and morphologically diverse stream.

Part IV of the thesis focuses on the hyporheic flow in the sediment of the lowland river Tern. Flow directions and velocities were measured at the scale of few cm at the depth of approximately 20 cm within the sediment. The measurements were conducted at three different sites along a well investigated river section, which were known to be characterized by substrate of different hydraulic conductivity and homogeneity. In this study, a novel heat pulse technique was applied to measure hyporheic flow vectors in the sediment using heat as a tracer.

The four parts of the thesis cover the entire chain of subsurface runoff generation from infiltration, subsurface lateral flow at the hillslope, groundwater inflow into the stream and flow processes in the streambed sediment. They all aimed to observe flow processes directly, using innovative techniques with a special focus on tracers. The observed flow processes were compared to spatial structures at various scales and the role of the these structures was evaluated. Thus, the link between dominant structural scale and process scale was to be revealed, answering the question of how large-scale structures influence small-scale flow processes and *vice versa*. These inter-scale dependencies are the basis for the concluding discussion and synthesis,

which aimed at a general conceptual representation of flow processes involved in creating hillslope stream connectivity across scales.

1.3.3 Study sites

Holtz catchment (Luxembourg)

Field work of parts I, II and III was conducted in the Holtz headwater catchment in the Grand Duchy of Luxembourg. The catchment has an area of 62 ha and is drained by an approximately 650 m long tributary of the Colpach in the Attert catchment. The field site is located at the southern verge of the Ardennes Massiv with its geology being dominated by strongly inclined Devonian schist bedrock. Agriculturally used plateaus, steep forested slopes and marginal floodplains only along the larger tributaries form the characteristic landscape of the region. The soils are shallow and characterized by periglacial slope deposits which can be found at most hillslopes in the Holtz catchment and the general area. Between 2011 and 2014 the average annual precipitation was 965 mm and annual average air temperature was 8.8 °C.

The Holtz headwater stream itself is narrow (0.3 m to 2.0 m), partially deeply incised, but also features several braided sections. The streambed substrate is highly heterogeneous, ranging from clay to boulders and solid bedrock. The stream channel shows an almost straight or mildly meandering course and its longitudinal profile is characterized by numerous steps and step-pool formations and has an overall gradient of 0 8.

Tern river (UK)

The field site of part IV on streambed sediment flow paths is located at the lowland river Tern, as this site has a very good data base on internal streambed structures. The river Tern is an 852 km^2 tributary of the river Severn in the UK. The research focused on an approximately 250 m long, meandering section of the river with 5 m to 8 m channel width. The floodplain is dominated by arable land and pasture with low relief energy. The geology of the region is dominated by Permotriassic Sherwood Sandstone and is part of one of the UK's major groundwater aquifers.

The streambed topography of the investigated river reach is characterized by multiple pool-riffle-pool sequences and vegetated side bars. The sediment consists

of postglacial drift deposits with grain sizes varying from midsized gravels over different sizes of sands to fine silty material. A characteristic feature of the streambed architecture are clay and peat lenses of variable thickness layered in between the sediments [Krause et al., 2012a]. These lenses show a significantly lower hydraulic conductivity in contrast to the surrounding substrate and are unevenly distributed across the research area [Krause et al., 2012a]. Such peat or clay layers are frequently found in the streambed of lowland rivers and the study site is deemed a representative example for characteristic lowland rivers in similar geological setting.

Chapter 2

Hillslope connectivity

Form and function in hillslope hydrology: characterization of subsurface flow based on response observations

Lisa Angermann, Conrad Jackisch, Niklas Allroggen, Matthias Sprenger, Erwin Zehe, Jens Tronicke, Markus Weiler, and Theresa Blume

Originally published in:
Hydrology and Earth System Sciences, 21(7):3727 – 3748, 2017.

Abstract

The phrase *form and function* was established in architecture and biology and refers to the idea, that form and functionality are closely correlated, influence each other and co-evolve. We suggest transferring this idea to hydrological systems to separate and analyze their two main characteristics: their form, which is equivalent to the spatial structure and static properties, and their function, equivalent to internal responses and hydrological behavior. While this approach is not particularly new to hydrological field research, we want to employ this concept to explicitly pursue the question of what information is most advantageous to understand a hydrological system. We applied this concept to subsurface flow within a hillslope, with a methodological focus on function: we conducted observations during a natural storm event and followed this with a hillslope-scale irrigation experiment. The results are used to infer hydrological processes of the monitored system. Based on these findings, the explanatory power and conclusiveness of the data are discussed. The measurements included basic hydrological monitoring methods, like piezometers, soil moisture, and discharge measurements. These were accompanied by isotope sampling and a novel application of 2D time-lapse GPR (ground-penetrating radar). The main finding regarding the processes in the hillslope was that preferential flow paths were established quickly, despite unsaturated conditions. These flow paths also caused a detectable signal in the catchment response following a natural rainfall event, showing that these processes are relevant also at the catchment scale. Thus, we conclude that response observations (dynamics and patterns, i.e., indicators of function) were well suited to describing processes at the observational scale. Especially the use of 2D time-lapse GPR measurements, providing detailed subsurface response patterns, as well as the combination of stream-centered and hillslope-centered approaches, allowed us to link processes and put them in a larger context. Transfer to other scales beyond observational scale and generalizations, however, rely on the knowledge of structures (form) and remain speculative. The complementary approach with a methodological focus on form (i.e., structure exploration) is presented and discussed in the companion paper by Jackisch et al. [2017].

2.1 Introduction

Characterizing subsurface flow is the aim of many hydrological field and modeling studies. In hillslopes with steep slopes and structured soils, subsurface flow is controlled by high gradients and high heterogeneity of hydraulic properties of the soil, resulting in a highly heterogeneous flow field and preferential flow paths [e.g., Scaini et al., 2017]. The specific challenge of investigating preferential flow lies in its manifestation across scales, its high spatial variability, and pronounced temporal dynamics. A considerable number of experimental and model approaches have been proposed to investigate the issue [Beven and Germann, 1982, Šimůnek et al., 2003, Gerke, 2006, Weiler and McDonnell, 2007, Köhne et al., 2009, Beven and Germann, 2013, Germann, 2014]. However, rapid flow in structured soils is still a challenge to current means of observation, process understanding, and modeling.

In previous studies at the hillslope scale, the focus was often on lateral flow processes and the establishment of overall connectivity. Hillslope-scale excavations yield information on spatial extent and characteristics of preferential flow paths in 3D [Anderson et al., 2009, Graham et al., 2010], but are highly destructive and lack the temporal component. Hillslope-scale tracer experiments in contrast resolve temporal dynamics and velocities [Wienhofer et al., 2009, McGuire and McDonnell, 2010] but lack the spatial information. Hillslope-scale experiments are usually very labor intensive and require high technical effort, and most studies are concentrated on well-monitored trenches [McGlynn et al., 2002, Tromp-Van Meerveld and McDonnell, 2006, Vogel et al., 2010, Zhao et al., 2013, Bachmair and Weiler, 2012]. Blume and van Meerveld [2015] give a thorough review of investigation techniques for subsurface connectivity and find experimental studies on this topic underrepresented in hydrological field research.

In recent years, a trend towards non-invasive methods for hillslope-scale observations has emerged [Gerke et al., 2010], which has been an important improvement with regard to repeatability and spatial and temporal flexibility of observations [Beven and Germann, 2013]. In this context various geophysical methods have been applied for subsurface exploration [e.g., Wenninger et al., 2008, Garré et al., 2013, Hübner et al., 2015]. From all applied geophysical techniques ground-penetrating radar (GPR) is known as the tool providing the highest spatial and temporal resolution. GPR provides information on subsurface structures at minimal invasive cost [e.g., Lambot et al., 2008, Bradford et al., 2009, Jol, 2009, Schmelzbach et al., 2011, 2012, Steelman et al., 2012]. Its short measurement times and high sensitiv-

ity towards soil moisture predestine GPR for monitoring subsurface flow processes. Nevertheless, only few field studies exist which have successfully applied surface-based GPR for the investigation of preferential flow paths or subsurface flow in general [Truss et al., 2007, Haarder et al., 2011, Guo et al., 2014, Allroggen et al., 2015b]. Previous GPR monitoring studies rely on two different principles. The first approach relies on interpreting selected reflection surfaces and comparing this interpretation result between the individual recorded GPR surveys [Truss et al., 2007, Haarder et al., 2011]. The result is a shift in GPR signal travel time, which can be interpreted in terms of soil moisture changes, using a petrophysical relation [e.g., the CRIM model, Allroggen et al., 2015b]. The second approach relies on calculating difference images between individual GPR surveys [e.g., Birken and Versteeg, 2000, Trinks et al., 2001, Guo et al., 2014, Allroggen and Tronicke, 2016] and thereby highlighting areas of increased changes in the subsurface. Due to the usually high noise level of field data, such difference calculations are critical and require sophisticated processing techniques [Guo et al., 2014, Allroggen and Tronicke, 2016].

Especially in structured soils, where subsurface flow is likely dominated by preferential flow paths, methods are required which are capable of covering the existing heterogeneity. Point measurements and integrated observations alone are barely able to meet this requirement. Structural changes of the subsurface as revealed by difference images obtained from GPR measurements, in contrast, reveal spatially discrete flow paths. We therefore applied and tested time-lapse GPR measurements to investigate subsurface flow processes within a hillslope with shallow and highly structured soils.

We chose a combination of conventional hydrological methods and non-invasive GPR measurements to explore flow processes by means of observations at the hillslope [hillslope-centered approach according to Blume and van Meerveld, 2015]. This hillslope-centered approach was supported by stream-centered process observations, including a basic hydrograph analysis and surface water stable isotope sampling during the natural rainfall event. Besides the 2D time-lapse GPR measurements, the hydrological methods at the hillslope include surface runoff collectors, a dense network of soil moisture observation profiles, stable isotope samples, and piezometers.

All methods and experimental results were subsumed under the framework of form and function as shown in Fig. 2.1. This framework was developed to analyze the explanatory power of the different observations. The idea of the form and

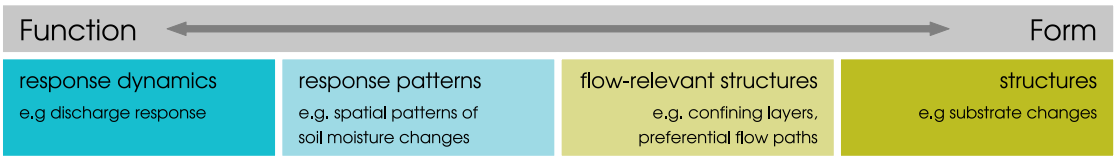

Figure 2.1: The concept of form and function applied to observations in hillslope hydrology. Four different categories which can be applied to data as well as the data sources.

function dualism was established in architecture [*form follows function* Sullivan, 1896], is commonly used in biology [e.g., Thompson, 1942], and describes the link, mutual influence, and co-evolution of the outer appearance and functional purpose of a (research) object.

In our case, form includes all static properties and spatial structures, such as topography, geology, and subsurface structures, but also porosity, hydraulic conductivity, and stone content of the soil. Function summarizes all dynamics and processes, including soil moisture dynamics, discharge behavior, and preferential flow. These two are closely related and co-evolve. Based on this idea, Sivapalan [2005] suggested that patterns, responses, and functions are the basic key to understanding and describing a hydrological system, as they incorporate the morphogenetic processes that led to the spatial structures. While this approach refers to the larger scale and the development of a general theory, our aim is to apply the form–function framework to observations at the local scale.

Starting on the left side of the spectrum presented in Fig. 2.1, we focus on the observation of response dynamics and response patterns. The potential of the methods for the investigation of subsurface flow processes at the hillslope scale and the characterization of typical runoff generation mechanisms are discussed and possible further improvements suggested. Based on these findings, the informative power and conclusiveness of the data will be discussed. To complement the functional perspective on the investigation of subsurface flow, the companion paper by Jackisch et al. [2017] concentrates on the spatial characteristics of subsurface flow from the point to hillslope scale with a specific focus on subsurface structures.

Following the form and function framework, the hypotheses focus on the potential of response observations for hillslope hydrological field research and the application of time-lapse GPR measurements in this context.

H1 Response observations (discharge, TDR and GPR data) are sufficient to characterize subsurface flow within the hillslope. (function described without form)

H2 Response patterns can be used to deduce flow-relevant structures in the subsurface. (function reveals form)

H3 Time-lapse GPR measurements visualize subsurface flow dynamics and patterns and can replace hillslope trenches.

2.2 Methods

2.2.1 Study site

The investigated area is located at the south-eastern edge of the Ardennes Massif in western Luxembourg. It consists of a number of nested sub-catchments of the Colpach River catchment, which is part of the Attert River basin. The landscape of this area is characterized by Devonian schist bedrock [Colbach and Maquil, 2003]. The soils are young and composed of eolian loess deposits and weathered schist debris. Under periglacial conditions, the weathered rocks were relocated by solifluction, causing an often horizontal or slope parallel orientation of the saprolite [Juilleret et al., 2011]. The periglacial deposit layer (basal layer) is overlain by shallow top soil (upper layer). The soil is classified as Haplic Cambisol [CM, IUSS Working Group WRB, 2006]. Saturated hydraulic conductivity of the soil was found to be highly heterogeneous, exceeding the measuring range of the capacity of the constant head permeameter of $10^{-8}\,\mathrm{m\,s^{-1}}$ to $10^{-3}\,\mathrm{m\,s^{-1}}$. While depth profiles of hydraulic conductivity measured in the area did not show a specific pattern of conductive layers, measurements at the investigated hillslope indicated higher conductivity at a depth of 0.7 m [Jackisch et al., 2017].

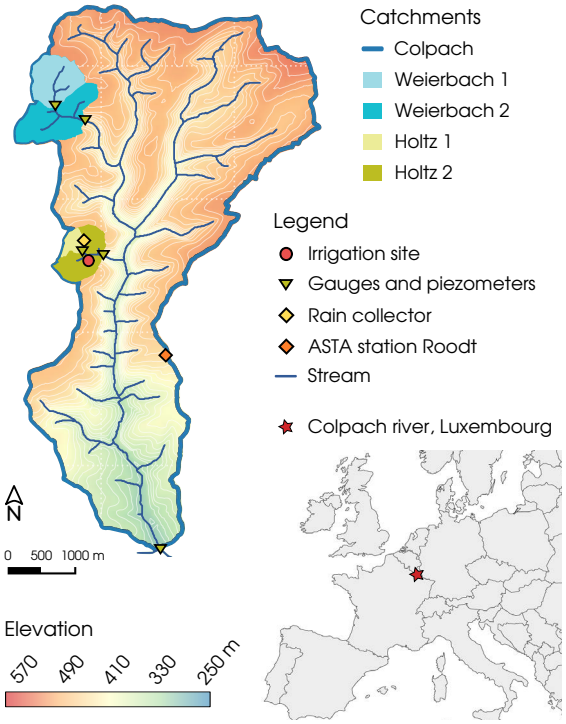

Figure 2.2: Map of the investigated Colpach River catchment and the four gauged sub-catchments. The site of the hillslope-scale irrigation experiment is located in the Holtz 2 catchment and is indicated in red.

The schist bedrock below is strongly inclined, with almost vertical foliation, and is considered impermeable but with fractures which can function as a complex flow network with local storage in the rock cracks when saturated [Van den Bos et al., 2006, Kavetski et al., 2011]. The subsurface structures are of predominantly geogenic origin and are considered temporally persistent.

Within this landscape, a typical hillslope consists of agriculturally used elevated plateaus and forested valleys with steep slopes (15° to 25°). While the headwater catchments of the investigated area are usually narrow with marginal floodplains, the main Colpach River network is characterized by wider valleys with more pronounced floodplains.

The average annual precipitation between 2011 and 2014 was 965 mm; the annual average air temperature was 8.8 °C. These data stem from a meteorological station from ASTA (administration des services techniques de l'agriculture de Luxembourg) close to Roodt, approximately 2 km from the experimental site.

The experimental work conducted in the framework of this study focused on a north-facing hillslope in the Holtz headwater catchment. The experiment was supplemented by hydrological data from five neighboring headwater catchments of different sizes. All sub-catchments as well as the location of the irrigation site are shown in Fig. 2.2.

2.2.2 Experimental approaches

The experimental approach consists of two parts. The hillslope-centered approach concentrates on local observations at the hillslope. It includes soil moisture profile measurements and 2D GPR measurements, pore water and piezometer isotope data, and measurements of surface runoff. These data were collected during a natural summer storm event on 20 June 2013 and a hillslope-scale irrigation experiment one day later on 21 June 2013.

The stream-centered approach focuses on the discharge response and stream water stable isotope signal during the same summer storm event as mentioned above. The stream-centered approach focuses on the integrated response of a catchment. While hydrographs and stream tracer dynamics have been studied and discussed extensively elsewhere [Wrede et al., 2015, Martínez-Carreras et al., 2016, in the same area], we wanted to use these data to position our hillslope observation in

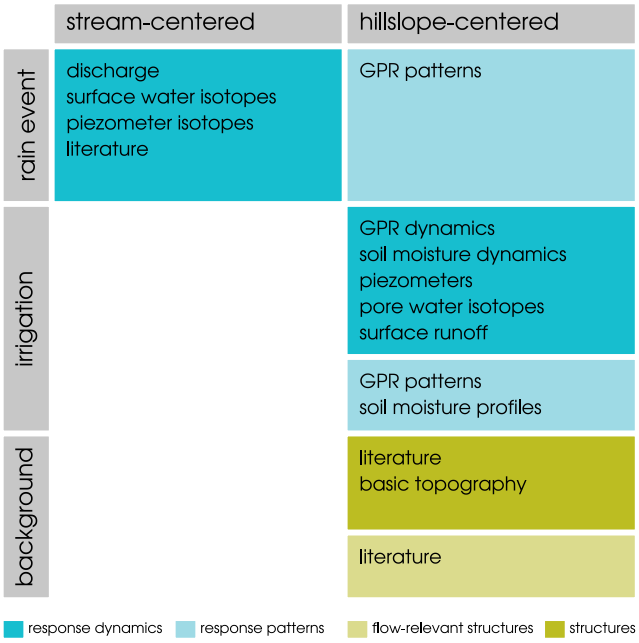

Figure 2.3: Experimental methods applied in the stream-centered and hillslope-centered approaches, divided into the sampling during the natural rain event and the irrigation experiment. Additionally, some structural background information was obtained from the literature and a digital elevation model.

the bigger picture of the catchment-scale dynamics. An overview of approaches, methods, and their foci is given in Fig. 2.3.

2.2.3 Stream-centered approach

The hydrological response behavior of several nested sub-catchments was investigated. At four locations v-notch or trapezoidal gauges were installed and equipped with pressure transducers, measuring water level, electric conductivity, and temperature (CTD sensors, Decagon Devices Inc.). Water levels were measured every 15 min. Precipitation was monitored with tipping buckets (Davis Instruments Corp.) in the Holtz 1 headwater. All data were logged with CR1000 data loggers (Campbell Scientific Inc.).

At the same locations and additionally close to the source of the Holtz River (Holtz 1 in Fig. 2.2), water samples were taken with auto samplers (ISCO 3700, Teledyne). The bottles of the auto samplers were pre-filled with styrofoam beads to avoid evaporation from the sample bottles. Samples were then transferred to glass bottles and analyzed in the laboratory at the Chair of Hydrology, University of Freiburg. The isotopic composition ($\delta^{18}O$ and $\delta^{2}H$) of the water samples was measured by wavelength-scanned cavity ring-down spectrometry (Picarro L2120-iWS-CRDS). The results are given in δ-notation in [‰], describing the deviation of the ratio between heavy and light isotopes ($^{2}H/^{1}H$ and $^{18}O/^{16}O$) relative to the

ratio of the Vienna Standard Mean Ocean Water (VSMOW). For liquid analysis the accuracy is given as 0.1 ‰ for $\delta\ ^{18}O$ and 0.5 ‰ for $\delta\ ^{2}H$ (according to the manufacturer).

In addition to the stream water, rainfall water was sampled. Bulk samples were collected during the rainfall events right next to the experimental site. The water from the saturated zone was manually sampled on a monthly basis over the course of one year from piezometers close to the sub-catchment gauges. Samples were taken with a peristaltic pump from fresh water flowing into the piezometers, after they had been pumped empty (Fig. 2.2).

To calculate the event water contribution, we applied a simple hydrograph separation [Pearce et al., 1986]. Equation 2.1 shows the calculation of the discharge attributed to the natural rain event Q_e, based on the isotopic composition of the base flow (c_b) three days before the storm event, the river water during and after the event (c_t), and the rain water (c_e) [Leibundgut et al., 2011]. Q_t is the total discharge during and after the event, which is constituted of the two components Q_e and Q_b, i.e., event water and pre-event water.

$$Q_e = Q_t \cdot \frac{c_t - c_b}{c_e - c_b} \tag{2.1}$$

2.2.4 Hillslope-centered approach

Irrigation setup

The plot of the hillslope-scale irrigation experiment was located on the bottom 8 % to 13 % of the 238 m long investigated hillslope, which was defined by a slope of more than 6°, excluding the plateaus. The plot had a slope of ∼14°. While vegetation at the hillsope is dominated by beech forest (*fagus sylvatica*) of mixed age, the irrigation plot is placed in an area with no major trees. Except for a few young trees with breast height diameters below 0.1 m in the downhill monitoring area, all shrubs were cut to facilitate GPR measurements and allow for uniform irrigation. The entire investigated hillslope section covers an area of approximately 260 m^2.

Four circular irrigation sprinklers (Wobbler, Senninger Irrigation Inc.) were arranged in a 5 m by 5 m square in the upper part of the experimental site (Fig. 2.4). The sprinklers had a nominated sprinkler radius of 4 m and were installed approximately 0.7 m (two uphill sprinklers) and 1.5 m (two downhill sprinklers) above ground surface. The level difference between the uphill and downhill sprinklers

was 0.5 m. The 25 m^2 area spanned by the four sprinklers is referred to as the core area, with a homogeneous irrigation intensity of ∼30.8 mm h^{-1} over 4:35 h. Total water input at the core area was 141 mm. These settings aimed at activating all potential flow paths and were chosen on the basis of an *a priori* simulation of the experiment [see the Appendix of Jackisch et al., 2017]. Transferring this amount of water from the irrigated core area (5 m hillslope parallel length) to the entire hillslope uphill of the rain shield (219 m), this intensity compares to a rain event of 3.2 mm precipitation. While these irrigation settings do not mimic natural conditions, this relation allows us to compare and evaluate observations under experimental and natural conditions regarding lateral subsurface flow.

The surrounding area functioned as a buffer of about 4 m with less intense irrigation, thus mitigating boundary effects. A rain shield defined the lower boundary of the core area as a sharp transition to the non-irrigated area below. The water from the rain shield was collected with a gutter and routed away from the investigated area. The overall irrigation area (including core area and buffer) covered ∼120 m^2.

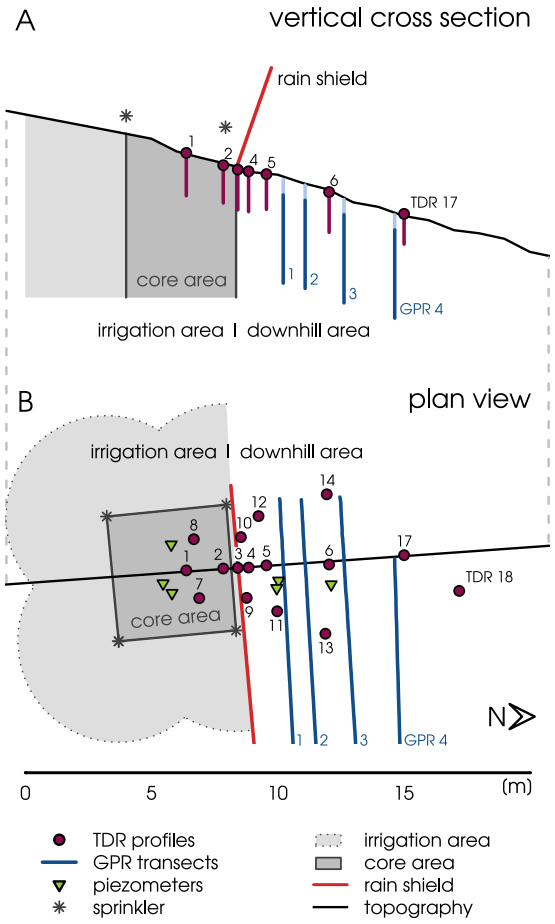

Figure 2.4: Vertical cross section (A) and plan view (B) of the experimental setup and sprinkler array. Location and depth of the TDR profiles and GPR transects are given in purple and blue, respectively. The black line along the central TDR transect in B marks the vertical cross section depicted in A.

To monitor the irrigation, we used a flow meter at the main water supply of the irrigation system to measure the absolute water input. Furthermore, one tipping bucket was used to quantify the temporal variability of applied irrigation, and 42 mini rain collectors, evenly distributed across the core area, covered the spatial distribution of the irrigation amount. The topography of the experimental site as well as all devices and installations were mapped with a total station (Leica Geosystems AG).

The experiment took place on 21 June 2013. After one week of dry weather, two natural rainfall events of 20.2 mm and 21.2 mm occurred on 20 June. The first one had a mean intensity of 2.9 mm h^{-1} and ended 29:33 h before the irrigation experiment; the second rainfall event had a mean intensity of 9.0 mm h^{-1} and ended 19:22 h before the experiment.

Process monitoring

The monitoring of hydrological processes during and after the irrigation period was accomplished with a combination of methods: a dense array of soil moisture profiles for time domain reflectrometric (TDR) measurements arranged as diverting transects along the slope line, and four GPR transects located downhill of the core area and oriented parallel to the contour lines and the rain shield for time-lapsed GPR measurements. The latter yielded vertical cross sections of the subsurface.

A surface runoff collector was installed across 2 m at the lower boundary of the core area. Surface runoff was collected by a plastic sheet installed approximately 1 cm below the interface between the litter layer and the Ah horizon of the soil profile and routed to a tipping bucket.

An array of 16 access tubes for manual soil moisture measurement with TDR probes (Pico IPH, IMKO GmbH) covered the depth down to 1.7 m below ground. The layout consisted of three diverging transects with four TDR profiles in the lower half of the core area, the highest density of profiles just downslope of the rain shield, and the furthest profile about 9 m downhill (Fig. 2.4B). This setup allows for the separate observation of predominantly vertical flow at the core area and lateral flow processes at the downhill monitoring area.

Soil moisture was measured manually. To increase the temporal resolution of the measurements, three probes were used in parallel. While these probes were identical with regard to measuring technique and manufacturing, they differed slightly in

sensor design: two TDR probes had an integration depth (i.e., sensor head length) of 0.12 m, and one probe had an integration depth of 0.18 m. These sensors were manually lowered to different depths into the 16 access tubes, where they measured the dielectric permittivity of the surrounding soil in the time domain through the access tubes. Given a mean penetration depth of 5.5 cm and a tube diameter of 4.2 cm, this yields an integration volume of ∼0.72 L and 1.05 L, respectively. The manual measurements were conducted in 0.1 m depth increments and followed a flexible measuring routine with regard to the sequence in which the access tubes were measured. Thus, active profiles were covered with higher frequency.

In addition to the hydrological methods, GPR was used to monitor the shallow subsurface. Two-dimensional time-lapse GPR measurements were conducted along four transects across the downhill monitoring area. The transects had distances of ∼2 m, 3 m, 5 m and 7 m to the lower boundary of the core area and were arranged approximately perpendicularly to the topographic gradient. Each transect was measured nine times. One measurement was taken before irrigation started and the last one about 24:00 h after irrigation start.

The GPR system consisted of a pulseEKKO PRO acquisition unit (Sensors and Software Inc.) equipped with shielded 250 MHz antennas. The data were recorded using a constant offset of 0.38 m, a sampling interval of 0.2 ns, and a time-window of 250 ns. For accurate positioning, a kinematic survey strategy was employed. The positioning was based on a self-tracking total station (Leica Geosystems AG), which recorded the antenna coordinates as described by Boeniger and Tronicke [2010]. To guarantee the repeatability of the 2D time-lapse GPR measurements, all four transects were defined by wooden guides for an exact repositioning of the antennas. The measurement of one transect took approximately 2 min and measurements of all four transects were taken every 40 min to 120 min during and after irrigation as well as 18:00 h and 24:00 h after irrigation start.

Isotope sampling

The stable isotope sampling included samples taken from five soil cores (pore water), piezometers (percolating pore water), as well as irrigation and rain water (input water). The soil cores were taken with a percussion drill with a head diameter of 7 cm and split into 5 cm increments to get depth profiles of the stable isotopic composition ($\delta^{18}O$ and $\delta^{2}H$) of the pore water. Two profiles were taken before the rainfall events, one after the first minor rain event on 20 June, and two more after

the irrigation experiment (at the core area and the downhill monitoring area). All profiles covered a depth of ∼1.7 m below ground.

At the locations of the pre-irrigation soil cores, piezometers were installed. Additionally, three more piezometers were installed at a depth of ∼1.0 m. This depth was chosen based on observations in the core samples, which showed wet areas at the depths between 0.8 m and 1.2 m, right above the C_v horizon. All piezometers consist of PVC tubes of 5 cm diameter and were screened at the bottom 20 cm. They were equipped with pressure transducers (CTD sensors, Decagon Devices Inc.). As only a few mL of water were seeping into the piezometers, water tables could not be properly monitored, and the data will not be shown. However, the water could be sampled using a peristaltic pump. In addition to the pore water and piezometer samples, bulk samples of rainfall water were collected during the rainfall events prior to the irrigation experiment and directly next to the irrigation plot. Water samples were also taken from the irrigation water reservoir five times during irrigation.

The soil samples were prepared following the direct equilibration method as proposed by Wassenaar et al. [2008] and described in detail by Sprenger et al. [2016]. The precision for the method is reported to be 0.31 ‰ for $\delta\ ^{18}O$ and 1.16 ‰ for $\delta\ ^{2}H$ [Sprenger et al., 2015a]. All water samples were analyzed following the same procedure as described in Sec. 2.2.3.

2.2.5 Data analysis

TDR data analysis

Almost 5000 individual soil moisture measurements were taken during the irrigation experiment. As the three TDR probes had different integration depths (0.12 m and 0.18 m), the measurements had a different depth offset relative to the ground surface when referenced to the center of the probe, and had to be aligned. To do so, the measurements, which were originally taken in 0.1 m increments, were resampled at depths by linear interpolation. Due to the potentially short correlation length of soil moisture [Zehe et al., 2010], inverse distance interpolation between two locations is generally not appropriate. In the case of the vertical profiles, however, the integration depths of the probes exceeded the measuring increments. Due to the resulting overlap of the integration volumes, this procedure was assumed to be adequate. The measurement of one depth increment took between approximately

10 s to 30 s. While the data were interpolated in time for better visualization, all data analyses were performed with the uninterpolated data.

All TDR measurements were referenced to the last measurement before irrigation. The resulting data set of relative soil moisture changes $\Delta\theta$ was used for the discussion of soil moisture dynamics and response velocity calculation (Sec. 2.2.5). The storage changes (mm) in the top 1.4 m of the core area were estimated based on the four core area TDR profiles TDR1, TDR2, TDR7, and TDR8, by multiplying the $\Delta\theta$ (%) of each depth increment by the respective depth interval (mm). Together with the time series of water input, these data were used to estimate the mass balance dynamics of the core area.

Data processing of 2D time-lapse GPR measurements

The time-lapse GPR survey comprised repeated recordings of vertical 2D GPR data along the four transects. The data processing of each measurement relied on a standard processing scheme, including bandpass filtering, zero time correction, exponential amplitude preserving scaling, inline fk-filtering, and a topographic migration approach, as presented by Allroggen et al. [2015a]. The GPR data were analyzed using an appropriate constant velocity and gridded to a 2D transect with a regular trace-spacing of 0.02 m.

There is no standard interpretation procedure for the analysis of time-lapse GPR data. Most approaches are based on calculating trace-to-trace differences [Birken and Versteeg, 2000, Trinks et al., 2001] or picking and comparing selected reflection events in the individual time-lapse transects [Allroggen et al., 2015b, Haarder et al., 2011, Truss et al., 2007]. In the context of this study, however, both approaches provided only limited interpretable information. Considering the methodological uncertainty, the highly heterogeneous soil did not provide reflectors which were a suitable reference. Therefore, we used a time-lapse structural similarity attribute presented by Allroggen and Tronicke [2016], which is based on the structural similarity index known from image processing [Wang et al., 2004]. This approach incorporates a correlation-based attribute for highlighting differences between individual GPR transects and has been shown to improve imaging, especially for noise data and limited survey repeatability.

The calculated structural similarity attributes are a qualitative indicator of relative deviations from the reference state. The GPR data indicated remaining water from the natural rain event when the experiment was started. Therefore, the last

acquisition time 24:00 h after irrigation start was chosen as the reference time for all GPR transects. Based on the assumption that the reference state is the one with the lowest water content, decreasing structural similarity was interpreted as an increase in soil moisture.

To convert GPR two-way travel time (TWT) into depth, we used the average measured GPR propagation velocity of $0.07\,\mathrm{m\,ns^{-1}}$. This velocity is based on additional common midpoint data and the assumption of static conditions during the experiment. Using this velocity, the GPR transects covered a TWT of 120 ns, which corresponds to a depth of $\sim$4.2 m below ground surface. Approximately the first 20 ns of each transect are influenced by the interfering arrival of the direct wave and the ground wave. Consequently, we observe no interpretable reflected energy in the uppermost time window. Thus, the 2D GPR measurements imaged the subsurface between $\sim$0.7 m to 4.2 m depth below ground.

Comparison of a natural event vs. irrigation based on 2D GPR data

To interpret the structural similarity attribute images, we discriminated between the signal of the natural rain event and the irrigation. The discrimination was based on the temporal dynamics of each pixel of the GPR transects (i.e., every single value in the matrix of distance along the GPR transect and depth/TWT). The first GPR measurements were taken 12:52 h after the end of the second rainfall event (i.e., 6:30 h before irrigation start) and the observed responses were attributed to the natural rain event. Once the structural similarity attribute value of a pixel decreased more than 0.15 after irrigation start, the signal of that pixel was attributed to the irrigation. The threshold of 0.15 was chosen based on the noise of the last measurement 18:00 h after irrigation start and exceeds the standard deviation of that measurement by a factor of 3. The same procedure was applied to infer the time of first response to the irrigation signal, which was used to calculate response velocities.

This procedure yields 2D maps of response patterns, with each pixel being attributed to either the irrigation or the natural rain event. The structural similarity values are a semi-quantitative measure of soil moisture and thus no reliable indicator to directly compare actual soil moisture responses recorded at different locations or at different times. We therefore used the areal share of the monitored cross sections to compare the impact of the two input events. To do so, all pixels of one of the two categories (natural rainfall or irrigation) which fell below the value of 0.85

(i.e., maximum similarity 1 minus threshold 0.15) were counted and expressed as a fraction of the entire cross section. The resulting areal share does not represent the actual share of activated flow paths, but is a semi-quantitative indicator of the hillslope cross section impacted by active flow paths.

Response velocity calculation

As no tracers were used for irrigation, dynamic processes had to be inferred from changes in state. For TDR measurements, the time of first response was defined as an increase in soil moisture by 2 vol% relative to initial conditions. This threshold was chosen based on the standard deviation of measurements under presumably constant conditions. The time of first response was identified for each TDR profile and depth increment.

Due to the experimental setup, soil moisture dynamics on the core area were dominated by vertical processes, while lateral processes controlled the dynamics at the downhill monitoring area. Accordingly, vertical and lateral response velocities were calculated from core area and downhill monitoring area TDR profiles, respectively.

As a continuous wetting of the soil profile could not be assumed, all response velocities were calculated for the entire depth (or distance) instead of depth increments. Response velocities are therefore integrated values describing processes in the entire soil column above. This procedure also accounts for heterogeneous processes and preferential flow paths, which may bypass shallower depths without leaving a detectable soil moisture signal.

Lateral response velocities account for the depth and distance between soil surface at the rain shield and TDR profile in question and, therefore, integrate lateral and vertical flow. They were calculated for every depth of the soil moisture profiles at the downhill monitoring area. The time of the very first response signal measured on the core area was used as reference time t_0, which was 15 min after irrigation start. Due to the slope-parallel or horizontal orientation of the saprolite, we assumed that the water flows either vertically or laterally rather than diagonally. Based on this assumption, the distances were calculated from the slope parallel distance of each profile from the lower boundary of the core area plus the depth of every measuring point. The distance assumptions for both, vertical and lateral velocity calculations, do not resolve tortuosity of flow paths and, therefore, drastically reduce the complexity of the flow path network to its integral behavior. The calculated response

velocities are thus not to be interpreted as *in situ* flow velocities in the flow paths, but rather as the minimum necessary velocity explaining the observed arrival of the wetting signal.

The same holds true for the lateral response velocities calculated from GPR data. In accordance with the separation of the natural rain event signal and the irrigation signal, the first decrease in structural similarity of more than 0.15 was interpreted as the arrival of the irrigation signal. Single structures and flow paths are not the focus of this article and will be discussed in the companion study by Jackisch et al. [2017]. Here, we therefore simplified the 2D patterns to a depth distribution of occurring response velocities. To do so, all areas that were newly activated at the time of one measurement were accumulated by depth and given as a portion of the entire width of each GPR transect. Comparable to the procedure applied to the TDR data, the response velocities were then calculated from the respective measuring time, the distance between transect and irrigation area, and the depth. The resulting patterns show the spatial fraction of the depth increment which is connected to flow paths of the calculated velocity or faster and give an idea of the spatial distribution of the GPR response velocities.

2.3 Results

2.3.1 Response to the natural rainfall

Stream-centered approach: hydrograph and surface water isotopes

In response to the summer storm event just before the hillslope-scale irrigation experiment, all gauged sub-catchments showed double-peak hydrographs, with one immediate short peak, and one prolonged peak delayed by several hours (second rainfall event, Fig. 2.5). In the headwater catchments (Holtz 2, Weierbach 1 and 2), the first peak occurred almost instantly, while the more distant Colpach gauge showed a delay of approximately 3:00 h. The second response was prolonged, with a maximum approximately 36:00 h after the event. The strength and ratio of the two peaks varied across different sub-catchments and according to hydrological conditions, but the general pattern is characteristic of the hydrological behavior of the Colpach River catchment. Similar behavior was also reported by Fenicia et al. [2014], Wrede et al. [2015], and Martínez-Carreras et al. [2016], whose investigations focused on the Weierbach 1 catchment.

A simple mass balance calculation revealed that the first peak constituted 7.5 % of the total event runoff at gauge Holtz 2. The total event runoff coefficient was 0.44. Referenced to the precipitation amount, about 3.3 % of the input left the headwater within 7:00 h after the rain event. In the neighboring Weierbach catchment and the Colpach, the first peak contributed more strongly to the total event runoff (14.2 % and 12.9 % at Weierbach 1 and 2, and 19.7 % at Colpach).

The $\delta^{18}O$ signature of the stream water is indicative of the origin of the water. It showed strong dynamics during the discharge response to the rain events on 20 June (Fig. 2.5). The results of the hydrograph separation show that the event water contributed up to 67.6 % to the event runoff during the response to the first rain event in the morning of 20 June (Colpach, 6:00h). After that, the total discharge dropped again, with the event water contribution decreasing to 31.6 %. With the onset of the first peak caused by the second rain event at 19:20h, the event water contribution increased again and reached values of over 50 % (58.0 % in the Colpach at 22:00h, 55.2 % in the Holtz 1 catchment, 21:51h on 20 June; Fig. 2.5). $\delta^{18}O$ values then declined, indicating event water contributions of around 20.0 % (24.8 % at 4:00h in the Colpach, 18.1 % at 13:03h in Holtz 2, and 16.2 % at 21:28h in Holtz 1 on 21

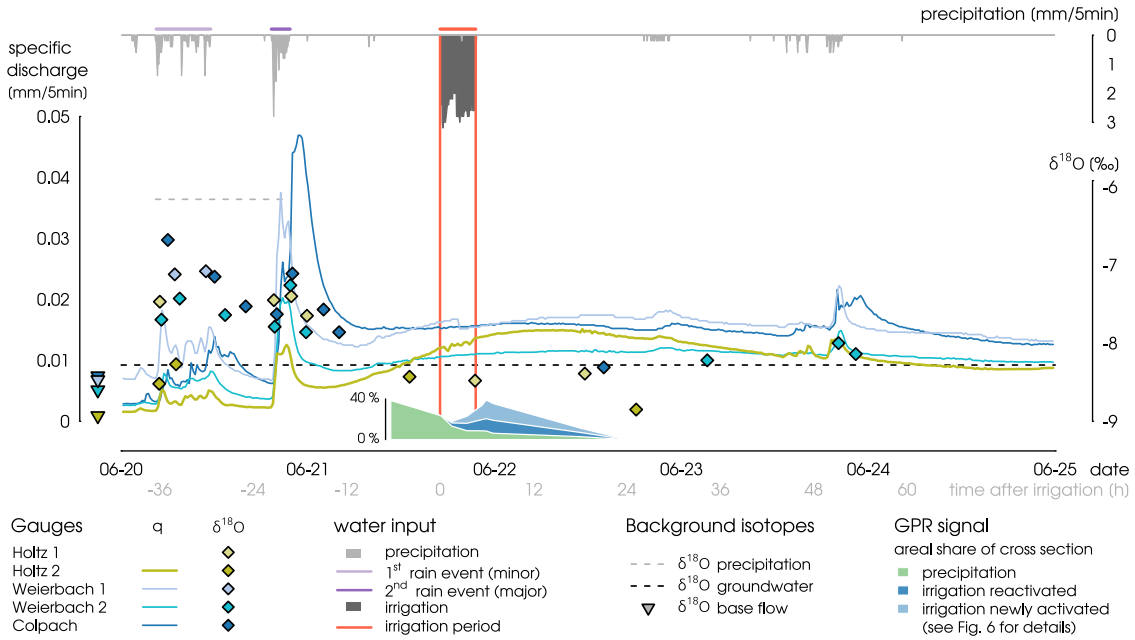

Figure 2.5: The figure shows the natural storm event on 20 June 2013 in the Colpach River catchment and the local intensity of the irrigation on 21 June 2013. The hydrographs below show the discharge response of four nested catchments (solid lines), in combination with the dynamics of the $\delta^{18}O$ isotopic composition of the surface water (dots). The isotopic composition of the groundwater (annual mean) and the precipitation (daily values) are given by the dashed lines. Furthermore, the dynamic response of the GPR signal to natural and artificial rainfall is given in green and blue. While the first minor rain event caused only a weak response, the second event caused a double-peak discharge response in all sub-catchments. The irrigation experiment took place 19:22 h after the rain event.

June). Weierbach 1 and 2 showed the same pattern, with event water contributions well above 50.0 % for the first peak of the second rainfall event.

Uncertainty in hydrograph separation was caused by the uncertainty of the stable isotopic composition of the precipitation input. The uncertainty due to spatial variability of the precipitation input was kept minimal for Holtz 2, by sampling the precipitation within the small catchment (45.9 ha). While we could not sample the isotopic input at high temporal frequency, the bulk sample of the precipitation data represents a weighted average of the input isotopic signal.

Hillslope-centered approach: subsurface response patterns

The 2D time-lapse GPR measurements yield images of structural similarity referenced to the last measurement, which were taken 24:00 h after irrigation start, which translates to 43:22 h after the second natural rain event. The first GPR measurements were taken about 12:52 h after the second rain event and can be

Table 2.1: Overview of sub-catchment size and accumulated specific discharge as a percentage of the precipitation amount (%)

		accumulated discharge *	
Gauge	size (ha)	after 7:00 h	after 72:00 h
Holtz 1	9.2	-	-
Holtz 2	45.9	3.3	43.8
Weierbach 1	45.1	8.8	61.8
Weierbach 2	106.3	5.3	41.4
Colpach	1903.3	12.6	64.1

* expressed as percentage of the precipitation amount (%)

interpreted as the subsurface response patterns of this event (Fig. 2.6A). The subsequent GPR measurements furthermore show the temporal dynamics of the rainfall signal, overlain by the irrigation signal. The high initial signals, as well as the high but decreasing areal share of active regions in all transects during the first measurements until 1:30 h after irrigation start (Fig. 2.6B), indicated free water remaining from the preceding natural rain event which slowly disappeared.

While transect 1 showed only a weak signal of the natural event in the first measurement, transects 2 and 3 exhibited stronger and longer lasting signals. The areal share of active regions of the four transects in the measurement preceding the irrigation experiment was 38.5 %, 51.6 %, 64.4 % and 50.5 % from upslope to downslope. Except for transect 2, which even showed a slight increase in the areal share of active regions between the first and second measurements, the signal of the natural rain event was continuously vanishing (Fig. 2.6B).

2.3.2 Hillslope-scale irrigation experiment

Core area water balance dynamics

The irrigation intensity was relatively constant over time, with only weak fluctuations due to gradual clogging of the intake filter. The spatial distribution of the irrigation intensity on the core area was influenced by the sprinkler setup and the slope of the experimental site. The mean intensity on the core area was $30.8 \pm 7.3\,\mathrm{mm\,h^{-1}}$, with slightly higher values in the vicinity of the four sprinklers. Surface runoff at the lower boundary of the core area started 20 min after irrigation start and ceased with the same time lag. In total, surface runoff amounted to 0.5 L, which equals only 0.02 % of the water balance.

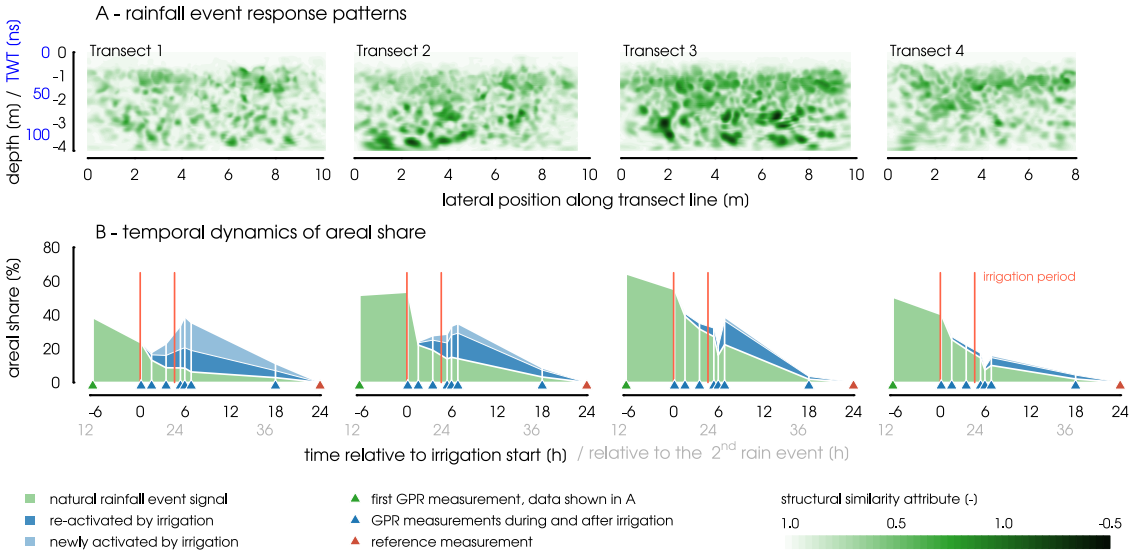

Figure 2.6: A: Two-dimensional GPR data showing the subsurface response patterns caused by the natural rainfall event. The data show the structural similarity between the first GPR measurement (approximately 6:30 h before irrigation start and 12:52 h after the second rainfall event) and the last one. Low values of structural similarity are interpreted as high changes in soil moisture. B: Temporal dynamics of the areal share of active regions attributed to the natural rain event and the irrigation. Activated regions were identified by a structural similarity attribute of less than 0.85. Data were interpolated linearly between the measurements for visualization. The measurements shown in A show the data used to calculate the first data point shown in B.

The core area mass balance is shown in Fig. 2.7, depicting the storage increase in the top 1.4 m of the soil. All profiles showed a mass recovery of more than 100 % (i.e., higher storage increase than water input at measuring time; see Fig. 2.7) in the first 60 min of the irrigation period. In profiles TDR1, TDR2, and TDR8 mass recovery then decreased and dropped below 100 %, while TDR7 increased further, with a maximum overshoot of almost 50 % approximately 2:00 h after irrigation start. The last measurement during irrigation was taken approximately 50 min before the end of the irrigation period. At this time, the average storage increase was more than 20 % lower than the input mass.

The first measurement after irrigation (6 min to 19 min after irrigation stopped) showed a mean deficit of 54.7 %, indicating that on average 31.2 % (between 18.6 % and 43.9 %) of the water that has been recorded at the last measurements before irrigation stop was freely percolating and had left the monitored depth immediately. After this fast instantaneous reaction, the water content decreased equably. Mean total mass recovery dropped to 8.9 % after 18:24 h after irrigation start and almost returned to initial conditions (1.6 %) after 24:00 h after irrigation start.

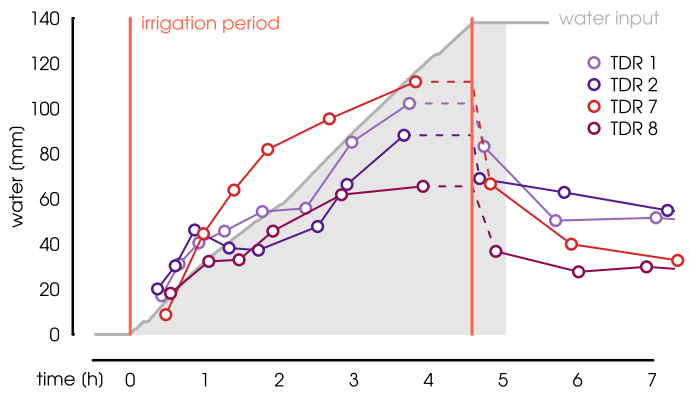

Figure 2.7: Water balance of the top 1.4 m of the soil column for the four core area TDR profiles. Dashed lines indicate the storage increase at the last measurement before irrigation ended. The variability between the four profiles shows the high heterogeneity and causes uncertainty regarding the average mass balance of the core area.

Soil moisture profiles and dynamics

The high variability in soil moisture dynamics observed in the TDR profiles is summarized in Fig. 2.8. The four uppermost panels (rows 1 and 2) show the core area profiles. Columns represent the three diverging TDR transects. The general pattern observed at the core area was characterized by a strong and comparably fast response in the top 0.4 m of the soil and below the depths of approximately 1.2 m. The response in between these active layers was more diverse and generally weaker.

Soil moisture in the top 0.4 m of the soil of TDR1, TDR7, and TDR8 quickly stabilized around constant values, indicating the establishment of quasi steady-state conditions. After the end of the irrigation, the soil moisture quickly declined down to a $\Delta\theta$ of 4 vol%, indicating a very fast response to the dynamics of the water input. In contrast to the fast establishment of steady-state conditions and the fast decline, a slightly increased water content of up to 4 vol% above initial conditions was persistent in distinct depth increments and was also measured even 24:00 h after irrigation stopped.

The soil moisture patterns at the downhill monitoring area were more diverse. Profiles located directly below the rain shield (TDR9, TDR3, and TDR10 with a distance to core area of 0.2 m to 0.5 m, Fig. 2.8) exhibited dynamics that resemble the reaction at the core area, but with mostly lower intensities and higher variability in depth. More distant TDR profiles however showed a highly variable picture. Distinct layers in variable depths were activated, while no change in the water content was seen at the other soil depths. Especially noteworthy are TDR10 and TDR11, which showed a strong soil moisture increase of up to 18 vol% below 1.4 m depth and around 10 vol% in the top 0.3 m of the soil. Profiles TDR13, TDR6, and TDR14 showed only weak signals, with the strongest response below 1.4 m below ground in TDR6.

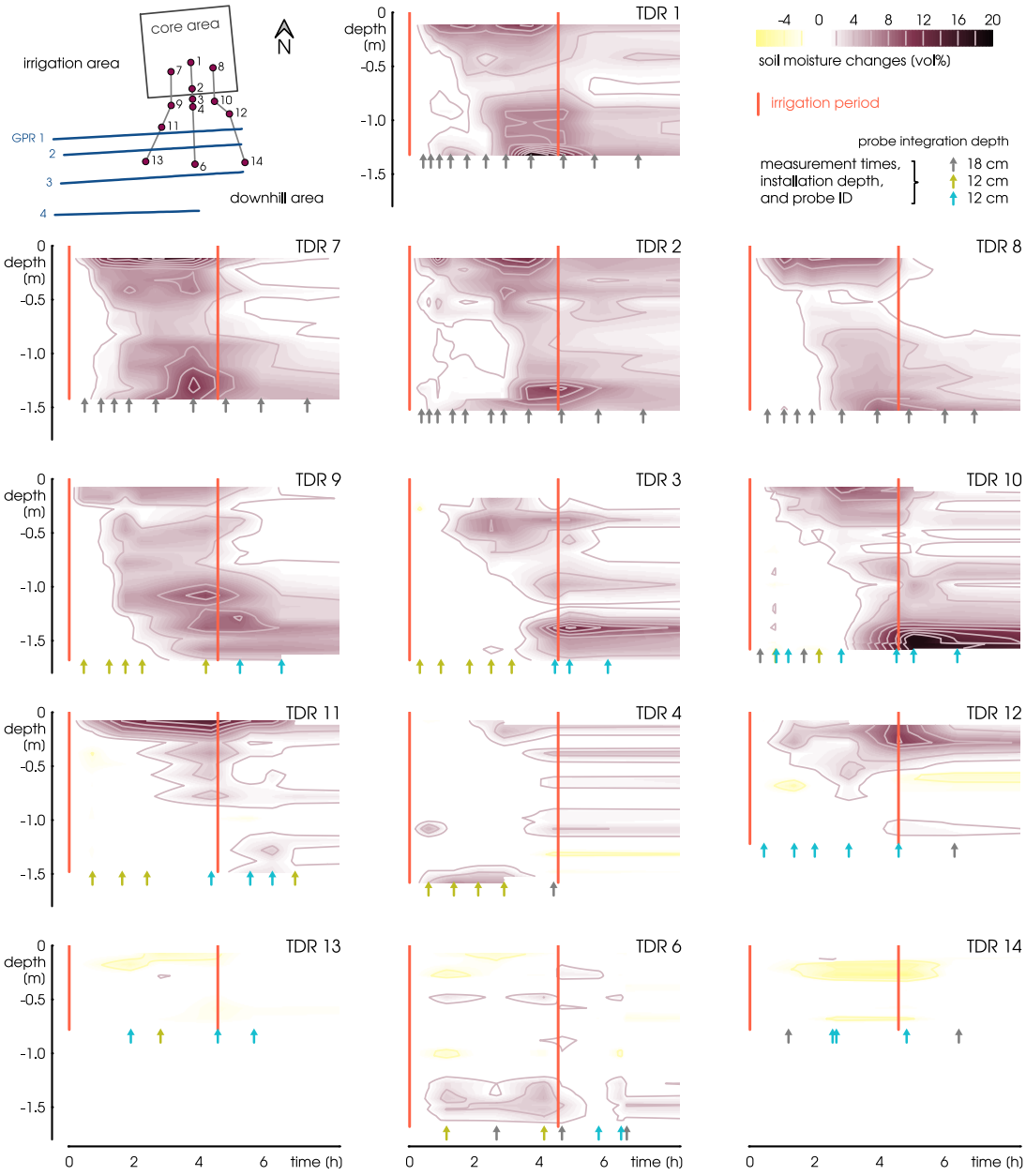

Figure 2.8: The soil moisture data measured at the TDR profiles at the irrigation site, showing the soil moisture dynamics in depth. The top four plots show all four core area profiles; columns are arranged according to the three diverging transects in the downhill direction. Rows are approximately at the same contour line. Measurements were taken at 0.1 m increments. While data analysis was based on non-interpolated data, soil moisture measurements were here interpolated linearly for better visualization. The plots cover the time from irrigation start until 9:00 h after irrigation start to focus on the first soil moisture response. Arrows indicate the measurement times and installation depth of each TDR profile. Time is given in hours after irrigation start.

Profiles TDR6, TDR12, TDR13, and TDR14 showed the strongest decrease in soil moisture over the course of the measurements, with $\Delta\theta$ of -3.7 vol%, -4.4 vol%, -5.8 vol% and -6.2 vol% at certain depths, indicating vanishing free water from the storm event. While the results from the left (TDR7, TDR9, and TDR11) and right (TDR8, TDR10, and TDR12; see Fig. 2.8) transects suggested lateral flow at different depths, the central transect (TDR1 through TDR6) did not indicate lateral flow.

Time-lapse GPR dynamics

The TDR measurements at the downhill monitoring area were complemented by the 2D time-lapse GPR measurements (see Fig. 2.4), yielding 2D images of structural similarity attributes referenced to the last measurement 24:00 h after irrigation start (Fig. 2.9). The first irrigation signals (shown in blue in Fig. 2.9) appeared in the first measurement after irrigation start (1:28 h), with transect 1 showing the clearest response. After about 3:23 h strong, localized signals occurred and increased in intensity over time. The general maximum was reached approximately 5:18 h after irrigation start, showing distinct activated flow paths. Most signals started to decline after 6:45 h, which is 2:10 h after the end of the irrigation period.

In transect 2 some weak signals appeared at the depth below 2.5 m 1:30 h after irrigation start. At this time, the signal was close to the noise level, but the pattern became stronger and more distinct in the following measurements. At transect 3, the persisting signal of the natural rain event made it difficult to identify the irrigation-induced response. However, a weak irrigation signal appeared after 1:28 h and reached its maximum at 6:45 h after irrigation start. At transect 4 the structural similarity attribute values were generally low, which indicates a low deviation from the reference state. Either the mobile water showed low dynamics (with regard to total mass over time) or water was less confined to specific structures and local changes are less pronounced. Both interpretations suggest that this transect was generally wetter due to its proximity to the river. Overall, the experiment does not appear to have affected this transect much.

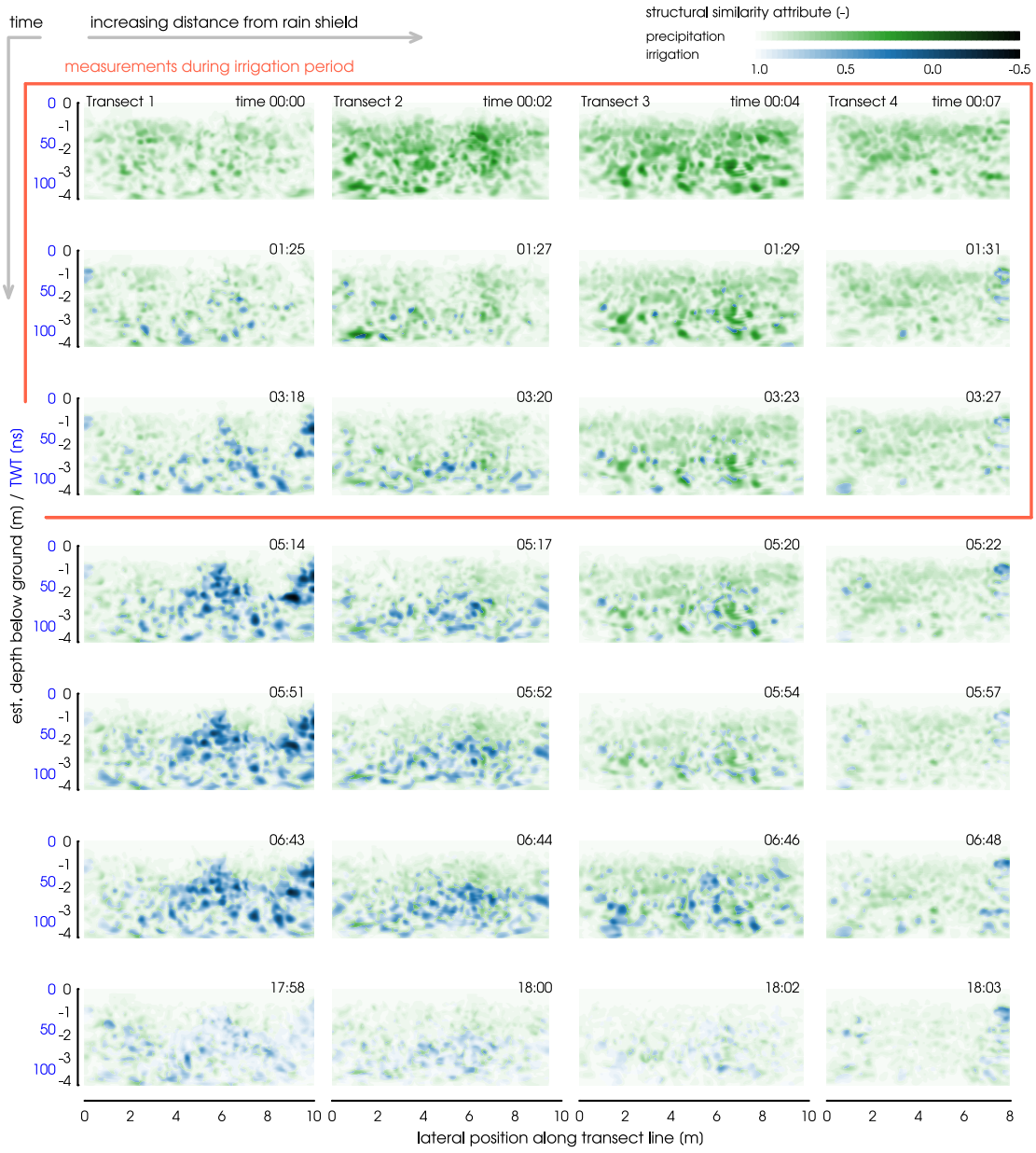

Figure 2.9: Structural similarity attributes calculated from time-lapse GPR data. All measurements were referenced to the last one 24:00 h after irrigation start, indicating changes in the GPR reflection patterns associated with soil moisture changes. A structural similarity attribute value of 1 indicates full similarity, lower values signify higher deviation from the reference state. Water from the preceding natural rain event (green) was identified by constant or increasing structural similarity attributes. Water from the experimental irrigation (blue) was identified by decreasing values after irrigation start by more than 0.15. Within one column the rows give a sequence over time (after irrigation start). Columns proceed downhill, with increasing distance from the rain shield.

The overviews of all GPR measurements in Fig. 2.6B and 2.9 visualize the dynamics of the hillslope section. The green natural rainfall signal faded from uphill to downhill, with the highest intensity and duration in transect 3. After irrigation start, the blue irrigation signal appeared, gradually propagating downhill and eventually overpowering the natural rain signal; 18:00 h after irrigation start (i.e., 13:25 h after irrigation ended), no changes in the GPR signal could be observed anymore. This suggests steady soil moisture conditions and, thus, the absence of highly mobile water in all transects. The mobile water either left the monitored area or dispersed by diffusion into the matrix surrounding preferential flow paths, where it remained beyond the time of the reference measurement and thus would not have been visible by means of structural similarity attributes.

Pore water and piezometer isotope responses

The temporal dynamics of the stable isotope compositions of the pore water (selected depths shown as circles in Fig. 2.10A) partially traced the signals of the rainfall and irrigation water input (green lines and blue triangles). The high δ^2H signal of the first minor rainfall event (dark green) was clearly visible in the top 10 cm below ground in the profile sampled at the downhill monitoring area 24:00 h before irrigation and 5:30 h after this rainfall event (Fig. 2.10A and C). Similarly, the isotope signal of the second rainfall event (light green) and the irrigation water (blue triangles) could be seen in the top 10 cm of the soil at the core area and the downhill monitoring area, respectively. Especially the isotope profile taken at the core area after irrigation showed an increase in δ^2H in the top 0.85 m below ground, showing the influence of both, the irrigation and the event water. Below the depth of approximately 1.2 m of all profiles, the soil water isotope composition seemed not to be impacted by the rainfall events and irrigation.

Only a few mL of water were seeping into the piezometers, with piezometer B being the only one that could be sampled more than once. Piezometer B was sampled first shortly after the irrigation ended (0:20 h) and showed a composition that was close to the irrigation water. The other two samples 1:32 h and 13:38 h after irrigation ended showed a decrease in δ^2H, towards the composition of the soil water (red and orange diamonds in Fig. 2.10).

Piezometers A, C, G, and H were sampled once 13:38 h after irrigation ended. The water sampled from piezometers at the core area (A and C, pink diamonds) showed the same composition as the irrigation water. Piezometers located at the

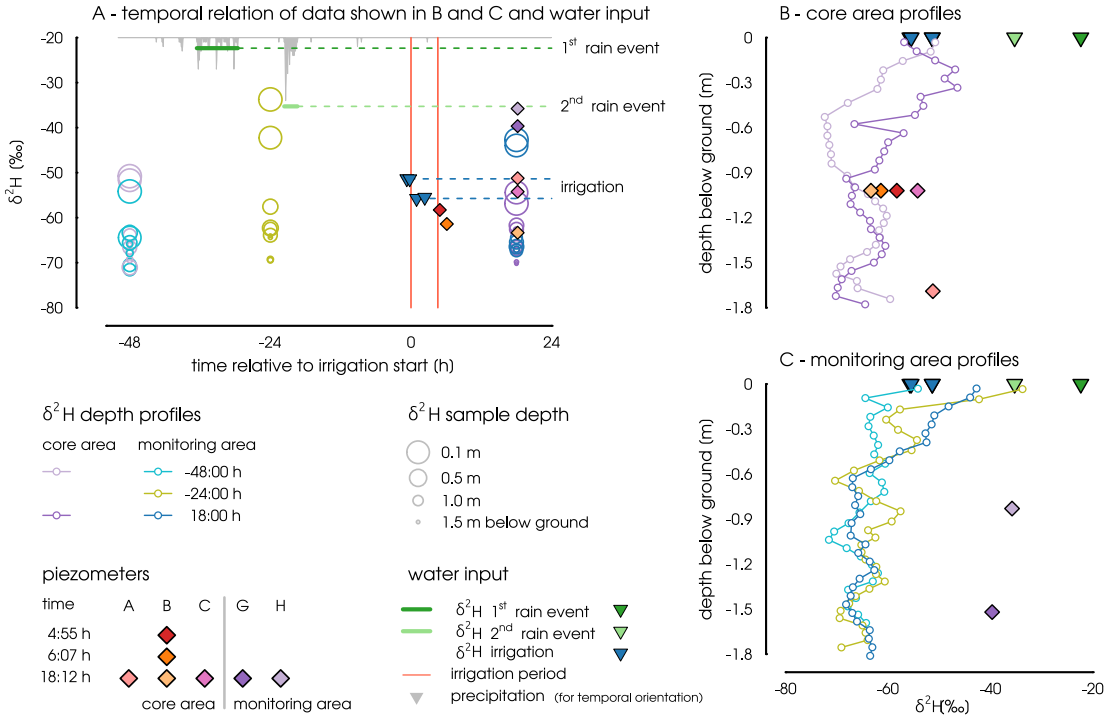

Figure 2.10: Stable isotope data from precipitation, irrigation, piezometers and pore water samples. A: Temporal dynamics of pore water and piezometer δ^2H in relation to water input by precipitation (bulk samples) and irrigation. Pore water data are shown only for the depths of piezometer filters (compare with panels B and C for depths) and the top soil (0.1 m below ground). The graph shows the direct impact of the water input (dashed lines) on the pore water isotope composition of the top soil. B and C: Pore water and piezometer data over depths, separated by core area and monitoring area. Water input stable isotope data are indicated at the soil surface.

downhill monitoring area (G and H, purple diamonds) in contrast, showed an isotopic composition similar to the rainfall water and different to the pore water in the depth profiles.

Response velocities

The results of the calculated response velocities from TDR and GPR measurements are summarized in Fig. 2.11. The top row shows the depth distribution of observed response velocities for all GPR transects. The bottom row shows the TDR-based results, separated into core area profiles and the three diverging transects. In addition to the TDR-based response velocities, GPR-based velocities observed in 0.5 m wide sections of the GPR transects which were closest to the displayed TDR profiles are shown in the plot.

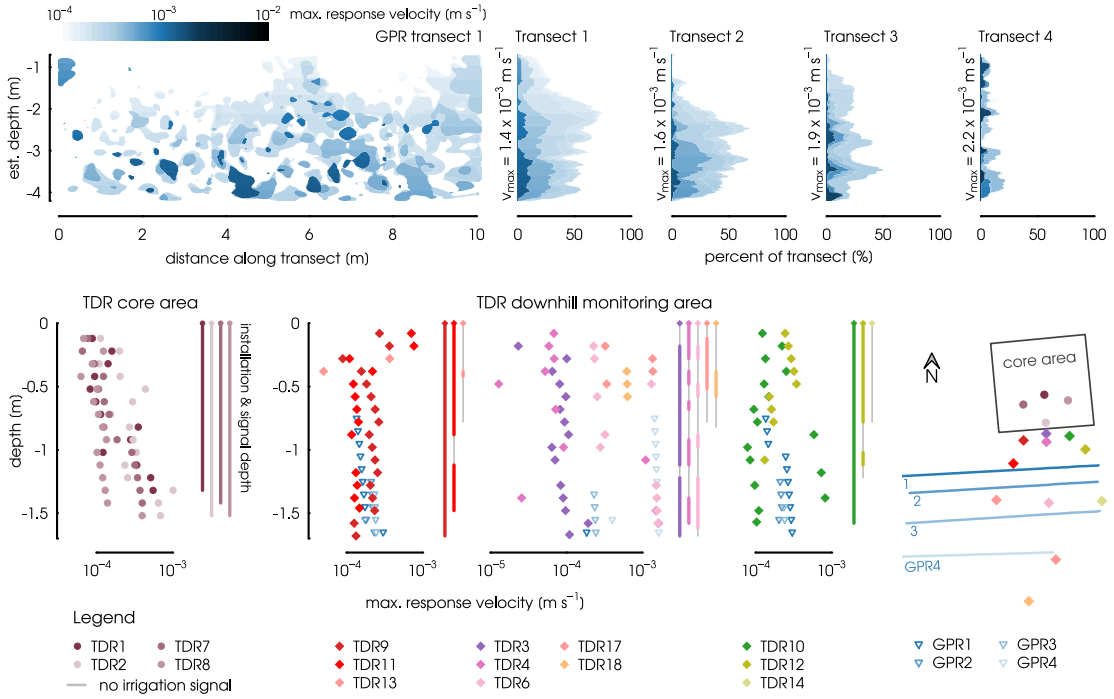

Figure 2.11: Top: depth distributions of response velocities calculated for the four time-lapse GPR transects. The blue scale indicates the response velocity calculated from the time of first arrival for each pixel. White areas did not show any irrigation water signal. The left plot shows the 2D results for GPR transect 1. The margin plots on the right show the same data accumulated to 1D depth profiles for all four GPR transects. Bottom: response velocities calculated from TDR measurements at the core area (strictly vertical), and the downhill monitoring area (vertical and lateral). The three plots showing the TDR profiles at the downhill monitoring area are sorted according to the three diverging transects. Lines within the right margins of each TDR plot show the installation depths of the TDR profiles. Grey sections indicate depth increments that did not show a change in soil moisture. Additionally, GPR-based velocities derived from 0.5 m wide sections of the GPR transects close to the TDR profiles are shown.

At the core area, the dominating vertical response velocity was around $10^{-4}\,\mathrm{m\,s^{-1}}$, with a tendency to increasing velocities with depth (Fig. 2.11, bottom left). As response velocities were calculated for the entire soil profile above the measuring depth, this increase indicates a bypass of intermediate depths through preferential flow paths, and a limited and slow interaction with the matrix. The highest observed vertical velocity was $10^{-3}\,\mathrm{m\,s^{-1}}$ at the depth of 1.4 m below ground. The respective soil moisture signal was recorded in the very first measurement after irrigation start, which indicates that we might have even missed the first response.

Similar to the vertical response velocities, the dominant TDR-based response velocity at the downhill monitoring area was on the order of magnitude of $10^{-4}\,\mathrm{m\,s^{-1}}$ (Fig. 2.11, bottom row). Response velocities of around $10^{-3}\,\mathrm{m\,s^{-1}}$ were observed in

six profiles all over the downhill monitoring area, of which the highest values (1.0 to $1.6 \times 10^{-3}\,\mathrm{m\,s^{-1}}$, TDR4, TDR6, TDR11, and TDR17) are based on signals observed during the first profile measurements after irrigation start. The fastest response was observed in the top 0.5 m (TDR4, TDR6, and TDR10) of the soil and below a depth of 1 m (TDR6, TDR11, TDR17, and TDR18).

The GPR-based response velocities were calculated for the entire width of the GPR transects down to the maximum depth of approximately 4.2 m (Fig. 2.11, top), and for single sections close to the TDR profiles down to a depth of 1.7 m for comparison of the methods (Fig. 2.11, bottom). All GPR transects showed slight and localized irrigation water signals in the data collected 1:28 h after irrigation start (see Fig. 2.9). This translates to response velocities between $1.4 \times 10^{-3}\,\mathrm{m\,s^{-1}}$ to $2.2 \times 10^{-3}\,\mathrm{m\,s^{-1}}$, depending on the distance to the irrigation area and signal depth. The data of the next measurements suggest response velocities between $5.4 \times 10^{-4}\,\mathrm{m\,s^{-1}}$ to $8.2 \times 10^{-4}\,\mathrm{m\,s^{-1}}$. Given the fact that this is a conservative estimate due to the even sparser temporal resolution in comparison to the TDR measurements, response velocities are likely to be similar to or even higher than those calculated from TDR results.

2.3.3 Comparison of natural event and irrigation response patterns in 2D GPR images

The signal of the natural rainfall event could be observed throughout the entire transects. The highest signal density (i.e., areal share) was found between 0.9 m to 1.7 m depth in all transects (Fig. 2.12). The strongest response signal (i.e., lowest structural similarity) appeared below the depth of 2.5 m in transects 2 and 3 (Fig. 2.6A). Transect 4 showed generally higher structural similarity and thus lower response signals.

The areal share of the irrigation signal was highest in transect 1, with 30.5 %, and decreased downhill, with 20.8 % in transect 2, 16.4 % in transect 3, and 6.0 % in transect 4 (Fig. 2.12). Transect 1 also shows irrigation signals in regions which were not (or no longer) active after the rain event. Especially the depth between 1.7 m to 2.2 m, which showed a comparably low natural rain signal, was activated by the irrigation. Overall 39.7 % of the regions activated by irrigation in transect 1 have already been active before irrigation. In contrast, most of the irrigation signals observed in the three downhill transects consist of re-activated flow paths. Here,

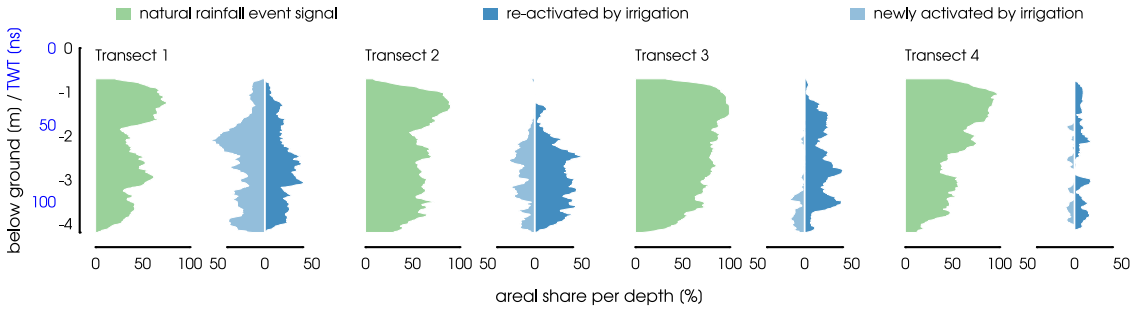

Figure 2.12: Areal share of activated regions per depth for the natural rainfall event and the irrigation. Areal share of activated regions for the natural rain event were calculated from the first measurement only (also depicted in Fig. 2.6A). Values for the irrigation experiment were accumulated over all GPR measurements, counting every pixel that had been activated after irrigation start. Here, we distinguish between pixels that had been active before irrigation start and were re-activated again, and pixels which have not been active previously and were newly activated by irrigation.

72.5 %, 86.1 % and 76.7 % of the irrigation patterns were activated both by the natural event as well as by the irrigation experiment (Fig. 2.6 and Fig. 2.12).

The natural rain signal in the GPR data was overpowered by the irrigation signal at 20:27 h, which is 3:23 h after irrigation start and about 22:45 h after the natural rain event. The dampened dynamics in transect 4 were due to its proximity to the river and therefore generally wetter conditions and less capacity for additional wetting.

The comparison of the two different response patterns shows that both the natural rain event and the irrigation caused advective flow in discrete flow paths more or less evenly distributed over the hillslope cross section. An (ephemeral) groundwater body or specific flow layers could not be identified in the top 4.2 m of the subsurface. Furthermore, the artificial irrigation had only a minor impact in comparison to the natural rain event, despite higher local input. Water that was supplied from upslope areas was therefore more important for the downhill soil moisture response than irrigation intensity or duration.

2.4 Discussion

2.4.1 Process interpretation

Irrigation experiment

The mass balance at the core area showed an overshoot in calculated mass recovery (i.e., higher mass recovery than water input) during the first 1:00 h of the irrigation period (Fig. 2.7). This might have been related to the spatial heterogeneity in irrigation intensity or lateral redistribution of water in the shallow subsurface. After that, the four core area profiles behaved differently. While the storage change in TDR7 was continuous, the other three profiles showed a stepwise behavior with abrupt stagnation and storage increase. This behavior is interpreted as a stepwise activation of flow paths in the vertical or lateral direction. The decrease in mass recovery, which started approximately 1:00 h after irrigation start, signified a loss of water from the core area (0 m to 1.4 m depth). Thus, flow paths towards greater depth and the downhill monitoring area were established around this time.

The fast soil moisture response in depth and at the downhill monitoring area (Fig. 2.8), as well as the immediate drainage after irrigation stop according to the mass balance (Fig. 2.7), also suggested a high fraction of mobile water. The mobile water is not bound to the matrix and is likely subject to advective flow with high velocities of over $10^{-3}\,\mathrm{m\,s^{-1}}$, according to the first response observed in TDR and GPR measurements (Fig. 2.11). The order of magnitude of the response velocities agreed with the *in situ* measurements of hydraulic conductivity (up to $10^{-3}\,\mathrm{m\,s^{-1}}$ and higher), but clearly exceeded the potential of matrix flow for the silty matrix material [Jackisch et al., 2017]. Similarly high preferential flow velocities are reported for the well-studied MaiMai hillslope, with initial breakthrough velocities between $6.7 \times 10^{-3}\,\mathrm{m\,s^{-1}}$ and $3.3 \times 10^{-2}\,\mathrm{m\,s^{-1}}$, being at least two orders of magnitude higher than matrix flow, which ranges between $3.8 \times 10^{-6}\,\mathrm{m\,s^{-1}}$ and $1.04 \times 10^{-4}\,\mathrm{m\,s^{-1}}$ [Graham et al., 2010].

Various studies report a concentration of lateral preferential flow at a more or less impermeable bedrock interface for other sites [e.g., Graham et al., 2010, Tromp-Van Meerveld and McDonnell, 2006], which has also been hypothesized for the Colpach River catchment [e.g., Fenicia et al., 2014]. However, none of the piezometers showed a significant reaction, despite being installed at the depths with the highest observed soil moisture responses. Instead, both TDR profiles and the time-lapse GPR tran-

sects revealed very heterogeneous patterns and a soil moisture response at multiple depths, as was also reported by Wienhofer et al. [2009] for a mountainous hillslope with young, structured soils.

The heterogeneous flow patterns observed with both TDR and GPR (Fig. 2.8 and 2.9) and the delayed signal at the intermediate depth at the core area (Fig. 2.8) suggest a heterogeneous network of preferential flow paths which bypassed a large portion of the unsaturated soil [see also Jackisch et al., 2017]. The water passes either through the intermediate depth outside of the monitored soil volume or through small preferential flow paths. If the volume of these flow path is small in comparison with the soil volume monitored by the TDR probes, they will only become visible (by means of soil moisture changes) if the water leaks into the surrounding matrix and effectively increases the soil moisture content of the integration volume. While we can not distinguish between these processes by means of the data presented here, both are preferential flow processes acting at different scales.

The pore water and piezometer stable isotope composition at the core area showed that the freely percolating water on the core area was predominantly constituted of irrigation water (Fig. 2.10A and B). Piezometer B, which was the only piezometer to be sampled more than once, showed a trend from irrigation water composition shortly after irrigation stop, towards pore water composition, 14:38 h later. This trend indicates that the irrigation water first percolated through the preferential flow paths without significant mixing with old water. After the water supply ended, however, preferential flow is (partially) fed by pore water, suggesting mixing and interaction between matrix and preferential flow as suggested by Klaus et al. [2013].

The piezometers at the downhill monitoring area in contrast had water with the same isotopic composition as rain water of the second rainfall event prior to the irrigation experiment (Fig. 2.10C). This water has been re-mobilized, as it only seeped into the piezometers after the irrigation, but shows no signs of interaction with the soil matrix or the irrigation water. While this observation has previously been made in other soils with well-developed macropore systems [Leaney et al., 1993], it contradicts the observations at the core area and rather suggests dual flow domains.

Natural rainfall event observations

The subsurface response patterns revealed by the GPR measurements after the natural rainfall events were similar to the irrigation-induced patterns with regard to their patchiness, but showed a slightly different spatial distribution (Fig. 2.12). The GPR measurements 12:52 h after the second rain event showed the highest density of response signals at the depths between 0.8 m to 1.7 m (Fig. 2.6A and 2.12). The higher response in the shallow depth could be interpreted as the signal of vertically infiltrating rain water, which did not occur during the irrigation. This depth also correlated with high stone content and the periglacial cover beds, which are characteristic of the area [e.g, Juilleret et al., 2011] and have also been observed at the investigated hillslope [Jackisch et al., 2017]. While no (transient) water table could be detected at the monitored depth, the patterns indicated a concentration of preferential flow paths at this depth.

The TDR measurements also showed high initial soil moisture content and a strong reaction to irrigation at this depth. However, except for the slight decrease in soil water content in the most downhill located TDR profiles TDR6, TDR13, and TDR14 (Fig. 2.8), soil moisture values barely fell below the initial values measured between rainfall events and irrigation. This means that the signal of the natural rainfall events was already gone in the shallow subsurface, and that no information on the natural rain signal in this depth can be derived from the TDR data.

The timing of the response dynamics observed with the GPR measurements and the discharge response also shed light on the prevalent processes. The natural rainfall events ended at 21:35 h on 20 June 2013. The first GPR measurements were taken at 10:42 h on 21 June, 12:52 h later (Fig. 2.5). Located at the lower section of the hillslope, they were interpreted to show a declining soil moisture signal, which was mostly gone 37:22 h after the second rainfall event (i.e., 18:00 h after irrigation start; see Fig. 2.6B and 2.9). Following the hypothesis of a top-to-bottom drainage of the hillslope, and considering the downslope location of the study site, the recorded signal represented the tailing of the shallow subsurface flow response to the natural rainfall event.

At the time of the first GPR measurements, the first peak of the hydrograph was already gone, while the second peak was on its rising limb and reached its maximum 12:00 h later (24:52 h after the rainfall event, Fig. 2.5). Thus, the following decline in subsurface response was observed after the first peak and coincided with the rise of the second peak. This timing provides strong evidence that the second hydrograph

peak was not primarily caused by the activation of the observed preferential flow paths in the shallow subsurface.

Several studies investigated the double-peak hydrographs of the Weierbach catchment. Wrede et al. [2015] used dissolved silica and electrical conductivity and found that the first peak was dominated by event water, while the second peak mainly consisted of pre-event water and strongly depended on antecedent conditions. Based on these observations, the first peak was attributed to fast overland flow from near-stream areas, while the second peak was attributed to subsurface flow where antecedent water was mobilized. Fenicia et al. [2014] came to a similar conclusion and identified a riparian zone reservoir as the origin of the first peak.

The stable isotope data collected during the rainfall event prior to the irrigation experiment showed the same dynamics as observed by Wrede et al. [2015] (Fig. 2.5). The isotopic composition of the first peak suggested a mixture of event water and pre-event water, while the composition of the second peak indicated the dominance of pre-event water. A simple water balance revealed that the total mass of the first peak of the event runoff accounted for about 4 % of the precipitation amount in the Holtz 2 catchment. Based on a rough delineation, the existing wetland patches in the catchment amounted to approximately 800 m^2 in the source area. Thus, the specific discharge of the first peak exceeded the existing wetland patches or riparian zones of the headwater catchment by a factor of 27, and suggests that overland flow from near-stream saturated areas could not solely explain the observed discharge response.

Synthesis: functioning of the investigated hillslope

Comparison of the GPR data of the natural rainfall event and the irrigation reveals that the response in the shallow subsurface was stronger after the natural event, even though the input per square meter was much lower than during the irrigation (Fig. 2.6 and 2.12). The observed response could only be caused by the accumulated water input of the (entire) hillslope draining through the shallow subsurface. Thus, the hillslope is prone to a substantial amount of lateral flow, which quickly ceases after water supply stops.

In combination with this finding, the high potential response velocities (Fig. 2.11) revealed by the TDR and GPR measurements show that fast, lateral subsurface flow is an important process in the investigated hillslope. The timing of the hydrograph dynamics and the declining response in the GPR measurements described above

(Fig. 2.5), as well as the freely percolating rainfall event and irrigation water shown by the stable isotope data, give further evidence that the activation of preferential flow paths within the shallow subsurface was contributing to the first immediate peak of the stream hydrograph.

Preferential flow paths were established quickly, and high response velocities have the potential to route water from the hillslopes towards the river within a few hours. The presence of preferential flow paths and the steep slopes in the Colpach River catchment were reported to enable subsurface runoff, even at times when the soil and weathered zone are not yet at field capacity [Van den Bos et al., 2006, Nimmo, 2012].

Many catchments reportedly showing double-peak hydrographs are headwater catchments with predominantly steep slopes and shallow soils, and in many cases with periglacial slope deposits [Burt and Butcher, 1985, Onda et al., 2001, Graeff et al., 2009, Birkinshaw and Webb, 2010, Fenicia et al., 2014, Wrede et al., 2015, Martínez-Carreras et al., 2016]. Such systems are characterized by pronounced subsurface structures and therefore are prone to heterogeneous flow patterns and preferential flow at the plot and hillslope scale. The results on subsurface structures presented in the companion study by Jackisch et al. [2017] also support this interpretation. Inter-aggregate flow paths at the scale of 5×10^{-3} m to 5×10^{-2} m are the reason for the highly variable hydraulic conductivities found in the investigated area and enable such high flow velocities. While these structures could only be revealed at the plot scale by excavation and direct observation, the related patterns in soil moisture response were similar (with regard to activated depths and response velocities) also for lateral flow at the hillslope scale.

The processes causing the second peak could not be resolved with this study, but it is hypothesized that deep percolating water from hillslopes and plateaus caused the delayed response. This hypothesis is backed by a study comparing catchments of different geology [Onda et al., 2001], where the prolonged response of double-peak hydrographs was identified as an indicator of deeply percolating subsurface flow through bedrock fissures. This theory might also apply to the catchment investigated here, with its fractured schist bedrock [Kavetski et al., 2011]. Furthermore, deep subsurface storage overflow [Zillgens et al., 2007] and fast groundwater displacement [Graeff et al., 2009] are processes which may play a role in the behavior of the Colpach River catchment. These hypotheses are also backed by the isotopic composition of the second peak, suggesting the dominance of pre-event water (Fig. 2.5), and

could explain the dependency of the occurrence of double-peak hydrographs on groundwater storage as described by Wrede et al. [2015] and Martínez-Carreras et al. [2016]. They could only observe the second peak if the groundwater storage was sufficiently filled, resulting in a hysteretic threshold behavior for the occurrence of the second, delayed peak.

2.4.2 Form and function in hillslope hydrology

Similar to the categorization of methods and observation data (see Fig. 2.1 and 2.3), we also subsumed the results and findings under the form–function framework. By doing so it becomes clear that observations and results are not always linearly obtained. Function is not necessarily described best by mere process observations. We thus want to discuss our data with regard to their value for the findings of the different form and function categories.

As elaborated in Sec. 2.4.1, the response observations described the functioning of the investigated hillslope well. Response dynamics and their temporal relation to each other across scales were a valuable source of information, shedding light on the characteristics of subsurface flow processes. Response patterns and spatially distributed point observations helped to develop a conceptual idea of the spatial organization and the functional network of the system. They were necessary to provide the spatial context to calculate response velocities and develop an idea of the establishment of flow paths. As such, response observations were the major key step towards understanding the investigated system, supporting hypothesis H1: the function of a system can be described by response observations.

In addition to pure response observations, however, basic knowledge about structures and local characteristics strongly improved the interpretability of our data and reduced the ambiguity. This basic knowledge included the presence of periglacial slope deposits, the downslope position at the hillslope, and the hydraulic conductivity of the subsurface. This information was easily obtained from the literature and in the field. It allowed us to close gaps in observational scales and link local observations at the hillslope to the overall system responses, and was thus the basis for more reliably relating the hillslope-centered and stream-centered observations.

Without this structural knowledge, the informative scope of response observations is limited to their observation scale. Transfer to other scales remains speculative. This fact is an important aspect, as most *in situ* response observations suffer

from limitations in spatial resolution. Soil moisture measurements are restricted by their integration volume, and point measurements in general struggle to cover the entire domain. Here, more detailed information on (flow-relevant) structures might greatly improve our process understanding.

While we were able to describe processes in the investigated hillslope in great detail, new findings on structure inferred from function observations are scarce. Any details on the form of the investigated hillslope we concluded from our observations were mere confirmations of previous knowledge. The only knowledge on form that was gained from the presented data concerns the flow relevance of structures. We found that substantial lateral flow occurrde at an intermediate depth, most likely associated with the periglacial slope deposits. We also found that distinct preferential flow paths occurred at greater depth, which suggests that the bedrock might not be as impermeable as previously assumed [Fenicia et al., 2014].

A conclusive picture of subsurface structures could not be drawn from process observations, partially refuting hypothesis H2, stating that function helps to reveal form. However, the observed response patterns were helpful in characterizing known structures with regard to their flow relevance.

2.4.3 Two-dimensional time-lapse GPR measurements as link between form and function

The interpretation of the time-lapse 2D GPR measurements is difficult and requires a thorough understanding of the expressiveness of the recorded data. While changes in structural similarity at a short timescale can be attributed to variations in soil moisture content, these changes contain no information on where the soil moisture content increases or decreases (i.e., the direction of the change). The data are qualitative information only, and supportive measurements are necessary. Therefore, GPR measurements need to be accompanied by other methods, such as TDR measurements or trenches, to provide a reference for the observed changes.

In the correct setup, however, 2D time-lapse GPR measurements are a very powerful tool and a valuable source of information. Especially in highly heterogeneous hillslopes, the spatial context provided by the GPR measurements is crucial for the investigation of complex preferential flow networks. This spatial context can not be provided by ERT measurements with their comparably low spatial resolution, highly invasive excavated trenches, or by point measurements. In contrast to dye

tracer excavations, time-lapse GPR measurements yield the temporal component, which is necessary to cover the highly dynamic processes. They allow for repetitions in time and space and avoid the manipulating effect of a trench face on flow through the unsaturated zone [Atkinson, 1978].

Within the form and function framework, the biggest advantage of 2D time-lapse GPR measurements lies in the combination of high-resolution spatial response patterns and dynamics. The method provides a direct link between flow-relevant structures and processes. It allows us to map response patterns in high temporal resolution, without manipulating the subsurface flow field. As such, this method has the potential to visualize the gradual establishment of flow paths, localize them, and calculate response velocities (hypothesis H3).

2.5 Conclusions

The study site is an example of headwater catchments with steep slopes and young and highly structured soils, typical of landscapes that formed under periglacial conditions. Here, preferential flow paths quickly developed in the unsaturated zone within minutes after the onset of an intense irrigation or rain event, causing lateral flow across the hillslope. In combination with the high response velocities of up to $10^{-3}\,\mathrm{m\,s^{-1}}$ or faster, and the large fraction of mobile water, these flow paths have the potential to quickly route water from the hillslopes towards the stream.

The strong dynamics and high spatial variability of preferential flow challenge the investigation of these processes. While we were able to describe the overall flow dynamics, the spatio-temporal resolution of our monitoring setup was not sufficient to reliably quantify the maximum response velocities. Our study has furthermore shown that causes and importance of observations can only be evaluated if the necessary context is known. This context includes knowledge of the spatio-temporal patterns on the one hand, and relevant process scales on the other.

The spatio-temporal context is provided by a combination of quantitative point measurements (TDR), qualitative mapping of patterns and dynamics (2D time-lapse GPR), and the observation of integrated system response (hydrographs and stable isotopes). Either of these approaches provides a substantial piece to the puzzle, while neither of them on its own would have provided the full picture. The experiment has shown that time-lapse GPR measurements are a powerful tool which provides new perspectives for the investigation of preferential flow processes in hillslopes. The methodology's flexibility and minimally invasive character allow for repetitions in time and space, and, thus, the direct observation of processes under driven conditions. Depending on the research question, the method can replace labor-intensive trenches and increase the observation density.

The observation of response patterns and dynamics by means of the TDR and GPR measurements was shown to suffice to characterize subsurface flow within the hillslope. Processes were identified and characterized without any concrete information about spatial structures. However, despite the high number of TDR observations and the 2D response patterns obtained from the time-lapse GPR measurements, our observations were methodologically limited in spatial and temporal resolution. Measuring intervals and integration volumes of the methods are restricted, and interpretations beyond observation scale remain speculative. Conclusions on or

links to larger or smaller scales are not reliable. Here, more detailed information on spatial structures and their impact on flow processes might improve our understanding of the investigated area and will improve the ability to transfer findings to other scales by providing the physical basis behind the observed processes.

The observed response patterns, revealed by the GPR and TDR measurements, allowed us to develop a conceptual description of the flow path network, which is linked to subsurface structures. Certain spatial characteristics of the flow path network such as layers prone to preferential flow could be inferred from the response patterns. However, actual structural features, such as the delineation of the deposit layer or the bedrock interface, could not be localized. All structure-related conclusions merely confirm previous findings. The topic of structural exploration is taken up in the companion paper by Jackisch et al. [2017] to further elaborate and discuss the methodological aspects linking form and function in hillslope hydrology.

Acknowledgements We are grateful to Marcel Delock, Lisei Köhn and Marvin Reich for their support during fieldwork, as well as Markus Morgner and Jean Francois Iffly for technical support, Britta Kattenstroth for hydrometeorological data acquisition and isotope sampling and Barbara Herbstritt and Begoña Lorente Sistiaga for laboratory work. Laurent Pfister and Jean-Francois Iffly from the Luxembourg Institute of Science and Technology (LIST) are acknowledged for organizing the permissions for the experiments and providing discharge data for Weierbach 1 and Colpach. We also want to thank the three anonymous reviewers, whose thorough remarks greatly helped to improve the manuscript. This study is part of DFG-funded CAOS project *From Catchments as Organised Systems to Models based on Dynamic Functional Units* (FOR 1598).

Chapter 3

Connectivity in the soil profile and at the hillslope

Form and function in hydrology: in situ imaging and characterization of flow-relevant structures from point to hillslope scale

Conrad Jackisch, Lisa Angermann, Niklas Allroggen, Matthias Sprenger, Theresa Blume, Jens Tronicke, and Erwin Zehe

Originally published in:
Hydrology and Earth System Sciences, 21(7):3749 – 3775, 2017.

Abstract

The study deals with the identification and characterization of rapid subsurface flow structures through pedophysical and geophysical measurements and irrigation experiments at the point, plot and hillslope scale. Our investigation of flow-relevant structures and hydrological responses refers to the general interplay of form and function, respectively. To obtain a holistic picture of the subsurface, a large set of different laboratory, exploratory and experimental methods was used at different scales. For exploration these methods included drilled soil core profiles, *in situ* measurements of infiltration capacity and saturated hydraulic conductivity, and laboratory analyses of soil water retention and saturated hydraulic conductivity. The irrigation experiments at the plot scale were monitored through a combination of dye tracer, salt tracer, soil moisture dynamics, and 3D time-lapse ground penetrating radar (GPR) methods. At the hillslope scale the subsurface was explored by a 3D GPR survey. A natural storm event and an irrigation experiment were monitored by a dense network of soil moisture observations and a cascade of 2D time-lapse GPR "trenches". We show that the shift between activated and non-activated state of the flow paths is needed to distinguish structures from overall heterogeneity. Pedophysical analyses of point-scale samples are the basis for sub-scale structure inference. At the plot and hillslope scale 3D and 2D time-lapse GPR applications are successfully employed as non-invasive means to image subsurface response patterns and to identify flow-relevant paths. Tracer recovery and soil water responses from irrigation experiments deliver a consistent estimate of response velocities. The combined observation of form and function under active conditions provides the means to localize and characterize the structures (this study) and the hydrological processes [companion study Angermann et al., 2017, this issue].

3.1 Introduction

3.1.1 Form–function relationship in hydrological sciences

From a general perspective the interplay of processes and spatial structures [Grayson and Blöschl, 2001] manifests itself as patterns in dynamics [Sivapalan, 2005] and self-organization [Zehe et al., 2013]. This interplay can be expressed as a form–function relationship, which is addressed in many disciplines. Especially in systems biology, the form–function relations are deeply rooted [e.g., Aristotle in Blits, 1999, Thompson, 1942] and under debate until today [e.g., Mugler et al., 2011]. In abstract terms the relation of form and function is fundamental for the concept that we can predict the behavior of a system under different forcings by knowing its constructive properties. In this respect we understand form as the shape and material property of the soil domain, whereas function refers to the dynamic behavior of water within the same. Although the existence and importance of form–function relationships are generally agreed upon, it is not clear to what extent form follows or reveals function and *vice versa*. In a soil-hydrological context of soil–water interactions the retention curve relates the pore size distributions and their covariance structure to storage of water against gravity and root water uptake. The hydraulic conductivity curve relates the pore size distribution and the interconnectedness of the pores to the conductance/release function of water depending on the wetting state. These are classic examples of form–function relations at the Darcy scale. However, the established relation does not directly translate to water displacement and contact angles at the actual pore scale [Armstrong et al., 2016]. At larger scales, accepted form–function relations turn out to be incomplete when preferential flow paths become important, as observed at plots of different soil types [Flury et al., 1994] and in most catchments [Uhlenbrook, 2006]. Form–function relations for plots and hillslopes should reflect how macropore density and connectivity in conjunction with the rainfall forcing and initial state control initiation and interaction of macropore flow with the soil matrix and thus ultimately export and redistribution of water from or within the control volume. In either case determining topology and connectivity (form) and understanding their implication for soil water transport (function) is seen as the "forefront of multiphase flow research" [Armstrong et al., 2016].

It is a long-standing vision in eco-hydrology to observe and characterize the form and function of all possible different flow paths in the subsurface. However, this is hindered by a lack of observation techniques which are capable of measuring and

visualizing flow paths across the relevant range of scales in a continuous manner. In this study, we address the challenge of *in situ* observation, identification and characterization of flow-relevant structures through a series of complementary methods at the point, plot and hillslope scale.

3.1.2 Identification and characterization of flow-relevant structures in the subsurface

While heterogeneity is seen as a purely random variation of soil properties, organized heterogeneity implies a spatial covariance of these properties and connected flow paths. As such we define structure based on their functional implication in line with Gerke [2012] and others. While such structures can be classical macropores like earthworm burrows [Palm et al., 2013, Blouin et al., 2013, Van Schaik et al., 2014], decayed root channels [Nadezhdina et al., 2010] or cracks and geogenic structures like voids in periglacial cover beds [Heller and Kleber, 2011], we also attribute connected inter-aggregate pores to structure. They have in common that gravity induced preferential subsurface flow is facilitated through the directed drainage paths, partially bypassing large sections of the soil. Beven and Germann [1982] initiated a discussion about macropores and preferential flow and more recently repeated that the topic is still not given the attention appropriate to its significance in all areas of soil and catchment hydrology [Beven and Germann, 2013].

Despite observation of fast responses through such macroporous networks, e.g., as tracer breakthroughs [Schotanus et al., 2012, Klaus et al., 2013] or in multimodal reactions [Martínez-Carreras et al., 2016], it was shown that quick responses of catchments are often fed by pre-event water [Neal and Rosier, 1990, Jones et al., 2006] which is known as the "old water paradox" [Kirchner, 2003].

Due to limited direct observability of subsurface flow, most evidence is either inferred from integral responses or derived from model applications: in the field, a large spectrum of methods is applied to investigate subsurface connectivity [Bishop et al., 2015, Blume and van Meerveld, 2015] and to quantify preferential flow [Allaire et al., 2009]. Dye staining has evolved as common practice since its first applications [presumably Bouma and Dekker, 1978] for a retrospective imaging of preferential flow paths. Even though Anderson et al. [2009] extended dye staining to the hillslope scale, the technique is usually limited to plot-scale applications. Another drawback is the requirement to excavate and thereby destroy the system, which prohibits

analyses of function under variable forcing. Application of salt tracers in the vadose zone adds a quantitative measure, but at lower spatial resolution than dye staining. It also suffers from the a posteriori inference about the retention of the solutes.

Furthermore, breakthrough curves of precipitation or irrigation events at trenches or springs are commonly used [e.g., McDonnell et al., 1996, Tromp-Van Meerveld and McDonnell, 2006, Bachmair and Weiler, 2014]. In combination with fluorescent, salt and natural tracers they can provide quantitative information over the course of rapid flow events at this scale [e.g., Wienhofer et al., 2009]. However, such measurements can only capture spatially integral signals and require one to infer the form by the observed function.

So far, relatively few studies managed to actually *in situ* image spatially distributed subsurface flow paths at larger scales. On the one hand, applicability is also often technically limited to very small scales: Schlüter et al. [2016] examined multiphase flow with time-lapse X-ray microtomography in a sample of 4.2 mL. Koestel and Larsbo [2014] presented an X-ray tomography study with a sample of 258 mL undisturbed soil. Gerke [2012] analyzed the pore fractions in two samples of 785 mL undisturbed soil through a medical CT X-ray scanner. Wehrer and Slater [2015] report findings from tracer breakthrough experiments in laboratory lysimeters accompanied by 3D time-lapse electrical resistivity tomography (ERT). Guo et al. [2014] conducted a multi-2D time-lapse ground penetrating radar (GPR) survey to identify preferential flow structures *in situ* in a $2\,m^2$ section of a hillslope. On the other hand, the lack of a unified theory of advective and diffusive soil water redistribution, mixing, storage and release [Beven and Germann, 2013] adds to unclearness about appropriate observation strategies.

Hydrological "standard approaches" attempt to explore parameters like soil layer depth, porosity and hydraulic conductivity based on distributed point-scale measurements. Also, state and flux monitoring most often consists of a set of point observations, e.g., of hydro-meteorological conditions and soil moisture. An appropriate sampling design is substantial for the statistical inference [e.g., De Gruijter et al., 2006]. Thus there is also a conceptual issue arising from the fact that such samples necessarily integrate over sub-scale structures, such as inter-aggregate pore networks. At the same time such an integral may not necessarily allow inference about structures at the larger scale exceeding the support of the observation. When the respective sampled set and the subsurface setting is basically unknown, spatial scaling of soil moisture [Western and Blöschl, 1999] and other observed variables

becomes problematic. In a "Special Section" on preferential flow, Gerke et al. [2010] highlighted that further analyses need to focus on the quantification of flow-relevant structures. They continue that experimentally non-invasive and imaging techniques are needed for research and model testing. We will take up these issues in the discussion section.

3.1.3 Hypotheses and overall aims of the study

The rationale of this study is to analyze insights into flow-relevant subsurface structures based on qualitative and quantitative measurements at the point, plot and hillslope scale. Specifically, we hypothesize that a combination of quantitative field methods and *in situ* imaging of subsurface response patterns with dye staining and time-lapse GPR provides the missing link between form of the flow structures and how their interactions determine rapid subsurface flow and thus function.

We test this hypothesis by addressing three main research questions.

Q1 What kind of information on sub-scale flow-relevant structures, their characteristics and their distribution can be inferred from a large set of direct point-scale measurements of soil hydraulic properties?

Q2 How do salt tracer data, dye tracer stains, soil moisture response patterns, and 3D time-lapse GPR compare with respect to inference on vertical flow channels and apparent flow velocities at the plot scale?

Q3 How do methods and identified structures convey to the hillslope scale?

The study approaches the identification and characterization of flow-relevant subsurface structures as the aspect of *form*. The alternative starting point towards hillslope process understanding is taken in the companion study by Angermann et al. [2017, this issue] with the aspect of *function*.

3.2 Experimental approaches and study methods

The study at hand approached the topic on three complementary scales with a range of different methods: as a standard reference, results from auger exploration and *in situ* measurements of hydraulic conductivity and infiltration capacity were collected. They were extended with pedo-physical laboratory examination of 250 mL

undisturbed ring samples for bulk density, porosity, texture, soil water retention characteristics, and saturated hydraulic conductivity. We then broadened the perspective to the plot scale with irrigation experiments accompanied by TDR (time domain reflectometry) measurements of soil moisture dynamics in a 1D profile, 3D time-lapse GPR imaging, and tracer recovery of dye, salt and stable isotopes. At the hillslope scale, 3D GPR was used to identify flow-relevant structures in a static survey. For dynamic investigation, an irrigation experiment specifically designed to identify lateral flow structures was observed by a dense network of TDR soil moisture profiles and a series of trench-like 2D time-lapse GPR transects.

3.2.1 Study site description

The study is situated in the headwaters of the Colpach River, a tributary of the Attert which has been investigated by several studies before [Pfister and Hoffmann, 2002, Hellebrand et al., 2011, Jackisch, 2015]. Located at the southern edge of the schistose Ardennes Massif, the soils are characterized by eolian loess deposits and weathered schist debris. The hydrological setting of quick catchment reaction to precipitation especially during the non-vegetated season has been subject to some process hypotheses related to the periglacial deposit layers and flow at the bedrock interface [Van den Bos et al., 2006, Fenicia et al., 2014, Wrede et al., 2015, Loritz et al., 2017]. Our measurements and experiments focus on two forested hillslopes (mostly managed stands of beech, *Fagus sylvatica*, with mixed shrubs; some measurements took place in stands of spruce, *Picea abies*). The agriculturally used plateaus at the hilltops are not examined here. Figure 3.1 presents a map of the area and the location of the respective measurements and experiments.

3.2.2 Pedo-physical exploration

The soil physical exploration addressed our research question Q1 using an intentionally large set of hydrological and geophysical methods to survey the subsurface. The sampling is guided by a network of hydro-meteorological monitoring stations measuring all relevant fluxes and states in the atmospheric boundary layer and the subsurface (research project “Catchments As Organized Systems”, Zehe et al., 2014a).

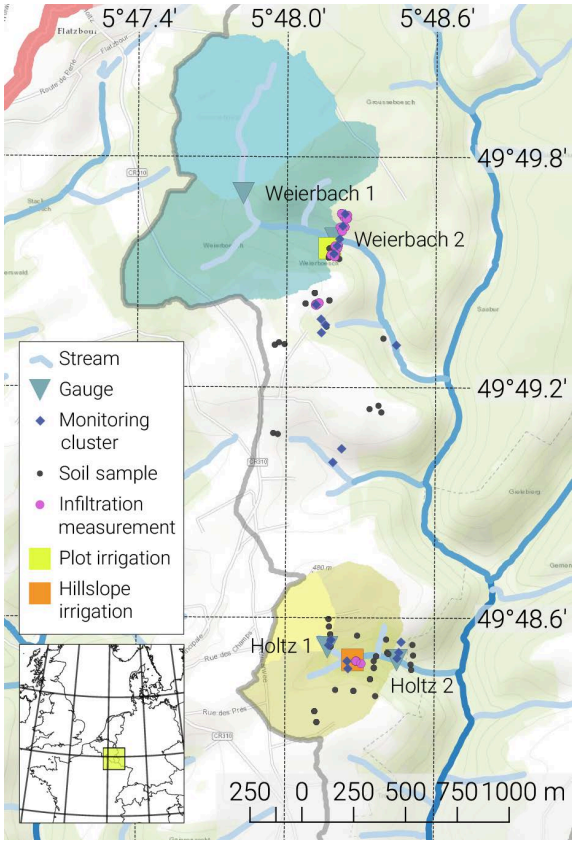

Figure 3.1: Map of the study sites in the upper Attert basin, Luxembourg.

Sampling design

Aligned with the monitoring stations, infiltration capacity and saturated hydraulic conductivity were measured. In order to address plot-scale (few meters) and hillslope-scale (a few hundred meters) heterogeneity, the design consisted of clustered sets of point measurements along two catenas plus one set at the site of the hillslope irrigation experiment presented in Sect. 3.2.4. A detailed map is included as Fig. A.1 in the Appendix.

The distance between the clustered sets was 80 m to 200 m. In each, three nested sets with a lag distance of 10 m to 20 m along and perpendicular to the contour line were defined. In such a nested set at least one measurement of infiltration capacity and two profiles (laterally spaced 1 m) of saturated hydraulic conductivity in different depth levels were conducted. To complete the scale triplet [Blöschl and Sivapalan, 1995], the respective support is given in the description of each technique.

In addition to the point measurements, a series of percussion drilled profiles (drill head diameter of 4 cm) as 1D profiles were drawn and 250 mL ring samples were taken within the top 0.6 m for laboratory analyses.

Exploration techniques

Infiltration capacity was measured at 40 points with a Hood Tension-Infiltrometer (IL-2700, UGT GmbH). It employs a tension chamber (12.4 cm radius) as infiltration water supply. Inside the chamber, a defined low negative pressure head is established, which allows a precise measurement of infiltration capacity at different tensions. Three to five tension levels between the 0 cm and 5.5 cm water column were applied at each spot.

In addition to infiltration capacity at the surface, we used a Compact Constant Head Permeameter (CHP, Ksat Inc.) for determination of saturated hydraulic conductivity in 32 borehole profiles with 3 to 7 depth levels of about 20 cm increments, with the lowest level at a depth where further hand-drilling was inhibited by stones. The permeameter establishes a constant water level (10.5 cm in our cases) above the bottom of a borehole (here 5 cm in diameter). The outflow is measured to calculate saturated hydraulic conductivity [Amoozegar, 1989].

The 63 undisturbed soil ring samples were analyzed for bulk density, porosity (assumed to be equal to saturated soil water content), soil water retention properties (Hyprop, UMS GmbH and WP4C Decagon Devices Inc.), saturated hydraulic conductivity (Ksat, UMS GmbH), and soil texture (ISO 11277, wet sieving and pipette method sedimentation).

3.2.3 Imaging and quantification of rapid flow in plot-scale irrigation experiments

In order to explore the network of flow-relevant structures and patterns of rapid subsurface flow, we conducted three plot-scale irrigation experiments. This relates to our second research question Q2. The general setup is very similar to the one described by Allroggen et al. [2015b], Van Schaik [2009], Öhrström et al. [2004] and Kasteel et al. [2002]. Marked on the map in Fig. 3.1, the three plots are located on a forested mid slope near gauge Weierbach 2 (see also Fig. A.1).

Experimental design and multi-method approach

Three plots of 1 m^2 size were irrigated each for 1 h with an intensity of 50 $mm\,h^{-1}$, 30 $mm\,h^{-1}$ and 50 $mm\,h^{-1}$ on 30 October, 1 November, and 2 November 2013, re-

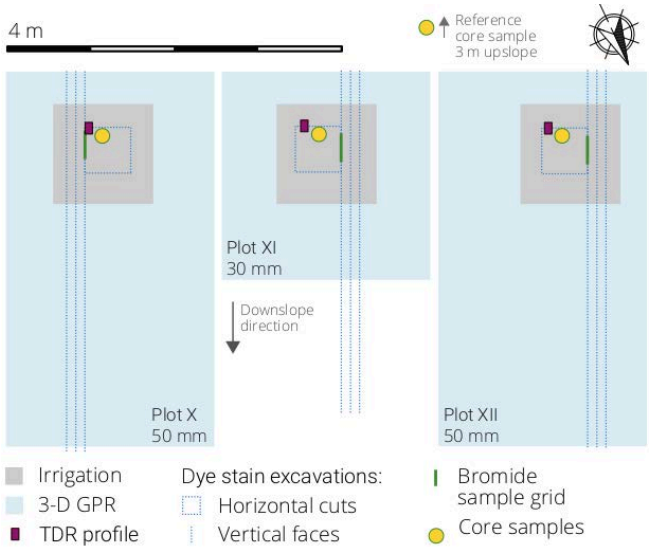

Figure 3.2: Plan view of the plot-scale irrigation experiments. Three irrigation plots (1 m^2, gray squares) are monitored by 3D time-lapse GPR (blue rectangles) and TDR (soil moisture tube probe, red box). The plots are sampled for tracer recovery by percussion drilled core samples (yellow dot) and in a grid on the last of three vertical faces (dashed blue line). Moreover, dye stains are excavated at horizontal cuts in the center of the irrigation area (dashed blue square). A pre-irrigation reference for porewater stable isotope composition is sampled as a fourth core 3 m upslope.

spectively. The relatively high rates were chosen to activate the potential flow paths, thereby establishing connectivity. A layout of the experiment is presented in Fig. 3.2.

The irrigation was accomplished by spray irrigation (full-cone nozzle Spraying Systems Co.) using a wind-protection tent. Brilliant Blue dye tracer (4 g L^{-1}) and bromide salt (5 g L^{-1} potassium bromide) were used for qualitative and quantitative reference, respectively.

In addition, temporal dynamics of soil moisture along a selected profile was monitored throughout the experiments through continuous TDR measurements in an access tube (Pico IPH, IMKO GmbH) down to 1.5 m depth and with a diameter of 4.2 cm. This technique is chosen to minimize the impact of sensor installation (percussion drilling and installation of the tubes from the surface) and to avoid interference with the GPR (sensor probe was removed during GPR measurements). The sensor measured an integral of about 1 L (depth increment of 18 cm , mean signal penetration of 5.5 cm). It was manually lowered in the tube to the respective depth for each reading. Each measurement took about 10 s. Hence the whole procedure added up to 4 min to 10 min per profile record. The procedure was continuously repeated until 1.5 h after irrigation onset in line with the findings of Germann and al Hagrey [2008] and Germann and Karlen [2016]. They propose that film flow in soil structures disperses into the matrix after 1.5 times the duration of a constant plot irrigation.

Two hours after the end of each irrigation, a percussion drilled soil core was taken (drill head diameter of 8 cm) and sampled in 5 cm depth increments down to 1 m. The plot was excavated 24 h after irrigation for vertical and horizontal recovery

of Brilliant Blue stains. This was done by successive digging of three vertical faces into the plot (aligned with the slope line, 0.1 m distance starting from the lateral edge) and five to seven horizontal cuts in different depth levels down to the first deposit layer (0.5 m by 0.5 m^2 in the center of the plot). On the third vertical face in the center of the plot core samples of 66 mL soil were taken in a 5 cm grid with 5 columns and 14 to 21 rows. In order to minimize time lags in the 70 to 105 individual samples, a quick-sampler (see Appendix A.1) was developed, allowing for precise and nearly undisturbed sampling.

Bromide recovery and stable isotope analysis

All samples were analyzed for bromide Br^-. This was done by oven drying the samples and consecutively suspending them in 150 mL de-ionized water (72 h in an overhead shaker at nine rotations per minute). The samples were then left 4 d for sedimentation to exfiltrate the excess through (a) filtration paper (5 µm to 13 µm) and (b) 0.45 µm PP micro-filter. The extracts were analyzed in an ion chromatograph (Metrohm 790 Personal IC) with an anion separation column (Metrosep A Supp 4 – 250/4.0) for Br^- concentration.

A recovery coefficient (RC) is calculated as proportion of recovered mass of Br^- in the soil samples scaled to the total irrigated area times the depth of the lowest sample. Through this we neglect lateral flow from the irrigation spot and further percolation in the calculation. We also assume the samples to be representative for the whole affected soil volume.

Prior to the bromide analysis, the percussion drilled soil core samples were also analyzed for their stable isotopic composition ($\delta^{18}O$ and δ^2H) of the porewater. See Appendix A.4 for details and results, which are given in comparison to the bromide recovery.

Calculation of apparent vertical flow velocity

The quantitative measurements allow one to infer apparent vertical flow velocity along the profiles. For bromide we employ a cumulative curve method [Leibundgut et al., 2011]. The distribution of the advective velocity v_{advect} is set to the depth distribution of the tracer concentration at the time of fixation t_{fix}. For the profile we assume apparent velocities:

$$v = z/t_{\text{fix}}. \tag{3.1}$$

Relating to our third research question Q3, they are projected to the recovered distribution of tracer concentration:

$$\Phi(v_{\text{advect},z}) = c_{\text{tracer},z} / \sum_{z=0}^{z_{\max}} c_{\text{tracer}}, \tag{3.2}$$

where z is depth and Φ is the cumulative distribution function. Obviously, the estimated travel velocity distribution depends strongly on the selection of t_{fix} somewhere between irrigation and excavation. This can scale v several orders of magnitude. Again, the reference of 1.5 times the irrigation duration is chosen [Germann and Karlen, 2016]. For Br^- in the sampled grids, each column was treated as an individual 1D profile. The calculation further assumes full tracer recovery.

Analysis of soil moisture responses

The individual TDR soil moisture measurements (θ) were projected to a regular grid of 0.1 m depth increments and 10 min time increments for visualization of changes compared to the initial records. As an alternative and independent estimate of vertical response velocities (research question Q3), we calculated the distribution of first exceedance of soil moisture by ≥ 2 vol% in each depth level z:

$$v_{\text{response}} = z/t_{\Delta\theta \geq 0.02} \tag{3.3}$$

For this the un-interpolated measurements were used.

3D time-lapse GPR

GPR is known as geophysical imaging technique with high spatial resolution [Huisman et al., 2003, Binley et al., 2015]. Applied at the shallow subsurface it has been proven as potential means to locate and characterize soil layers and subsurface structures [Holden, 2004b, Gormally et al., 2011, Steelman et al., 2012, Klenk et al., 2015]. GPR is also capable to monitor subsurface fluid migration in time-lapse approaches [Birken and Versteeg, 2000, Trinks et al., 2001]. Our experiments were monitored by 3D time-lapse GPR measurements as described by Allroggen et al. [2015b]. We employed a PulseEKKO Pro GPR system (Sensors and Software Inc.)

equipped with 500 MHz shielded antennas with constant offset of 0.18 m. Sampling interval was set to 0.1 ns, recording a total trace length of 100 ns in 8 internal stacks. Since precise positioning and accurate repeatability are key requirements, we used a kinematic survey approach relying on an automatic-tracking total station (Leica Geosystems AG, providing sub-centimeter coordinates) in combination with a portable measuring platform [Allroggen et al., 2015b].

Using this setup, we acquired one 3D GPR data cube before irrigation, one directly after the end of irrigation, and a last one about 20 h after irrigation for each plot. One survey took about 45 min. Allroggen and Tronicke [2016] have shown that a pixel-to-pixel comparison of the radar amplitudes (A) is not suitable for analyzing time-lapse GPR data in the presence of limited repeatability and noisy data. They propose a structural similarity attribute inspired by Wang et al. [2004] calculated in a moving window. It normalizes the cross-correlation $c_{x,y}$ of the residuals $(A - \mu_A)$ of two different acquisition times (x, y) by the product of their standard deviations (σ_A). They further introduced a as 10 % of the maximum amplitude to avoid numerical instabilities with near-zero σ values:

$$S_{\text{struct}}(x, y) = \frac{c_{x,y} + a}{\sigma_{A_x}\sigma_{A_y} + a} \tag{3.4}$$

In our study we calculate the structural similarity attribute S_{struct} of the pre-irrigation reference and the two post-irrigation records using a local Gaussian window of 2.5 ns along the vertical axis and 0.1 m along the horizontal axis. The attribute ranges between 1 and -1, with 1 being highly similar and -1 referring to most dissimilar. Points of low similarity indicate deviations that arise from changes in dielectric permittivity which likely reflect changes in local soil water content.

As an additional estimate of vertical response velocities, the same approach as for the soil moisture responses (Sect. 3.2.3) was employed with a threshold of the similarity attribute of zero between pre- and post-irrigation records.

3.2.4 Lateral subsurface flow paths in the hillslope

In order to examine the characteristics of flow-relevant structures and the periglacial deposit layers at the hillslope scale, we conducted an experiment on 21 June 2013 at a close-by hillslope. The experiment was specifically designed to explore the response in lateral preferential flow paths and to replicate the plot-scale experiments without tracer application. The site had to be chosen for facilitation reasons (permissions,

accessibility, collaboration within the CAOS research project). With reference to its hydrological responses [companion paper Angermann et al., 2017, this issue], vegetation, slope, soils and hydraulic properties, we consider the hillslopes to be very similar.

3D GPR survey of the hillslope

As an additional reference to the soil core profiles, a 3D GPR survey of the hillslope was conducted prior to the natural event and the irrigation. The GPR data processing relies on a standard processing scheme including bandpass filtering, zero time correction, envelope-based automatic scaling, gridding to a regular 0.03 m by 0.1 m grid, inline fk-filtering and a 3D topographic migration approach as presented by Allroggen et al. [2015a], using an appropriate constant velocity of $0.07\,\mathrm{m\,ns^{-1}}$.

For structural analysis, the processed data are imported into the OpenDtect software (dGB Earth Sciences). Under heterogenous soil conditions the derived data cube is dominated by complex reflection patterns which prohibit a classical structural analysis based on picking reflectors [as done in a study with a similar cope but different setting by Gormally et al., 2011]. Therefore, we support our interpretation of the 3D GPR data and picking of potential flow-relevant horizons by a dip-corrected semblance attribute. The attribute calculates the spatial coherency and highlights areas of coherent reflections [Marfurt et al., 1998]. Low semblance indicates more complex reflection patterns caused by high internal heterogeneity, possibly influencing the subsurface flow regime.

Experimental design

The experimental site is located at the lower part of a north facing hillslope. Vegetation is dominated by mixed beech forest. However, the experimental site is placed in an area with no major trees. Except for few young trees at the downhill monitoring area, all shrubs were carefully removed from the experimental site to accomplish GPR measurements and allow for undisturbed and homogeneous irrigation. The topographic gradient is about 14°.

The experimental layout is given in Fig. 3.3. Irrigation intensity, the duration of the experiment and the spacing of the observation profiles have been decided based on *a priori* modeling scenarios as described in Appendix A.3. The experiment was preceded by two strong storm events of 43 mm in total on 20 June. The events ended

20 h before irrigation onset. The irrigation of 141 mm in 4.5 h was fed from stream water and was realized by four circular sprinklers (Wobbler, Senninger Irrigation Inc.) arranged to overlap at a 5 m by 5 m core area with relatively homogeneous intensity. While boundary effects were mitigated by an irrigated buffer zone of about 4 m at the uphill and lateral borders of the core area, the downhill boundary was defined by a rain shield. This established a sharp transition to the non-irrigated area below. Water collected by the rain shield was routed off the experimental site. Irrigation was monitored by a flow meter to measure the absolute water input, one tipping bucket to monitor the temporal variability, and 42 mini rain collectors evenly distributed across the core area to check spatial heterogeneity of the intensity.

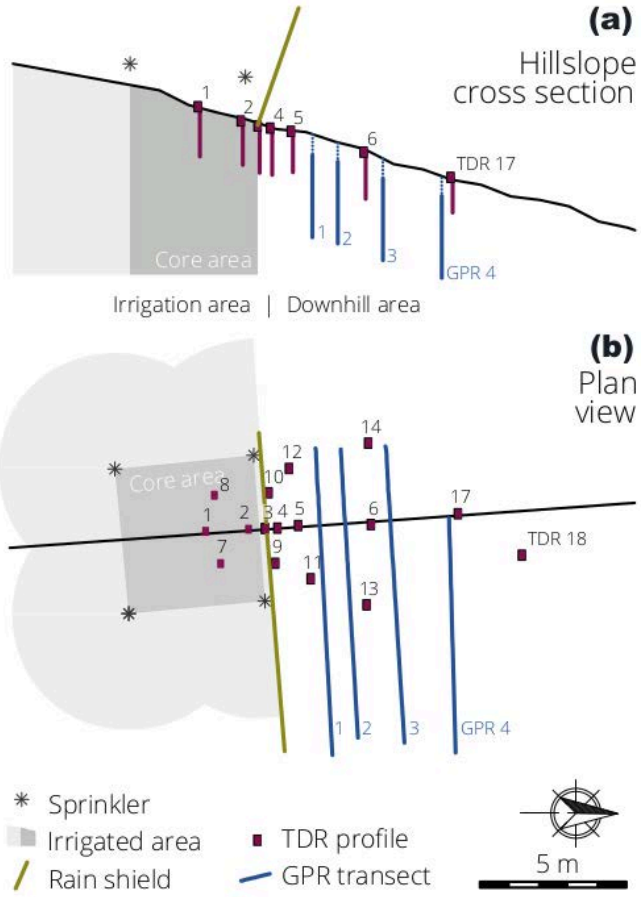

Figure 3.3: Layout of the hillslope-scale irrigation experiment as vertical view **(a)** and plan view **(b)**. The hillslope is divided into an irrigation area and a downhill area by a rain shield. Sixteen access tubes for TDR measurements of soil moisture profiles are arranged in three diverting transects. Parallel to the contour lines, four transects of 2D time-lapse GPR are recorded.

Moreover, a surface runoff collector was installed across 2 m of the lower boundary of the core area. It was built from a plastic sheet installed approximately 1 cm below the interface between litter layer and Ah horizon of the soil profile. At the downhill end of the sheet, the water was captured by a buried and covered gutter. An in-ground tube was attached to the deepest point of the gutter to conduct the

water to a tipping bucket downhill of the investigated area. The tube had been filled with water prior to the experiment to ensure an immediate reaction to the occurrence of surface runoff.

We monitored soil moisture dynamics in a setup of 16 access tubes with 3 manual TDR probes like in the plot-scale experiments (Imko GmbH, two with 12 cm integration depth and one with 18 cm). Measurements required manual positioning of the sensor probes for each reading. We continuously recorded the states in all tubes in 10 cm depth increments, realizing revisiting intervals of 520. The tubes were installed to reach to a depth of about 1.7 m. The layout consisted of three diverging transects with four TDR profiles in the lower half of the core area, the highest density of profiles just downhill from the rain shield, and the furthest profile about 9 m downhill.

Four 2D time-lapse GPR transects were treated as *GPR-inferred, non-invasive trenches* parallel to the contour lines located 2 m, 3 m, 5 m and 7 m downhill from the rain shield. Here, the GPR acquisition unit was equipped with shielded 250 MHz antennas. The data were recorded using a constant offset of 0.38 m, a sampling interval of 0.2 ns and a time window of 250 ns. Wooden guides and the automatic tracking total station guaranteed accurate and repeatable positioning.

Analysis of TDR data

In order to synchronize the almost 5000 individual TDR soil moisture records to a regular grid in time and depth, interpolation and resampling were required. To do so, we generated an intermediate grid of high data density onto which linearly interpolated versions of the time series of each profile were projected. We then resampled from this intermediate grid to derive a synchronized version of the records at 0.1 m depth and 15 min time increments. With this the spatial aggregation remains below the integration length of the TDR probes.

The temporal resampling and the therefore necessary linear interpolation is close to the acquisition timing of one profile (4 min to 10 min each). Since the correlation length of distributed soil moisture observations is rather short and because we explicitly aim to analyze the responses of preferential flow structures, the issue of interpolation needs special attention and will be discussed in Sect. 3.4.2.

All soil moisture measurements are converted to changes in soil moisture referenced to the state previous to irrigation onset to identify activated flow paths.

Lateral interpolation between different TDR profiles over distances of about 1 m and above is unfeasible. Soil moisture as extensive state variable is discontinuous at interfaces. The found subsurface setting does not exhibit any isotropic continuum required for such interpolations.

As in the plot irrigation experiments, vertical response velocities are calculated for the TDR profiles in the core area. The calculation of lateral response velocities is given in the companion study [Angermann et al., 2017, this issue].

GPR transects and structural similarity attribute interpretation

The 2D time-lapse GPR data are derived from nine repeated recordings along the four vertical GPR transects. Each record is processed after a standard processing scheme of bandpass filtering, zero time correction, exponential amplitude preserving scaling, inline fk-filtering, topographic migration with constant velocity ($0.07\,\mathrm{m\,ns^{-1}}$), and consecutive gridding to a 2D transect with regular trace-spacing of 0.02 m.

Most time-lapse GPR data analyses are based on calculating trace-to-trace differences [Birken and Versteeg, 2000, Trinks et al., 2001] or picking and comparison of selected reflection events in the individual time-lapse transects [Allroggen et al., 2015b, Haarder et al., 2011, Truss et al., 2007]. Like in the 3D time-lapse GPR applications, the radargrams in the young, highly heterogeneous soils do not exhibit explicit reflectors as suitable references. In addition, the limited repeatability of the measurements and the desired identification of lateral flow structures require an alternative approach.

Like for the plot-scale experiments, we use the time-lapse structural similarity attribute [Allroggen and Tronicke, 2016, and Sect. 3.2.3]. It is calculated using a local Gaussian window of 2 ns along the vertical axis and 0.06 m along the horizontal axis.

Due to the presence of remaining event water from the preceding storm event [Angermann et al., 2017, this issue], all measurements are referenced to the last acquisition time approximately 23 h after irrigation start and about 19 h after irrigation. The resulting structural similarity attribute images are used as a qualitative indicator of relative deviations from the reference state.

Discriminating the natural storm event and the irrigation experiment

The experiment was preceded by two strong storm events of 43 mm in total on 20 June 2013. In reference to the gauge reaction the experiment was conducted shortly before the second peak of the runoff reaction to the preceding storm events [see Fig. 2.5 in Angermann et al., 2017, for details]. Accordingly, the structural similarity attributes, which compare the distributed states at the respective acquisition time to the last record, identify responses to both drivers, the natural storm event and the irrigation experiment. To discriminate the signals we analyze the dynamics of each pixel in the GPR transects over time. The first two structural similarity attribute transects 7.5 h before and directly at irrigation start are attributed to the natural event and show an increasing structural similarity (towards accordance with the reference state). Once the attribute value of a pixel decreases again (increasing deviation from the reference state) it is attributed to the irrigation. For stability reasons, a threshold of 0.15 was introduced for the attribute to be exceeded to detect changes. This is discussed in Appendix A.6.

3.3 Results

3.3.1 Soil physical exploration

Point samples show high heterogeneity

The *in situ* point measurements of infiltration capacity and saturated hydraulic conductivity showed high variability without clear relationships with simple morphological descriptors like depth, hillslope position or topographic flow gradient (details given in Fig. A.1). Infiltration capacity ranged from $10^{-5}\,\mathrm{m\,s^{-1}}$ to $10^{-3}\,\mathrm{m\,s^{-1}}$. The values for saturated hydraulic conductivity ranged from $10^{-8}\,\mathrm{m\,s^{-1}}$ to $10^{-3}\,\mathrm{m\,s^{-1}}$ and even exceeded the measuring range of the constant head permeameter. Only at the site of the hillslope-scale experiment was a pattern of elevated conductivity at about 0.6 m depth found. The strong heterogeneity and large spread of values was also depicted in the analyses of undisturbed soil samples (Fig. 3.4).

On average the area is dominated by silty soils [see also Juilleret et al., 2011]. This was corroborated by texture analyses and the mean retention characteristics. However, the measurements of saturated hydraulic conductivity are on average 2 orders of magnitude larger than what might be expected for these soils given their

texture [Schaap et al., 2001]. Also, porosity clearly exceeds the expected values, while bulk density is smaller than expected (compare Fig. 3.4, middle row of the panels). All measurements exhibited a large spread of values which does not correlate well with simple morphological variables like depth or hillslope position (Fig. 3.4, bottom row). The high hydraulic conductivity and large porosity is maybe explained by aggregation of fine silty material in conjunction with a network of rapidly draining inter-aggregate pores.

Soil core profile snapshots

The soil core profiles (Fig. 3.5) generally confirmed the presence of the periglacial slope deposits by gravel bands but also showed a high degree of heterogeneity. The thickness of the horizons was variable, with a humidified mineral A-horizon of up to 0.3 m. The gravel content gradually increased over depth in the Bw-horizon and further increased in the C-horizon, starting between 0.4 m and 1.1 m depth. Below the depth of 0.5 m, scattered layers of weathered rock with usually horizontal orientation were found in some soil cores. Percussion drilling was often inhibited at a depth between 1.5 m and 2.0 m (lower end of the bars in Fig. 3.5) due to even higher stone content with a more and more vertical orientation of the weathered rocks. In core 7, concretions of iron and manganese oxides were found at depths between 1.6 m and 1.9 m below ground, indicating hydromorphic conditions.

Based on these standard techniques the overall setting of a heterogeneous silty soil deviating from expected low hydraulic conductivity was revealed. So far gained insight is limited to the general existence of periglacial deposit layers (high gravel content in soil profiles), rapid flow paths (hydraulic conductivity several orders of magnitude above literature references), and some integral retention properties. However, details about its spatial organization and the detection of specific and potentially continuous structures remained obscured by high heterogeneity.

3.3.2 Plot-scale flow path activation and vertical velocities

Irregular patterns of dye stains

In the plot-scale tracer experiments the Brilliant Blue dye stains identified patchy infiltration patterns partially bypassing large sections of the soil without clear traces of the actual flow path (Fig. 3.6). For all experiments stained patches were found

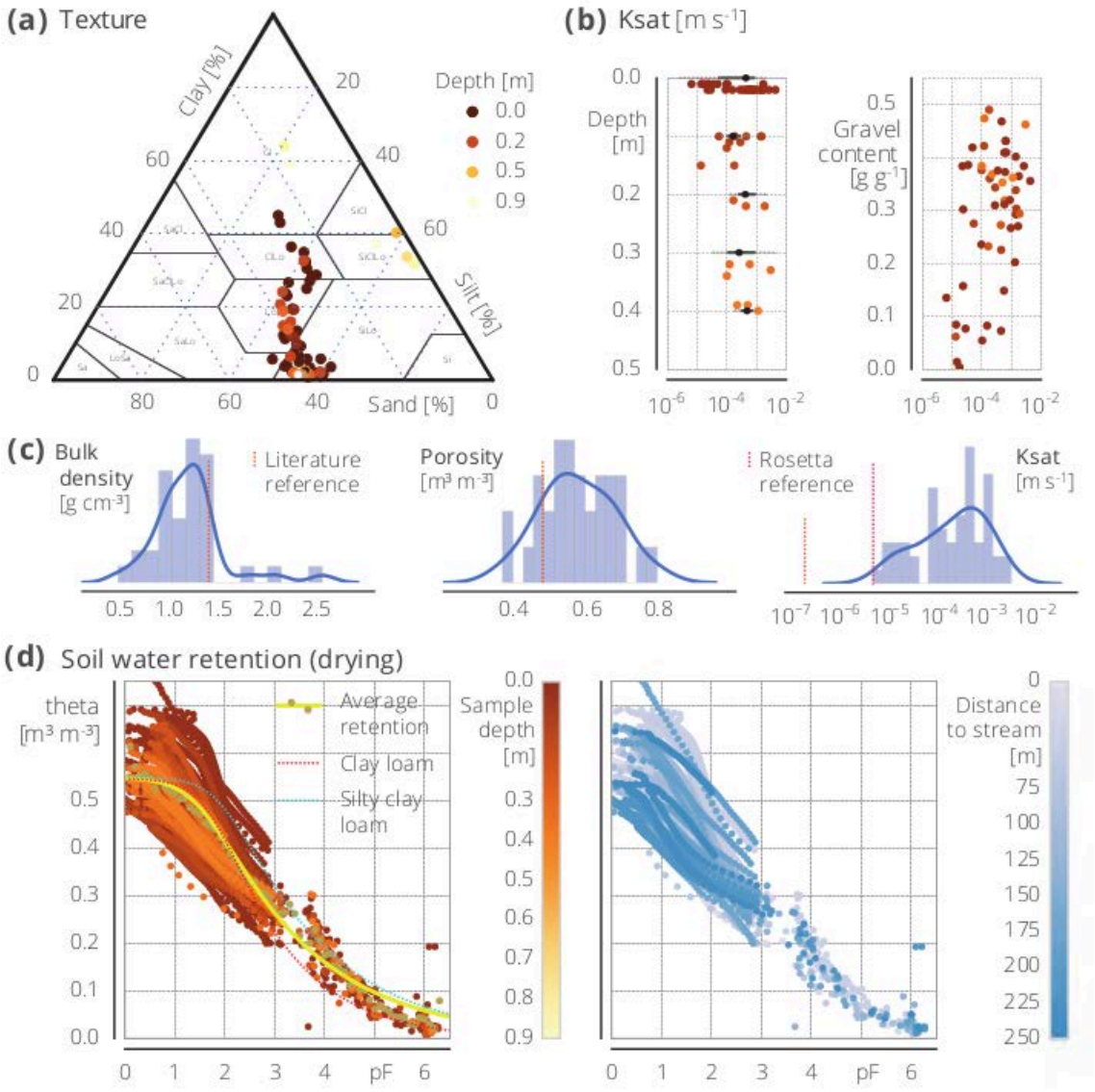

Figure 3.4: Results of laboratory analyses of 63 undisturbed 250 mL ring samples. **(a)** Soil texture analyzed with wet sieving and sedimentation (pipette method). Saturated hydraulic conductivity (Ksat) measured with the Ksat apparatus and plotted against sample depth and gravel content. Black box plots of the respective depth increment. **(b)** Histograms and kernel density estimate of measured bulk density, porosity and Ksat. Reference as the mean silty loam value from the literature [Hillel, 1980, Rawls et al., 1982, Carsel and Parrish, 1988]. Rosetta [Schaap et al., 2001] reference based on mean values of samples (15.7 %, 47.9 % and 36.4 % sand, silt, clay and BD 1.1 g cm^{-3}). **(c)** Soil water retention relation measured with the Hyprop and the WP4C apparatus with the average retention estimate (respective mean of each 0.05 pF-bin and fitted van Genuchten model) and literature references [Carsel and Parrish, 1988] scaled to measured average θ_s and θ_r.

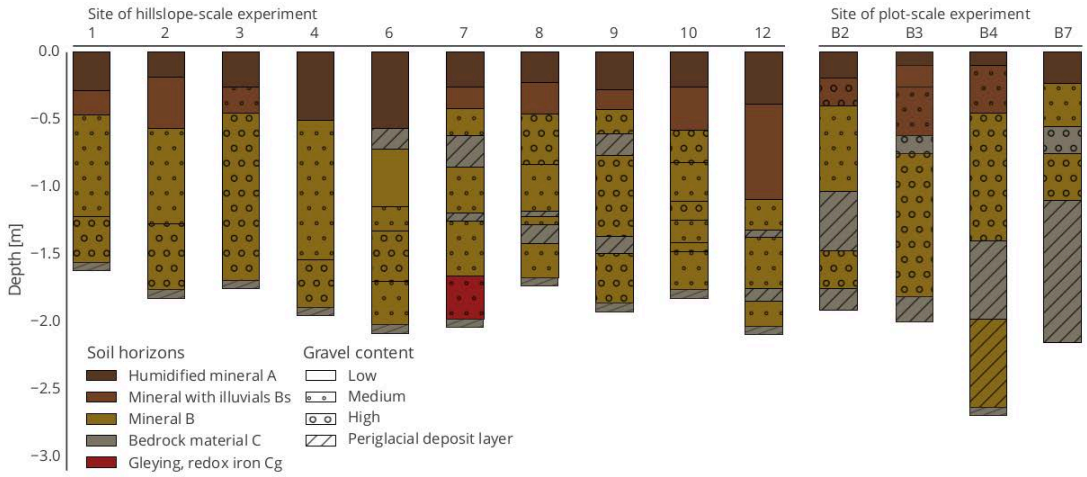

Figure 3.5: Soil core profiles from the upper Colpach River basin. See Figs. 3.3 and A.1 for positions. Bar depth is the maximum drilling depth of the cores restricted by large stones or bedrock.

down to the periglacial deposit layer at 0.6 m to 0.8 m depth. During the excavation apparently isolated dye traces were recovered even several meters downhill from the irrigation spot (4 m downslope, 1 m deep). The stains did not reveal a network of large macropores but an irregular mesh of connected inter-aggregate voids. This is in line with the observed hydraulic capacity (Fig. 3.4).

Bromide breakthrough to the periglacial deposit layer

The connectedness and large transport capacity of this network of inter-aggregate pores is corroborated by the distributions of bromide tracer recovery (Fig. 3.7, top row). All plots suggested a relatively strong response at a depth of approximately 0.6 m. This depth correlates with the upper boundary of the first layer of periglacial deposits found in core profile B3 (Fig. 3.5) and in the excavated soil profiles. This response is contrasted by low bromide concentration at a shallower depth. Even plot XI, where only 30 mm were applied, showed the same pattern with a clear breakthrough to the periglacial deposit layer.

At plot XII we found a stronger interaction with the soil matrix, which led to more dye staining and a higher bromide recovery. Overall, tracer recovery was incomplete (0.45, 0.38, and 0.83 for plots X to XII, respectively) and even declined when including the core samples (0.24, 0.3, 0.63) once more, pointing to strongly irregular soil water redistribution.

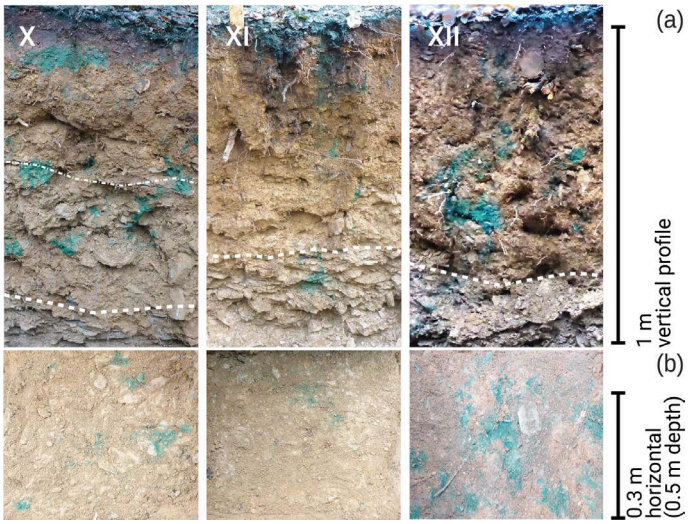

Figure 3.6: Recovered dye patterns in plot irrigation experiments. **(a)** Excavated vertical faces, **(b)** horizontal cuts in 0.5 m depth. Dashed lines indicate level of periglacial deposit layer.

Quick soil moisture response in greater depth

The observed soil moisture changes (Fig. 3.7, bottom row) corroborated the results from the tracer data. Especially at plot X and XII we found a relatively quick and strong response at 0.7 m and 0.5 m depth, respectively. This even preceded soil moisture changes in shallower layers in plot X. Hence the records highlighted an important characteristic of the identified flow-relevant structures, although the signal had a much lower spatial resolution than the tracer results. In contrast to the recovered tracers, we did not observe significant changes in soil moisture in plot XI. This can be explained by its position offset from the main flow field (Fig. A.4 in Appendix A.4).

3D view on soil water redistribution

The structural similarity attribute of the 3D time-lapse GPR measurements provided qualitative information of changes in soil moisture in a spatial context. At all plots the response patterns of low structural similarity pointed out quick vertical flow to a depth of 80 ns or about 1.4 m within 1.5 h after irrigation start (Fig. 3.8, and Figs. A.4 and A.5 in Appendix A.4). Also here, strongest deviations were recorded in the mid horizon between 40 ns and 60 ns two way travel time (TWT) corresponding to approximately 0.7 m to 1 m depth. The top horizon between 20 ns and 40 ns (0.35 m to 0.7 m) had comparably high similarity. Measurements above that depth were technically not possible. Patches of low structural similarity until 20.5 h after irrigation start suggested further lateral redistribution in the later course of the experiment at plot X. At plot XI with 30 mm irrigation further vertical transport

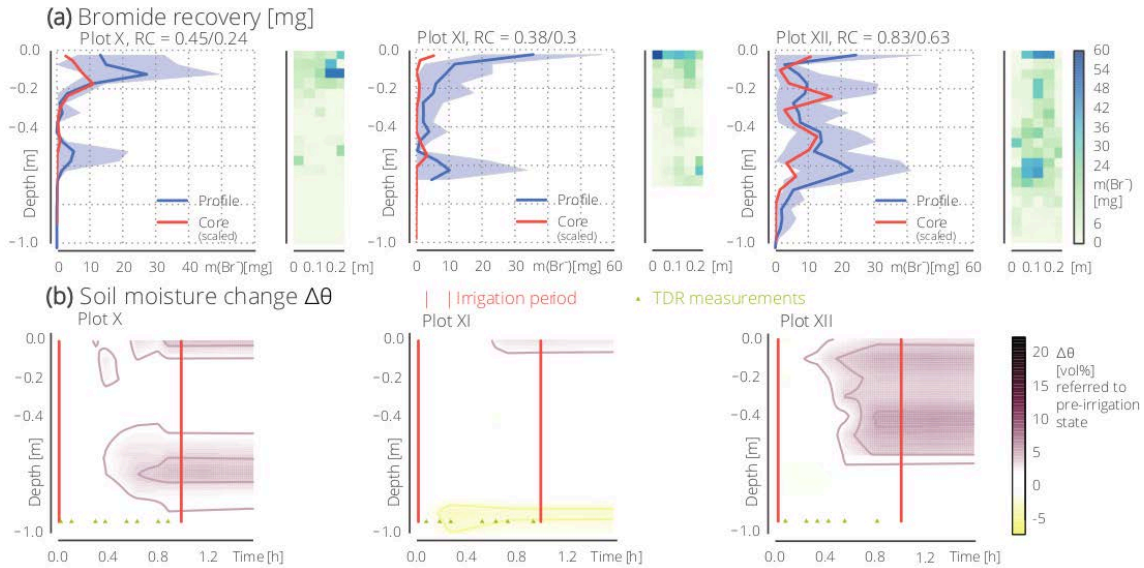

Figure 3.7: Results from plot-scale irrigation experiments with 50 mm, 30 mm and 50 mm spray irrigation for 1 h. **(a)** Recovered bromide mass profiles and grids (5 × 5). Blue line as mean and shaded area between min/max for each depth of the sampling grid. Orange line is the mass recovered in drilled profile samples (scaled to the same volume reference). Recovery coefficient (RC) calculated for the profile samples (first value) and the profile and core samples (second value). **(b)** Observed soil moisture change referenced to the first measurement shortly before onset of the irrigation. Individual measurement times marked with triangles.

with a slight lateral component was recorded. Plot XII had a very high similarity between the first and third acquisitions. This is a sign of stronger macropore–matrix interaction and dispersive redistribution.

The contrasting attribute distributions over time and comparing plots X and XII not only revealed diverse patterns. It also highlighted the qualitative nature of the analytical method of the GPR data. Although visual interpretation of the radargrams (top rows in Figs. 3.8, A.4 and A.5) is very difficult, they show how the structural similarity attribute highlighted areas where radar patterns changed. Due to the complex reflection energy patterns it is not suitable to trace individual reflectors. This prevents a quantitative interpretation as shown by Allroggen et al. [2015b].

For the identification of structures, the results did not exhibit specific macropores like the dye stains, but areas of response to the irrigation. Nevertheless, the patchy characteristic of the found response patterns was very similar to that of Brilliant Blue.

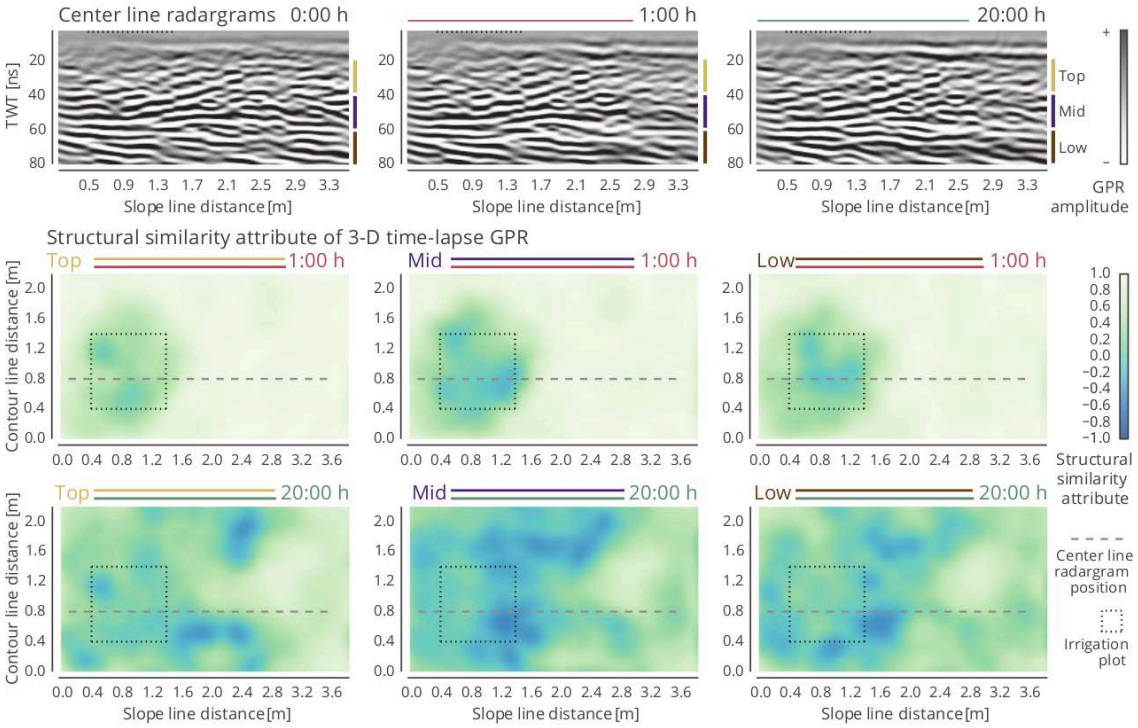

Figure 3.8: Time-lapse 3D GPR of irrigation experiment at plot X. Center line radargrams at the marked transect (gray dashed line in lower panels) for the three acquisition times (before 0:00 h, directly after irrigation 1:00 h, 20:00 h after irrigation) are given in the top row. Two way travel time (TWT) is given as original depth reference. The structural similarity attribute of the 3D data cube is given in three different depth layers (top 20 ns to 40 ns, mid 40 ns to 60 ns, low 60 ns to 80 ns) in the lower panels. The irrigation plot is marked by a black dashed box/line. Slope line distance is increasing downslope.

Derivation of vertical flow velocities

Based on all applied techniques, hydraulic conductivity and apparent vertical flow velocities were calculated (kernel density estimates plotted in Fig. 3.9). The many point-scale measurements (left panel based on 63 ring samples, 40 infiltrometer points, 102 individual permeameter measurements) resulted in disagreeing distributions stretching across a large spectrum of flow velocities. The reason for this spread stems from the fact that the measurements consist of matrix flow and flow in structures. The response-related methods of the irrigation experiments were in much better accordance because they all relate to the same processes. They revealed an apparent vertical velocity of $10^{-3.5}\,\mathrm{m\,s^{-1}}$ (Fig. 3.9b, based on bromide recovery with an estimated time of fixation (t_{fix}) after 1.5 h, first excess of TDR recorded soil moisture $\geq 2\,\mathrm{vol\%}$, and GPR structural similarity attributes below zero between pre-irrigation and the first post-irrigation records).

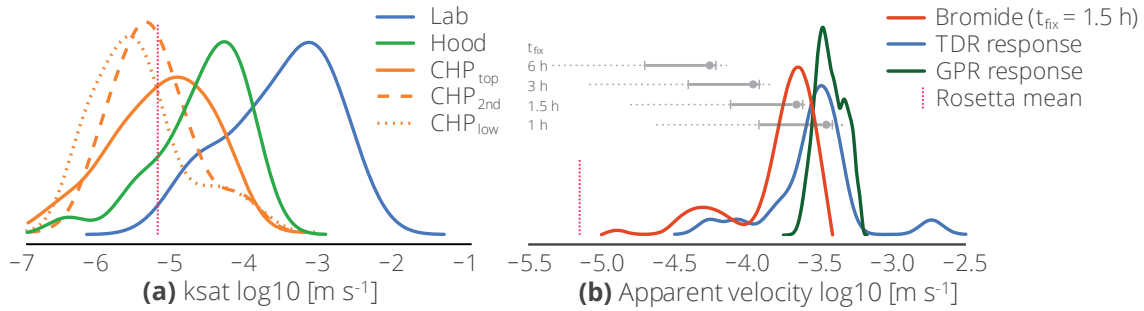

Figure 3.9: Saturated hydraulic conductivity and apparent vertical flow velocity kernel density estimates. **(a)** Point-scale measurement results (Lab: Ksat apparatus; Hood: Hood Tension-Infiltrometer; CHP: Constant Head Permeameter in different depth levels); **(b)** results from plot-scale irrigation experiments. Horizontal gray lines are box plots of velocity distribution based on different t_{fix} estimates for bromide. Rosetta [Schaap et al., 2001] reference based on mean values of ring samples (15.7 %, 47.9 % and 36.4 % sand, silt, clay and BD 1.1 g cm^{-3}).

All results ranged several orders of magnitude above the literature reference of 2.5×10^{-7} m s^{-1} [mean of reported values for silt and silty loam – Hillel, 1980, Rawls et al., 1982, Carsel and Parrish, 1988] and the Rosetta value of 2×10^{-6} m s^{-1} [Schaap et al., 2001]. The role of inter-aggregate pores facilitating this quick redistribution even through comparably small voids without noticeable dye staining was corroborated.

3.3.3 Hillslope-scale detection of lateral flow paths

3D GPR survey suggests a fragmentary layer

The 3D GPR survey at the site of the hillslope experiment identified fragmented structures at about 1.5 m depth (Fig. 3.10). This is in accordance with the soil core profile depth (Fig. 3.5). Especially profile 7 suggested an impermeable layer just below that depth. Although a potential structure can be identified, it remains unclear to which degree this area of high spatial inhomogeneity in terms of radar reflection characteristics is flow-relevant, unless a reaction to an event is observed.

Hillslope responses

The results of the hillslope-scale irrigation experiment can be distinguished into the core area observations with TDR profiles only and observations at the downhill monitoring area, including TDR profiles as well as 2D GPR transects. The change in soil moisture at the core area (TDR 2 and 8 in Fig. 3.11) was very much in

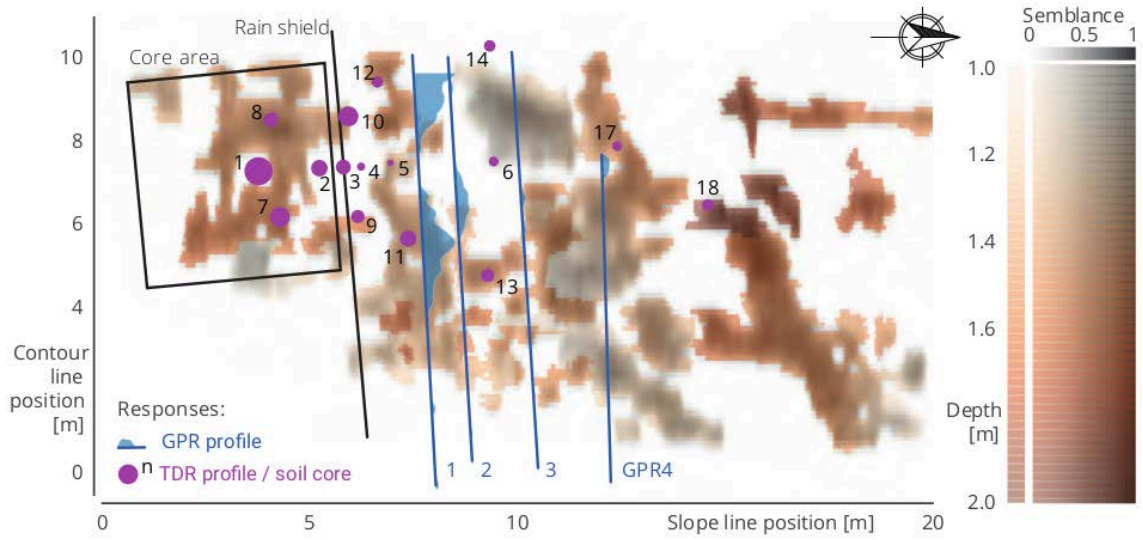

Figure 3.10: Potential subsurface structures from 3D GPR survey and setup of hillslope experiment. Structure identification guided by the dip corrected semblance attribute. Depth estimated based on mean measured effective radar velocity in soil of $0.07\,\mathrm{m\,ns^{-1}}$. Summary of the hillslope experiment given by locations of TDR profile tubes (purple, also location of respective soil cores in Fig. 3.5) and GPR transects (blue). Dot size of TDR scaled to maximum of observed change in soil moisture. Along GPR transects lateral marginals of the structural similarity attribute as proxy for recorded advection. Note that the picked potential subsurface structures are located in different depth (white to black) and that variations in spatial contrast can be seen in the semblance attribute (white to orange). Where more than one horizon has been identified the top one is plotted.

line with the findings from the plot-scale experiments. Given sufficient irrigation, both experiments showed a quick and clear response at greater depth, even before intermediate layers responded. While the patterns were similar, the signal was much stronger during the hillslope-scale experiment, which is due to the higher irrigation amount and duration. The calculated apparent vertical response velocities (lower right panel in Fig. 3.11) had a wider spread towards the faster end but ranged around the values identified in the plot irrigation experiments. The downhill profiles (e.g., TDR 9 and 11 in Fig. 3.11) showed a more diverse response. With greater distance to the core area the reaction was more and more limited to single depth levels. But since the depth levels and responses were highly diverse, it remains rather ambiguous to determine their lateral connection. Overall changes in soil moisture as a maximum at each TDR profile did not corroborate the potential subsurface structures identified in the 3D GPR survey (compare identified potential structures with dot sizes in Fig. 3.10). The full set of profiles is reported in the companion study [Angermann et al., 2017, this issue].

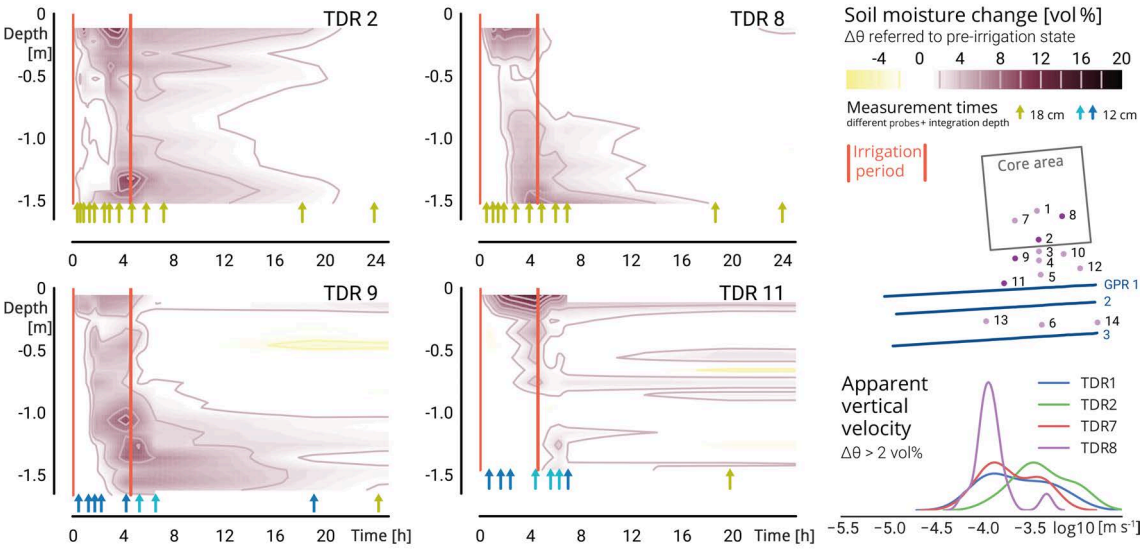

Figure 3.11: Development of soil moisture in TDR profiles during and after the hillslope irrigation experiment.Exemplary transect with changes referenced to pre-irrigation conditions and attributed to irrigation water. Time is given in hours after irrigation start. Individual measurements and probe reference marked with triangles. More data and explanation in Angermann et al. [2017], this issue. Right bottom: apparent vertical flow velocity as first excess of $\Delta\theta \geq 2$ vol% at core area profiles.

The four successive GPR transects across the downhill monitoring area provided spatially distributed images of hillslope-scale flow patterns and boundary fluxes. The structural similarity attribute of storm event water (green) and irrigation water (blue) revealed distinct, heterogeneously distributed patterns (Fig. 3.12) pointing to discrete connected flow paths instead of an irregular network of inter-aggregate pores. The measurements suggested that lateral flow takes place in a very diverse network with very low similarity between the transects. Moreover, the responses to the irrigation decayed with distance to the core area.

The patches which reacted to the storm event are mostly different ones than the structures used to drain the irrigation water. Apparently the irrigation experiment initiated flow in more shallow structures (compare transect 1 irrigation reaction with transect 3 storm water in Fig. 3.12). Areas of high temporal dynamics of the similarity attribute were identified as regions of such flow-relevant structures (Fig. 3.12, bottom row). Note that the last recorded difference 18 h to 23 h after irrigation start (not shown) exhibited high similarity in all profiles. The mean of the attribute was between 0.93 and 0.96 and standard deviation between 0.076 and 0.048 for GPR transects 1 and 3, respectively. Apparently, the system had reached

a steady state without much further change in soil moisture (see Appendix A.6 for more details).

The patchy structures at the transects highlighted the irregularly distributed nature of lateral preferential flow paths which was similarly observed in the plot experiments. Although some areas exert a higher density of reacting flow paths than others, no continuous patterns could be specified throughout the hillslope. We also saw a decay of the signal strength and areal share with distance from the core area. As the patterns from transect 1 did not simply propagate further downslope, the flow paths must be tortuous and leaky. Hence inferring the configuration of the connection between the four transects in the downhill direction is not feasible. A comparison of the suggested structures of the 3D GPR survey to the overall response to irrigation recorded at the GPR transects did not correlate well (compare identified potential structures with the reaction summary at GPR transects in Fig. 3.10).

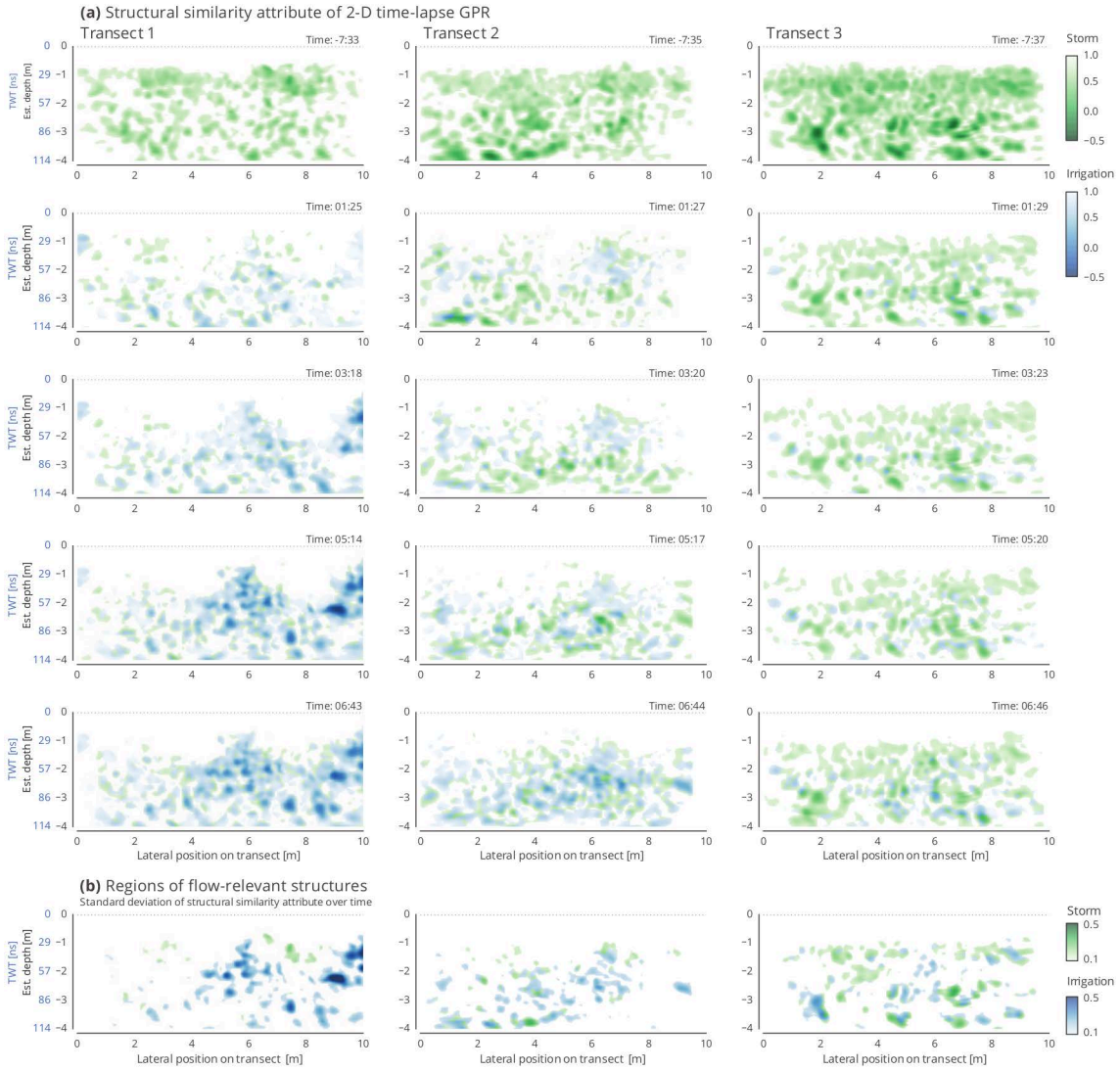

Figure 3.12: Structural similarity attribute in time-lapse 2D GPR transects. Blue: irrigation event water; green: storm event water. Columns: time series in one transect; rows: different transects at the same time. **(b)** Identified regions of rapid subsurface flow based on the standard deviation of all structural similarity attributes at one transect over time. Note: the structural similarity attribute calculates similarity between the radargram at the respective time to the last record 23 h past irrigation. A threshold of 0.15 is applied to identify significant changes. It is a qualitative measure based on the assumption that the last record is in steady state and that all differences are induced by soil water redistribution.

3.4 Discussion

3.4.1 Identification of flow-relevant structures across scales

Our results have shown that the silty soils coincide with high porosities and high hydraulic conductivity at the point scale. Such a coincidence is not what is expected for cohesive, fine textured soils and can be explained by a setting of aggregated fine material in conjunction with a network of inter-aggregate pores. With respect to our research question Q1, the pedo-physical analyses initiated the recognition of these sub-scale structures. However, neither their position nor their general setup can be identified based on point observations because of its scale below the support of the measurements. *Vice versa*, methods at the next scale do not provide information about porosity and bulk density.

Irrigation experiments at the plot scale visualized that a network of these inter-aggregate voids connects the surface to the periglacial deposit layer and is responsible for highly diverse soil water redistribution. These structures are different from what we usually expect (cracks, worm burrows, roots channels) at this scale. This could be depicted from dye tracer stains (Fig. 3.6), which still have the highest spatial resolution on the cost of a lack of temporal insight. It requires strong assumptions about macropore–matrix interaction, time of fixation and dye supply, retention and recoverability. Despite all uncertainty about what process caused staining, the technique allows to identify the structures activated by irrigation and to infer much about their setting where dye has been retained. Although dye stains are closely related to actual flow and thus function, they only reveal the potential pathways and thus form as the actual processes and timing remain unknown. When irrigation intensity and irrigation amount ranges near the hydraulic capacity of the macropore network while still avoiding ponding or macropore clogging, the entire network of flow-relevant structures is marked. 3D time-lapse GPR has proven to be capable to detect similar response patterns. However, the spatial and temporal resolution of the method is still insufficient to detect the flow-relevant inter-aggregate voids marked by dye stains. Some of the structures have not even been traced with dye, nor could GPR identify them. Nevertheless, the overall characteristics of the structures as patchy responses are depicted well and in a non-invasive, spatially continuous manner. Thus most of the point-sampling related issues (Sects. 3.3.1 and 3.4.4) are resolved. Regarding research question Q2, the visualization of flow structures based on responses to irrigation succeeded at the plot scale. They are in good coherence

with the quantitative findings from salt tracers, stable isotopes and soil moisture dynamics. Interestingly, the found vertical response velocity distributions correspond well to the saturated hydraulic conductivity measurements in soil samples, although their distribution is much more tight.

At the hillslope scale (Q3), applications of 3D time-lapse GPR are technically impossible due to the long acquisition times. Consequently we altered the setup to four trench-like 2D time-lapse GPR profiles to facilitate the required high temporal resolution. The responses suggest structures similar to but less diverse than the found inter-aggregate voids at the plot scale. They are spatially persistent and leaky and apparently feed from diverse sources. As such the irrigation experiment caused a similar response in different structures than the previous storm event. Moreover, the relatively high input rates have proven adequately chosen to identify lateral subsurface flow paths. At this scale the capability of point-based methods for structure identification is even more limited as the dense network of soil moisture profile observations did not allow the derivation of a conclusive picture.

3.4.2 Event response patterns reveal flow-relevant structures

Interestingly, static methods failed to unravel structures from overall heterogeneity. This corroborates our idea that responses to an event are required for the identification of flow-relevant structures. Furthermore, it confirms that a combined assessment of form and function is needed to mutually reduce ambiguity. This is also shown in the companion study with a focus on function and processes at the hillslope scale [Angermann et al., 2017, this issue].

Soil moisture responses

In our case TDR measurements through access tubes were employed as low-impact means to monitor soil water dynamics in order to detect areas of quick and strong response. Structures in general and the inter-aggregate voids in our case cover only a very small fraction of the measured volume. We may underestimate detected flow paths when they do not alter the total volumetric soil water content much (bypassing). This can explain the observed patterns of low response in the topsoil and changes in regions where the fast flow is decelerated at some kind of bottleneck. Referring to the theoretical integration volume of 1 L, it would require a macropore

of about 1 cm diameter within the support of the sensor to be filled to just reach a threshold of 2 vol%. Adding this 20 mL of water diffusively would result in the same measurement. This shows that soil moisture measurements exhibit a conceptual bias towards the diffusive fraction of the soil water.

The quantification of advective water from the recorded changes in soil moisture has been proven as not feasible. Given the insight of the discretely structured flow domain and the high lateral response velocities identified in the companion study [Angermann et al., 2017, this issue], the soil moisture measurements leave us with many questions. Comparing the identified regions of flow structures (Fig. 3.12) with the support of a TDR sensor quickly reveals that even a large number of point observations remains highly uncertain if the overall spatial context is unknown. This is especially the case at the hillslope scale. At the plot scale, the issue is less pronounced, as we have shown with the good correspondence between tracers, GPR and soil moisture reaction at plot X (Figs. 3.7 and 3.8). However, at plot XI with less intense irrigation, the soil moisture profile did not react despite the evidence of quick vertical redistribution in all the other methods. Apparently, the TDR records were simply not close enough to the relatively few activated flow paths (Fig. A.4).

Time-lapse GPR patterns

The potential horizons identified by the static 3D GPR survey do not coincide with the observed responses (Fig. 3.10). This is another example for the requirement of a shift between active and non-active state to identify flow-relevant structures. The structural similarity attributes derived from time-lapse GPR reveal the patterns of soil water redistribution. The continuous 2D and 3D data allow to relate temporal changes in space as images of the subsurface as proposed by Gerke et al. [2010], Beven and Germann [2013] and others.

The comparison of radargrams in time needs further attention: In other time-lapse GPR applications for soil water dynamics in structured domains [Truss et al., 2007, Haarder et al., 2011, Allroggen et al., 2015b, Klenk et al., 2015] analysis is guided by reference to a reflector and its apparent displacement can be used to calculate changes in soil moisture. Alternatively, a wetting front could generate a reflector [Léger et al., 2014]. In our case none of these existed.

On the one hand, we minimized methodological problems concerning the noise arising from the imperfect positioning of repeated GPR measurements by using a measuring platform at the plot scale, transect guides at the hillslope scale, and

an automatic-tracking total station [Allroggen et al., 2015b]. On the other hand, we base our analysis on the structural similarity attribute instead of pixel-to-pixel comparison or picked reflectors. The disadvantage is that this allows only for a qualitative measure. The advantage is that the spatial organization of areas with changing reflection and transmission properties (which are attributed to changes in soil moisture) can be revealed even in heterogeneous soils. The 3D applications at the plot scale avoid strong assumptions about the continuity of preferential flow paths when inferring 3D networks from 2D measurements [Gormally et al., 2011, Guo et al., 2014]. Especially in the soils under study, the found response patterns (Fig. 3.8) and the excavated stained soil profiles (Fig. 3.6) show highly tortuous flow paths. Thus we refrain from interpolations between the multi-2D transects at the hillslope scale.

Although the 3D time-lapse attribute data of the plot irrigation experiments are of low spatial resolution (blur due to similarity attribute method and long duration of one acquisition) and limited temporal resolution (few acquisition times), they are suitable to identify regions of flow-relevant structures and their characteristics. In the multi-2D transects resolution was enhanced (short duration of one acquisition and many repeated measurements) which depicted the structures much better. Hence, time-lapse GPR can especially be improved by enhancing the acquisition time and frequency.

The observation of changes during activation of flow-relevant structures generated the required contrast to overall heterogeneity. For large structures, this led to precise identification and localization. Smaller flow paths cannot be fully resolved. Nevertheless, the continuous 2D and 3D images of the subsurface response patterns provide means to non-invasively study the form–function relationship *in situ* and to overcome some of the restrictions of retrospective and destructive tracer methods. However, quantitative interpretation of time-lapse GPR data remains challenging.

3.4.3 Methodological assessment

In contrary to our first expectation, the value of pedo-physical analyses of soil core samples has been relatively high even for characteristics of flow facilitated by the revealed paths at larger scales. Structure identification is not only obscured in heterogeneity as one would expect, but properties deviating from the standard situation (fine texture, low bulk density and high porosity) gave rise to the identification of the inter-aggregate flow paths. However, the spatial organization of structures below

and above the support of the samples cannot be revealed. This is also the reason for the relatively low information which could be drawn from the *in situ* infiltration measurements: The observed flow rates are largely affected by the capacity of the connected flow paths draining the measurement point. This adds to the critical assumption of homogeneity [Langhans et al., 2011].

Besides the high information gain through the state shift of flow-relevant structures in irrigation experiments, the employed methods at the plot scale have very specific advantages and disadvantages: Especially the laborious and costly analysis of salt tracers and stable isotopes is contrasted by relatively little additional information. Moreover, the lack of a temporal information about when the solute or water molecule was retained in a certain depth is seen problematic. Soil moisture profile dynamics and time-lapse GPR do not suffer this drawback. Both can be employed with very low or even no impact on the subsurface from the surface. While GPR requires to be operated in higher temporal resolution (see Sect. 3.4.5), soil moisture profiles lack the desired spatial discretization. Dye staining delivers the highest spatial resolution to reveal subsurface structures on the cost of blindness about the temporal dynamics. Furthermore, a tomographic excavation of the stains has proven very difficult.

Under strongly structured conditions as at the hillslope under study, point observations remain a needle in a haystack. Unlike for vertical structures at the plot scale, the dense network of soil moisture profiles could not depict the lateral flow paths well. Here, the *GPR-inferred trenches* have shown to be a valuable surrogate for massive staining like in the study by Anderson et al. [2009]. In addition, the temporal dynamics of the hillslope reaction could be observed.

With regard to our *a priori* model application, the combination of vertical and lateral flow paths (identified in the irrigation experiments) with layers of low permeability just below the structures (observed in the soil core profiles) could refine the domain towards more lateral soil water transport. The mean retention properties (derived from pedo-physical analyses) are adequate. Hence, the combination of data from all scales can contribute to a refinement of the model.

3.4.4 Heterogeneity versus structure

Based on numerous point-scale measurements, the overall layering and mean property of a heterogeneous soil with periglacial deposit layers were described in Sect. 3.3.1.

Given the large effort to conduct such an exploration, this result appears rather unsatisfactory and could have been achieved with much simpler means [e.g., compare Heller and Kleber, 2016, in a similar setting]. However, they have been key to the identification of the sub-scale inter-aggregate structures which convey to vertical drainage paths and a lateral network in the subsurface. Without high supply rates from the point scale, subsurface storm flow in the lateral structures could not be sustained.

At the hillslope scale, the attribute supported picking of potential structures in the 3D GPR data cube also had high discrepancies from the actual relevant structures (see the differences between potential subsurface structures and recorded reaction in the TDR and GPR profiles in Fig. 3.10), which is in contrast to similar GPR applications by Gormally et al. [2011] and explains the large spread in the results of the hydrological measurements.

It has proven particularly difficult to distinguish heterogeneity and structure. This has conceptual implications: as introduced, we regard statistical heterogeneity as random small-scale changes in hydrological soil properties [De Marsily et al., 2005] and structure as spatially organized flow paths [Gerke, 2012] and their connectedness [Tetzlaff et al., 2010] or persistent spatial covariance of high conductivity values. Hence the structures require multivariate or topological characterization. To infer on the directed flow in subsurface structures, a spatially continuous observation of the reaction to an event is required. This has proven especially challenging as the spatial scale of flow-relevant structures (5×10^{-3} m to 5×10^{-2} m) is several orders of magnitude below the support of standard soil samples and hydrological measurements (10^{-4} m^3 to 10^{-3} m^3).

In more general terms, heterogeneity can be seen as deviation of the found reality from the concept of quasi-homogeneous elementary volumes. If this deviation concerns only the apparent parameters of the same physical process, more samples are adequate to determine their distribution. In cases (like here) where this deviation also means a shift in the physical processes, heterogeneity may introduce bias as it becomes a scale problem: any measurement will consist of an unknown subset of connected or non-connected flow paths. This makes it impossible to unravel the properties of the different flow domains without knowing the composition of the explored ensemble of each measurement. Hence the point-based techniques cannot determine the super-scale organization outside the support of the measurement.

Without detection of organization and thus flow-relevant structures, they can only recover heterogeneity independently of the number of samples.

3.4.5 Outlook on structure identification with time-lapse GPR

In the context of preferential flow studies in watersheds around the globe and in many different models, our results open new ways to visualize subsurface flow and to facilitate more field studies to understand stormflow generation [as recently found in a meta-analysis by Barthold and Woods, 2015]. Although we cannot fully resolve macropores as needed in spatially explicit representation of macropores as vertically and laterally connected flow paths [Vogel et al., 2005, Sander and Gerke, 2009, Klaus and Zehe, 2011], our findings provide the experimental basis to further develop such models. More implicit approaches like stochastic stream tubes [Jury et al., 1990], the scale way idea [Vogel and Roth, 2003], or dual porosity and permeability approaches [Gerke, 2006] could be extended by providing spatial and temporal context, which is one of their assumptions. Also, more integrating concepts like the representative watersheds [REW, Reggiani et al., 1998, Tian et al., 2006, Lee et al., 2007] could define zones for quick drainage based on repeated response observations in vertical and lateral structures.

In the form and function framework one implication of the study is that a disjunct analysis of the two is a source of unnecessary ambiguity and susceptibility to bias. Although the conjugated nature of form and function is very much in line with the general findings and perception of early studies [Aristotle in Blits, 1999, Thompson, 1942, Wittgenstein, 1922, and others], it contradicts a general notion in hydrological surveying and modeling to separate the two. In most models different flow paths are defined in a lumped manner using effective parameters after all. These domains and their parameters could be determined based on irrigation experiments and time-lapse GPR measurements.

While models require specific parameters about the site under study which are coherent with their conceptual assumptions or modeler's perception [Holländer et al., 2014], the experiments are also strongly shaped by our perceptual model about the processes. With this, the matter of model adequacy is not restricted to numerical aspects alone [Gupta et al., 2012]. Methodologically, the *in situ* imaging of subsurface flow processes can be used to reduce ambiguity of measurements and to constrain

the process conceptualization in heterogeneous and structured soils. In our case, the *a priori* model overestimated deep percolation and underestimated the velocity of lateral soil water redistribution through subsurface flow paths. Based on field information about the overall distribution of flow paths or quickly reacting areas, sampling and monitoring could be guided. This would reduce the limitations of point-scale methods with relatively little effort.

3.5 Conclusions

In the hillslopes under study, silty, cohesive soils coincide with high porosity and high flow velocities at the Darcy scale. This motivated in depth investigation of flow-relevant structures explaining this. We have shown that subsurface heterogeneity and the mismatch of observation and process scales obscured the identification of flow-relevant structures under static conditions without a shift between active and non-active states. The pedo-physical analyses initiated the recognition of these sub-scale inter-aggregate structures. The point-scale exploratory methods could quantify the general characteristics of the subsurface only within a wide spectrum of the respective target properties. However, they failed to identify flow-relevant structures in terms of position, distribution and capacity at larger scales. Measurements of infiltration capacity and hydraulic conductivity require special attention, because they integrate over an unknown set of advective and diffusive flow paths. The discrepancy between results from the soil core profiles and a 3D GPR survey on the one hand and the time-lapse approaches on the other hand indicates that structures identified from inhomogeneities are not necessarily flow-relevant pathways.

Joint application of tracers and time-lapse GPR during irrigation experiments revealed details about the structures and their activation by flow. At the plot scale a network of inter-aggregate pores enables fast soil water redistribution in a less directed manner and at much finer scales than usually expected in macropores like cracks, worm burrows or root channels. This facilitates high apparent vertical flow velocities ranging around $10^{-3.5}\,\mathrm{m\,s^{-1}}$, while operating in fine pores at scales very difficult to identify even with dye staining. The combination of tracer and time-lapse GPR methods enabled the more holistic view into the subsurface which was further applied to the hillslope scale. There persistent lateral pathways connecting along the hillslope have been identified through *GPR-inferred trenches*.

Our findings show that form and function in hydrological systems operate in conjugated pairs. This implies that it is very difficult to observe them separately and that their projections are inherently non-unique and scale-dependent. Besides the fine scale of the inter-aggregate voids, form needs to be addressed in its context to reveal information about its structure and characteristics, but addressing function also needs details about the spatial circumstances to be conclusive. Overly strong assumptions about structures or processes can be avoided by the presented non-invasive time-lapse GPR method, which can visualize and localize response patterns at the plot and hillslope scale. They compare well with soil moisture dynamics and tracer recovery. As such the localization of responses provides the missing link to relate form to function [taken up in the companion study by Angermann et al., 2017, this issue] and to guide more specific investigation, monitoring and modeling of subsurface processes.

All data used in this study is foreseen to be openly published in Earth System Science Data (ESSD) as concise outcome from the research project. Until then they are available from the authors on request.

Acknowledgements Thanks to Elly Karle and the Engler-Bunte-Institute, KIT, for the IC measurements of bromide. We are grateful to Selina Baldauf, Marcel Delock, Razije Fiden, Barbara Herbstritt, Lisei Köhn, Jonas Lanz, Francois Nyobeu, Marvin Reich and Begoña Lorente Sistiaga for their support in the lab and during fieldwork, as well as Markus Morgner and Jean Francois Iffly for technical support and Britta Kattenstroth for hydrometeorological data acquisition. Laurent Pfister and Jean-Francois Iffly from the Luxembourg Institute of Science and Technology (LIST) are acknowledged for organizing the permissions for the experiments. Moreover, we thank Markus Weiler (University of Freiburg) for his strong support during the planning of the hillslope experiment and the preparation of the manuscript. This study is part of the DFG-funded CAOS project "From Catchments as Organised Systems to Models based on Dynamic Functional Units" (FOR 1598). The manuscript was substantially improved based on the critical and constructive comments of the anonymous reviewers, Christian Stamm and Alexander Zimmermann, and the editor Ross Woods during the open review process, which is highly appreciated.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

Chapter 4

Hillslope-stream connectivity

The competing impact of stream morphology and catchment topography on hillslope-stream connectivity

Lisa Angermann, Markus Weiler, Theresa Blume

This article has not been published or submitted yet.

Abstract

Hydrological connectivity summarizes all processes of water mediated transport within or between elements of a landscape. As such, it is key to understanding internal processes and the behavior of catchments. Structural properties have been acknowledged to be an important control on catchment-scale connectivity. However, structural controls may act at multiple scales and their relative importance can be difficult to assess. This study aims at evaluating the role of local stream channel morphology and catchment-scale topography for the connectivity between hillslopes and streams. To do so, discharge patterns along the stream of a small headwater catchment were measured at multiple sales and interpreted as the fingerprint of the hierarchical organization of structures in the riparian zone and the catchment. Discharge patterns were measured by sequential salt dilution measurements along the stream at different scales in combination with radon sampling as tracer for groundwater inflow to the stream. These measurements were repeated under different discharge conditions and the resulting patterns were correlated with topographic and morphological measures. The results showed that small-scale discharge patterns were controlled by stream channel morphology at all times, with gains being almost exclusively concentrated at steps in the longitudinal streambed profile. At larger scales, the patterns changed with discharge conditions which indicates different controls with changing system states. During summer base flow conditions, channel morphology was the main control for the connectivity between riparian groundwater and the stream. With increasing discharge under winter base flow conditions, the lateral contributing area was a better predictor for discharge patterns along the stream. Under medium and high flow conditions, discharge signals of the lateral contributing area appeared to be shifted downstream, which indicated the activation of parafluvial flow through preferential flow paths within the riparian zone. Thus, reach-scale connectivity was controlled by topography and discharge conditions, while small-scale spatial patterns of connectivity were modulated by structures in the riparian zone and the stream channel.

4.1 Introduction

4.1.1 Background and hypotheses

Considering surface water and groundwater as one continuum is widely accepted and influences research and management practices alike. In this context, the concept of hydrological connectivity summarizes all flow and transport processes from hillslopes to streams [Bracken et al., 2013]. While often applied in order to address internal catchment processes and their spatial or functional characteristics, it also provides the conceptual link between terrestrial processes on the one hand and stream behavior and groundwater discharge patterns on the other [Blume and van Meerveld, 2015]. The spatial organization of flow paths at the hillslopes and in the riparian zone controls discharge patterns in the stream and with that ultimately spatial patterns of hillslope-stream connectivity.

Various structures in the catchment are crucial for the spatial organization of flow paths. These structures include topography [Beven and Kirkby, 1979, McGuire et al., 2005] and hillslope morphology [Anderson and Burt, 1978, Bogaart and Troch, 2006, Jencso et al., 2009], subsurface characteristics, such as bedrock topography [e.g Freer et al., 2002, Tromp-Van Meerveld and McDonnell, 2006], periglacial slope deposits [Angermann et al., 2017], soil structures [Wienhofer et al., 2009], and macropores [McDonnell, 1990]. Some of these (e.g., soil structures and macropores) can be strongly linked to ecological factors, such as vegetation cover and land use [Nippgen et al., 2015], or control the physical characteristics of the soil, like saturated hydraulic conductivity and water retention characteristics [Van Schaik et al., 2014].

While the stable structural properties of a catchment are widely acknowledged to control the contribution of water from the hillslopes and groundwater flow towards the stream, geomorphological properties of the riparian zone and the streambed can control the interaction between the stream and its immediate surrounding [Harvey and Bencala, 1993]. However, the role of local-scale (below 10 m) morphological control on flow processes is most extensively studied in the field of hyporheic hydrology and often has a strong focus on in-stream or in-sediment processes and water quality.

Studies on hillslope-stream connectivity, usually focus on scales of a couple of 100 m or more. Hyporheic flow processes are deemed to be irrelevant for the larger picture of hillslope-stream connectivity, leading to a lack of knowledge on the bi-directional interactions between riparian and hyporheic flow processes [Ward et al., 2012]. The concept of parafluvial flow intends to close this gap, describing flow

paths "parallel to the stream" in the riparian zone, connecting upstream losing and downstream gaining stream reaches over distances between meters and several 100 m [Cartwright and Hofmann, 2016] However, this concept is usually only applied in the context of a stream-riparian zone continuum rather than of the whole catchment. Located "in between" the catchment and the stream, streambeds and riparian zones have been shown to strongly affect the chemical composition of water routed from the hillslopes towards the stream [Hooper et al., 1998, Weiler and McDonnell, 2006]. At the same time, they have the potential to also alter or interrupt flow processes that connect the catchment and its stream [Burt, 2005]. This potential is especially large in headwaters with shallow and highly structured soils and morphologically diverse streams, where preferential flow in the riparian zone and the streambed might superimpose larger-scale flow patterns.

Investigating these processes of hillslope-stream connectivity requires covering both, catchment processes and hyporheic or parafluvial flow. However, exchange processes at the scale between 10 m to 1 km are "inherently messy" [Bencala et al., 2011] and hard to assess. The here investigated system is large enough to be subject to spatial variability of catchment processes, yet small enough to be influenced by hyporheic and riparian processes [Bencala et al., 2011]. Under such conditions, experimental approaches are challenged to provide a sufficiently high spatial resolution, while at the same time cover large-scale patterns as well.

With respect to the spatial organization of the catchment (i.e., the structurally controlled, spatial characteristics of flow paths through the catchment) as major control of hillslope-stream connectivity, the ultimate measure is the signal within the stream. Spatial patterns of groundwater discharge are the "fingerprint" of connectivity [Blume and van Meerveld, 2015] which integrate over all processes in the catchment and the riparian zone.

The analysis of hydrographs and environmental tracer dynamics is a commonly used approach to investigating catchment processes via transit time distributions [e.g., McGuire et al., 2005] and source area connectivity [e.g., Tetzlaff et al., 2007]. Similarly, groundwater discharge patterns contain information about the spatial organization of flow paths within the catchment and the connectivity between hillslopes and stream. These groundwater discharge patterns can be used to answer the question of how different spatial structures at multiple scales (such as topography, soil and subsurface structures at the hillslopes and in the riparian zone, streambed

morphology, etc.) influence the hydrological processes within the catchment and eventually catchment behavior.

The approach is limited by the potentially very complex interplay of processes and the repeating superposition of signals from different structural features and flow processes. The quality of information thus depends on the structural complexity and temporal dynamics of flow processes in the catchment and riparian zone and in certain cases might be limited to the signal generated in the riparian zone and the streambed. Nevertheless, this is valuable information for the mechanistic understanding of hillslope-stream connectivity.

In order to identify the relevant patterns, the spatial resolution of the observations has to be adapted to the revalent structures in the catchment. Thus, a nested approach is required, which is capable of capturing both types of patterns: those induced by small-scale stream or near-stream morphology and those induced by large-scale topography.

To assess groundwater discharge patterns in a nested approach, a flexible methodology is needed to measure groundwater inflow along stream sections of varying length and heterogeneity. Salt dilution gauging is the most suitable method to measure discharge in small and heterogeneous streams, and allows for calculating absolute gains and losses at the scale of stream reaches of 50 m length or less [Payn et al., 2009].

The radioactive noble gas radon (here referring only to its most stable isotope ^{222}Rn, $t_{\frac{1}{2}} = 3.8\,\mathrm{d}$) has been used for the investigation of surface water-groundwater interaction in various settings and research questions. It is a product of the radioactive decay series of uranium and is enriched in water traveling through the subsurface, with equilibrium concentrations being reached after approximately 25 d. Its low solubility leads to quick degassing and low concentrations in surface water [Porcelli and Swarzenski, 2003]. The resulting pronounced differences in concentrations of surface water and groundwater and its inert nature make radon a powerful tracer for groundwater-surface water interactions.

Applications of radon as environmental tracer include the investigation of groundwater discharge into coastal areas [e.g., Kim and Hwang, 2002, Burnett and Dulaiova, 2003] and lakes [e.g., Schmidt et al., 2010, Dimova et al., 2013]. In streams it was used to quantify groundwater inflow into lowland rivers [e.g., Cook et al., 2003, Mullinger et al., 2007] or surface water infiltration into the banks or the groundwater body [e.g., Hoehn and Von Gunten, 1989, Stellato et al., 2013], or to study

the bi-directional interactions between surface water body and groundwater [e.g., Ellins et al., 1990, Stellato et al., 2008]. Most of these studies focus on water bodies with comparably large volumes and low internal heterogeneity, such as lakes and lowland rivers. The application of radon as an environmental groundwater tracer in small and turbulent streams however, is less common [e.g., Genereux and Hemond, 1990]. The potentially high variability in turbulence and gains along small headwaters poses additional challenges. However, the distinct concentrations in surface water and groundwater and its comparably fast and cheap measurement (also *in situ*) make it a promising tracer.

4.1.2 Hypotheses

Based on these elaborations, we formulated the following hypotheses:

H1 In addition to the spatial structures of the catchment stream morphology has a strong impact on hillslope-stream connectivity in highly structured headwater catchments.

H2 Multi-scale patterns of hillslope-stream connectivity can be used to decipher the hierarchical interplay of spatial structures within the catchment, especially with focus on the relative importance of local-scale streambed and riparian zone morphology *versus* catchment-scale topography for runoff generation.

H3 The combination of radon as an environmental groundwater tracer and differential salt dilution gauging is a useful approach to investigate patterns of groundwater-surface water interactions in small streams at multiple scales.

4.1.3 Approach

The hypotheses were tested in a small headwater catchment in Luxembourg. We measured absolute gains and losses along in the headwater streams every 50 to match major topographic features of the catchment. Additional measurements of absolute gains and losses were taken every 2.5 m to 15 m, based on gemorphological features along the stream. These measurements were accompanied by water sampling for radon and an artificial tracer study, applying sulfur hexafluoride (SF_6) to quantify radon loss to the atmosphere. These data were used to quantify the spatial variability of gains and losses during four campaigns exemplary of different discharge

conditions. We furthermore calculated the mixing ratio between new and old water of the quantified gains and discussed the uncertainties in using radon as tracer for groundwater inflow to streams. The resulting groundwater discharge patterns were compared and correlated with mapped geomorphological features (1 m to 10 m scale) along the headwater and catchment topography (50 m scale).

4.2 Methods

4.2.1 Study site

The Holtz catchment is a sub-catchment of the Attert experimental basin. It is located at the south eastern edge of the Ardennes Massif in western Luxembourg. Strongly inclined Devonian schist bedrock is the foundation of the characteristic landscape in this area [Van den Bos et al., 2006]. The young soils are relatively shallow and developed under periglacial conditions, when weathering and solifluction led to the development of a saprolite layer or periglacial deposit layer [Juilleret et al., 2011], which can be found at almost all hillslopes in the area. The Holtz catchment is a narrow headwater with relatively steep forested slopes (15° to 25°) and agriculturally used plateaus. Except for some wetland patches in the proximity to the springs, there is no significant floodplain. The forest at the hillslope is dominated by spruce (*Picea abies*), beech (*Fagus sylvatica*) or mixed deciduous forest.

The stream is approximately 600 m long and shows a straight or mildly meandering course. Its morphology is characterized by step-pool sequences and several braided sections. At some sections the stream is deeply incised with stream banks of up to 1.4 m height, revealing the loamy substrate and the saprolite deposits of the underlying bedrock.

Two branches (north and south) make up the spring area and merge approximately 550 m from the outlet. Both springs are located in swampy areas, and the exact location varies according to hydro-meteorological conditions. The northern branch originates on a pasture and has a diffuse source area. The spring of the southern branch is more clearly defined. One intermittent tributary is activated only occasionally during strong rainfall events and joins the main stream approximately 400 m from its outlet.

Two V-notch gauges were installed for continuous discharge measurements. Gauge U was located 265 m from the outlet, gauge R was located shortly downstream the spring of the norther branch. Water levels were monitored in 5 min intervals with CTD sensors (Campbell Scientific Inc.), which were installed in perforated PVC tubes in the streambed.

The stream was divided into 50 m sections for further investigation. Labeling of the sections started at the outlet, proceeding upstream as H00, H01, H02 etc.. The two branches were labeled with the suffixes “N” and “S” for north and south,

respectively. The 50 m stream sections were used for general orientation, the large-scale discharge and tracer measurements (basic measurements, Sec. 4.2.4), as well as the terrain analyses (Sec. 4.2.3).

4.2.2 Morphological mapping

The topographic gradient of the streambed was measured along the main stream, using a digital level (Sprinter 250M, Leica Geosystems AG). The spatial resolution was adapted to the slope along the stream and ranged between 0.5 m to 10 m. Additionally, steps or step-pool formations with a drop of more than 15 cm were documented.

Every 5 m along the main stream, we measured the height of the left and right stream bank. To obtain an incision measure for the 50 m sections, the accumulated bank area was calculated as the product of stream length and bank height at both sides of the stream. At the same locations we measured streambed width (based on the morphological shape) and stream width (actually submerged streambed). All morphological mapping was conducted between March 2013 and January 2014.

4.2.3 Terrain analyses

A terrain analysis was performed using a DEM (digital elevation model) of 10 m resolution. The 50 m sections mentioned in Sec. 4.2.1 formed the basis of this analysis. The lateral contributing area of every stream section was calculated as the difference between the accumulated uphill area of the downstream and upstream position of each stream section. Thus, the lateral contributing area is defined as the catchment area contributing water to one stream section only [Bergstrom et al., 2016].

The *extended contributing area* is the lateral contributing area of a stream reach, extending upstream of the investigated stream section for a certain distance. We introduced this measure to check for potential effects superposing the discharge generation at the hillslope, such as parafluvial flow in the valley bottom, which could lead to a displacement of gains in downstream direction. The extended contributing area can be calculated for reaches of any length l and any upstream extension length x and is named ECA_{l+x} accordingly. The extended lateral contributing area for a 50 m section extending 80 m upstream is thus calculated for a 130 m long stream section and called ECA_{50+80}.

Lateral contributing areas and ECAs were both calculated as the differences between the accumulated uphill area of the upstream and downstream end of the reference stream reach (l or $l+x$, respectively). Uphill area was accumulated on the basis of multiple triangular flow directions as presented by Seibert and McGlynn [2007]. All terrain analyses were conducted with SAGA GIS (version 6.3.0) and RSAGA (version 1.0.0).

4.2.4 Sequential discharge measurements and sampling

Field campaigns

Discharge, absolute gains and losses along the stream were measured repeatedly and with varying spatial resolution. This was done by sequential salt dilution measurements (see Sec. 4.2.4 for measuring procedure), which were accompanied by sampling of radon as naturally occurring tracer (Sec. 4.2.4), and in some cases artificial tracer application of sulfur hexafluoride and uranine (Sec. 4.2.4). An overview over the six campaigns and measurements is given in Tab. 4.1).

Four campaigns focused on the entire stream, with measurements integrating over the 50 m long stream sections described in Sec. 4.2.1 (basic measurements). These campaigns were conducted in November 2012, March 2013 (twice) and August 2013. Based on the antecedent discharge conditions, these campaigns were classified into four categories.

- Summer baseflow (campaign 2013/08) was characterized by low baseflow after a long dry period. During this campaign, the upper part of the stream fell dry, which was observed during few days or weeks a year.

- Winter baseflow (2012/11) was characterized by generally moist conditions but without significant precipitation amounts previously or during the campaign.

- Medium flow (2013/03a) was similar to the previous campaign, but experienced a minor snow melting event a few days earlier, causing medium flow conditions.

- High flow (2013/03b) took place during a major snowmelt event resulting in a strong increase in discharge. During this campaign, the intermittent tributary was active.

Table 4.1: Overview over all campaigns and the according measurements (bf=baseflow).

campaign	scope	conditions	SF_6	radon	salt
2012/11	basic	winter bf	–	x	x
2013/03a	basic	medium flow	–	x	x
2013/03b	basic	high flow	–	x	x
2013/08	basic	summer bf	–	x	x
2014/07	high res	summer bf	x	x	x
2015/02	high res	medium flow	x	x	x

Two more campaigns were conducted in July 2014 and February 2015. Here, measurements focused on several short stream sections of 36 m to 60 m length (high resolution measurements). Each of these sections included several distinct morphological features, such as step-pool-formations, braided stream sections or sections characterized by a well defined narrow streambed, which were used to define separately sampled and measured subsections. The length of single subsections ranged from 2.5 m to 15 m. Flow conditions during these campaigns were similar to campaign 2013/08 (summer baseflow) and 2013/03a (medium flow).

Salt dilution measurements

The sequential salt (NaCl) dilution discharge measurements started at the most downstream position. A salt solution was injected instantaneously at least one mixing length (10 m to 15 m) upstream of the measuring point. Breakthrough curves of electric conductivity were measured with two EC-meters (WTW 3100 or Multi 350i, WTW, Xylem Inc.), to assure proper mixing. The sensors were placed 10 mm to 20 mm above the streambed sediment in freely flowing water [Moore, 2005].

In order to determine absolute gains and losses along the section upstream of this point, the next salt dilution measurement was conducted with two additional EC-meters placed at the upstream end of this section. Several EC-meters were installed in between at intermediate positions. Salt injection was repeated exactly as described above and breakthrough curves were recorded at three to eight locations, depending on the number of intermediate positions. Depending on the distance between salt injection point and measuring location, the temporal resolution of the EC measurements was set to 1 s or 5 s.

During the basic campaigns, each salt dilution measurement was conducted over a 100 m stream section. Thus, one measurement covered two 50 m stream sections at a time, with one intermediate measuring location in between. The upstream end

of one measurement was used as the downstream location of the next salt dilution measurement, continuing sequentially from the outlet towards the springs.

The high resolution measurements covered stream sections of 36 m to 60 m length. Besides the downstream and upstream locations, up to 7 intermediate measuring locations were included to increase the spatial resolution along the morphological features described in Sec. 4.2.4.

To calculate the salt concentration from EC data, 1 L water samples were taken at the most upstream measuring point on each day of salt dilution gauging. All sensors were subsequently calibrated in the laboratory. The resulting concentration data were analyzed using the method presented by Payn et al. [2009]. This approach allows us to calculate absolute gain (as opposed to net gain, "gain" without specification will be used to refer to absolute gain in the following) and loss along the measured stream section. The intermediate measurements additionally provide the absolute gain along the measured subsection.

Equations 4.1 and 4.2 show the calculation of loss along the stream reach on basis of the lost tracer mass M_{loss} between the upstream and downstream end of the stream reach and the tracer concentration c_{salt} at the upstream (subscript u) and downstream (subscript d) end of the reach. The latter one depends on the sequence of gaining and loosing locations along the stream reach, which requires the differentiation of the two extreme scenarios. The two scenarios yield the potential minimum and maximum loss: $Q_{loss\,min}$ for upstream loss and subsequent gain, and $Q_{loss\,max}$ for upstream gain and subsequent loss.

$$Q_{loss\,min} = \frac{M_{loss}}{\int_0^t c_{salt\,u}(\tau)\; d\tau} \tag{4.1}$$

$$Q_{loss\,max} = \frac{M_{loss}}{\int_0^t c'_{salt\,d}(\tau)\; d\tau} \tag{4.2}$$

$$Q_{gain} = \Delta Q - Q_{loss} \tag{4.3}$$

Minimum and maximum absolute gain are calculated from loss and net change in discharge ΔQ (Eq. 4.3, applies to both, $Q_{gain\,min}$ and $Q_{gain\,max}$). A more detailed description of the measurements, analysis routine and equations is given by Payn et al. [2009].

To account for potential changes in discharge between measurements of the four basic campaigns, the results of the salt dilution measurements (discharge, gains and losses) were corrected based on the time series measured at the stream gauges. This was done by increasing or reducing the measured fluxes proportionally to the change in discharge observed at the gauge between the time of measurement and the reference time [Szeftel et al., 2011]. The reference time was defined as the time of the discharge measurement at H05 for each of the four campaigns. Data of the campaigns 2013/03a, 2013/03b, and 2013/08 were referenced on basis of the discharge at gauge U, while data of campaign 2012/11 were referenced on basis of the data from gauge R due to sensor failure.

Radon measurements

Surface water radon concentrations were measured to calculate the old and new water mixing ratio of the absolute gains that were quantified by the aforementioned salt dilution measurements. Samples were taken shortly before each salt dilution measurement at each location where EC was measured (upstream, downstream and intermediate locations). In order to avoid turbulence and air contact, and thus loss of radon to the atmosphere, samples were taken with a peristaltic pump. The intake hose was placed 10 mm to 30 mm above the streambed sediment in freely flowing water. Additionally to the surface water samples, springs and lateral inflow were sampled during each campaign. Lateral inflow includes permanent springs along the stream channel as well as event-based seepage that could be sampled from the stream banks.

As sample size directly influences the precision and detection limit of the radon measurements, it was adapted to the expected concentration. 40 mL samples were taken from springs. A sample size of 250 mL was used for most surface water samples. Concentrations below $400\,\mathrm{Bq\,L^{-1}}$ were expected at the stream sections closest to the outlet. Here, a sample size of 1000 mL was used to reliably measure such low concentrations.

The samples were analyzed with a Rad7 Radon Detector (Durridge Company, Inc.), in combination with the RadH2O setup for the 40 mL and 250 mL samples, and the Big Bottle system for 1000 mL (both Durridge Company, Inc.). Both systems are based on the aeration of water samples with a closed air loop. The air is pumped through the sample bottle, dried with desiccant granulate (Drierite, W.A. Hammond Drierite Company Ltd.) and routed to the measuring chamber of the Rad7.

The samples measured with the RadH2O setup (40 mL and 250 mL) were measured for 40 min in 5 min cycles. The 1 L samples were measured for 60 min in 10 min cycles with the Big Bottle system. The first 10 min of each measurement were used for equillibration and were dismissed. Before every measurement, the Rad7 device was purged with fresh, dried air to remove all radon from the measuring chamber and reduce the relative humidity to values below 10 %.

The measured in-air concentrations had to be transferred into water concentrations. The relationship between these values is based on the air and water volume as well as the temperature-dependent *Weigel* coefficient α [Weigel, 1978].

Sulfur hexafluoride & uranine as additional artificial tracers

To reliably calculate the hypothetical radon concentration of the gained water, the loss of radon to the atmosphere had to be estimated. Especially along the short and morphologically distinct stream sections of the high resolution measurements, the degassing rates were expected to vary substantially. Therefore, the high resolution measurements were supported by steady-state gas tracer experiments to assure a precise quantification of radon loss to the atmosphere along the short stream sections. The highly volatile and inert gas sulfur hexafluoride (SF_6) was used to quantify degassing rates of radon along the different stream (sub-)sections. Uranine dye tracer was used in addition to correct for dilution by groundwater inflow. SF_6 samples and fluorescence measurements were taken at the same locations as the radon and salt measurements.

Both tracers were injected continuously approximately 5 m upstream of the investigated stream section, where a small pool allowed for the application of the SF_6 injection described below. A uranine solution was released with a Marrionette's Bottle to ensure a constant injection rate of approximately 100 mL min^{-1}. The concentration of the uranine solution was adapted to the estimated discharge, aiming at a uranine concentration between 5 µg L^{-1} to 50 µg L^{-1} in the stream.

Uranine measurements were conducted *in situ* with two fluorescence probes (DataBank^{TM} Handheld Datalogger, Turner Designs). One of them was placed at the most downstream measuring point and monitored until a constant concentration was established. As soon as steady state was established, the second probe was used to take manual measurements at every observation point. Measurements along the stream were repeated at least three times, while the downstream probe measured continuously.

SF_6 injection was conducted with a setup consisting of semi-permeable silicon tubing loosely attached to a mesh of 0.5 m by 1 m. The injection tube was 35 m long and had an inner diameter of 4 mm with a wall thickness of 0.4 mm. A PVC tube connected the injection tube to a gas cartridge with a pressure regulator.

The injection setup was placed in a small pool of at least 50 mm water depth. It was submerged with stones from the streambed in a manner that the mesh and tubing were located in the slowly flowing water. Prior to the experiment, the tubing was rinsed with SF_6. Afterwards, the pressure was set to approximately 1.5 bar. As soon as the tubing was covered with tiny bubbles, the uranine injection was started and used as a reference for steady tracer concentrations along the investigated stream section.

Water samples for SF_6 analysis were taken with a peristaltic pump and at the same time and locations as the radon samples and uranine measurements. The 1000 mL sample bottles were filled and allowed to spill for a couple of seconds, to replace the water that has been in contact with the air in the bottle during filling. Afterwards, an extraction tube was inserted into the sample bottle. The extraction tube was a gas-permeable silicone tube of 300 mm length, 13 mm inner diameter and 0.5 mm wall thickness, which was sealed with septa at both ends.

The samples were incubated at around 21 °C for at least 7 days, to establish equilibrium between sample water and the gaseous phase within the extraction tube. After the incubation period, each sample bottle was weighed, the temperature was measured and the extraction tube removed from the sample. The SF_6 concentration of the gaseous phase in the extraction tube was measured immediately with the setup described below.

To stabilize the pressure in the extraction tube and the mass spectrometer during measurement, a pressure equalizer (a miniature Marriotte's bottle) was filled with sample water and the outlet connected to one side of the extraction tube, using a syringe. The other end of the extraction tube was connected to the mass spectrometer (OmniStar^{TM}, Pfeiffer Vacuum GmbH), using another syringe. A two-point calibration of the mass spectrometer was conducted, using fresh air and a 99.3 standard as references.

Finally, the SF_6 concentration in the gaseous phase had to be translated into the original concentration of the water sample. The exact water and air volume were determined by weighing the sample bottles and the extraction tubes completely dry and entirely filled with water (including the extraction tube). The original water

concentration is the equilibrium concentration in the water phase, plus the SF_6 mass in the gaseous phase, referenced to the water volume.

The experimentally quantified degassing rate of SF_6 (K_{SF_6}) can be converted into the degassing rate of Rn (K_{Rn}) refering to the Schmidt numbers of the two gases (Sc_{SF_6} and Sc_{Rn}) [Benson et al., 2014]. The Schmidt number is a dimensionless factor describing the ratio of kinematic water viscosity and the diffusion coefficient of the respective gas. The calculation is shown in Eq. 4.4

$$K_{Rn} = \left(\frac{Sc_{Rn}}{Sc_{SF_6}}\right)^{-0.5} \cdot K_{SF_6} \tag{4.4}$$

Degassing is the only radon sink we considered, because loss by radioactive decay is comparably low and thus, marginal at the observed time scales [Genereux and Hemond, 1990]. In-stream radon production is also considered marginal, since radium (^{226}Ra) concentration in surface water is usually low [Genereux and Hemond, 1990].

4.2.5 Data analysis

The composition of the gained water was calculated based on a radon mass balance. In a first step, the hypothetical radon concentration of the section's gain (c_{gain}) was calculated from section gain (Q_{gain}), discharge at the upstream and downstream section end (Q_u and Q_d), the radon concentration at the upstream and downstream section end (c_u and c_d), and under considering the atmospheric loss of Radon as a function of travel time ($exp\,(bt)$). Two scenarios were considered:

1. Upstream gain scenario: 100 % of the entire section's absolute gain happens at the uppermost location. Thus, all gained water is subject to subsequent radon degassing loss along the entire stream length (maximum gain and highest possible radon loss).

$$c_{gain} = \frac{\frac{Q_d \cdot c_d}{\exp(bt)} - Q_u \cdot c_u}{Q_{gainmax}} \tag{4.5}$$

2. Downstream gain scenario: 100 % of the section's absolute gain happens at the most downstream location. Here, Q_u is reduced by the amount of the measured loss and c_u is reduced due to degassing before the surface water mixes with the gained water.

$$c_{gain} = \frac{Q_d \cdot c_d - Q_u \cdot c_u \cdot \exp(bt)}{Q_{gainmin}} \tag{4.6}$$

This procedure is in accordance with the distinction of cases in absolute gain and loss calculation described in Sec. 4.2.4 [Payn et al., 2009, Szeftel et al., 2011]. The resulting two extreme values of c_{gain} were used in a second step to calculate the ratio of new to old water for the gaining water, using equation 4.7 [Genereux and Hemond, 1990].

$$f_{new} = \frac{c_{old} - c_{gain}}{c_{old} - c_{new}} \tag{4.7}$$

In equation 4.7, f_{new} is the fraction of new water and c_{gain}, c_{new}, and c_{old} are the radon concentrations of the absolute gain, new and old water, respectively. The end-member concentrations were estimated based on data from lateral inflow and spring samples.

For better comparability between the different campaigns, the 50 m section gains were translated into the deviation from the average section gain of each campaign in order to highlight the relative spatial patterns. The resulting contribution factors [-] of the entire dataset (for general patterns) as well as single campaigns (for condition-specific patterns) were tested for linear correlations with the lateral contributing area, ECAs and morphological characteristics of the respective stream section. The morphological characteristics include the number of steps per section, bank area, streambed width and streambed gradient.

In a first step, the 50 m contribution factors were correlated with the extended contributing areas from ECA_{50+0} up to ECA_{50+150}, proceeding in 10 m increments. For each campaign, the ECA showing the best regression was used for further analysis.

From the resulting set of potential predictors, the best models for the entire dataset as well as the single campaigns were identified by a stepwise reduction of predictors. The linear models were evaluated using the Akaike information criterion [AIC, Akaike, 1974], resulting in mono- or multivariate models for each single campaign and all campaigns combined.

4.3 Results

4.3.1 Degassing

Degassing rates along the short morphologically identified subsections were very variable (Fig. 4.1). In direct section-by-section comparison, summer measurements usually showed lower values than winter measurements. The degassing weakly correlated with flow velocity, but also stream morphology.

The directly measured data were used to analyze the high resolution measurements (campaigns 2014/07 and 2015/02) and estimate adequate degassing coefficients for the basic measurements (campaigns 2012/11, 2013/03a, 2013/03b and 2013/08). To do so, we used the regression between measured section degassing coefficients and median flow velocities from the SF_6 gas tracer experiment ($R^2 = 0.74$).

While median flow velocities measured during the summer and winter baseflow and medium flow campaigns (2013/08, 2012/11, and 2013/03a) were mostly in the velocity range covered by the SF_6 gas tracer experiments, velocities measured during the high flow campaign (2013/03b) exceeded this range. The reliability of the modeled degassing coefficients is thus the lowest for this campaign (see Fig. 4.1).

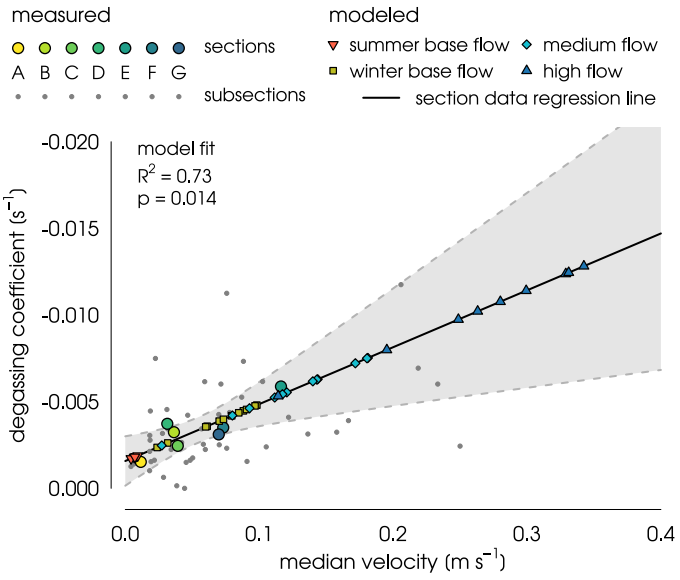

Figure 4.1: Measured degassing rates for short morphological stream sections (subsections) and entire sections measured during the high resolution measuring campaigns. The degassing rates for the basic campaigns were estimated using a linear model (the regression line) fitted to the degassing coefficients and flow velocities measured for the 36 m to 60 m long sections during the high resolution campaigns.

4.3.2 Background concentrations

The radon concentrations of the two targeted end-members (new and old water) were estimated based on 64 water samples taken from springs and lateral inflows to the stream channel. These samples displayed radon concentrations ranging from $0.67\,\mathrm{Bq\,L^{-1}}$ to $40.85\,\mathrm{Bq\,L^{-1}}$, varying more or less independently from location and campaign. For simplicity and conformity, the range of lateral inflow radon concentrations was taken as end-member concentrations for all campaigns.

Based on these end-member conentrations, the degassing rates presented in Sec. 4.3.1, surface water radon concentrations, and section gains, old and new water contributions were calculated.

4.3.3 Discharge, gain and groundwater contribution

50 m sections

The repeated measurements of 50 m section gains during the four basic campaigns (2012/11, 2013/03a, 2013/03b and 2013/08) revealed spatially and temporally highly variable patterns along the stream (Fig. 4.2). Gains ranged between $0.2\,\mathrm{L\,min^{-1}}$ to $1.5\,\mathrm{L\,min^{-1}}$ (per 50 m section) during summer baseflow, and $1.8\,\mathrm{L\,min^{-1}}$ to $12.0\,\mathrm{L\,min^{-1}}$ during high flow.

The measurements in March 2013 (2013/03a and 2013/03b, medium flow and high flow) showed highest absolute gains between sections H05 and H08 as well as the source areas. This patterns was especially pronounced during high flow, when these sections increased gain by factors up to 0, while sections H09 and H10 did not show a significant change in section gain. Campaign 2012/11 (winter baseflow) showed a different and almost inverted picture, with the highest gains at sections H03, H04 and H08. During summer baseflow (2013/08), half of the stream fell dry. It emerged shortly downstream of H07 and showed the strongest gain at H02. Summarizing the results of all campaigns, the absolute gain along the stream does not exhibit temporally stable patterns.

The losses, which could only be quantified for 100 m section due to the salt tracer setup, were high during all campaigns and did not show strong variability at such a large scale. The ratio of absolute gain to loss accumulated over the entire stream was the lowest during summer baseflow (2.7) and highest during high flow (6.0).

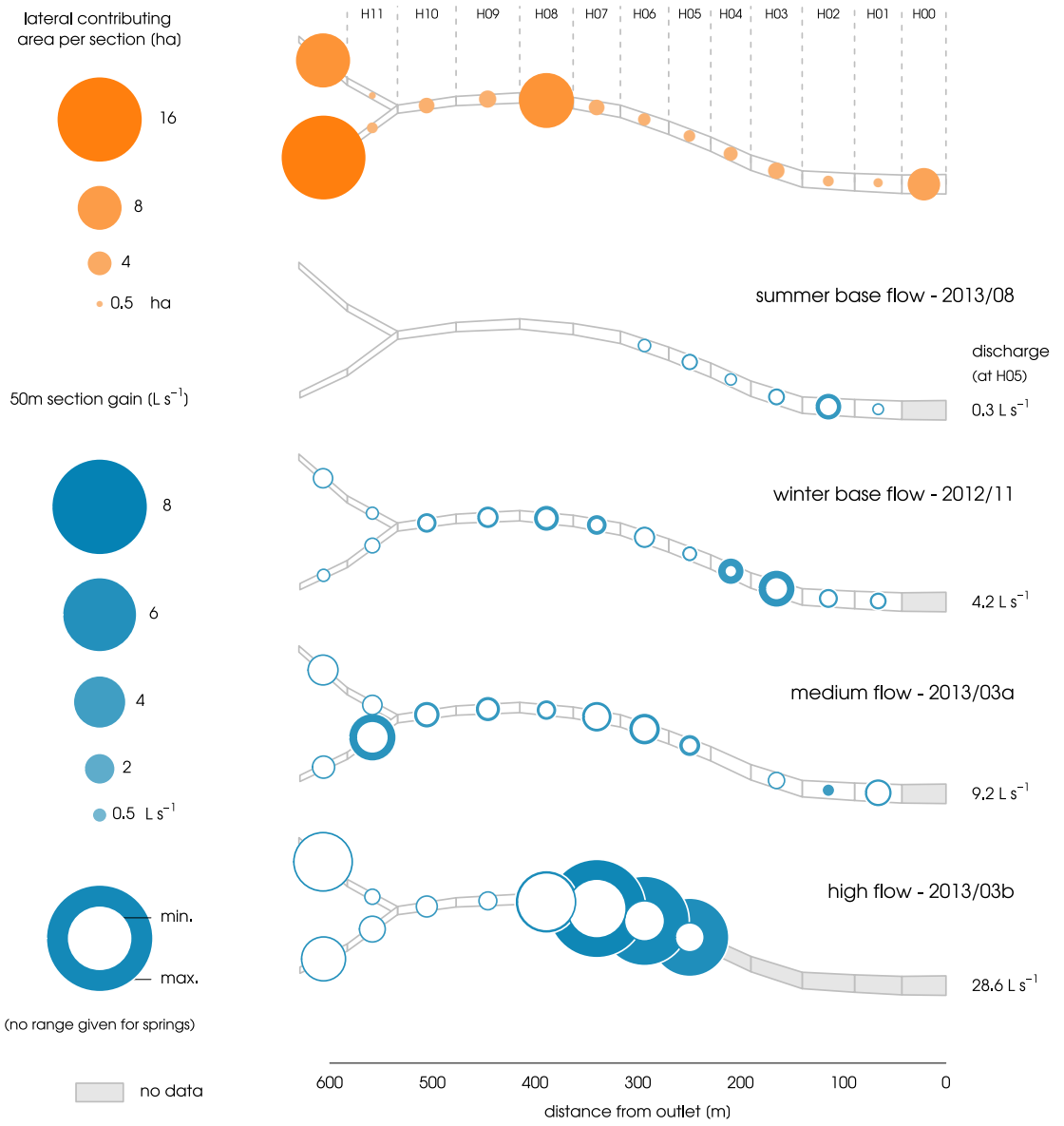

Figure 4.2: Lateral contributing area and 50m section gain. The minimum and maximum gains were calculated from the two scenarios described in Sec. 4.2.4, depending on the order of gains and losses along the stream section. All gains were corrected for changes in discharge during the campaign measurements.

The radon mass balance was calculated for each stream section and was used to estimate the ratio of new and old water of the gain along the respective stream section. Along with the two scenarios of upstream and downstream section gain calculated from the salt dilution measurements, the calculation of the water compositions also varies, resulting to two very different pictures. Figure 4.3 shows the gains and ratios for the two scenarios. Given the enormous differences between the

two scenarios, meaningful interpretations of the 50 m patterns are scarce. However, there is a tendency towards higher portions of new water in the downstream 200 m of the stream and a generally higher portion of new water during the high flow campaign (2013/02b). Furthermore, the high portion of new water in springs during all campaigns is remarkable and indicated mixing and degassing in the unsaturated zone.

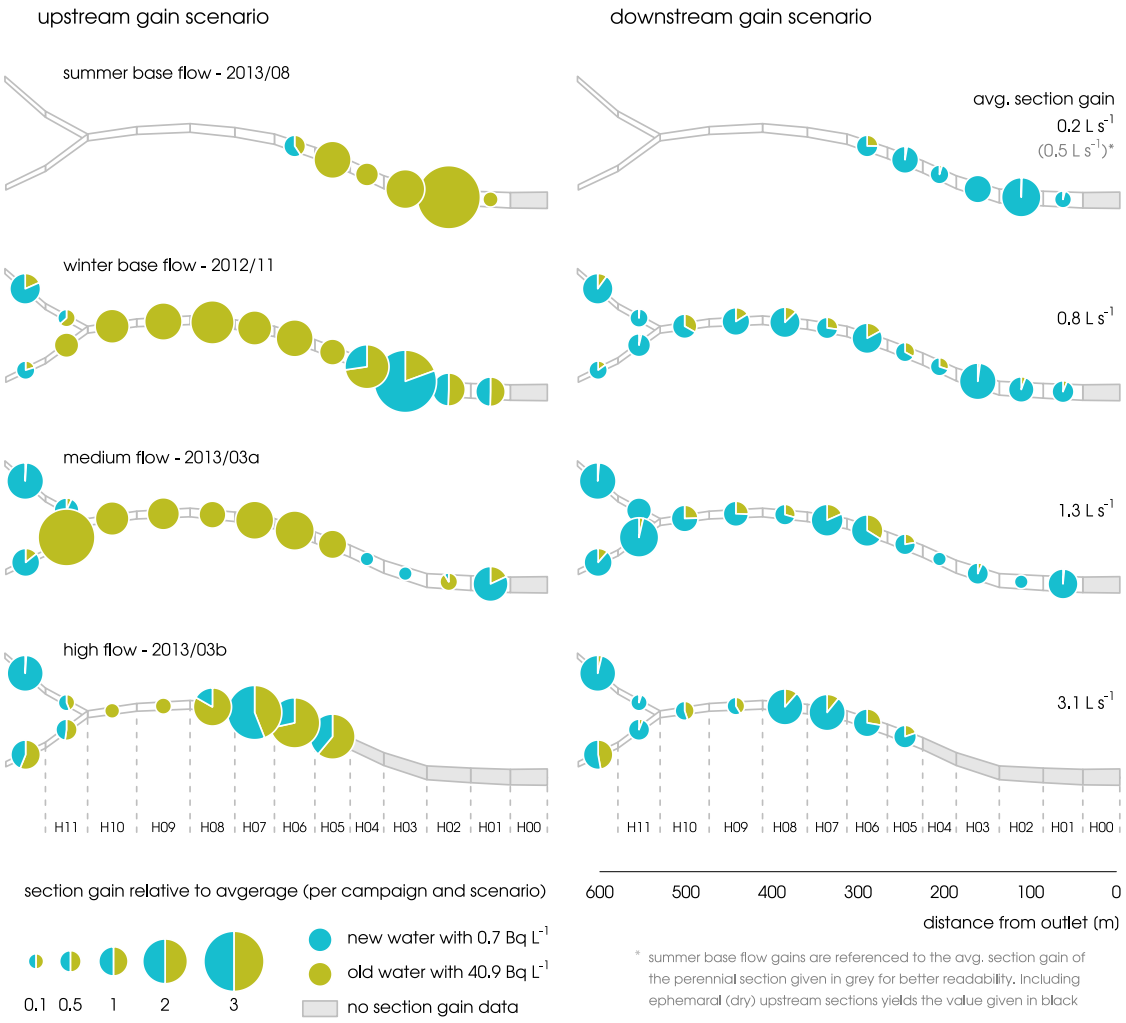

Figure 4.3: Old and new water fractions of 50 m section gains shown for the two extreme scenarios described in Sec. 4.2.4 and 4.2.5: all gain at the upstream end of each section and subsequent loss (left) *versus* loss first and gain at the downstream end of each section (right). As the order of gains and losses along the stream section has implications for both, the gain calculation (see Fig. 4.2) as well as the radon degassing, this leads to relatively high uncertainties. Please note that these scenarios are to be interpreted as per-section scenarios, which are independent from each other.

Morphologically identified subsections

For the short (1 m to 10 m), morphologically identified stream sections the spatial patterns were highly variable in space but quite stable in time (Fig. 4.4). The absolute gain per subsection reached values as high as $0.24\,\mathrm{L\,s^{-1}\,m^{-1}}$. The gains were usually concentrated at few subsections, with others showing no gain at all. Aside from few exceptions, all bigger gains occurred at sudden drops like steps and step-pool formations. Gains were usually dominated by new water, even during summer baseflow conditions.

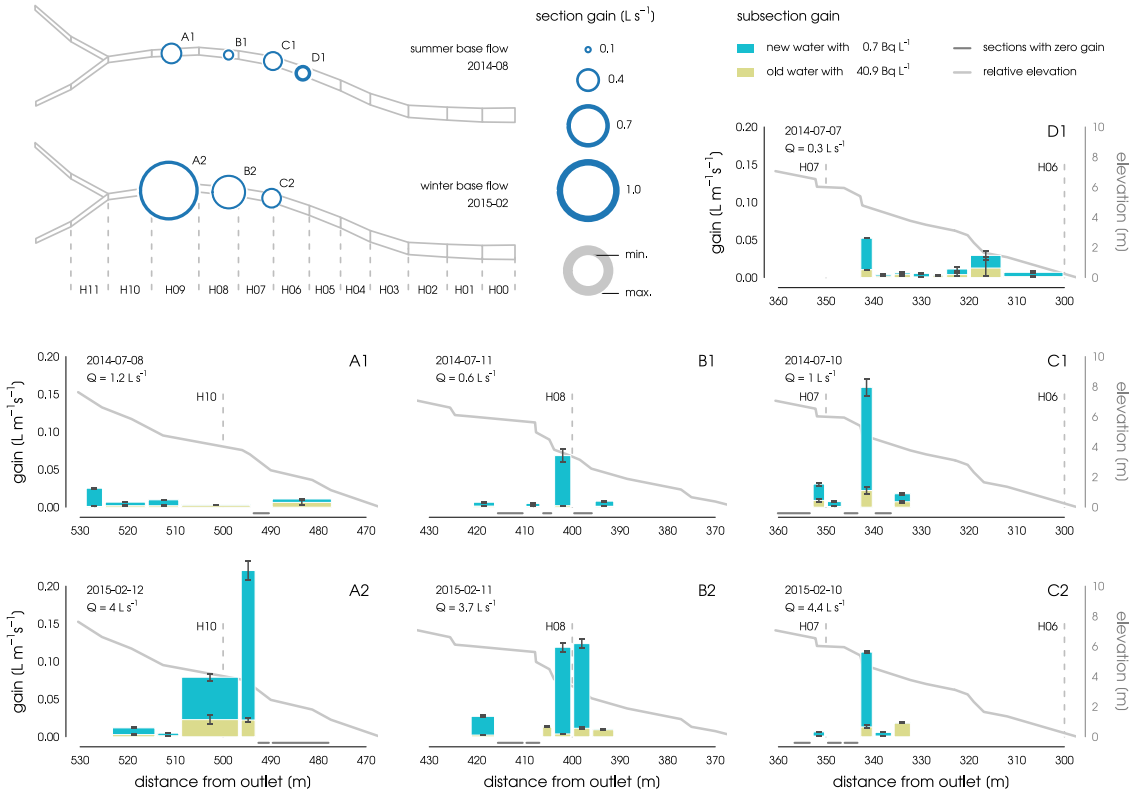

Figure 4.4: Old and new water fractions for short section gains. Despite the different discharge conditions during the two campaigns, the spatial patterns were quite similar, with most gains being concentrated at steps in the longitudinal streambed profile.

4.3.4 Topography and morphology

The Holtz headwater had a absolute size of 61.8 ha and was subdivided into 15 sub-catchments (or catchment segments) by the 50 m section (Fig. 4.2 top), which make up the lateral contributing area of the respective stream section. The sub-catchments varied in size between 0.67 ha to 15.93 ha. Besides the two big sub-

catchments forming the northern and southern spring area (9.99 ha and 15.93 ha, north and south, respectively), H08 had the largest sub-catchment along the main stream. With 10.21 ha H08 included the area of the intermittent tributary.

The longitudinal streambed gradient was relatively homogeneous along the stream (Fig. 4.5 top). The overall gradient of the main stream was −0.11, the average gradient of the 50 m sections ranged between −0.08 and −0.12. Locally, the variability was higher due to numerous steps and step-pool formations, but also several meter long sections of different slopes.

The total number of steps along the 560 m long stream was 55, unevenly distributed along the entire length of the stream. The highest number of steps was found in sections H02, H03, and H05, while the most steps with drops of more than 30 cm were located further upstream (Fig. 4.5 top).

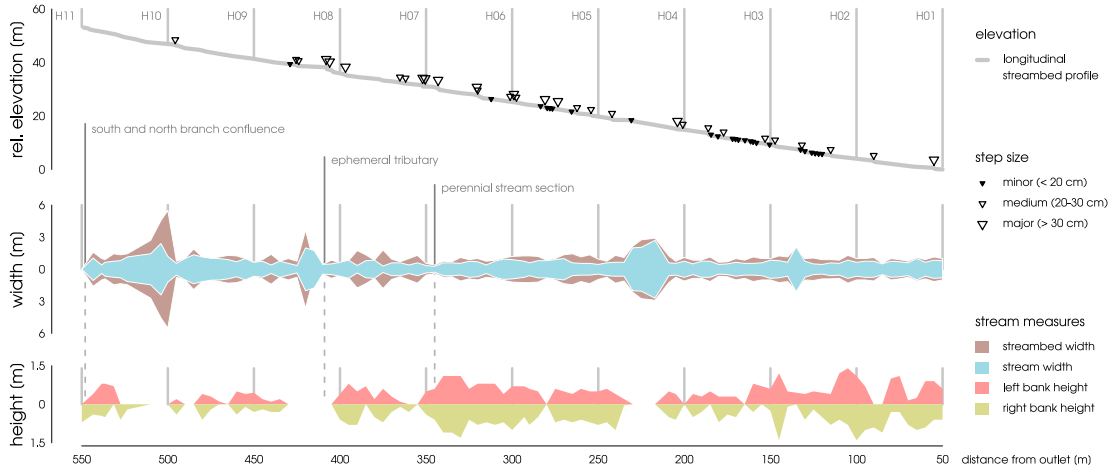

Figure 4.5: Elevation, steps, stream bank height, and streambed width along the main stream.

The streambed and (inundated) stream width varied along the entire stream, but did not show a trend at the 50 m scale. The pattern was dominated by three braided or impounded stream reaches of different length (located in H04, H08, and H10), which are characterized by high stream width and even broader streambeds (Fig. 4.5 center).

The bank height varied strongly along the stream, with a maximum measured height of 1.4 m (Fig. 4.5 bottom). The stream sections upstream of 350 m from the outlet (H07) showed no or only shallow incision. Bank height increased strongly shortly downstream of H07, coinciding with a major step-pool formation. The accumulated bank area was calculated for every 50 m section as the product of stream length and bank height for both sides of the stream. The accumulated bank areas

of all sections upstream of H07 ranged between 10.9 m^2 and 38.2 m^2, and between 28.5 m^2 and 81.5 m^2 downstream of H07.

4.3.5 Correlation analysis

50 m section gain and contributing area

As a first step in the correlation analysis between 50 m section gains and topographical or morphological predictors, the importance of the lateral contributing area and different ECAs were compared. We found that 50 m section gains did not correlate significantly with the lateral contributing areas (or ECA_{50+0}). For all four campaigns, R^2 remained below 0.2. The correlation with the ECA_{50+0} through ECA_{50+150}, however, strongly varied between campaigns and yielded different results for different discharge conditions (Fig. 4.6).

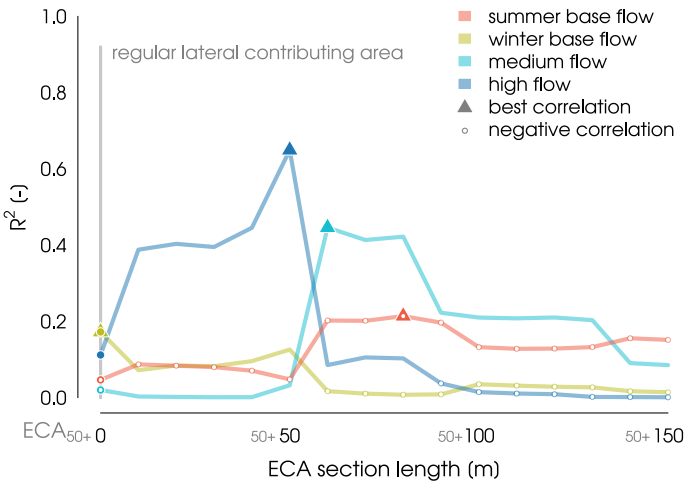

Figure 4.6: The development of R^2 of the mono-variate correlation between 50 m section gain (referenced to the average section gain per campaign, i.e., contributing factors) and a range of extended contributing areas (ECA_{50+0} through ECA_{50+150}, $n = 12$ for medium flow, winter and summer baseflow conditions and $n = 8$ for high flow conditions). The maximum R^2 indicated the ECA that best predicts the gain patterns and which was used for further analysis. White dots mark ECAs with negative correlations.

The entire range of ECAs did not show a satisfying correlation for the winter baseflow campaign (2012/11), with a maximum R^2 of 0.17 for ECA_{50+0}, i.e., the regular lateral contributing area. During medium flow (2013/03a), correlations with ECA_{50+60} to ECA_{50+80} yielded R^2 values up to 0.45. A R^2 value 0.21 was found with ECA_{50+80} for summer baseflow conditions (campaign 2013/08), however, with a negative correlations over the entire range of ECAs. The most distinct result was shown for high flow conditions, where the correlation improved from an R^2 of 0.11 for the correlation with the regular lateral contributing area to 0.65 for ECA_{50+50}. The ECAs were included in the following multivariate correlation analysis in the

next step. To do so, for each campaign only the ECA showing the best correlation was chosen for further analysis to evaluate their significance relatively to and in combination with other morphological parameters.

50 m section gain and morphology

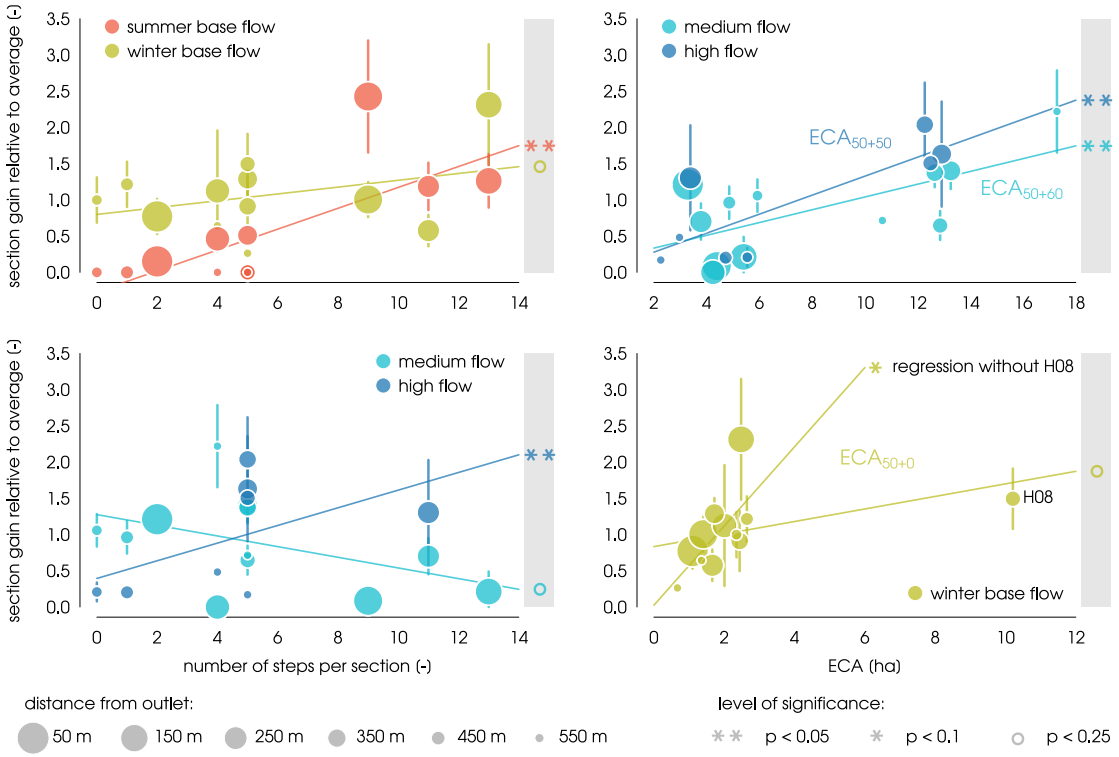

Figure 4.7: 50m section gains of the four campaigns plotted against their best predictors. While the correlations between summer baseflow and high flow patterns are strong, intermediate flow conditions (winter baseflow and medium flow) are less defined (see also Tab. 4.2). Especially during winter flow conditions, the variability among the 50m section gains is low.

The best predictors for 50 m gain patterns for all campaigns together as well as for each campaign separately were identified by a stepwise model selection, resulting in mono- or multivariate models. Prior to model selection, however, streambed width and bank area were dismissed from the list of predictors due to their significant correlation with other morphological properties (gradient and number of steps per section, $p < 0.1$).

Correlations between the 50 m section gains of all campaigns and the set of morphological parameters were weak. The number of steps per section was identified to be the only significant predictor for the entire dataset ($p = 0.015$) but showed a weak correlation of $R^2 = 0.13$.

Table 4.2: Overview over best models identified for each campaign

campaign	steps	$\frac{\Delta x}{\Delta h}$	ECA	R^2	p-value
high flow	**	^	**	0.95	0.004
medium flow	°		**	0.57	0.021
winter bf	°		°	0.31	0.193
summer bf	**		°	0.66	0.008
all	**			0.13	0.015
Significance:	0 **	0.05 *	0.1 °	0.25 ^	0.35

The picture became clearer when single campaigns were correlated separately. The stepwise model selection returned regressions of reasonable quality for summer baseflow, medium flow and high flow conditions, with multiple R^2s of 0.66, 0.57 and 0.95, respectively.

Models for all campaigns included the number of steps per section and the campaign-specific ECA, but with varying significance (Tab. 4.2). For high flow conditions, the average section gradient was also included, but showed a very low p-value of 0.341, while number of steps and ECA_{50+50} showed p-values of 0.009 and 0.007, respectively. Section gains (relative to average) are plotted against all significant predictors in Fig. 4.7.

Especially for medium and high flow conditions, the campaign-specific ECAs appeared to explain most of the gain patterns, while during summer baseflow the number of steps per section prevailed (based on statistical significance). The low R^2 for the campaign under winter baseflow conditions (0.31, $p = 0.193$) might be due to generally relatively low spatial variability during this campaign. The regression could be drastically improved by excluding section H08 (including the intermittent tributary) from the dataset, resulting in an R^2 of 0.44 ($p = 0.026$). This could be explained by an incomplete connection of the large H08 tributary catchment. Under this assumption, the lateral contributing area (ECA_{50+0}) was the most significant control, with parafluvial flow paths (intersecting hillslope flow paths) being of minor importance.

The models included two cases of negative correlations: between gain patterns during summer baseflow and ECA_{50+80}, and between medium flow gain patterns and the number of steps per section. The first one could be explained by the fact that all sections with well-above-average lateral contributing areas (H08, and the two springs) were located upstream of the perennial stream section. The extended or lateral contributing areas were therefore not decisive for the maintenance of baseflow

under very dry conditions. However, both show a low statistical significance (p-values of 0.115 and 0.133, respectively).

Correlations between water losses measured for the 100 m sections and morphology (number of steps per section and topographic gradient), are non-existent.

4.4 Discussion

4.4.1 Methodological discussion

Small headwater streams are often characterized by a highly diverse channel geomorphology. This diversity impacts the hydrological behavior of the stream and controls its connectivity to the hyporheic and riparian zone [Wondzell, 2006] as well as in-stream processes, such as flow velocity and turbulence. In such heterogeneous systems, relevant scales and scales of interest are in the range of few meters to several 10 m. To investigate headwater catchments, methods are required which can handle these small scales and are robust over a broad range of different flow conditions.

Differential salt dilution gauging

The method of differential salt dilution gauging and mass recovery analysis was first applied along 200 m long reaches of a 2.6 km long mountain headwater stream with substantial channel and valley structural variability [Payn et al., 2009, Stringer Creek, part of the Tenderfoot Creek Experimental Forrest in Montana, USA]. Instead of measuring net changes only, the method allows us to quantify concurrent absolute gains and losses, which can be substantial [e.g., Covino and McGlynn, 2007] and are indicative for the hyporheic interaction within and between contiguous stream reaches [Payn et al., 2009]. The method is especially well suited for small, turbulent streams with high structural variability, as it does not require close to laminar flow or uniform channel geometry and rather benefits from turbulent flow and thus quick mixing [e.g., Day and Day, 1977, Moore, 2005].

There are, however, a few things to consider, which are not necessarily limited to small, turbulent streams, but may be more problematic than in larger systems. As described in Sec. 4.2.4 and elaborated by Payn et al. [2009], the data analysis requires a distinction of cases, representing the potential sequence of gains and losses along the reach resulting in two potential extreme cases (all losses first, then gain, and *vice versa*). Especially in systems with high losses, this creates a lot of uncertainty. With strongly localized gain (as observed in the Holtz stream, see Fig. 4.4) and potentially also localized loss, this uncertainty is even more immanent than in systems with more evenly distributed gains and losses, were average values might be an adequate approximation. Thus, a better localization of gains and losses is desirable, meaning a refinement of resolution. However, at such small scales (few

m) the small gains (and losses) can not be resolved anymore. Here, steady-state tracer injections appeared to yield more reliable results for gains (or dilution) than Dirac injection, but do not allow for loss calculation.

Radon as groundwater tracer

Radon is increasingly used as tracer for groundwater inflow to streams. However, most applications were performed in lowland rivers or larger streams and focus on scales of few 10 km to up to several 100 km [e.g., Cook et al., 2003, 2006, Mullinger et al., 2007, Unland et al., 2013, Cartwright et al., 2014, Atkinson et al., 2015, Frei and Gilfedder, 2015]. While the methodologies applied in those studies vary, they are all based on in-stream radon mass balances and usually use radon in order to directly estimate groundwater inflow. Common problems and uncertainties were identified to arise from radon input through hyporheic exchange [Cook et al., 2006], variable groundwater background concentrations [Hoehn and Von Gunten, 1989], and radon loss to the atmosphere. In small, turbulent streams, the same factors are relevant and will be discussed in the following.

While in-stream radon production is considered irrelevant by the vast majority of researchers, radon enrichment of water subject to hyporheic exchange was identified as a significant source, that can lead to overestimation of groundwater gain of up to 70 % if disregarded [Cook et al., 2006]. To compensate for this, a hyporheic production term was introduced into mass balance models, which is based on the assumption of strictly vertical exchange flow and calculated from porosity, depth of the hyporheic zone, residence time of the water, and radon production rate in the hyporheic zone [Cook et al., 2006, Frei and Gilfedder, 2015]. This simplification is legitimate when the investigated stream reaches are (a) homogeneous enough to describe its hyporheic zone and exchange flow with constant parameters, and (b) significantly longer than potential lateral hyporheic or parafluvial flow paths. Both conditions are unlikely to be met by morphologically diverse headwater streams, especially when investigated at the scale of meters.

In the present study, radon input by means of hyporheic flow was accounted for by measuring absolute gains separately and using radon for end-member separation only. This way, hyporheic flow was represented as a part of the new water end-member and allowed to travel laterally.

Another source of uncertainty lies in the difficult quantification of end-member concentrations for both, new and old water. The definition of the new water com-

ponent on the basis of one tracer is vague and includes hyporheic/parafluvial water, interflow, and water from the vadose zone. Accordingly, many different processes controlling the radon concentration are involved. Hyporheic/parafluvial water is stream water that becomes enriched with radon while travelling through the mostly saturated subsurface, interflow is rain water being enriched in the unsaturated zone, and water from the vadose zone might have had groundwater concentrations but is subject to degassing. The equilibrium concentrations of these processes might not necessarily be the same, but are not sufficiently distinct to distinguish them as different end-members. However, they are all significantly lower than groundwater equilibrium concentrations. Combination with other environmental tracers such as stable isotopes, tritium, and chloride [e.g., Atkinson et al., 2015], have been suggested to refine the definition of end-members.

The processes controlling the old water radon concentration are more uniform, namely, the equilibrium between radon production and decay in the saturated zone. However, radon production rates are very sensitive to the lithological characteristics of the aquifer, such as radium activity, porosity, and specific surface. This sensitivity leads to a very high spatial variability even at the scale of meters [Hoehn and Von Gunten, 1989]. Water samples from near-stream Piezometers [Cook et al., 2006] or direct measurements of seepage concentrations as shown here might therefore be more reliable than regional groundwater values.

In our study, end-member radon concentrations were estimated based on water samples taken from lateral inflow and permanent spring water, totaling 64 samples taken at 18 locations during 6 campaigns. These data revealed a high spatiotemporal variability of groundwater radon concentrations and was not sufficient to identify any spatial or temporal patterns. At larger scales, the average values are more representative than at small scales, where the length of investigated stream reaches might be similar to the scale of radon concentration variability. However, the big difference between the two end-members (new *versus* old water, i.e., short *versus* long subsurface residence times, with $0.7\,\mathrm{Bq\,L^{-1}}$ and $40.9\,\mathrm{Bq\,L^{-1}}$, respectively) makes the data analysis relatively insensitive to the exact accuracy of end-member concentrations.

The biggest source of uncertainty at small scales is the gas exchange with the atmosphere. While degassing rates are generally assumed to be critical for the in-stream radon mass balance [Genereux and Hemond, 1990, e.g.,] and hard to quantify, this factor needs special attention in small systems. This is because the gas exchange is generally higher in shallow, fast flowing streams due to the continu-

ous surface water replacement [Raymond and Cole, 2001]. Furthermore, the high morphological variability not only prevents the application of established models [Genereux and Hemond, 1992], but also increases the spatial variability of degassing rates, which which have been shown to be a function of flow depth and velocity [Streeter and Phelps, 1925] and thus, controlled by stream channel slope and topography [Tsivoglou and Neal, 1976]. The results of the SF_6 tracer clearly showed this variability at the scales of few meters (Fig. 4.1). Generalized assumptions could lead to significant errors in the calculation of the radon mass balance for such short and diverse stream sections. A direct measurement of degassing rates by artificial gas tracer such as SF_6 strongly reduces the uncertainty here, and becomes more important the more divers the stream's morphology and the shorter the investigated stream subsections. Degassing rates over the longer distances (approximately 50 m) appeared to be more stable, as the meter-scale variability averages out. However, estimation of degassing rates still is more challenging than for bigger systems, and direct measurements are recommended.

Another aspect that needs consideration is the impact of the two scenarios regarding the spatial sequence of gains and losses along the investigated stream reach (all gains first, subsequent loss or *vice versa*). This distinction yields extreme values of absolute gains and losses [Payn et al., 2009] and also impacts the radon mass balance and thus the resulting ratio of new and old water. This is not only by the different volumes of mixing waters, but also by the different time and flow length available for degassing before the downstream measuring point. As mentioned above, this theoretical calculation of extreme values is even more important in systems with localized gains, and the differences between the two scenarios increase proportionally with radon degassing rates.

The strong effect of degassing in combination with the extreme case scenarios becomes strikingly visible in the results of the 50 m sections (Fig. 4.3). The end-member contributions of the two scenarios are very different, making meaningful interpretations impossible. This effect increases with decreasing discharge: while the results from the high flow campaign (2013/03b) return different values for the old and new water composition of the gained water, summer baseflow (2013/08) results yield ranges of 0 % to 100 % for almost all sections, which contain no valuable information at all.

The results from the morphologically identified subsections show a more concise picture. Due to the spatial constriction of gains, the uncertainty was drastically

reduced. Here, most of the uncertainty stems from the gain calculation itself rather than the potential radon degassing, and results on water composition show significant differences.

Summarizing these observations, we conclude that the usability of radon as groundwater tracer is limited by the importance of degassing for the overall radon mass balance. The combination of radon and differential salt dilution gauging is therefore a useful approach to investigate patterns of groundwater-surface water interactions in small streams **(H3)** only if the observation scale is adapted to the size and discharge condition of the stream under study. At the Holtz headwater, the method worked well at the scale of around 10 m under summer baseflow and medium flow conditions. The results of the salt dilution measurements on their own were well suited for the investigation of this small headwater and the main basis for the process understanding gained in this study.

4.4.2 Controls of hillslope-stream connectivity: riparian morphology *versus* catchment topography

The observed patterns of gains (and losses) along the investigated headwater stream were interpreted and compared with the riparian morphology and catchment topography in order to identify the structural controls and relevant scales. Gain patterns showed a different behavior at different scales: the short morphologically identified stream subsections showed high spatial variability, with gains being highly localized. These small-scale patterns were temporally stable, while the large-scale gain patterns were more variable over time. This scale-dependent variability indicates, that processes at different scales superimpose each other and vary in importance according to the hydro-meteorological conditions.

At the meter-scale (morphologically identified subsections), stable patterns indicate that the connectivity between subsurface water and the stream channel was controlled by the stream morphology at all times. In most cases, high gains occurred at step-pool formations or subsections with steep streambed gradients (Fig. 4.4). This means that all subsurface water, be it previously lost hyporheic water, event water or groundwater, preferably enters the stream channel at those locations.

Consistently, the number of steps also was the best predictor for the gain patterns at the larger scale (50 m stream sections), when all campaigns were considered. However, the correlation was very weak ($R^2 = 0.13$, $p = 0.015$). When 50 m gain

patterns of single campaigns were correlated separately, different predictors were eligible and appeared to overlay the small-scale effect of stream channel morphology. The changing importance of different predictors with changing hydro-meteorological conditions will be discussed in more detail in the following.

Under baseflow conditions, the stream channel morphology had the highest impact on gain patterns. The number of steps per section was the best predictor for 50 m section gain during summer baseflow conditions (2013/08, Fig. 4.7). Also, the perennial stream reach begins at H07, 350 m from the outlet. Upstream of this point, the stream fell dry occasionally (few days to weeks during summer). The streambed was less deeply incised and had a lower number of steps per section in comparison with the downstream located perennial stream reach (Fig. 4.5).

During the medium and high flow campaigns in March 2013 (2013/03a and 2013/03b) the picture was different. During both campaigns, the highest gains were observed downstream of the confluence with the intermittent tributary at H08. The same locations also showed the strongest increase following the thawing event preceding 201303b, which also activated the intermittent tributary. Under such wet conditions, the extended contributing areas of ECA_{50+60} for medium flow and ECA_{50+50} for high flow were better predictors for gain patterns than the number of steps per section. This suggests, that water was routed from the hillslopes towards the valley, with subcatchments contributing to event discharge proportionally to their size. Instead of entering the stream immediately, however, hillslope flow paths were intersected by parafluvial flow paths, shifting the hillslope/topography signal downstream.

The losses we measured were high. In combination with the high percentages of new water (Fig. 4.4), this observation indicates high hyporheic interaction along the entire stream. However, the presented experimental setup yielded absolute losses in a lower spatial resolution than absolute gains. Due to this reduced resolution, no patterns or correlations could be identified for stream losses.

From these observations of gains and losses, we concluded that parafluvial preferential flow in subsurface structures in the riparian zone was a key process defining the gain patterns along the stream. The identified ECAs for medium and high flow conditions (ECA_{50+60} and ECA_{50+50}) indicated, that a fraction of the event water from the hillslopes traveled up to 60 m downstream before entering the stream channel when these flow paths were connected to the hillslopes. The structures enabling the preferential flow are weathered bedrock and saprolite deposits, layered into the

otherwise loamy substrate of the alluvial. Similar structures (periglacial slope deposits) can be found at hillslopes throughout the catchment and were reported to cause preferential flow in the hillslopes [Angermann et al., 2017, Jackisch et al., 2017]. With velocities of up to $10^{-3}\,\mathrm{m\,s^{-1}}$ under otherwise unsaturated conditions, the hillslope structures were found to impact the overall catchment behavior of the Holtz stream [Angermann et al., 2017]. In the frame of this study, the riparian zone structures were not investigated or mapped in detail. However, deposit layers similar to those reported at the hillslopes could be found at many locations along the stream channel where they were excavated by steps or deeply incised stream banks. The water percolated through these high conductivity deposit layers and entered the stream at locations where these layers were interrupted.

While the elaborations above describe the role of parafluvial flow through high conductivity deposit layers in the riparian zone for event runoff, the same mechanisms are likely to control the connectivity between groundwater and riparian zone and the stream channel. In earlier studies, losses were found to correlate with the number of steps per section [Harvey and Bencala, 1993, Bergstrom et al., 2016], while at the same time, steps were identified as hotspots for groundwater or hyporheic water upwelling [e.g., Harvey and Bencala, 1993]. Depending on the overall structural properties of the stream channel and the riparian zone, nested series of subsurface flow paths develop [Wondzell, 2006]. Payn et al. [2012] described losses from upstream reaches traveling along intermediate flow paths reemerging downstream in the channel and contributing to more discharge than expected from the downstream lateral areas. While hyporheic exchange could not be observed in more detail here, the processes described in earlier studies are likely to also apply to the Holtz stream and can explain the high portions of new water gained during baseflow conditions.

With the riparian zone being located between the stream channel and the catchment, parafluvial flow paths intersect with flow paths routing water from the hillslopes and the groundwater towards the stream [Tóth, 1963]. Thus, hyporheic flow, surface water groundwater interactions and event water contribution are subject to the same mechanisms. The riparian zone is a buffer not only for water chemistry, but also for hydrological processes [Hill, 1996, Hooper et al., 1998, Burt, 2005].

The relevance of hyporheic and parafluvial flow processes for the (in-stream) water balance depends on the scale of interest and the process scale. Bergstrom et al. [2016] found that the lateral contributing area was a significant predictor for

net discharge only at scales of 1.5 km and larger. They hypothesized, that flow paths from the contributing areas "intersected longer, more intermediate flow paths in the valley system [Tóth, 1963]", and also cite Payn et al. [2012], describing the shift of losses from upstream reaches reemerging downstream.

Parafluvial flow processes are relevant, if the amount of water shifted in or out of the targeted stream reach is comparably large in relation to the contribution from the lateral contributing area related to the investigated reach. This is more likely the more water travels through the parafluvial flow paths, and the longer the parafluvial flow paths are in relation to the observation scale. In the case of the Holtz stream, the regression analysis suggested that water can travel in parafluvial flow paths up to 60 m downstream. This can affect discharge patterns at the scale between 10 m and a couple of 100 m, while it is irrelevant for the scale of the entire headwater catchment. Depending on the scale of interest, the parafluvial flow paths controlled by stream morphology can have a strong impact on hillslope-stream connectivity in highly structured headwater catchments **(H1)**.

4.4.3 Groundwater discharge patterns as fingerprint of catchment organization and hillslope-stream connectivity

From a conceptual point of view, discharge patterns in the stream channel are always the result of catchment processes. However, the superposition of signals are limitations, which define the informational value of the channel signal. The stronger the influence of the riparian zone processes, the weaker or more obscure are the signals from the hillslopes and *vice versa*. The informational value of channel discharge patterns can thus be summarized as follows:

1. discharge patterns are the fingerprint of processes in the riparian zone.
2. they also always contain information about the relative importance of riparian zone *versus* hillslope or groundwater flow processes.
3. depending on the relative importance of the riparian zone signal, they might also contain information about more distant processes.

In practice, the extraction of the relevant information, i.e., the identification of patterns and their interpretation, is not trivial. A multi-scale approach is necessary to identify the impact of different processes and structures. Observation scales have

to be chosen according to the processes to be identified and results are limited by assumptions and the experimental design.

Under consideration of these conceptual and practical limitations, groundwater water discharge patterns can be a valuable source of information about runoff generation processes and the hierarchical interplay of spatial structures within the catchment **(H2)**. Especially the combination with structural data helps to formulate well informed hypotheses and develop an experimental design that matches the relevant processes in the catchment, but also helps to interpret the patterns more deeply.

4.5 Summary and conclusions

To characterize hillslope-stream connectivity and its controls we measured discharge gains for consecutive stream reaches of 2 m to 50 m in length. We then correlated the results with structural properties in order to identify the relative importance of stream channel morphology and catchment topography on hillslope-stream connectivity. We found that local morphological features, such as step-pool formations, dominated local discharge patterns (at the scale of 2.5 m to 15 m), and also influenced reach-scale patterns by controlling the connectivity between parafluvial flow paths and stream channel. The spatial distribution of event discharge contributions was primarily controlled by topography, but the signal was modulated and shifted downstream. The discharge patterns were hierarchically controlled, with large-scale signals shaped by topography being superposed by processes at consecutively smaller scales in the riparian zone.

Our experiments also showed that stream reach discharge patterns can be used to identify internal catchment processes to a certain extent. The patterns are the fingerprint of processes and structures in the catchment, the riparian zone and the streambed, however, the informational value of processes "further away" from the stream channel strongly depends on the degree to which the catchment signal is altered in the riparian zone. Even if under certain conditions the riparian zone and stream channel may be the only factors controlling the discharge patterns, obscuring all information about more distant processes, the relative importance of different processes can always be evaluated.

Regarding the methodological approach, we found that the application of differential salt dilution gauging in combination with radon as environmental tracer was suited to identify spatial discharge patterns, but comes with a number of limitations which need to be considered. These challenges include the high variability of end-member background concentrations and degassing rates. They apply to most applications of radon as groundwater tracer, but are especially pronounced at small scales and for heterogeneous streams. Nevertheless, the distinct concentrations in surface water and groundwater are an advantage especially at small spatial scales and make radon a powerful tracer for surface water-groundwater interactions. In combination with other environmental or artificial tracers to quantify degassing rates, absolute gain, or more clearly define different end-members, radon is a cheap and flexible tool to investigate groundwater discharge into small and turbulent streams.

Acknowledgements We are grateful to Sibylle Hassler, Lisei Köhn-Reich, Begoña Lorente Sistiaga, Daniel Mack, Christina Tecklenburg, and Heiko Vollmann for their support during fieldwork. Britta Kattenstroth is acknowledged for hydrometeorological data acquisition. We also would like to thank Martin Zimmer, who greatly contributed to the conceptualization of the SF_6 sampling and analysis, as well as Brian McGlynn for extensive discussions on the data. Laurent Pfister and Jean-Francois Iffly from the Luxembourg Institute of Science and Technology (LIST) are acknowledged for organizing the permissions for the experiments This study is part of the DFG funded CAOS project 'From Catchments as Organised Systems to Models based on Dynamic Functional Units' (FOR 1598).

Chapter 5

Connectivity between groundwater, hyporheic zone, and surface water

Application of heat pulse injections for investigating shallow hyporheic flow in a lowland river

Lisa Angermann, Stefan Krause, and Jörg Lewandowski

Originally published in:
Water Resources Research, 48(12), 2012.

Abstract

Hyporheic zone processes can have significant impact on groundwater and surface water resources. Detailed knowledge of exchange flow patterns is crucial for understanding the ecohydrological and biogeochemical functioning of river corridors. In particular, small-scale hyporheic exchange flow is still poorly understood, partially because of the lack of adequate in situ monitoring technology. This paper investigates the spatial heterogeneity of hyporheic exchange flow in a lowland river at multiple scales. It demonstrates the conjunctive use of active heat pulse tracing at shallow depths (15 cm) and vertical hydraulic gradients (VHG) at 120 cm to 150 cm streambed depth for improving the understanding of hyporheic exchange flow processes. Generally positive VHG indicated a regional dominance of groundwater up-welling. High and temporally variable VHG were used to identify confined conditions caused by low conductivity layers in the subsurface (low connectivity), while locations with lower and temporally less variable VHG indicated free groundwater up-welling (high connectivity) in highly conductive sediments. A heat pulse sensor (HPS) was applied for identifying shallow hyporheic flow at three locations representative for high versus low streambed connectivity. Shallow hyporheic flow patterns were found to be spatially heterogeneous. Subsurface flow could only partially be explained by streambed topography. Surface water infiltration and horizontal flow coincided with inhibited groundwater up-welling, whereas locations with high streambed connectivity were characterized by increased up-welling. The combined information of spatiotemporal VHG variability and flow vector frequency distribution by HPS has the potential to improve the understanding of impacts of streambed topography and subsurface stratification on hyporheic flow patterns.

5.1 Introduction

The understanding of groundwater and surface water systems in hydrological sciences has experienced a paradigm shift in recent decades, progressing from defining rivers and aquifers as discrete, separate entities toward an understanding of groundwater and surface water as integral components of an aquifer-river continuum with strong mutual influences between river, aquifer, and the interconnecting hyporheic zone (HZ) [Brunke and Gonser, 1997, Bencala, 1993, 2000, Sophocleous, 2002, Krause et al., 2011b].

Exchange fluxes of water, solutes, and heat between surface and groundwater environments strongly impact biogeochemical and ecohydrological processes in the HZ [Dole-Olivier et al., 1997, Malard et al., 2003, Datry et al., 2005, Boulton et al., 1998, 2008]. The HZ, as the saturated interface between surface water and groundwater, has been recognized as an important buffer and refugial zone which provides fundamental ecosystem services and functioning [Brunke and Gonser, 1997, Boulton, 2007, Krause et al., 2011b]. The often enhanced chemical reactivity found at aquifer-river interfaces [Boulton et al., 1998, Fisher et al., 1998, Mulholland et al., 2000, Krause et al., 2009, Pinay et al., 2009, Lewandowski and Nützmann, 2010] is controlled by (1) steep redox-gradients (2) high abundances of organic matter and microorganisms [Jones et al., 1995, Fisher et al., 1998, Chafiq et al., 1999, Duff and Triska, 1990, 2000, Hinkle et al., 2001, Findlay et al., 2003, Hill and Cardaci, 2004, Storey et al., 2004], and (3) hyporheic flow paths and hyporheic residence times [Bencala, 1993, Jones et al., 1995, Fisher et al., 1998, Duff and Triska, 2000, Zarnetske et al., 2011]. The comprehensive understanding of the ecohydrological and biogeochemical functioning of hyporheic streambed environments, therefore, requires detailed knowledge of magnitude, spatial patterns, and temporal dynamics of groundwater-surface water exchange within the streambed at multiple scales.

Exchange flow at the aquifer-river interface is controlled by hydraulic head gradients and by the hydraulic conductivity of the streambed sediments. These parameters are determined by a variety of processes including streambed geomorphology, such as pool-riffle-step formations [Kasahara and Wondzell, 2003, Storey et al., 2003, Boano et al., 2007, Käser et al., 2009] and stream meanders [Boano et al., 2006, Cardenas, 2009], the hydrogeological setting [Fleckenstein et al., 2006, Frei et al., 2009], natural and artificial flow obstacles [Kasahara and Hill, 2008], and microtopography induced advective pumping [Thibodeaux and Boyle, 1987, Boano et al., 2007, Cardenas and Wilson, 2007b,a, Tonina and Buffington, 2007]. Drivers and controls

of exchange fluxes at the aquifer-river interface may act on variable spatiotemporal scales, resulting in exchange flow patterns that are often spatially very complex and vary from small (cm) to large (km) scales [White, 1993, Sophocleous, 2002, Storey et al., 2003, Krause et al., 2011b]. The experimental investigations of such complex flow patterns, therefore, require multidimensional exploration methods, covering a range of spatial and temporal scales [Palmer, 1993, White, 1993, Krause et al., 2011b]. Under field conditions the identification of potentially interacting processes controlling aquifer-river exchange fluxes and flow patterns in particular at small scales and in superficial sediments remains a challenge [Endreny and Lautz, 2012, Krause et al., 2012b,a, Munz et al., 2011].

5.1.1 Streambed stratification versus topography controls on hyporheic exchange flow patterns

Hyporheic exchange fluxes at stream reach scales have been frequently described as being driven by the variability in pressure distributions in relation to streambed topography and channel bed form [Boano et al., 2007, Cardenas and Wilson, 2007b,a, Krause et al., 2011b,a]. Experimental [Lautz et al., 2010, Endreny et al., 2011] and model-based investigations [Boano et al., 2007, Cardenas and Wilson, 2007b,a] found that spatial patterns of hyporheic exchange were strongly influenced by dynamic pressure fields and advective pumping related to streambed topography and surface flow turbulences [Thibodeaux and Boyle, 1987, Elliott and Brooks, 1997a,b].

Previous investigations of the hydrogeological heterogeneity in lowland river streambeds found that the architecture of streambed strata can also significantly impact on groundwater-surface water exchange via the aquifer-river interface [Fleckenstein et al., 2006, Schornberg et al., 2010, Krause et al., 2012a]. At the River Tern (Fig. 5.1), a characteristic lowland river in the UK, recent heat tracer studies at multiple scales, covering a pool-riffle-pool sequence [Krause et al., 2011a] and a larger 250 m long stream reach [Krause et al., 2012a], identified spatially heterogeneous patterns in exchange fluxes between aquifer and river to be controlled by the impact of streambed topography as well as streambed hydraulic conductivity. The combination of vertical hydraulic gradient (VHG) observations and fiber-optic distributed temperature sensing (FO-DTS) at the streambed [Krause et al., 2012a] provided evidence that groundwater up-welling patterns at stream reach scale can be strongly controlled by low conductivity peat and clay lenses in the streambed

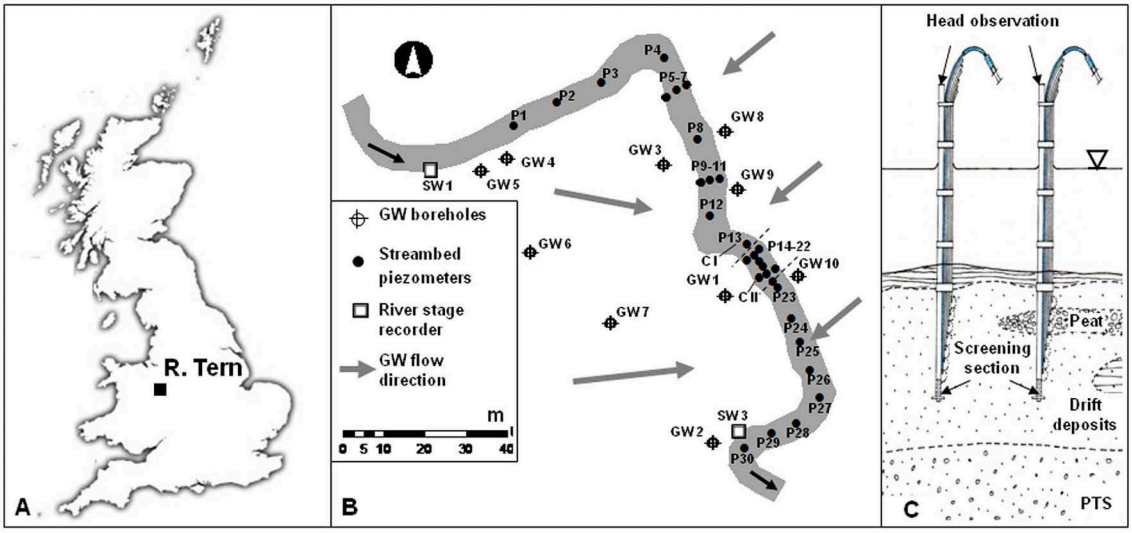

Figure 5.1: A: Location of the River Tern field site in the UK. B: Instrumentation at the River Tern with groundwater boreholes (GW), river stage recorders (SW), the streambed piezometer (P) network and indication of the location of selected representative sediment cores CI and CII. C: Piezometer experimental design. PTS means Permotriassic sandstone.

(Fig. 5.2), resulting in up-welling inhibition at locations with flow confining layers but also preferential fluxes of enhanced groundwater up-welling where low conductivity strata were interrupted. Comparable heterogeneity in streambed stratification and conductivity has been found in other lowland river streambeds [Krause et al., 2008, 2011b] where spatially rather homogeneous groundwater up-welling from non-fractured aquifers is modulated by the spatially heterogeneous hydraulic conductivities of drift deposits in the streambed. However, although the combination of state of the art VHG observations, and FO-DTS monitored streambed temperature patterns has been successfully used for identifying groundwater up-welling in response to discrete changes in streambed conductivity (Fig. 5.2), the applied methodologies did not account for possible implications on shallow hyporheic exchange flow patterns. As aforementioned, shallow hyporheic exchange flow in the uppermost streambed sediments, describing the downwelling and re-emerging of surface water from the hyporheic zone can be driven by multiple processes at a variety of spatial and temporal scales and is often characterized by a strong horizontal flow component. It may interfere or superimpose the rather vertical transit flow of for instance up-welling groundwater resulting in complex, three-dimensional flow pattern at the aquifer-river interface. As shallow hyporheic exchange is controlled by hydrostatic pressure head distributions and hydrodynamic forcing as advective pumping, it is

likely that in particular large contrasts in groundwater up-welling patterns can have a significant effect on the spatial occurrence and extend of shallow hyporheic fluxes.

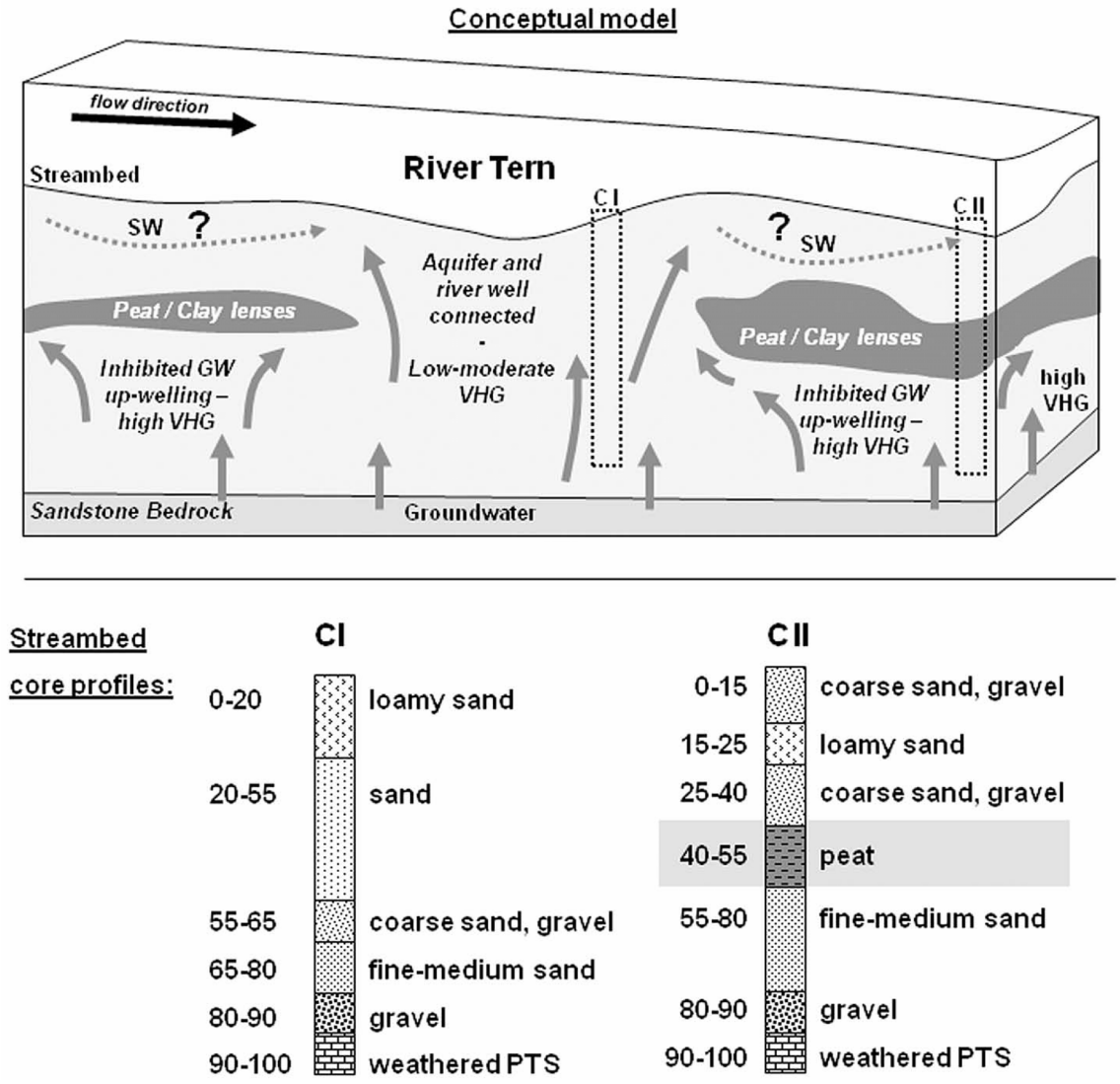

Figure 5.2: Conceptual model of streambed hydrofacies impact on groundwater (GW) up-welling and hyporheic flow of surface water (SW) at the River Tern (top). Streambed core profiles (in cm) of selected representative streambed cores (for sampling locations see Fig. 5.1), PTS means Permotriassic sandstone (bottom).

5.1.2 Aims and objectives

It is hypothesized that a reduction of groundwater up-welling in lowland rivers as caused by flow confining low conductivity streambed strata (peat or clay lenses) several tens of cm below the streambed surface may induce increased surface water downwelling and horizontal hyporheic flow into shallow streambed sediments

above the low conductivity streambed strata. With the reduced forcing of up-welling groundwater, streambed topography induced pressure head distributions and advective pumping would become more important. Therefore, low conductivity streambed strata have the potential to significantly impact on overall transient storage in shallow streambed sediments with potential implications for riverine and riparian biogeochemical and ecohydrological functioning.

The present study demonstrates the combined application of a novel heat pulse tracer methodology and observations of VHG for investigating the extent and spatial variability of shallow hyporheic flow at a lowland river section. The objectives of this study are to:

1. Investigate the spatial patterns of groundwater up-welling for a stream section based on observation of VHG patterns and temporal variability to discriminate between streambed locations with intensive groundwater up-welling (associated with highly conductive streambed strata and lower VHG) versus locations with up-welling inhibition (corresponding with confining streambed strata and increased VHG of higher temporal variability) and provide a selection of representative sites for further investigation of shallow horizontal hyporheic flow.

2. Apply a novel heat pulse sensor (HPS) for investigating shallow hyporheic flow in the top 15 cm of the streambed at locations previously identified by VHG observations to represent characteristic conditions for groundwater up-welling and inhibited up-welling.

3. Compare flow vector frequency distributions observed by HPS in shallow streambed sediments of designated locations with VHG derived information on groundwater up-welling to test the hypothesis that local surface water infiltration and an increased hyporheic flow component may be expected at locations of inhibited groundwater up-welling.

4. Revise the conceptual understanding of streambed topography and streambed stratification as drivers of multi-dimensional hyporheic flow processes in lowland rivers.

5.2 Materials and methods

5.2.1 Study area and field site

The field site (2°53′ W, 52°86′ N) is located at the River Tern, an 852 km^2 tributary of the River Severn in the UK (Fig. 5.1A). The local geology is dominated by Permotriassic Sherwood Sandstone, which represents one of the UK's major groundwater aquifers. The investigated stream reach covers an approximately 250 m meandering section of the River Tern with its 5 m to 8 m wide channel and the adjacent floodplain (Fig. 5.1B). The meandering stream section with its streambed topography dominated by pool-riffle-pool sequences and partly vegetated side bars provides a representative example for characteristic lowland rivers in similar geological setting. The River Tern was selected by the UK Natural Environment Research Council (NERC) as a study area representative for lowland sandstone rivers under the Lowland Catchment Research Programme (LOCAR; Wheater and Peach [2004]); and it is monitored by the Environment Agency as part of the Shropshire Groundwater Scheme [Streetly and Shepley, 2002]. Parts of the monitoring infrastructure of these previous projects provide baseline data for this study (groundwater observation boreholes, borehole logs, flow gauge).

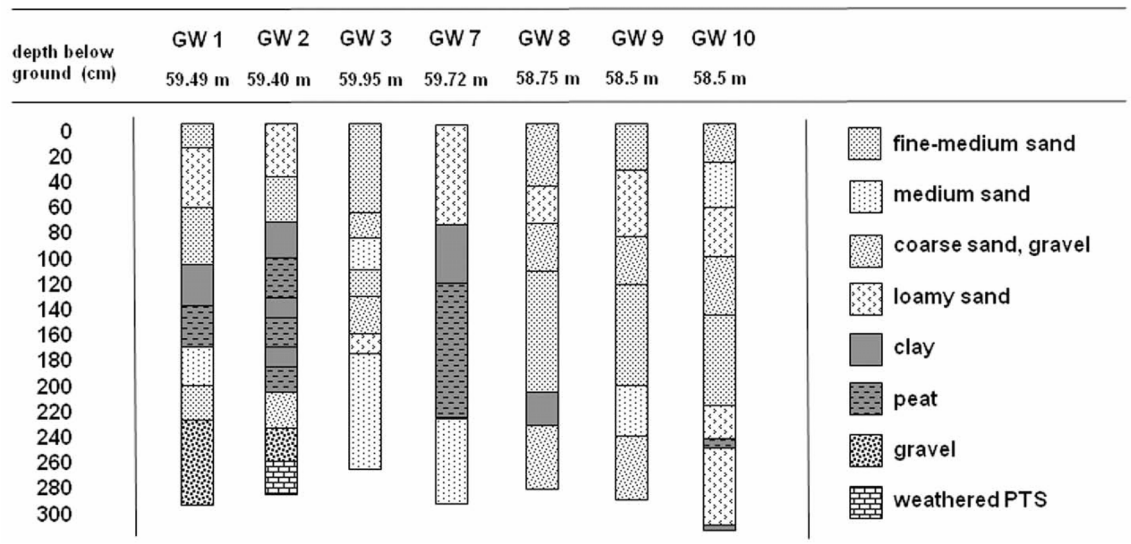

Figure 5.3: Borelogs for seven sediment cores of the riparian floodplain coinciding with the groundwater boreholes (GW1, GW2, GW3, GW7, GW8, GW9, GW10) of Fig. 5.1. The stratification of the cores adjacent to the HPS focus sites (Fig. 5.8) is shown for up to 300 cm below ground with ground levels in meters above sea level indicated for each location.

Riparian and streambed sediments are characterized by intensive spatial variability in material properties and stratification and a wide range of hydraulic conductivities [Krause et al., 2012a]. While the majority of sediments in the research area vary from midsized gravels over different sizes of sands to fine silty material with hydraulic conductivities in the range of $10^{-3}\,\mathrm{m\,s^{-1}}$ to $10^{-5}\,\mathrm{m\,s^{-1}}$, hydraulic conductivity of clay and peat layers in the streambed was significantly lower with $10^{-8}\,\mathrm{m\,s^{-1}}$ to $10^{-9}\,\mathrm{m\,s^{-1}}$ [Krause et al., 2012a]. As coring in the streambed would have permanently disturbed the investigated conditions and in particular altered the impact of peat and clay layers, Fig. 5.2 displays the abundance, thickness, and depth of peat or clay structures in the research area for two example streambed sediment profiles. The high variability of low conductivity peat and clay structures in the riparian drift deposits as result of the postglacial depositional history is shown for seven riparian corelogs in Fig. 5.3. Structure and geometry of the streambed peat or clay layers, which are common streambed features in lowland rivers, have been successfully identified by combined FO-DTS and VHG application at the field site [Krause et al., 2012a]. Given their low hydraulic conductivities, these structures have been shown to have the potential to reduce the regional groundwater up-welling [Krause et al., 2012a], causing flow confinement and potentially increased streambed residence times (Fig. 5.2).

The research of this study focused on summer base flow conditions in July and August 2009. Meteorological data were recorded at the nearby Keele weather station (2°16′12.90″ W, 52°59′55.86″ N). Mean annual precipitation close to the field site is 583 mm, although increased rainfall of up to 740 mm has been observed in the northern headwaters of the River Tern [Hannah et al., 2009]. Precipitation during the observation period was dominated by an extended wet period at the end of July and early August 2009 (Fig. 5.4C) with monthly rainfall exceeding 100 mm. Several short but intensive storm events in August 2009 exceeded $10\,\mathrm{mm\,h^{-1}}$ of precipitation.

Air temperature during the observation period varied by more than 20 °C, with minimum temperatures of 5.2 °C at night in the end of June 2009 and maximum temperatures of 26.4 °C during the day in mid July 2009 (Fig. 5.4A). Diurnal air temperature amplitudes varied substantially with maximum night-day temperature differences of 12.5 °C in June and minimum night-day temperature differences of 2.3 °C in the end of July. Diurnal surface water temperature amplitudes followed air temperature patterns and daily temperature oscillations were slightly dampened

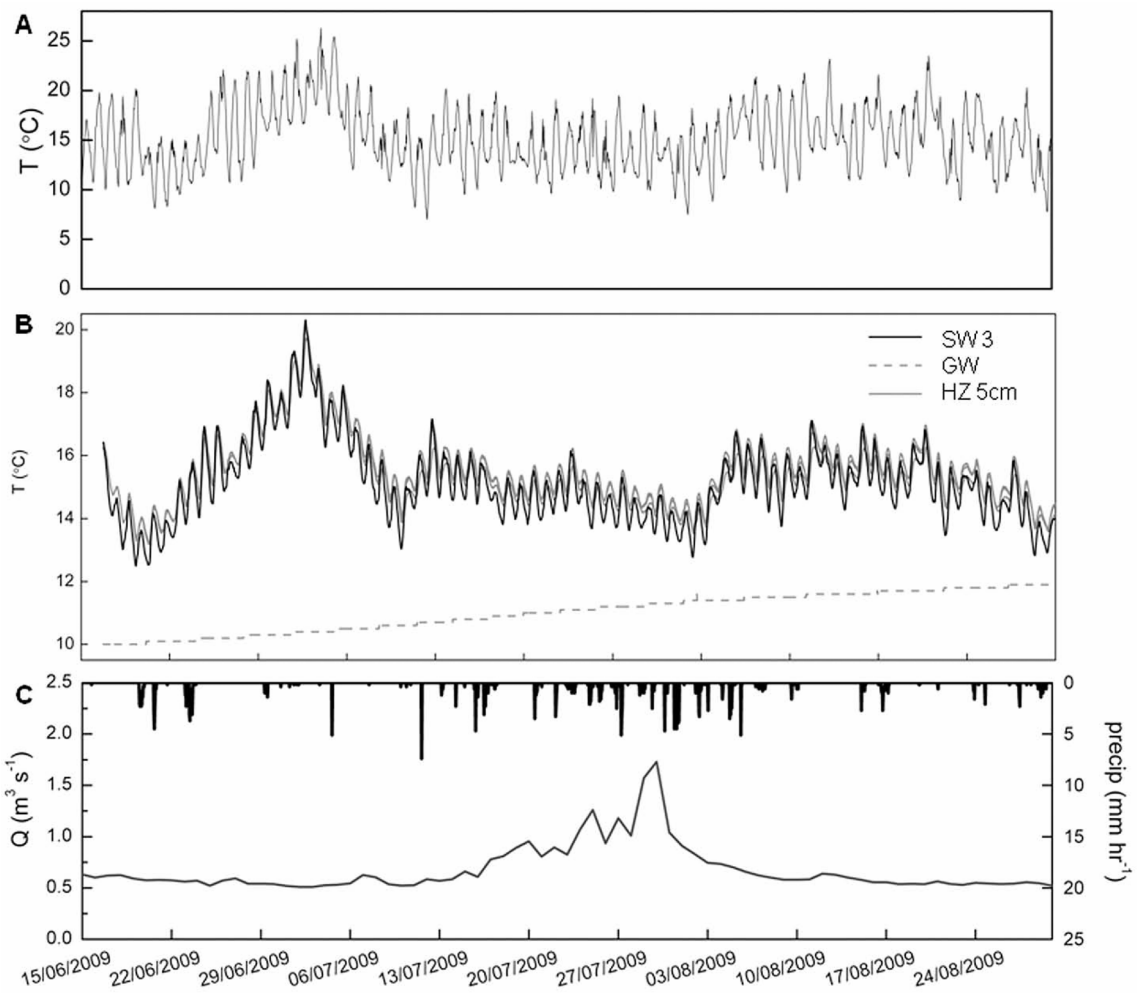

Figure 5.4: Hydrometeorological conditions at the field site with A: air temperatures; B: surface water (SW3), groundwater (GW), and interstitial pore water at 5 cm depth (HZ 5 cm) temperatures; C: precipitation and river discharge (EA Ternhill gauging station) for the period of 15 June 2009 to 30 August 2009.

in shallow streambed sediments (5 cm) and not detectable within the groundwater (Fig. 5.4B).

Stream discharge data were provided by the Environment Agency from the nearby gauging station at Ternhill (2°55′12″ W, 52°87′92″ N) which covers a catchment area of 92 km^2. Mean discharge is 0.9 $m^3 s^{-1}$ with a 95 % exceedance (Q_{95}) of 0.4 $m^3 s^{-1}$ and a 10 % exceedance (Q_{10}) of 1.39 $m^3 s^{-1}$ (data period 1972–2010, UK National River Flow Archive). Summer base flow conditions usually occur from May to October. During the observation period, mean daily river discharge for Ternhill (Q) over the observation period averaged 0.91 $m^3 s^{-1}$ (Fig. 5.4C). Discharge dynamics varied significantly in time, ranging from minimum base flow discharges of 0.51 $m^3 s^{-1}$ to maximum storm discharges of 1.77 $m^3 s^{-1}$ (Fig. 5.4C). The general

summer base flow period was interrupted by a major discharge event in the end of July to early August with >20 days with discharges >1.0 $m^3 s^{-1}$ resulting from the prolonged rainfalls July/August 2009 (Fig. 5.4C).

5.2.2 Data collection

Topographical surveys deploying differential GPS were carried out at the field site in June 2009. They cover the channel, streambed, and riparian zone. The resulting high-resolution digital elevation model of streambed and floodplain topography has a vertical precision of 1 cm and a horizontal precision of 25 cm. Differential GPS was also used for surveying the exact heights of installed groundwater boreholes and piezometers.

During the observation period hydraulic heads and temperatures in groundwater and surface water were automatically recorded by combined pressure transducers and temperature probes (Solinst) which monitor surface water or groundwater head (i.e., water depth) at 5 min to 15 min intervals (Fig. 5.1B). Groundwater boreholes comprised a network of ten 300 cm deep boreholes covering the riparian groundwater within the floodplain drift deposits (Fig. 5.1B). Surface water levels were monitored at a downstream located stream gauge (Fig. 5.1B). Monitored groundwater and surface water pressure heads were corrected for barometric pressure fluctuations using an atmospheric pressure sensor located at groundwater borehole site GW7 (Fig. 5.1B).

Streambed mini piezometers were installed in the streambed sediments at depths between 120 cm and 150 cm below streambed surface (Fig. 5.1B and 5.1C) to observe interstitial pore water pressure head distributions. They were installed in a longitudinal transect along the stream reach with several cross-sectional extensions toward the river banks (Fig. 5.1B). Hydraulic heads in streambed piezometers were monitored manually, on an approximately fortnight basis throughout June and August 2009 using a graduated electric contact meter (dipmeter). In order to provide quality assurance for automatically logged pressure heads, manual dipmeter sampling was also carried out at the network of shallow riparian groundwater boreholes. As piezometers were installed to depths of 120 cm to 150 cm within the streambed, it is assumed that they are providing general information on aquifer-river exchange flow but do not account for obstacle driven or hydraulic pumping controlled shallow hyporheic exchange fluxes (Fig. 5.1C).

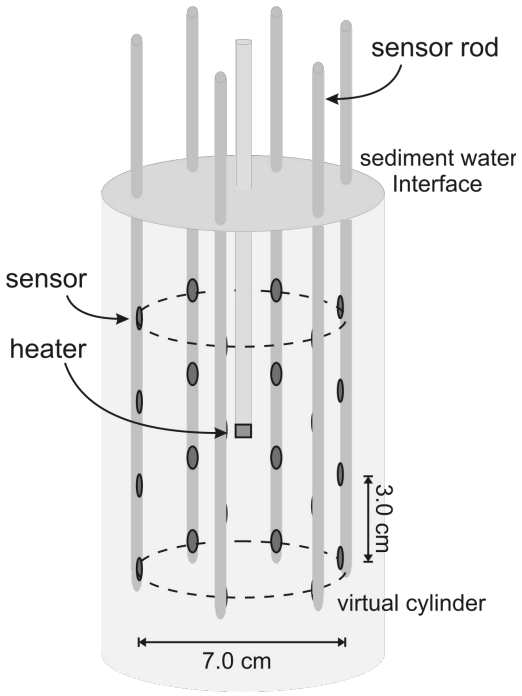

Figure 5.5: Instrumental setup of the heat pulse sensor. The heat source in the center is surrounded by an array of 24 temperature sensors, tracing the propagation of the heat pulse. Heater and sensors are attached to rods which are positioned and stabilized by a frame above the sediment water interface (not shown).

Vertical hydraulic gradients, indicating the potential strength and direction of exchange fluxes between groundwater and surface water, were determined from hydraulic head measurements in the streambed. VHG were calculated by $\Delta h/\Delta l$, with Δh given by the elevation difference of the water tables observed inside and outside the piezometer and Δl given by the distance between the midscreen depth and the surface of the water-sediment interface. The accuracy of dipmeter based hydraulic head observations was approximately 63 mm head and accounts for uncertainties in the measurements introduced by turbulent flow conditions around the piezometers, which can affect the outside head estimates [Krause et al., 2009, Käser et al., 2009].

Although VHG measurements without further knowledge of the streambed hydraulic conductivity represent poor indicators of aquifer-river exchange fluxes, in combination with FO-DTS observed streambed temperature anomalies, variable patterns in streambed VHG at the study site have been proven to successfully indicate spatial patterns in streambed sediment conductivities, which vary by five orders of magnitude [Krause et al., 2012a].

By comparative analysis with FO-DTS observed streambed temperature anomalies, Krause et al. [2012a] found that increased VHG (>0.4) correlate with areas of inhibited groundwater up-welling underneath flow confining streambed structures of low hydraulic conductivity (Fig. 5.2) and low VHG (<0.4) coincide with highly conductive streambed sediments. The combined application of VHG observations

and FO-DTS monitored temperature patterns lead to the development of a conceptual model of vertical exchange fluxes between aquifer and river [Krause et al., 2012a] that is believed to be representative for lowland rivers such as the River Tern (Fig. 5.2). It assumes that within regionally rather homogeneous groundwater flow fields, strength and temporal variability of VHG can present an indicator of the contrasting streambed conductivity patterns with inhibited groundwater up-welling due to flow-confining streambed strata and preferential groundwater up-welling at locations with high groundwater-surface water connectivity. However, VHG interpretations consider groundwater-surface water exchange processes as one dimensional. They give evidence about connectivity and interaction between river and aquifer only and do not provide any indication of potential shallow hyporheic exchange with possible increased horizontal flow in streambed sediments above flow-confining low conductivity layers.

Therefore, measurements with a specially designed heat pulse sensor were conducted to extend process investigations into the shallow subsurface not covered by other methods. The HPS enables for the three-dimensional analysis of small-scale fluxes in the shallow subsurface of streambed sediments. It has been specifically developed for application in the hyporheic zone of sand-bed streams and covers a scale of few cm (diameter of investigated sediment cylinder 7 cm, Fig. 5.5). The measuring principle is based on a short heat pulse emitted by a point source directly into the shallow sediment in 5 cm to 10 cm depth. Heating was applied with a power of 12 W for an interval of 60 s, resulting in an energy input of 720 J [Lewandowski et al., 2011]. The movement of the heat pulse through the sediment is traced by an array of 24 thermocouple-sensors surrounding the heat source three dimensionally. Each four sensors are attached to one separate sensor rod (Fig. 5.5). This setup allows measuring hyporheic flow directly in the sediments with minimal disturbance of the sediment specimen. The peak temperature increase ΔT_{max} at the sensors was 0.22 °C to 1.20 °C depending on flow velocity. Temperature changes were measured with an accuracy of approximately 0.005 °C. The resulting breakthrough curves are analyzed with an analytical solution of the convective-dispersive heat transport equation in a cylindrical coordinate system (Eq. 5.1) and provide evidence of pore water flow direction and velocity in the investigated streambed sediments [Angermann et al., 2012]:

$$\Delta T_{(d,r,t)} = \frac{Q_i}{8\pi \cdot \frac{D_t}{R} \cdot t \cdot \rho_B c_B \cdot \sqrt{\pi \frac{D_t}{R} t}} \cdot exp\left(-\frac{R(d - tv_w)^2}{4D_l t} - \frac{Rr^2}{4D_t t}\right) \tag{5.1}$$

Within equation 5.1 ΔT (°C) is the temperature increase relative to the ambient streambed temperature. d and r (m) are the longitudinal and radial coordinates of the cylindrical coordinate system that give the sensor's position relative to the fitted flow direction. Q_i is the thermal energy emitted by the heat source (J). For simplicity, the volumetric heat capacity of the bulk material $\rho_B c_B$ and the retardation factor R were assumed to be constant throughout the investigation site and defined to equal medium sand with a small fraction of organic matter. $\rho_B c_B$ was $3.3 \times 10^6\,\mathrm{J\,m^{-3}\,°C^{-1}}$, R was 1.7. The time after heat pulse injection is given with t (s). The longitudinal and transversal dispersion coefficients D_l and D_t ($\mathrm{m^2\,s^{-1}}$) summarize the effects of dispersion and heat conduction. Both coefficients are velocity dependent and generally defined as the products of velocity and longitudinal and transversal dispersivity (m), respectively. Dispersivities are specific to the substrate and hard to quantify. They are fitted during the analysis procedure [Angermann et al., 2012] and mean values were found to be 0.25 m (longitudinal, with a standard deviation of 0.43) and 0.11 m (transversal, with a standard deviation of 0.14). v_w ($\mathrm{m\,s^{-1}}$) is the effective flow velocity (sometimes also called pore velocity; Darcy velocity divided by porosity).

As discussed in detail by Lewandowski et al. [2011] and Angermann et al. [2012], some limitations of the methodology originate from the geometrical setup of the thermistors, and in particular the cylindrical alignment of the sensors around the heat source. As result of this setup, the propagation of the heat pulse is not covered ideally for vertical flow. When the vertical angle exceeds ±50 °C, only the periphery of the heat pulse is recorded instead of its center. Flow velocity and direction have to be calculated from the break-through curves of the periphery of the heat pulse and the accuracy of the method is reduced from 5 °C in horizontal flow to 20 °C in vertical flow direction. However, this precision is assumed to be sufficient for *in situ* measurements so that all data were included in the evaluation of the measurement results.

The analysis routine requires precise positions of the temperature sensors relative to the heat source, as errors in distances lead to a misinterpretation of the breakthrough curves. In comparison to the device presented by Lewandowski et al.

[2011] and Angermann et al. [2012], who positioned a single sensor rod with a stencil, in the present study these rods were assembled to one device in order to cope with higher water depths. This setup ensures very precise positioning, however, bending of single sensor rods due to subsurface obstacles or coarse substrate will not be prevented. Visual observation is necessary to prevent wrong measurements, and substrate properties limit the applicability of the device.

A further problem exists at very low velocities, when the effect of heat conduction may exceed the convective heat transport. In such conditions it is not possible to discriminate both processes and quantify flow velocity with acceptable accuracy anymore. The threshold, i.e., the detection limit of the method, depends on the water content and the thermal properties of the sediment and can be recognized by the observation of the fitted parameters during data analysis [Angermann et al., 2012]. According to this procedure, the detection limit was identified to be $1.0 \times 10^{-5}\,\mathrm{m\,s^{-1}}$. 21 % of the analyzed heat pulse measurements at River Tern were below this threshold and are therefore of limited reliability but still considered in the analysis. Above this threshold, the method performs with an accuracy of approximately 10 % in flow velocity.

Heat pulse measurements were conducted at designated streambed sections of approximately 10 m length which were identified by VHG observations to be representative for characteristic streambed conditions, ranging from high aquifer-river conductivity to low conductivity and up-welling inhibition due to flow confining streambed structures. At each of the sites 18 to 21 HPS measurements were conducted. Measurements were carried out between 21.07.2009 and 19.08.2009 (Tab. 5.1) and arranged along several transects (named TR1 to TR7, enumerated upstream to downstream), most of them oriented across the streambed to cover the whole site. With this monitoring setup, the HPS observations aimed to provide a representative sample of the characteristic flow vector frequency distribution at the specific streambed locations but did not aim to monitor the entire hyporheic flow network in the research area which would have required a significantly higher number of HPS measurements.

5.3 Results and discussion

5.3.1 Hydraulic head patterns in groundwater and surface water

Temporal variability in time series of surface water levels observed at stream gauges coincided with groundwater levels monitored at groundwater observation boreholes (Fig. 5.6). Water levels at selected riparian groundwater boreholes (GW1, GW3, GW7) and river gauges (SW3) were generally low during the base flow conditions with average levels ranging between 58.6 m to 58.9 m asl (above sea level) (Fig. 5.6). Base flow conditions were interrupted by a three-week episode of increased groundwater and surface water levels reaching values of up to 59.38 m asl (surface water) and 59.27 m asl (groundwater) in July 2009 (Fig. 5.6), caused by precipitation events as seen in Fig. 5.4B.

Table 5.1: Timing of HPS Measurements in 2009[a]

Date		HPS Measurements
21	July 2009	3.1–3.7
29	July 2009	2.16
31	July 2009	2.10–2.14
3	August 2009	2.3–2.9
6	August 2009	2.15, 2.17, 2.18
7	August 2009	2.1, 2.2
10	August 2009	3.8–3.13
12	August 2009	3.14–3.21
13	August 2009	1.10–1.17
19	August 2009	1.1–1.9

[a] For respective meteorological and hydrological conditions please see Fig. 5.4 and 5.6.

Groundwater levels in the monitored boreholes GW1 and GW3 close to the stream as well as GW7 at further distance (Fig. 5.1B) exceeded surface water levels for most of the observation period, indicating a general flow direction from riparian groundwater toward the stream. Inverse head gradients, indicating reversed flow conditions were only observed during the July/August storm event (Fig. 5.4C), when surface water levels rose faster and higher than associated groundwater levels (Fig. 5.6), causing surface water infiltration into the riparian groundwater. However, such flow inversions did not reach the central parts of the floodplain as indicated at GW7.

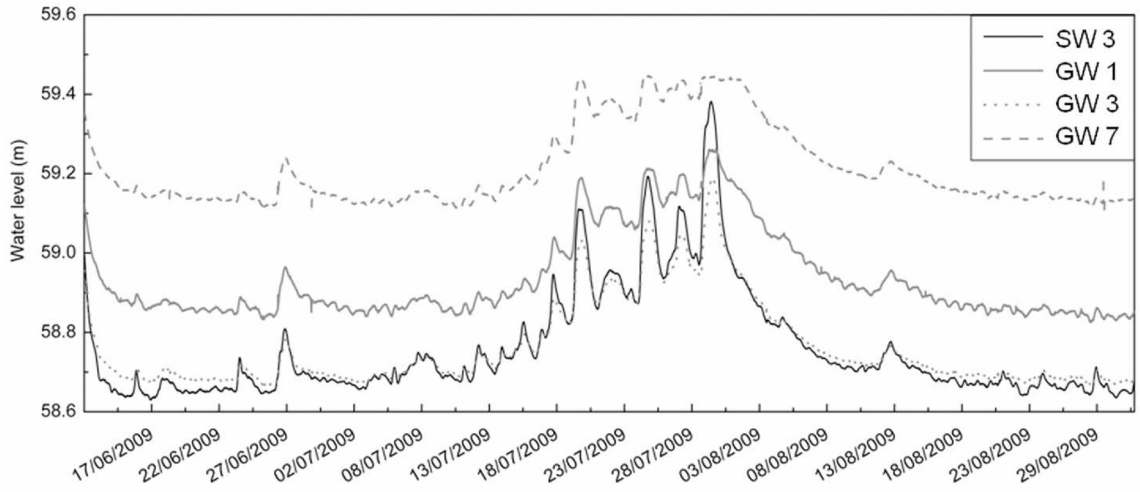

Figure 5.6: Surface water (SW3) and selected groundwater levels (GW1, GW3, GW7) at the River Tern field site (Fig. 5.1) for the investigation period from 15 June 2009 to 30 August 2009.

5.3.2 VHG for characterizing spatial patterns in groundwater up-welling

Vertical hydraulic gradients observed at streambed piezometers were positive throughout the observation period, indicating a constant up-welling of groundwater into the investigated stream reach. With values ranging between 0.08 and 0.8, VHG in the research area were found to be spatially highly variable (Fig. 5.7). Although absolute values of VHG at streambed piezometer locations varied at the three monitoring dates, spatial patterns remained rather constant throughout the observation period (Fig. 5.7). The spatially and temporally very homogeneous VHG in the east-west oriented up-stream (P1-P3) and down-stream (P28-P30) river sections (5.1B) were generally low with values never exceeding 0.2. In contrast, within the north-south oriented central section of the investigated stream reach (P4-P27), VHG were spatially more variable with some dominant hotspots of increased VHG reaching values of up to 0.8 around the central section (P14-P22).

The observed complex spatial patterns in monitored VHG are likely to result from the wide range and high spatial variability in hydraulic conductivities of the drift deposits forming the streambed sediments of the research area. As the sediment properties of the nonfractured Permotriassic sandstone aquifer are spatially very homogeneous at the investigation scale, observed VHG patterns are more likely to relate to variable material properties of the drift sediments than to result from local or regional groundwater flow variability. Therefore, VHG observations of spa-

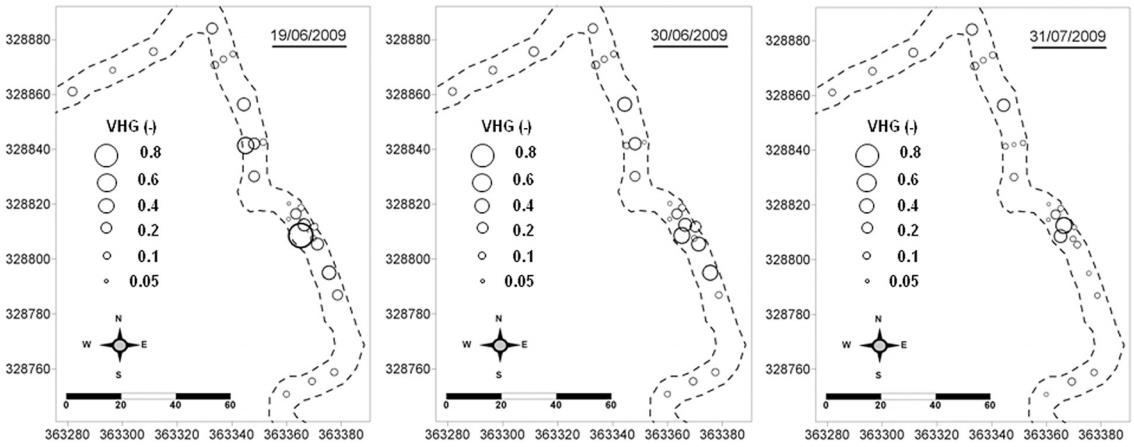

Figure 5.7: Spatial patterns of observed vertical hydraulic gradients in the River Tern streambed for three selected sampling dates (19 June 2009, 30 June 2009, and 31 July 2009) during the 15 June 2009 to 30 August 2009 observation period.

tially isolated particularly high values of up to 0.8 (as in the central stream section) were interpreted as indicator of local inhibition of groundwater up-welling by flow confining streambed peat and clay lenses, whereas up-welling inhibition was considered to be less likely at locations with lower and spatially more homogeneous VHG (as for instance in the most up-stream and down-stream sections of the study area). VHG within the central river section expressed higher temporal variability with values varying by up to 0.3 between monitoring dates, whereas VHG at piezometer locations in the most up-stream and most down-stream river sections did not vary by more than 0.05 for different observation dates. Furthermore, given the depth of piezometer screening sections of between 120 cm and 150 cm below the streambed surface, a relevant impact of streambed topography induced advective pumping effects on the observed head gradients is not likely.

Higher temporal variability in VHG within the central river section correlated with meteorological conditions, suggesting a higher impact of hydroclimatological forcing at the respective piezometer locations. As discussed in comparison with FO-DTS observed streambed temperatures by Krause et al. [2012a], such dynamics may be explained by the fact that flood conditions like at the end of July 2009 (Fig. 5.4C) cause variable responses in groundwater and surface water (Fig. 5.6) with peaks in groundwater heads usually tailing off slower than in surface water due to retention by riparian storage. In highly conductive sediments such induced changes in head differences between groundwater and surface water are expected to cause alteration in exchange fluxes over the aquifer-river interface which quickly equilibrate VHG. In contrast, underneath flow confining streambed structures alter-

ation of VHG is likely to be more persistent as head differences are less compensated by exchange fluxes. As confirmed by comparison with streambed temperature observations [Krause et al., 2012a], it can therefore be assumed, that VHG temporal variability qualifies as a further indicator of the existence of flow confining structures (e.g., in the central river section), whereas temporally stable VHG (as in the most up-stream and most down-stream sections) are likely to indicate highly conductive streambed sediments where groundwater-surface water head differences are faster equilibrated by exchange fluxes.

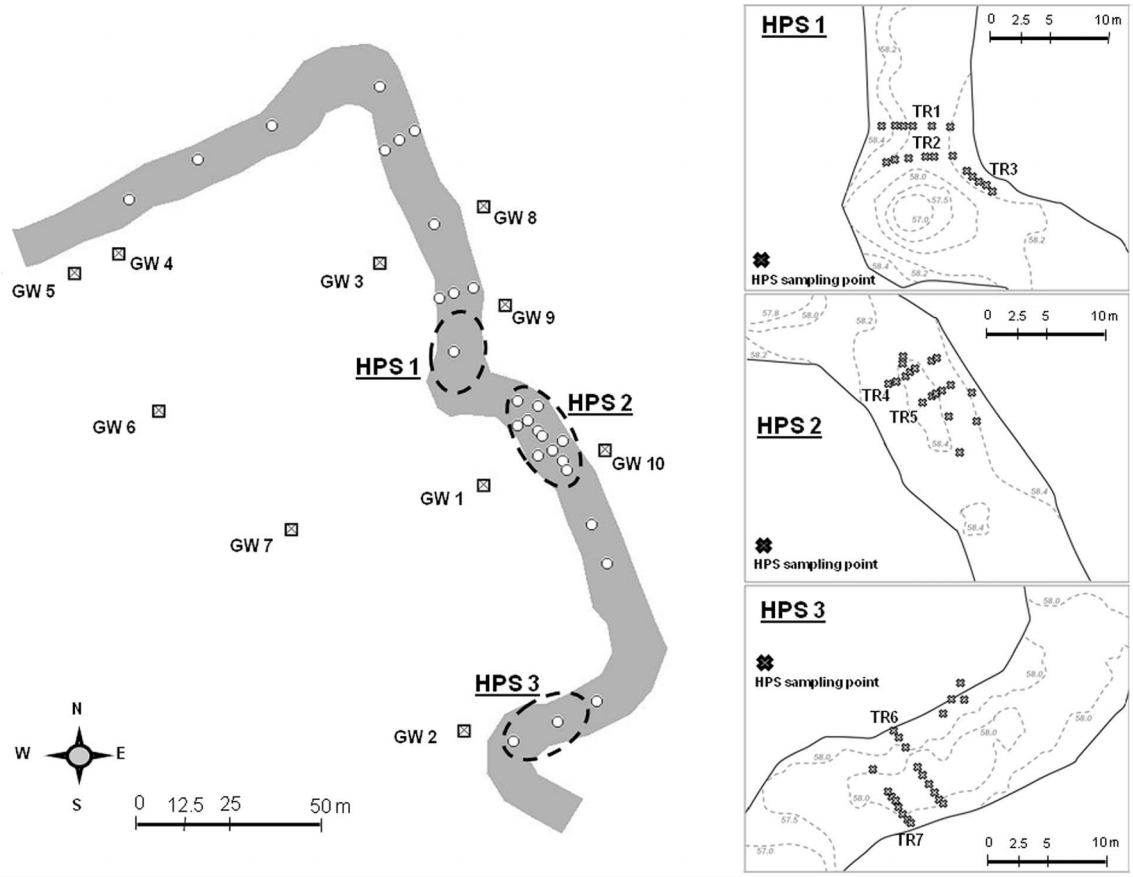

Figure 5.8: (left) Location of the three identified HPS focus sites selected on the basis of characteristic VHG measurements (open circles) to represent distinctive aquifer-river connectivity conditions. (right) Alignment of HPS monitoring points (mostly along transects TR1 through TR7) and bathymetry, indicated by dashed gray lines within the three focus sites.

As vertical hydraulic gradients were determined by head differences between river and 120 cm to 150 cm depth within the streambed, they do not account for potential local shallow hyporheic exchange as for instance streambed topography induced surface water downwelling and mixing with groundwater. In order to test if shallow horizontal hyporheic fluxes are more pronounced at locations where the

forcing of up-welling groundwater is inhibited, three key locations were selected based on VHG observations (Fig. 5.7) in order to represent different streambed architecture and related exchange fluxes (Fig. 5.8):

Site HPS1 – representing an isolated location of moderate and temporally stable VHG (suggesting groundwater up-welling through relatively conductive sediments) which is surrounded by up-stream and down-stream sections of high VHG suggesting flow confinement (Fig. 5.7). HPS sampling points are located in direct vicinity of a major pool with >2 m water depth (Fig. 5.8).

Site HPS2 – representing an area of complex VHG patterns with the highest and temporally most dynamic VHG's observed in close vicinity to locations with lower and temporally more stable VHG, suggesting spatially very heterogeneous streambed properties with low conductive and highly conductive sediments. Parts of the streambed are characterized by a pool-riffle-pool sequence (Fig. 5.8) of which a longitudinal transect has been subject of investigation by Krause et al. [2011a].

Site HPS3 – representing a larger area (>20 m stream length) of low and temporally relatively stable VHG, suggesting intensive groundwater up-welling through highly conductive streambed sediments. The center of HPS3 covers an extended central riffle section and a partly vegetated mid-channel bar (Fig. 5.8).

5.3.3 HPS monitoring

The applied HPS is a new technology that has only been tested under laboratory conditions and in the sandy streambed of the River Schlaube (Germany) prior to this study [Angermann et al., 2012]. In addition to general limitations of the HPS heat sensor technology, size and site conditions at the investigated stream reach of the River Tern represented some practical challenges for the application and data analysis of the HPS.

During the investigations at the River Tern, measurements in some locations were inhibited or disturbed by high water levels, coarse substrate, macrophyte growth, and high surface flow velocities. The HPS was deployed into the sediment manually and its correct position and the occurrence of sediment transport due to turbulence caused by the installation were observed via a hydroscope. Inclination of the device and the disturbance of the measurement by too high surface flow velocities could be identified and resulting errors accounted for. Areas with the above mentioned distinct properties could not be sampled and had to be ignored even though they

might be a characteristic part of the investigated site, like, e.g., the deep and coarse thalweg at site HPS2. The negligence of entire structural units of a river section might affect the validity of the obtained flow vector frequency distribution.

Due to the size and heterogeneity of the investigated river section, heat pulse measurements were conducted in different substrates reaching from fine sand at site HPS1 and the left bank of HPS2 over medium sand at HPS3 and the sand bank of HPS2 to coarse sand and fine gravel in the thalweg and downstream of HPS2. Substrate conditions at HPS1 and HPS3 were relatively homogeneous while HPS2 was characterized by a high diversity of substrate. Varying sediment types entail different parameter sets required for the calculation of hyporheic flow velocities, i.e., the volumetric heat capacity of the sediment $\rho_B c_B$ and the retardation factor R. As it is practically not achievable to accurately quantify all of the required parameters for the entire area, as a simplification, average conditions for all parameter sets (equaling conditions of medium sands) were assumed for all measurements conducted in this study. The parameters used in this study where $3.3 \times 10^6\,\mathrm{J\,m^{-3}\,°C^{-1}}$ for volumetric heat capacity and a retardation factor of 1.7. This parameter set equals conditions of medium sand with a small fraction of organic matter, while the range of possible values for the sediment types present in the investigated sites is approximately $2.6 \times 10^6\,\mathrm{J\,m^{-3}\,°C^{-1}}$ to $3.6 \times 10^6\,\mathrm{J\,m^{-3}\,°C^{-1}}$ for volumetric heat capacity (dependent on water and organic matter content) and 1.6 to 2.6 for the retardation factor, respectively. Such simplification has the potential to cause minor errors in flow velocity estimates and reduces the comparability between measurements of different sites, in particular if site specific fluxes differ only insignificantly. However, this limitation affects mainly site HPS2 with its spatially heterogeneous conditions while the comparability within the more homogeneous sites HPS1 and HPS3 is not compromised.

5.3.4 HPS for identifying shallow hyporheic flow paths

As site HPS1 was located close to a pool of more than 2 m depth and as the HPS device had to be manually deployed at the sediment, experiments were restricted to locations where surface water depths were not exceeding 120 cm. Thus, heat pulse measurements were conducted along two cross-sectional transects up-stream of the pool (TR1 and TR2, Fig. 5.8) and one longitudinal transect TR3 parallel to the eastern river bank (in flow direction).

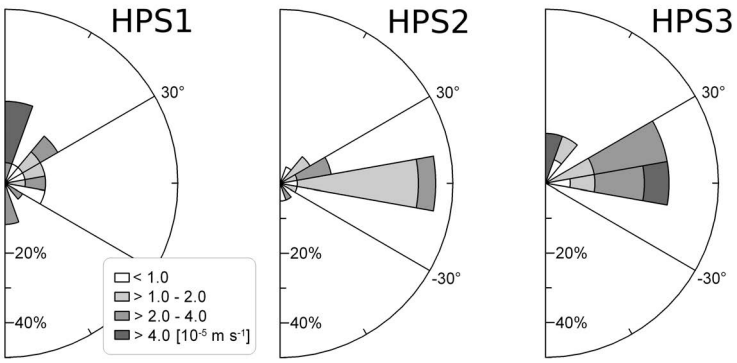

Figure 5.9: Relative frequencies (percentage of all measurements at one site (%), radial axis) of absolute hyporheic flow velocities (gray scale) and vertical angles (in 20° clusters) measured with the heat pulse sensor at the three HPS sites.

Site HPS1 was found to be dominated by upward directed flow (Fig. 5.9). Highest flow velocities ($>4 \times 10^{-5}\,\mathrm{m\,s^{-1}}$) occurred in upward direction. 24 % of the measurements showed angles steeper than +70° relative to a horizontal plane and 47 % of the measurements showed a vertical angle β steeper than +30°, whereas only 18 % showed clear surface water infiltration (i.e., $\beta < -30°$). Nevertheless, 35 % of the measured flow directions were between −30° and +30° (Fig. 5.9) and thus, were categorized as horizontal flow. Hence, although the up-welling was dominant at the study site, horizontal flow was found to be still relevant. Figure 5.10 shows the horizontal orientation of hyporheic flow which in case of HPS1 is dominated by a clockwise deviation from the reference flow direction (i.e., the approximated main surface flow direction).

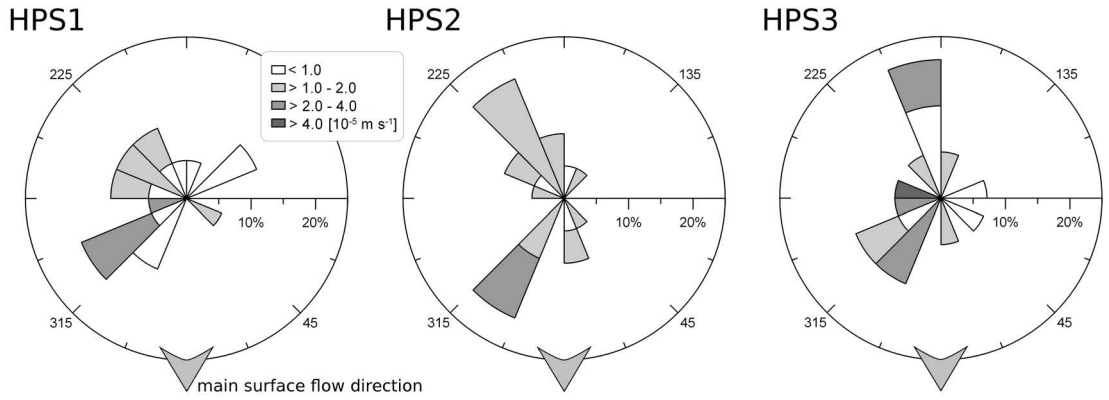

Figure 5.10: Relative frequencies (percentage of the whole amount of measurements at one site (%), radial axis) of horizontal hyporheic flow components, i.e., the horizontal projections of flow vectors. The gray scale indicates horizontal velocity and the angles show the horizontal deviation (in 20° clusters) from the main surface flow direction, acting as reference for every single HPS site.

The HPS measurements indicated that the downstream tailing end and the thalweg side of a sand bar located directly up-stream of the pool (Fig. 5.11) were char-

acterized by up-welling with angles of +38° (measurement 1.2 in transect TR1, Fig. 5.11) and close to +90° (+80° and +88° , measurements 1.3 and 1.4 of the same transect). Measurements 1.5 and 1.6 of transect TR1 were located in the thalweg and showed horizontal flow parallel to surface flow, whereas the measurements 1.7, 1.8, and 1.9 of transect TR2 indicated surface water infiltration into the sediment of the tailing end of the sand bank perpendicular to surface flow direction (−26° , −75° and −73°). Heat pulse measurements 1.10, 1.11, and 1.12 of transect TR2, located in the thalweg and at the Eastern bank, as well as 1.13 to 1.17 along the longitudinal transect TR3 close to the Eastern bank indicated lateral groundwater exfiltration from the bank into the river (Fig. 5.11).

The observed patterns of surface water downwelling into flow obstructions beside the thalweg, predominantly along the Western sand bar as well as groundwater up-welling from the down-stream Eastern sandy river bank can be largely explained by topography induced pressure field variation.

Due to coarser substrate in the area of the main thalweg close to the right bank at site HPS2, heat pulse measurements were restricted to manageable sediment conditions where the HPS could be inserted into the streambed. Hence, HPS measurements were concentrated on the remaining area of site HPS2, which was characterized by a pool-riffle-pool section with a sandy central riffle bar which separated the main (Western) thalweg from a minor, secondary thalweg on the eastern side of the stream. Downstream of the site, the streambed widened with a general flattening of the streambed topography (Fig. 5.8 and 5.11).

Hyporheic flow velocities at site HPS2 were found to be the lowest of all three investigated sites. The average flow velocity at site HPS2 was $1.6 \times 10^{-5}\,\mathrm{m\,s^{-1}}$ in comparison to $2.4 \times 10^{-5}\,\mathrm{m\,s^{-1}}$ and $2.8 \times 10^{-5}\,\mathrm{m\,s^{-1}}$ at site HPS1 and HPS3, respectively. Flow in the streambed was dominated by horizontal fluxes (72 %), with only 11 % of downward and 17 % of upward flow (Fig. 5.9). Slow horizontal flow with vertical angles β around 0° was found in front of the upstream crest of a sandy riffle around the central part of transect TR4 (Fig. 5.11). At these locations, the main flow direction in the streambed was surprisingly found to be directed against the surface water flow direction with a slight tendency toward the main thalweg (measurements 2.1 and 2.2; 2.5, 2.6, and 2.7 of transect TR4 and 2.10 of transect TR5). The further downstream part of the streambed including transect TR5 was characterized by an area of moderate up-welling, with angles between 14° and 36° (measurements 2.11, 2.12, 2.13 of transect TR5, and 2.16). This area included the

central section of the sandy riffle between main and secondary thalweg. Negative vertical angles indicating downwelling fluxes of surface water were mainly found within the area of the secondary thalweg (measurements 2.9, 2.14, 2.17, and 2.18).

Low hyporheic flow velocities at a central sand bank in contrast to downwelling in the secondary thalweg, as shown in Fig. 5.11, were also reported by Angermann et al. [2012] and Rosenberry and Pitlick [2009]. Both reported the same relation between hyporheic flow in sand bank and thalweg in different rivers. These patterns are indicative for the impact of surface flow velocities on hyporheic flow, dominating over the pressure of up-welling groundwater. The moderate up-welling at the central crest and the Western area of the sandy riffle coincided with expectations of dominant flow directions caused by topography induced hyporheic fluxes. However, observations of dominant groundwater up-welling at the head of the riffle section and downwelling down-stream of the riffle tail at the field site did not coincide with the previously described typical exchange flow patterns caused by topography induced pressure field distributions [Kasahara and Wondzell, 2003, Boano et al., 2007, Tonina and Buffington, 2007, Kasahara and Hill, 2008]. VHG observations at site HPS2 indicated significantly increased VHG in piezometers P19-P23 (Fig. 5.1B and 5.7) which has been interpreted as result of confined groundwater conditions under flow confining streambed peat and clay strata. These locations spatially coincided with the HPS measurements at transect TR5 and downstream (Fig. 5.8). The coincidence of VHG evidence for confined up-welling and HPS observed downward fluxes provided strong evidence that the existence of low conductivity streambed strata in this area induced increased surface water infiltration and horizontal flow and thus increased shallow hyporheic transient storage. The observed hyporheic fluxes of opposite direction than the surface water flow in front of the riffle head cannot be explained by streambed topography but were related to preferential groundwater up-welling in the neighborhood of flow confining streambed strata. This implies that the preferential groundwater up-welling acted against and exceeded the topography induced flow at the riffle head.

Overall, spatial patterns of streambed flow at site HPS2 can only partially be explained by streambed topography induced pressure field distribution but, in comparison with VHG observations, provide evidence for the strong impact of streambed stratification on hyporheic exchange fluxes, including preferential surface water infiltration and horizontal fluxes above flow confining low conductivity streambed strata.

Site HPS3, located in the most down-stream part of the research area, was confined by two sharp river bends. The site was a characterized by a clearly defined thalweg of about 1 m water depth on the Southern side, a mid-channel sandy riffle section and two sections of dense macrophyte growth (dominated by Ranunculus) in the central section of the sandy riffle and the sandy Southern river bank (Fig. 5.8 and Fig. 5.11). Dense macrophyte roots partially restricted HPS measurements at some locations by preventing a correct deployment and spacing of the sensors.

Heat pulse measurements in the HPS3 area were conducted during two periods of different flow conditions. Measurements 3.1 to 3.7 at the northern end of transect TR6 (Fig. 5.11) and the up-stream sampling locations were carried out during the major discharge event in the end of July (Fig. 5.4C). Due to high surface flow velocities and water levels, HPS measurements during this period were restricted to an area close to the Northern river bank. The four most upstream measurements 3.1 to 3.4 were located at a gently sloped section of the bank which was not completely submerged during base flow conditions. During flood conditions these sediments were submerged and formed a bay at the concave bank that acted as dead water zone. Streambed pore water fluxes in this area were found to be weakly up-welling, but also showed a strong horizontal flow tendency in opposite direction than surface flow in the main channel. Streambed flow in the northern end of transect TR6 (3.5, 3.6, and 3.7) were weakly up-welling and down-welling at one location with a horizontal flow component following the surface water flow direction. The majority of HPS measurements at site HPS3 were conducted during base flow conditions at the same time as measurements at sites HPS1 and HPS2. Measurements of this period covered the two major transects TR6 and TR7 across the entire river (measurements 3.8 to 3.21). 29 % of these measurements indicated upward flow, 71 % showed horizontal flow whereas no measurement showed angles steeper than $-10°$ downward flow (Fig. 5.9).

The steepest positive vertical angles of $+84°$ and $+64°$ at transect TR6 and $+86°$ and $+62°$ at transect TR7 were found at the slopes from the central riffle toward the thalweg (3.11 and 3.13 at transect TR6, 3.17 and 3.21 at transect TR7), while horizontal flow with moderate positive angles characterized the bottom of the thalweg. Measurement 3.12 located at the deepest point along transect TR6 showed a vertical angle β of $+29°$. The deepest measurement of transect TR7 and the two adjacent measurements, 3.18, 3.19, and 3.20, showed values of $+5°$, $0°$, and $+16°$, respectively.

The general up-welling flow at site HPS3 was influenced by macrophyte growth at the streambed surface, representing a flow barrier in the surface water body, forcing a flow separation into two subchannels. High velocities in the thalweg force water into the sediment and cause a diversion of the vertical flow of up-welling groundwater. Strong horizontal fluxes in the dead water zone during the high discharge period were interpreted as the results of backflow and eddy formation in this area and represent further indication of the impact of streambed morphological structures on hyporheic flow and the propagation of surface flow patterns into the subsurface.

5.3.5 Intersite heterogeneity of shallow hyporheic flow paths

Heat pulse measurements at all three sites detected high heterogeneity of fluxes in the streambed at the investigated spatial scale and shallow hyporheic flow opposed the main surface flow direction in several cases. HPS measurements indicated that, in contrast to flow patterns observed by 120 cm to 150 cm deep streambed piezometers, fluxes in the shallow streambed were not only characterized by groundwater up-welling but complex patterns of upward and downward fluxes as well as a substantial horizontal flow component, indicating a strong impact of streambed topography and heterogeneous patterns of streambed hydraulic conductivity on hyporheic exchange flow. At all three sites, horizontal flow and downwelling were detected mainly at the deepest points of the thalweg and in case of HPS1 and HPS3, steep vertical angles were found at the slopes of the thalweg.

Although in many cases monitored hyporheic flow vectors can be at least partially explained by streambed topography, we observed significant differences in flow vector frequency distributions between the three investigated sites that were related to small-scale streambed structural heterogeneity and related spatially diverse patterns of streambed hydraulic conductivities.

As sites HPS1 and HPS3 showed both relatively low and temporally stable VHG, it was assumed that groundwater and surface water at these sites were well connected with no evidence of flow confining streambed strata. At both sites, this assumption was partly confirmed by HPS measurements which identified predominately positive vertical flow, indicating groundwater exfiltration. According to nonparametric Wilcoxon-Mann-Whitney tests for non-normally distributed sample populations, HPS2 and HPS3 sides were found to differ significantly in regard to their absolute flow velocities. While HPS1 and HPS 2 differed significantly in their vertical flow velocities and absolute vertical angles, difference between HPS1 and HPS3

were not significant (all tests for 95 % significance levels). With a mean absolute vertical angle $|\beta|$ of $+29°$, site HPS3 showed a stronger horizontal flow component in comparison to HPS1 with a $|\beta|$ of $+46°$. This fact has been interpreted as an indication of a more immediate interaction between surface water and groundwater at site HPS1 with predominantly vertical exchange and some lateral groundwater exfiltration through the bank. As the highly conductive sediments at site HPS1 are known to be in direct vicinity to confining streambed strata restricting exchange between groundwater and surface water up-stream and down-stream of the site, the increased hydraulic conductivity of streambed sediments at site HPS1 presents a window of preferential exchange flow between aquifer and river. The HPS results coincide with observed streambed temperature anomalies at this location observed in Krause et al. [2012a]. In contrast, HPS3 represents an extended stream section with no known presence of flow confining peat layers in the streambed and evidently lower VHG than HPS1, implying high hydraulic conductivity but also no particular hydraulic forcing and preferential groundwater up-welling due to nearby flow restriction as seen at HPS1. Thus, with a hydrologically less constrained situation, weaker driving forces of hyporheic exchange as for instance macrophytes structures were likely to gain more impact.

HPS2, in contrast, represented an area of complex VHG pattern, suggesting spatially very heterogeneous streambed properties with a mosaic of highly conductive sediments as well as flow confining peat or clay layers. The observed dominance of horizontal flow with only few cases of vertical flow and relatively slow hyporheic flow velocities supports the hypothesis of potential increase in surface water infiltration and horizontal flow above flow confining streambed strata which inhibit or at least reduce the pressure of up-welling groundwater in the uppermost streambed sediments overlying low conductivity streambed structures.

A number of previous studies emphasized the impact of streambed topography induced dynamic pressure fields and advective pumping on hyporheic exchange fluxes [Boano et al., 2007, Cardenas and Wilson, 2007b,a, Tonina and Buffington, 2007, Endreny et al., 2011, Endreny and Lautz, 2012]. However, the combined VHG and HPS measurements of this study provide evidence that at the research site, spatial patterns of high versus low conductivity streambed strata, and in particular the size and extent of confining peat and clay lenses, can have significant impact on near surface aquifer-river exchange fluxes in the streambed and may in some cases even superimpose the impact of streambed topography on hyporheic exchange fluxes. The

fact that streambed conductivity significantly controls groundwater-surface water exchange has been previously investigated in experimental and model-based studies [Palmer, 1993, Wondzell et al., 2009a, Krause et al., 2011b]. However, the results of this study indicate that in particular small-scale patterns of streambed conductivity and the combined impact of complex streambed stratification and topography on hyporheic exchange flow patterns are not sufficiently understood. The results of this study improved the conceptual understanding of drivers and controls of hyporheic exchange fluxes (Fig. 5.2) and in particular highlight the impact of flow confining low conductivity streambed strata on shallow hyporheic flow. With its capacity to improve understanding of flow processes on local scales of cm to dm, the combination of HPS and VHG provides a useful approach that complements other methods as the previously suggested combination of FO-DTS and VHG [Krause et al., 2012a] which targets the identification of flow pattern at a spatial scale of several hundred meters. The knowledge gain might be useful for a further process-based upscaling and to build up profound models.

These enhancements in process understanding will need to be considered for the improvement of quantitative assessments of multiscale implications of different drivers and controls of groundwater-surface water exchange fluxes. In order to ensure the efficient design of streambed restoration measures aiming at alteration of streambed topography or streambed structure and stratification [Boulton et al., 1998, Boulton, 2007, Kasahara and Hill, 2008], future research is necessary to investigate the interaction of streambed structure and surface topography in controlling hyporheic exchange fluxes. As demonstrated in this study, the application of HPS technology has a great potential to complement the portfolio of experimental methods for improving mechanistic process understanding in particular of small-scale flow patterns in streambed sediments.

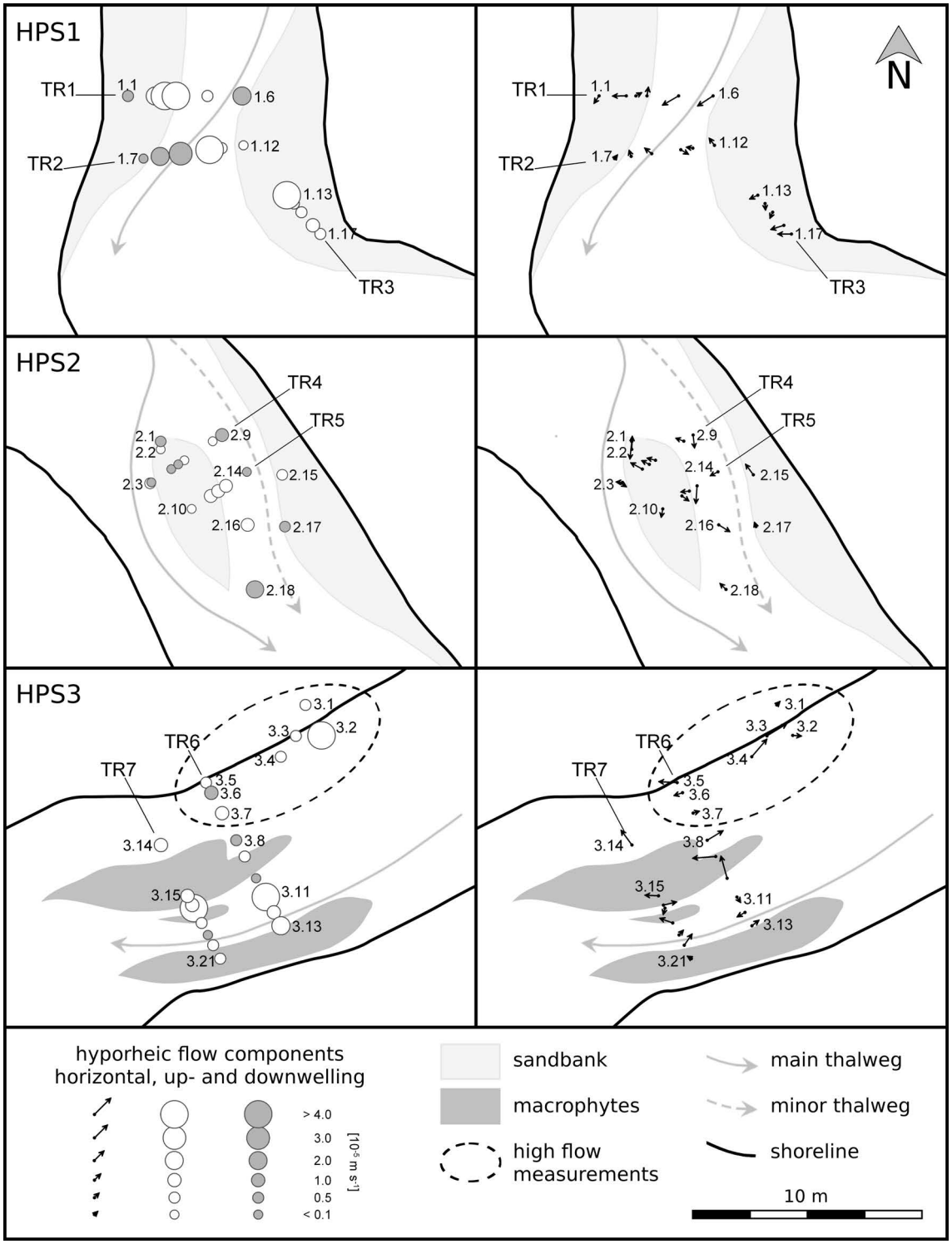

Figure 5.11: Spatial pattern of shallow hyporheic flow at the three HPS sites (HPS1 to HPS3) for 7 transects (TR1 to TR7). Symbol size indicates the (left) vertical component of hyporheic flow, arrows show horizontal projections of the flow vector, i.e., (right) horizontal flow components (compare Fig. 5.10). Maps in the background show schematics of the local streambed topography (HPS1 and HPS2) and macrophyte stands (HPS3). Heat pulse measurements are labeled, with the first digit referring to the HPS location and the digit right of the period referring to the respective sample location, numbered sequentially for every site up-stream to down-stream.

5.4 Conclusions

The results of this study provide evidence that, in addition to pressure head distributions and topography induced advective pumping, streambed stratification and in particular spatial patterns of high versus low conductivity streambed strata can have substantial impact on exchange flow patterns between groundwater and surface water. Combination of VHG observations and novel HPS measurements of small-scale streambed fluxes indicate that low conductivity lenses of peat and clay at streambed depths of several tens of cm can locally reduce the impact of up-welling groundwater pressure on streambed pore waters above the confining structures and thus, support preferential surface water downwelling and horizontal streambed flow. Hence, the research results confirm the hypothesis that low conductivity streambed sediment structures not only inhibit groundwater up-welling but can also support the enhancement of surface water down-welling and horizontal pore water flow at locations where the up-welling force of the groundwater is impeded. Such increase in surface water down-welling and horizontal streambed fluxes does enhance hyporheic transient storage with potential implications for the biogeochemical cycling and eco-hydrological functioning of hyporheic zones. The results furthermore indicate the possibility of enhanced groundwater up-welling where highly conductive streambed sediments in close vicinity to confining streambed structures provide a preferential flow path for fast up-welling of the semiconfined groundwater. The results of this study provide new mechanistic understanding of aquifer-river exchange flow processes in dependence of streambed stratification and topography. They provide evidence that in lowland rivers with complex patterns of variable hydraulic conductivities in streambed as at the investigated field site, the impact of flow confining streambed strata can superimpose the effects of streambed topography induced hyporheic exchange. This improved process understanding has substantial implications for river restoration and management. As residence time distribution and mixing ratios of groundwater and surface water in the streambed provide a unique hyporheic environment where structural alteration can have significant implications for its eco-hydrological and biogeochemical functioning, consideration of the improved process knowledge will be crucial in particular for river restoration programs altering the streambed sediment structure.

The field application of the HPS demonstrated the potential as well as limitation of this novel technology for tracing streambed flow patterns at small scales. Additionally, heat pulse measurements have been found to provide valuable insight

into drivers and controls of shallow streambed pore water fluxes at the aquifer-river interface when compared with VHG observations at greater depth. While HPS observations did not in every case support a fully conclusive explanation of measured single flow vectors, the identified flow vector frequency distributions indicate that shallow hyporheic flow is strongly controlled by streambed conductivity and patterns in groundwater up-welling, supporting initial hypothesis formulated on the basis of VHG observations.

The HPS identified patterns of shallow hyporheic flow contribute highly valuable information to the study of groundwater-surface water interaction and demonstrate a high potential to complement larger scale pressure head or passive heat tracing methods. The comparison of HPS and VHG results provided further evidence that in particular the combination of experimental methods at different scales has the potential to significantly improve mechanistic process understanding, whereas the application of each singular method would have not allowed conclusive differentiation between fluxes, their drivers, and controls. Future research will be required to implement the improved process knowledge into existing modeling approaches for quantification and up-scaling to management relevant scales.

Acknowledgements The authors wish to acknowledge substantial field work support by E. Naden (formerly University of Keele) for some of the HPS measurements as well as from M. Munz (University of Potsdam) and C. Tecklenburg (Helmholtz Centre for Geoscience Research Potsdam) for the installation of the streambed piezometer network and sediment coring. We thank the UK Natural Environment Research Council (NERC-NE/I016120/1) as well as the Royal Geographical Society for funding parts of the research presented in this paper. The authors would like to thank the editor Graham Sanders as well as the anonymous reviewers for their helpful comments and constructive criticism.

Chapter 6

Synthesis

6.1 Summary of process results: hydrological connectivity from hillslopes to streams

The four presented studies investigated the flow processes involved in connecting hillslopes, riparian zones and streams and the importance of spatial structures for response patterns and hydrological behavior. The studies focused on different spatial scales and different sections of the flow path from hillslopes to streams: flow processes in the soil column and in hillslopes, the spatial patterns of groundwater discharge along a small headwater stream, and the connectivity between groundwater, hyporheic zone, and surface water in the streambed of a lowland river. The results of these studies showed a strong link between characteristic spatial structures and hydrological behavior.

6.1.1 Subsurface structures create soil profile and hillslope connectivity

The first two parts of the thesis featured a hillslope-scale irrigation experiment that was accompanied by subsurface structural explorations and several plot-scale irrigation experiments (Chapter 3), and catchment-scale discharge dynamics (Chapter 2). Flow processes within the soil profile were controlled by the highly heterogeneous substrate of the young and shallow soil. The mixture of silty matrix material and non-uniform sequences of layered periglacial deposits caused highly irregular and patchy infiltration patterns. The multi-porous substrate allowed for the establish-

ment of preferential flow paths in the entire soil profile. As a consequence, we observed fast soil moisture responses at greater depths (below 0.5 m below ground surface), while large portions of the soil were by-passed and barely interacted with the infiltrating water. At the hillslope, heterogeneous subsurface structures created connectivity by linking mm scale flow paths across the entire hillslope. The result was lateral subsurface flow through a preferential flow path network, with impact on the overall catchment behavior (double-peak hydrographs).

6.1.2 Topography, subsurface structures, and stream morphology control connectivity between catchment and stream

In the same study area, a combination of incremental discharge gauging, radon measurements, and topographic and geomorphological analyses was applied to identify spatial patterns of catchment-stream connectivity and their structural controls (Chapter 4). The hydrological connectivity between hillslopes, riparian zone and stream was shaped by topography and the dualism of characteristic subsurface structures that could be found across hillslopes and in the riparian zone, and the geomorphological heterogeneity of the stream channel. The stream was fed from groundwater and event water that was routed from the hillslopes towards the valley bottom. The yield of a section of the stream was found to be proportional to the catchment area during winter base flow conditions, but during medium and high flow, the pattern was modified by the local geomorphology of the riparian zone and the stream channel. The patterns observed during these campaigns were shifted downstream, which could be explained by parafluvial flow through deposit layers in the riparian zone intersecting flow paths from the hillslopes. The highest gains were found at locations downstream of reaches with large lateral catchment areas and where the parafluvial flow paths in the riparian zone were interrupted by step-pool morphology in the stream channel. Thus, the groundwater discharge pattern along the stream was controlled by the catchment topography and modulated by riparian zone and channel morphology.

6.1.3 Streambed structures modulate groundwater upwelling and hyporheic exchange

Part IV of this thesis used vertical hydraulic gradient observations and advective heat transport measurements to investigate the influence of substrate heterogeneity on flow processes in the streambed of the lowland river Tern (Chapter 5). The spatial arrangement of flow-impeding streambed strata were found to control hyporheic flow and surface water-groundwater connectivity. While vertical hydraulic gradients indicated upwelling conditions along the entire study area, the strength and variability over time suggested that different processes were controlling the connectivity between groundwater, hyporheic zone and surface water: Differences in streambed hydraulic conductivity were found to modulate the upwelling of groundwater. At the same time, streambed strata strongly controlled the interaction between surface water and the hyporheic zone. While the impact of local streambed hydraulic conductivity is an obvious control on vertical flow processes, the study also showed that the larger-scale spatial arrangement of flow-impeding structures also impacted local flow processes. Narrow high conductivity "windows" in otherwise less permeable substrate focused vertical flow, while wide-stretched areas of the same highly conductive substrate allowed for lateral flow and behaved differently despite similar sediment profiles.

6.2 Synthesis of process results

The results of the four studies show that in both environments and at all scales, preferential flow played an important role. Preferential flow processes controlled response patterns and hydrological behavior, not only in the highly heterogeneous headwater, but also the lowland river Tern. The experiments also showed that the spatial organization, i.e., the spatial arrangement and connectivity of structures across scales, was important to understanding internal processes. On the one hand, we found small-scale structures that seemed local, but connected across hillslopes, and thus influenced hillslope-scale flow processes and catchment behavior (Chapter 2 and 3, plot and hillslope irrigation experiments). On the other hand, large-scale response patterns were modulated through medium-scale or small-scale structures. At the lowland river Tern, heterogeneous streambed strata changed the large-scale groundwater flow pattern (Chapter 5, streambed heat pulse injections). At the Holtz

headwater, the topographic signal was superimposed by the effect of deposit layers in the riparian zone and streambed morphology (Chapter 4, radon and sequential discharge measurements).

6.3 Practical and conceptual challenges

The process results of the four studies show that the spatial context of structures and the resulting preferential flow processes is key for understanding hydrologic connectivity. Thus, hydrologic connectivity of a system cannot be assessed properly at a single scale, the scale of interest alone. For an ideal investigation or representation, all relevant scales have to be identified and considered. Depending on the complexity of the investigated system, this means extensive effort – be it in experimental work or model complexity – which is practically not feasible and leads to over-parameterization of models [Blöschl, 2001]. The problem of scaling, i.e., the identification of relevant scales and the transfer of information across scales (see Sec. 1.2.2), is an important and recurring topic in (catchment) hydrology.

Depending on the scale of interest, different foci regarding the spatial organization of flow processes that lead to connectivity are commonly applied: At the scale of soil specimens the focus is on pore geometry [e.g., Vogel and Roth, 2001] and pore size distribution [e.g., Tyler and Wheatcraft, 1990], processes at the pore-matrix-interface [e.g., organic coatings in biogene pores, Leue et al., 2015], or in-pore processes like capillary pressure dynamics and air clogging [Helmig et al., 2007, Snehota et al., 2015]. At the plot scale, the impact of initial conditions on the establishment of preferential flow [e.g., Flury et al., 1994] and also connectivity of structures [e.g., Klaus et al., 2013] are investigated. The connectivity of subsurface structures appears to be the main factor at the hillslope scale [e.g., Tromp-Van Meerveld and McDonnell, 2006, McGuire and McDonnell, 2010, Wienhöfer and Zehe, 2014], while topographic connectivity is often used as proxy for functional organization [Jencso and McGlynn, 2011] at the catchment scale.

The investigation of catchment-scale SW-GW interaction focuses on lithological or hydrogeological settings [e.g., Valett et al., 1996, Malcolm et al., 2005] and large-scale topographical features creating groundwater recharge and discharge areas [e.g., Covino and McGlynn, 2007]. At the reach scale, a typical control of SW-GW interaction is riparian morphology, like stream meanders [e.g., Boano et al., 2006], and the position of the stream channel within the groundwater system [e.g., Sophocleous,

2002], and channel-sized flow obstacles like beaver dams [e.g., Hill and Duval, 2009]. At the plot scale, hyporheic exchange is controlled by channel morphology such as gravel and point bars, step-pool formations [e.g., Buffington and Tonina, 2009], and in-channel flow obstacles like wooden debris [e.g., Wondzell et al., 2009b] in combination with hydro-static or dynamic pressure heads. At the small scale, the focus is on streambed structures and advective pumping [e.g., Packman et al., 2004].

The scale-dependent research foci show that different processes are relevant and take effect at different scales. For a holistic system understanding, however, and to identify relevant processes with regard to catchment-scale connectivity, observations at different scales need to be integrated and merged into one unifying conceptual representation of the system. To do so, a conceptual framework is required that represents the basic architecture of the system, but is flexible enough to allow for the complexity of a spatially diverse and dynamic system.

One concept to cope with the matter of scaling is the scaleway concept. This concept was developed to "handle multi-scale heterogeneity in order to predict flow and transport processes at a specific site and at a given scale" [Vogel and Roth, 2003]. It is based on the idea that soil is a hierarchical heterogeneous medium, in which the material properties at each possible scale are determined by the characteristics at the next smaller scale [Roth et al., 1999].

6.4 The scaleway concept in catchment hydrology

The scaleway concept was originally suggested for modeling approaches with focus on soil profiles. Extending it to the hillslope scale and further may also be possible, but entails increasing challenges regarding the identification of controlling structures and requires upscaling within a structurally increasingly complex system [Schulz et al., 2006].

The same concept is applicable to experimental field research and is comparable to the approach presented in Chapter 2 and 3 (plot and hillslope irrigation experiments) of this thesis: Knowledge about structures and processes in the soil column was used to explain and interpret hillslope-scale observations. The insight gained on hillslope-scale processes, in turn, were linked to the catchment behavior. This way, we were able to develop a conceptual model of the investigated hillslope that includes processes at multiple scales and leading to an improvement of our mechanistic understanding of the internal processes.

6.5 The reverse scaleway?

The approaches and results presented in the two other parts of this thesis (Chapter 4, radon and sequential discharge measurements, and Chapter 5, streambed heat pulse injections), however, do not fit into the conceptual frame of the scaleway. Both studies did not reveal small-scale processes that assemble and connect to larger patterns, but rather *vice versa*: large-scale patterns were altered and modulated by smaller-scale structures.

Depending on the scale of interest, the large-scale conditions could be summarized and treated as the boundary condition for a more detailed investigation or representation of the smaller scale. This procedure is intuitive and quite frequently applied especially by modelers but also experimental hydrologists. Catchments are delineated and general groundwater flow directions are estimated based on spatially distributed measurements of groundwater levels and taken as basis for further observations. Another example are vertical hydraulic gradients which are used as boundary conditions for a streambed profile model.

In analogy to the scaleway concept as suggested by Vogel and Roth [2003], this approach can be described as a *reverse scaleway*. This means that the starting point of any investigation or modeling approach is at the upper end of the range of scales covered by the study. In the next step, models or observations will be iteratively refined according to structures that may alter the hydrological behavior.

In the case of hyporheic connectivity at the lowland river Tern (Chapter 5, streambed heat pulse injections), the relevant scales are relatively discrete and distinguishable: Groundwater levels and vertical hydraulic gradients were used to approximate the groundwater flow field at the scale of 100 m but streambed structures at the m scale also influenced SW-GW interaction. In such relatively clearly structured environments, the refinement can be based on a distinction of cases. Large-scale observations and structures are used to develop well informed hypotheses for the structural control and behavior at the smaller scale. Afterwards, representative landscape units are identified which are suitable to test these hypotheses.

In case of the hillslope-riparian zone-stream connectivity of the Holtz headwater (Chapter 4, radon and sequential discharge measurements), the situation is more complicated. Here, the scales of different structures (topography, subsurface structures at hillslopes and in the riparian zone, channel-morphology) and related processes overlap, and the scale hierarchy is not as easily identifiable as in the aforemen-

tioned example. The separation of different processes is not easily possible and the approach needs to be adapted accordingly. Instead of identifying and investigating representative landscape units, a multi-parametric approach is required.

6.6 The parabolic scale series

The scaleway and reverse scaleway concept are similar in the way that they both acknowledge that every process or structure is assembled of smaller units in the sense of scale levels and holons as described in Sec. 1.2.2. The main difference is the direction of causality. According to the scaleway concept small-scale structures and processes are required to create the larger-scale effect, while the reverse scaleway concept requires the large-scale processes to drive the smaller-scale effect.

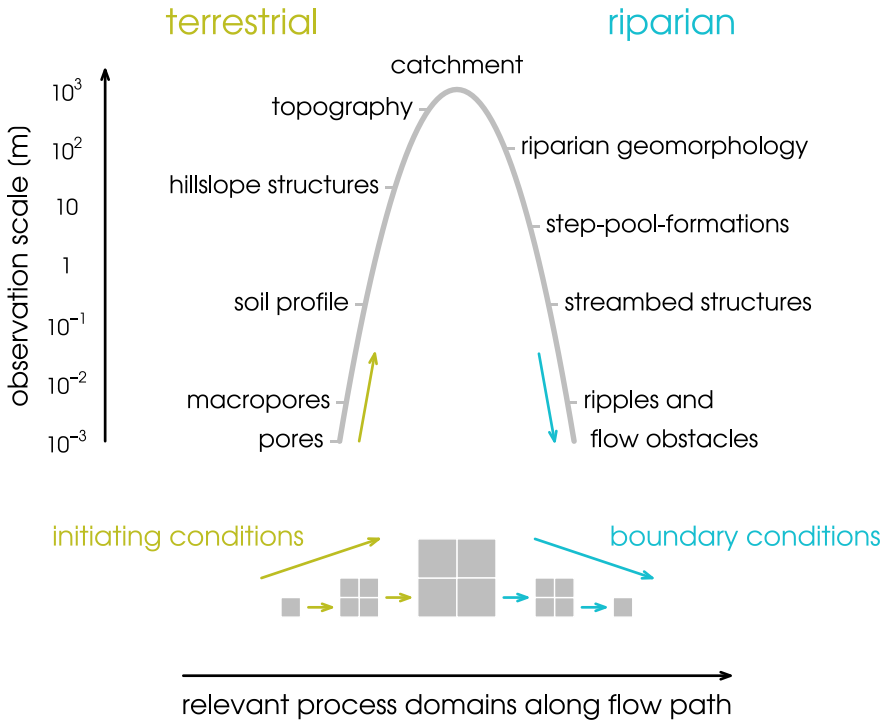

Figure 6.1: The parabolic scale series

In the case of catchment hydrology, the direction of causality equals the direction of flow [Larsen et al., 2012]. Accordingly, the scaleway concept is most applicable to research questions focusing on the soil and hillslopes. Here, the terrestrial part of the water cycle and the flow path through the catchment begins with the infiltration

into the pores of the soil. The tiny flow paths merge and combine to larger flow path networks and flow fields, dominating soil column and hillslope behavior. Here, patterns of connectivity are *created*. While constraints exerted by the upper level do exist in soils and hillslopes (e.g., as gradients in matrix potential or topography), the initiating condition, mechanisms and components provided by the lower scale level [Wu, 1999] are better suited to decompose the system behavior.

The reverse scaleway concept is suited best for research questions focusing on surface waters and their interaction with the landscape. From a process-oriented surface-water perspective, large-scale drivers are setting the frame for smaller-scale structures and processes that alter patterns and behavior. The connectivity patterns created at the hillslope and within the catchment are *modulated*. Here, the boundary conditions exerted by the upper level set the context for processes at the lower level [Wu, 1999].

Both concepts are ways to decompose the system into observable units and decipher the spatial organization of structures and processes at certain scales. Combining them the way described above allows us to do so from hillslopes to streams. By ascending and descending the "scaling ladder" along the flow path through the catchment, the prevailing spatial organization of a catchment could be revealed. Starting with pore-scale processes at the soil surface and proceeding towards the larger scale, as Roth et al. [1999] and Vogel and Roth [2003] suggested, processes in the soil profile are analyzed. This procedure can basically be continued up to the scale of hillslopes and small catchments. At some point, however, the flow paths stop to converge, but diverge instead. The maximum unifying scale is reached, when the significantly different structures and mechanisms of the riparian zone have to be included into the scale domain. The reverse scaleway begins and when the catchment signal is modified and superposed by riparian and streambed structures and processes (Fig. 6.1).

Every scale level consists of a multitude of holons which represent the processes and structures within a catchment. Examples for parallel holons are different soil types, concave or convex hillslopes, or geomorphological features along the stream. The holons of one scale level may or may not interact with each other and are themselves contained in one or multiple holons of the next larger scale level. This is a way to integrate information across scales [Vogel and Roth, 2003], but at the same time the amount of holons per scale level and their ramifications across scales also tells us a lot about the spatial organization of the catchment.

6.6.1 The parabolic scale series in the literature

The very concept of the scaleway and the reverse scaleway applied to hillslopes and riparian zones, respectively, can be found in hydrologic literature in the form of readily implemented models, conceptual approaches and characterization schemes.

The scaleway [Roth et al., 1999, Vogel and Roth, 2003] has been discussed in detail already and partially inspired the development of the concept of the parabolic scale series presented here (Fig. 6.1). Another example that resembles the terrestrial limb of the parabolic scale series is the idea of elementary functional units that dynamically reassemble and form lateral topological units, depending on external drivers [Zehe et al., 2014b]. While this approach is not explicitly bound to a representation by scale levels, its concept is nonetheless implicitly based on the idea of small systems that assemble to larger systems to better understand and represent internal processes.

An example that is analogous to the reverse scaleway, i.e., the riparian limb of the parabolic scale series (Fig. 6.1), is a process-oriented hierarchical typology for riparian landscapes with regard to SW-GW interactions [Dahl et al., 2007]. The typology was developed as a basis for water management assessments according to the European water framework directive [European Comission, 2000]. The goal of this typology is a reproducible characterization of the "functional linkages between several hydrological components" [Dahl et al., 2007] based on landscape type (catchment scale), hydrogeological setting (reach scale), and riparian flow path types (local scale).

6.6.2 Application of the parabolic scale series

The four parts of this thesis can be located at different sections of the two limbs of the scale parabola. The observation scales of the experiments complement one another, partially overlap, and leave gaps at scales that were not covered (Fig. 6.2).

In the Holtz headwater catchment, the plot-scale irrigation experiments (Chapter 3) covered the scales between mm to m at the terrestrial limb of the parabola. Subsurface structures, soil moisture and tracer response patterns in the soil profile at the dm scale and m scale, tracer recovery and GPR imaging at the cm scale, and the visual identification of tiny flow paths at the mm scale. The hillslope-scale irrigation experiment (Chapter 2 and 3) covered the scales between dm and 10 m: soil moisture response at the dm scale for single measurements (limited by the inte-

gration volume of the soil moisture measurement), at the m scale for single profiles, and at the 10 m scale for multiple, spatially distributed profiles. Time-lapse GPR imaging covered the dm scale and the 10 m scale.

The hydrograph and stable isotope data collected at the gauges of the Holtz headwater and neighboring streams (Chapter 2) are the integrated signal of the entire catchment and represent processes at the 10^3 m scale. The scales between 10 m and km are thus not directly covered.

On the riparian limb of the scale parabola, structures are covered at the catchment scale by GIS analyses and at the m scale to the 10 m scale by the data on stream channel morphology. Processes were observed in the form of discharge patterns at the scales between 1 m to 100 m (Chapter 4). While response patterns were covered more or less continuously in that range, there is a lack of knowledge about structures at this intermediate scale.

According to this analysis, the logical next step in the investigation of the Holtz headwater catchment is an assessment of structures within the riparian zone, and the spatial coverage of hillslope subsurface structures throughout. Especially the knowledge about the connectivity and similarity between specific characteristics of these structures at the same scale (10 m to 100 m) for both the hillslopes and in the riparian zone would be of great value. This knowledge could be used to test the conclusions on the importance of parafluvial flow for discharge patterns along the headwater stream (Chapter 4) and put them on a physically sound basis. Alternatively, a literature review of the spatial distribution of periglacial deposit layers and their occurrence in floodplains could be used to bridge this gap.

It is important to note that the knowledge collected in the scale parabola in Fig. 6.2 is not a complete representation of structures and processes even for the scales that are covered. There might be a number of alternative holons, especially at the lower scale levels. As an example, hillslope studies all focused on forested hillslopes, while the catchment also features a significant portion of arable plateaus at its periphery. The structures and processes related to these plateaus were not considered and would contribute additional holons at the same scale levels as the plot-scale and hillslope-scale experiments. Similarly, processes under radiation driven conditions are not represented in the current conceptual model of the terrestrial limb of the parabola.

The observations from the lowland river Tern are also depicted in the parabola in Fig. 6.2. As the two study areas are very different in their characteristics, there is not

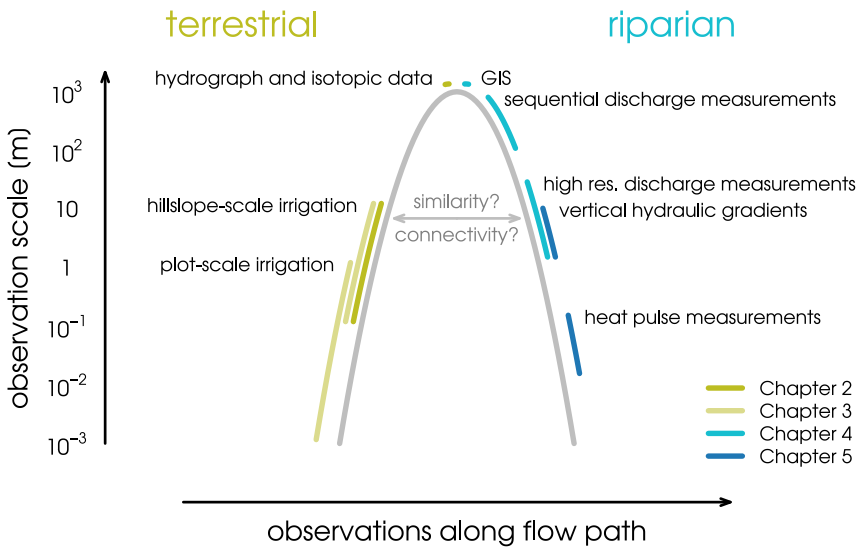

Figure 6.2: The observations of the four studies classified along the parabolic scale series

much knowledge to transfer from one catchment to the other. However, knowledge of structures and processes can be put into relation to each other by locating them along the scale parabola. This is a way to compare the spatial organization of catchments and to identify differences and similarities. Based on this knowledge, the transferability of process understanding can be evaluated.

6.6.3 Limitations

The concept of the parabolic scale series is based on the assumption that the dominant process direction is from hillslopes towards the stream. This assumption exerts restrictions to the range of representable landscapes, as the flow direction can only be taken for granted if the catchment's topography is clearly structured. As elaborated in Sec. 6.5, processes within the hyporheic zone are generally better represented by the reverse scaleway (see also Carpter 5 on processes in the streambed sediment of the lowland river Tern). Similarly, processes in the soil profile can be described by the scaleway as described by [Vogel and Roth, 2003]. However, the hillslope-centered view of hillslope processes being the link to catchment-scale behavior might not be adequate. Instead of hillslopes, different-scale groundwater flow systems [regional, intermediate, local Tóth, 1963] might be a better link between terrestrial and riparian processes and thus a proxy for connectivity in low relief landscapes.

Another limitation is mainly of practical nature: while it is theoretically possible to represent topographically diverse catchments with a high number of nested sub-catchments, this diversity drastically increases the complexity of the resulting parabola (i.e., holons per scale level) beyond practicability. Rather than fitting everything into one scheme, catchments should be broken down into sub-catchments of manageable complexity. The scheme is therefore best suited for small-scale to medium-scale catchments with a moderate degree of landscape variability.

6.7 Outlook

The concept of the parabolic scale series was developed to relate observations of different nature and scale made in the frame of the four parts of this thesis. The concept is based on ideas that had been presented before [Roth et al., 1999, Vogel and Roth, 2003, Dahl et al., 2007] and it was evaluated with regard to its applicability to other observations in the fields of hillslope hydrology, SW-GW interactions and catchment-stream connectivity in general. However, the elaborations above are only a first proof of concept, and a detailed application of the parabolic scale series is still pending.

The next step is to systematically apply the parabolic scale series concept. Different procedures are possible, depending on the objective and the available data base. A possible strategy is to apply the concept to one well-studied catchment and collect static properties to represent the structural organization of a catchment. All related dynamic processes could then be organized in separate parabolas, representing the functional organization in dependency on different driving conditions [e.g., radiation driven *versus* precipitation-driven conditions, Zehe et al., 2014a]. The resulting set of parabolas documents and visualizes the functional diversity of the represented system and is a tool to elucidate knowledge gaps. If applied to multiple catchments, the parabolic scale series could serve as a scheme to assess and compare spatial and functional organization of different systems and identify similarities.

As such extensive data bases rarely exist for single catchments, an alternative approach is to merge knowledge gained from studies in multiple catchments into one scale parabola. To do so, the targeted system has to be defined with regard to size, complexity, spatial properties and/or behavior, e.g., headwater catchments that are characterized by periglacial slope deposits or double-peak hydrographs. Based on these criteria a thorough literature review is conducted and the results and obser-

vations structured along the limbs of the scale parabola. The resulting conceptual model could reveal dominant structures or processes for the targeted catchment type or help to identify characteristic structure-process combinations across scales. As such, it provides a scheme for transferring knowledge and developing well-informed hypotheses for ungauged systems of similar morphogenetic origin.

The applications described above all refer to the organization and categorization of field data. Last but not least, the parabolic scale series could also be a blueprint for the development of coupled models that link multiple scales, similar to the scaleway model presented by Vogel and Roth [2003].

The End

Acknowledgements

First of all, I am heartily thankful to my advisers. Theresa Blume took the time whenever I needed advice or support and her door was always open to me. She provided new perspectives when I was stuck and was always prepared to discuss new ideas. Most importantly, she was always backing and supporting me when waters got rough. Bruno Merz gave me the opportunity to take my doctor's degree at the Hydrology section at the German Research Centre for Geosciences (GFZ). I am very grateful to him for his counseling and being an objective adviser. My work on this thesis was funded by the German Research Centre for Geosciences (GFZ), the German research foundation (DFG, FOR 1598), the German Academic Exchange Service (DAAD), and the German social welfare system.

Beyond that, many other people were involved in crating a cooperative and pleasant working environment and transforming my PhD project into a fruitful and exciting endeavor. In this context, special thanks go to the entire CAOS team, my co-authors, and all the wonderful people who have contributed their knowledge, sweat, and time during and in preparation of the field campaigns: Marcel Delock, Sibylle Hassler, Jean-Francois Iffly, Britta Kattenstroht, Lisei Köhn-Reich, Begoña Lorente Sistiaga, Daniel Mack, Markus Morgner, Emma Naden, Lena Reiber, Marvin Reich, Christina Tecklenburg, Heiko Vollmann, and Martin Zimmer.

Brian McGlynn gave me the great opportunity to be part of his lab group for the period of five months; a time that I will always remember as an inspiring experience and which spawned many interesting discussions with him and his team. I owe my gratitude to Jan Fleckenstein and Jörg Lewandowski, who have been mentoring and supporting me in the early days of my scientific career and until today. My heartfelt thanks go to my colleague and dear fellow Christina Tecklenburg. As my office mate and ally she made this project so much more joyful.

Most importantly, none of this would have been possible if it was not for the loving support by my family. Especially the emotional support and child care provided by my mother Renate, and Daniel Mack's dedication, ranging from his patience and affirmation throughout the entire period, to him keeping me free of ties, his linguistic advice, and the healthy food he provided me with in the final phase of this endeavor.

Thank you!

Appendix A

A.1 Quick sampler for fast undisturbed core sampling on excavated profiles

In addition to the dye tracer stain records, quantitative analysis of salt tracer recovery distribution in the excavated profiles underneath the irrigation plots was done. One challenge to address was the required time to collect an adequate array of such soil samples with known volumetric reference. We developed a reloadable core sampler with a calibrated sample volume of 66 mL.

The sampler is applied like a ring sample with an attached hammering adaptor. In order to minimize time and impact on the profile we enabled a pull-withdrawal of the sample. For this, the sampler is about 15 mm longer than the desired sample. The irregular open edge is scraped off by a calibration twist drill. The prepared and accurate to volume sample is finally pushed out by a piston from the sampler into a sealable brown glass bottle for further treatment in the laboratory.

A.2 Detailed results of hydrological measurements

In situ measurements of infiltration capacity and saturated hydraulic conductivity had a highly heterogeneous distribution. To detail the respective records and found profiles, Fig. A.1 shows them in a spatial context.

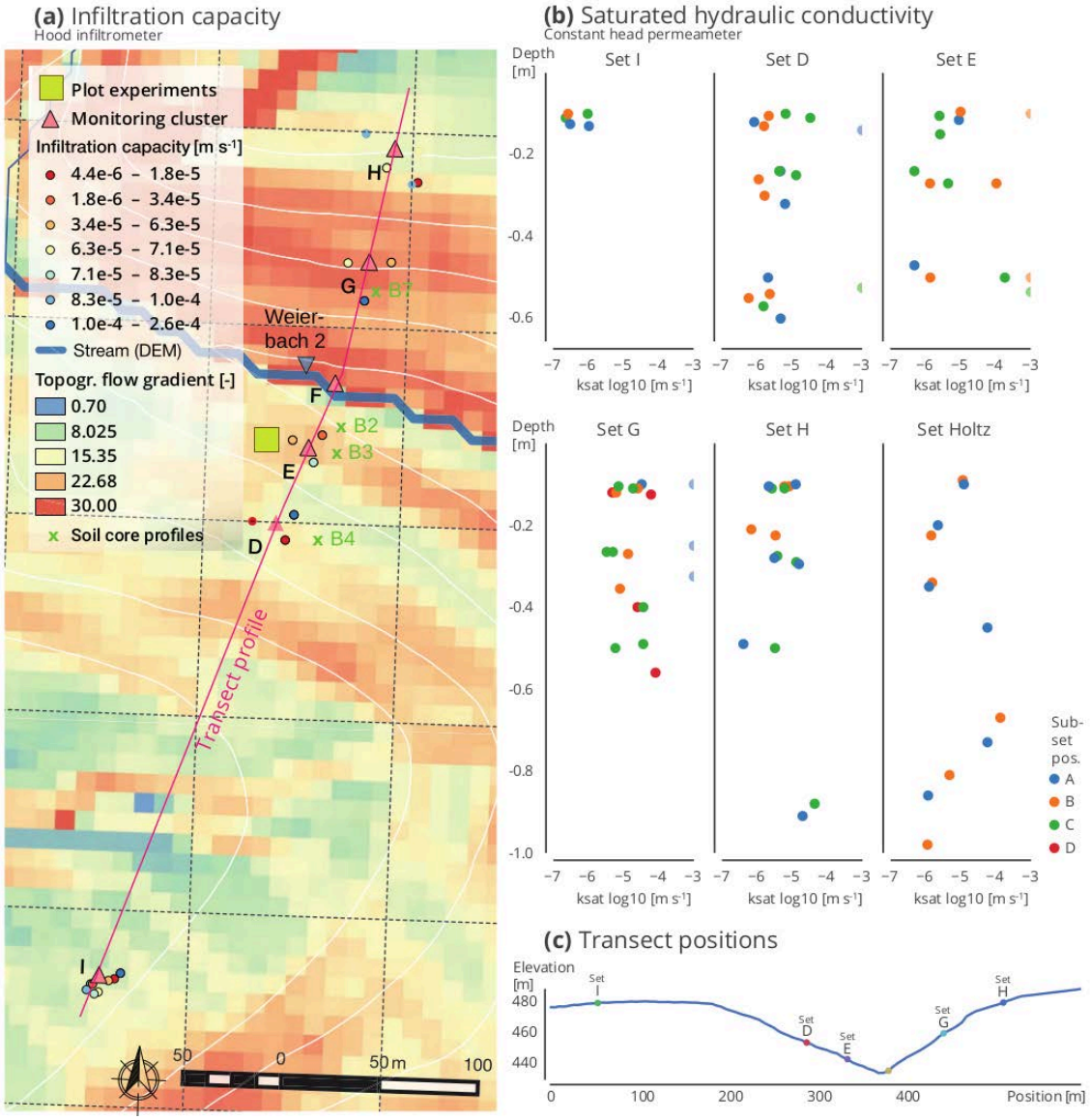

Figure A.1: Hydrological exploration results. **(a)** Infiltration capacity (values color coded) measured with hood infiltrometer. Base map with topographic flow gradient $G = \arctan(\partial z_{rel}/\partial d_{rel})$, topography contour lines and positions of hydro-meteorological monitoring clusters. **(b)** Saturated hydraulic conductivity measured with a constant head permeameter. Individual profile position in the respective nested set color coded (two measurement holes with dist = 1 m). Values exceeding the device capacity set to $10^{-3}\,\mathrm{m\,s^{-1}}$. **(c)** Elevation profile of the transect as a characteristic landscape feature in the sub-basin.

A.3 A priori model reference

Based on the findings of the pedo-physical exploration, we setup the 2D process model CATFLOW [Zehe et al., 2001] as representative hillslope for hypothetical a priori simulation of the experiment in order to determine the required irrigation intensity, the spatial extent of the observation network, the temporal resolution, and

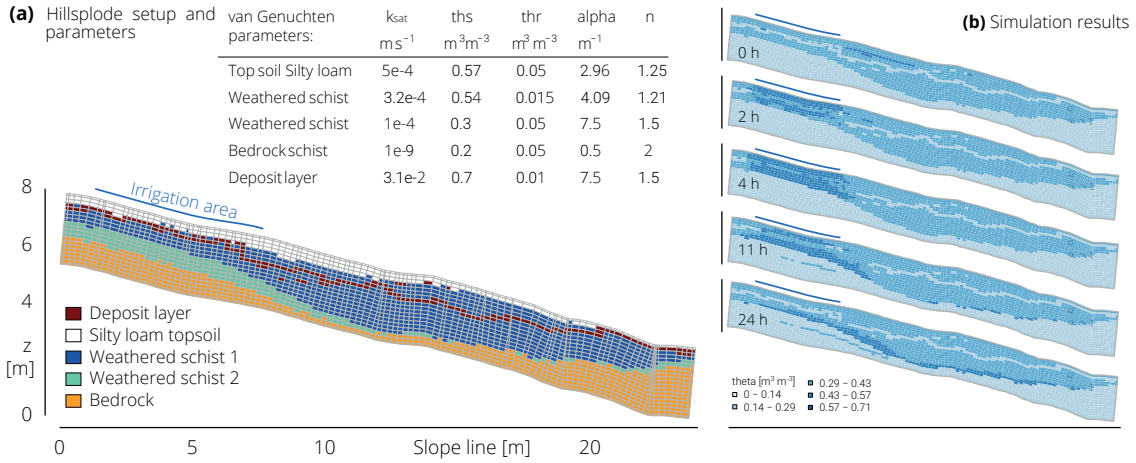

Figure A.2: CATFLOW model reference of the hillslope experiment. **(a)** Table of used soil definitions and hypothetical hillslope setup. **(b)** Simulated soil moisture with 30 mm h^{-1} irrigation for 4 h.

the duration of the monitoring. The model domain was set up based on the soil property estimates from the soil physical exploration assuming a fractured periglacial deposit layer as conductive layer in the hillslope (Fig. A.2). In a series of scenarios, the one with 30 mm h^{-1} irrigation for 4 h turned out to be well balanced with respect to anticipated hillslope reaction given a limited source area. Fast soil water redistribution was modeled to last for 12 h.

Comparing the results from the model with the experiment shows strong deviation in terms of the activation of a conductive layer. However, this could be improved by adding a layer of low permeability below, since the modeled reaction on the bedrock interface is quite similar but slower than the observed dynamics.

A.4 Porewater stable isotope analysis in plot irrigation experiments

In addition to bromide as a conservative salt tracer, the same percussion drilled core samples were analyzed for their stable isotopic composition ($\delta^{18}O$ and $\delta^{2}H$) of the porewater. This was done with the direct equilibration method as proposed by Wassenaar et al. [2008] and described in detail by Sprenger et al. [2015b] using a wavelength-scanned cavity ring-down spectrometer (Picarro Inc.). The precision for the method is reported to be 0.31 ‰ for $\delta^{18}O$ and 1.16 ‰ for $\delta^{2}H$ [Sprenger et al., 2016]. The measured isotopic signal is given relative to the Vienna Standard Mean Ocean Water. As a pre-experiment reference, a fourth reference core was sampled prior to the experiments about 3 m upslope.

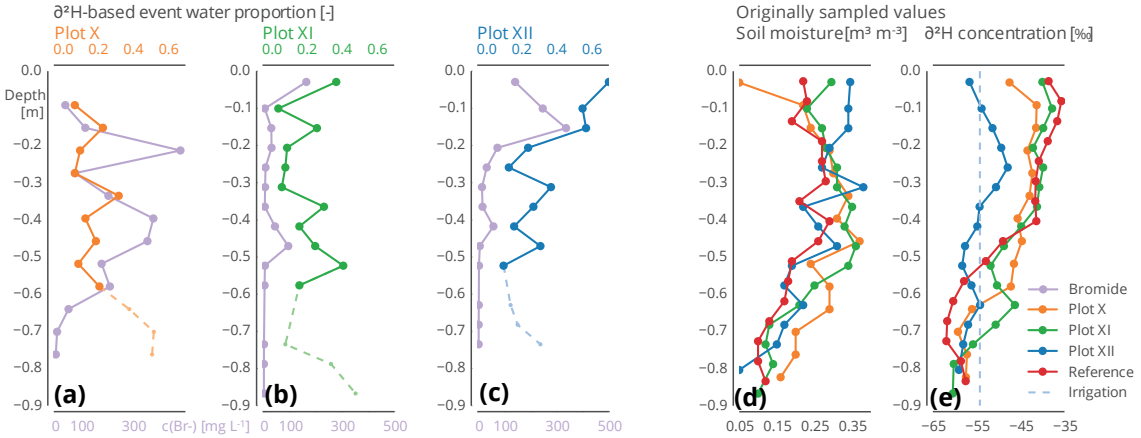

Figure A.3: Plot-scale irrigation experiments. Proportion of event water derived from deviations in concentrations of deuterium and bromide in the soil water of sampled cores to the reference **(a–c)**. Absolute measured soil water content **(d)**. Deuterium concentration in samples and signature of irrigation water **(e)**. Note that the δ^2H concentration profiles and the signature of the irrigation water show low deviation and even coincide at 0.5 m depth. Thus the interpretation needs precaution, especially below that level (dashed lines, **a–c**).

We calculate the volumetric event water portion (-) in the soil water as

$$\frac{\Theta_{event}}{\Theta_{2h}} = \frac{\Theta_{2h} \cdot \delta^2 H_{2h} - \Theta_{pre} \cdot \delta^2 H_{pre}}{\Theta_{2h} \cdot \delta^2 H_{event}}, \tag{A.1}$$

with δ^2H as deuterium composition (‰) in the pre-event reference sample ("pre"), in the core sample 2 h after irrigation start (2 h), and in the irrigation water ("event"). The amount of soil water is given as Θ (SI).

Figure A.3a–c show the depth profile of irrigation water as a portion of total water content, calculated from the deviation in δ^2H concentration between the reference and 2 h past irrigation core samples. The results are also compared to the bromide concentrations in the soil water phase of the same samples, showing slight correlation. However, the values are rather noisy due to low difference of the isotopic composition of the soil water and the non-enriched irrigation water. Figure A.3d–e highlight the very weak soil moisture signal and low deviation between the respective soil cores close to the method's precision. Especially interpretation of the peak at about 0.5 m depth and signals below may be erroneous, because the signature of the reference core coincides with the irrigation water there.

In line with the findings of Klaus et al. [2013] the isotopic signal of non-enriched water required strong assumptions for its interpretation. In our case this specifically applies to the plot-scale core samples where we calculated the difference to the pre-experiment core regardless of the fact that soil water and irrigation water deviated only slightly (≥ 15) and even had the same values at 0.5 m depth. Moreover, the reference core was required to be at a different location. Hence, flow paths and thus the initial isotope profile are not necessarily the same as at the respective plots. However, as an assumably ideal tracer, the stable isotope data allowed for an

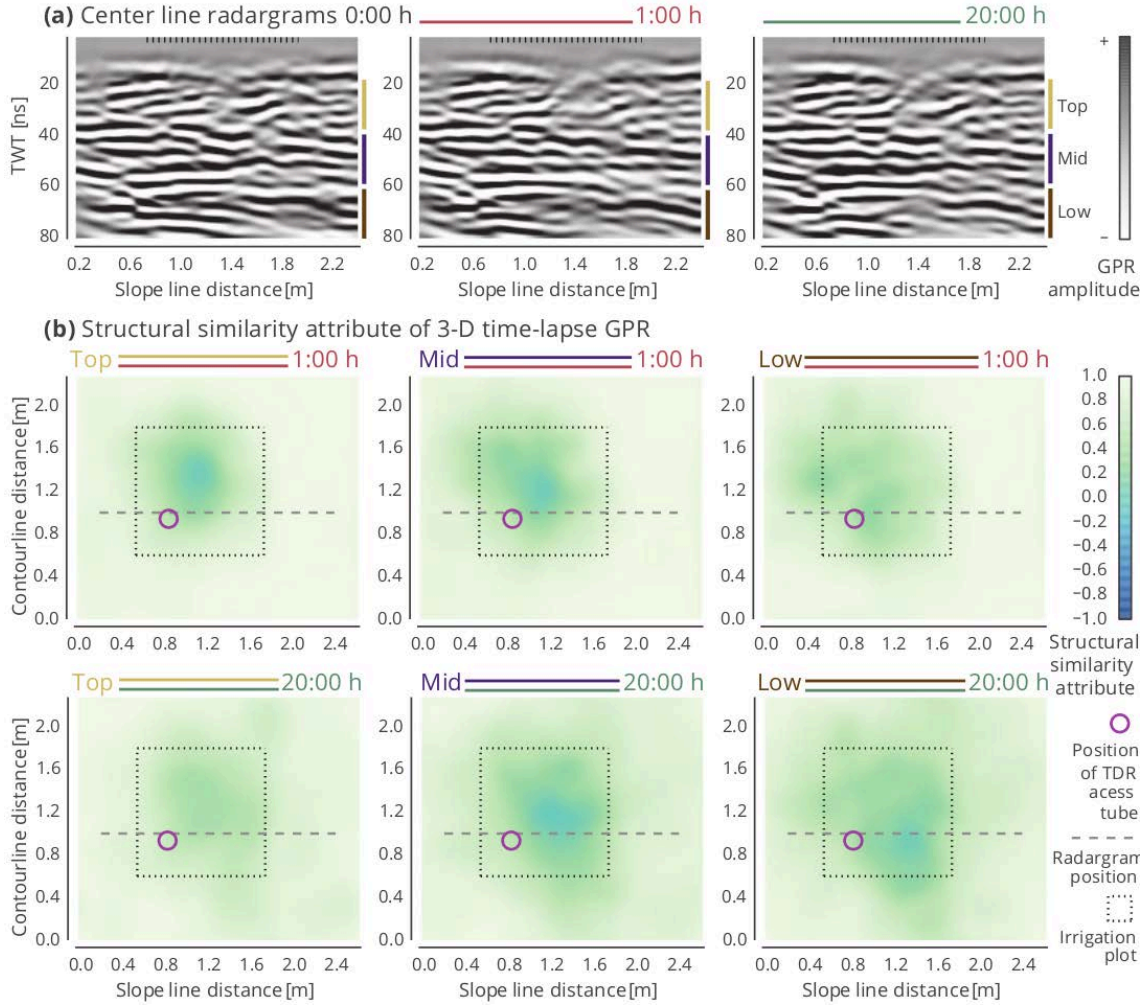

Figure A.4: Time-lapse 3D GPR of the irrigation experiment at plot XI. Center line radargrams at the marked transect (gray dashed line in the lower panels) for the three acquisition times (before 0:00 h, directly after irrigation 1:00 h, 20:00 h after irrigation) are given in the top row. The structural similarity attribute of the 3D data cube is given in three different depth layers (top 20 ns to 40 ns, mid 40 ns to 60 ns, low 60 ns to 80 ns) in the lower panels. The irrigation plot is marked by a black dashed box/line. Slope line distance is increasing downslope.

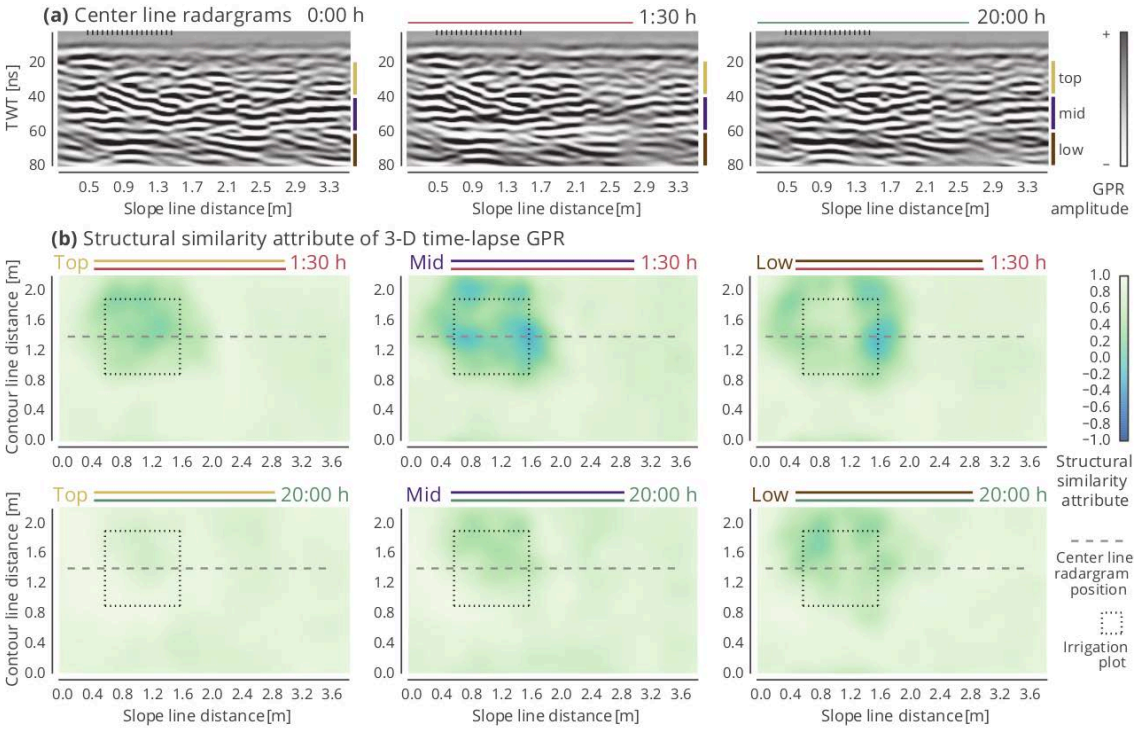

Figure A.5: Time-lapse 3D GPR of the irrigation experiment at plot XII. Center line radargrams at the marked transect (gray dashed line in the lower panels) for the three acquisition times (before 00:00 h, directly after irrigation 01:30 h, 20:00 h after irrigation) are given in the top row. The structural similarity attribute of the 3D data cube is given in three different depth layers (top 20 ns to 40 ns, mid 40 ns to 60 ns, low 60 ns to 80 ns) in the lower panels. The irrigation plot is marked by a black dashed box/line. Slope line distance is increasing downslope.

additional and coherent measurement. With respect to the overall findings of rapid flow in discrete structures the assumption is justified.

A.5 Results for the 3D time-lapse GPR at plots XI and XII

In addition to the results in Sect. 3.3.2, here the radargrams and structural similarity attributes for the other two plot-scale experiments are given in Figs. A.4 and A.5. In plot XI with less intense irrigation the lateral spread of water is less pronounced. As found by the tracer methods, interaction with the soil matrix was elevated in plot XII. Moreover, the acquisition of the GPR data took longer at this plot.

A.6 Technical concerns of the time-lapse GPR and the structural similarity attribute

The demands on the precision of the repeated acquisition with spatial determination and antenna contact to the ground are very high and are assumed to be nearly perfect within our experiments. Under field conditions precision is limited due to

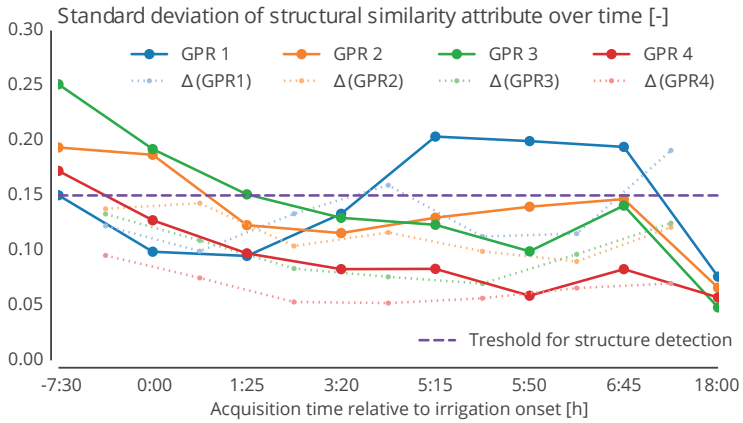

Figure A.6: Standard deviation of the structural similarity attribute at the different GPR transects in the hillslope experiment over time (solid lines) and standard deviation of the differences of two successive attribute distributions (dotted lines). The used threshold for the detection of flow-relevant structures is marked as the dashed purple line.

numerous effects like micro-topography, topsoil conditions, signal attenuation and even weather. The lack of distinguished reflectors also inhibited any estimation of quantitative values. Further, the referenced depths in Fig. 3.12 are only estimates based on a constant mean GPR velocity which can also vary in time and space depending on the initial conditions.

The highlighted assumptions clearly frame the limits of the technique. The overall sensitivity of the approach can be judged from the structural similarity attribute of the last pairs of records in the hillslope experiment when we assume the soil water to be in equilibrium again. Figure A.6 presents the development of the standard deviations of the structural similarity attribute of the respective transects over time. In dotted lines we plotted the standard deviations of the stepwise attribute differences. The standard deviations of the attribute for the last pairs of records is 0.06. Using this value as methodological noise reference, it implies that weak responses and local effects must not be over interpreted. Hence, the introduced threshold of 0.15 for irrigation signal detection appears to be a reasonable choice for qualitative interpretation in our case.

Another limit is the interpretability of changes in the radargrams, as water can have different effects under different situations. A wetted well-defined surface may quickly become a reflector which is easy to detect. However, tortuous flow paths may not be as ideal. Small structures might be well below the limits of detectability in the complex reflection pattern. As such the structural similarity attribute can only detect zones of significant changes which can be induced by many lumped small structures, one big flow path, or even a favorably oriented stone which gets wetted.

References

H. Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974. doi: 10.1109/TAC.1974.1100705.

P. Ala-Aho, C. Soulsby, O. Pokrovsky, S. Kirpotin, J. Karlsson, S. Serikova, S. Vorobyev, R. Manasypov, S. Loiko, and D. Tetzlaff. Using stable isotopes to assess surface water source dynamics and hydrological connectivity in a high-latitude wetland and permafrost influenced landscape. *Journal of Hydrology*, 556: 279–293, 2018. doi: 10.1016/j.jhydrol.2017.11.024.

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

S. E. Allaire, S. Roulier, and A. J. Cessna. Quantifying preferential flow in soils: A review of different techniques. *Journal of Hydrology*, 378(1-2):179–204, 2009. doi: 10.1016/j.jhydrol.2009.08.013.

N. Allroggen and J. Tronicke. Attribute-based analysis of time-lapse ground-penetrating radar data. *Geophysics*, 81(1):H1–H8, 2016. doi: 10.1190/geo2015-0171.1.

N. Allroggen, J. Tronicke, M. Delock, and U. Böniger. Topographic migration of 2D and 3D ground-penetrating radar data considering variable velocities. *Near Surface Geophysics*, 13:253–259, 2015a. doi: 10.3997/1873-0604.2014037.

N. Allroggen, L. van Schaik, and J. Tronicke. 4D ground-penetrating radar during a plot scale dye tracer experiment. *Journal of Applied Geophysics*, 118:139–144, 2015b. doi: 10.1016/j.jappgeo.2015.04.016.

A. Amoozegar. Comparison of the glover solution with the simultaneous-equations approach for measuring hydraulic conductivity. *Soil Science Society of America Journal*, 53(5):1362–1367, 1989. doi: 10.2136/sssaj1989.03615995005300050010x.

C. Amoros and G. Bornette. Connectivity and biocomplexity in waterbodies of riverine floodplains. *Freshwater Biology*, 47(4):761–776, 2002. doi: 10.1046/j.1365-2427.2002.00905.x.

A. E. Anderson, M. Weiler, Y. Alila, and R. O. Hudson. Dye staining and excavation of a lateral preferential flow network. *Hydrology and Earth System Sciences*, 13 (6):935–944, 2009. doi: 10.5194/hess-13-935-2009.

M. Anderson and T. Burt. The role of topography in controlling throughflow generation. *Earth Surface Processes and Landforms*, 3(4):331–344, 1978. doi: 10.1002/esp.3290030402.

L. Angermann, J. Lewandowski, J. H. Fleckenstein, and G. Nützmann. A 3d analysis algorithm to improve interpretation of heat pulse sensor results for the determination of small-scale flow directions and velocities in the hyporheic zone. *Journal of hydrology*, 475:1–11, 2012. doi: 10.1016/j.jhydrol.2012.06.050.

L. Angermann, C. Jackisch, N. Allroggen, M. Sprenger, E. Zehe, J. Tronicke, M. Weiler, and T. Blume. Form and function in hillslope hydrology: characterization of subsurface flow based on response observations. *Hydrology and Earth System Sciences*, 21(7):3727–3748, 2017. doi: 10.5194/hess-21-3727-2017.

R. T. Armstrong, J. E. McClure, M. A. Berrill, M. Rücker, S. Schlüter, and S. Berg. Beyond darcy's law: The role of phase topology and ganglion dynamics for two-fluid flow. *Physical Review E*, 94(4):043113, 2016. doi: 10.1103/PhysRevE.94.043113.

A. Atkinson, I. Cartwright, B. Gilfedder, H. Hofmann, N. Unland, D. Cendón, and R. Chisari. A multi-tracer approach to quantifying groundwater inflows to an upland river; assessing the influence of variable groundwater chemistry. *Hydrological processes*, 29(1):1–12, 2015. doi: 10.1002/hyp.10122.

T. Atkinson. Techniques for measuring subsurface flow on hillslopes. *Hillslope Hydrology*, pages 73–120, 1978.

S. Bachmair and M. Weiler. Hillslope characteristics as controls of subsurface flow variability. *Hydrology and Earth System Sciences*, 16(10):3699–3715, 2012. doi: 10.5194/hess-16-3699-2012.

S. Bachmair and M. Weiler. Interactions and connectivity between runoff generation processes of different spatial scales. *Hydrological Processes*, 28(4):1916–1930, 2014. doi: 10.1002/hyp.9705.

F. K. Barthold and R. A. Woods. Stormflow generation: A meta-analysis of field evidence from small, forested catchments. *Water Resources Research*, 51(5):3730–3753, 2015. doi: 10.1002/2014WR016221.

K. E. Bencala. A perspective on stream-catchment connections. *Journal of the North American Benthological Society*, pages 44–47, 1993.

K. E. Bencala. Hyporheic zone hydrological processes. *Hydrological Processes*, 14(15):2797–2798, 2000. doi: 10.1002/1099-1085(20001030)14:15⟨2797::AID-HYP402⟩3.0.CO;2-6.

K. E. Bencala, M. N. Gooseff, and B. A. Kimball. Rethinking hyporheic flow and transient storage to advance understanding of stream-catchment connections. *Water Resources Research*, 47(3), 2011. doi: 10.1029/2010WR010066.

A. Benson, M. Zane, T. E. Becker, A. Visser, S. H. Uriostegui, E. DeRubeis, J. E. Moran, B. K. Esser, and J. F. Clark. Quantifying reaeration rates in alpine streams using deliberate gas tracer experiments. *Water*, 6(4):1013–1027, 2014. doi: 10.3390/w6041013.

A. Bergstrom, K. Jencso, and B. McGlynn. Spatiotemporal processes that contribute to hydrologic exchange between hillslopes, valley bottoms, and streams. *Water Resources Research*, 2016. doi: 10.1002/2015WR017972.

K. Beven and P. Germann. Macropores and water flow in soils. *Water Resources Research*, 18(5):1311–1325, 1982. doi: 10.1029/WR018i005p01311.

K. Beven and P. Germann. Macropores and water flow in soils revisited. *Water Resources Research*, 49(6):3071–3092, 2013. doi: 10.1002/wrcr.20156.

K. J. Beven and M. J. Kirkby. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, 24(1):43–69, 1979. doi: 10.1080/02626667909491834.

A. Binley, S. S. Hubbard, J. A. Huisman, R. André, D. A. Robinson, K. Singha, and L. D. Slater. The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. *Water Resources Research*, 51:3837–3866, 2015. doi: 10.1002/2015WR017016.

R. Birken and R. Versteeg. Use of four-dimensional ground penetrating radar and advanced visualization methods to determine subsurface fluid migration. *Journal of Applied Geophysics*, 43:215–226, 2000. doi: 10.1016/S0926-9851(99)00060-9.

S. J. Birkinshaw and B. Webb. Flow pathways in the Slapton Wood catchment using temperature as a tracer. *Journal of Hydrology*, 383(3):269–279, 2010. doi: 10.1016/j.jhydrol.2009.12.042.

J. M. Bishop, M. V. Callaghan, E. E. Cey, and L. R. Bentley. Measurement and simulation of subsurface tracer migration to tile drains in low permeability, macroporous soil. *Water Resources Research*, 51(6):3956–3981, 2015. doi: 10.1002/2014WR016310.

K. C. Blits. Aristotle: form, function, and comparative anatomy. *The Anatomical Record*, 257(2):58–63, 1999. doi: 10.1002/(SICI)1097-0185(19990415)257:2⟨58::AID-AR6⟩3.0.CO;2-I.

G. Blöschl. Scaling issues in snow hydrology. *Hydrological processes*, 13(14-15):2149–2175, 1999. doi: 10.1002/(SICI)1099-1085(199910)13:14/15⟨2149::AID-HYP847⟩3.0.CO;2-8.

G. Blöschl. Scaling in hydrology. *Hydrological Processes*, 15(4):709–711, 2001. doi: 10.1002/hyp.432.

G. Blöschl and M. Sivapalan. Scale issues in hydrological modelling: a review. *Hydrological processes*, 9(3-4):251–290, 1995. doi: 10.1002/hyp.3360090305.

M. Blouin, M. E. Hodson, E. A. Delgado, G. Baker, L. Brussaard, K. R. Butt, J. Dai, L. Dendooven, G. Pérès, J. Tondoh, et al. A review of earthworm impact on soil function and ecosystem services. *European Journal of Soil Science*, 64(2): 161–182, 2013. doi: 10.1111/ejss.12025.

T. Blume and H. I. van Meerveld. From hillslope to stream: methods to investigate subsurface connectivity. *Wiley Interdisciplinary Reviews: Water*, 2(3):177–198, 2015. doi: 10.1002/wat2.1071.

F. Boano, C. Camporeale, R. Revelli, and L. Ridolfi. Sinuosity-driven hyporheic exchange in meandering rivers. *Geophysical Research Letters*, 33(18), 2006. doi: 10.1029/2006GL027630.

F. Boano, R. Revelli, and L. Ridolfi. Bedform-induced hyporheic exchange with unsteady flows. *Advances in water resources*, 30(1):148–156, 2007. doi: 10.1016/j.advwatres.2006.03.004.

U. Boeniger and J. Tronicke. On the Potential of Kinematic GPR Surveying Using a Self-Tracking Total Station: Evaluating System Crosstalk and Latency. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10):3792–3798, 2010. doi: 10.1109/TGRS.2010.2048332.

P. Bogaart and P. Troch. Curvature distribution within hillslopes and catchments and its effect on the hydrological response. *Hydrol. Earth Syst. Sci*, 10:925–936, 2006.

A. J. Boulton. Hyporheic rehabilitation in rivers: restoring vertical connectivity. *Freshwater Biology*, 52(4):632–650, 2007. doi: 10.1111/j.1365-2427.2006.01710.x.

A. J. Boulton, S. Findlay, P. Marmonier, E. H. Stanley, and H. M. Valett. The functional significance of the hyporheic zone in streams and rivers. *Annual Review of Ecology and Systematics*, pages 59–81, 1998. doi: 10.1146/annurev.ecolsys.29.1.59.

A. J. Boulton, G. D. Fenwick, P. J. Hancock, and M. S. Harvey. Biodiversity, functional roles and ecosystem services of groundwater invertebrates. *Invertebrate Systematics*, 22(2):103–116, 2008. doi: 10.1071/IS07024.

J. Bouma and L. W. Dekker. A case study on infiltration into dry clay soil i. morphological observations. *Geoderma*, 20(1):27–40, 1978. doi: 10.1016/0016-7061(78)90047-2.

L. Bracken, J. Wainwright, G. Ali, D. Tetzlaff, M. Smith, S. Reaney, and A. Roy. Concepts of hydrological connectivity: research approaches, pathways and future agendas. *Earth-science reviews*, 119:17–34, 2013. doi: 10.1016/j.earscirev.2013.02.001.

L. J. Bracken, L. Turnbull, J. Wainwright, and P. Bogaart. Sediment connectivity: a framework for understanding sediment transfer at multiple scales. *Earth Surface Processes and Landforms*, 40(2):177–188, 2015. doi: 10.1002/esp.3635.

J. H. Bradford, J. T. Harper, and J. Brown. Complex dielectric permittivity measurements from ground-penetrating radar data to estimate snow liquid water content in the pendular regime. *Water Resources Research*, 45(8):1–12, 2009. doi: 10.1029/2008WR007341.

D. D. Brammer and J. J. McDonnell. An evolving perceptual model of hillslope flow at the maimai catchment. *Advances in hillslope processes*, 1:35–60, 1996.

M. A. Briggs, L. K. Lautz, and J. M. McKenzie. A comparison of fibre-optic distributed temperature sensing to traditional methods of evaluating groundwater inflow to streams. *Hydrological Processes*, 26(9):1277–1290, 2012. doi: 10.1002/hyp.8200.

M. Brunke and T. Gonser. The ecological significance of exchange processes between rivers and groundwater. *Freshwater biology*, 37(1):1–33, 1997. doi: 10.1046/j.1365-2427.1997.00143.x.

J. M. Buffington and D. Tonina. Hyporheic exchange in mountain rivers ii: effects of channel morphology on mechanics, scales, and rates of exchange. *Geography Compass*, 3(3):1038–1062, 2009. doi: 10.1111/j.1749-8198.2009.00225.x.

W. C. Burnett and H. Dulaiova. Estimating the dynamics of groundwater input into the coastal zone via continuous radon-222 measurements. *Journal of environmental radioactivity*, 69(1-2):21–35, 2003. doi: 10.1016/S0265-931X(03)00084-5.

D. A. Burns, J. J. McDonnell, R. P. Hooper, N. E. Peters, J. E. Freer, C. Kendall, and K. Beven. Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the panola mountain research watershed (georgia, usa). *Hydrological Processes*, 15(10):1903–1924, 2001. doi: 10.1002/hyp.246.

T. Burt. A third paradox in catchment hydrology and biogeochemistry: decoupling in the riparian zone. *Hydrological Processes*, 19(10):2087–2089, 2005. doi: 10.1002/hyp.5904.

T. Burt and D. Butcher. Topographic controls of soil moisture distributions. *Journal of Soil Science*, 36(3):469–486, 1985. doi: 10.1111/j.1365-2389.1985.tb00351.x.

J. Buttle. Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins. *Progress in Physical Geography*, 18(1):16–41, 1994.

M. B. Cardenas. Stream-aquifer interactions and hyporheic exchange in gaining and losing sinuous streams. *Water Resources Research*, 45(6), 2009. doi: 10.1029/2008WR007651.

M. B. Cardenas and J. Wilson. Hydrodynamics of coupled flow above and below a sediment–water interface with triangular bedforms. *Advances in water resources*, 30(3):301–313, 2007a. doi: 10.1016/j.advwatres.2006.06.009.

M. B. Cardenas and J. L. Wilson. Dunes, turbulent eddies, and interfacial exchange with permeable sediments. *Water Resources Research*, 43(8), 2007b. doi: 10.1029/2006WR005787.

R. F. Carsel and R. S. Parrish. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5):755–769, 1988. doi: 10.1029/WR024i005p00755.

I. Cartwright and H. Hofmann. Using radon to understand parafluvial flows and the changing locations of groundwater inflows in the avon river, south-east australia. *Hydrology and Earth System Sciences*, 20(9):3581, 2016. doi: 10.5194/hess-20-3581-2016.

I. Cartwright, H. Hofmann, B. Gilfedder, and B. Smyth. Understanding parafluvial exchange and degassing to better quantify groundwater inflows using 222 rn: the king river, southeast australia. *Chemical Geology*, 380:48–60, 2014. doi: 10.1016/j.chemgeo.2014.04.009.

G. Cassiani, V. Bruno, A. Villa, N. Fusi, and A. M. Binley. A saline trace test monitored via time-lapse surface electrical resistivity tomography. *Journal of Applied Geophysics*, 59(3):244–259, 2006. doi: 10.1016/j.jappgeo.2005.10.007.

M. Chafiq, J. Gibert, and C. Claret. Interactions among sediments, organic matter, and microbial activity in the hyporheic zone of an intermittent stream. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(3):487–495, 1999. doi: 10.1139/f98-208.

R. Colbach and R. Maquil. *Carte géologique du Luxembourg 1:25000-Feuille n° 7 Rédange*. Service Géologique du Luxembourg, Luxembourg, 2003.

P. Cook, G. Favreau, J. Dighton, and S. Tickell. Determining natural groundwater influx to a tropical river using radon, chlorofluorocarbons and ionic environmental tracers. *Journal of Hydrology*, 277(1-2):74–88, 2003. doi: 10.1016/S0022-1694(03)00087-8.

P. Cook, S. Lamontagne, D. Berhane, and J. Clark. Quantifying groundwater discharge to cockburn river, southeastern australia, using dissolved gas tracers 222rn and sf6. *Water Resources Research*, 42(10), 2006. doi: 10.1029/2006WR004921.

P. G. Cook. Estimating groundwater discharge to rivers from river chemistry surveys. *Hydrological Processes*, 27(25):3694–3707, 2013. doi: 10.1002/hyp.9493.

T. P. Covino and B. L. McGlynn. Stream gains and losses across a mountain-to-valley transition: Impacts on watershed hydrology and stream water chemistry. *Water Resources Research*, 43(10), 2007. doi: 10.1029/2006WR005544.

M. Dahl, B. Nilsson, J. Langhoff, and J. Refsgaard. Review of classification systems and new multi-scale typology of groundwater–surface water interaction. *Journal of Hydrology*, 344(1-2):1–16, 2007. doi: 10.1016/j.jhydrol.2007.06.027.

H. Darcy. *Les fontaines publiques de la ville de Dijon: exposition et application...* Victor Dalmont, 1856.

T. Datry, F. Malard, and J. Gibert. Response of invertebrate assemblages to increased groundwater recharge rates in a phreatic aquifer. *Journal of the North American Benthological Society*, 24(3):461–477, 2005. doi: 10.1899/04-140.1.

T. Day and T. Day. Field procedures and evaluation of a slug dilution gauging method in mountain streams. *Journal of Hydrology (New Zealand)*, pages 113–133, 1977.

J. De Gruijter, D. J. Brus, M. F. Bierkens, and M. Knotters. *Sampling for natural resource monitoring*. Springer Science & Business Media, 2006.

G. De Marsily, F. Delay, J. Gonçalvès, P. Renard, V. Teles, and S. Violette. Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13(1):161–183, 2005. doi: 10.1007/s10040-004-0432-3.

M. Delorme, B. Atfeh, V. Allken, and B. Bourbiaux. Upscaling improvement for heterogeneous fractured reservoir using a geostatistical connectivity index. *Geostats 2008*, 2008.

N. T. Dimova, W. C. Burnett, J. P. Chanton, and J. E. Corbett. Application of radon-222 to investigate groundwater discharge into small shallow lakes. *Journal of Hydrology*, 486:112–122, 2013. doi: 10.1016/j.jhydrol.2013.01.043.

M.-J. Dole-Olivier, P. Marmonier, and J.-L. Beffy. Response of invertebrates to lotic disturbance: is the hyporheic zone a patchy refugium? *Freshwater Biology*, 37(2):257–276, 1997. doi: 10.1046/j.1365-2427.1997.00140.x.

J. H. Duff and F. J. Triska. Denitrifications in sediments from the hyporheic zone adjacent to a small forested stream. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(6):1140–1147, 1990. doi: 10.1139/f90-133.

J. H. Duff and F. J. Triska. *Nitrogen biogeochemistry and surface-subsurface exchange in streams*. Academic, 2000.

K. J. Edwardson, W. B. Bowden, C. Dahm, and J. Morrice. The hydraulic characteristics and geochemistry of hyporheic and parafluvial zones in arctic tundra streams, north slope, alaska. *Advances in Water Resources*, 26(9):907–923, 2003. doi: 10.1016/S0309-1708(03)00078-2.

K. K. Ellins, A. Roman-Mas, and R. Lee. Using 222rn to examine groundwater/surface discharge interaction in the rio grande de manati, puerto rico. *Journal of Hydrology*, 115(1-4):319–341, 1990. doi: 10.1016/0022-1694(90)90212-G.

A. H. Elliott and N. H. Brooks. Transfer of nonsorbing solutes to a streambed with bed forms: Laboratory experiments. *Water Resources Research*, 33(1):137–151, 1997a. doi: 10.1029/96WR02783.

A. H. Elliott and N. H. Brooks. Transfer of nonsorbing solutes to a streambed with bed forms: Theory. *Water Resources Research*, 33(1):123–136, 1997b. doi: 10.1029/96WR02784.

T. Endreny and L. Lautz. Comment on "M. Munz, S. Krause, C. Tecklenburg, and A. Binley. Reducing monitoring gaps at the aquifer–river interface by modelling groundwater–surfacewater exchange flow patterns.". *Hydrological Processes*, 26 (10):1586–1588, 2012. doi: 10.1002/hyp.8410.

T. Endreny, L. Lautz, and D. Siegel. Hyporheic flow path response to hydraulic jumps at river steps: Flume and hydrodynamic models. *Water Resources Research*, 47(2), 2011. doi: 10.1029/2009WR008631.

European Comission. Water Framework Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, 2000.

K. D. Fausch, C. E. Torgersen, C. V. Baxter, and H. W. Li. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes: A continuous view of the river is needed to understand how processes interacting among scales set the context for stream fishes and their habitat. *AIBS Bulletin*, 52(6):483–498, 2002. doi: 10.1641/0006-3568(2002)052[0483:LTRBTG]2.0.CO;2.

F. Fenicia, D. Kavetski, H. H. G. Savenije, M. P. Clark, G. Schoups, L. Pfister, and J. Freer. Catchment properties, function, and conceptual model representation: is there a correspondence? *Hydrological Processes*, 28(7):2451–2467, 2014. doi: 10.1002/hyp.9726.

S. E. Findlay, R. L. Sinsabaugh, W. V. Sobczak, and M. Hoostal. Metabolic and structural response of hyporheic microbial communities to variations in supply of dissolved organic matter. *Limnology and oceanography*, 48(4):1608–1617, 2003. doi: 10.4319/lo.2003.48.4.1608.

S. G. Fisher, N. B. Grimm, E. Martí, R. M. Holmes, and J. B. Jones Jr. Material spiraling in stream corridors: a telescoping ecosystem model. *Ecosystems*, 1(1): 19–34, 1998. doi: 10.1007/s100219900003.

J. H. Fleckenstein, R. G. Niswonger, and G. E. Fogg. River-aquifer interactions, geologic heterogeneity, and low-flow management. *Ground water*, 44(6):837–852, 2006. doi: 10.1111/j.1745-6584.2006.00190.x.

J. H. Fleckenstein, S. Krause, D. M. Hannah, and F. Boano. Groundwater-surface water interactions: New methods and models to improve understanding of processes and dynamics. *Advances in Water Resources*, 33(11):1291–1295, 2010. doi: 10.1016/j.advwatres.2010.09.011.

M. Flury, H. Flühler, W. A. Jury, and J. Leuenberger. Susceptibility of soils to preferential flow of water: A field study. *Water resources research*, 30(7):1945–1954, 1994. doi: 10.1029/94WR00871.

J. Freer, J. J. McDonnell, K. Beven, N. E. Peters, D. A. Burns, R. Hooper, B. Aulenbach, and C. Kendall. The role of bedrock topography on subsurface storm flow. *Water Resources Research*, 38(12), 2002. doi: 10.1029/2001WR000872.

S. Frei and B. Gilfedder. Finiflux: An implicit finite element model for quantification of groundwater fluxes and hyporheic exchange in streams and rivers using radon. *Water Resources Research*, 51(8):6776–6786, 2015. doi: 10.1002/2015WR017212.

S. Frei, J. Fleckenstein, S. Kollet, and R. Maxwell. Patterns and dynamics of river–aquifer exchange with variably-saturated flow using a fully-coupled model. *Journal of Hydrology*, 375(3):383–393, 2009. doi: 10.1016/j.jhydrol.2009.06.038.

K. A. Fryirs, G. J. Brierley, N. J. Preston, and M. Kasai. Buffers, barriers and blankets: the (dis) connectivity of catchment-scale sediment cascades. *Catena*, 70 (1):49–67, 2007. doi: 10.1016/j.catena.2006.07.007.

S. Garré, I. Coteur, C. Wongleecharoen, T. Kongkaew, J. Diels, and J. Vanderborght. Noninvasive Monitoring of Soil Water Dynamics in Mixed Cropping Systems: A Case Study in Ratchaburi Province, Thailand. *Vadose Zone Journal*, 12(2):–, 2013. doi: 10.2136/vzj2012.0129.

D. P. Genereux and H. F. Hemond. Naturally occurring radon 222 as a tracer for streamflow generation: Steady state methodology and field example. *Water resources research*, 26(12):3065–3075, 1990. doi: 10.1029/WR026i012p03065.

D. P. Genereux and H. F. Hemond. Determination of gas exchange rate constants for a small stream on walker branch watershed, tennessee. *Water Resources Research*, 28(9):2365–2374, 1992. doi: 10.1029/92WR01083.

H. H. Gerke. Preferential flow descriptions for structured soils. *Journal of Plant Nutrition and Soil Science*, 169(3):382–400, 2006. doi: 10.1002/jpln.200521955.

H. H. Gerke. Macroscopic representation of the interface between flow domains in structured soil. *Vadose Zone Journal*, 11(3), 2012. doi: 10.2136/vzj2011.0125.

H. H. Gerke, P. Germann, and J. Nieber. Preferential and Unstable Flow: From the Pore to the Catchment Scale. *Vadose Zone Journal*, 9(2):207, 2010. doi: 10.2136/vzj2010.0059.

P. F. Germann. Preferential Flow. Geographica Bernensia, Institute of Geography, University of Bern, Bern, 2014. ISBN 9783905835342.

P. F. Germann and S. A. al Hagrey. Gravity-driven and viscosity-dominated infiltration into a full-scale sand model. *Vadose zone journal*, 7(4):1160–1169, 2008. doi: 10.2136/vzj2007.0172.

P. F. Germann and M. Karlen. Viscous-flow approach to in situ infiltration and in vitro saturated hydraulic conductivity determination. *Vadose zone journal*, 15(2), 2016. doi: 10.2136/vzj2015.05.0065.

M. N. Gooseff and B. L. McGlynn. A stream tracer technique employing ionic tracers and specific conductance data applied to the maimai catchment, new zealand. *Hydrological processes*, 19(13):2491–2506, 2005. doi: 10.1002/hyp.5685.

M. N. Gooseff, S. M. Wondzell, R. Haggerty, and J. Anderson. Comparing transient storage modeling and residence time distribution (rtd) analysis in geomorphically varied reaches in the lookout creek basin, oregon, usa. *Advances in Water Resources*, 26(9):925–937, 2003. doi: 10.1016/S0309-1708(03)00105-2.

K. H. Gormally, M. S. McIntosh, A. N. Mucciardi, and G. W. McCarty. Ground-penetrating radar detection and three-dimensional mapping of lateral macropores: Ii. riparian application. *Soil Science Society of America Journal*, 75(4):1236–1243, 2011. doi: 10.2136/sssaj2010.0342.

T. Graeff, E. Zehe, D. Reusser, E. Lück, B. Schröder, G. Wenk, H. John, and A. Bronstert. Process identification through rejection of model structures in a mid-mountainous rural catchment: observations of rainfall–runoff response, geophysical conditions and model inter-comparison. *Hydrological Processes*, 23:702–718, 2009. doi: 10.1002/hyp.7171.

C. B. Graham, R. A. Woods, and J. J. McDonnell. Hillslope threshold response to rainfall:(1) a field based forensic approach. *Journal of Hydrology*, 393(1-2):65–76, 2010. doi: 10.1016/j.jhydrol.2009.12.015.

R. Grayson and G. Blöschl. *Spatial patterns in catchment hydrology: observations and modelling*. CUP Archive, 2001.

L. Guo, J. Chen, and H. Lin. Subsurface lateral preferential flow network revealed by time-lapse ground-penetrating radar in a hillslope. *Water Resources Research*, 50(12):9127–9147, Oct. 2014. doi: 10.1002/2013WR014603.

H. V. Gupta, M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye. Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, 48(8), 2012. doi: 10.1029/2011WR011044.

E. B. Haarder, M. C. Looms, K. H. Jensen, and L. Nielsen. Visualizing unsaturated flow phenomena using high-resolution reflection ground penetrating radar. *Vadose Zone Journal*, 10(1):84–97, 2011. doi: 10.2136/vzj2009.0188.

R. Haggerty, S. M. Wondzell, and M. A. Johnson. Power-law residence time distribution in the hyporheic zone of a 2nd-order mountain stream. *Geophysical Research Letters*, 29(13), 2002. doi: 10.1029/2002GL014743.

J. Harvey and K. Bencala. The effect of streambed topography on surface-subsurface water exchange in mountain catchments. *Water Resources Research*, 29(1):89–98, 1993. doi: 10.1029/92WR01960.

H. Hellebrand, C. Müller, P. Matgen, F. Fenicia, and H. Savenije. A process proof test for model concepts: modelling the meso-scale. *Physics and Chemistry of the Earth, Parts A/B/C*, 36(1-4):42–53, 2011. doi: 10.1016/j.pce.2010.07.019.

K. Heller and A. Kleber. Einfluss periglazialer Deckschichten auf die oberflächennahen Fließwege am Hang – eine experimentelle Prozessstudie im Osterzgebirge, Sachsen. In *Hydrologie & Wasserwirtschaft-von der Theorie zur Praxis. Forum für Hydrologie und Wasserbewirtschaftung*, volume 30, pages 50–56, 2011.

K. Heller and A. Kleber. Hillslope runoff generation influenced by layered subsurface in a headwater catchment in ore mountains, germany. *Environmental Earth Sciences*, 75(11):943, 2016. doi: 10.1007/s12665-016-5750-y.

R. Helmig, A. Weiss, and B. I. Wohlmuth. Dynamic capillary effects in heterogeneous porous media. *Computational Geosciences*, 11(3):261–274, 2007. doi: 10.1007/s10596-007-9050-1.

E. T. Hester and M. N. Gooseff. Moving beyond the banks: Hyporheic restoration is fundamental to restoring ecological services and functions of streams, 2010.

A. R. Hill. Nitrate removal in stream riparian zones. *Journal of environmental quality*, 25(4):743–755, 1996. doi: 10.2134/jeq1996.00472425002500040014x.

A. R. Hill and M. Cardaci. Denitrification and organic carbon availability in riparian wetland soils and subsurface sediments. *Soil Science Society of America Journal*, 68(1):320–325, 2004. doi: 10.2136/sssaj2004.3200.

A. R. Hill and T. P. Duval. Beaver dams along an agricultural stream in southern ontario, canada: their impact on riparian zone hydrology and nitrogen chemistry. *Hydrological processes*, 23(9):1324–1336, 2009. doi: 10.1002/hyp.7249.

D. Hillel. *Fundamentals of soil physics*. Academic press, 1980.

S. Hinkle, J. Duff, F. Triska, A. Laenen, E. Gates, K. Bencala, D. Wentz, and S. Silva. Linking hyporheic flow and nitrogen cycling near the willamette river—a large river in oregon, usa. *Journal of Hydrology*, 244(3):157–180, 2001. doi: 10.1016/S0022-1694(01)00335-3.

E. Hoehn and H. Von Gunten. Radon in groundwater: A tool to assess infiltration from surface waters to aquifers. *Water Resources Research*, 25(8):1795–1803, 1989. doi: 10.1029/WR025i008p01795.

J. Holden. Hydrological connectivity of soil pipes determined by ground-penetrating radar tracer detection. *Earth Surface Processes and Landforms*, 29(4):437–442, 2004a. doi: 10.1002/esp.1039.

J. Holden. Hydrological connectivity of soil pipes determined by ground-penetrating radar tracer detection. *Earth Surface Processes and Landforms*, 29(4):437–442, 2004b. doi: 10.1002/esp.1039.

H. Holländer, H. Bormann, T. Blume, W. Buytaert, G. Chirico, J.-F. Exbrayat, D. Gustafsson, H. Hölzel, T. Krauße, P. Kraft, et al. Impact of modellers' decisions on hydrological a priori predictions. *Hydrology and Earth System Sciences*, 18(6): 2065, 2014. doi: 10.5194/hess-18-2065-2014.

R. P. Hooper. Applying the scientific method to small catchment studies: a review of the panola mountain experience. *Hydrological Processes*, 15(10):2039–2050, 2001. doi: 10.1002/hyp.255.

R. P. Hooper, B. T. Aulenbach, D. A. Burns, J. McDonnell, J. Freer, C. Kendall, and K. Beven. Riparian control of stream-water chemistry: implications for hydrochemical basin models. *International Association of Hydrological Sciences, Publication*, 248:451–458, 1998.

L. Hopp and J. McDonnell. Connectivity at the hillslope scale: Identifying interactions between storm size, bedrock permeability, slope angle and soil depth. *Journal of Hydrology*, 376(3-4):378–391, 2009. doi: 10.1016/j.jhydrol.2009.07.047.

R. Hübner, K. Heller, T. Günther, and A. Kleber. Monitoring hillslope moisture dynamics with surface ERT for enhancing spatial significance of hydrometric point measurements. *Hydrology and Earth System Sciences*, 19(1):225–240, 2015. doi: 10.5194/hess-19-225-2015.

J. A. Huisman, S. S. Hubbard, J. D. Redman, and A. P. Annan. Measuring soil water content with ground penetrating radar. *Vadose zone journal*, 2(4):476–491, 2003. doi: 10.2136/vzj2003.4760.

IUSS Working Group WRB. World reference base for soil resources. *World Soil Resources Report*, 103, 2006.

C. Jackisch. Linking structure and functioning of hydrological systems – How to achieve necessary experimental and model complexity with adequate effort, 2015. Dissertation.

C. Jackisch, L. Angermann, N. Allroggen, M. Sprenger, T. Blume, J. Tronicke, and E. Zehe. Form and function in hillslope hydrology: in situ imaging and characterization of flow-relevant structures. *Hydrology and Earth System Sciences*, 21(7):3749–3775, 2017. doi: 10.5194/hess-21-3749-2017.

K. G. Jencso and B. L. McGlynn. Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation. *Water Resources Research*, 47(11), 2011. doi: 10.1029/2011WR010666.

K. G. Jencso, B. L. McGlynn, M. N. Gooseff, S. M. Wondzell, K. E. Bencala, and L. A. Marshall. Hydrologic connectivity between landscapes and streams: Transferring reach-and plot-scale understanding to the catchment scale. *Water Resources Research*, 45(4), 2009. doi: 10.1029/2008WR007225.

H. M. Jol, editor. *Ground penetrating radar theory and applications*. Elsevier Science, Amsterdam, 1 edition, 2009.

J. Jones, E. Sudicky, A. Brookfield, and Y.-J. Park. An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow. *Water Resources Research*, 42(2), 2006. doi: 10.1029/2005WR004130.

J. B. Jones, R. M. Holmes, S. G. Fisher, N. B. Grimm, and D. M. Greene. Methanogenesis in arizona, usa dryland streams. *Biogeochemistry*, 31(3):155–173, 1995. doi: 10.1007/BF00004047.

J. Juilleret, J. Iffly, L. Pfister, and C. Hissler. Remarkable Pleistocene periglacial slope deposits in Luxembourg (Oesling): pedological implication and geosite potential. *Bulletin de la Société des Naturalistes Luxembourgeois*, 112:125–130, 2011.

W. A. Jury, K. Roth, et al. *Transfer functions and solute movement through soil: theory and applications*. Birkhäuser Verlag AG, 1990. ISBN 3764325097.

E. Kalbus, F. Reinstorf, and M. Schirmer. Measuring methods for groundwater-surface water interactions: a review. *Hydrology and Earth System Sciences Discussions*, 10(6):873–887, 2006.

T. Kasahara and A. R. Hill. Modeling the effects of lowland stream restoration projects on stream–subsurface water exchange. *Ecological Engineering*, 32(4): 310–319, 2008. doi: 10.1016/j.ecoleng.2007.12.006.

T. Kasahara and S. M. Wondzell. Geomorphic controls on hyporheic exchange flow in mountain streams. *Water Resources Research*, 39(1), 2003. doi: 10.1029/2002WR001386.

D. H. Käser, A. Binley, A. L. Heathwaite, and S. Krause. Spatio-temporal variations of hyporheic flow in a riffle-step-pool sequence. *Hydrological Processes*, 23(15): 2138–2149, 2009. doi: 10.1002/hyp.7317.

R. Kasteel, H.-J. Vogel, and K. Roth. Effect of non-linear adsorption on the transport behaviour of brilliant blue in a field soil. *European Journal of Soil Science*, 53(2): 231–240, 2002. doi: 10.1046/j.1365-2389.2002.00437.x.

D. Kavetski, F. Fenicia, and M. P. Clark. Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment. *Water Resources Research*, 47(5), 2011. doi: 10.1029/2010WR009525.

G. Kim and D.-W. Hwang. Tidal pumping of groundwater into the coastal ocean revealed from submarine 222rn and ch4 monitoring. *Geophysical Research Letters*, 29(14), 2002. doi: 10.1029/2002GL015093.

J. W. Kirchner. A double paradox in catchment hydrology and geochemistry. *Hydrological Processes*, 17(4):871–874, 2003. doi: 10.1002/hyp.5108.

M. Kirkby, L. Bracken, and S. Reaney. The influence of land use, soils and topography on the delivery of hillslope runoff to channels in se spain. *Earth Surface Processes and Landforms*, 27(13):1459–1473, 2002. doi: 10.1002/esp.441.

J. Klaus and J. McDonnell. Hydrograph separation using stable isotopes: Review and evaluation. *Journal of Hydrology*, 505:47–64, 2013. doi: 10.1016/j.jhydrol.2013.09.006.

J. Klaus and E. Zehe. A novel explicit approach to model bromide and pesticide transport in connected soil structures. *Hydrology and Earth System Sciences*, 15 (7):2127, 2011. doi: 10.5194/hess-15-2127-2011.

J. Klaus, E. Zehe, M. Elsner, C. Külls, and J. J. McDonnell. Macropore flow of old water revisited: experimental insights from a tile-drained hillslope. *Hydrology and Earth System Sciences*, 17(1):103–118, Jan. 2013. doi: 10.5194/hess-17-103-2013.

P. Klenk, S. Jaumann, and K. Roth. Quantitative high-resolution observations of soil water dynamics in a complicated architecture using time-lapse ground-penetrating radar. *Hydrology and Earth System Sciences*, 19(3):1125–1139, 2015. doi: 10.5194/hess-19-1125-2015.

K. Koch, J. Wenninger, S. Uhlenbrook, and M. Bonell. Joint interpretation of hydrological and geophysical data: electrical resistivity tomography results from a process hydrological research site in the black forest mountains, germany. *Hydrological Processes*, 23(10):1501–1513, 2009. doi: 10.1002/hyp.7275.

J. Koestel and M. Larsbo. Imaging and quantification of preferential solute transport in soil macropores. *Water Resources Research*, 50(5):4357–4378, 2014. doi: 10.1002/2014WR015351.

A. Koestler et al. *The ghost in the machine*. Hutchinson London, 1967.

J. M. Köhne, S. Köhne, and J. Šimůnek. A review of model applications for structured soils: a) Water flow and tracer transport. *Journal of Contaminant Hydrology*, 104(1-4):4–35, 2009. doi: 10.1016/j.jconhyd.2008.10.002.

G. Kondolf, A. Boulton, S. O'Daniel, G. Poole, F. Rahel, E. Stanley, E. Wohl, A. Bång, J. Carlstrom, C. Cristoni, et al. Process-based ecological river restoration: visualizing three-dimensional connectivity and dynamic vectors to recover lost linkages. *Ecology and society*, 11(2), 2006.

S. Krause, J. Jacobs, A. Voss, A. Bronstert, and E. Zehe. Assessing the impact of changes in landuse and management practices on the diffuse pollution and retention of nitrate in a riparian floodplain. *Science of the total Environment*, 389 (1):149–164, 2008. doi: 10.1016/j.scitotenv.2007.08.057.

S. Krause, L. Heathwaite, A. Binley, and P. Keenan. Nitrate concentration changes at the groundwater-surface water interface of a small cumbrian river. *Hydrological Processes*, 23(15):2195–2211, 2009. doi: 10.1002/hyp.7213.

S. Krause, D. Hannah, and T. Blume. Heat transport patterns at pool-riffle sequences of an uk lowland stream. *Ecohydrol. J*, 4(4):549–563, 2011a.

S. Krause, D. Hannah, J. Fleckenstein, C. Heppell, D. Kaeser, R. Pickup, G. Pinay, A. Robertson, and P. Wood. Inter-disciplinary perspectives on processes in the hyporheic zone. *Ecohydrology*, 4(4):481–499, 2011b. doi: 10.1002/eco.176.

S. Krause, T. Blume, and N. Cassidy. Application of fibre-optic dts to identify streambed controls on aquifer-river exchange fluxes in lowland rivers. *Hydrol. Earth Syst. Sci*, 16(6):1775–1792, 2012a. doi: 10.5194/hess-16-1775-2012.

S. Krause, M. Munz, C. Tecklenburg, and A. Binley. The effect of groundwater forcing on hyporheic exchange: Reply to comment on 'Munz M, Krause S, Tecklenburg C, Binley A. Reducing monitoring gaps at the aquifer-river interface by modelling groundwater-surfacewater exchange flow patterns. Hydrological Processes. *Hydrological Processes*, 26(10):1589–1592, 2012b. doi: 10.1002/hyp.8080.

S. Lambot, A. Binley, E. Slob, and S. Hubbard. Ground Penetrating Radar in Hydrogeophysics. *Vadose Zone Journal*, 7(1):137, 2008. doi: 10.2136/vzj2007.0180.

S. Lane, C. Brookes, M. Kirkby, and J. Holden. A network-index-based version of TOPMODEL for use with high-resolution digital topographic data. *Hydrological processes*, 18(1):191–201, 2004. doi: 10.1002/hyp.5208.

S. Lane, S. Reaney, and A. L. Heathwaite. Representation of landscape hydrological connectivity using a topographically driven surface flow index. *Water Resources Research*, 45(8), 2009. doi: 10.1029/2008WR007336.

M. Lang, O. McDonough, G. McCarty, R. Oesterling, and B. Wilen. Enhanced detection of wetland-stream connectivity using LiDAR. *Wetlands*, 32(3):461–473, 2012. doi: 10.1007/s13157-012-0279-7.

C. Langhans, G. Govers, J. Diels, A. Leys, W. Clymans, A. Van den Putte, and J. Valckx. Experimental rainfall–runoff data: Reconsidering the concept of infiltration capacity. *Journal of hydrology*, 399(3-4):255–262, 2011. doi: 10.1016/j.jhydrol.2011.01.005.

L. G. Larsen, J. Choi, M. K. Nungesser, and J. W. Harvey. Directional connectivity in hydrology and ecology. *Ecological applications*, 22(8):2204–2220, 2012. doi: 10.1890/11-1948.1.

L. K. Lautz, N. T. Kranes, and D. I. Siegel. Heat tracing of heterogeneous hyporheic exchange adjacent to in-stream geomorphic features. *Hydrological processes*, 24 (21):3074–3086, 2010. doi: 10.1002/hyp.7723.

F. Leaney, K. Smettem, and D. Chittleborough. Estimating the contribution of preferential flow to subsurface runoff from a hillslope using deuterium and chloride. *Journal of Hydrology*, 147(1-4):83–103, 1993. doi: 10.1016/0022-1694(93)90076-L.

H. Lee, E. Zehe, and M. Sivapalan. Predictions of rainfall-runoff response and soil moisture dynamics in a microscale catchment using the crew model. *Hydrol. Earth Syst. Sci.*, 11:819–849, 2007. doi: 10.5194/hess-11-819-2007.

E. Léger, A. Saintenoy, and Y. Coquet. Hydrodynamic parameters of a sandy soil determined by ground-penetrating radar inside a single ring infiltrometer. *Water Resources Research*, 50(7):5459–5474, 2014. doi: 10.1002/2013WR014226.

C. Leibundgut, P. Maloszewski, and C. Külls. *Tracers in hydrology*. John Wiley & Sons, 2011.

M. Leue, H. H. Gerke, and S. C. Godow. Droplet infiltration and organic matter composition of intact crack and biopore surfaces from clay-illuvial horizons. *Journal of Plant Nutrition and Soil Science*, 178(2):250–260, 2015. doi: 10.1002/jpln.201400209.

J. Lewandowski and G. Nützmann. Nutrient retention and release in a floodplain's aquifer and in the hyporheic zone of a lowland river. *Ecological Engineering*, 36 (9):1156–1166, 2010. doi: 10.1016/j.ecoleng.2010.01.005.

J. Lewandowski, L. Angermann, G. Nützmann, and J. H. Fleckenstein. A heat pulse technique for the determination of small-scale flow directions and flow velocities in the streambed of sand-bed streams. *Hydrological Processes*, 25(20):3244–3255, 2011. doi: 10.1002/hyp.8062.

I. Lexartza-Artza and J. Wainwright. Hydrological connectivity: Linking concepts with practical implications. *Catena*, 79(2):146–152, 2009. doi: 10.1016/j.catena.2009.07.001.

R. Loritz, S. K. Hassler, C. Jackisch, N. Allroggen, L. van Schaik, J. Wienhöfer, and E. Zehe. Picturing and modeling catchments by representative hillslopes. *Hydrology and Earth System Sciences*, 21(2):1225, 2017. doi: 10.5194/hess-21-1225-2017.

F. Malard, D. Ferreira, S. Dolédec, and J. Ward. Influence of groundwater upwelling on the distribution of the hyporheos in a headwater river flood plain. *Archiv für Hydrobiologie*, 157(1):89–116, 2003. doi: 10.1127/0003-9136/2003/0157-0089.

I. Malcolm, C. Soulsby, A. Youngson, and D. Hannah. Catchment-scale controls on groundwater–surface water interactions in the hyporheic zone: implications for salmon embryo survival. *River Research and Applications*, 21(9):977–989, 2005. doi: 10.1002/rra.861.

K. J. Marfurt, R. L. Kirlin, S. L. Farmer, and M. S. Bahorich. 3-d seismic attributes using a semblance-based coherency algorithm. *Geophysics*, 63(4):1150–1165, 1998. doi: 10.1190/1.1444415.

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

Á. G. Mayor, S. Bautista, E. E. Small, M. Dixon, and J. Bellot. Measurement of the connectivity of runoff source areas as determined by vegetation pattern and topography: a tool for assessing potential water and soil losses in drylands. *Water Resources Research*, 44(10), 2008. doi: 10.1029/2007WR006367.

J. J. McDonnell. A rationale for old water discharge through macropores in a steep, humid catchment. *Water Resources Research*, 26(11):2821–2832, 1990. doi: {10.1029/WR026i011p02821}.

J. J. McDonnell, J. Freer, R. Hooper, C. Kendall, D. Burns, K. Beven, and J. Peters. New method developed for studying flow on hillslopes. *EOS, Transactions American Geophysical Union*, 77(47):465–472, 1996. doi: {10.1029/96EO00306}.

B. L. McGlynn and J. J. McDonnell. Quantifying the relative contributions of riparian and hillslope zones to catchment runoff. *Water Resources Research*, 39 (11), 2003. doi: {10.1029/2003WR002091}.

B. L. McGlynn, J. J. McDonnell, and D. D. Brammer. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. *Journal of Hydrology*, 257(1-4):1–26, 2002. doi: {10.1016/S0022-1694(01)00559-5}.

K. McGuire, J. J. McDonnell, M. Weiler, C. Kendall, B. McGlynn, J. Welker, and J. Seibert. The role of topography on catchment-scale water residence time. *Water Resources Research*, 41(5), 2005. doi: {10.1029/2004WR003657}.

K. J. McGuire and J. J. McDonnell. Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlinearities. *Water Resources Research*, 46(10), 2010. doi: 10.1029/2010WR009341.

Merriam-Webster. Online Dictionary. URL `https://www.merriam-webster.com/dictionary/connectivity`. Accessed: 2018-06-17.

K. Michaelides and A. Chappell. Connectivity as a concept for characterising hydrological behaviour. *Hydrological Processes*, 23(3):517–522, 2009. doi: {10.1002/hyp.7214}.

D. R. Montgomery. Process domains and the river continuum. *JAWRA Journal of the American Water Resources Association*, 35(2):397–410, 1999. doi: {10.1111/j.1752-1688.1999.tb03598.x}.

R. Moore. Slug injection using salt in solution. *Streamline Watershed Management Bulletin*, 8(2):1–6, 2005.

A. Mugler, B. Grinshpun, R. Franks, and C. H. Wiggins. Statistical method for revealing form-function relations in biological networks. *Proceedings of the National Academy of Sciences*, 108(2):446–451, 2011. doi: 10.1073/pnas.1008898108.

P. J. Mulholland, J. L. Tank, D. M. Sanzone, W. M. Wollheim, B. J. Peterson, J. R. Webster, and J. L. Meyer. Nitrogen cycling in a forest stream determined by a 15n tracer addition. *Ecological Monographs*, 70(3):471–493, 2000. doi: 10.1038/nature06686.

P. J. Mulholland, A. M. Helton, G. C. Poole, R. O. Hall, S. K. Hamilton, B. J. Peterson, J. L. Tank, L. R. Ashkenas, L. W. Cooper, C. N. Dahm, et al. Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature*, 452(7184):202–205, 2008. doi: 10.1038/nature06686.

N. Mullinger, A. Binley, J. Pates, and N. Crook. Radon in Chalk streams: Spatial and temporal variation of groundwater sources in the Pang and Lambourn catchments, UK. *Journal of Hydrology*, 339(3-4):172–182, 2007. doi: 10.1016/j.jhydrol.2007.03.010.

M. Munz, S. Krause, C. Tecklenburg, and A. Binley. Reducing monitoring gaps at the aquifer–river interface by modelling groundwater–surface water exchange flow patterns. *Hydrological Processes*, 25(23):3547–3562, 2011. doi: 10.1002/hyp.8080.

A. Musolff, C. Schmidt, B. Selle, and J. H. Fleckenstein. Catchment controls on solute export. *Advances in water resources*, 86:133–146, 2015. doi: {10.1016/j.advwatres.2015.09.026}.

N. Nadezhdina, T. S. David, J. S. David, M. I. Ferreira, M. Dohnal, M. Tesař, K. Gartner, E. Leitgeb, V. Nadezhdin, J. Cermak, et al. Trees never rest: the multiple facets of hydraulic redistribution. *Ecohydrology*, 3(4):431–444, 2010. doi: 10.1002/eco.148.

C. Neal and P. T. Rosier. Chemical studies of chloride and stable oxygen isotopes in two conifer afforested and moorland sites in the british uplands. *Journal of Hydrology*, 115(1-4):269–283, 1990. doi: 10.1016/0022-1694(90)90209-G.

J. R. Nimmo. Preferential flow occurs in unsaturated conditions. *Hydrological Processes*, 26(5):786–789, 2012. doi: 10.1002/hyp.8380.

F. Nippgen, B. L. McGlynn, and R. E. Emanuel. The spatial and temporal evolution of contributing areas. *Water Resources Research*, 51(6):4550–4573, 2015. doi: {10.1002/2014WR016719}.

P. Öhrström, Y. Hamed, M. Persson, and R. Berndtsson. Characterizing unsaturated solute transport by simultaneous use of dye and bromide. *Journal of Hydrology*, 289(1-4):23–35, 2004. doi: 10.1016/j.jhydrol.2003.10.014.

Y. Onda, Y. Komatsu, M. Tsujimura, and J. I. Fujihara. The role of subsurface runoff through bedrock on storm flow generation. *Hydrological Processes*, 15(10): 1693–1706, 2001. doi: 10.1002/hyp.234.

A. I. Packman, M. Salehin, and M. Zaramella. Hyporheic exchange with gravel beds: basic hydrodynamic interactions and bedform-induced advective flows. *Journal of Hydraulic Engineering*, 130(7):647–656, 2004.

J. Palm, N. L. M. van Schaik, and B. Schröder. Modelling distribution patterns of anecic, epigeic and endogeic earthworms at catchment-scale in agro-ecosystems. *Pedobiologia*, 56(1):23–31, 2013. doi: 10.1016/j.pedobi.2012.08.007.

M. A. Palmer. Experimentation in the hyporheic zone: challenges and prospectus. *Journal of the North American Benthological Society*, pages 84–93, 1993.

R. Payn, M. Gooseff, B. McGlynn, K. Bencala, and S. Wondzell. Channel water balance and exchange with subsurface flow along a mountain headwater stream in Montana, United States. *Water Resources Research*, 45(11), 2009. doi: {10.1029/2008WR007644}.

R. Payn, M. Gooseff, B. McGlynn, K. Bencala, and S. Wondzell. Exploring changes in the spatial distribution of stream baseflow generation during a seasonal recession. *Water Resources Research*, 48(4), 2012. doi: {10.1029/2011WR011552}.

A. J. Pearce, M. K. Stewart, and M. G. Sklash. Storm runoff generation in humid headwater catchments: 1. Where does the water come from? *Water Resources Research*, 22(8):1263–1272, 1986. doi: {10.1029/WR022i008p01263}.

N. E. Peters, J. Freer, and B. T. Aulenbach. Hydrological dynamics of the panola mountain research watershed, georgia. *Groundwater*, 41(7):973–988, 2003. doi: {10.1111/j.1745-6584.2003.tb02439.x}.

L. Pfister and L. Hoffmann. Experimental hydroclimatological atlas of the alzette river basin-grandduchy of luxembourg. *Centre de Recherche Rublic-Gabriel Lippmann, Cellule de Recherche en Environnement et Biotechnologies, Luxembourg*, 2002.

G. Pinay, T. C. O'Keefe, R. T. Edwards, and R. J. Naiman. Nitrate removal in the hyporheic zone of a salmon river in alaska. *River Research and Applications*, 25 (4):367–375, 2009. doi: 10.1002/rra.1164.

G. C. Poole. Stream hydrogeomorphology as a physical science basis for advances in stream ecology. *Journal of the North American Benthological Society*, 29(1): 12–25, 2010. doi: {10.1899/08-070.1}.

D. Porcelli and P. W. Swarzenski. The behavior of u-and th-series nuclides in groundwater. *Reviews in Mineralogy and Geochemistry*, 52(1):317–361, 2003. doi: {10.2113/0520317}.

C. M. Pringle. Hydrologic connectivity and the management of biological reserves: a global perspective. *Ecological Applications*, 11(4):981–998, 2001. doi: {10.1890/1051-0761(2001)011[0981:HCATMO]2.0.CO;2}.

C. M. Pringle. What is hydrologic connectivity and why is it ecologically important? *Hydrological Processes*, 17(13):2685–2689, 2003. doi: {10.1002/hyp.5145}.

W. J. Rawls, D. L. Brakensiek, and K. Saxtonn. Estimation of soil water properties. *Transactions of the ASAE*, 25(5):1316–1320, 1982. doi: 10.13031/2013.33720.

P. A. Raymond and J. J. Cole. Gas exchange in rivers and estuaries: Choosing a gas transfer velocity. *Estuaries*, 24(2):312–317, 2001.

P. Reggiani, M. Sivapalan, and S. M. Hassanizadeh. A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Advances in Water Resources*, 22(4):367–398, 1998. doi: 10.1016/S0309-1708(98)00012-8.

L. A. Richards. Capillary conduction of liquids through porous mediums. *physics*, 1(5):318–333, 1931.

D. O. Rosenberry and J. Pitlick. Local-scale variability of seepage and hydraulic conductivity in a shallow gravel-bed river. *Hydrological Processes*, 23(23):3306–3318, 2009. doi: 10.1002/hyp.7433.

K. Roth, H. Vogel, and R. Kasteel. The scaleway: a conceptual framework for upscaling soil properties. *Modelling of Transport Processes in Soils at Various Scales in Time and Space*, pages 24–26, 1999.

T. Sander and H. H. Gerke. Modelling field-data of preferential flow in paddy soil induced by earthworm burrows. *Journal of contaminant hydrology*, 104(1-4):126–136, 2009. doi: 10.1016/j.jconhyd.2008.11.003.

H. Savenije. Hess opinions "topography driven conceptual modelling (flex-topo)". *Hydrology and Earth System Sciences*, 14(12):2681–2692, 2010. doi: {10.5194/hess-14-2681-2010,2010.}.

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

M. G. Schaap, F. J. Leij, and M. T. Van Genuchten. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of hydrology*, 251(3-4):163–176, 2001. doi: 10.1016/S0022-1694(01)00466-8.

S. Schlüter, S. Berg, M. Rücker, R. Armstrong, H.-J. Vogel, R. Hilfer, and D. Wildenschild. Pore-scale displacement mechanisms as a source of hysteresis for two-phase flow in porous media. *Water Resources Research*, 52(3):2194–2205, 2016. doi: {10.1002/2015WR018254}.

C. Schmelzbach, J. Tronicke, and P. Dietrich. Three-dimensional hydrostratigraphic models from ground-penetrating radar and direct-push data. *Journal of Hydrology*, 398(3):235–245, 2011. doi: 10.1016/j.jhydrol.2010.12.023.

C. Schmelzbach, J. Tronicke, and P. Dietrich. High-resolution water content estimation from surface-based ground-penetrating radar reflection data by impedance inversion. *Water Resources Research*, 48(8), 2012. doi: 10.1029/2012WR011955.

A. Schmidt, J. Gibson, I. R. Santos, M. Schubert, K. Tattrie, and H. Weiss. The contribution of groundwater discharge to the overall water budget of two typical Boreal lakes in Alberta/Canada estimated from a radon mass balance. *Hydrology and Earth System Sciences*, 14(1):79, 2010.

C. Schornberg, C. Schmidt, E. Kalbus, and J. H. Fleckenstein. Simulating the effects of geologic heterogeneity and transient boundary conditions on streambed temperatures–implications for temperature-based water flux calculations. *Advances in Water Resources*, 33(11):1309–1319, 2010. doi: 10.1016/j.advwatres.2010.04.007.

D. Schotanus, M. van der Ploeg, and S. van der Zee. Quantifying heterogeneous transport of a tracer and a degradable contaminant in the field, with snowmelt and irrigation. *Hydrology and Earth System Sciences*, 16(8):2871, 2012. doi: 10.5194/hess-16-2871-2012.

K. Schulz, R. Seppelt, E. Zehe, H. Vogel, and S. Attinger. Importance of spatial structures in advancing hydrological sciences. *Water Resources Research*, 42(3), 2006. doi: {10.1029/2005WR004301}.

J. Schwoerbel. Über die lebensbedingungen und die besiedlung des hyporheischen lebensraumes. *Archiv für Hydrobiologie, Supplement*, 25:182–214, 1961.

J. Seibert and B. L. McGlynn. A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models. *Water resources research*, 43(4), 2007. doi: {10.1029/2006WR005128}.

J. S. Selker, L. Thévenaz, H. Huwald, A. Mallet, W. Luxemburg, N. Van De Giesen, M. Stejskal, J. Zeman, M. Westhoff, and M. B. Parlange. Distributed fiber-optic temperature sensing for hydrologic systems. *Water Resources Research*, 42(12), 2006. doi: {10.1029/2006WR005326}.

R. C. Sidle. Field observations and process understanding in hydrology: essential components in scaling. *Hydrological Processes*, 20(6):1439–1445, 2006. doi: 10.1002/hyp.6191.

R. C. Sidle, S. Noguchi, Y. Tsuboyama, and K. Laursen. A conceptual model of preferential flow systems in forested hillslopes: Evidence of self-organization. *Hydrological Processes*, 15(10):1675–1692, 2001. doi: {10.1002/hyp.233}.

H. Simon. The architecture of complexity. *Proceedings of the American Philosophical Society*, 106:467–482, 1962.

M. Sivapalan. Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. *Encyclopedia of hydrological sciences*, 2005. doi: {10.1002/0470848944.hsa012}.

M. Snehota, V. Jelinkova, J. Sacha, M. Frycova, M. Cislerova, P. Vontobel, and J. Hovind. Experimental Investigation of Preferential Flow in a Near-saturated Intact Soil Sample. *Physics Procedia*, 69:496–502, 2015. doi: 10.1016/j.phpro.2015.07.070.

M. Sophocleous. Interactions between groundwater and surface water: the state of the science. *Hydrogeology journal*, 10(1):52–67, 2002. doi: {10.1007/s10040-001-0170-8}.

M. Sprenger, B. Herbstritt, and M. Weiler. Established methods and new opportunities for pore water stable isotope analysis. *Hydrological Processes*, 29(25):5174–5192, 2015a. doi: 10.1002/hyp.10643.

M. Sprenger, T. H. Volkmann, T. Blume, and M. Weiler. Estimating flow and transport parameters in the unsaturated zone with pore water stable isotopes. *Hydrology and Earth System Sciences*, 19(6):2617, 2015b. doi: 10.5194/hess-19-2617-2015.

M. Sprenger, S. Seeger, T. Blume, and M. Weiler. Travel times in the vadose zone: Variability in space and time. *Water Resources Research*, 52(8):5727–5754, 2016. doi: {10.1002/2015WR018077}.

C. M. Steelman, A. L. Endres, and J. P. Jones. High-resolution ground-penetrating radar monitoring of soil moisture dynamics: Field results, interpretation, and comparison with unsaturated flow model. *Water Resources Research*, 48(9):1–17, 2012. doi: 10.1029/2011WR011414.

L. Stellato, E. Petrella, F. Terrasi, P. Belloni, M. Belli, U. Sansone, and F. Celico. Some limitations in using 222 rn to assess river–groundwater interactions: the case of castel di sangro alluvial plain (central italy). *Hydrogeology Journal*, 16(4): 701–712, 2008. doi: {10.1007/s10040-007-0263-0}.

L. Stellato, F. Terrasi, F. Marzaioli, M. Belli, U. Sansone, and F. Celico. Is 222Rn a suitable tracer of stream–groundwater interactions? A case study in central Italy. *Applied geochemistry*, 32:108–117, 2013. doi: {10.1016/j.apgeochem.2012.08.022}.

M. Stieglitz, J. Shaman, J. McNamara, V. Engel, J. Shanley, and G. W. Kling. An approach to understanding hydrologic connectivity on the hillslope and the implications for nutrient transport. *Global Biogeochemical Cycles*, 17(4), 2003. doi: {10.1029/2003GB002041}.

R. G. Storey, K. W. Howard, and D. D. Williams. Factors controlling riffle-scale hyporheic exchange flows and their seasonal changes in a gaining stream: A three-dimensional groundwater flow model. *Water Resources Research*, 39(2), 2003. doi: 10.1029/2002WR001367.

R. G. Storey, D. D. Williams, and R. R. Fulthorpe. Nitrogen processing in the hyporheic zone of a pastoral stream. *Biogeochemistry*, 69(3):285–313, 2004. doi: 10.1023/B:BIOG.0000031049.95805.ec.

H. Streeter and E. Phelps. A study of the pollution and natural purification of the ohio river. iii. factors concerned in the phenomena of oxidation and reaeration. *Public Health Bulletin*, 146, 1925.

M. Streetly and M. Shepley. East shropshire permo-triassic sandstone groundwater modelling project, task 2: Literature review. *Environment Agency Technical Report CR/02*, 176, 2002.

L. H. Sullivan. The tall office building artistically considered. *Lippincott's Magazine*, 57(3):406, 1896.

P. Szeftel, R. D. Moore, and M. Weiler. Influence of distributed flow losses and gains on the estimation of transient storage parameters from stream tracer experiments. *Journal of hydrology*, 396(3-4):277–291, 2011. doi: {10.1016/j.jhydrol.2010.11.018}.

P. D. Taylor, L. Fahrig, K. Henein, and G. Merriam. Connectivity is a vital element of landscape structure. *Oikos*, pages 571–573, 1993. doi: {10.2307/3544927}.

D. Tetzlaff, C. Soulsby, P. Bacon, A. Youngson, C. Gibbins, and I. Malcolm. Connectivity between landscapes and riverscapes – a unifying theme in integrating hydrology and ecology in catchment science? *Hydrological Processes*, 21(10): 1385–1389, 2007. doi: {10.1002/hyp.6701}.

D. Tetzlaff, S. Carey, H. Laudon, and K. McGuire. Catchment processes and heterogeneity at multiple scales—benchmarking observations, conceptualization and prediction. *Hydrological Processes*, 24(16):2203–2208, 2010. doi: 10.1002/hyp.7784.

L. J. Thibodeaux and J. D. Boyle. Bedform-generated convective transport in bottom sediment. *Nature*, 325(6102):341, 1987.

D. W. Thompson. *On growth and form*. Cambridge Univ. Press, 1942.

F. Tian, H. Hu, Z. Lei, and M. Sivapalan. Extension of the representative elementary watershed approach for cold regions via explicit treatment of energy related processes. *Hydrology and Earth System Sciences Discussions*, 10(5):619–644, 2006. doi: 10.5194/hess-10-619-2006.

D. Tonina and J. M. Buffington. Hyporheic exchange in gravel bed rivers with pool-riffle morphology: Laboratory experiments and three-dimensional modeling. *Water Resources Research*, 43(1), 2007. doi: 10.1029/2005WR004328.

J. Tóth. A theoretical analysis of groundwater flow in small drainage basins. *Journal of geophysical research*, 68(16):4795–4812, 1963.

I. Trinks, H. Stümpel, and D. Wachsmuth. Monitoring water flow in the unsaturated zone using georadar. *First break*, 19(12):679–684, 2001. doi: 10.1046/j.1365-2397.2001.00228.x.

H. Tromp-van Meerveld and J. McDonnell. Threshold relations in subsurface stormflow: 1. a 147-storm analysis of the panola hillslope. *Water Resources Research*, 42(2), 2006. doi: {10.1029/2004WR003778}.

H. J. Tromp-Van Meerveld and J. J. McDonnell. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resources Research*, 42(2):1–11, Feb. 2006. doi: 10.1029/2004WR003800.

S. Truss, M. Grasmueck, S. Vega, and D. A. Viggiano. Imaging rainfall drainage within the Miami oolitic limestone using high-resolution time-lapse ground-penetrating radar. *Water Resources Research*, 43(3):W03405, 2007. doi: 10.1029/2005WR004395.

E. Tsivoglou and L. Neal. Tracer measurement of reaeration: Iii. predicting the reaeration capacity of inland streams. *Journal (Water Pollution Control Federation)*, pages 2669–2689, 1976.

S. W. Tyler and S. W. Wheatcraft. Fractal processes in soil water retention. *Water Resources Research*, 26(5):1047–1054, 1990. doi: {10.1029/WR026i005p01047}.

T. Uchida, I. Tromp-van Meerveld, and J. J. McDonnell. The role of lateral pipe flow in hillslope runoff response: an intercomparison of non-linear hillslope response. *Journal of Hydrology*, 311(1-4):117–133, 2005. doi: {10.1016/j.jhydrol.2005.01.012}.

S. Uhlenbrook. Catchment hydrology – a science in which all processes are preferential. *Hydrological Processes*, 20:3581–3585, 2006. doi: 10.1002/hyp.6564.

N. Unland, I. Cartwright, G. Rau, J. Reed, B. Gilfedder, A. Atkinson, H. Hofmann, et al. Investigating the spatio-temporal variability in groundwater and surface water interactions: a multi-technique approach. *Hydrology and Earth System Sciences*, 17(9):3437, 2013. doi: {10.5194/hess-17-3437-2013}.

H. M. Valett, J. A. Morrice, C. N. Dahm, and M. E. Campana. Parent lithology, surface–groundwater exchange, and nitrate retention in headwater streams. *Limnology and oceanography*, 41(2):333–345, 1996. doi: {10.4319/lo.1996.41.2.0333}.

R. Van den Bos, L. Hoffmann, J. Juilleret, P. Matgen, and L. Pfister. Regional runoff prediction through aggregation of first-order hydrological process knowledge: a case study. *Hydrological Sciences Journal*, 51(6):1021–1038, Dec. 2006. doi: 10.1623/hysj.51.6.1021.

L. Van Schaik, J. Palm, J. Klaus, E. Zehe, and B. Schröder. Linking spatial earthworm distribution to macropore numbers and hydrological effectiveness. *Ecohydrology*, 7(2):401–408, 2014. doi: 10.1002/eco.1358.

N. Van Schaik. Spatial variability of infiltration patterns related to site characteristics in a semi-arid watershed. *Catena*, 78(1):36–47, 2009. doi: 10.1016/j.catena.2009.02.017.

H.-J. Vogel and K. Roth. Quantitative morphology and network representation of soil pore structure. *Advances in water resources*, 24(3-4):233–242, 2001. doi: {10.1016/S0309-1708(00)00055-5}.

H.-J. Vogel and K. Roth. Moving through scales of flow and transport in soil. *Journal of Hydrology*, 272(1):95–106, 2003. doi: {10.1016/S0022-1694(02)00257-3}.

H.-J. Vogel, I. Cousin, O. Ippisch, and P. Bastian. The dominant role of structure for solute transport in soil: Experimental evidence and modelling of structure and transport in a field experiment. *Hydrology and Earth System Sciences Discussions*, 2(5):2153–2181, 2005. doi: 10.5194/hess-10-495-2006.

T. Vogel, M. Sanda, J. Dusek, M. Dohnal, and J. Votrubova. Using Oxygen-18 to Study the Role of Preferential Flow in the Formation of Hillslope Runoff. *Vadose Zone Journal*, 9(2):252–259, 2010. doi: 10.2136/vzj2009.0066.

J. Šimůnek, N. J. Jarvis, M. T. van Genuchten, and A. Gardenas. Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone. *Journal of Hydrology*, 272:14–35, 2003. doi: 10.1016/S0022-1694(02)00252-4.

J. Wainwright, L. Turnbull, T. G. Ibrahim, I. Lexartza-Artza, S. F. Thornton, and R. E. Brazier. Linking environmental regimes, space and time: Interpretations of structural and functional connectivity. *Geomorphology*, 126(3-4):387–404, 2011. doi: {10.1016/j.geomorph.2010.07.027}.

Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. doi: 10.1109/TIP.2003.819861.

A. S. Ward, M. Fitzgerald, M. N. Gooseff, T. J. Voltz, A. M. Binley, and K. Singha. Hydrologic and geomorphic controls on hyporheic exchange during base flow recession in a headwater mountain stream. *Water Resources Research*, 48(4), 2012. doi: {10.1029/2011WR011461}.

J. V. Ward. An expansive perspective of riverine landscapes: pattern and process across scales. *GAIA-Ecological Perspectives for Science and Society*, 6(1):52–60, 1997. doi: {10.14512/gaia.6.1.6}.

L. Wassenaar, M. Hendry, V. Chostner, and G. Lis. High resolution pore water δ^2H and $\delta^{18}O$ measurements by H_2O (liquid)- H_2O (vapor) equilibration laser spectroscopy. *Environmental science & technology*, 42(24):9262–9267, 2008. doi: 10.1021/es802065s.

M. Wehrer and L. D. Slater. Characterization of water content dynamics and tracer breakthrough by 3-d electrical resistivity tomography (ert) under transient unsaturated conditions. *Water Resources Research*, 51(1):97–124, 2015. doi: 10.1002/2014WR016131.

F. Weigel. Radon. *Chemiker Zeitung*, 102(9):287–299, 1978.

M. Weiler and J. McDonnell. Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes. *Water Resources Research*, 43(3):1–13, 2007. doi: {10.1029/2006WR004867}.

M. Weiler and J. J. McDonnell. Testing nutrient flushing hypotheses at the hillslope scale: A virtual experiment approach. *Journal of Hydrology*, 319(1-4):339–356, 2006. doi: {10.1016/j.jhydrol.2005.06.040}.

J. Wenninger, S. Uhlenbrook, S. Lorentz, and C. Leiundgut. Identification of runoff generation processes using combined hydrometric, tracer and geophysical methods in a headwater catchment in South Africa. *Hydrological Sciences Journal*, 53(1): 65–80, 2008. doi: {10.1623/hysj.53.1.65}.

A. W. Western and G. Blöschl. On the spatial scaling of soil moisture. *Journal of hydrology*, 217(3-4):203–224, 1999. doi: 10.1016/S0022-1694(98)00232-7.

A. W. Western, R. B. Grayson, G. Blöschl, G. R. Willgoose, and T. A. McMahon. Observed spatial organization of soil moisture and its relation to terrain indices. *Water resources research*, 35(3):797–810, 1999. doi: {10.1029/1998WR900065}.

A. W. Western, G. Blöschl, and R. B. Grayson. Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, 37(1):83–97, 2001. doi: {10.1029/2000WR900241}.

H. S. Wheater and D. Peach. Developing interdisciplinary science for integrated catchment management: the UK lowland catchment research (LOCAR) programme. *International Journal of Water Resources Development*, 20(3):369–385, 2004. doi: {10.1080/0790062042000248565}.

D. S. White. Perspectives on defining and delineating hyporheic zones. *Journal of the North American Benthological Society*, pages 61–69, 1993.

J. Wienhöfer and E. Zehe. Predicting subsurface stormflow response of a forested hillslope–the role of connected flow paths. *Hydrology and Earth System Sciences*, 18(1):121–138, 2014. doi: {10.5194/hess-18-121-2014}.

J. Wienhofer, K. Germer, F. Lindenmaier, A. Färber, and E. Zehe. Applied tracers for the observation of subsurface stormflow at the hillslope scale. *Hydrology and Earth System Sciences*, 13(7):2961–3006, 2009. doi: 10.5194/hess-13-1145-2009.

L. Wittgenstein. *Tractatus logico-philosophicus*. Kegan Paul, Trench, Trubner Co., Ltd., 1922.

W. W. Woessner. Stream and fluvial plain ground water interactions: rescaling hydrogeologic thought. *Groundwater*, 38(3):423–429, 2000. doi: {10.1111/j.1745-6584.2000.tb00228.x}.

S. M. Wondzell. Effect of morphology and discharge on hyporheic exchange flows in two small streams in the Cascade Mountains of Oregon, USA. *Hydrological Processes*, pages 267–287, 2006. doi: {10.1002/hyp.5902}.

S. M. Wondzell, J. LaNier, and R. Haggerty. Evaluation of alternative groundwater flow models for simulating hyporheic exchange in a small mountain stream. *Journal of Hydrology*, 364(1):142–151, 2009a. doi: {10.1016/j.jhydrol.2008.10.011}.

S. M. Wondzell, J. LaNier, R. Haggerty, R. D. Woodsmith, and R. T. Edwards. Changes in hyporheic exchange flow following experimental wood removal in a small, low-gradient stream. *Water Resources Research*, 45(5), 2009b. doi: {10.1029/2008WR007214}.

S. Wrede, F. Fenicia, N. Martínez-Carreras, J. Juilleret, C. Hissler, A. Krein, H. H. G. Savenije, S. Uhlenbrook, D. Kavetski, and L. Pfister. Towards more systematic perceptual model development: a case study using 3 Luxembourgish catchments. *Hydrological Processes*, 29:2731–2750, 2015. doi: 10.1002/hyp.10393.

K. K. Wright, C. V. Baxter, and J. L. Li. Restricted hyporheic exchange in an alluvial river system: implications for theory and management. *Journal of the North American Benthological Society*, 24(3):447–460, 2005.

J. Wu. Hierarchy and scaling: extrapolating information along a scaling ladder. *Canadian journal of remote sensing*, 25(4):367–380, 1999. doi: {10.1080/07038992.1999.10874736}.

J. P. Zarnetske, R. Haggerty, S. M. Wondzell, and M. A. Baker. Dynamics of nitrate production and removal as a function of residence time in the hyporheic zone. *Journal of Geophysical Research: Biogeosciences*, 116(G1), 2011. doi: 10.1029/2010JG001356.

E. Zehe, T. Maurer, J. Ihringer, and E. Plate. Modeling water flow and mass transport in a loess catchment. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(7-8):487–507, 2001. doi: 10.1016/S1464-1909(01)00041-7.

E. Zehe, T. Blume, and G. Blöschl. The principle of 'maximum energy dissipation': a novel thermodynamic perspective on rapid water flow in connected soil structures. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1545):1377–1386, 2010. doi: {10.1098/rstb.2009.0308}.

E. Zehe, T. Blume, A. Kleidon, U. Ehret, U. Scherer, and M. Westhoff. A thermodynamic approach to link self-organization, preferential flow and rainfall–runoff behaviour. *Hydrology and Earth System Sciences*, 17:4297–4322, 2013. doi: 10.5194/hess-17-4297-2013.

E. Zehe, U. Ehret, L. Pfister, T. Blume, B. Schröder, M. Westhoff, C. Jackisch, S. Schymanski, M. Weiler, K. Schulz, N. Allroggen, J. Tronicke, P. Dietrich, U. Scherer, J. Eccard, V. Wulfmeyer, and A. Kleidon. HESS Opinions: From response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments. *Hydrology and Earth System Sciences*, 18(11):4635–4655, 2014a. doi: 10.5194/hess-18-4635-2014.

E. Zehe, U. Ehret, L. Pfister, T. Blume, B. Schröder, M. Westhoff, C. Jackisch, S. J. Schymanski, M. Weiler, K. Schulz, N. Allroggen, J. Tronicke, P. Dietrich, U. Scherer, J. Eccard, V. Wulfmeyer, and A. Kleidon. Functional units: a novel framework to explore the link between spatial organization and hydrological functioning of intermediate scale catchments. *Hydrology and Earth System Sciences Discussions*, 11(3):3249–3313, 2014b. doi: 10.5194/hessd-11-3249-2014.

P. Zhao, X. Tang, P. Zhao, C. Wang, and J. Tang. Tracing water flow from sloping farmland to streams using oxygen-18 isotope to study a small agricultural catchment in southwest China. *Soil and Tillage Research*, 134:180–194, 2013. doi: 10.1016/j.still.2013.08.005.

B. Zillgens, B. Merz, R. Kirnbauer, and N. Tilch. Analysis of the runoff response of an alpine catchment at different scales. *Hydrology and Earth System Sciences*, 11 (4):1441–1454, 2007.

Author's declaration

I prepared this dissertation without illegal assistance. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. This dissertation has not been presented to any other University for examination, neither in Germany nor in another country.

Berlin, June 2018 Lisa Angermann