# THE ROLE OF ENERGY AND INFORMATION IN HYDROLOGICAL MODELING

RALF LORITZ

# THE ROLE OF ENERGY AND INFORMATION IN HYDROLOGICAL MODELING

zum Erlangen des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)

von der KIT-Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften

des Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von
Ralf Loritz
aus Karlsruhe

Tag der mündlichen Prüfung:
12. Dezember 2019

REFERENT: Prof. Dr. Erwin Zehe
KORREFERENT: Dr. Markus Hrachowitz

Karlsruhe 2019

Dedicated to Nina and Emil,
my beloved little family
and
to my mother,
for the sleepless nights she had during my youth.

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABSTRACT

In the past decades, hydrology has evolved from an engineering discipline to a fully established earth system science build upon physical principles. This evolution is mirrored by the research questions hydrology is dealing with today which range from infiltration experiments at the pore scale to simulations of the water cycle at a global scale. While hydrology made great scientific progress in the last decades, many hydrological models, especially outside the realm of science, remain largely empirical and founded on strong physical simplifications (e.g. unit hydrograph). If sufficiently long observation time series are available and if stationary conditions can be assumed, using such models for practical reasons seems fair. This changes, however, as soon as the focus shifts to predictions under change or to ungauged basins. Here are models which represent the physics of a hydrological system in a more exhaustive manner a more promising choice. While this is rarely questioned in hydrology there is currently neither consensus on how an appropriate physical system description in hydrology might look like nor which details a hydrological model needs to represent to be called "physically-based".

James Dooge observed the lack of a commonly accepted theoretical foundation in catchment hydrology already in his seminal paper in 1986. He argued that in order to minimize the necessity of model calibration and empiricism to improve its ability to make predictions, hydrology needs to identify and derive a comprehensive theoretical framework on which its models and its corresponding hypothesis are based. My thesis is motivated by exactly this search for an improved theoretical underpinning of hydrology with an emphasis on hydrological models for meso-scale catchments. The goal of this thesis is thereby not to propose a fundamentally new theoretical framework for hydrology but to acknowledge and advance the existing theoretical basis. I draw largely from two related scientific theories, information theory and thermodynamics, and use the concepts of information and energy to shed new light on well-established hydrological research questions.

In chapter 2, I develop a model concept called "representative hillslope" which was designed to represent meso-scale catchments in an effective top-down manner, however with a 2d bottom-up hillslope model. To test the concept, two representative hillslope models are set up based on an extensive environmental database in two catchments that differ distinctively with respect to their dominant runoff

processes. The models are based on macroscopic model parameters which were derived from a mosaic of point measurements as well as from qualitative information, for instance about the occurrence and relevance of macropores or subsurface water ponds. Both models are optimized within a few trial and error runs on streamflow and tested carefully against soil moisture measurements and sap flow velocities. The results indicate that representative hillslopes are a promising way to represent meso-scale catchments with bottom-up models, but they also point out several limitations of the approach itself and the chosen model. Nevertheless, the study shows that representative hillslopes are a logical starting point for any distributed modeling task and a first attempt to bridge the gap between top-down and bottom-up models.

In chapter 3 I continue with one of the main objectives of this thesis to improve the way meso-scale catchments are represented with hydrological models. More specifically, I attempt to identify functional units that could serve as elementary spatial units for a distributed model or a hydrological measurement network. I approach this question by dividing a catchment into 105 hillslopes and representing each hillslope with a 2d bottom-up model. These 105 models differ only with respect to parameters that are derived from a digital elevation model (DEM) - the hillslope's form and aspect – to examine the specific influence that surface topography has on the function of each sub-unit. The 105 models are run with the same climatic forcing for one hydrological year and the discharge as well as storage simulations are analyzed by means of the Shannon entropy. The results show that the simulations are highly redundant and that a compressed catchment model that consists of six instead of 105 hillslopes was able to produce similar discharge simulations as the entire ensemble. However, the compressibility (redundancy of the simulations) of the model ensemble and hence the "optimal" number of hillslopes varied strongly in time and depended on the target variable (streamflow or water storage). This highlights that hydrological similarity is controlled not only by the function but also by the state of a given landscape element. An "optimal" spatial division of a catchment into hydrologically similar units needs hence to be specific to a given function and to a given state. The latter requires a flexible approach which can adapt itself in time.

In chapter 4 I identify hydrological similar landscape units based on the available geo-data without the detour and uncertainties using a distributed model entails. The goal of chapter 4 is hence to develop similarity index which is able to identify functional units directly on the available topographic information. The resulting index rDUNE (reduced dissipation per unit length) is an energy-centered reinter-

pretation and enhancement of the well-established index HAND (height above the nearest drainage). It is fundamentally based on the observation that the majority of the incoming potential energy of precipitation is dissipated when rainfall becomes runoff. rDUNE is tested against catchment-wide distributions of HAND and of the topographic wetness index (TWI). The results show that rDUNE is indeed superior to the other indices if the goal is to identify catchments with similar dominant runoff processes based on their topography. Our analysis indicates that accounting for both, the driver and resistance term of flux generation law, provides a promising approach to develop similarity indices in hydrology. The study ends with the questions why the majority of the potential energy is dissipated at the hillslope scale although it is frequently reported that hydrological structures evolve in a way that energy dissipation is minimized.

# ZUSAMMENFASSUNG

In den letzten Jahrzehnten hat sich die Hydrologie von einer Ingenieurdisziplin zu einer etablierten Erdsystemwissenschaft entwickelt, die auf physikalischen Prinzipien beruht. Diese Entwicklung spiegelt sich auch in den aktuellen, hydrologischen Fragestellungen wider die von Infiltrationsexperimenten auf der Porenskala bis hin zu Simulationen des weltweiten Wasserkreislaufs reichen. Während die Hydrologie in den letzten Jahren große wissenschaftliche Fortschritte gemacht hat und immer vielfältiger wurde, sind viele, besonders in der Praxis angewendeten hydrologische Modelle (z. B. die Einheitsganglinie), größtenteils noch immer empirische Konzepte, die auf starken physikalischen Vereinfachungen beruhen. Solche Ansätze eignen sich für Fragestellungen bei denen ausreichend lange Beobachtungszeitreihen vorliegen und stationäre Bedingungen angenommen werden können. Sie sind jedoch tendenziell ungeeignet, sobald der Fokus auf instationären Bedingungen oder auf Gebiete mit wenigen oder schlechten Daten gerichtet wird. In solchen Fällen sind stärker physikalisch-basierte Modelle eine vielversprechende Alternative. Während letzteres in der Hydrologie kaum angezweifelt wird, gibt es bis heute jedoch keinen Konsens darüber, wie genau ein adäquat physikalisch-basiertes Model aussieht und welche Details und physikalische Gesetze in ihm abgebildet werden müssen.

Diese Uneinigkeit kann zumindest teilweise darauf zurückgeführt werden, dass es in der Hydrologie bis heute keine allgemein akzeptierten theoretischen Grundlage gibt. James Dooge erkannte das bereits im Jahr 1986 und argumentierte, dass die hydrologische Gemeinschaft zusammen einen klaren theoretischen Rahmen ableiten müsse, auf dem ihre Modelle sowie die entsprechenden Hypothesen beruhen. Diese gemeinsame theoretische Grundlage würde dann die Notwendigkeit von Modellkalibrierungen und die damit verbundene Empirie in hydrologischen Modellen minimieren und schlussendlich zu belastbareren Vorhersagen führen.

Diese Dissertation knüpft an diesen Punkt an und beschäftigt sich mit der Verbesserung der theoretischen Grundlage hydrologischer Modelle, mit Schwerpunkt auf mesoskalige Einzugsgebiete. Ziel dieser Arbeit ist es dabei nicht, eine umfassende, neue theoretischen Grundlage einzuführen, sondern die bestehenden theoretischen Grundlagen a) anzuerkennen und b) weiter zu entwickeln. Dazu nutze ich weitgehend zwei verwandte wissenschaftliche Theorien, die Informationstheorie und die Thermodynamik und verwende

damit die Konzepte von Information und Energie, um Forschung zu betreiben sowie neue Blickwickel auf etablierte Fragestellungen zu erhalten.

In Kapitel 2 entwickle ich das Modellkonzept "Representative Hillslopes", welches eine Kombination aus top-down und bottom-up Model darstellt. Um das Konzept zu testen, werden zwei repräsentative Hangmodelle in zwei meso-skaligen Einzugsgebieten auf grundlage einer umfrangreichen Umweltdatenbank aufgesetzt. Die Modelparameter stützen sich dabei überwiegend auf makroskopische Gebietseigenschaften, die aus einem Mosaik von Punktmessungen abgeleitet wurden sowie auf qualitativen Informationen, wie beispielsweise Informationen über das Vorkommen und die Relevanz von Makroporen oder von unterirdischen Wasserspeichern. Beide Modelle wurden mittels weniger Testläufe anhand des Abflusses kalibriert und sorgfältig in Bezug auf Bodenfeuchte- (storage) und Saftflussmessungen (transpiration) getestet. Die Ergebnisse zeigen zum einen, dass repräsentative Hänge ein vielversprechender Weg sind, mesoskalige Einzugsgebiet mit bottom-up Modellen abzubilden, aber auch, dass Limitationen im entwickelten Ansatz und gewählten Modell, wie die verwendeten Fließgesetze in den Drainagestrukturen, vorliegen. Schlussendlich zeigt sich aber, dass repräsentative Hänge ein logischer Startpunkt sind um a) räumlich verteilte bottom-up Modelle aufzusetzen und b) ein erster Versuch sein könnten, um die Kluft zwischen top-down und bottom-up Modellen zu schließen.

Im Kapitel 3 fahre ich mit einem der Hauptziele dieser Arbeit fort, Simulationen und die generelle Abbildung mesoskaliger Einzugsgebiete mittels hydrologischer Modelle zu verbessern. Dafür versuche ich funktionale Einheiten zu identifizieren, die als elementare räumliche Basis für verteilte Modelle oder ein hydrologisches Messnetzwerk dienen könnten. Dazu teile ich ein Einzugsgebiet in 105 Hänge ein und repräsentiere jeden Hang mit einem 2d bottom-up Modell. Diese 105 Hangmodelle unterscheiden sich dabei ausschließlich durch Parameter, die aus einem digitalen Höhenmodell (DEM) abgeleitet wurden - die Hangform und dessen geographische Ausrichtung - um den spezifischen Einfluss der Oberflächentopographie auf die Funktion der einzelnen Untereinheiten (Hänge) zu untersuchen. Die 105 Modelle werden mit den gleichen Klimadaten für ein hydrologisches Jahr simuliert und die Abfluss- sowie Wasserspeichersimulationen mit Hilfe der Shannon-Entropie analysiert. Die Ergebnisse zeigen, dass die Simulationen stark redundant sind und dass ein komprimiertes Einzugsgebietsmodell, das aus sechs anstelle von 105 Hängen besteht, ähnliche Simulationen wie das gesamte Ensemble erzeugen kann. Die Kompressibilität (Redundanz der Simulationen) des Ensembles und damit die "optimale"Anzahl der Hänge variierte dabei jedoch zeitlich sowie je nach Zielvariable (Abfluss oder Speicherung) stark. Dies zeigt, dass hydrologische Ähnlichkeit von der Funktion und vom Zustand eines Landschaftselements abhängt. Eine ”optimale”räumliche Unterteilung eines Einzugsgebiets in hydrologisch ähnliche Einheiten muss daher für eine bestimmte Funktion aber auch für unterschiedliche Zustände in der Zeit angepasst werden.

In Kapitel 4 versuche ich hydrologisch ähnliche Landschaftseinheiten auf der Grundlage der verfügbaren Geodaten zu identifizieren, ohne die Umwege und Unsicherheiten, die mit der Verwendung eines Modelles einhergehen. Ziel von Kapitels 4 ist es, einen Ähnlichkeitsindex zu entwickeln, mit dem räumliche Einheiten direkt anhand der verfügbaren topografischen Informationen identifiziert werden können. Der entwickelte Index (reduced dissipation per unit length; rDUNE) ist dabei eine energiezentrierte Neuinterpretation des etablierten Index HAND (height above the nearest drainage) und basiert im Wesentlichen auf der Beobachtung, dass der Großteil der potentiellen Energie des Niederschlags dissipiert wird, wenn Niederschlag zu Abfluss wird. rDUNE wird gegen einzugsgebietsweite Verteilungen von HAND und vom topographic wetness index (TWI) getestet. Die Ergebnisse zeigen, dass rDUNE den anderen getesteten Indizes in der Tat überlegen ist, wenn das Ziel darin besteht, Einzugsgebiete mit ähnlich dominanten Abflussprozessen anhand ihrer Topographie zu identifizieren. Die Analyse zeigt auch, dass die Einteilung von Landschaftsfaktoren in Treiber und Widerstandsterme einen vielversprechenden Ansatz darstellt, um Ähnlichkeitsindizes in der Hydrologie zu entwickeln. Die Studie endet mit der Frage, warum sich Landschaften so entwickeln, dass der größte Teil der potenziellen Energie auf der Hangskala dissipiert wird, obwohl schon oft gezeigt wurde, dass Energiedissipation innerhalb von Flussnetzen minimiert wird.

Part I

# INTRODUCTION

# 1 INTRODUCTION

## 1.1 MOTIVATION

The increased probability of extreme precipitation events and consequently the higher chance of flash floods and erosion of fertile lands, the depletion of available water stocks by irrigation as well as water quality issues related to the excessive use of pesticides are only a few challenges hydrology is facing in the $21^{st}$ Century. What these challenges have in common is that they are all caused or at least highly influenced by human activities. This implies that the impact they have on society can be reduced if the socio-hydrological management of available water resources is improved (e.g. Montanari et al., 2013). In this context it is interesting to note that many management decisions in hydrology are designed and outlined neither at the plot scale of several square meters nor on the scale of entire large scale watersheds ($> 250\,\mathrm{km}^2$) but on the scale of meso-scale catchments (5 to $250\,\mathrm{km}^2$). This is the case as, at least in humid regions, water amounts and fluxes in meso-scale catchments have reached dimensions that they are large enough that relevant infrastructure like power plants or drinking water wells either depend or are impacted by their function and state.

*The importance of meso-scale catchments in hydrology.*

Despite of the importance of meso-scale catchments for management decisions it is exactly this spatial scale on which up to date neither a unifying hydrological theory (Sivapalan, 2003) nor a general accepted modeling approach (Hrachowitz and Clark, 2017) has been established. It was James Dooge who first acknowledged this discrepancy in his landmark publication in 1986 where he stated that hydrology as a science should shift its focus to the derivation of new hydrological laws and modeling approaches for meso-scale catchments rather than investing its energy on model calibration (Dooge, 1986). Although Dooge's seminal paper has been published more than 30 years ago his general statements are still relevant and valid today. In the following I will hence shortly re-visit the arguments proposed in Dooge (1986) from an updated, however obviously subjective position.

*Looking for hydrologic laws.*

### *Systems of organized complexity*

The line of argument in the commentary of Dooge (1986) is to a large extent founded on the work of Weinberg (1975), a system theorist, who classified physical systems into three types of systems to better explain why it is so challenging to adapt theoretical solutions to real

*Weinberg's system categories.*

*Systems of organized simplicity.*

world problems. Weinberg's first class of systems is thereby defined as *systems of organized simplicity*. The dynamics of these systems can be approximated in a deterministic, reversible and continuous way, under the condition that the initial states are known, that it is possible to keep track of all state changes in time as well as that the dynamic laws describing the state changes are known. Arguably the most prominent example of this class is our solar system and the somewhat idealized movements of the planets around the sun. For instance, by observing the current trajectories (position and velocity) of the planets in our solar system we are able to apply the law of gravity and make astonishingly precise approximations about the state of the system in the future as well as in the past. For instance, the next total solar eclipse over Karlsruhe (Southern Germany) is going to be on the third of September in the year 2081 at 9:07 a.m. These kind of predictions are based on a concept which is sometimes referred to as locality. Locality means in this context that we can neglect certain forces and features in our consideration if they only slightly change our approximations. Despite these simplifications predictions in systems of organized simplicity are remarkably accurate, especially if compared to our current ability to predict the next flood or drought in hydrology.

*The origin of statistical mechanics and thermodynamics.*

The majority of the history of physics, starting with the qualitative description of nature from Thales, Herakles and Archimedes in ancient Greece up to the early $18^{th}$ century, including the development of the famous Newtonian laws in the $17^{th}$ century, is concerned almost exclusively with systems of organized simplicity (branch of physics referred to as classical mechanics). However, this exclusive way of examining systems changed rather fundamentally in the late $18^{th}$ and early $19^{th}$ century as a result of the experiments and theoretical essays trying to explain the phenomena of temperature and pressure (*historical context* e.g. Ben-Naim, 2008). The early $19^{th}$ century marks thereby a clear milestone in the history of physics. Especially the work of Maxwell, Boltzmann, Gibbs and later Heisenberg and Bohr fundamentally introduced the concept of probability and consequently also the concept of uncertainty into physics (Lindley, 2007) and Feynman once stated appropriately: "*In its efforts to learn as much as possible about nature, modern physics has found that certain things can never be "known" with certainty. Much of our knowledge must always remain uncertain. The most we can know is in terms of probabilities.*" (Feynman, 1963). The two terms, uncertainty and predictability, are hence always inseparably linked to each other and the knowledge about their connection is especially in the field of statistical mechanics which deals after Weinberg with *systems of unorganized complexity* the key to make predictions.

The most common example for systems of unorganized complexity is a cylinder filled with millions of moving molecules. Describing the evolution of such a system contains the paradox that the movements as well as the interactions of its single features still follows classical mechanics, but the sheer quantity of features prevent bookkeeping of their trajectories making predictions in a classical mechanics way impossible (e.g. Popper, 1959). A requirement to make predictions in such systems is hence a switch in perspective, from the micro- to the macroscale. This means that the focus needs to shift from the single particle to the entire gas filled cylinder and that the description of the system are based on aggregated (macroscopic) states rather than on fixed microscopic features and states. Furthermore, if it is possible to assume that the behavior of the microscopic features is to a certain extent random it enables to apply statistical concepts such as the "Law of Large Numbers" from Jakob Bernoulli, following the convention that "[...] *the larger the population involved the more likely we are to observe values that are close to the predicted average values*" (Dooge, 1986). This makes predictions in systems of unorganized complexity - similar to the case of systems of organized simplicity - quite accurate if it is possible to identify the integral system properties (e.g. the average kinetic energy and mass of the molecules) and the macroscopic driving potentials (e.g. the concentration gradient), for instance to estimate the time until two liquids are perfectly mixed by diffusion.

*Systems of unorganized complexity.*

In contrast to the great advances modern physics has made to describe systems of organized simplicity and those of unorganized complexity, most systems an engineer or earth scientist may work with neither belong to one of the two categories. This is, to put it mildly, unfortunate as almost all physical theories and consequently most physical laws an earth scientist might apply were established for these two types of systems. Weinberg and Dooge argue hence that there is a third category of systems called *systems of organized complexity*. These systems fill the space between systems of organized simplicity and unorganized complexity and can neither be treated exclusively in a mechanistic nor statistical manner. This is the case because the number of features within the system is already too high that the system can be treated in a deterministic way, but not high enough that it can be assumed that the behavior of the individual compartments is entirely random. This hampers the identification of macroscopic gradients and system properties. Dooge (1986) argued that most hydrological systems belong to exactly this category and particularly meso-scale catchments "[...] *exhibit a considerable degree of both spatial organization and stochastic heterogeneity, being too large for a fully deterministic treatment yet too small for a simplified conceptual treatment.*" as re-visited by Zehe et al. (2014). There is hence a clear *lack of theory on the scale on which most hydrological decisions are made*. This lack of a distinct theoretical foundation explains to some extent

*Hydrological systems belong to neither of the two above sketch categories.*

*Systems of organized complexity.*

why mid- and long-term predictions as well as uncalibrated model predictions are so uncertain in hydrology (e.g. Hrachowitz et al., 2013), especially if changing boundary conditions and evolving systems need to be considered. It is hence, at least since the work of Dooge (1986), a long standing vision in hydrology that these limitations can only be overcome if we take one step back and re-visit existing theory as well as develop new laws and modeling approaches for meso-scale catchments (Clark et al., 2016; Dooge, 1986; Sivapalan et al., 2003; Zehe et al., 2014).

*Lack of a general accepted theoretical basis in catchment modeling.*

This thesis is motivated and framed by exactly this quest to improve the theoretical underpinning of hydrology as a science with the overall goal to enhance our system understanding as well as our ability to make predictions at the catchment scale. The emphasis is thereby on the development and analysis of hydrological models and on the question how to build, test and distribute models for catchments of the meso-scale. The thesis is structured as follows:

*Chapter II: Representative hillslopes - A modeling concept for the meso-scale.*

*Chapter III: Identifying hydrological similarity to build better distributed models.*

*Chapter IV: Topographic similarity to explain hydrological function.*

I introduce a modeling concept called "representative hillslopes" in chapter 2. This concept was developed for catchments that are already too large to be represented in a fully-distributed way (lack of computer power) but in which a significant body of data is available, making the use of more parsimonious empirical approaches impractical as they are unable to process the available information about the system. I continue the journey to improve our ability to represent meso-scale catchments by models chapter 3. Here, I investigate how the concept of landscape organization and hydrological similarity can support our search for an optimum spatial model structure. I use novel concepts taken from information theory which have not or rarely been applied in hydrology and show how the connection between data compression and landscape organization can be a guidance when setting up environmental models. Founded on the findings in chapter 3 I develop a topographic index based on a straightforward thermodynamic argumentation to improve our ability to divide a landscape into hydrologically similar units in chapter 4. This index can eventually be the foundation for new spatially distributed models and can help us to reduce model complexity a-priori. Finally, in chapter 5, I discuss and synthesize the key findings from the different topics in chapters 2-4 and I identify challenges and opportunities for future research.

### *CAOS - catchments as organized systems*

*The CAOS project.*

This thesis is embedded and motivated within the DFG research unit "Catchments As Organized Systems" (CAOS, Zehe et al., 2014). Within this project, a 256 km$^2$ research catchment located in western Luxembourg (the Attert basin) was extensively monitored for a period of more than 6 years. Among the observation network were 46 meteorological stations, 400 soil moisture probes, 200 piezometers and up to 116 sap flow sensors (Hassler et al., 2017). Furthermore, several geo- and soil-physical measurement campaigns were carried out in the years 2012, 2013 and 2014 collecting in total more than 150 soil cores samples (Jackisch, 2015), around 100 drilling cores (Sprenger et al., 2015) as well as various 2d profiles of the subsurface measured by means of electric resistivity tomography (ERT), refraction seismic and ground penetrating radar (GPR).

*The Attert catchment, a close to "perfect" natural research environment.*

The Attert basin was chosen as research environment because of its special geological and climatic properties. The catchment is divided into three main geological settings (schist, marl and sandstone; Bos et al., 1996) which are all characterized by distinctly different dominant runoff processes (Wrede et al., 2015) while contrasts in the climatic forcing play only a minor role (Pfister et al., 2018). These properties make the Attert catchment a close to perfect natural environment for hydrological research with a focus on runoff generation as the climatic differences between the contrasting geological units do not superimpose their hydrological functions.

All hypotheses formulated in this thesis are tested within the Attert basin. A detailed site description can hence be found in each chapter with an individual focus on the different sub-basins used for the research question in the corresponding chapter. In the following, I introduce the different parts of my thesis in more detail.

## 1.2 CHAPTER 2: THE COMPLEMENTARY MERITS OF TOP-DOWN AND BOTTOM-UP MODELS

*Top-down and bottom-up models.*

All models are simplified system representations, and can accordingly be classified – similar to physical systems - into models starting either with a macroscopic (top-down) or microscopic (bottom-up) system perspective on catchments (e.g. Gao et al., 2019; Hrachowitz and Clark, 2017; Savenije, 2009).

*Bottom-up, physically-based or reductionist models.*

Models following a bottom-up perspective (also referred to as physically-based or reductionist models in hydrology) are based on the concept that each process in a catchment is represented in an explicit manner (Freeze and Harlan, 1969). A typical starting point for such models is a classical mechanistic, eulerian description of the flow system in a soil column. Other processes are then added to the soil domain until the desired level of process diversity is reached. Bottom-up models in hydrology are typically based on the Darcy-Richards equation to simulate water flow in the unsaturated zone, the

Penman-Monteith equation to approximate the net evapotranspiration, a kinematic wave approach to simulate surface runoff and the Darcy law to simulate water fluxes in the saturated zone. To represent the function of an entire catchment with a bottom-up model, numerous soil columns are coupled and spatially distributed using different techniques and resolutions, thereby increasing the dimensions to 2 d or 3 d. In general, the driving potential differences (e.g. resulting from topographic differences) as well as the corresponding dissipative energy losses (e.g. related to the friction at the water-solid interface) of a given process depend on the local system configuration and on the individual location of the soil column with respect to its neighbors. Both the driver and corresponding resistance term can hence vary greatly within a given landscape. At least theoretically it is possible to identify the model parameters of bottom-up model based on measurements and observations. However, as a result of our inability to measure especially the subsurface at the scale of interest in combination with the fact that bottom-up models are founded on laws not appropriate under some conditions (e.g. Darcy-Richards; Beven and Germann, 1982, 2013), bottom-up models often have numerous model parameters which are difficult to assess a-priori by observations. This ambiguity and uncertainty of model parameters makes calibration of bottom-up models often a necessary evil especially if the goal is to produce results with an acceptable performance. Prominent examples of classical bottom-up models are MIKESHE (Hughes and Liu, 2008), HydroGeoSphere (Brunner and Simmons, 2012) and CATFLOW (Zehe et al., 2001). The latter is used in chapter 2 and 3.

*How physically-based can a bottom-up models be?*

Modeling approaches following a top-down perspective (also referred to as conceptual or bucket models in hydrology) start at the other end of the spectrum with the delineation of catchment. The most elementary type of a top-down model is a single linear reservoir representing an entire catchment which embodies a closed control volume with respect to the mass balance. Usually, several reservoirs and lag functions are combined until the desired level of system and process complexity is reached (Gupta et al., 2012). As a result of the macroscopic system representation, driving potential differences as well as the corresponding resistance terms are also of macroscopic nature. For instance, the water table differences within a linear reservoir is can act as a driver for runoff generation, encountered by model parameters which represent integrated system properties that hamper the flow of water in a catchment. As these macroscopic potentials and controls cannot be measured directly on the scale of interest they normally need to be inferred from observed time series of system forcing and response, typically streamflow and rainfall, again requiring model calibration, as it is the case for most bottom-up models. Prominent examples of top-down models are HBV (Bergström and Forsman,

*Top-down, conceptual or bucket models.*

*The necessity to calibrate top-down models.*

1973), FLEX-topo (Fenicia et al., 2011) or TOPMODEL (Beven and Kirkby, 1979).

It is important to note that the border between the different modeling approaches is at best vague, and the plurality of hydrological models certainly spans the entire range between bottom-up and top-down approaches. However, the majority of hydrological models still follows either a top-down or bottom-up perspective especially considering the roots of their development (Gao et al., 2019).

To improve our ability to make predictions on the catchment scale Savenije (2009) and Hrachowitz and Clark (2017) proposed that a promising way forward is to better link and combine top-down (macroscopic perspective) with bottom-up (microscopic perspective) model concepts . This means to identify models which balance complexity with simplicity (Savenije, 2010; Schoups et al., 2008; Zehe et al., 2014), to find better way to integrate and use new observations in hydrological models and finally to facilitate the communication between the two somewhat separated scientific modeling communities in hydrology.

*The combination of top-down and bottom-up models. A way forward?*

*The concept of representative hillslopes*

Following the line of thoughts proposed by Hrachowitz and Clark (2017) I introduce the modeling concept of representative hillslopes in chapter 2. This modeling concept tries to bridge the gap between top-down and bottom-up model approaches by investigating that a single spatially-aggregated 2 d numerical hillslope model can represent an entire meso-scale catchment in a top-down approach using a bottom-up model. Chapter 2 is thereby divided into two main sections:

*Representative hillslopes are bottom-up models which are set up in a top-down manner.*

In the first section, I develop two perceptual models which represent the hydrological functioning of two distinctly different catchments in a top-down manner. The perceptual models are thereby grounded on numerous field observations (e.g. Jackisch et al., 2016), on reported findings about the functioning of the catchment (e.g. Martínez-Carreras et al., 2016) as well as on the qualitative knowledge of field hydrologists, geo-physicists and biologists which worked in the Attert basin during the CAOS project.

In the second section, I translate the two perceptual models into two bottom-up hillslope models. This is done by deriving macroscopic model parameters based on the available mosaic of point measurements as well as on qualitative information, for instance about the occurrence of macropores or subsurface water ponds. These macroscopic relationships cannot be measured directly within the

*Parameterization of the representative hillslopes.*

two research environments but where derived based on the diversity of point measurements available with the overall goal to represent the function of the two catchments. The hillslope models are set up and tested against discharge time series within a few trial and error runs and are finally compared against sap flow velocities (proxy for catchment transpiration) and soil moisture observations (proxy for catchment storage). The results of chapter 2 highlight that the concept of representative hillslopes is indeed promising, especially in catchments where more information about the system is available than rainfall and runoff. However, they also point out the limitations of our spatially aggregated modeling concept as well as of the used bottom-up model itself.

## 1.3 CHAPTER 3: FUNCTIONAL UNITS AND LANDSCAPE ORGANIZATION

The fact that the hydrological functioning of a 20 $km^2$ large catchment can be represented with a single spatially aggregated model is impressive, especially considering the seemingly overwhelming heterogeneity of the subsurface properties in the two experimental catchments (e.g. Jackisch et al., 2017). The results are, however, in line with the findings of chapter 4 where I show that geologically similar sub-basins also share almost identical specific discharges despite the fact that their areas range from 0.5 to 30 $km^2$. The two findings that we are on one hand able to represent the runoff generation of an entire catchment with a spatially-aggregate hydrological model and on the other hand that the hydrological functioning of a 0.5 $km^2$ catchment can be almost identically to 30 $km^2$ catchment give rise to the well established research question if catchments located in a similar natural area are also organized in a hydrological similar way. The identification of hydrological similar landscape entities is thereby the key for any "optimal" measurement network, the foundation of a representative hillslope as well as the basis for any distributed model strategy. This is the case as it allows us to transfer knowledge from one location to another and by that minimize the chances that we miss either important details or produce redundancy in our observations and simulations.

*Functional groups, the basis for representative hillslopes.*

In chapter 3, I investigate the question how the concept of hydrological similarity and the phenomenon of landscape organization can be used to build sophisticated hydrological models. Before, I shortly re-visit the origin of the term "organization" and explain its connection to the concept of "entropy" from a science historical perspective.

*Organization, entropy and hydrological similarity*

The term "organization" and the quantity "entropy" are scientifically inevitably linked to each other (Tolman, 1938). Entropy as a physical quantity was introduced by Clausius in 1854 when he studied the Carnot cycle, essentially trying to explain why heat flows from warm to cold. In its classical thermodynamic definition, entropy reflects the thermal energy per unit temperature of a system that cannot be used to perform mechanical work. The entropy of an isolated system can, following the second law of thermodynamics, only increase which is as shown by Boltzmann a matter of probability, as any system will evolve to the most probable state for which its inner gradients get depleted as long as it is not forced otherwise by its environment. The depletion of gradients and the maximization of entropy are thereby manifested through a re-organization of a system from an unlikely to a more likely state. The term "organization" refers thus to (a) a *process* of how a system will adapt its internal states or structure to given persisting gradient (self-organization) and to (b) a state of its structural properties measured by distance of the system from its entropy maximum (degree of organization).

*The link between entropy and organization.*

*Organization can refer to a process as well as to a state.*

Similar to the example with the cylinder filled with a gas, catchments can be seen as a box consisting of functionally partly independent sub-units (e.g. Zehe et al., 2014). The number of functional elements needed to represent a catchment depends thereby on the degree of landscape organization. In a wide range of catchments around the world it was found that this degree can be surprisingly high and that the number of needed functional units to represent a catchment low (Tucker and Bras, 2000; Yoshida and Troch, 2016). This is the case as land-use, geology and soil properties always evolve together as a team rather than being independent variables meaning that the degree of freedom a hydrological landscape has to evolve its function and structure on the macroscale might be smaller than expected on first glance (see also section 1.4).

*The limited degrees of freedom of landscape evolution.*

A series of different approaches have been proposed in hydrology to divide a landscape into similarly organized sub-units like the representative elementary area (REA) concept by Sivapalan et al. (1987), the hydrological response units (HRU) by Flügel (1996), the representative elementary watersheds (REW) by Reggiani et al. (1998) or the elementary functional units (EFU) by Zehe et al. (2014). While these concepts are all motivated and based on different assumptions, the key idea remains the same: to identify spatial entities of a landscape which function hydrologically similar so that a single unit can represent its entire functional group. The advantage of this spatial division is two-fold. First, a valid classification can improve

*Why functional units?*

our ability to set up distributed hydrological models which are spatially as parsimonious as possible but as complex as necessary (Zehe et al., 2014). Second, functional units can help to identify a minimum number of necessary observation points to capture the hydrological functioning of a landscape, thus avoiding redundant measurements. However, despite these promises, no approach to identify and divide a landscape into hydrological similar units has yet become a standard in hydrology.

*The dynamic nature of hydrological similarity*

*Data compression to identify functional similarity.*

In chapter 3 I shift my focus to the research question how functionally similar landscape units can be identified. To assess this issue I divide a catchment into 105 hillslopes based on an existing river network and represent each hillslope with a hydrological model. These models differ only with respect to parameters that are typically derived from a digital elevation model (DEM) - the hillslope form and aspect - with the goal to examine the specific influence that surface topography has on the function of each sub-unit. The simulations are analyzed by means of Information Theory (Cover and Thomas, 2005) and the results show that the combination of information theory and thermodynamic reasoning provides a promising set of methods to identify hydrological similar landscape units. This is further underpinned by the fact that the identified functional groups could be used to set up a compressed catchment model which consists of 6 instead of 105 hillslopes without losing predictive power. Finally do the results indicate that the definition of hydrologically similar areas cannot be time invariant as hydrological systems move from unorganized to organized states in time. An "optimal" spatial division of a catchment into hydrological similar units therefore also needs to be time variant.

## 1.4 CHAPTER 4: HYDROLOGICAL SIMILARITY EXPLAINED BY TOPOGRAPHIC SIMILARITY

*How to identify functional units directly on the available geo-data.*

The results presented in chapter 3 are encouraging, both for identifying hydrologically similar landscape units and for an improved definition of the term model complexity in hydrology. Especially the capabilities of information theory, as a diagnostic tool, and thermodynamics, for hydrological reasoning, proved to be exceedingly valuable. However, taking a practical viewpoint on model building, the task to set up, run and finally compress a distributed hydrological model as done in chapter 3 is very time-consuming. A long-standing vision in hydrology is hence to identify similar functioning units directly on the available geo-data (e.g. Beven and Kirkby, 1979) without the detour and uncertainties that come along when using a hydrological model. Especially topographic maps are promising here as they are available in a decent resolution around the world and reflect the function of a landscape to a certain extent.

In this regard it is interesting to recall that the different sub-basins located in the Attert catchment function hydrologically almost identically as long as they are located within the same geological setting. As stated above we might wonder if the different catchments within the same geological settings have evolved in a similar fashion and are hence also organized alike with respect to their function.

*Surface topography as fingerprint of a hilly landscape.*

The evolution of a landscape on long-time scales is a constant process of continental uplift and erosion. Its pace and magnitude is controlled by the amount of free energy which is available to perform work and by system properties like soil texture, vegetation cover and parent material (Kleidon et al., 2013). Surface topography reflects maybe the most obvious consequence of this constant evolution and the distribution of wetlands, river valleys and hillslopes mirrors the interaction between the past incoming mass and energy as well as the initial geological setting of a landscape. The energy balance of a catchment is thereby dominated by the solar radiation, nevertheless, in humid regions it is mainly the influx of energy by precipitation which drives landscape evolution and forms channel networks, erodes hillslopes or provides energy to create subsurface preferential flow paths (Zehe et al., 2013). The influx of energy added is thereby determined by the mass and the kinetic energy of precipitation as well as by the surface topography of a catchment. The latter implies that a water drop falling on a mountain top has more potential energy than a water drop falling close to outlet of a catchment. This spatial energy distribution mirrors a unique fingerprint of a landscape and not only reflects the past evolution of a landscape but also controls its recent development and hence its future appearance.

Landscape evolution, surface topography and the available energy to perform work are highly intertwined processes and structural properties (Langbein and Leopold, 1964). It seems hence reasonable to apply an energy-centered perspective on hydrological similarity if our goal is to identify hydrologically similar units based on the topography (e.g Zehe et al., 2014).

*Linking surface topography and runoff generation by means of energy dissipation*

In chapter 4 I derive a topographic index with the goal to identify functional groups based on surface topography. The new index (reduced Dissipation per unit length; rDUNE) is an energy-centered reinterpretation and enhancement of the well-established index HAND

*Energy-centered re-interpretation of HAND.*

(Height Above the Nearest Drainage; Rennó et al., 2008) and fundamentally based on the observation that the majority of the incoming potential energy of precipitation is dissipated when rainfall becomes runoff. I test rDUNE against catchment-wide distributions of the topographic wetness index (TWI, Kirkby, 1976) and HAND and show that rDUNE is indeed a meaningful tool to identify landscapes which are organized similarly and hence function alike. This shows once more, (Zehe et al., 2018), that an energy-centered perspective on hydrology could serve as a general scheme to define a set of closely related measures to improve our ability to identify functionally similar landscape units a priori a modeling task or measurement campaign.

Part II

# PICTURING AND MODELING CATCHMENTS BY REPRESENTATIVE HILLSLOPES

This study is published in the scientific journal Hydrology and Earth System Science (HESS). The remainder of part II is a reprint of:

*Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J. and Zehe, E. (2017): Picturing and modeling catchments by representative hillslopes, Hydrology and Earth System Sciences, 21(2), 1225-1249, doi:* 10.5194/hess-21-1225-2017

# 2 PICTURING AND MODELING CATCHMENTS WITH REPRESENTATIVE HILLSLOPES


## ABSTRACT

This study explores the suitability of a single hillslope as parsimonious representation of a catchment in a physically based model. We test this hypothesis by picturing two distinctly different catchments in perceptual models and translating these pictures into parametric setups of 2d physically based hillslope models. The model parametrizations are based on a comprehensive field data set, expert knowledge and process-based reasoning. Evaluation against stream flow data highlights that both models predicted the annual pattern of stream flow generation as well as the hydrographs acceptably. However, a look beyond performance measures revealed deficiencies in streamflow simulations during the summer season and during individual rainfall-runoff events as well as a mismatch between observed and simulated soil water dynamics. Some of these shortcomings can be related to our perception of the systems and to the chosen hydrological model, while others point to limitations of the representative hillslope concept itself. Nevertheless, our results corroborate that representative hillslope models are a suitable tool to assess the importance of different data sources as well as to challenge our perception of the dominant hydrological processes we want to represent therein. Consequently, these models are a promising step forward in the search of the optimal representation of catchments in physically based models.

## 2.1 INTRODUCTION

The value of physically based hydrological models has been doubted (e.g. Beven, 1989; Savenije and Hrachowitz, 2017) since their idea was introduced by Freeze and Harlan (1969). Physically based models like MikeShe (Refsgaard and Storm, 1995) or CATHY (Camporese et al., 2010) typically rely on the Darcy-Richards concept for soil water dynamics, the Penman–Monteith equation for soil-vegetation-atmosphere exchange processes and hydraulic approaches for overland and stream flow. Each of these concepts is subject to limitations arising from our imperfect understanding of the related processes and is afflicted by the restricted transferability of process descriptions from idealized laboratory conditions to heterogeneous natural systems (Grayson et al., 1992; Gupta et al., 2012).

Nevertheless the usefulness of physically based models as a learning tool to explore how internal patterns and processes control the integral behavior of hydrological systems has been corroborated in several studies. For example Pérez et al. (2011) used Hydrogeosphere (Brunner and Simmons, 2012), together with a regularization scheme for its calibration, to infer how changes in agricultural practices affect the streamflow generation in a catchment. Hopp and McDonnell (2009) explored the role of bedrock topography in the runoff generation using HYDRUS 3-D (Simunek et al., 2006) at the Panola hillslope. Coenders-Gerrits et al. (2013) used the same model structure to examine the role of interception and slope in the subsurface runoff generation. Bishop et al. (2015), Wienhöfer and Zehe (2014) and Klaus and Zehe (2011) used physically based models to investigate the influence of vertical and lateral preferential flow networks on subsurface water flow and solute transport, including the issue of equifinality and its reduction. These and other studies (e.g. Ebel et al., 2008; Scudeler et al., 2016) show that physically based models can be set up using a mix of expert knowledge and observed parameters and may be tested against a variety of observations beyond streamflow – such as soil moisture observations, groundwater tables or tracer breakthrough curves. Such studies are, on the one hand, an option to increase our limited understanding of the processes underlying physically based models (Loague and VanderKwaak, 2004), and on the other hand reveal whether a model allows consistent predictions of dynamics within the catchment and of its integral response behavior (Ebel and Loague, 2006).

Setting up a classical physically based model in a heterogeneous environmental system is, however, a challenge as it requires an enormous amount of highly resolved spatial data, particularly on subsurface characteristics. Such data sets are rare and only available in rather homogeneous systems or in environmental system simulators such as Biosphere 2 LEO (Hopp and McDonnell, 2009). Therefore, it has been a long standing vision to replace fully distributed physically based models with aggregated yet physically based model concepts, for instance the Hillslope Storage Boussinesq approach (HSB, Berne et al., 2005; Troch et al., 2003), the REW approach (Representative Elementary Watershed, e.g. Reggiani and Rientjes, 2005; Zhang and Savenije, 2005) or different dual-continuum approaches (Dusek et al., 2012). The key challenge in applying these concepts to real catchments is the assessment of a closure relationship, which parametrizes (a) hydrological fluxes (Beven, 2006a) and (b) soil water characteristics in an aggregated effective manner (Lee et al., 2007; Zehe et al., 2006). Furthermore, it is not completely clear whether the entire range of variability in subsurface characteristics is relevant for hydrological simulations (Dooge, 1986; Zehe et al., 2014). There are, however, promising concepts emerging, for example the work of Hazenberg et al. (2016), who recently developed a hybrid model consisting of the HSB model in combination with a 1-D representation of the Richards equation for the unsaturated zone.

Regardless of whether one favors physically based, hybrid or more statistical model approaches, a perfect representation of a hydrological system should balance the necessary complexity with the greatest possible simplicity (Zehe et al., 2014). The former is necessary to avoid oversimplification. The latter attempts to avoid the drawbacks of overparametrization (Schoups et al., 2008). In principle there are two ways one can try to reach this optimum model structure: either by starting with a complex system representation, for instance a full 3d catchment model, and simplifying the model structure as much as possible, or by starting at the other end of the spectrum, with the most parsimonious model structure, and proceeding towards higher complexity. In conceptual rainfall–runoff models which follow the HBV concept (Bergström and Forsman, 1973) the most parsimonious model structure for simulating the behavior of a catchment is a single reservoir. In the case of physically based models there is more than one starting point. In flatland catchments without dominant lateral flow processes in the soil one might choose a single soil column. This ”null model” could be refined into multiple parallel acting columns, to capture variability in vegetation and soil properties. This represents the first generation of land surface components in meteorological models (e.g. Niu et al., 2011) and the first generation of models for the catchment-scale dynamics of nitrate (Refsgaard et al., 1999).

However, in hilly or mountainous terrain the smallest meaningful unit is a hillslope including the riparian zone, because rainfall and radiation input depend on slope and aspect, as well as on downslope gradients which cause lateral fluxes in the unsaturated zone (e.g. Bachmair and Weiler, 2011; Zehe et al., 2007). This is the reason why hillslopes are often regarded as the key landscape elements controlling transformation of precipitation and radiation inputs into fluxes and stocks of water (e.g. Bronstert and Plate, 1997), energy (Zehe et al., 2010, 2013) and sediments (Mueller et al., 2010).

The most parsimonious representation of a small catchment in a physically based model could thus be a single representative hillslope. However, the challenge of how to identify such a hillslope has rarely been addressed. This reflects the fact that the identifiability of a representative hillslope has been strongly questioned since the idea was born. For example, (Beven, 2006b) argues that the hillslope form is not uniquely defined nor is it clear whether it is the form that matters, the pattern of saturated areas Dunne and Black (1970) or the subsurface architecture. The enormous spatial variability of soil hydraulic properties and preferential flow paths in conjunction with process non-linearity are additional arguments against the identifiability of representative hillslope models (Beven and Young, 2013). Nevertheless, hillslopes act as miniature catchments (Bachmair and Weiler, 2011), which made Zehe et al. (2014) postulate that structurally similar hillslopes act as functional units for the runoff generation and might thereby be a key unit for understanding catchments of organized complexity (Dooge, 1986). Complementarily, Robinson et al. (1995) showed that the behavior of catchments up to the lower mesoscale (5–50 $\mathrm{km}^2$) are strongly dominated by the hillslope behavior, and Kirkby (1976) highlighted that in catchments extending up to 50 $\mathrm{km}^2$ random river networks had the same explanative power for runoff generation as the real river network. He concluded that as long as river networks are not dominant, the characteristic areas of the catchment hold the key to understanding its functioning.

In this context it is of interest to which extent the parameters of a representative hillslope model can be derived by averaging various structural properties of several hillslopes or plots in a catchment. A promising avenue is to set up the representative hillslope based on a perceptual model which is in turn a generalized and simplified picture of the catchment structure and functioning. This is because perceptual models provide a useful means of facilitating communication between field researchers and modelers (Seibert and McDonnell, 2002) and additionally often represent catchments as hillslope-like cross sections. The general idea to translate a perceptual model into a model structure is not new and has already been applied within a conceptual rainfall–runoff model framework even within the same area (Wrede et al., 2015). The scientific asset of using a physically based model is that the perceptual model provides important information on typical ordinal differences in the hydraulic conductivity of different subsurface strata and the nature and qualitative locations of the dominating preferential flow paths. This information can be implemented in hillslope models in a straight-forward manner. The transformation of a qualitative model structure into a quantitative parametrization of the model depends, however, strongly on the chosen hydrological model and the quality and amount of available data.

*Objective and approach*

We hypothesize that a single hillslope in a physically based model is a parsimonious representation of a small hilly catchment. The objective of this study is to test this hypothesis in a two-step approach:

- First we derive a qualitative model structure of a representative hillslope from our perception of the dominant processes and the related dominant surface and subsurface characteristics in the catchment.
- In the second step we transform this qualitative model structure into a quantitative model structure without the use of an automatic parameter allocation.

The challenge in deriving a qualitative model structure lies in the separation of the important details from the idiosyncratic ones. This process is to a large extent independent of the chosen hydrological model and is strongly related to the available expert knowledge and quality of the data. The transformation of a qualitative to a quantitative model structure on the other hand depends on the chosen model and whether it is for example based on 2d or 3d hillslope module or how rapid flow paths are represented. For this reason the objective of our study is not to ″sell″ our particular model, but to share the way how we distilled the quantitative model setups in our target catchments from available data and to evaluate the ability of this parsimonious physically based model to accurately simulate multiple state and flux variables. During the model setup we intendedly avoided using an optimization algorithm to fit the model to the data. In contrary, we relied on various available observations, process-based reasoning, and appropriate literature data for conceiving our perceptual models and parameterizing the representative hillslope models as their quantitative analogues. More specifically, we use geophysical images to constrain subsurface strata and bedrock topography and derived representative soil-water retention curves from a large data set of undisturbed soil samples. Furthermore, we use observations from soil pits, dye staining experiments and observed leaf area indices (LAI) for our model parametrization. Finally, we benchmark the hillslope models against normalized double mass curves, the hydrograph as well as against distributed soil moisture and sap flow observations.

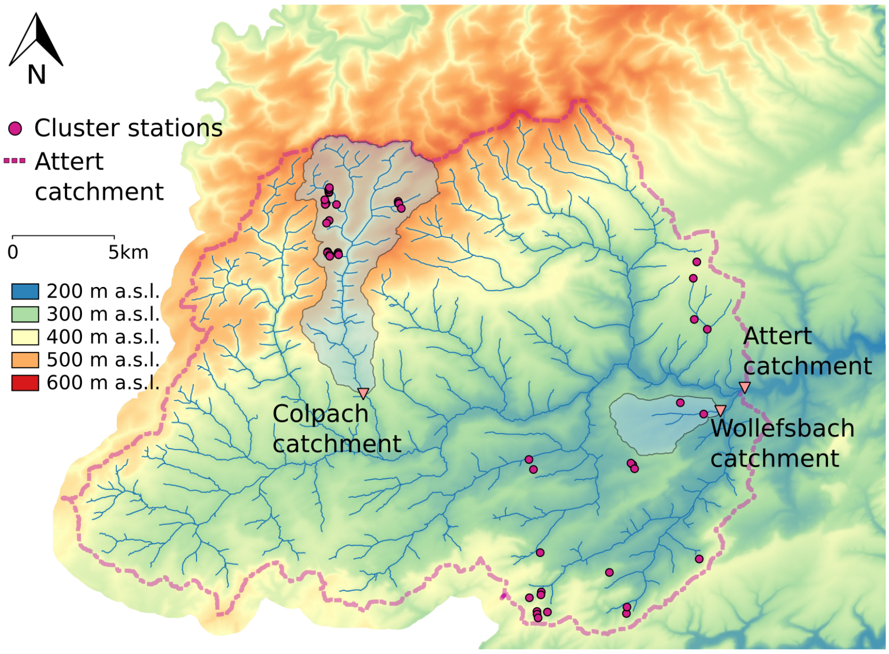


Figure 2.1: Map of the Attert basin with the two selected headwater catchments of this study (Colpach and Wollefsbach). In addition, the cluster sites of the CAOS research unit are displayed.

## 2.2 STUDY AREA, DATABASE AND SELECTED MODEL

We focus our model efforts on two different catchments, the Colpach and the Wollefsbach, located in the Attert experimental basins in Luxembourg (Fig. 2.1, Pfister et al., 2000). These sites offer comprehensive laboratory and field data collected by the CAOS (Catchments As Organized Systems) research unit (Zehe et al., 2014). Besides standard hydrometeorological data the model setup is based on (a) observed soil hydraulic properties of a large number of undisturbed soil cores, (b) 2d electric resistivity profiles in combination with soil pits and augering to infer on bedrock topography, and (c) flow patterns from dye staining experiments and soil ecological mapping of earthworm burrows, to infer the nature and density of vertical preferential flow paths. The representative hillslopes for the two catchments were each set up as a single 2d hillslope in the CATFLOW model (Zehe and Flühler, 2001). The following subsections will provide detailed information on the perceptual models and on the water balance of both catchments. We will shortly refer to the key data and those parts of the model which are relevant for the quantitative model setup, while the appendix provides additional details on both.

### 2.2.1 *The Attert experimental basin*

The Attert basin is located in the mid-western part of the Grand Duchy of Luxembourg and has a total area of 288 $km^2$. Mean monthly temperatures range from 18 °C in July to a minimum of 0 °C in January; mean annual precipitation in the catchment varies around 850 mm (1971–2000; Pfister et al., 2000). The catchment covers three geological formations, the Devonian schists of the Ardennes massif in the northwest, Triassic sandy marls in the center and a small area of sandstone (Jurassic) in the southern part of the catchment (Martínez-Carreras et al., 2012). Our study areas are headwaters named Colpach in the schist area and Wollefsbach in the marl area. As both catchments are located in distinctly different geologies and land use settings, they differ considerably with respect to runoff generation and the dominant controls (e.g. Bos et al., 1996; Fenicia et al., 2013; Jackisch, 2015; Martínez-Carreras et al., 2012; Wrede et al., 2015).

*Colpach catchment: perceptual model of structure and functioning*

The Colpach catchment has a total area of 19.4 $km^2$ and elevation ranges from 265 to 512 m a.s.l. It is situated in the northern part of the Attert basin in the Devonian schists of the Ardennes massif (Fig. 2.1 a). Around 65 % of the catchment is forested, mainly the steep hillslopes (Fig. 2.2 a). In contrast, the plateaus at the hilltops are predominantly used for agriculture and pasture. Several geophysical experiments and drillings showed that bedrock and surface topography are distinctly different. The bedrock is undulating and rough with ridges, depressions and cracks (compare the perceptual model in Fig. 2.3 a and the ERT image in Fig. 2.6 b). Depressions in the bedrock interface are filled with weathered, silty materials which may form local reservoirs with a high water holding capacity. These reservoirs are connected by a saprolite layer of weathered schist which forms a rapid lateral flow path on top of the consolidated bedrock. Rapid flow in this "bedrock interface" is the dominant runoff process (Wrede et al., 2015) and the specific bedrock topography is deemed to cause typical threshold-like runoff behavior similar to the fill-and-spill mechanism proposed by Tromp-Van Meerveld and McDonnell (2006). Further indication that fill-and-spill is a dominant process is given by the fact that the parent rock is reported as impermeable, which makes deep percolation through unweathered schist layers into a large groundwater body unlikely (Juilleret et al., 2011). Furthermore, surface runoff has rarely been observed in the catchment, except along forest roads, which suggests a high infiltrability of the prevailing soils (Bos et al., 1996). This is in line with distributed permeameter measurements and soil sampling performed by (Jackisch, 2015). Moreover, numerous irrigation and dye staining experiments highlight the important role of vertical structures in rapid infiltration and subsequent subsurface


Figure 2.2: (a) Typical steep forested hillslope in the Colpach catchment; (b) soil profile in the Colpach catchment after a Brilliant Blue sprinkling experiment was conducted. The punctual appearance of blue color illustrates the influence of vertical structures on soil water movement in this schist area. (c) Plain pasture site of the Wollefsbach catchment; (d) soil profile in the Wollefsbach catchment after a Brilliant Blue experiment showing the influence of soil cracks and vertical structures on the soil water movement.

runoff formation (Jackisch, 2015; Fig. 2.2). These vertical preferential flow paths, the saprolite layer on top of the impermeable bedrock, the bedrock topography as well as the absence of a major groundwater body are regarded as the dominant structures for the representative hillslope model (Fig. 2.3 a and c).

*Wollefsbach catchment: perceptual model of structure and functioning*

The Wollefsbach catchment is located in the Triassic sandy marls formation of the Attert basin. It has a size of 4.5 km$^2$ and low topographic gradients, with elevation ranging from 245 to 306 m a.s.l. The catchment is intensively used for agriculture and pasture (Fig. 2.2 c); only around 7 % are forested. Hillslopes are often tile-drained (compare the perceptual model sketch in Fig. 2.3 b). The heterogeneous marly

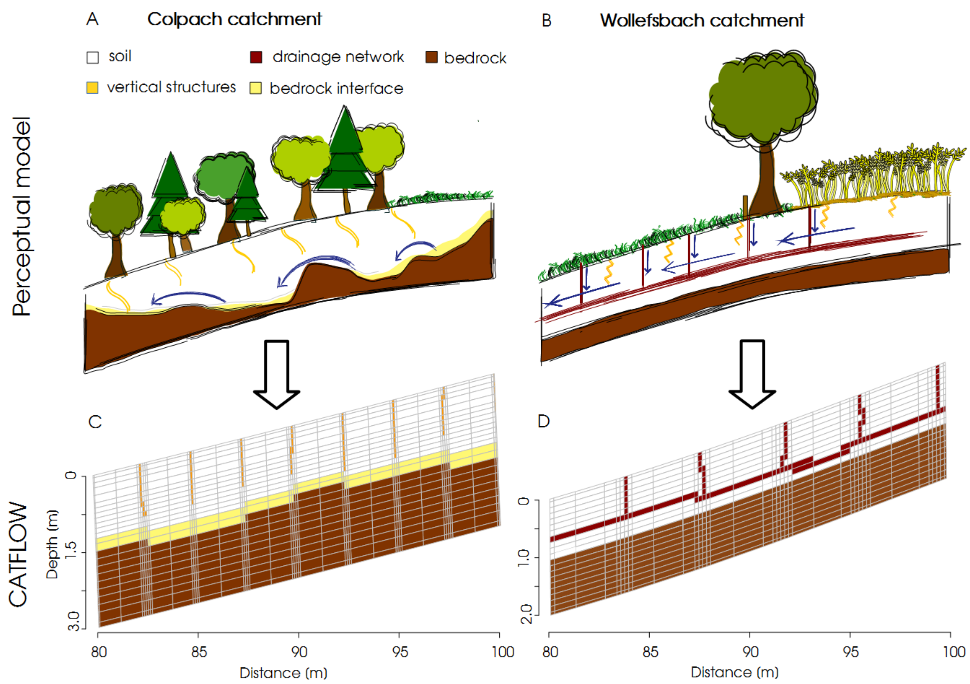


Figure 2.3: Perceptual models of the (a) Colpach and (b) Wollefsbach and their translation into a representative hillslope model for CATFLOW. It is important to note that only small sections of the model hillslope are displayed (C Colpach; D Wollefsbach) and not the entire hillslope.

soils range from sandy loams to thick clay lenses and are generally very silty with high water holding capacities. Similar to the Colpach catchment, vertical preferential flow paths play a major role in the runoff generation, their origin, however, is distinctly different between the seasons. Biogenic macropores are dominant in spring and autumn due to the high abundance of earthworms. Because earthworms are dormant during midsummer and winter, their burrows are partly disconnected by ploughing, shrinking and swelling of the soils (Fig. 2.2 d; see also Fig. 2.4). Soil cracks emerge during long dry spells in midsummer due to the considerable amount of smectite clay minerals in these soils, which drastically increase soil infiltrability in summer (Fig. 2.4). The seasonally varying interaction of both types of preferential flow paths with a dense man-made subsurface drainage network is considered the reason for the flashy runoff regime of this catchment, where discharge rapidly drops to baseflow level when precipitation events end. This is the key feature that needs to be captured by the representative hillslope model. However, as the exact position of the subsurface drainage network and the worm burrows as well as the threshold for soil crack emergence are unknown, the specific influence of each structure on runoff generation in a hydrological model is difficult to estimate.

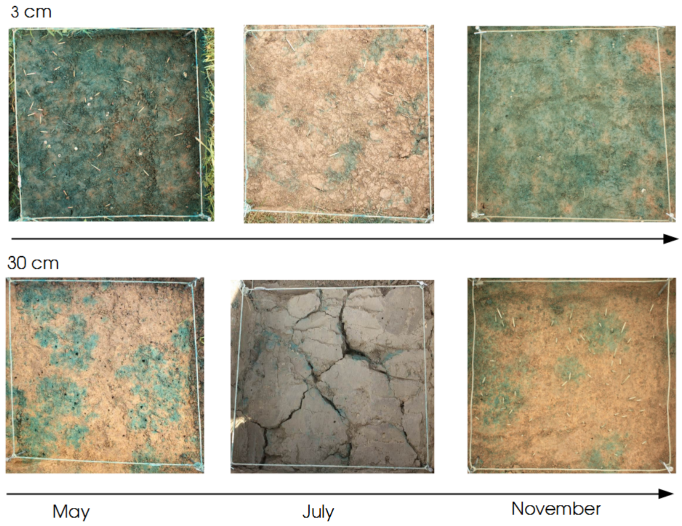


Figure 2.4: Emergent structures in the Wollefsbach catchment for the sampling dates (plot size is 1 $m^2$). In May macropore flow through earthworm burrows dominates infiltration, while in July clearly visible soil cracks occur. In contrast, a more homogenous infiltration pattern is visible in November, especially at 3 cm depth. Figure

*Water balance and seasonality*

The water balance of the Colpach and Wollefsbach catchments for several hydrological years is presented in Fig. 2.5 as normalized double mass curves. Normalized double mass curves relate cumulated runoff to cumulated precipitation, both divided by the sum of the annual precipitation (Pfister et al., 2002; Seibert et al., 2016). Annual runoff coefficients in the Colpach catchment vary around $0.51 \pm 0.06$ among the 4 hydrological years (Fig. 2.5 a). Annual runoff coefficients are smaller in the Wollefsbach catchment than in the Colpach catchment, and vary across a wider range, from 0.26 to 0.46 (Fig. 2.5 b). In both catchments the winter period is characterized by step-like changes which reflect fast water release during rainfall events partly due to rapid subsurface flow. In contrast, the summer regime is characterized by a smooth and almost flat line when vegetation is active. Accumulated rainfall input is not transformed into additional runoff, but is either stored in the system or released as evapotranspiration (Jackisch, 2015). As suggested by Seibert et al. (2017), we used a temperature index model from Menzel et al. (2003) to detect the bud break of the vegetation and to separate the vegetation-controlled summer regime from the winter period in these curves.

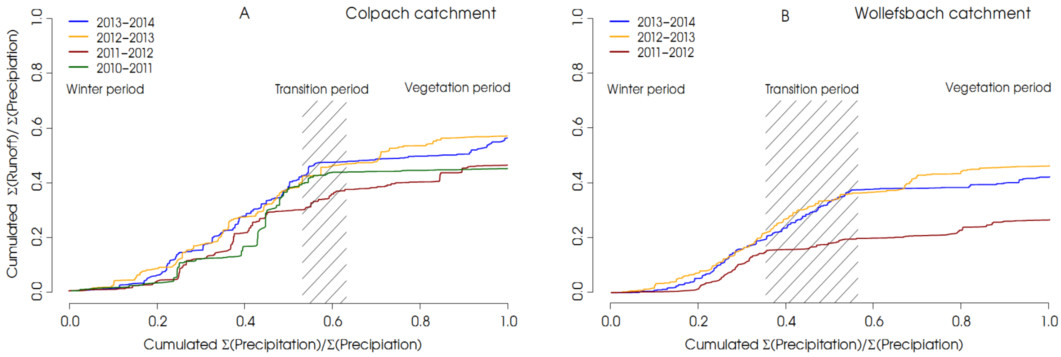


Figure 2.5: Normalized double mass curves for each hydrological year from 2010 to 2014 in the Colpach catchment (a) and from 2011 to 2014 in the Wollefsbach catchment (b). The transition period marks the time of the years when the catchment shifts from the winter period to the vegetation period. The separation of the seasons is based on a temperature index model from Menzel et al. (2003). Since the season shift varies between the hydrological years the transition period is displayed as an area.

### 2.2.2 *Database*

*Surface topography and land use*

Topographic analyses are based on a 5 m LIDAR digital elevation model which was aggregated and smoothed to 10 m resolution. Land use data from the Occupation Biophysique du Sol are based on CORINE land use classes analyzed by color infrared areal images published in 1999 by the Luxembourgian surveying administration, Administration du cadaster et de la Topographie, at a scale of 1 : 15 000.

*Subsurface structure and bedrock topography*

We used hillslope-scale 2d electrical resistivity tomography (ERT) in combination with augerings and soil pits to estimate bedrock topography in the schist area. Our auger profiles revealed, in line with Juilleret et al. (2011) and Wrede et al. (2015) that the vertical soil setup comprises a weathered silty soil layer with a downwards increasing fraction of rock fragments, which is underlain by a transition zone of weathered bedrock fragments and by non-weathered and impermeable bedrock. Based on a robust inversion scheme as implemented in Res2Dinv (Loke, 2003) and additional expert knowledge, the subsurface was subdivided into two main layers of unconsolidated material and solid bedrock. The bedrock interface was picked by the 1500 $\Omega$ m isoline, as explained in detail in the appendix. For our study we used seven ERT profiles from the Colpach area (for an example, see Fig. 2.6 b). Due to the very different geological setting in the marl region (high clay content and alternating sedimentary layering), we could not estab-

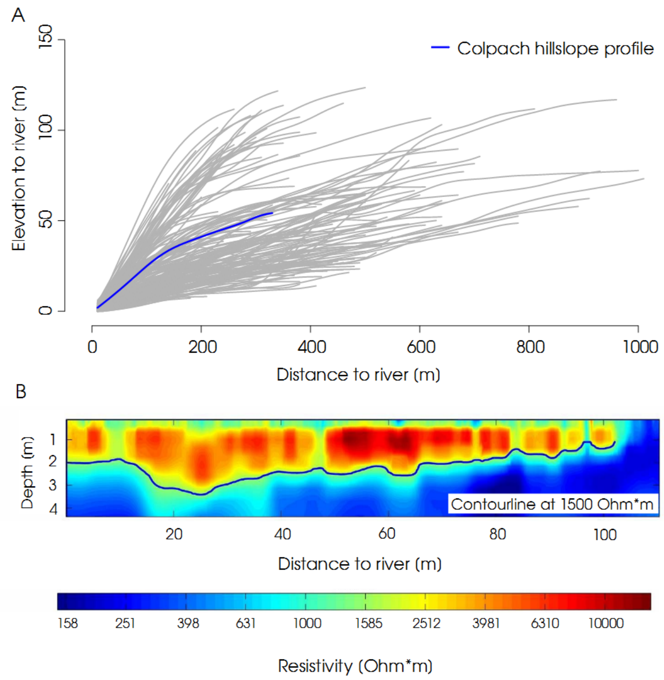


Figure 2.6: (a) Profile of all hillslopes extracted from a DEM in the Colpach catchment. Hillslope profile we used in this study high- lighted in blue. (b) Bedrock topography of a hillslope in the schist area measured using ERT. The contour line displays the 1500 Ωm isoline which is interpreted as the soil–bedrock interface.

lish a relation between bedrock depth and the electrical conductivity data for this region. Therefore, the available ERT data do not provide information on depth to bedrock for this geological setting and we had to rely on auger profiles to estimate the average soil depth.

*Soil hydraulic properties*

We determined soil texture, saturated hydraulic conductivity and the soil water retention curve for 62 soil samples in the schist area and 25 in the marl area. Particularly for the soil hydraulic functions, Jackisch (2015) and Jackisch et al. (2017) found large spatial variability, which was neither explained by slope position nor by the soil depth at which the sample was taken (Fig. 2.7). As our objective was to assess the most parsimonious representative hillslope model, we neglected this variability but used effective soil water characteristics for both catchments instead. These were not obtained by averaging the parameters of the individual curves, but by grouping the observation points of all soil samples for each geological unit and averaging them in steps of 0.05 pF. We then fitted a van Genuchten–Mualem model using a maximum likelihood method to these averaged values (Table 2.1 and Fig. 2.7). The appendix provides additional details on measurement devices and on the dye staining experiments.

*Meteorological forcing and discharge*

Meteorological data are based on observations from two official meteorological stations (Useldange and Roodt) provided by the Administration des services techniques de l'agriculture Luxembourg. Air temperature, relative humidity, wind speed and global radiation are provided with a temporal resolution of 1 h, while precipitation data are recorded at an interval of 5 min. Precipitation was extensively quality checked against six disdrometers which are stationed within the Attert basin and by comparing several randomly selected rainfall events against rain radar observations, both using visual inspection. Discharge observations are provided by the Luxembourg Institute of Science and Technology (LIST).

*Sap flow and soil moisture data*

The Attert basin is instrumented with 45 automated sensor clusters. A single sensor cluster measures inter alia rainfall and soil moisture in three profiles with sensors at various depths. In this study we use 38 soil moisture sensors located in the schist area and 28 sensors located in the marl area, at depths of 10 and 50 cm. Furthermore we use sap flow measurements from 28 trees at 11 of the sensor cluster sites. The measurement technique is based on the heat ratio method (Burgess et al., 2001); sensors are East 30 Sensors three needle sap flow sensors. As a proxy for sap flow we use the maximum sap velocity of the measurements from three xylem depths (5, 18 and 30 mm) as recorded by each sensor. To represent the daytime flux, we use 12 h daily means between 08:00 and 20:00 lt. For further technical details on the sap flow measurements, see Hassler et al. (2017).

### 2.2.3 *Physically based model CATFLOW*

Model simulations were performed using physically based hydrological model CATFLOW (Maurer, 1997; Zehe and Flühler, 2001). CATFLOW consists of a 2d hillslope module which can optionally be combined with a river network to represent a catchment (with several hillslopes). The model employs the standard physically based approaches to simulate soil water dynamics, optional solute transport, overland and river flow and evapotranspiration, which were already mentioned in the introduction and are described in more detail in the Appendix. In the following we will only explain the implementation of rapid flow paths in the model, as this aspect differs greatly from model to model.

Table 2.1: Hydraulic and transport parameter values used for different materials in the model setups.

| Type of structure | Hydrau. conductivity $K_s (m\,s^{-1})$ | Total porosity $\theta_s(-)$ | Residual para. $\theta_r(-)$ | Alpha value $\alpha(m^{-1})$ | Shape para. $n(-)$ |
|---|---|---|---|---|---|
| Colpach | | | | | |
| Soil layer | $5 \times 10^{-4}$ | 0.57 | 0.05 | 2.96 | 1.25 |
| Macropores & soil bedrock interface | $1 \times 10^{-3}$ | 0.25 | 0.1 | 7.5 | 1.5 |
| Bedrock | $1 \times 10^{-9}$ | 0.2 | 0.05 | 0.5 | 2 |
| Wollefsbach | | | | | |
| Soil layer | $2.92 \times 10^{-4}$ | 0.46 | 0.05 | 0.66 | 1.05 |
| Drainage system | $1 \times 10^{-3}$ | 0.25 | 0.1 | 7.5 | 1.5 |
| Bedrock | $1 \times 10^{-9}$ | 0.2 | 0.05 | 0.5 | 2 |

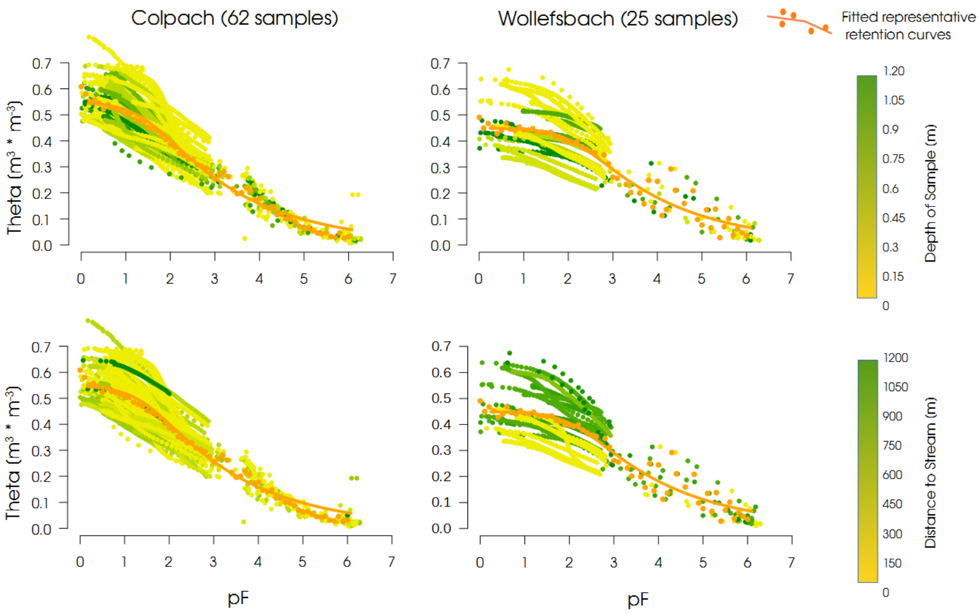


Figure 2.7: Fitted soil water retention curves and measured soil water retention relationships for the Colpach (a) and Wollefsbach (b) catchments.

*Generation of rapid vertical and lateral flow paths*

Vertical and lateral preferential flow paths are represented as a porous medium with high hydraulic conductivity and very low retention. This approach has already been followed by others (Castiglione et al., 2003; Lamy et al., 2009; Nieber and Sidle, 2010; Nieber and Warner, 1991), and is one of many ways to account for rapid flow paths in physically based models. However, it is important to note that such a macropore representation is obviously not an image of the real macropore configuration given the typical grid size of a few centimeters, but a conceptualization to explicitly represent parts of the subsurface with prominent flow paths and the adjacent soil matrix in an effective way. The approach includes the assumption that preserving the connectedness of the rapid flow network (Fig. 2.3) is more important than separating rapid flow and matrix flow into different domains. Implementations of this approach with CATFLOW were successfully used to predict hillslope-scale preferential flow and tracer transport in the Weiherbach catchment, a tiledrained agricultural site in Germany (Klaus and Zehe, 2011), and at the Heumöser hillslope, a forested site with fine textured marly soils in Austria (Wienhöfer and Zehe, 2014). The locations of vertical macropores may either be selected based on a fixed distance or via a Poisson process based on the surface density of macropores. From these starting points the generator stepwise extends the vertical preferential pathways downwards to a selected depth, while allowing for a lateral step with a predefined probability of typically 0.05 to 0.1 to establish tortuosity. Lateral preferential flow paths to represent either pipes at the bedrock interface or the tile drains are generated in the same manner: starting at the interface to the stream and stepwise extending them upslope, again with a small probability of a vertical upwards or downwards step to allow for tortuosity (Fig. 2.3 c and d).

## 2.3 PARAMETRIZATION OF THE REPRESENTATIVE HILLSLOPE MODELS

### 2.3.1 *Colpach catchment*

*Surface topography and spatial discretization*

We extracted 241 hillslope profiles based on the available DEM in the Colpach catchment using Whitebox GIS (Lindsay J.B., 2014) following the LUMP approach (Landscape Unit Mapping Program, Francke et al., 2008). Based on these profiles (Fig. 2.6 a) we derived a representative hillslope with a length of 350 m, a maximum elevation of 54 m above the stream, and a total area of 42 600 m$^2$. The hillslope has a mean slope angle of 11.6° and faces south (186°), similar to the average aspect of the Colpach catchment. The first step in generating the representative hillslope profile was to calculate the average distance to the river of all 241 extracted hillslope profiles as equal to 380 m. In the next step all elevation and width values of the profiles were binned into 1 m "distance classes" from the river ranging up to the average distance of 380 m. For each class the median values of the (a) elevation above the stream and (b) the hillslope width were derived and used for the representative hillslope profile (Fig. 2.6 a). For numerical simulation the hillslope was discretized into 766 horizontal and 24 vertical elements with an overall hillslope thickness of 3 m. The vertical grid size was set to 0.128 m, with a reduced vertical grid size of the top node of 0.05 m. Grid size in the downslope direction varied between 0.1m within and close to the rapid flow path and 1 m within reaches without macropores (Fig. 2.3 c). The hillslope thickness of 3 m was chosen to reflect the average of the deepest points of the available bedrock topographies extracted from ERT profiles, which was 2.7 m. Boundary conditions were set to the atmospheric boundary at the top and the no flow boundary at the right margin. At the left boundary of the hillslope we selected the seepage boundary condition, where outflow only occurs under saturated and no flow under unsaturated conditions. A gravitational flow boundary condition was established for the lower boundary. We used spin-up runs with initial states of 70 % saturation for the entire hydrological year of interest and used the resulting soil moisture pattern for model initialization. This initialization approach was also used for the Wollefsbach catchment.

*Land use and vegetation parametrization*

According to the land use maps, the hillslopes are mostly forested. As the hilltop plateaus account for only a very small part of the representative hillslope, the land use type for the entire hillslope is set to forest (Fig. 2.2 a). The start and end of the vegetation period were defined using the temperaturedegree model of Menzel et al. (2003), which allowed successful identification of the tipping point between the winter and vegetation season in the double mass curves of the Colpach and of the Wollefsbach (compare Fig. 2.5 a and b).We further used observed LAI to parametrize the evapotranspiration routine. However, since only 14 single measurements at different positions are available for the entire schist area and vegetation period, we use the median of all LAI observations from August as a constant value of 6.3 for the vegetation period. To account for the annual pattern of the vegetation phenology we interpolate the LAI for the first and last 30 days of the vegetation period linearly between zero and 6.3, respectively. The other evapotranspiration parameters are displayed in Table 2.2 and were taken from Breuer et al. (2003) or Schierholz et al. (2000).

*Bedrock topography, permeability and soil hydraulic functions*

We used the shape of the bedrock contour line of the ERT image (Fig. 2.6) to constrain the relative topography of the bedrock interface in the hillslope model as follows. We scaled the 100 m of bedrock topography to the hillslope length of 380 m. We then used the average depth to bedrock from all seven available ERT measurements (2.7 m) to scale the maximum depth to bedrock in our model. To this end we divided the average depth of 2.7 m by the deepest point of the bedrock in Fig. 2.6 b (3.3 m) and used the resulting factor of 0.88 to reduce the bedrock depth of Fig. 2.6 b relatively at all positions. As a result, the soil depths to the bedrock interface vary between 1 and 2.7 m, with local depressions that form water holding pools. Since no major groundwater body is suspected and no quantitative data on the rather impermeable schist bedrock in the Colpach are available, we use a relatively impermeable bedrock parametrization suggested by Wienhöfer and Zehe (2014) (Table 2.1). It is important to note that due to this bedrock parametrization water flow through the hillslope lower boundary tends to zero. The silty soil above the bedrock was modeled with the representative hydraulic parameters obtained from field samples listed in Table 2.1. Since there was no systematic variation of hydraulic parameters of the individual soil samples with depth, soil hydraulic parameters were set constant over depth, except for porosity, which was reduced to a value of 0.35 $m^3\,m^{-3}$ at 50 cm depth to account for the increasing skeleton fraction of around 40 % in deeper soil layers.

*Rapid subsurface flow paths*

Macropore depths were drawn from a normal distribution with a mean of 1 m and a standard deviation of 0.3 m. These values are in agreement with the mean soil depth and correspond well to the results of dye staining experiments performed by Jackisch (2015) and Jackisch et al. (2017). Additionally, macropores were slightly tortuous, with a probability of a lateral step of 5 %. Since no observations for the macropore density were available, we use a fixed macropore distance of 2 m. The macropore distance was chosen rather arbitrarily to reflect their relative density in the perceptual model and to establish a partly connected network of vertical and lateral rapid flow paths. The vertical flow paths were parametrized using an artificial porous medium with high hydraulic conductivity and low retention properties proposed by Wienhöfer and Zehe (2014) (Table 2.1). Also, the weathered periglacial saprolite layer which is represented by a 0.2m thick layer above the bedrock was parametrized as a porous medium following (Wienhöfer and Zehe, 2014). The estimated saturated hydraulic conductivity of $1 \times 10^3$ $m\,s^{-1}$ corresponds well to the velocities described by Angermann et al. (2017). This ensures that the Reynolds number is smaller

Table 2.2: Vegetation parameter values for the different land use forms in the model setup.

| | Start-end of the veg. period | LAI | Root depth | Through-fall rate | Plant height | Inter-ception | Max. stomata conduc-tance | Albedo |
|---|---|---|---|---|---|---|---|---|
| | (doy) | (-) | (m) | % | (m) | (mm) | (mm $s^{-1}$) | (-) |
| Colpach | | | | | | | | |
| Forest (Fagus sylvatica) | 97-307 | 6.3[4] | 1.8 | 95 | 24[4] | 2 | 5 | 0.2 |
| Wollefsbach | | | | | | | | |
| Corn (Zea mays) | 97-307 | 4[2] | 1.2[1] | 100 | 2 | 3 | 2.5 | 0.2 |
| Drainage system | 97-274 | 6[2] | 1.3[3] | 100 | 0.4 | 3.1[3] | 2.5 | 0.2 |

[1]Value for gley brown soils [2] mean value (Breuer et al., 2003) [3] Trifolium spec. [4] observed

than 10, implying that flow can be considered laminar and that the application of Darcy's law is still appropriate (Bear, 1972).

### 2.3.2 *Wollefsbach catchment*

#### *Surface topography and spatial discretization*

Since only eight relatively similar hillslope profiles were derived from the DEM in the Wollefsbach, we randomly chose one of those with a length of 653 m, a maximal elevation above the river of 53 m and an area of 373 600 $m^2$. The hillslope has a mean slope angle of 8.1 ° and faces south (172°). The hillslope was discretized into 553 horizontal and 21 vertical elements with an overall hillslope thickness of 2 m (Fig. 2.3 d). The vertical grid size was set to 0.1 m, with a reduced top and bottom node spacing of 0.05 m. Grid size in the lateral direction varied between 0.2 m within and close to the rapid flow paths and 2 m within reaches without macropores (Fig. 2.3 b and d).

#### *Land use and vegetation parametrization*

Land use was set to grassland within the steeper and lower part of the hillslope, and set to corn for larger distances to the creek ($> 325$ m). Due to the absence of local vegetation data we used tabulated data characterizing grassland and corn from Breuer et al. (2003). The start and end points of the vegetation period for the grassland and the start point for the corn cultivation were again identified by the temperature index model of Menzel et al. (2003). The vegetation period for the corn cultivation ends at the beginning of October since this is the typical period for harvesting. The intra-annual vegetation dynamics were taken from Schierholz et al. (2000).

*Bedrock topography, permeability and soil hydraulic functions*

In contrast to the Colpach, geophysical measurements and augerings revealed bedrock and surface to be more or less parallel. Soil depth was set to a constant 1m and the soil was parametrized using the representative soil retention curves shown in Fig 2.7. The bedrock was again parametrized according to values Wienhöfer and Zehe (2014) proposed for the impermeable bedrock at the Heumöser hillslope in Austria (Table 2.1), which is also in a marl geology.

*Rapid subsurface flow paths*

Based on the perceptual model (Fig. 2.3 b and d) and the reported vertical and lateral drainage structures in the catchment, we generated a network of fast flow paths. The depths of the vertical flow paths were drawn from a normal distribution with a mean of 0.8 m and a standard deviation of 0.1 m. The tile drain was generated at the standard depth of 0.8 m extending 400 m upslope from the hillslope–creek interface. Due to the apparent changes in soil structure either by earthworm burrows or emergent soil cracks (Fig. 2.4), we used different macropore setups for the winter and vegetation seasons. For the winter setup we implemented vertical drainage structures every 4 m. In the summer setup we added fast flow paths every 2 m to account for additional cracks and earthworm burrows. The positions of the conceptual macropores were selected again arbitrarily to create an image of the perceptual model and to ensure that the soil surface and the tile drain were well connected. Vertical flow paths and the tile drain were parametrized similarly to the Colpach with the same artificial porous medium (Table 2.1). Boundary conditions of the hillslope, initialization and the spin-up phase were the same as described for the Colpach model.

### 2.3.3 *Model scenarios*

Both hillslope models were set up within a few test simulations to reproduce the normalized double mass curves in both catchments of the hydrological year 2014. Within those trials we compared for instance setups without and with an arbitrary selected density of macropores, but we did not perform an automated parameter allocation as stated above. We choose the normalized double mass curves as a fingerprint of the annual pattern of runoff generation since it is particularly suitable for detecting differences in the inter-annual and seasonal runoff dynamics of a catchment (Jackisch, 2015). Model performance was judged by visual inspection as well as by using the

Kling–Gupta efficiency (KGE, Gupta et al., 2009).

In a second step we compared the simulated overland flow and subsurface storm flow across the left hillslope boundary to observed discharge. Water leaving the hillslope through the lower boundary was neglected from the analysis because in both setups the total amount was smaller than 1 % of the overall hillslope outflow.We compared the specific discharge of the hillslopes to the observed specific discharge of the two catchments in mm $h^{-1}$ by dividing measured and simulated discharge by the area of the catchments and the hillslopes. Our goal was to test whether our hillslope models represented the typical subsurface filter properties which are relevant for the runoff generation in both selected hydrological landscapes (schist and marl areas in the Attert basin). We measured the model performance with respect to discharge, again based on the KGE. Since it is advisable to calculate and display various measures of model performance (Schaefli and Gupta, 2007), we calculated the Nash–Sutcliffe efficiency (NSE; a measure of model performance with emphasis on high flows) and the logarithmic NSE (logNSE; a performance measure suited for low flows). As both catchments are characterized by a strong seasonality, we further separated the simulation period into winter and vegetation periods and calculated the KGE, NSE as well as the logNSE separately for each of the seasons. In addition, we followed Klemeš (1986) and performed a proxy-basin test to check whether the runoff simulation is transposable within the same hydrological landscape and conducted a split sampling to examine whether the models also work in the hydrological year of 2013. Finally, we judged the model goodness visually for selected rainfall–runoff events.

In a third step we evaluated the model setups against available soil moisture observations. A natural starting point for a modeling study would be to classify the available soil moisture observations for instance by their landscape position. However, similar to the case of the soil water retention properties, the small-scale variability of the soil properties seems to be too dominant, as grouping according to hillslope position was not conclusive (Jackisch, 2015; Appendix A4). We therefore extracted simulated soil moisture at 20 virtual observation points at different downslope positions at the respective depths of the soil moisture observations (10 and 50 cm), and compared the median of the simulated virtual observations against the 12 h rolling median of the observed soil moisture using the KGE and the Spearman rank correlation. Finally, we analyzed simulated transpiration of the Colpach model by plotting it against the 3-day rolling median of the daily sap flow velocities observed in the schist area of the Attert basin. As sap flow is a velocity and transpiration is a normalized flow, they are not directly comparable. This is why we normalized both observed

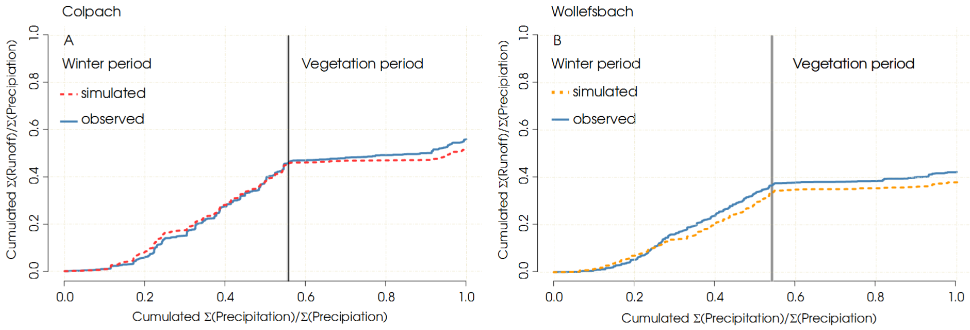


Figure 2.8: Simulated and observed normalized double mass curves of (a) the Colpach catchment and (b) the Wollefsbach catchment. The double mass curves are separated into a winter period and a vegetation period following Menzel et al. (2003)

sap flow and simulated transpiration by dividing their values by their range and only discuss the correlation among the normalized values. The visual inspection shows additionally to which extent maximum and minimum values of both normalized time series coincide. This cannot be inferred from the correlation coefficient.

## 2.4 RESULTS

### 2.4.1 *Normalized double mass curves and discharge*

The hillslope models reproduce the typical shape of the normalized double mass curves – the steep, almost linear increase in the winter period and the transition to the much flatter summer regime - in both catchments very well (Fig. 2.8 a and b). In both catchments subsurface flow is at 99 % in the Colpach and at 94 % in the Wollefsbach, the dominant form of simulated runoff.

The KGEs of 0.92 and 0.9 obtained for the Colpach and the Wollefsbach, respectively, confirm that within the error ranges both double mass curves are explained well by the models. As a major groundwater body is unlikely in both landscapes, a large inter-annual change in storage is not suspected and we hence state that the hillslope models closely portray the seasonal patterns of the water balance of the catchments. This is further confirmed by the close accordance of simulated and observed annual runoff coefficients. We obtain 0.52 compared to the observed value of 0.55 in the Colpach and 0.39 compared to an observed value of 0.42 in the Wollefsbach.

In addition to the seasonal water balances, both models also match observed discharge time series in an acceptable manner (KGE 0.88 and 0.71; Table 2.3). A closer look at the simulated and observed runoff time series (Figs. 2.9 and 2.10) reveals that the model performance differs in both catchments between the winter and summer seasons. Generally we observe a better model accordance during the wet winter season, when around 80 % of the overall annual runoff is generated in both catchments. In contrast, there are clear deficiencies during dry summer conditions. This is also highlighted by the different performance measures which are in both catchments higher during the winter period than during the vegetation period (Table 2.3).

The Colpach model misses especially the steep and flashy runoff events in June, July and August, and underestimates discharge in summer. It also misses the characteristic double peaks of the catchment as highlighted by runoff events 2 and 3 (Fig. 2.9). Although the model simulates a second peak, it is either too fast (event 2) or the simulated runoff of the second peak is too small (event 3). This finding suggests that our perceptual model of the Colpach catchment needs to be revised, as further elaborated in the discussion. Another shortcoming is the missing snow routine of CATFLOWwhich can be inferred from event 1 (Fig. 2.9, top left panel). While snow is normally not a major control of runoff generation in the rather maritime climate of the Colpach catchment, the runoff event 1 happened during temperatures below zero and was most likely influenced by snowfall and subsequent snowmelt, which might explain the delay in the observed rainfall–runoff response.

In the Wollefsbach model the ability to match the hydrograph also differed strongly between the different seasons (Table 2.3; Fig. 2.10). The flashy runoff response in summer is not always well captured by the model, as for example for a convective rainfall event with rainfall intensities of up to 18 mm 10 min $s^{-1}$ in August (Fig. 2.10, event 2). On the contrary, runoff generation during winter is generally simulated acceptably (KGE = 0.74). Yet, the model strongly underestimates several runoff events in winter too (Fig. 2.10, event 1). As temperatures during these events were close to zero, this might again be a result of snow accumulation, which cannot be simulated with CATFLOW due to the missing snow or frozen soil routine. It is of key importance to stress that we only achieve acceptable simulations of runoff production in the Wollefsbach when using two different macropore setups for the winter and the summer periods to account for the emergence of cracks (Fig. 2.4) by using a denser 2 m spacing of macropores. When using a single macropore distance of either 2 m (summer setup) or 4 m (winter setup) in the entire simulation period, the model shows clear deficits with a KGE of 0.61 and 0.53, respectively. Furthermore, we are able to improve the performance of the Wollefsbach model if we use values of saturated hydraulic conductivity faster than $1 \times 10^{3}\,\mathrm{m\,s^{-1}}$

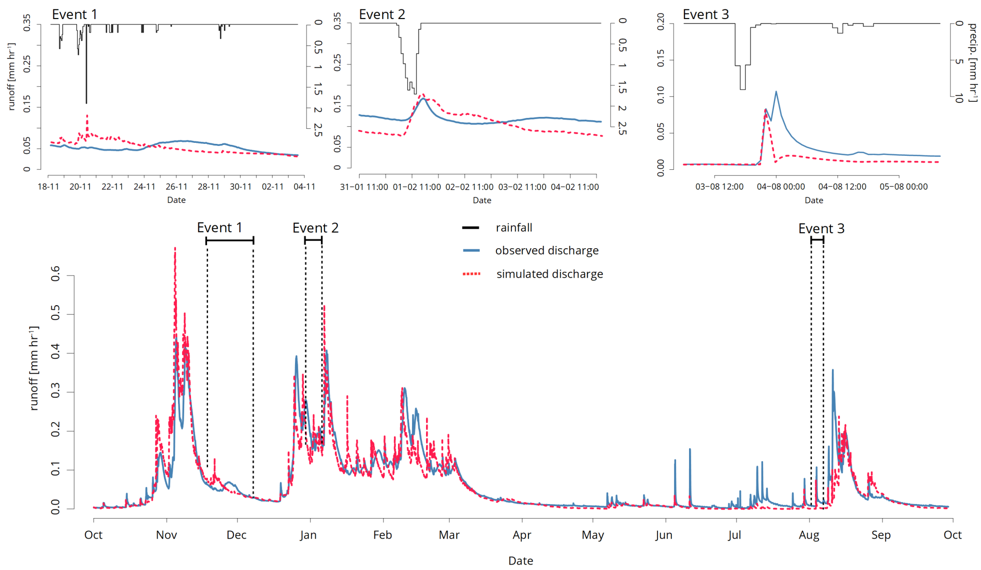


Figure 2.9: Observed and simulated runoff of the Colpach catchment. Moreover, three rainfall–runoff events are highlighted and displayed separately.

for the drainage structures. However, this violates the laminar flow assumption and the application of Darcy's law becomes inappropriate.

### 2.4.2 *Model sensitivities, split sampling and spatial proxy test*

Sensitivity tests for the Colpach reveal that the model performance of matching the double mass curves is strongly influenced by the presence of connected rapid flow paths. A complete removal of either the vertical macropores or the bedrock interface from the model domain decreases the model performance considerably (KGE 0.71 or 0.72, respectively). In contrast, reducing the density of vertical macropores from 2 to 3 or 4 m only leads to a slight decrease in model performance (KGE 0.85 and 0.82, respectively). In an additional sensitivity test we changed the bedrock topography from the one inferred from the ERT data to a surface parallel one, which reduces model performance with respect to discharge (KGE $< 0.6$).

The temporal split sampling reveals that the representative hillslope model of the Colpach also performs well in matching the hydrograph of the previous hydrological year 2012–2013 (KGE $= 0.82$). Furthermore, the parameter setup was tested within uncalibrated simulations for the Weierbach catchment (0.45 $km^2$), a headwater of the Colpach in the same geological setting. This again leads to acceptable results (KGE $= 0.81$, NSE $= 0.68$). The same applies to the representative hillslope model of the Wollefsbach, which also performs

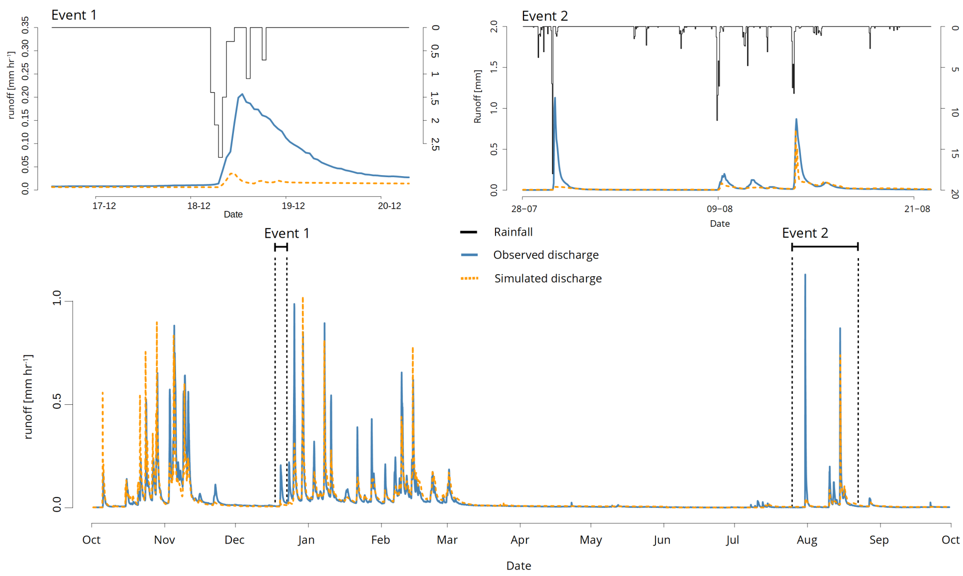


Figure 2.10: Observed and simulated runoff of the Wollefsbach catchment. Two rainfall–runoff events are highlighted and displayed separately

Table 2.3: Benchmarks for simulated double mass curves and simulated discharge for all model setups used in this study.

| Model setup | Double mass curve | Discharge | | |
|---|---|---|---|---|
| | KGE | KGE | NSE | logNSE |
| Colpach models | | | | |
| Reference Colpach model | 0.92 | 0.88 | 0.79 | 0.25 |
| Performance winter | 0.95 | 0.88 | 0.75 | 0.93 |
| Performance summer | 0.49 | 0.52 | 0.51 | 0.62 |
| Wollefsbach models | | | | |
| Reference Wollefsbach model | 0.9 | 0.71 | 0.68 | 0.87 |
| Performance winter | 0.85 | 0.74 | 0.7 | 0.84 |
| Performance summer | 0.74 | 0.28 | 0.33 | 0.57 |

well in matching the hydrograph of the previous year (KGE = 0.7). Furthermore, the parameter setup was tested within an uncalibrated simulation for the Schwebich catchment (30 km$^2$), a headwater of the Attert basin in the same geological setting as the Wollefsbach, and again with acceptable results (KGE = 0.81, NSE = 0.7).

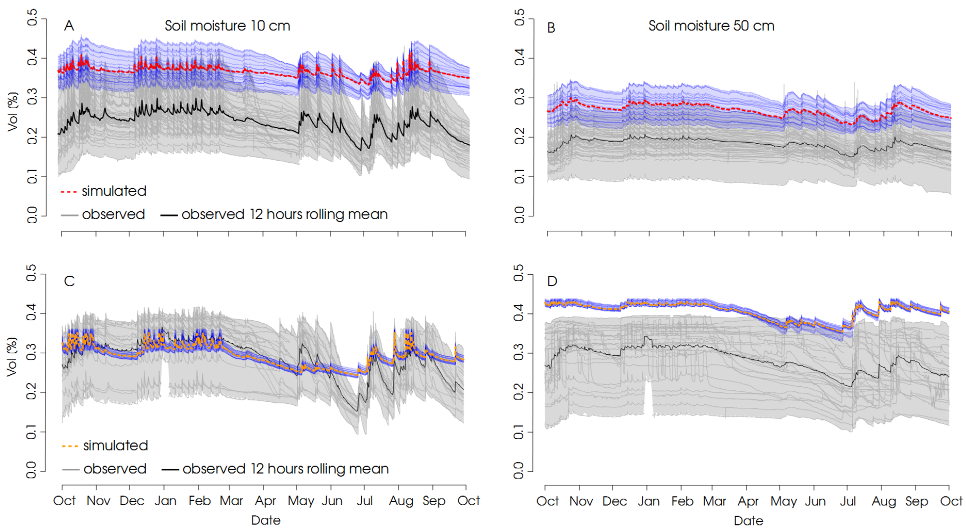


Figure 2.11: Observed soil moisture at 10 and 50 cm depths in the schist (a, b) and marl (c, d) areas of the Attert catchment. Additionally the 12 h rolling median (black) derived from the soil moisture observations and the simulated soil moisture dynamics at the respective depths (red Colpach; orange Wollefsbach) are displayed.

### 2.4.3 *Simulated and observed soil moisture dynamics*

We compare the ensemble of soil moisture time series from the virtual observation points to the ensemble of available observations (Fig. 2.11). In the Colpach, soil moisture dynamics are matched well (Spearman rank correlation $r_s = 0.83$). This is further confirmed when comparing this value to the median Spearman rank correlation coefficient of all sensor pairs ($r_s = 0.66$). However, simulated soil moisture at 10 cm depth was systematically higher than the average of the observations. The predictive power in matching the observed average soil moisture dynamics was small (KGE = 0.43; Fig. 2.11 a). In contrast to the positive bias, the total range of the simulated ensemble appears, at 0.1 m$^3$ m$^{-3}$, much smaller than the huge spread in the observed time series (0.25 m$^3$ m$^{-3}$). In line with the model performance in simulating discharge, the model has deficiencies in capturing the strong declines in soil moisture in June and July. Simulated soil moisture at 50 cm depth exhibits a strong positive bias and again underestimates the spread in the observed time series. The predictive power is slightly

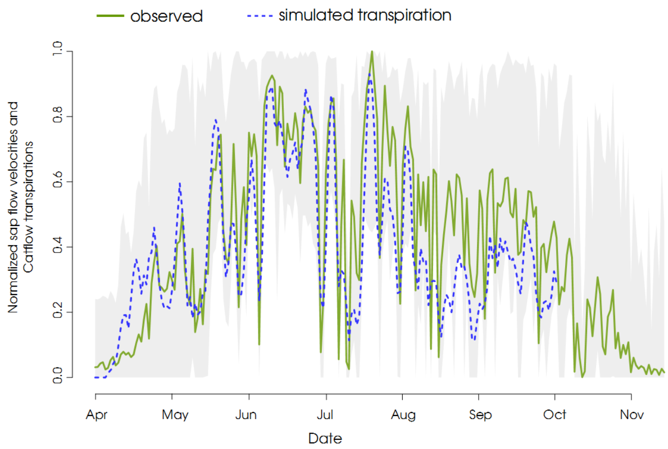


Figure 2.12: Normalized observed average sap velocities of 28 trees in the Colpach catchment (green) and normalized simulated transpiration from the Colpach model smoothed with a 3-day rolling mean (dashed blue). Additionally the ensemble of all 28 sap flow measurements is displayed in grey.

better (KGE = 0.51), while simulated and observed average dynamics are in good accordance ($r_s = 0.89$). In contrast to what we found for the Colpach, the ensemble of simulated soil moisture at 10 cm for the Wollefsbach falls into the state space spanned by the observations; it only slightly underestimates the rolling median of the observed soil moisture (Fig. 2.11 c). The predictive power is higher (KGE = 0.67) than in the Colpach, while the match of the temporal dynamics is slightly lower ($r_s = 0.81$). Again the model fails to reproduce the strong decline in soil moisture between May and July. It is, however, interesting to note that the model is nearly unbiased during August and September. This is especially interesting since the Wollefsbach model does not perform too well in simulating discharge during this time period. Simulated soil moisture at 50 cm depth shows similar deficiencies as found for the Colpach, while the predictive power was slightly smaller (KGE = 0.44), and the dynamics is also matched slightly worse ($r_s = 0.79$). When recalling the soil water retention curves (Fig. 2.7), one can infer that a soil water content of $0.2\,\mathrm{m}^3\,\mathrm{m}^{-3}$ corresponds to pF around 3.8 in the Colpach and to pF around 4.1 in the Wollefsbach. That in mind it is interesting to note that some observed soil moisture values are below this threshold throughout the entire year. This is particularly the case for soil moisture observation at 50 cm depth in the Colpach, where almost 50 % of the sensors measure water contents close to the permanent wilting point throughout the wet winter period. This also holds true for eight sensors at 10 cm depth.

### 2.4.4 *Normalized simulated transpiration versus normalized sap flow velocities*

As sap flow provides a proxy for transpiration, we compared normalized, averaged sap flow velocities of beech and oak trees to the normalized simulated transpiration of the reference hillslope model of the Colpach. The 3-day rolling mean of sap flow data stays close to zero until the end of April and starts to rise after the bud break of the observed trees. The Colpach model is able to match the bud break of the vegetation well. Furthermore, the simulated and observed transpiration fluxes and observations are in good accordance during midsummer. In the period between August and October the simulations underestimate the observations, while in April and May the simulations are too high (Fig. 2.12). Nevertheless, the model has some predictive power (KGE = 0.65), and is able to mimic the dynamics well ($r_s = 0.75$).

## 2.5 DISCUSSION

The results partly corroborate our hypothesis that single representative hillslopes might serve as parsimonious and yet structurally adequate representations of two distinctly different lower mesoscale catchments in a physically based model. The setups of the representative hillslopes were derived as close images of the available perceptual models and by drawing from a variety of field observations, literature data and expert knowledge. The hillslope models were afterwards tested against streamflow data, including a split sampling and a proxy basin test, and against soil moisture and against sap flow observations. From the fact that streamflow simulations were acceptable in both catchments when being judged solely on model efficiency criteria, one could conclude that the hillslopes portray the dominant structures and processes which control the runoff generation in both catchments well. A look beyond streamflow-based performance measures revealed, however, clear deficiencies in streamflow simulations during the summer season and during individual rainfall–runoff events as well as a mismatch in simulated soil water dynamics. In the next sections we will hence discuss the strengths and weaknesses of the representative hillslope model approach. More specifically, in Sect. 5.1 we will focus on the role of soil heterogeneity, preferential flow paths and the added value of geophysical images. In Sect. 5.2 we will discuss the consistency of both models with respect to their ability to reproduce soil moisture and transpiration dynamics. Finally, in Sect. 5.3 we discuss whether the general idea to picture and model a catchment by a single 2d representative hillslope is indeed appropriate to simulate the functioning of a lower-mesoscale catchment.

### 2.5.1 *The role of soil heterogeneity in discharge simulations*

By using an effective soil water retention curve, instead of accounting for the strong variability of soil hydraulic properties among different soil cores (Sect. 2.2.3), we neglect the stochastic heterogeneity of the soil properties controlling storage and matrix flow. This simplification is a likely reason why the model underestimates the spatial variability in soil moisture time series (compare Sect. 5.2.1). However, our approach does not perform too badly in simulating the normalized double mass curves as well as the runoff generation, at least to some extent, in both catchments. Especially during the winter, when around 80 % of the runoff is generated, runoff is reproduced acceptably well. As our models do not represent the full heterogeneity of the soil water characteristics but are still able to reproduce the runoff dynamics in winter, we reason in line with Ebel and Loague (2006) that heterogeneity of soil water retention properties is not too important for reproducing the streamflow generation in catchments. In this context it is helpful to recall the fact that hydrological models with three to four parameters are often sufficient to reproduce the streamflow of a catchment. This confirms that the dimensionality of streamflow is much smaller than one could expect given the huge heterogeneity of the retention properties. This finding has further implications for hydrological modeling approaches as it once more opens the question on the amount of information that is stored in discharge data and how much can be learned when we do hydrology backwards (Jakeman and Hornberger, 1993). Our conclusion should, however, not be misinterpreted that we claim the spatial variability of retention properties to be generally unimportant. The variability of the soil properties of course plays a key role as soon as the focus shifts from catchment-scale runoff generation to, e.g., solute transport processes, infiltration patterns or water availability for evapotranspiration.

### 2.5.2 *The role of drainage structures and macropores in discharge simulations*

By representing preferential flow paths as connected networks containing an artificial porous medium in the Richards domain, we assume that preserving the connectedness of the network is more important than the separation of rapid flow and matrix flow into different domains. The selected approach was successful in reproducing runoff generation and the water balance for the winter period in the Wollefsbach and Colpach catchments. Simulations with a disconnected network, where either the saprolite layer at the bedrock interface or the vertical macropores were removed, reduced the model performance in the Colpach model from KGE = 0.88 to KGE = 0.6 and KGE = 0.71, respectively. We hence argue that capturing the topology and connectedness of rapid flow paths is crucial for the simulation of streamflow release with representative hillslopes. We furthermore showed that a reduction in the spatial density of macropores from a 2 to 4 m spacing did not strongly alter the quality of the discharge simulations. This insensitivity can partly be explained by the fact that several configurations of the rapid flow network may lead to a similar model performance. From this insensitivity and the equifinality of the network architecture (Klaus and Zehe, 2010; Wienhöfer and Zehe, 2014) we conclude that it is not the exact position or the exact extent of the macropores which is important for the runoff response, but the bare existence of a connected rapid flow path (Jakeman and Hornberger, 1993).

However, our results also reveal limitations of the representation of rapid flow paths in CATFLOW. For instance, model setups with higher saturated hydraulic conductivities ($>10^3\,\mathrm{m\,s^{-1}}$) of the macropore medium clearly improved the model performance in the Wollefsbach but violated the fundamental assumption of Darcy's law of pure laminar flow. This was likely one reason why capturing rapid flow was much more difficult with the selected approach for the Wollefsbach. Another reason was the emergence of cracks, implying that the relative importance of rapid flow paths for runoff generation is not constant over the year, as highlighted by the findings of dye staining experiments (Fig. 2.4). Given this non-stationary configuration of the macropore network it was indispensable to use a summer and winter configuration to achieve acceptable simulations. This indicates that besides the widely discussed limitations of the different approaches to simulating macropore flow, another challenge is how to deal with emergent behavior and related non-stationary hydrological model parameters. This is in line with the work of Mendoza et al. (2015), who showed that the agility of hydrological models is often unnecessarily constrained by using static parametrizations. We are aware that the use of a separate model structure in the summer period is clearly only a quick fix, but it highlights the need for more dynamic approaches to account for varying morphological states of the soil structure during long-term simulations.

### 2.5.3 *The role of bedrock topography and water flow through the bedrock*

The Colpach model was able to simulate the double peak runoff events which are deemed typical for this hydrological landscape. However, the model did not perform satisfactorily with regard to peak volume and timing. A major issue that hampers the simulation of these runoff events is that the underlying hydrological processes are still under debate. While Martínez-Carreras et al. (2015) attribute the first peak to water from the riparian zone and the second to subsurface storm flow, other researchers (Angermann et al., 2017; Graeff et al., 2009) suggested that the first peak is caused by subsurface storm flow and the second one by release of groundwater. The representative hillslope model in its present form only allows simulation of overland flow and subsurface storm flow and not the release of groundwater because of the low permeability of the bedrock medium of $10^{-9}\,\mathrm{m\,s^{-1}}$. The deficiency of this model in reproducing double peak runoff events shows that neglecting water flow through the bedrock is possibly not appropriate (Angermann et al., 2017) and that both the perceptual model and the setup of the representative hillslope for the Colpach need to be refined. We hence suggest that the representative hillslope approach provides an option for a hypothesis-driven refinement of perceptual models, within an iterative learning cycle, until the representative hillslope reproduces the key characteristics one regards as important. The importance of bedrock topography for the interplay of water flow and storage close to the bedrock was further highlighted by the available 2d electric resistivity profiles. A model with surface-parallel bedrock topographies performed considerably worse in matching streamflow in terms of the selected performance measures and particularly did not produce the double peak events. This underlines the value of subsurface imaging for process understanding, and is a hint that the Colpach is indeed a fill-and-spill system (Tromp-Van Meerveld and McDonnell, 2006). It also shows that 2d electric resistivity profiles can be used to constrain bedrock topography in physically based models (Graeff et al., 2009), which can be of key importance for simulating subsurface storm flow (Hopp and McDonnell, 2009; Lehmann et al., 2007). Although we used constrained bedrock topography only in a straightforward, relative manner in this study, our results corroborated the added value of ERT profiles for hydrological modeling in this kind of hydrological landscape. Nevertheless, we are aware of the fact that a much more comprehensive study is needed to further detail this finding.

### 2.5.4 *Integration and use of multi-response and state variables*

*Storage behavior and soil moisture observations*

Both hillslope models reveal much clearer deficiencies with respect to soil moisture observations. While average simulated and observed soil moisture dynamics are partly in good accordance, both models are biased, except for the Wollefsbach model at 10 cm depth. In the Wollefsbach catchment this might be explained by the fact that we use a uniform soil porosity for the entire soil profile, although porosity is most likely lower at larger depths, for instance due to a higher skeleton fraction. This is no explanation for the Colpach catchment as porosity was reduced in deeper layers with respect to the skeleton

fraction. In this context it is interesting to note that quite a few of the soil moisture observations are suspiciously low, with average values of around 0.2. The resulting pF values of around 3.8 and 4.1 in the Colpach and Wollefsbach, respectively, indicate dry soils even in the wet winter period. This fact has two implications: the first is that the chosen model is almost not capable of simulating such small values, because root water uptake stops at the permanent wilting point and is small at these pF values. The second is that these sensors may have systematic measurement errors, possibly due to entrapped air between the probe and the soil. This entrapped air decreases the dielectric permittivity close to the sensor (Graeff et al., 2010), which implies that measured values will be systematically too low. From this we may conclude that the average soil moisture dynamics in both catchments might be higher and the spatial variability of soil moisture time series in turn lower, as it appears from the measurements. The obvious mismatch between the observed moisture maxima and the laboratory measurements could justify a reduction of the porosity parameter in the models, which would lead to even better fits.
In addition to the mismatch of the soil moisture simulations, the model fails in reproducing the strong decline in observed soil moisture between May and July 2014. A likely reason for this is that plant roots in the model extract water uniformly within the root zone, while this process is in fact much more variable (Hildebrandt et al., 2016).

*Simulated transpiration and sap velocities*

It is no surprise that evapotranspiration in our two research catchments is – with a share of around 50 % of the annual water balance – equally important as streamflow. It is also no surprise that evapotranspiration is dominated by transpiration, as both catchments are almost entirely covered by vegetation. However, measuring transpiration remains a difficult task, and a lack of reliable transpiration data often hinders the evaluation of hydrological models with respect to this important flux. While it is possible to calculate annual or monthly evapotranspiration sums based on the water balance, more precise information about the temporal dynamics of transpiration is difficult to obtain. Therefore we decided to evaluate our transpiration routine with available sap flow velocity data, because although the absolute values are somewhat error-prone, the dynamics are quite reliable. We tried to account for the uncertainties of the measurements by deriving a 3-day rolling median of 28 observations instead of using single sap flow velocity measurements. As we are comparing sap flow velocity to the simulated transpiration as a normalized flow, we only compare the dynamics of both variables. It is remarkable that despite the uncertainties in the sap flow velocity measurements and our ad hoc parametrization of the vegetation properties, the comparison of sap flow velocity and simulated transpiration provides additional information, which cannot be extracted from the double mass curve or discharge data. For example, based on the comparison with sap flow velocities we were able to evaluate whether the bud break of the dormant trees was specified correctly by the temperature index model of Menzel et al. (2003); this was not the case when using the default and pre-defined vegetation table of CATFLOW (not shown). Additionally, we could identify that the spring and autumn dynamics of transpiration, in April as well as in August and September, are matched poorly by the model, while the pattern corresponds well in May, June and July. We attribute this discrepancy to the lack of measured LAI values in spring and autumn and to our simple vegetation parametrization which includes several parameters like root depth or plant albedo that are held constant throughout the entire vegetation period.We are aware that this comparison of modeled transpiration with sap flow velocity is only a first, rather simple test; however, it encourages the use of sap flow measurements for hydrological modeling. It shows furthermore that the concept of a representative hillslope offers various opportunities for integrating diverse field observations and testing the model's hydrological consistency, for example evaluating it against soil water retention data and sap flow velocities.

### 2.5.5 *The concept of representative hillslope models*

The attempt to model catchment behavior using a 2d representative hillslope implies a symmetry assumption in the sense that the water balance is dominated by the interplay of hillslope parallel and vertical fluxes and the related driving gradients (Zehe et al., 2014). This assumption is corroborated by the acceptable yet seasonally dependent performance of both hillslope models with respect to matching the water balance and the hydrographs. We particularly learn that the timing of runoff events in these two catchments is predominantly controlled by the structural properties of the hillslopes. This is remarkable for the Colpach catchment, which has a size of 19.4 $km^2$, but in line with Robinson et al. (1995), who showed that catchments of up to 20 $km^2$ can still be hillslopedominated. An example of the limitations of our single hillslope approach is the deficiency of both models in capturing flashy rainfall–runoff events in the vegetation period. Besides the existence of emergent structures, these events might likely be caused by localized convective storms, probably with a strong contribution of the riparian zones (Martínez-Carreras et al., 2015) and forest roads in the Colpach catchment, and by localized overland flow in the Wollefsbach catchment (Martínez-Carreras et al., 2012). Such fingerprints of a non-uniform rainfall forcing are difficult to capture by a simulation with a spatially aggregated model, and might require an increase in model complexity. Nevertheless, we suggest that a representative hillslope model provides the right start-up for parametrization of a functional unit when setting up a fully distributed catchment model consisting of several hillslopes and an interconnecting river network. Simulations with distributed rainfall and using the same functional unit parametrization for all hillslopes would tell how the variability in response and storage behavior can be explained compared to the single hillslope. If different functional units are necessary to reproduce the variability of distributed fluxes and storage dynamics, these can for example be generated by stochastic perturbation. We further conclude that the idea of hillslope-scale functional units, which act similarly with respect to runoff generation and might hence serve as building blocks for catchment models, has been corroborated. This is particularly underpinned by the fact that the parametrization of both models was – without tuning – successfully transferred to headwaters in the same geological setting and also worked well for other hydrological years.

## 2.6 CONCLUSIONS

The exercise to picture and model the functioning of an entire catchment by using a single representative hillslope proved to be successful and instructive. The picturing approach allowed us to consider both quantitative and qualitative information in the physically based modeling process. This concept made an automated parameter calibration unnecessary and led to overall acceptable streamflow simulations in two lower-mesoscale catchments. A closer look, however, revealed limitations arising from the drawn perceptual models, the chosen hydrological model or the applicability of the concept itself. Distilling a catchment into a representative hillslope model obviously cannot reflect the entire range of the spatially distributed catchment characteristics. But as the streamflow dynamics of the catchments were simulated reasonably well and the models were even transferable to different catchments, it seems that the use of physically based models and the large heterogeneities in subsurface characteristics must not prevent meaningful simulations. Additionally, our results highlight the importance of considering non-stationarity of catchment properties in hydrological models on seasonal timescales and emphasize once more the value of multiresponse model evaluation. A representative hillslope model for a catchment is, hence, perhaps less accurate than a fully distributed model, but in turn also requires considerably less data and reduced efforts for setup and computation. Therefore, this approach provides a convenient means to test different perceptual models, and it can serve as a starting point for increasing model complexity through a combination of different hillslopes and a river network to model a catchment in a more distributed manner.

Part III

# ON THE DYNAMIC NATURE OF HYDROLOGICAL SIMILARITY

This study is published in the scientific journal HESS. It is part of the special issue "Thermodynamics and optimality in the Earth system and its subsystems"; a joint special issue between the scientific journal Earth system dynamics (ESD) and HESS. The remainder of part III is a reprint of:

*Loritz, R., Gupta, H., Jackisch, C., Westhoff, M., Kleidon, A., Ehret, U., and Zehe, E. (2018). On the dynamic nature of hydrological similarity. Hydrology and Earth System Sciences, 22(7), 3663-3684. doi:* 10.5194/hess-22-3663-2018

# 3 ON THE DYNAMIC NATURE OF HYDROLOGICAL SIMILARITY


## ABSTRACT

The increasing diversity and resolution of spatially distributed data on terrestrial systems greatly enhances the potential of hydrological modeling. Optimal and parsimonious use of these data sources requires, however, that we better understand a) which system characteristics exert primary controls on hydrological dynamics and b) to what level of detail do those characteristics need to be represented in a model.

In this study we develop and test an approach to explore these questions that draws upon information theoretic and thermodynamic reasoning, using spatially distributed topographic information as a straightforward example. Specifically, we subdivide a meso-scale catchment into 105 hillslopes and represent each by a two dimensional numerical hillslope model. These hillslope models differ exclusively with respect to topography related parameters derived from a digital elevation model; the remaining setup and meteorological forcing for each are identical. We analyze the degree of similarity of simulated discharge and storage among the hillslopes as a function of time by examining the Shannon information entropy. We furthermore derive a ”compressed” catchment model by clustering the hillslope models into functional groups of similar runoff generation using normalized mutual information as a distance measure.

Our results reveal that, within our given model environment, only a portion of the entire amount of topographic information stored within a digital elevation model is relevant for the simulation of distributed runoff and storage dynamics. This manifests through a possible compression of the model ensemble from the entire set of 105 hillslopes to only 6 hillslopes, each representing a different ”functional group”, which leads to no substantial loss in model performance. Importantly, we find that the concept of hydrological similarity is not necessarily time-invariant. On the contrary, the Shannon entropy as measure for diversity in the simulation ensemble shows a distinct annual pattern, with periods of highly redundant simulations, reflecting coherent and organized dynamics, and periods where hillslopes operate in distinctly different ways.

We conclude that the proposed approach provides a powerful frame-

work for understanding and diagnosing how and when process organization and functional similarity of hydrological systems emerges in time. Our approach is neither restricted to the model, nor to model targets or the data source we selected in this study. Overall, we propose that the concepts of hydrological systems acting similarly (and thus giving rise to redundancy) or displaying unique functionality (and thus being irreplaceable) are not mutually exclusive. They are in fact of complementary nature, and systems operate by gradually changing to different levels of organization in time.

## 3.1 INTRODUCTION

### 3.1.1 *Motivation*

This paper addresses the question "How important is spatial variability of terrestrial system characteristics and meteorological forcing when viewed from the perspective of stream flow generation and distributed water storage?" While this question has motivated hydrologists since the early days of our science, it gained substantial attention with the development of distributed hydrological models, and it seems fair to say that attempts to address the question still lie at the heart of every distributed model application (e.g. Beven, 1989; Freeze and Harlan, 1969; Hrachowitz and Clark, 2017; Refsgaard, 1997). Needless to say, this question has not found easy answers. Besides the lack of sufficient process understanding (in part due to the difficulty of gathering relevant data about hydrologic systems), there is also the uncertainty we unavoidably encounter when dealing with the steadily growing and changing pool of geo-information (Musa et al., 2015). For instance land surface digital elevation information is now available at a resolution of 25 m globally (Farr et al., 2007). Similarly, weather radar coverage is available for large parts of Europe, providing accumulated 15 min precipitation estimates at 4 km resolution (Huuskonen et al., 2014). Despite the huge potential for model improvement provided by these new and diverse pools of information, a danger associated with their use is that we can "miss the forest for the trees" unless we are able to determine which information contained in the data is of relevance to the questions we seek to answer. We therefore now face the problem of how to discriminate important details about the hydrological landscapes from idiosyncratic ones, and hence must deal with the challenge of how to identify which characteristics explain hydrological similarity (Blöschl and Sivapalan, 1995). This study is largely motivated by the "power" view introduced by Wagener and Gupta (2005) which advocates "a need to develop better methods for characterizing and extracting relevant information from data" (see also Gupta and Nearing, 2014). Our specific objective is to propose an approach addressing this issue, by drawing upon an information theoretic perspective to extract and quantify the relevant information for spatially distributed hydrological modeling, and by using thermodynamic reasoning to explain why only a portion of the full information content available in the data is relevant.

### 3.1.2 *Background*

From a thermodynamic perspective, streamflow generation is driven by differences in potential energy between the upslope catchment areas and the stream channel. The majority of this available energy is dissipated during runoff concentration and infiltration, while the remaining part is exported from the catchment as the kinetic energy of streamflow (Kleidon and Renner, 2013). These potential energy differences depend largely on catchment topography, and on the space-time patterns of precipitation (Zehe et al., 2013). Accordingly, we might be naturally drawn to expect that large spatial variations in both characteristics will result in large spatial variations in runoff generation. However, when exactly should spatial variation be considered "large" enough that we need to explicitly account for it?

In the context of spatially distributed rainfall, this latter question has received considerable attention (e.g. Arnaud et al., 2002; Das et al., 2008; Obled et al., 1994; Tetzlaff et al., 2005; Zehe et al., 2005). In general, the predominant view that seems to emerge from these studies is that the impact (on runoff simulations) of spatial distribution in rainfall increases with size of the area considered. This is often traced back to the growing importance of flood routing, in combination with the average spatial extent of typical rain storms (e.g. Lobligeois et al., 2014; Smith et al., 2004). Nevertheless, no consensus has yet emerged as to whether this statement is generally valid, and no guidelines exist regarding under which conditions the use of information regarding the spatially distributed nature of rainfall becomes inevitable (Emmanuel et al., 2015). Similarly, the question of how strongly the spatial resolution of a DEM affects the results of a distributed model application has been investigated in various studies (e.g. Schoorl et al., 2000; Sørensen et al., 2006; Thompson et al., 2001). For instance Zhang and Montgomery (1994) varied the resolution of their DEM and reported that spatial resolutions finer than 10 m did not result in significant improvements to the simulation results of their hydrological model. Chaubey et al. (2005) tested the influence of DEM spatial resolution on simulation results of the Soil Water and Assessment Tool (SWAT) and reported that grid size has a significant influence on different watershed responses, as well as on the sub-basin classification implemented in SWAT. However, as with the case of distributed rainfall, the results of these studies do not point to a generic approach, nor to any general conclusions regarding the importance of DEM-resolution for distributed hydrological modeling.

Overall, this lack of a coherent image certainly reflects the varying sensitivities of different model structures (Das et al., 2008), the dependence on scope and scale of the model exercise (Blöschl and Sivapalan, 1995) and on differences among hydrological landscapes (Beven, 2000). It seems, therefore, that an investigation of the role of distributed information in hydrological modeling may benefit from a more generic and systematic approach, one that may be generalized to different spatially distributed data sources and models, and that is able to cope with interactions among them in a straightforward manner. In contrast to much of the aforementioned work, which has relied primarily on statistical methods, the purpose of the work reported here is to investigate the extent to which information theory (Cover and Thomas, 2005) is able to provide instructive measures that are suitable for this purpose.

More specifically the main objective of this study is to present and test an approach to quantify the relevance of spatially distributed data sources for hydrological simulations drawing from information theory. We exemplify this approach using catchment topography as distributed information source as well as stream flow and soil water storage as modeling targets, however, the general mindset of the approach is applicable to any distributed information source such as spatially distributed rainfall or geology as well as to a wide range of arbitrary model target and different distributed models.

### 3.1.3 *The role of surface topography in hydrological modeling*

Despite the fact that DEM's provide the basis for identifying watershed boundaries, river networks and potential energy differences in the landscape, several studies have concluded that topography alone is a weak descriptor for inferring similarity in hydrological behavior. For instance, Zehe et al. (2005) showed that the topographic wetness index Beven and Kirkby (1979), a popular topographic similarity measure, failed to explain soil moisture variability and similarity in runoff generation in a lower mesoscale catchment. Fenicia et al. (2016) and Jackisch (2015) showed that topography alone might be a poor guide for subdividing a 256 km$^2$ catchment into different functional units, and questioned the explanatory power of the topography in this respect. Our own work, Loritz et al. (2017), has shown that an ”effective” representation of two different catchments by a single representative hillslope was able to provide successful simulations of their inter-annual runoff responses and annual storage dynamics. Together, these findings suggest that an informationally ”compressed” representation of the topographic map may be able to preserve the relevant information regarding geopotential differences that drive runoff generation.

In line with these findings, we therefore pose the hypothesis that ”although a highly-resolved DEM contains a large amount of information about topography, not all of this spatially distributed information is relevant for the generation of hydrological predictions”. Following Weijs et al. (2013a), it seems reasonable that information theory may provide a natural framework for dealing with such compression of information in hydrologic science. The term ”compression” was originally coined by Claude Shannon to refer to the quantification, storage and communication of information (Shannon, 1948). In environmental science, information-theoretic concepts such as the "Shannon entropy" have found widespread use in various applications (e.g. (Brunsell, 2010; Weijs et al., 2013b; Yakirevich et al., 2013), ranging from uncertainty assessment in 3-D geological models (Schweizer et al., 2017) to the delineation of water resource zones in Japan (Kawachi et al., 2001). For an introduction to, and detailed review of, information theoretic concepts we refer the reader to Cover and Thomas (2005), Singh (2013), and Weijs et al. (2013a).

With respect to the above finding it is important to note that compressibility relates to order or organization (Davies, 1990). The identification of relevant information within distributed system characteristics is therefore closely linked to the identification of spatial organization and thus with the identification of hydrological similar functioning areas (Sivapalan, 2005). As pointed out by Zehe et al. (2014), these "functional units" may be straightforwardly defined in thermodynamic terms as any flux is driven by a specific gradient while it performs work against a specific flow resistance. Similarity of both the relevant drivers and the resistance terms is a sufficient criterion to expect that two systems behave similarly with respect to the generation of a flow, and with regard to the associated entropy production. If we transfer this concept to runoff generation, differences in the geopotential (topography) act as driver since runoff is driven by gravity. The resistance term, on the other hand depends either on surface roughness (and thus for instance on the vegetation in case of overland flow), on the pattern of subsurface conductance, apparent preferential pathways and in case of matrix flow on the capacity of the system to store water. Yet, the gradient flux-resistance relation is non-unique, because a twice as large driver in combination with a twice as large resistance results in exactly the same flux. It is this non-uniqueness, which explains why two hillslopes with distinctly different topographies may still produce the same runoff when these differences are compensated by their associate resistances.

However, while a physical explanation of the phenomena "landscape organization" is crucial to our understanding, for practical modeling applications we need to step beyond that and actually identify these functional units in the landscape. One avenue is surely to detect these gradients and resistance terms directly based on the available landscape characteristics (Seibert et al., 2017). However, it is often difficult to know a-priori which characteristics dominate the function of a landscape element (Oudin et al., 2010). Another approach is, hence, to identify functional units a-posteriori directly based on their function, and to subsequently identify which characteristics dominate the hydrological processes, and at which scale (Sivapalan, 2003). It is exactly here that an information theoretic perspective might be particularly valuable as, despite the more qualitative and descriptive nature of the concept of landscape organization, compressibility is actually quantifiable. For instance two hillslopes showing a similar function with respect to a given process can be compressed and hence combined into a larger landscape element without losing information about the spatial distribution of processes in a catchment. The identification of functional similar areas is hence directly connected to both statistical physics (organization) and information theory (compressibility). For this reason we believe that concepts such as maximum (Shannon)-entropy (Jaynes, 1957) and information theoretic variables like the "mutual information" and "Kullback-Leibler divergence" (Cover and Thomas, 2005; Weijs et al., 2013a; Weijs and Giesen, 2013) provide an excellent framework for connecting the generic informational concepts of statistical inference and compression of data with the specific domain concepts of landscape organization and hydrological similarity.

### 3.1.4 *Objectives and scope*

The main objective of this study is to propose and test a generic approach, based on information theory and to quantifying the relevance and value of spatially distributed data sources for hydrological simulations. Our approach is developed and tested using catchment topography as the source of spatially distributed information, and stream flow and soil water storage as the modeling targets. Specifically, we subdivide a 19.4 $km^2$ catchment into 105 hillslopes and represent each of these contributing spatial units with a hydrological hillslope model. Following Loritz et al. (2017), the hillslope models are identically parametrized with respect to soils, bedrock topography and vegetation, and differ only with respect to the values of their topography dependent parameters such as aspect, slope and elevation above and distance to the river. Each of these hillslope models is driven by the same meteorological forcing for one hydrological year yielding 105 independent runoff and storage time series. In the first part of this manuscript we analyze the distributions of runoff and storage simulations at each time step by means of the Shannon information entropy. With this approach we are able to reveal different levels of redundancy in our simulated output in time and try to answer the question whether we can identify the necessary spatial complexity of our chosen model structure. In the second part of this manuscript we evaluate the similarities of the runoff time series simulated by the hillslope models in terms of their mutual information. We use this as a basis for compressing them into a smaller set of functional groups, such that in each group the members are to a certain extent predictable from each other in terms of runoff generation. Here we choose the average Shannon entropy of the simulation period to determine the number of different functional groups. Based on this we construct different time invariant realizations of a compressed catchment model and test those against observations and the simulation with the uncompressed model. Finally, we reiterate that the overall approach presented here is applicable to a variety of different spatially distributed information such as spatially distributed rainfall or land-use, as well as to most modeling target and to a wide range of spatially distributed hydrological models available. This paper is, however, restricted to development and testing of the approach using only catchment topography and one numerical hillslope model.

## 3.2 STUDY AREA AND MODEL REALIZATIONS

In this section we introduce the study area, the database used, and the general model setup of the different hillslopes.

### 3.2.1 *The Colpach catchment*

The 19.4 $km^2$ Colpach catchment is situated in the northern part of the Attert basin in the Devonian schists of the Ardennes massif, and has an elevation ranging from 265 to 512 m a.s.l. (Figure 3.1 a). Approximately 65 % of the catchment is forested, mainly on the steep hillslopes. In contrast, the plateaus at the hilltops are predominantly used for agriculture and pasture. The dominant runoff process is rapid flow in a highly permeable saprolite layer above the bedrock, and the catchment is characterized as a fill-and-spill system (Wrede et al., 2015). Besides the importance of lateral flow along the bedrock, several irrigation and dye staining experiments have highlighted the role of vertical structures for infiltration and subsequently for subsurface runoff formation (Jackisch et al., 2017). For a more detailed description please see Loritz et al. (2017) and Wrede et al. (2015) and Jackisch (2015).

### 3.2.2 *The CATFLOW model*

The spatially-distributed hillslope-scale model CATFLOW (Maurer, 1997; Zehe and Flühler, 2001) is based on the subdivision of a catchment into several hillslopes connected by a drainage network. Each hillslope is discretized along a 2-dimensional cross section using curvilinear orthogonal coordinates. Each surface model element extends over the width of the hillslope, and these widths may vary along the hillslope. Evapotranspiration is represented using an advanced SVAT approach based on the Penman-Monteith equation, which accounts for tabulated vegetation dynamics, albedo as a function of soil moisture, and the impact of local topography on wind speed and radiation. Soil water dynamics and solute transport are simulated based on the mixed form of the Darcy-Richards equation, solved using mass conservative Picard iteration and adaptive time stepping (Celia et al., 1990). The hillslope module is designed to simulate infiltration excess runoff, saturation excess runoff, re-infiltration of surface runoff, lateral water flow in the subsurface, return flow and solute transport.

### 3.2.3 *Hillslope setup, forcing and model evaluation*

The topographic analysis was based on a 5 m Lidar digital elevation model, aggregated and smoothed to 10 m resolution. GRASS GIS (Neteler et al., 2012) was used to subdivide the catchment into 105 hillslopes (Figure 3.1 a) using a classical hydrological terrain analysis algorithm r.watershed. This approach generates a stream network after the user sets a threshold for the minimum size of an exterior watershed basin. We identified this value by varying this threshold across a range of values trying to reproduce the official stream network which was available from the Luxembourg Institute of Technology (LIST) by visual inspection. Following the standard procedure of r.watershed each stream segment has two corresponding hillslopes (left and right side of the stream). We use the landscape units mapping program (LUMP; Francke et al., 2008) and again GRASS GIS to derive the hillslope profiles, including properties such as the elevation and distance to the river, and the mean aspect and width function of each hillslope (Figure 3.1 b). On average the hillslopes lie 67 m above the river, are 446 m wide, and cover an area of 0.16 km$^2$. The maximum area of a hillslope is 0.86 km$^2$ while the smallest hillslope covers an area of 0.12 km$^2$.

With respect to soils, bedrock topography and vegetation, the 105 hillslope models were identically parameterized using a parameter set, macropore distribution and subsurface stratification tested and derived by Loritz et al. (2017) when representing the entire Colpach catchment by a single effective hillslope model. Accordingly the hillslopes differ only in the values of parameters that are extracted from the digital elevation model (hillslope profile and length, width and aspect). All hillslope models are 2 m deep, where the upper 1 m is classified as soil followed by a 0.2 m lateral saprolite layer and an 0.8 m deep almost impermeable bedrock (see soil parameter and structure in Tab. 1 in Loritz et al. (2017)). The porosity of the upper 1 m of soil is assumed to reduce linearly with depth, with the lowest value being 0.3 at a depth of one meter from the surface. In order to account for reported preferential flow in this area (Jackisch et al., 2017) we added additionally, every 4 m, a 0.1 m wide rapid flow path (vertical flow structure) with an depth of 1 m. The entire soil setup follows the findings of Loritz et al. (2017) in which it was shown that a representative hillslope was able to provide successful simulations of various hydrological fluxes. The discretization of the hillslope in the downslope direction varies between a maximum of 1 m and minimum of 0.1 m, where the latter occurs close to rapid flow paths. The vertical grid size was set to 0.1 m, with a reduced vertical grid size of the top node of 0.05 m (Figure 3.1 c).

Boundary conditions were set to an atmospheric boundary at the top, no flow boundary conditions at the upslope, and a gravitational flow boundary condition at the lower boundary. At the hill foot of the hillslope we selected a seepage interface for the upper 0.4 m, where outflow only occurs under saturated and no flow under unsaturated conditions. For the lower 1.6 m of the downslope boundary we selected a no flow boundary to mimic a saturated zone close to the river. All of the hillslopes are covered entirely by forest and the evapotranspiration routine is parameterized similarly to the one described in detail in Loritz et al. (2017). Figure 3.1 c shows an example of a typical CATFLOW hillslope grid and soil setup divided into soil, rapid flow paths and bedrock.

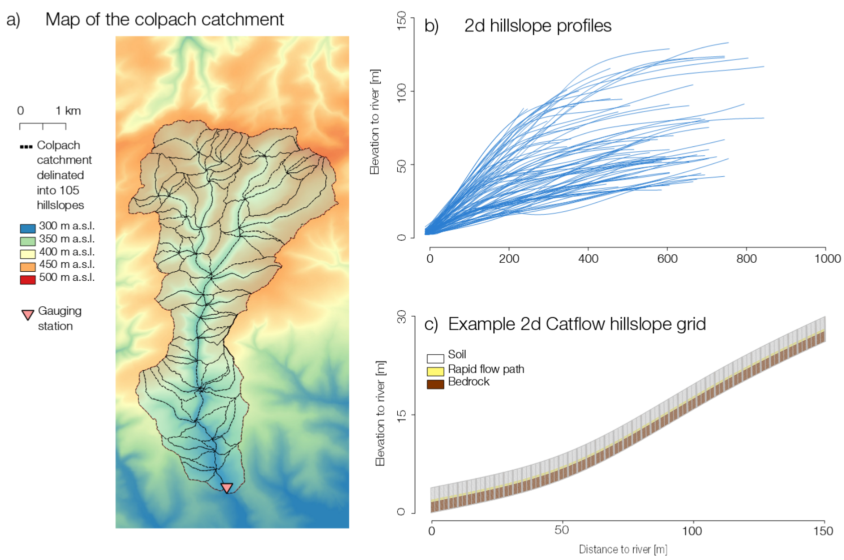


Figure 3.1: a) Digital elevation model of the Colpach catchment and its delineation into 105 hillslopes b) all hillslope profiles extracted using the LUMP approach c) example of a CATFLOW hillslope grid.

*Model forcing and application*

Meteorological input data are recorded at an official meteorological station (Roodt), and were provided by the "Administration des Services Techniques de l'Agriculture Luxembourg". All hillslope models were forced with identical meteorological inputs. This implies, for instance, that we neglect observed variations of rainfall and wind speed within the catchment. We compared simulated and observed specific runoff by dividing the respective values by the relevant contributing areas; i.e., either by the area of the hillslope or of the Colpach catchment. Similarly, we calculated the area specific water storage (average water content per $m^2$) for each hillslope. The simulation period is the hydrological year 2014 from October 2013 to October 2014. This is preceded by a model spin-up of one year with initial states of 70 % saturation.

*Model evaluation*

The intention of the model evaluation performed here was not to infer whether we have identified the best performing model structure, but to evaluate and quantify differences in modelled runoff and storage arising from underlying differences in hillslope topography. Therefore, while this exercise does not require a comparison to observations, we nevertheless do so to demonstrate that the different models (and in particular the entire ensemble) produces meaningful simulations that are consistent with observed hydrological storage and streamflow dynamics. We inspected the runoff simulations both visually and by comparison to the observed specific discharge using the normalized mutual information (NMI, specified below; see also Michaels et al., 1998). In addition, we use the Kling–Gupta efficiency (KGE, Gupta et al., 2009) to highlight that the NMI provides a consistent picture and is able to identify differences between hydrographs. Furthermore, we use the NMI in our functional classification because it is symmetric and satisfies the mathematical requirements of a distance metric (see section 2.6; for a further comparison of the NMI as well as the Appendix C). Additionally, we calculated the KGE and NMI between the area weighted median of the runoff simulations and the observed specific discharge of the catchment. By simply using the area weighted median instead of a river network routing scheme we assume, in line with Robinson et al. (1995) and our own findings (Loritz et al., 2017), that the Colpach catchment is hillslope dominated and that the timing of the routing is small enough to be neglected. With respect to the storage dynamics, we estimated the average amount of water within the hillslope (in mm for each hillslope) and compared these values against the median of storage estimates calculated from available soil moisture measurements in 10, 30 and 50 cm, which have been collected at different locations throughout the catchment (for detailed information of the soil moisture sensors and observations please see Loritz et al. (2017)). As the model and the observations estimates are based at largely different scales, we believe that any comparison more detailed than the comparison of their temporal dynamics is in-appropriate.

## 3.3 THEORETICAL BACKGROUND, APPROACH AND METHODS

In the following section we provide a detailed review of the important concepts from information theory, and discuss how we used these concepts to address the study objectives.

### 3.3.1 *Information theory and Shannon information entropy*

The field of Information theory originally developed within the context of communication engineering, deals with the quantification of information with respect to a concept called "surprise" (Applebaum, 1996). For a discrete random variable $X$ that can take on several values $X \in x_1; x_2; x_3 ... x_i$ with associated prior probabilities $p(x_1); p(x_2); p(x_3) ... p(x_i)$ the surprise or information content of receiving/observing a specific value $X = x_i$ is defined as:

$$I = -log_k(p(x_i)) \tag{3.1}$$

where I is the information content, $k$ is the base of the logarithm and $p(x_i)$ the prior probability that $X$ can exist in the state $x$. The logarithm in this definition assures that information is an additive quantity. When the base $k$ of the logarithm is chosen to be 2, information is measured in "bits" (abbreviated from binary digit). While different $k$ values can be used to calculate the information content of a random discrete variable, here we stick with the logarithm to the base 2.
To calculate the average information content associated with the random variable $X$ we can estimate the Shannon entropy $H(X)$ defined (by taking its expectation) as:

$$H(X) = -\sum_{x \in X} p(x_i)\, log_2(p(x_i)) \tag{3.2}$$

where $p(x_i)$ is again the probability that $X$ can be in the state $x$. In this study we computed the Shannon entropy of the probability distribution of the 105 runoff and storage simulations for each hourly time step. In addition to computing the Shannon entropy for a single random variable (also called self-information), we compute the joint entropy $H(X, Y)$ of a set of variables $X$ and $Y$ as follows:

$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x_i, y_i)\, log_2(p(x_i, y_i)) \tag{3.3}$$

where $p(x_i, y_i)$ is the joint probability. The joint entropy is used to estimate the mutual information (described below) between two random variables. For more detailed discussion of information theoretic concepts and variables please see Applebaum (1996) and Cover and Thomas (2005).

### 3.3.2 *The appropriate binning for estimating discrete probability density functions*

A crucial step in the computation of Shannon entropy and/or mutual information of discrete distribution (see section 3.1 and 3.2) is a careful choice of the bin widths used to construct the probability density functions (pdf; Gong et al., 2014; Pechlivanidis et al., 2016). Various guidelines are available regarding how to properly estimate the bin width from the viewpoint of statistical rigor (e.g. Scott, 1979). However, Weijs et al. (2013a) also point out that the bin width for a pdf should always be chosen based on considerations related to the question one wishes to answer. For instance, hydrologists often evaluate their models against measured soil moisture or discharge data. As such observations always imply the existence of measurement errors, observational differences smaller than the typical size of such errors should not be afforded physically meaningful importance. (To infer on the sensitivity of the Shannon entropy to different bin width please see the appendix B).

Accordingly, for calculation of the entropy of the runoff and the storage simulations we propose that the smallest meaningful bin width should be greater than or equal to the measurement error. Consequently, we choose the mean relative error of the rating curve (8.5 %, see appendix A) to estimate the Shannon entropy of the runoff simulations and the measurement error of the installed capacitive soil moisture probe soil moisture probes of 1 Vol. % for the storage simulations (Decagon 5TE; ± 1 - 2 % volumetric water content for calibrated soils; manufacture information). For the runoff simulations, we started with a bin width of 0.01 mm and then progressively increased the bin width by a factor of 8.5 %. This results in a non-uniform bin width distribution with constantly increasing bin sizes for larger discharge values as the uncertainty in the measurements increases with higher flows. In contrast, for the storage simulations, we used a constant bin width of 10 mm because the measurement errors of our soil moisture probes do not depend on the magnitude of the measured value. We transferred the error of the soil moisture probes to our storage simulations as follows. The 1 m thick soil domain has a porosity of 0.57 $m^3 m^{-3}$, having a total storage volume of 570 mm. We hence use a constant bin width of 10 mm, corresponding to 1 % vol, with bins ranging from 10 mm (1 % vol) to 570 mm (57 % vol)

### 3.3.3 *Upper and lower boundary of the Shannon entropy – perfect versus no organization*

Isolated systems evolve, according to the second law of thermodynamics, to a state of maximum entropy in which all gradients are depleted and each microstate of the system is equally likely (Kondepudi and Prigogine, 1998). This implies maximum uncertainty about the microstate and the absence of any organization/order in the system. Jaynes (1957) transferred this fundamental insight into a method of statistical inference, stating "when making inferences based on incomplete information, the best estimate for the probabilities is the distribution that is consistent with all information, but maximizes uncertainty". This condition is reflected by a uniform distribution where all outcomes are equally likely (Weijs et al., 2010). With respect to our model ensemble, a state of maximum entropy implies that each of the 105 hillslopes models produces a unique output that cannot be guessed given knowledge regarding the output of any other hillslope. Accordingly, we can calculate the theoretic maximum entropy for our model as:

$$H_{max} = log_2(N) \tag{3.4}$$

where $N = 105$ is the number of hillslope models. This maximum reflects a theoretical state of zero spatial organization in the catchment, where each hillslope provides a unique contribution to stream flow and storage dynamics due to its specific. A further compression of the catchment subdivision, for instance by leaving out or merging certain hillslopes, is not possible without losing precision. At the other end of the spectrum, one may have a state of perfect spatial organization in which all 105 hillslope models are within the error margin of observations perfectly predictable from each other. This would correspond to zero entropy and implies that the compression of the spatially distributed model is trivial as any arbitrarily selected hillslope will represent it equally well.

It is important to note that $H_{max}$ is (in our virtual experiment) a theoretical upper limit as the hillslope models would, given our bin width, need to simulate discharge values as high as 48.3 mm $hr^{-1}$ to reach this theoretical limit. We thus distinguish between the maximum entropy of our model ensemble given the spatial discretization of the model and the maximum entropy of our experiment given the uncertainties and physical limits of our discharge and storage simulations and observations. The difference becomes clear if one imagines a simple thought experiment in which one would like to study a dice with six possible outcomes. The maximum entropy of this dice is linked to the number of possible states of the "system" and hence is $log_2(6) = 2.58$. Now depending on our investigation, we might change our question and only ask for values larger or smaller than 3. In this case the maximum entropy of our "experiment" would, with two possible outcomes, be $log_2(2) = 1$.

### 3.3.4 *Mutual information as similarity measure*

To compare simulated runoff time series generated by different hillslopes, we calculate their pair-wise mutual information of each simulated runoff time series as a similarity measure. Mutual information $I(X, Y)$ between two discrete random variables $X$ and $Y$ is a measure of the strength of their informational correspondence, defined by Cover and Thomas (2005) as:

$$I(X, Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y)\, log_2 \frac{p(x, y)}{p(x)p(y)} \tag{3.5}$$

where $p(x, y)$ is the joint probability of $X$ and $Y$ and $p(x)$ and $p(y)$ are their marginal probabilities. Equivalently, mutual information can also be calculated directly as a difference between the sum of the entropies of $X$ and $Y$ minus the joint entropy of $X$ and $Y$ (Figure 3.2).

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \tag{3.6}$$

While Shannon entropy is used to determine the information redundancy or compressibility between the 105 simulated discharge time series at a certain time steps, we now show how mutual information can be used to see how similar or dissimilar two discharge simulations are.

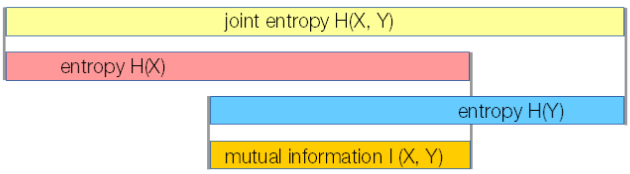


Figure 3.2: Sketch of the relation between information entropy, joint entropy and mutual information displayed as bar diagram.

Mutual information quantifies the amount of information that one variable reveals about another and thus the strength of their co-dependence. If the mutual information is zero, the two variables are independent while larger values correspond to stronger relationships. When using the binary logarithm mutual information, Shannon entropy and joint entropy share the same unit ”bits”. Here, we seek to use the mutual information between different hillslope runoff simulations as a measure of similarity or distance between the hillslope models. However, since the value of mutual information depends on the absolute magnitude of joint entropy between the two chosen variables, it is not appropriate to use mutual information directly as a distance function for relative comparisons (if the joint entropy of two variables is low the value of mutual information will also be low even if the two variables are perfectly related). Hence, following Michaels et al. (1998), we normalize $I(X, Y)$ using the larger of the entropies of the two random variables $X$ and $Y$. It is important to note that this normalization can also be done using the smaller of the entropies of the two random variables $X$ and $Y$ or the joint entropy of $X$ and $Y$. Depending on the objective this can be an important choice (see appendix C). In this study we follow the avenue recommended by Michaels et al. (1998) and use the maximum.

$$NMI(X,Y) = \frac{I(X,Y)}{max[H(X), H(Y)]} \tag{3.7}$$

Accordingly, the normalized mutual information $(NMI)$ ranges from 0 to 1, with higher values corresponding to stronger relationships (higher mutual information content). Further, to make the $NMI$ easier interpretable we subtract the $NMI$ from 1 as typical distance functions are normally closer to zero in case of a stronger similarity (see Appendix C for a comparison of the $NMI$ with the Pearson correlation coefficient and the Euclidean distance).

### 3.3.5 *Functional classification of hillslopes with similar runoff behavior*

Using $NMI$ as distance metric, we classified the 105 hillslope models into functional groups of similar runoff behavior based on the 105 runoff time series, using a hierarchical cluster analysis based on Ward's minimum variance method (Hastie et al., 2009; Murtagh and Legendre, 2014). As a first guess of a physically meaningful number of functional groups we used the mean annual entropy of all 105 discharge simulations (further discussed in section 4.2).

$$No. of\ functional\ groups = 2^{mean\ annual\ entropy} \tag{3.8}$$

This choice is inspired by the fact that the Shannon entropy of a random variable $X$ is closely related to the maximum compressibility of the information about this variable. This is because, when the Shannon entropy is calculated using the binary logarithm, it relates to the minimum number of binary "yes or no questions' necessary to determine the actual value of $x_i$ from $X$. In the special case where the distribution of the random variable is dyadic, the value of the Shannon entropy $H(X)$ and the expected minimum number of questions are equivalent, while if this is not the case the expected number of questions lies between the computed value of the entropy $H$ and its increment $H + 1$ (for further details see Cover and Thomas, 2005).

$$H(X) \leq Expected\ Questions < H(X) + 1 \tag{3.9}$$

So, in general, if the entropy of a discrete random variable $X$ is $H(X) = 2$, we know that the expected number of binary (Yes/No) questions needed to quantify $x$ lies between 2 and 3. This implies that the number of possible outcomes lies somewhere between $2^2 = 4$ and $2^3 = 8$, as every binary question can have two possible answers.

### 3.3.6 *Compression of the catchment model based on functional groups*

Having grouped the hillslope models into time-invariant functionally similar groups, we test whether this grouping provides a solid basis to compress the model structure of 105 hillslopes into a less redundant one that yet produces results of similar quality as the full set of hillslopes but at much smaller computational cost. There are at least three avenues to do so. The first one is to simply calculate the area weighted median or average of all runoff simulations within a functional group. This, however, means that all 105 runoff simulations are necessary to build this compressed model and we cannot run the compressed model in a forward mode. The second avenue is to take functionally united hillslopes and derive for each functional unit an effective, spatially aggregated hillslope in a similar fashion as done in Loritz et al. (2017). Though this is most likely the most promising way to come up with a compressed catchment model, it is beyond the scope of this manuscript. Instead, to simplify this attempt in this study we use a third option and develop a compressed model structure using a bootstrap-like approach. For this we randomly select a single hillslope from each functional group, and calculate the area weighted median of the simulated discharge time series of the six randomly selected hillslope models (Compressed catchment model; Figure 3.3). The weight assigned to each of the selected discharge time series corresponds to the areal fraction of all hillslopes in the respective functional group. This assures mass conservation because runoff of each hillslope is equal to its area times the simulated specific discharge. We use random selection because each group member is regarded as equivalent to represent the runoff generation of the corresponding functional group. To account for sampling variability, as simulated runoff differs slightly among the hillslopes within a functional group, we repeat this random selection 1000 times. In a final step, we compare those values individually as well as the median of all realizations against the observed runoff of the Colpach using the KGE. This reveals the performance spread of the randomly generated compressed models compare to the area-weighted median of the entire 105 hillslopes.

## 3.4 RESULTS

### 3.4.1 *Runoff and storage simulations*

The overall model performance of the area weighted median of all hillslopes is decent, with a KGE of 0.76. The ability of different

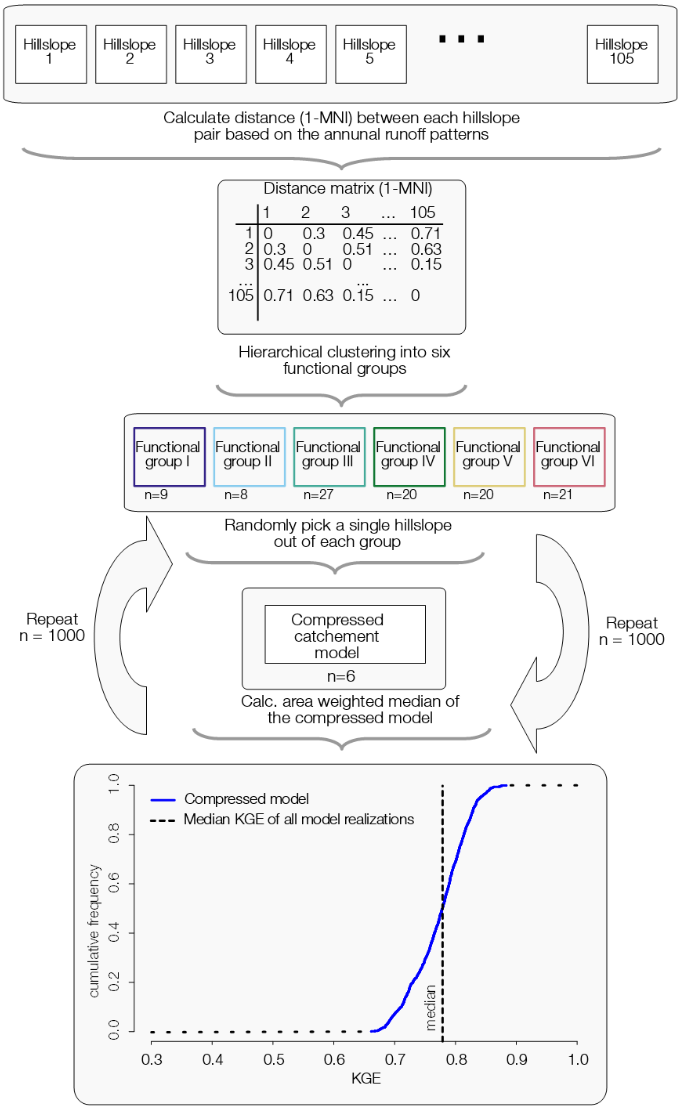


Figure 3.3: Sketch of the approach for compression and performance evaluation for the compressed catchment models.

hillslope models to reproduce the observed runoff dynamics of the Colpach catchment varies substantially (see Figure 3.4 a), with KGE values ranging between 0.44 and 0.92. This apparent spread in model performance among the hillslopes corroborates the sensitivity of simulated discharge to those parameters derived from the DEM. A similar pattern is revealed when model "goodness" is expressed by means of the normalized mutual information (NMI) between each hillslope model and the observed runoff. NMI values range from 0.51 to 0.71 and show a strong linear correlation to the corresponding KGE values (with a Pearson correlation coefficient of 0.89). This good correspondence of NMI with the KGE performance measure reinforces the notion that NMI is a suitable measure of similarity, or difference, between time series of hydrological variables.

The temporal patterns of total area specific storage for each hillslope model are shown in Figure 3.4 b. The skill of different hillslopes to reproduce the temporal dynamics of observed median storage is rather stable, with a Spearman rank correlation coefficient ranging from 0.77 to 0.86, with the ensemble median having a value of 0.82. Visual comparison of the simulated storage time series reveals that differences in hillslope topography result mainly in a parallel shift of the respective time series. This parallel spreading is stronger during the wet season and less pronounced during dry conditions. The latter might be due to the identical vegetation parameterization of each hillslope and hence a result of highly similar root water uptake which dominates storage dynamics during dry conditions in summer.

### 3.4.2 *Entropy of the model simulations*

If all 105 of the hillslope models were to produce unique simulations of equal importance, their entropy would be the theoretical maximum value of $log_2(105) = 6.7$. However, in our study the maximum entropy of our discharge simulations given the chosen binning size and the maximum simulated discharge value of 0.75 mm hr$^{-1}$ is $log_2(54) = 5.7$ and for the storage simulation given a minimum simulated soil moisture close to 200 mm and a maximum around 400 mm $log_2(21) = 4.4$. On the other side of the spectrum the minimum of the Shannon entropy associated with a perfectly redundant set of hillslopes, is 0.

As seen in Figure 3.4 c and d, the entropy of the ensemble of runoff simulations starts at a rather low value at the beginning of our simulation period, increases with the first rainfall events in autumn, stays at a high level (ranging between 3 and 4) during the winter period, and starts to decrease towards 0 in May. During the summer, the entropy reacts much more strongly to the different rainfall events than in winter, and peaks at a value of 4.9 in August (35 from 54

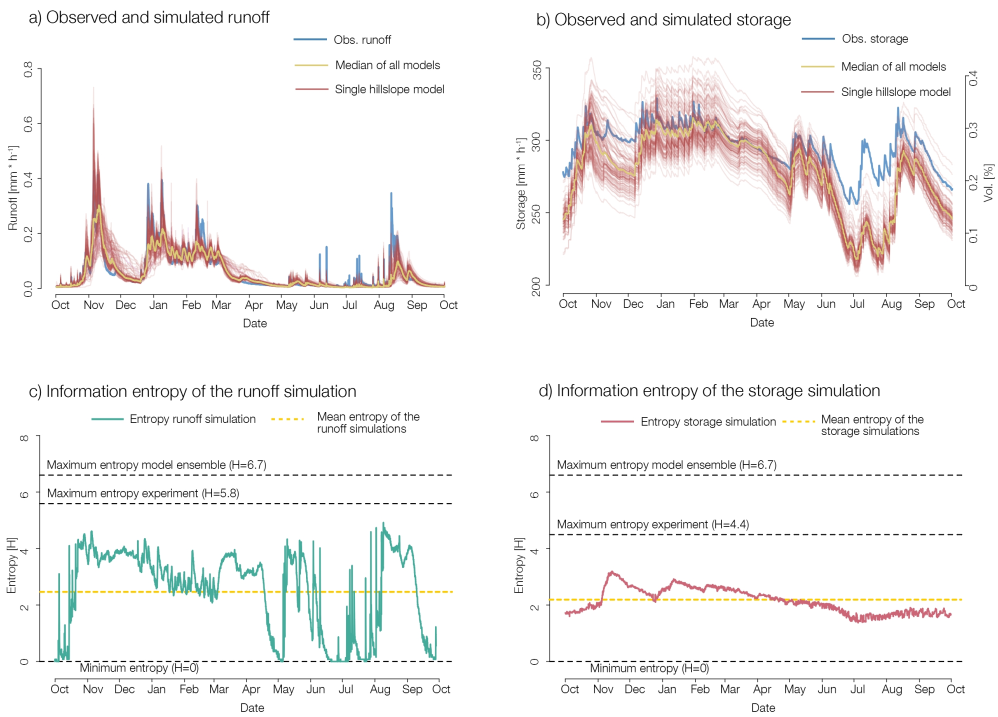


Figure 3.4: (a) Observed and simulated runoff of the Colpach catchment. The red lines correspond to individual hillslope models and the yellow line to area weighted median of all hillslopes. (b) Simulated total area specific storage of each hillslope in red and the median of all models in yellow. The median of the 141 observed soil moisture time series is smoothed with a 12 hour rolling mean (for more detail to the soil moisture observation we refer to Loritz et al. (2017) (c) Shannon entropy in turquoise for the runoff simulations as well as the corresponding mean and (d) a similar plot for the storage simulations (red).

bins allocated) when stream flow production grows again after a long dry period of low flow. It is interesting to note that the entropy in simulated stream flow is highly dynamic in time, implying that the required structural resolution of the model changes with time, with the 105-hillslope model structure being less redundant during periods of high entropy and more strongly redundant when entropy approaches 0 (see also Appendix D).

For the ensemble of storage simulations, the entropy varies between 1.5 and 2.9, which indicates less temporal variability compared to runoff. This is consistent with the visual impression that differences in topography result mainly in a parallel shift of the time series to a different annual mean. Nonetheless, the entropy time series exhibits weak annual dynamics, with a peak in mid-November when the wet season starts. This peak coincides with the entropy peak of the runoff simulations. In spring and summer, the entropy decreases slowly until it reaches the overall minimum of 1.71 in October. Note that this could be very different in case of (for instance) land-use differences or distributed rainfall among the hillslopes causing a likely increase of entropy during summer and autumn.

### 3.4.3 *Functional group and their typical runoff and storage dynamics*

The mean annual entropy of the runoff simulations is 2.5 (Figure 3.4 c), which implies that (on average) the number of functional groups or bins that can be distinguished lies between $22.5 \approx 6$ and $23.5 \approx 10$. In line with one of our goals to use information theoretic measures to define similar acting landscape elements and to compress the full catchment model into functional groups without substantial loss of information we took the lower value and used a hierarchical cluster analysis to classify the hillslopes into six functional groups using normalized mutual information (1-NMI) as distance metric. The median discharge for each functional group is shown in Figure 3.5 a, while the corresponding set of hillslope profiles is displayed in Figure 3.5 b. In general it seems that the functional groups 1, 2 and 6 exhibit the strongest differences with respect to their median runoff time series as well as with respect to the geopotential profiles whereas the classes 3, 4 and 5 appear much more similar in both aspects. The median of the storage simulation of each functional group is displayed in Figure 3.5 c. Consistently with simulated runoff, the storage time series of functional groups 1, 2 and 6 show the greatest differences. However, in contrast to the runoff simulations also the functional groups 3, 4 and 5 are better separable at least during the wet period. Consistent with the decline of the Shannon entropy in Figure 3.4 d these differences diminish in summer. Especially in June, July and August all of the functional groups simulate essentially identical storages as their differences are getting closer to the error margins of the soil moisture measurements. Again, we stress that this convergence could be explained by the dominant role of evapotranspiration and the identical land-use parameterization of all hillslopes. Note that functional group 6, showing the strongest and fastest overall runoff reaction and has the lowest overall storage simulation. Consistent with that, functional group 1 and 2, showing the slowest runoff reaction are characterized by the highest overall storage.

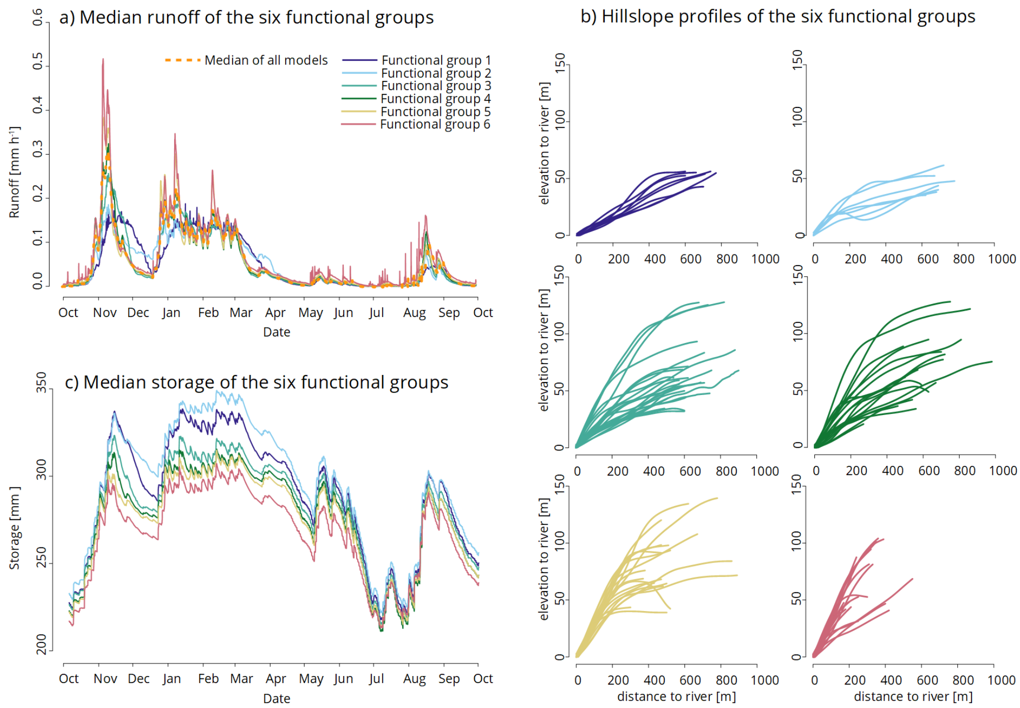


Figure 3.5: Median runoff of the six functional groups; b) corresponding hillslope profiles with the elevation to river on the y axis and distance to river on the x axis for each functional group. c) Median storage of the six functional groups.

### 3.4.4 *Performance of the compressed catchment models*

Figure 3.6 shows the cumulative frequency distribution of KGE values for the 1000 randomly selected model compressions using the aforementioned functional groups of similar runoff generation (Table 1). The median of all 1000 KGE values of all trials is 0.78 and corroborates that the compressed model structures perform on average slightly better than the area weighted median of the 105 hillslope models, which has a KGE of 0.76. However, the range of 0.66 to 0.88 in the KGE values indicates that the performance of a particular single realization of the compression depends on the actual combination of hillslopes

Table 3.1: Number of member as well as the mean and max values of the runoff simulation of each functional group.

| Functional group | Gr. 1 | Gr. 2 | Gr. 3 | Gr. 4 | Gr. 5 | Gr. 6 |
|---|---|---|---|---|---|---|
| member ($n$) | 9 | 8 | 27 | 20 | 20 | 21 |
| mean annual runoff ($mmh^{-1}$) | 0.051 | 0.052 | 0.053 | 0.054 | 0.056 | 0.065 |
| max runoff ($mmh^{-1}$) | 0.22 | 0.34 | 0.42 | 0.43 | 0.64 | 0.75 |
| mean storage ($mm$) | 289.6 | 295.7 | 281.7 | 277.1 | 273.7 | 267.7 |
| max storage ($mm$) | 338.6 | 349.1 | 323.7 | 316.2 | 312.8 | 307.2 |

selected for each group. As each realization of the compressed catchment model would in principle only use six hillslope models and if we assume that all hillslopes have the same run time this could, in theory, reduce the computational costs of our model application by a factor of 17.5.

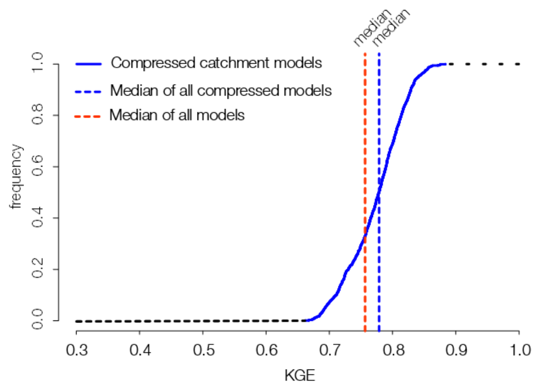


Figure 3.6: Distribution of model performances of the different realizations of the compressed catchment model (blue). The two dashed lines correspond to the median of the KGE values of all realization of the compressed catchment model (blue) as well as to the area weighted median of all 105 hillslope models (red).

## 3.5 discussion

The results presented above provide strong evidence that information theoretic concepts are powerful tools to quantify and explain the relevance of different system characteristics for distributed modeling. Following this overall result, we will start to discuss our main finding that the amount of topographic information relevant for distributed modeling is not constant but time variant. Furthermore in a second step, we address the closely related issue that we are able to compress the ensemble of hillslope models into functionally similar groups, and that a stronger compressibility implies a higher degree of functional organization in a heterogeneous environment. This discussion leads naturally to a short reflection on the advantages that concepts from information theory offer for exploring and explaining how spatial complexity and functional similarity of hydrological systems are connected. Finally, we conclude by revisiting the seeming antagonism between landscape organization (Dooge, 1986) and functional similarity (Wagener et al., 2007) against the recurring finding of heterogeneity and randomness and hence uniqueness of hydrological places (Beven, 2000) and provide an outlook on how to generalize the approach presented here.

### 3.5.1 *Temporarily varying importance of topography for distributed modeling*

The relevance of spatially variable but yet time-invariant topographic information on hydrological simulations was found to be strongly time dependent. The different topographic information used within the models led to complex temporal dynamics of the information content of the probability distribution of the discharge and storage simulations at a given time step. These temporal dynamics were furthermore distinctly different for the two target variables. The Shannon entropy of the discharge simulations revealed that there are alternating periods of high redundancy and of high diversity among the hillslope responses. This resulted in several local maxima and minima of the Shannon entropy in time. These maxima and minima are not easily explained by simply attributing them to high and low flow conditions (see Appendix D). For example the global maximum of 4.9 (close to the theoretical maximum of our experiment 5.8) was observed in August, when the system rapidly switched from low to high streamflow conditions in response to a strong convective rainfall event. In contrast, the Shannon entropy of storage simulation exhibited a distinctly different pattern compared to the discharge simulations with a much stronger autocorrelation, two clear identifiable maxima in winter, and overall lower values of the Shannon entropy in summer.

The overall differences between the two target variables, the dynamics of the information content within the discharge and storage simulations, and hence the changing maximal compressibility of the model ensemble, highlights that the relevant topographic information for distributed modeling depends firstly on the modeling target and secondly on the time, and thus on the prevailing forcing as well as on the state of the system. In other words, spatially distributed information about topography has a time varying impact on the model ensemble. Hence, the necessary complexity (Schoups et al., 2008) of a distributed model to capture this information is time dependent as well.

If we try to generalize and transfer this finding from the model world to a real hydrological system keeping in mind all the issues that go along such an approach, these results imply that *different landscape entities may either function similarly or dissimilarly depending on the time*. Hydrological similarity can therefore, rather than being static, be a dynamic attribute that depends on the "hydrological context". Interestingly, this context dependence can be straightforwardly explained by the generally dissipative nature of hydrological processes (Kleidon, 2010). Rainfall and radiation push and pull the hillslopes away from their local thermodynamic equilibrium, thereby generating internal system gradients in either potential energy or capillary binding energy. These gradients get depleted during system relaxation towards the equilibrium either through release of water from hillslopes to the stream or through recharge and capillary rise (Zehe et al., 2014). However, the generation and depletion of these gradients is controlled by a large variety of meteorological and hydrological processes interacting across a hierarchy of spatial and temporal scales (Blöschl and Sivapalan, 1995). Exactly the varying dominance of these processes, and hence the changing importance of the corresponding landscape control, is the key to understanding the time varying relevance of different system characteristics for distributed hydrological modeling, and explains the varying relevance of (in our case) topography for hydrological modeling even though topography is quasi static at classical hydrological time scale.

### 3.5.2 *Compressibility of time series and functional similarity of hillslopes*

As indicated in the section above, both of the target variables, storage and discharge, never reached the theoretical maximum value of the Shannon entropy implying that the model ensemble was producing redundancy and thus was compressible during the entire year. Based on this general finding we came up with the idea of a compressed catchment model which was built upon a straightforward clustering of all hillslope models into functional groups of similar annual runoff behavior. This compressed model consisted in a single realization of 6 instead of 105 hillslopes, which were then randomly drawn from each functional group. It is of interest that by reducing the model ensemble to a smaller set of hillslope models we were still able to match on average the observed annual streamflow in the catchment. This result agrees with the findings of Fenicia et al. (2016) who stated that spatial variations of the geopotential are too small in this landscape to have a dominant influence on the annual runoff generation, and with the findings of a foregoing study where we show that the annual runoff dynamics of the Colpach catchment can be simulated using a single effectively compressed hillslope model (Loritz et al., 2017).

Neglecting all the issues that occur when we compare distributed model applications with spatially aggregated models (e.g. Beven and Freer, 2001; Obled et al., 1994; Pokhrel et al., 2012) our comparison of the differently strong compressed catchment models matches with the conclusion of Pokhrel and Gupta (2010) that as long as we are not interested in the representation of the spatial distribution of hydrological fluxes or state variables, a spatially aggregated model which compresses the spatial variability of the landscape properties might be sufficient for predicting macroscopic variables (Hrachowitz and Clark, 2017). However, as soon as our focus shifts to the representation of the spatial distribution of a hydrological process, information entropy bears the key to defining and diagnosing the minimum adequate complexity of a distributed model (Schoups et al., 2008), particularly as it could help guide an approach to reducing computational costs without losing information (in our case by a factor of almost 17.5).
However, the assessment of a meaningful compression that leads to a less redundant and yet well performing distributed model structure is not at all a straightforward exercise. This is corroborated by the strongly variable performance of the 1000 randomly generated compressions, which highlights that the individual performance depends strongly on the model realization. From this we conclude that, contrary to our assumption, not each hillslope model represents stream flow generation of a functional unit equally well, as our classification is based on mutual information between the annual discharge time series. The fact that two hillslope models may yet act differently at certain time steps explains why every random realization of the model compression performs slightly different. The second and maybe more general shortcoming is that our proposed compression is based on a fixed number of groups, inferred from the average annual entropy. As the average annual entropy of simulated streamflow reflects the annual average maximal compressibility of the discharge simulation, our choice for the number of functional groups seems legitimate as a first attempt on an annual scale. However, as shown in Figure 3.4 c the Shannon entropy of the discharge simulations deviates substantially from this value. This implies that our model structure is either too simple in periods where the entropy is larger than the average or redundant in periods where the entropy is smaller. A best possible compression of a distributed catchment model, defined as the one that avoids any loss of information and also avoids any redundancy (also referred as lossless compression e.g. Weijs and Giesen, 2013) will therefore require a time variant number of functional groups. Such an effort to do simulations with a higher spatial model resolution in

times of high spatial complexity and with a coarser spatial model resolution in times of low spatial complexity, as is for example done with different adaptive time stepping schemes in numerical model implementations (e.g. Clark and Kavetski, 2010) or in adaptive model grid refinements (Faigle et al., 2014), points to new challenges that are not only beyond the scope of this study but likely also beyond the capabilities of most currently available model systems.

### 3.5.3 *Information theoretic measures to quantify similarity*

The venture to link complexity of spatially distributed catchment characteristics to functional similarity led us naturally to the concepts of information and (physical) entropy (Ben-Naim, 2008; Davies, 1990). Similarity of runoff, or storage of hillslopes, implies that their contribution to streamflow is redundant and hence does not change the information entropy within the simulations beyond its areal share (at least as long as the timing of the routing is not dominant). Removing this redundancy means to compress (Weijs et al., 2013b), and in our specific case to aggregate hillslopes to larger similar functioning landscape elements which we called functional groups in relation to the definition of functional units by Zehe et al. (2014). Although it is evident that this partitioning of similar acting units into larger groups does not require the use of information theory (e.g. Berghuijs et al., 2014; Sawicz et al., 2011; Wood et al., 1988, we believe that, besides the maybe more general assets of an information theoretic perspective on different hydrological issues (e.g. Ehret et al., 2014; Gupta and Nearing, 2014; Nearing et al., 2016; Weijs et al., 2013a), it has also major technical advantages for a variety of different tasks as shortly discussed in the following.

First, information theoretic measures like Shannon entropy and mutual information, when calculated with the same logarithmic base, share the same units, in our case "bits". This facilitates the intercomparison of the different variables, in our case storage and runoff, with respect to their diversity in the model ensemble. Furthermore, if calculated in the discrete form, a careful choice of the bin width according to the measurement error can also be interpreted as physical meaningful definition of the minimum separable difference between observations or simulations of the same state variable or flux. For instance, in this study, we used the inherent measurement errors of the soil moisture probes as well as the uncertainty in our rating curves to define the minimum separable differences of storage and runoff.

Another key advantage of the information theoretic perspective is that not only the minimum but also maximum information content and hence the maximal complexity or functional disorganization that a distributed model can produce in its responses is well defined. The latter corresponds to the state of maximum Shannon entropy which implies that each time series, either modelled or observed, contributes in a unique (non-redundant) fashion to the ensemble. We are therefore able to derive a theoretical upper and lower bound which reflects naturally the minimum and maximum reachable complexity of state/output response that our model can produce. The lower boundary represented by a zero entropy, corresponds to a situation where all model elements produce with respect to the corresponding observation error the same output and hence act identically. The upper boundary or maximum entropy, in our case 6.7, corresponds to a situation where all model units produce a unique output and to a situation of no redundancy at all. Given these two margins we can judge whether different model elements, in our case hillslopes, of a chosen model provide largely independent stream flow contributions.

## 3.6 CONCLUSION AND OUTLOOK

Based on the evidence presented here, we conclude that the proposed information theoretic measures and concepts provide a powerful framework for understanding and diagnosing how landscape organization and functional similarity of hydrological systems are connected. We are aware that the specific findings of the present work are necessarily constrained by the a-priori settings of the model ensemble, which exclusively focused on a spatially variable topography, while land-use, precipitation and the soil parameters were identical among the 105 hillslopes. The application of these concepts and the general mindset is, however, by no means restricted to this specific model neither to topography. On the contrary, it may be generalized either by additional data sources such as land-use, bedrock topography and distributed rainfall data as well as to any ensemble of time series, modeled or observed. This opens new opportunities to systematically explore how spatial variations of different landscape characteristics and meteorological forcing affect hydrological processes. Furthermore, as we only tested first order changes of topography and the influence on distributed modeling here, it also opens the possibility to test whether second order effects arise from combinations of several distributed characteristics.

Finally, in line with Clark et al. (2016) we argue that a comprehensive answer to the simple question stated in the introduction ”when is the spatial variation of a system characteristic large enough that we need to account for it” is not at all straightforward, but requires a solid theoretical framework. Following thermodynamic reasoning and information theory, the key to explain why hydrological systems often act so comprehensibly is that they are dissipative and highly organized (Zehe et al., 2014). This implies that organized simplicity might emerge when we move up to larger scales in space (Dooge, 1986; Savenije and Hrachowitz, 2017). Our results reveal, however, that simplicity manifests not only in space when moving to larger scales, but also manifests when ”the system moves through time” as functional similarity emerges in time. We therefore propose that the concepts of landscape areas that act either similarly and are thus redundant (Wagener et al., 2007) or show unique functioning and are thus irreplaceable (Beven, 2000) are consequently not mutually exclusive. They are in fact of complementary nature, and systems operate by gradually changing to different levels of organization in which their behaviors are partly unique and partly similar.

Part IV

# A TOPOGRAPHIC INDEX EXPLAINING HYDROLOGICAL SIMILARITY BY ACCOUNTING FOR THE JOINT CONTROLS OF RUNOFF FORMATION

This study is published in the scientific journal HESS. It is part of the special issue "Thermodynamics and optimality in the Earth system and its subsystems"; a joint special issue between the scientific journal Earth system dynamics (ESD) and HESS. The remainder of part III is a reprint of:

*Loritz, R., Kleidon, A., Jackisch, C., Westhoff, M., Ehret, U., Gupta, H., and Zehe, E. (2019). A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation. Hydrology and Earth System Sciences, 23(9), 3807–3821. doi:*10.5194/hess-23-3807-2019

# 4

# A TOPOGRAPHIC INDEX EXPLAINING HYDROLOGICAL SIMILARITY BY ACCOUNTING FOR THE JOINT CONTROLS OF RUNOFF FORMATION


## ABSTRACT

Surface topography is an important source of information about the functioning and form of a hydrological landscape. Because of its key role in explaining hydrological processes and structures, and also because of its wide availability at good resolution in the form of digital elevation models (DEM), it is frequently used to inform hydrological analyses. Not surprisingly, several hydrological indices and models have been proposed to link geomorphic properties of a landscape with its hydrological functioning; a widely used example is the "Height Above the Nearest Drainage" (HAND) index. From an energy-centered perspective HAND reflects the gravitational potential energy of a given unit mass of water located on a hillslope, with the reference level set to the elevation of the nearest corresponding river. Given that potential energy differences are the main drivers for runoff generation, HAND distributions provide important proxies to explain runoff generation in catchments. However, as expressed by the second law of thermodynamics, the driver of a flux explains only one aspect of the runoff generation mechanism, with the driving potential of every flux being depleted via entropy production and dissipative energy loss. In fact, such losses dominate when rainfall becomes runoff, and only a tiny portion of the driving potential energy is actually transformed into the kinetic energy of streamflow. In recognition of this, we derive a topographic index named reduced dissipation per unit length (rDUNE) by re-interpreting and enhancing the HAND index following a straight forward thermodynamic argumentation. We compare rDUNE with HAND, and with the frequently used topographic wetness index (TWI), and show that rDUNE provides stronger discrimination of catchments into groups that are similar with respect to their dominant runoff processes. Our analysis indicates that accounting for both the driver and resistance aspects of flux generation provides a promising approach to linking the architecture of a system with its functioning and hence develop similarity indices in Hydrology.

## 4.1 INTRODUCTION

The key role that surface topography plays in Hydrology has long been recognized (e.g. Horton, 1945). Topography provides information about the interplay between uplift, weathering and erosion, and hence about the past morphological development of a landscape. Further, it provides a strong constraint for future hydrological and geomorphic changes and, importantly for hydrology, is the key driver and control associated with runoff generation and several other hydrological processes.

This insight about the past, present and future roles played by topography is surely one reason why almost all key landscape entities in Hydrology, such as watershed boundaries, hillslopes and channel networks, are derived from properties of the land-surface topography. In support of this, digital elevation models (DEM) are available at fairly high resolution across the globe (Farr et al., 2007), helping to fuel the growing popularity of spatially explicit hydrologic models (e.g. Beven, 2001).

It is therefore no surprise that hydrology does not suffer from a lack of models or indices linking geomorphic properties of a landscape with its hydrological functioning. The most popular approach is arguably the topographic wetness index (TWI) proposed by Kirkby (1975) and Beven and Kirkby (1979). As a function of the local slope with the upslope contributing area per contour length, the TWI was originally developed to classify areas of similar functioning within a catchment and has been applied (e.g. Grabs et al., 2009), refined (e.g. Barling et al., 1994) and tested (e.g. Rodhe and Seibert, 1999) in numerous studies. However, other indices have also been proposed to link land surface topography with its runoff response. Hjerdt et al. (2004) developed the ”down slope topographic wetness index” (also called the $tan\beta$ index) that reflects the local hydraulic gradient in the case that flow is exclusively driven by gravity and under the assumptions of a fixed drop in elevation. They claimed that this index represents groundwater level gradients in a manner that is superior to the classical TWI approach, and showed it to be less sensitive to the quality of the DEM used to estimate the local slope. Adopting a hydraulics framework, Lyon and Troch (2010) developed an index called the catchment Péclet number, that is a volume or area weighted version of the hillslope Péclet number. The latter was derived by Berne et al. (2005) to characterize hillslopes by subsurface runoff formation, based on the relative importance of advective and diffusive flows, using the hillslope storage Boussinesq equation (Troch et al., 2003). Lyon and Troch (2010) showed that in a set of 400,000 synthetically generated and four real world catchments the catchment Péclet number provided a meaningful link between hydrological response and the geomorphic properties of a landscape.

An approach that has recently gained considerable attention is the "height above the nearest drainage" index (HAND) developed by Rennó et al. (2008), and under a different name "elevation difference (DZ)" by Crave and Gascuel-Odoux (1997). This approach assumes that water follows the steepest descent along the surface topography and, based on these drainage paths, the corresponding elevation of each raster cell above the nearest corresponding river cell is estimated. HAND has thereby been successfully applied and tested in numerous studies in a wide range of different landscapes. For instance, Gharari et al. (2011) compared a collection of hydrological similarity indices, their sensitivity to the DEM resolution as well as their ability to identify three visually pre-classified landscape types (wetlands, hillslopes, plateaus). Their results highlight the sensitivity of HAND to the chosen DEM resolution and show that HAND in combination with the slope lead to the "best" result with respect to match pre-classified observation points. Also Gao et al. (2014) used HAND in combination with the slope (additionally they also used the aspect) to identify hydrological similar areas in a model comparison study. They showed that a semi-distributed model setup which was based on a HAND landscape classification scheme outperformed a lumped and semi-distributed (based on the forcing data) hydrological model with respect to matching the hydrograph. The same leading author (Gao et al., 2019) further exploited the role surface topography plays when rainfall becomes runoff and used HAND to infer model parameters of a conceptual hydrological model showing that their developed runoff generation module performed almost as good as fully calibrated models. Finally, Zehe et al. (2018) used HAND as a proxy for the gravity potential for calculating potential energy of soil water and showed that their approach is "*well suited to distinguishing the typical interplay of gravity and capillarity controls on soil water dynamics in different landscapes.*"

The above mentioned studies highlight the large potential of the topographic index HAND and its relevance for hydrological research. From a theoretical point of view, HAND reflects thereby the gravitational potential energy of a given unit weight of water with the reference level set to the elevation of the nearest corresponding river. Given that differences in potential energy act as drivers for overland and subsurface storm flow, the distribution of HAND across a landscape represents a predominant control on the lateral distribution and redistribution of water in a catchment. However, because surface and subsurface water flows are also highly dissipative (e.g. Kleidon et al., 2013), similarity with respect to HAND distribution is not sufficient to ensure similarity with respect to runoff generation. This is due to the fact that the driving potential is only one of the important factors, with every flux encountering frictional losses along its flow path.

This latter insight recognizes the essential role of the second law of thermodynamics, based on which Zehe et al. (2014) postulated that equifinality is inherent to most of our governing equations, because every flux is unavoidably the result of the interplay between a driving potential and a resistance term. Accordingly, the overall flux through a system can remain unaffected when the driving potential is doubled if the corresponding frictional resistance losses are also doubled. From this perspective, only landscapes having similar combinations of characteristics controlling both the driver and resistance terms should satisfy a sufficiency condition for hydrological similarity (in terms of runoff generation). In recent years the importance of thermodynamic principles has increasingly gained attention in Hydrology. The Oxford dictionary defines thermodynamics as a "*branch of physical science that deals with the relations between heat and other forms of energy (such as mechanical, electrical, or chemical energy), and, by extension, of the relationships between all forms of energy.*" Given that all fluxes are driven by potentials, and that fluxes are necessarily "dissipative" (meaning that they produce entropy following the second law of thermodynamics (e.g. Kondepudi and Prigogine, 2014) it seems logical that thermodynamic concepts are relevant in Hydrology. However, although an energy-centered view has been applied to a variety of different issues in sub-disciplines such as groundwater hydrology (Hubbert, 1940) and soil physics (Babcock and Overstreet, 1955) it has not become established practice in classical rainfall-runoff centered surface water hydrology. This is likely due to the strong engineering context in which the understanding of surface hydrology was historically developed, with its overt focus on practical problem solving (Sivapalan, 2018). One interesting early exception is the work of Leopold and Langbein (1962), who showed that the concept of "entropy" in its probabilistic form (see Kondepudi and Prigogine, 2014) can be used in combination with a random walk term to infer the most probable state of a drainage network. Along the same lines, Howard (1990) and Rodríguez-Iturbe et al. (1992) showed how thermodynamic optimality principles can be used to derive realistic synthetic river networks. Such work motivated Hergarten et al. (2014) and others to apply similar concepts to explain subsurface flow patterns.

However, a thermodynamic perspective can be much more general, and is by no means limited to the explanation of optimal drainage densities. As examples, Zehe et al. (2013) showed that a thermodynamic optimum density of macropores maximizing dissipation of free energy during recharge events allowed an acceptable prediction of the rainfall-runoff response of a lower mesoscale catchment; Hildebrandt et al. (2016) used an energy-centered approach to explain how plants extract water from the soil, Zhang and Savenije (2018) how salt and fresh water mixing in estuaries can be described in energetic terms and Zehe et al. (2018) discussed how an energetic perspective on soil water movement can improve our general understanding of catchment hydrology. The above discussion highlights the considerable potential of a thermodynamic perspective to improve our understanding of hydrological functioning across a range of important issues. One reason that an energy-centered perspective on runoff generation remains the exception, rather than the rule, in catchment hydrology may be that the connection between the laws of thermodynamics and issues underlying questions of practical importance in hydrology is not always readily evident. A motivating rationale of this study is, therefore, to bridge this gap by showing how the fundamental concepts of thermodynamics can be applied to develop a solution to the classical hydrological question "*How can the geomorphic properties of a landscape be used to identify hydrological units that have similar hydrological functioning*".

In this study, we propose a topographic index that accounts for both the driving potential energy difference and the accumulated dissipative loss along the flow path following straightforward thermodynamic arguments. Our index, (reduced dissipation per unit length index) is thereby an energy-centered re-interpretation and enhancement of the well-established topographic index HAND. In the following, we discuss its similarities to other geomorphic indices used in Hydrology and test whether it provides sufficient information to enable distinguishing between two landscapes which differ distinctly with respect to their dominate runoff processes. Furthermore, are we comparing our index against a small subset of topographic indices namely its origin HAND and the frequently used TWI. Based on our findings we conclude that one meaningful way to build similarity indices in Hydrology is to acknowledge both the driving potential and the resistance term separately and hence identify the driving potentials and dissipative losses separately.

## 4.2 APPROACH AND METHODS

Here, we derive a topographic index based on the energy balance associated with runoff generation from a hillslope. This involves two steps: i) inferring which properties of a DEM provide information about the forces driving runoff generation, and ii) identifying how much resistance to the flow of water is offered by the landscape. As benchmarks for comparison, we briefly explain the well-established TWI and HAND indices.

### 4.2.1 *Energy balance of streamflow generation*

One of the most important steps in any thermodynamic approach is a proper system definition. Given that hillslopes are often described as the key landscape elements controlling runoff generation (e.g. Bachmair and Weiler, 2011), a starting point to describe the runoff generation of an entire catchment is to examine the energy balance of a hillslope with respect to the total energy of all fluids located on that hillslope. The total energy relevant for streamflow generation at the hillslope scale is thereby the sum of the influx of potential energy by water $J_{pot}$ (energy flux in W), the export of kinetic energy by water $J_{kin}$ (W), and the amount of energy D (W) dissipated due to friction along the flow path to the river (see Kleidon et al., 2013). In this regard, it is interesting to note that typically observed kinetic energies associated with overland flow are quite small compared to their driving potential energies. To get a sense of this, imagine a catchment having an average height above the runoff recording gauge of 20 m and a typical flow velocity of 1 m s$^{-1}$. In this case, only 0.5 % of the average potential energy is transformed into kinetic energy, while by far the largest amount (99.5 %) is dissipated due to friction at the fluid-solid interface along the flow path. This irreversible process implies an accumulative loss of free energy along its flow path, and hence a potential decrease in the ability of the fluid to perform work (Freeze and Cherry, 1979; Kleidon et al., 2013). The reason for this is that the potential and kinetic energies primarily determine how the fluid moves, while temperature differences within the fluid are of only minor importance. Accordingly, streamflow generation is accompanied by the conversion of potential energy into kinetic energy, and finally into heat (Currie, 2003; Song, 1992).

Fundamentally, the phenomenon of energy dissipation was first described through the second law of thermodynamics, which states that entropy can be produced but not consumed, implying that the sum of all processes in our universe proceed in a direction of entropy increase, meaning that they necessarily dissipate free energy and hence reduce the capacity of the system to perform work (Schneider and Kay, 1994). An elementary consequence of this is the negative sign in a diffusive flux law, which implies that heat flows from warm to cold temperatures, water flows downslope (more generally from higher to lower potential energy), and air moves from high-pressure to low-pressure. Mathematically this can be formulated as the flux gradient law, which states that any flux $\vec{q}$ is the product of a gradient $\nabla\varphi$ and the inverse of an effective resistance term $R$ which hampers the flux.

$$\vec{q} = \nabla\varphi R^{-1} \tag{4.1}$$

This equation was the basis for the statement by Zehe et al. (2014) that when dealing with the identification of hydrologically similar landscape entities we must consider the driving potential and the resistance terms separately. In the subsequent sections we explore each of these terms.

*The driving potential*

The main drivers for streamflow generation at the hillslope scale are the geo-potential differences between the upslope catchment areas and the stream channel, resulting from the gravitational energy of the mass of the water relative to its position (Bear, 1972; Kleidon, 2016). These potential energy differences driving streamflow generation are largely dependent on topographic differences, and on the space-time pattern of precipitation (Blöschl and Sivapalan, 1995). If the topography of a catchment is known, we can (in theory) calculate the potential energy associated to all water on the surface of a hillslope simply by applying Newtonian mechanics:

$$E_{pot} = mgh \tag{4.2}$$

where $E_pot$ is the potential energy of the water on the hillslope (J), m its mass (kg), g represents the gravitational acceleration (m s$^-2$), and h is the relative height of the water above a reference (m). Given Eq. 2 we can compute the influx of potential energy by water associated with a grid cell of a DEM by accounting for the spatial extension of the grid cell and the precipitation accumulated over a given time period. Accordingly, for each grid cell i of a DEM, we replace the mass term by the volumetric flux of water multiplied by its density $\rho$ (kg m$^-3$), the former computed as the summed total precipitation depth per time $P_i$ (m s$^-1$) within that grid cell multiplied with its area $A_i$ (m$^2$):

$$J_{pot,i} = P_i A_i \rho_i g_i h_i \tag{4.3}$$

$J_{pot,i}$ quantifies the influx of potential energy for a given grid cell $i$ and for a given time period. To finally calculate the influx of potential energy we need to set a reference level against which to quantify $h_i$. In this study, we will focus on catchments smaller than 50 km$^2$, and will therefore treat all of them as being "hillslope dominated", implying that channel routing is of only minor importance in the development of runoff generation (Kirkby, 1976; Robinson et al., 1995). By neglecting the stream network and assuming that water follows the surface topography along the steepest gradient, we can set the reference level to zero at the point where the hillslope connects to the nearest drainage, and thereby estimate $h_i$ for each cell in our DEM ($h_i$ = HAND). To summarize $J_{pot,i}$ quantifies the influx of potential energy by water within a given raster cell $i$, thereby providing an energy-centered interpretation of the well-established HAND concept.

The sum of $J_{pot,i}$ over a hillslope or catchment represents thereby the total influx of energy by water available to perform work in a given time period. It is hence straightforward to calculate $J_{pot,i}$ associated with, for instance, the long-term climatic precipitation if relevant information about the region of interest is available.

*Identifying the structures controlling dissipation*

While differences in geo-potential energy drive runoff generation, most of the available potential energy is dissipated during runoff generation. At the land surface this is controlled mainly by surface roughness (i.e. friction per unit length), which in turn depends on the nature of the vegetation, soil texture and the micro-topography. On the other hand, frictional losses within the subsurface are controlled by soil hydraulic conductivity, soil water content and (in case of deep percolation) by bedrock topography and conductivity. In both domains, additionally connected flow networks (such as rills, or vertical and lateral macropores) dramatically reduce frictional losses per flow volume, by providing a larger hydraulic radius (Hergarten et al., 2014; Howard, 1990).

The difficulty associated with estimating frictional losses, is that a variety of different runoff processes can occur within a hydrological year, all having different occurrence probabilities that are in turn controlled by different landscape properties of the hillslope. It is precisely this diversity of different spatio-temporal controls that makes it so difficult to upscale small scale processes to the scale of the entire catchment (Sivapalan et al., 2003). However, despite this variability, dissipation remains accumulative along the flow path (Kleidon et al., 2013; Rodríguez-Iturbe et al., 1992), offering the opportunity to define a "dissipation length" as a surrogate for the macroscopic flow resistance in the flux-gradient relationship (Eq. 1) as long as the pedo-geological setting does not change significantly along the flow path. For simplicity, we henceforth assume that the dissipation of the geo-potential energy during runoff production is proportional to the flow path length to the river. This assumption is in line with those made by Rodríguez-Iturbe et al. (1992) in the context of stream networks, and is based on the observation that the export of kinetic energy by water ($J_{kin}$) is often negligible small compared to the influx of potential energy by water ($J_{pot}$). The majority of available potential energy is hence dissipated (D) when rainfall becomes runoff:

$$D = J_{pot} \qquad \textit{given} \qquad J_{kin} \ll J_{pot} \tag{4.4}$$

A given mass of water traveling from a specific location (grid cell i) to the stream will dissipate its potential energy over its travel distance leading to:

$$\frac{D_i}{l_i} = P_i A_i \rho g \frac{h_i}{l_i} \tag{4.5}$$

With $l_i$ being the flow length of a given raster cell i to the nearest drainage (m) and $h_i$ the height above the nearest drainage (m) of that raster cell i. To assure that the developed index depend exclusively on information about the topography stored within a DEM we normalize Eq. (5) by the mass flux of precipitation and divide it by the gravity constant g, the resolution $A_i$, and by the density of water $\rho$, to obtain a dimensionless index:

$$rDUNE = \frac{h_i}{l_i} \tag{4.6}$$

This reduced dissipation per unit length index (rDUNE) is an estimate of the potential energy gradient at the surface topography of a given raster cell under the assumption of gravitational flow, and is similar to the index proposed by Hjerdt et al. (2004) but without the need to arbitrarily define the drop in elevation. Here we have chosen to multiply rDUNE with minus one as well as use the natural logarithm transformation to make the rDUNE more easily comparable with the TWI as well as to transform the skewed distributions to be more normally distributed and thereby make its patterns more easily interpretable.

$$rDUNE = -ln(\frac{h_i}{l_i}) \tag{4.7}$$

rDUNE is defined in a range from $-\infty$ till $\infty$, is zero if the flow length and height to the nearest drainage are equal, positive if the flow length is larger and negative if the flow length is shorter than the height to the nearest drainage. High rDUNE values mean that the dissipation of potential energy is reduced compared to landscapes with lower rDUNE values. This reduction could for instance stem from higher hydraulic conductivities of the prevailing soils or from the occurrence of different forms of preferential flow paths.

### 4.2.2 *Topographic wetness index (TWI) and height above the nearest drainage (HAND)*

We compute the frequency distributions of grid cell TWI and HAND indices for comparison with the rDUNE distributions. The TWI is defined for each raster cell as:

$$TWI = ln(\frac{a}{b}) \tag{4.8}$$

where $\alpha$ is the upslope accumulated area and $tan(\beta)$ the local slope angle (the TWI is usually divided by the resolution of the DEM before the logarithm is taken, to make it dimensionless). Meanwhile HAND is based on the concept that water follows the steepest gradient along the surface topography, and hence both a river network and as a flow direction map are required for its calculation. To better compare HAND with TWI and rDUNE, we again use its natural logarithm (ln(HAND)).

*Measuring divergences between distributions*

Measuring the similarity or dissimilarity of frequency distributions without resorting to statistical moments is not straightforward. Here we use a less well known measure, called Jensen-Shannon divergence (JSD, Lin, 1991) to estimate how similar catchments are with respect to their ln(HAND), TWI and rDUNE distributions. JSD is a non-negative, finite and bounded distance measure developed to quantify the divergence between probability distributions. It was introduced into Hydrology by Nicótina et al. (2008) and is strongly, but not necessarily, motivated by Information Theoretic considerations (for details on Information Theory please see Cover and Thomas (2005). JSD is based on the well-known Kullback-Leibler divergence (KLD; sometimes referred as relative entropy) defined as:

$$D_{KL}(X||Y) = \sum_{x \in X} p(x_i) log_2 \frac{p(x_i)}{p(y_i)} \tag{4.9}$$

where $p(x_i)$ and $p(y_i)$ are the probabilities that $X$ and $Y$ are respectively in the states $x_i$ and $y_i$. In brief, KLD quantifies the information loss when the probability density function of $Y$ is used in place of $X$, and has been applied in hydrology by Weijs et al. (2010) to evaluate hydrological ensemble predictions. However, because KLD is not a classical distance measure, being neither symmetric nor bounded (Majtey et al., 2005) it is not well suited to the simple comparison of distributions. To overcome this issue Lin (1991) and Rao (1982) developed a symmetric and bounded version of KLD that, when subjected to a square root transformation, satisfies the triangle inequality condition required of a distance metric (Endres and Schindelin, 2003). This is accomplished by computing the sum of the KLD of $(X||Y)$ and $(Y||X)$, thereby making it symmetric, as was originally proposed by Kullback and Leibler (1951) as the ”J divergence”. In its general form for N distributions, the J divergence can be written as:

$$J_{KL} = \sum_{i=1}^{N} (X_i||Y_i) \tag{4.10}$$

From this, the JSD is developed, by comparing each distribution to the "mid-point" distribution M, defined as:

$$M = \frac{1}{N}\sum_{i=1}^{N}(X_i + Y_i) \tag{4.11}$$

Accordingly, the JSD represents the average divergence of N probability distribution from their mid-point distribution, defined as:

$$JSD = \frac{1}{N}\sum_{i=1}^{N} D_{KL}(X_i||M) \tag{4.12}$$

If we calculate the JSD using logarithms to the base 2 the JSD associated with two distributions is bounded between zero and unity, while for N distributions it is bounded between zero and the maximum entropy $log_2 N$ (Jaynes, 1957). This is because the mid-point distribution M converges to a uniform distribution in the case of maximum dissimilarity between the distributions.

*Derivation of probability distributions*

To calculate the JSD it is necessary to convert the frequency distributions of ln(HAND), TWI and rDUNE into probability density functions. This step requires a careful choice of bin width (Gong et al., 2014). Various guidelines to properly estimate the bin width have been proposed, one of the earliest and most frequently used having been proposed by Scott (1979):

$$W = 3.49\sigma N^{-\frac{1}{3}} \tag{4.13}$$

where W is the bin width, $\sigma$ is the standard deviation of the distribution and N is the number of available samples belonging to the distribution. In our study, however, the optimal bin width turns out to be different for each distribution as a result of its shape and the number of samples (size of the catchment). This is inconsistent with the need to use the same binning for each case to facilitate comparisons of the different distributions. Accordingly, we decided to use only the largest bin width calculated for each similarity index – which is 0.5 for the TWI distributions, 0.2 for the ln(HAND) distributions and 0.15 for the rDUNE distributions (please note the JSD values between the distributions change only slightly if calculated with the smallest bin size; see appendix A1). Finally, as recommended by Darscheid et al. (2018), for any bin indicating zero probability (no data samples are found to fall in that bin) we treated it as though it contained a single sample, thereby associating that bin with a very low probability of occurrence.

## 4.3 STUDY AREA

The 288 km$^2$ Attert catchment, located in Luxembourg, has a mean annual precipitation of 850-1100 mm and mean monthly temperatures varying between 0°C in January to 18°C in July. Detailed descriptions of the climatology and hydrology of the catchment can be found in a series of studies (e.g. Bos et al., 1996; Jackisch, 2015; Martínez-Carreras et al., 2016; Wrede et al., 2015. An important – and particularly relevant – characteristic of the catchment is that it consists of two major geological formations. Devonian schists dominate the Ardennes massif in the northern and western part, and Triassic sandy marls dominate the rest of the catchment, interrupted by several small areas of sandstone in the south and north-west. To test the functional discrimination ability of the rDUNE, TWI and ln(HAND) indices, we selected six headwater catchments of different sizes (see Fig. 4.1), three in the Schist area (Platen 40 km$^2$; Colpach 19.4 km$^2$; and Weierbach 0.45 $km^2$), and three in the Marl area (Schwebich 30 km$^2$; Niederpallen 32.2 km$^2$; and Wollefsbach 4.4km$^2$).

### 4.3.1 *Hydrological regimes and runoff generation*

Important to this study is that the six catchments share similar hydroclimatic regimes (Jackisch, 2015), which can be separated into winter and vegetation seasons, during which either runoff or evapotranspiration respectively are the dominant water fluxes leaving the catchments (Loritz et al., 2017). Annual runoff coefficients vary from 30 - 60 % indicating distinct differences between the years; this is most likely the result of annual climatic variations (Pfister and Kirchner, 2017).

However, the way how the catchments transform rainfall to runoff, varies significantly between the different geological formations (Bos et al., 1996). The Schist region is characterized by a "fill and spill" runoff generation mechanism, wherein water flows along or within the bedrock comprise the dominant runoff process. On the other hand, in the Marl regions, saturated areas and preferential flow paths within macrospores and soil crack dominate how water is distributed.

Differences between the runoff regimes are highlighted in Fig. 4.2 for a series of rainfall-runoff events in the winter, summer, and autumn of 2012 and 2013. The runoff response in the Marl catchments is rather rapid and more peaked (but with less volume) than in the Schist catchments (Loritz et al., 2017). It is noteworthy that although all of the Marl catchments are of different size, they exhibit very similar patterns of runoff generation. On the other hand, the behaviours of the Schist catchments are quite different from each other, with Platen producing (over the long term) $\sim 30\,\%$ less discharge than the other two. A possible explanation for this is that around 30 % of the Platen catchment belongs to a sandstone formation that tends to be less responsive with regards to runoff and has deeper groundwater stores (Bos et al., 1996). Despite these differences, and in spite of the fact that their sizes differ by a factor of 10, the Schist catchments exhibit surprisingly similar runoff responses (with Spearman rank correlations above 0.9). This is highlighted by the characteristic double peaked nature of the runoff events in all three Schist catchments during the winter (Martínez-Carreras et al., 2016).

To summarize, the two geological formations share rather similar hydro-climatic regimes, but differ significantly with respect to dominant runoff processes and hence how they transform rainfall to runoff. We should therefore expect that any catchment similarity index, developed for the purpose of identifying and explaining differences in hydrological functioning (in terms of runoff generation), should be able to clearly distinguish these two geological areas from each other. It is important to note that we picked this set of catchments on purpose, because the climatic differences between the catchments are rather small and the corresponding catchments share a rather clear geological setting. This was possible due to the fact that the Attert catchment and sub-basins were setup for research purposes rather than for management reasons. Larger data sets with catchments fulfilling the conditions of comparable climatic and geological settings are rare, making the definition of functional similarity challenging in catchment comparative studies as well as our assumption that the pedo-geological setting does not change significantly along the flow path.

### 4.3.2 *Spatial analysis and the stream network*

For our topographic analyses we used a 5 m LIDAR digital elevation model, aggregated and smoothed to 10 m resolution. All spatial analysis were conducted using GRASS GIS (Neteler et al., 2012) and the GRASS GIS extension r.stream* (Jasiewicz and Metz, 2011). The latter was used to derive the distance-to-the-river and elevation-to-the-river (HAND) maps, used as the spatial basis for all subsequent analyses. Because the calculation of these maps is very sensitive to the extension and shape of the river network it is important to derive the stream network with care; for this analysis we used the stream network created by Loritz et al. (2018), by separately varying the minimum contributing area thresholds, depending on the geological setting, to match the official stream network available from the Luxembourg Institute of Technology (LIST). In addition, the stream network was evaluated against orthophotos and manually adjusted in close collaboration with field hydrologists working in the Attert region.

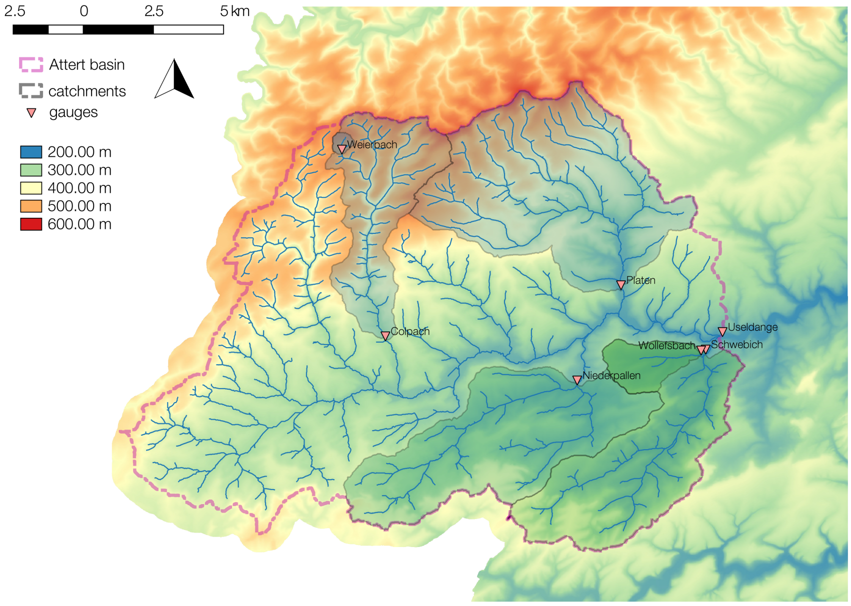


Figure 4.1: Map of the Attert basin with the six selected headwater catchments. In the northern part of the Attert catchment the three schist dominated catchments (blue: Platen, Colpach, Weierbach) are highlighted and in the southern part, the three marl dominated catchments (green: Schwebich, Niederpallen, Wollefsbach).

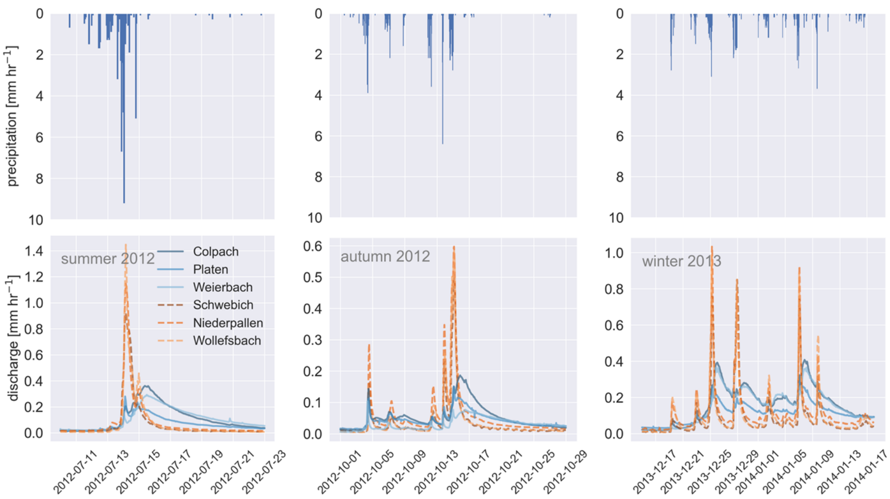


Figure 4.2: Observed specific discharge and precipitation with different ordinate scales for a time period in summer 2012 autumn 2012 and winter 2013 in the six catchments (orange: marl catchments and blue schist catchments.). This figure highlights that the two geological formations have a distinctly different hydrological function with respect to how they transform rainfall to runoff throughout the year.

## 4.4 RESULTS

Fig. 4.3 displays the frequency distributions and corresponding cumulative density functions of TWI, ln(HAND) and rDUNE for the six catchments examined in this study. In general, the TWI distributions do not indicate strong differences between the two geologies. For all six catchments, the distributions tend to be approximately Gaussian, with mean values close to 8 (see also Table 4.1). Visually, only the Platen and Colpach differ slightly from the other catchments, with distributions shifted somewhat to the left (lower means). That these six TWI distributions are indeed rather similar is also indicated by the JSD (Fig. 4.4), the values of which are all rather small indicating low divergence between the distributions. This similarity of the TWI distributions in spite of geological differences may, on first glance seem somewhat surprising given that the Schist catchments are generally much steeper than the Marl ones. However, in the Marl regions the water flow along the surface tends to be much less convergent, and consequently the flow accumulations tend to be lower than in the Schist regions.

The corresponding comparison of the ln(HAND) distributions indicates a greater degree of divergence between the two runoff regimes. In particular, the Platen and the Colpach catchments (both in the Schist region) differ from the other catchments with ln(HAND). This visual impression is reinforced by the average values of ln(HAND) (Tab. 4.1), with both the Colpach and the Platen catchments exhibiting similar average values close to 3 (ln(m)). In general, however, the index does not indicate a very distinct separation between the two geologies, and does not clearly distinguish between the Weierbach (Schist) and Niederpallen (Marl) catchments. The JSD values further reinforce the fact that the differences between the distributions tend to be quite small. For instance, the Platen (Schist) and Schwebich (Marl) catchments have very small JSD values ($\sim 0.042$), while the Wollefsbach catchment that is within the same geological formation (Marl) as the Schwebich has a JSD value of 0.11.

In contrast, the rDUNE distributions reveal a rather different picture. Visually, the rDUNE index clearly distinguishes between the two geologies. In particular, the shapes of the cumulative density functions indicate that the Marl catchments tend to have lower rDUNE values than the Schist catchments. The mean values of the rDUNE distributions (Tab. 4.1) are around 1.94 - 2.18 for the Schist catchments, and around 2.9-3.5 for the Marl catchments. Meanwhile, the JSD between all three Schist catchments are below 0.1, while being as large as 0.49 when computed against the Marl catchments.

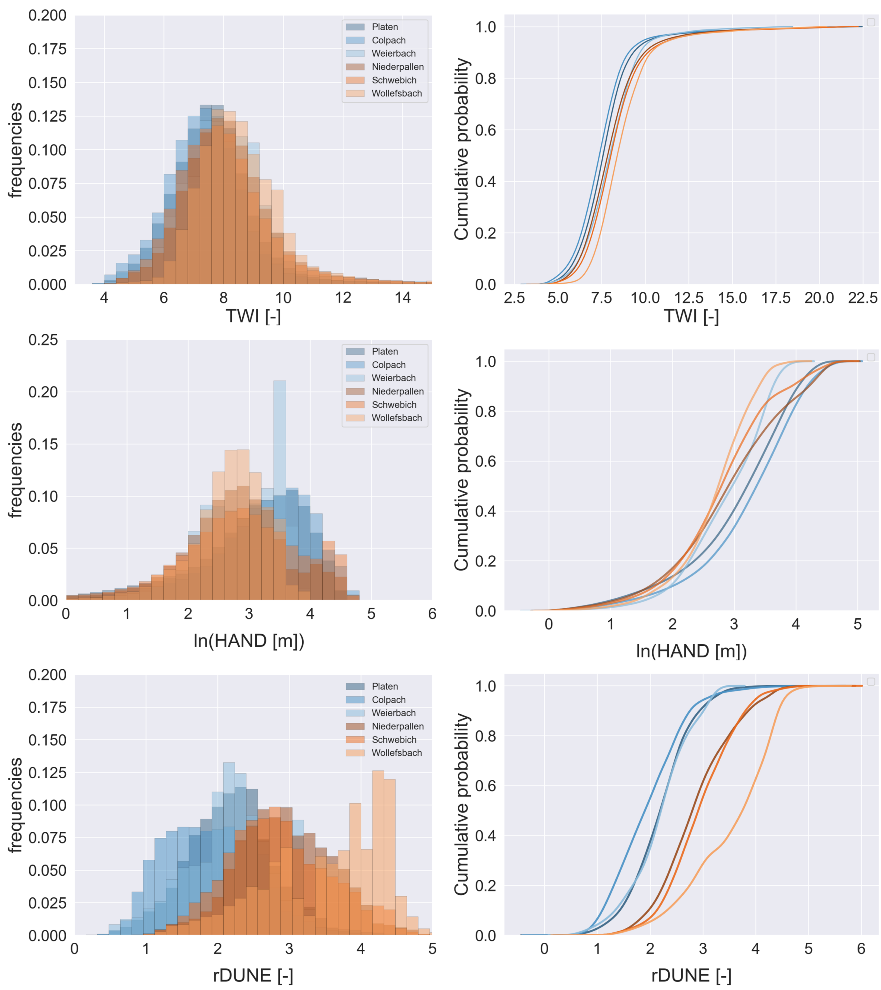


Figure 4.3: Frequency distributions and cumulative density functions of the TWI, ln(HAND) and rDUNE for the six research catchments. In blue the schist catchments (Platen ,Colpach, Weierbach) and in green the marl catchments (Schwebich, Niederpallen, Wollefsbach).

Table 4.1: Average (∅) and standard deviation (std) of the TWI, ln(HAND) and the rDUNE for each experimental catchment.

| | ∅ TWI+std [-] | ∅ ln(HAND)+std [ln(m)] | ∅ rDUNE+std [-] |
|---|---|---|---|
| **Schist** | | | |
| Platen | 7.77±1.9 | 3.03±0.9 | 2.18±0.5 |
| Colpach | 7.54±1.9 | 3.21±0.9 | 1.94±0.6 |
| Weierbach | 8.05±1.6 | 2.85±0.7 | 2.17±0.6 |
| **Marl** | | | |
| Niederpallen | 8.3±2 | 2.77±0.8 | 2.93±0.7 |
| Schwebich | 8.1±1.9 | 2.88±1.0 | 2.9±0.7 |
| Wollefsbach | 8.67±1.8 | 2.66±0.6 | 3.52±0.6 |

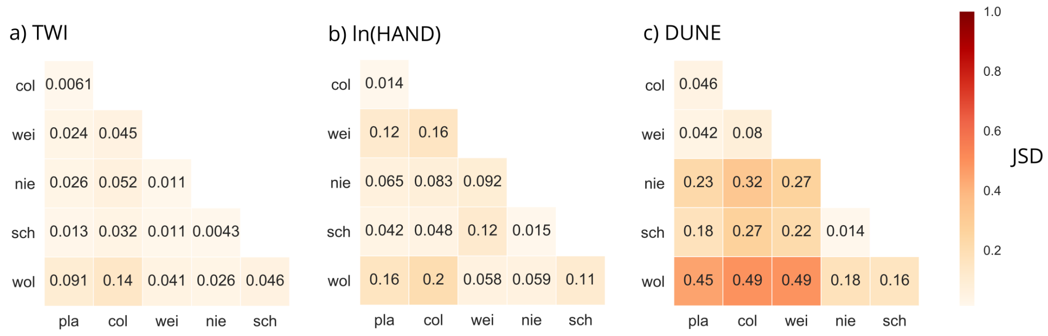


Figure 4.4: JSD values for the six research catchments (Schist: Platen (pla), Colpach (col), Weierbach (wei); Marl: Niederpallen (nie), Schwebich (sch), Wollefsbach (wol)). Panel a JSD of the TWI, b of the ln(HAND) and c of the DUNE. A high JSD value indicates a high divergence between the distributions with a maximum of 1.

## 4.5 DISCUSSION

### 4.5.1 *Potential energy differences as the driver for runoff generation*

The reduced dissipation per unit length index (rDUNE) is a straightforward energy based enhancement of the frequently used HAND approach (Rennó et al., 2008). The small, but significant, difference is that rDUNE is computed by dividing HAND by the flow path. This is motivated by the fact that almost all of the potential energy is dissipated within the runoff generation process. Though this extension might seem incremental, rDUNE thereby accounts for both the driving potential energy difference and the dissipative energy losses associated with the production of runoff. The latter is likely of particular importance when examining environments having a distinct topography where runoff generation is not limited by the available potential energy but by dissipation, and therefore facilitating preferential flow structures dominate surface and subsurface runoff generation. Accordingly, rDUNE should help to improve the classification of catchments into functionally similar spatial units, particularly for headwater catchments having moderate to steep topographies (Montgomery and Dietrich, 1988).

The first indication that rDUNE is indeed a useful addition to the variety of available topographic indices in Hydrology is highlighted by our results which show that rDUNE distributions stronger discriminate catchments with two distinctly different runoff regimes as it is the case using HAND or the TWI (Fig. 4.4). Furthermore, the fact that rDUNE values are on average higher in the marl region compared to schist catchments is physical reasonable considering the circumstances that soil cracks and worm burrows (in general preferential flow e.g. Loritz et al., 2017) play an important role in the way how the catchments transform potential energy in kinetic energy. This is the case as these structures reduce the dissipation of energy along the flow path and higher rDUNE values are expected in landscapes in which the dominant runoff process are characterized by flow through preferential flow path.

Despite these first promising results a full analysis of rDUNE and its sensitivity to different DEMs resolutions, flow direction algorithms as well as terrain smoothing functions is needed, as it has been shown in detail for the TWI and HAND by Gharari et al. (2011). However, as rDUNE is rather an extension and an energy-centered re-interpretation of HAND we would expect that the findings from Gharari et al. (2011) about the meaningful range of raster resolutions can be, at least partly, transferred to rDUNE. Exactly this relationship between HAND and rDUNE is hence rather a strength as a weakness and it would be an interesting avenue to test how and if the landscape classifications and model results of Gao et al. (2014) and Gao et al. (2019) change if HAND is replace by rDUNE.

Additionally, we speculate that our energy-centered re-interpretation of HAND may, besides improving its theoretical underpinning, further open the possibility to dynamically classify landscapes over time. This is because the incoming potential energy and the energy-centered foundation of rDUNE (Eq. 5 (J $m^{-3}$)) can be instead calculated with a mass flux rather than with a total mass, for instance using an hourly precipitation time series. As discussed in Loritz et al. (2018) this kind of dynamic classification may provide the key to successfully partitioning a catchment into similar functioning landscape entities, as hydrological systems move from complex to organized states. As a consequence, rDUNE in its current time invariant form will always be limited to identifying hydrological similar landscape units.

### 4.5.2 *Sensitivity to drainage density*

The fact that the rDUNE frequency distributions varied across the two geologies is clearly due to the fact that different accumulation values were used to derive the channel network in the different geologies. Changing the accumulation threshold means that water will start to flow sooner or later at the surface and hence that the flow length and the elevation to the nearest drainage will increase or decrease. The origination point of the channel network is thereby controlled by a variety of structural and climatic controls, and often varies depending on the prevailing season Montgomery and Dietrich (1992). However, varying the accumulation threshold within a reasonable range mainly changes the flow length in headwater catchments, and the flow length and elevations along the main river (where we are rather certain about the position of the channel network) will not change dramatically.

Another point, more specific to our tested geological formations, is that flow directions are more parallel in the Marl regions as a result of the smoother topography. Therefore, water will start to flow later at the surface within the stream network even if we choose the same flow accumulation threshold in both geologies. This can, of course, depends on the chosen flow direction algorithm (Seibert and McGlynn, 2007). Nevertheless, the fact that the accumulation area needed to form a channel is, in general, larger in the Marl region where slopes are more gentle compared to the Schist regions, matches the observation by Montgomery and Dietrich (1988) that there is a strong inverse relationship between the average length of a hillslope and its slope.

Finally, leaving aside the technical details of extracting a river network based on a DEM and the uncertainties that go along with such an approach, we note that the stream network we use in this study was carefully extracted based on an official stream network, and on several visits to the area, and was checked using orthophotos. This means that we are confident that we have correctly captured the overall picture of the perennial channel network, even if we are not able to examine every location where water under typical conditions begins to form a channel. The fact that the drainage densities of a catchment provide important information about the hydrological functioning of a landscape has been shown by several studies (e.g. Mutzner et al., 2016). This is because the extension of the stream network reflects the interplay of the climatic forcing and the hydro-pedological setting of a landscape and therefore the interaction of the driving potential of runoff generation and the resistance which works against it. This observation was previously made by Montgomery and Dietrich (1988), who postulated that it is logical to use the information stored within the extension of a channel network and the average hillslope length and height (slope) for developing models that try to explain hydrological similarity based on the topography.

### 4.5.3 *Topographic similarity and hydrological similarity*

Our comparison of TWI, ln(HAND) and rDUNE indicates that the rDUNE is more able to detecting differences between the two runoff regimes tested here. However, there exist a variety of other topography-based indices in use which we do not test in this study, ranging from simple comparison of the mean slopes of a catchment to approaches based on assumptions that are rather similar to those made in this study. A prominent example is the work of McGuire et al. (2005) who used the median flow path length (L), the median flow path gradient to the river (G) and the ratio of both (among other variables) to analyze how much of the inter-catchment variability of residence times of tracers can be explained by geomorphic properties. They found that the ratio of the flow path and the slope was superior to other variables in explaining hydrological dynamics. McGuire et al. (2005) stated that *"... the correlation of residence time with L/G is significantly better than the correlation of residence time with flow path length (L) or flow path gradient (G) individually. This suggests that both factors are important controls on residence time."*

Interestingly, Harman et al. (2009) gave exactly this index, under a different name, a theoretical basis when they derived the Boussinesq equation within their hillslope similarity study. It is remarkable that their topographic index (L/G) is rather similar to the rDUNE or the tan $\beta$ index (Hjerdt et al., 2004), although they use the median of the local slopes as proxy for the driving potential instead of the potential energy and further altered the ratio by dividing the flow path length by the gradient and not vice versa. The similarity between the three indices is, however, still evident as both include a surrogate for the driver of a flux and a surrogate for the friction term working against it.

In this context it is interesting to note that also the system properties represented in our governing equations are rarely independent but rather act in conjunction (Bárdossy, 2007). Because most similarity indices are derived upon those governing equations, we can find the aforementioned pattern in many other successful hydrological indices. For instance, also the TWI combines the driving potential (local slope) with an estimate of the conductivity of a given area (in the form of the upslope accumulation area). These assumptions might be appropriate for northern England (where TWI originally was developed) and may also work in many other environments, but will likely fail if the driver or the resistance term are not appropriately estimated. This highlights the fact that the concept of combining system properties driving a flow with properties that hamper flow might indeed be one meaningful way to link the hydrological functioning of a system with its architecture (Zehe et al., 2014). As the physical foundation for this perspective is based on thermodynamics it might be an advantage to routinely consider runoff generation not exclusively as a mass flux but as an energy driven and dissipative process, as this perspective may help us to better generalize our findings and identify the limitations of our concepts and models.

## 4.6 CONCLUSION

The dissipation per unit length index developed here is an energy-centered re-interpretation of the HAND index. Its use enabled us to use DEM data to detect differences between two sets of catchments having distinctly different dominant runoff processes and, in this regard exhibited superior performance to the TWI and HAND approaches. Our results indicate that a promising way to link system architectures with their functioning is to identify system properties in such a way that we can account separately for both the drivers of a flux and the properties that act to resist it.

The general idea behind this study is thereby the observation that the majority of the incoming potential energy associated with water flow within a hillslope is dissipated and only a fraction of it reaches the stream network as kinetic energy. This highlights the important role energy dissipation plays when rainfall is transformed to runoff within a catchment. Establishing a proxy for the structures that control energy dissipation is thus the key to functionally classifying environments that are not limited by the available potential energy and therefore have distinct topographies. Finally, by taking an energy-centered perspective on runoff generation, we can begin to address the question of why landscapes evolve in such a way that most of the potential energy is dissipated at the hillslope scale, although it is frequently reported that energy dissipation is minimized within river networks (Kleidon et al., 2013; Rinaldo et al., 1992; Rodríguez-Iturbe et al., 1992; Zehe et al., 2010).

# Part V

# SUMMARY AND SYNTHESIS

In following chapter, I condense the key findings and results I obtained in this thesis. Furthermore I propose opportunities for further research and discuss the key findings and their general relevance for hydrological modeling.

# 5 SUMMARY AND SYNTHESIS

## 5.1 SUMMARY OF THE KEY FINDINGS

*The ungauged basins problem.*

A large body of the hydrological research has well documented the role heterogeneity and process complexity play when the connection between hydrological function and structure is examined (McDonnell et al., 2007). A frequent finding is that particularly at the catchment scale it is challenging to distinguish between idiosyncratic and relevant system details (e.g. Sivapalan et al., 2003). This lack of fundamental system understanding and hence our inability to identify dominant hydrological system properties explain at least partly why the majority of available hydrological models need to be calibrated to produce meaningful results. This is, however, by no means a new observation and was already realized around 20 years ago when the hydrological community recognized that they were to a large extent unable to make reliable predictions in ungauged basins (e.g. Hrachowitz et al., 2013; Sivapalan et al., 2003). The issue of making predictions in ungauged basins is, although recognized two decades ago, still at the heart of hydrological research and with a ongoing climate change maybe more relevant than ever.

*Lack of theory explains partly why predictions in ungauged basins is still a challenge.*

One reason – among others – which partly explain the slow progress made in simulating hydrological fluxes and states in ungauged basins is owned by the fact that, even today, a large number of hydrological models remain essentially engineering concepts built upon strong physical simplifications and empiricism (Kirchner, 2006). While these concepts proved to be suitable tools for stationary hydrological predictions, for instance to design flood protection facilities, they are largely inappropriate to explore the influences of changing boundary conditions on hydrological systems or to simulate land-use changes as they are fundamentally built upon a stationary system understanding. A long-standing vision, since the work of Dooge (1986), is hence to improve the theoretical underpinning of hydrological models and to derive new hydrological laws and theory particular for the catchment scale (Clark et al., 2016; Kirchner, 2006; Sivapalan, 2003; Zehe et al., 2014).

My thesis is framed by this search for an improved theoretical underpinning of hydrological models with an emphasis on meso-scale catchments. In all chapters I discuss my research from a practical but also from a theoretical point of view with the objective that the

*Catchments are systems of organized complexity.*

findings can be generalized beyond the studied research environments. My arguments are thereby largely drawn from information theory and thermodynamics, and both serve as a general scientific framework and language in this thesis. In the following sections 5.1.1-5.1.3 I briefly summarize the key findings of this thesis and discuss their general implication for hydrology with a focus on hydrological modeling.

### 5.1.1 *Part II: Picturing and modelling catchments by representative hillslopes*

In my first study I design, implement and test a hydrological modeling concept - representative hillslopes - for meso-scale catchments. This concept is particularly interesting for catchments in which more information about the system is available than the classical hydrological rainfall-runoff data sets. As the amount of available and relevant data in hydrology is constantly increasing it seems fair to assume that this modeling approach will be increasingly possible in many regions around the world in the future.

*The role of point-scale soil water retention properties for hydrological modeling.*

The results of chapter 2 highlight that it is not necessary to map the entire heterogeneity of the observed soil water retention properties (Jackisch, 2015) into a hydrological model if the goal is to simulate macroscopic fluxes and state variables like streamflow or storage dynamics of a catchment. This finding is in line with the frequently shown result that rather simple mathematical models with less than four parameters are able to sufficiently reproduce the hydrograph of a catchment (e.g. Jakeman and Hornberger, 1993).

*Macropores and their influence on catchment-wide water balance simulations.*

A similar relationship applies to the observed diversity and density of the macropore network in the two study areas. I found that model parameters related to the number and location of macropores are fairly insensitive with respect to streamflow simulations as long as they are varied in a reasonable range. Contrastingly, perturbations in the topology of the preferential flow network significantly reduced the predictive performance of the models. Information about the exact location and number of preferential flow paths are hence only of minor importance if the goal is to make predictions on the catchment scale as long as the focus is on the water balance and not for instance on solute transport (Klaus and Zehe, 2011; Wienhöfer and Zehe, 2014). This should, however, not be misinterpreted in a way that preferential flow paths are unimportant, as model realizations without preferential flow network performed clearly worse than model realizations with perturbed networks. This model comparison shows that it is the information about the general occurrence as well as relative importance of macropores in a landscape which is relevant for a catchment modeler.

Apart from the relevance of macropores and soil water retention properties, storage volume of the bedrock as well as bud break of the dormant trees were key for successful simulations at the catchment scale. In both cases I used measurements rarely used in hydrological modeling, more specifically sap flow velocities and electrical resistivity tomography, to extract information about the dynamics of the vegetation and the bedrock topography. This highlights an advantage of the mixed top-down bottom-up model structure of representative hillslopes as the implementation and validation of these new measurements was straightforward.

*The role of bedrock topography and leaves sprout for catchment-wide water-balance simulations.*

The key findings of chapter 2 highlight that representative hillslopes are indeed a promising concept which needs to be further investigated in the future. While there are surely several limitations stemming from the chosen hydrological model or from the fact that representative hillslopes are necessarily spatially aggregated, it was shown that a single 2d bottom-up model can be used to represent an entire meso-scale catchment. The merits of such an approach are that it provides a possibility to merge the plot scale with the catchment scale and that it reflects the most elementary simplification of a hydrological model in which the distribution of potential energy along the flow path of a hilly landscape can be preserved. The derivation of representative relationships, as done for the representative soil water retention properties or the representative surface topography, is thereby the key for any successful modeling with a representative hillslope (see also Zehe et al., 2018).

*The concept of representative hillslopes.*

### 5.1.2 *Part III: The dynamic nature of hydrological similarity*

Hydrology has always been and still is a data scarce science (e.g. Beven, 2001). Most catchments around the world are ungauged and trustworthy information about the climatic forcing are rare. However, as already stated in section 5.1.1 this is changing slowly and the volume as well as frequency in which data is collected is constantly increasing. While this offers new and exciting opportunities for hydrological modeling -e.g. the concept of representative hillslopes- it also involves a series of new challenges, for example how to decide which of the information about the system under study is actually relevant for the targeted modeling purpose. In a few places around the world -mostly research catchments- this is already an issue and the amount but also the type of available data exceeds the capacity and flexibility of most hydrological models. If we want to use these data sources to improve and better constrain hydrological models we need to develop approaches and identify flexible model structures which are able to extract the relevant details about the system from these data sets. The second study (chapter 3) focuses on this quest.

*More information about the environmental system is available than ever before.*

The results of chapter 3 highlight that in our chosen research environment on average not more than six hillslope models are needed to produce essentially the same streamflow simulations as if 105 models were used. This strong reduction demonstrates that we are often overly optimistic about the sensitivity of our distributed models and how topographic differences and other landscape properties influence our simulations.

*The dynamic nature of hydrological similarity.*

The results show also that the identified number of six hillslopes is only a time average and that this number can vary between one and 30 in a hydrological year. This finding is somewhat counterintuitive at first glance as the 105 hillslope models differ only with respect to time-invariant topographic parameters, however, can be explained by the fact that the required amount of topographic information depends on the *current state of a landscape element* as well as on the *dominant hydrological processes related to this state*. For instance, after a convective rainfall event surface runoff might occur as a result of an exceeded infiltration capacity. This process is highly sensitive to the surface topography, the prevailing land-use and to the soil infiltrability. However, if time passed, a fraction of the rainfall might infiltrate and very different landscape controls such as the subsurface topography or the hydraulic conductivity will dominate runoff formation. Two hillslope models with different shapes can therefore behave similarly in certain conditions and differently in others. This means that the concept of hydrologically similarity cannot be time-invariant and that we need to identify hydrologically similar areas in a time-dependent manner. This means however also that in order to properly represent natural systems, hydrology models should also be able to account for this time-dependent similarity and adaptively adjust their model structure to be as complex as necessary but as parsimonious as possible (Savenije, 2010; Zehe et al., 2014).

*Data compression as tool to test the sensitivity of a distributed model to different data sources.*

The concept of analyzing model structures by means of data compression is a promising approach which was by no means fully exploited in this study. Especially the results of the compressed catchment model which consists of six instead of 105 hillslopes is encouraging although it is until now still a time-invariant concept. Especially the general approach to identify functional similar model elements by stepwise adding different sources of variability to a hydrological model and then analyzing the simulations by means of the Shannon entropy seems like an promising avenue to test the sensitivity of distributed hydrological models to different data sources. The results presented in chapter 3 have in general a large potential for further research which ranges from the identification of optimal measurement networks to model sensitivity approaches and finally to spatially adaptive modeling strategies (see also section 5.2.1).

### 5.1.3 *Part IV: A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation*

The third study in this thesis explores how straight-forward energy-centered arguments can be used to develop a topographic index (reduced Dissipation per unit length; rDUNE) to identify hydrologically similar landscape units in a forward mode (Seibert et al., 2017). The index is thereby an energy-centered re-interpretation and enhancement of the widely and successfully applied "*Height Above the Nearest Drainage*" index (HAND; Rennó et al., 2008) and based on the observation that most of the potential energy of water in a catchment is dissipated when rainfall becomes runoff.

*rDUNE a energy-centered re-interpretation of HAND.*

Our results show that rDUNE is capable to group catchments based on their surface topography into similar functional groups with respect to their runoff transformation. The performance of rDUNE proved thereby to be superior to a grouping based on the TWI or HAND. The results and discussion demonstrate that one meaningful way to derive similarity indices in hydrology is to capture both the driver and resistance term of a flux individually as proposed by Zehe et al. (2014). Furthermore, this study underpins that the extension and density of the channel networks as well as the average hillslope length are important factors for the classification of catchments as both variables store important information about the dissipation of potential energy when rainfall is transformed to runoff.

*Performance of rDUNE compared to the TWI and HAND.*

From a methodological point of view, this study shows once more that information theoretic measures, in this case the Jensen-Shannon divergence, are suitable tools to tackle a diversity of hydrological research problems. Furthermore, the energy-centered foundation of rDUNE opens up the possibly of a time-invariant grouping of hydrologically similar landscape units if rDUNE is calculated with a time series of precipitation. As highlighted in chapter 3 this could be the key to identify functionally similar units and to build spatially adaptive models. A promising way forward would hence be to test if rDUNE can be used to identify the six functional groups found in chapter 3 directly on the underlying topographic map. Finally, does the discussion about the role of energy dissipation in runoff generation open the research question if catchments have evolve in such a way that energy dissipation is maximized at the hillslope scale and minimize in the channel network and how this fact could be used to improve the way we do hydrological modeling.

*The role of energy dissipation for runoff generation.*

## 5.2 OUTLOOK

Following, I propose opportunities for future research which emerge from the key findings in chapters 2 - 4. More specifically, in subsection 5.2.1 I discuss the necessity of building spatially adaptive models in hydrology and discuss the merits of looking at the concept of landscape organization with a thermodynamic focus in subsection 5.2.2.

### 5.2.1 *The necessity of spatially adaptive models in the earth sciences*

*Perfection is achieved, not when there is nothing more to add, but when there is nothing left to be taken away. — Antoine de Saint-Exupéry*

A ”optimal” model represents a system in a manner that it balances necessary complexity with greatest possible parsimony (Savenije, 2010; Zehe et al., 2014). This means that no detail of a ”optimal” model can be taken away without losing relevant information in the simulations and no compartment can be added without producing redundancy in the simulations (Loritz et al., 2018). Simulations of a model are thereby only as reliable as adequately it represents the system under study (Dooge, 1973). As shown in chapter 3 catchments move from rather simple to complex states in time and space, a condition of a ”optimal” model is hence that it is able to adapt its internal structure to this change in a flexible way (Savenije, 2009). This rather theoretical perspective on model complexity shows that we need adaptive, (*flexible*) model approaches in Hydrology if we want to built models which balance complexity with parsimony. This has besides the theoretical merits also practical advantages.

#### *Practical reasons why adaptive models are needed in hydrology*

*The growing importance of intensity controlled processes in hydrology.*

A series of flash floods in south-west Germany (e.g. (Bronstert et al., 2017)) have highlighted the limitations of classical engineering concepts (e.g. unit hydrograph) with regard to predicting runoff responses resulting from high-intensity rainfalls. Especially when these high-intensity rainfall events are combined with dry soils, long-term rainfall-runoff relationships - the fundament of most empirical approaches in hydrology - do not represent the runoff formation of a catchment adequately. It is hence a long standing vision in hydrology to complement the empirical approaches used in operational flood forecasting with more physically-based approaches to improve our overall ability to predict the occurrence and magnitude of hydrological extremes.

In this regard, fully distributed bottom-up models like HydroGeoSphere (Brunner and Simmons, 2012) or MIKESHE (Refsgaard and Storm, 1995) are surely promising tools as they represent hydrological processes in a landscape in a state-of-the-art manner and are continuously refined and tested by a large community. However, bottom-up models are rarely used in operational forecasts as they are difficult to parameterize and rely on large amount of detailed spatial and temporal information about a landscape. Additionally, they demand large computational resources if applied to the scale of a catchment (>1 $km^2$). As already stated in chapter 3, data availability is slowly increasing, however, the high computation times for bottom-up models remain because the possibilities how fast we can calculate and parallelize certain numerical schemes are limited (Moore's law). For instance, Hopp and McDonnell (2009) used the state-of-the-art software package "HYDRUS 3D" (Šimůnek et al., 2016) and reported that they had computational simulation times ranging between 10 min up to 11 hrs when they simulated water fluxes and states at the Panola hillslope (area= 0.001250 $km^2$ (25 m x 50 m); maximal soil depths = 4 m; simulation time = 290 hrs.). As stated in the introduction, operational catchments usually range from 10 to 250 $km^2$. Applying bottom-up models in operational forecast, without violating the underlying physical foundation (Or et al., 2015), would hence result in a drastic increase of the simulation times, virtually preventing forward simulations.

*Computational limitation of bottom-up models.*

Recalling the findings of chapter 3 that only during a few short time periods during the year a higher spatial model complexity led to better simulations – as opposed to high redundancy in the other periods - we see that most of the time there is no need for a fully-distributed model to describe fluxes and state variables of a mesoscale catchments. This changes only as soon as we shift our attention to specific events where spatially explicit representation of the precipitation field and a distributed hydrological model are of considerable importance (Reid et al., 2005).

Following up on these findings it seems logical to use distributed rainfall and distributed hydrological models only in specific events when the length scale of the rainfall or of a specific geological control like the topography is indeed relevant. However, simply switching to the event scale is not a solution either as it is very difficult to approximate the initial conditions before an event given the degrees of freedom fully distributed models offer (e.g. Zehe et al., 2005). This is of considerable importance because particularly flash floods are highly sensitive to the actual state of the system such as current land-use management or the appearance of soil cracks.

*Uncertainties about the initial conditions make simulations on the event scale challenging.*

Adaptive modeling approaches are a way around this issue as they provide continuous hydrological simulations with higher model complexity and hence higher computational times only at time steps when they are actually needed. At least in theory building adaptive models is straightforward as there are only three main controls which influence the spatial complexity of a hydrological landscape.

*Identifying hydrological similarity with respect to the dominant processes.*

First, it is necessary to identify hydrological similar landscape units with respect to the hydrological processes which are dominant during different states. For instance, if surface runoff under high intensity rainfall is the main reason we need to distribute our model in space, we also should distribute our model based on hydrological units which reflect important controls of the surface runoff like the topography or the land-use. The first issue we need to address before we setup a adaptive model is hence: *Which processes dominate times of high computational needs and how can we identify corresponding hydrologically landscape units?* This question could be answered with an approach similar to the one presented in chapter 3.

*Spatial structure of the climatic forcing.*

Second, and maybe most obvious, the spatial structure of the forcing, for instance the precipitation field over a catchment, determines if a distributed hydrological model is needed or not. Simply speaking, if the forcing over a catchment is spatially homogeneous and the different landscape units within the catchment are in a similar state and share a comparable structure, there is no reason to assume that a landscape will react differently. It is only when the spatial structure of the forcing reaches a certain diversity it will cause diverse hydrological responses. This leaves us with the second issue we need to address if we want to build adaptive hydrological models: *When is the spatial variation of the climatic forcing large enough that it needs to be considered in hydrological simulations?*

*Hydrological memory of a landscape.*

The third control deals with the question how fast gradients which drive runoff generation get dissipated. For instance, two structurally similar hillslopes in a different state will react differently if they receive the same precipitation forcing. However, larger gradients get dissipated faster than smaller gradients in structurally similar systems, hence the two hillslopes will return to a similar state after a certain often surprisingly short period. The duration of this period is related to how fast the internal gradients which drive water flow of the hillslopes are dissipated. If we translate this finding to the model world we can assume that after an average dissipation time two structurally similar landscape elements which were forced differently can again be treated as one model element. Consequently, the second issue we need to address when building an adaptive hydrological model is: *How long do structurally similar landscape elements memorize a contrasting forcing in their distributed states and gradients?*

*Adaptive models - a way forward.*

To summarize, building spatially adaptive hydrological models means to i.) identify potentially hydrologically similar areas in a landscape, ii.) identify thresholds when differences in the climatic forcing matter and iii.) derive relationships which connect the dissipation of gradients with an average system memory to infer when two landscape units behave similarly again. All three issues are beyond the scope of this thesis. However, the methods and concepts taken from thermodynamics and information theory I use in my analyses are a suitable framework to build adaptive hydrological models.

### 5.2.2 *Maximum energy dissipation at the hillslope scale vs. minimum energy dissipation in the channel network*

The majority of the potential energy is dissipated along the flow path when rainfall is transformed to runoff (chapter 4). As this phenomenon has been observed in a wide range of catchments around the world (flow velocities rarely exceed $1\,\mathrm{m\,s^{-1}}$ despite quite large topographic differences; e.g. Leopold and Maddock, 1953), it opens room for speculations, and one might wonder if catchments have evolved to a state where energy dissipation at the hillslope is maximized. The latter contradicts however somewhat with the observation that energy dissipation in preferential flow networks such as rivers or macropores is often found to be minimized (Kleidon et al., 2013; Rinaldo et al., 1992; Rodríguez-Iturbe et al., 1992; Zehe et al., 2010). This discrepancy between minimizing and maximizing energy dissipation, however, only persist as long as we focus exclusively on energy dissipation and can be resolved if the perspective is shifted to the fraction of free energy which is available to perform work. In the following I will shortly explain the underlying idea behind the concept of minimizing work and will stress possibilities for further research.

*Free energy to perform work. A framework to describe the evolution of flow structures in a landscape?*

The configuration of hydrological structures like channel networks or hillslopes are often regarded as being in quasi-steady-state at least if view on hydrological timescales (e.g. Langbein and Leopold, 1964; Zehe et al., 2014). This does not mean that hydrological systems are in any sense stationary but that catchments are organized in a manner that times of significant change are linked to certain extremes in mass and energy inputs.

The concept of a quasi-steady-state has been underpinned by a series of studies founded on large data sets (e.g. Leopold and Maddock, 1953) which showed that rather simple mathematical, time-invariant, empirical power laws can be used to extract drainage network densities (Howard, 1990), predict mean discharges in relation to average channel slopes (Hack et al., 1957) or to identify discharges where critical erosion of the river bed is to be expected (Pfeiffer et al., 2017). The possibility to identify these simple mathematical relationships and link function and structure on long timescales gives rise to the research question if catchments around the world have evolved to a comparable state with respect to how they transfer, store and exchange energy and mass with their environment.The idea behind the concept of landscape evolution should thereby not be mistaken with an esoteric idea of a system wanting to be in a certain way but follows a simple physical reasoning.

*Limited degree of freedom to adapt to change.*

Hydrological systems have, despite the large diversity in their climatic and geological setting, limited possibilities to adapt their internal structure to incoming flows of energy and mass. This stems, once more, from the second law of thermodynamics which states in a somewhat abstract form that any flux is a ratio of a *driving potential* and a *resistance term* working against this flux (explained in more detail in chapter 4). To transfer this rather theoretical concept to landscape evolution lets assume that we increase the total volume of the long-term annual precipitation in a catchment. This means subsequently that we also increase the potential energy differences driving runoff generation in this catchment, simply because more mass at the same potential means also more energy. If we assume further that on longer timescales the larger potential energy differences do not result in an increased export of kinetic energy we can expect that the additional energy needs to be consumed within the system boundaries. This is done by performing work on the catchment structure by i) either changing the distribution of topographic potentials (reducing the average hillslope height by erosion) or ii) by creating or extending flow structures such as gullies or preferential flow network paths. The newly created structures can thereby only persist if free energy is continuously invested in their maintenance meaning that they on one hand reduce friction but on the other hand also need a constant investment of work to endure. Both processes, the reduction of the topographic potential and the creation or rearrangement of new flow structures have in common that eventually the additional added amount of energy, through an increased annual rainfall, is again consumed within the system borders.

To summarize, the assumption that hydrological systems have adapted to a certain long-term climate forcing is not far-fetched as the amount of work which is available to perform work is minimized when new flow structures are created or when potentials are eroded (Langbein and Leopold, 1964). A testimony of this line of thoughts is that at least in humid regions we only observe significant modifications on the catchment structure if we either alter the boundary conditions (e.g. climate change) and thereby increase or decrease the average energy input to a system or artificially alter dominant structures (e.g. deforestation, agriculture) and thereby invest energy to move a system away from its local equilibrium.

Besides being an intellectually interesting and challenging concept landscape organization has also practical implications for hydrology. For instance, if we assume that the structure of a landscape has adapted to a certain long-term climatic forcing we should be able to approximate a certain range of energy and mass input to the system beyond which we will not expect large erosion events or discharge outside the river bed. The idea of a local equilibrium could then be used to underpin extreme value statistics in hydrology with a physical theory instead of focusing exclusively on statistical properties of fitted distribution functions. This would mean a general shift in perspective as the focus would move to identifying of a local equilibrium state first and only after this specify if an event can be considered extreme and not vice versa.

*Energy and work as foundation of extreme value statistics in hydrology.*

## 5.3 SYNTHESIS

In the following section 5.3.1 I reviewing the connection between information theory and thermodynamcis and explain why both scientific frameworks together provide a set of tools and general mindset which is perfectly suited to improve the theoretical underpinning of hydrological models (Clark et al., 2016), a major goal of this thesis. In section 5.3.2, I finally conclude that information theory, thermodynamics and classical hydrological modeling concepts together provide a powerful combination of theoretical and practical approaches which could serve as foundation to tackle the challenges hydrology is facing in the 21$^{th}$ century.

### 5.3.1 *The relation between information and energy*

A lot has been written about the question whether or not there is a relationship between information theory and thermodynamics (e.g. Ben-Naim, 2008; Jaynes, 1957; Koutsoyiannis, 2014; Shannon, 1948). This discussion is certainly caused by Claude Shannon naming his measure of uncertainty "entropy", thereby borrowing a term typically linked to thermodynamics. Although thermodynamic and information entropy were introduced for very different reasons, the name entropy was not chosen by coincidence.

*The name entropy was proposed by von Neumann in a letter to Claude Shannon.*

It is the similarity of the mathematical formulas of the thermodynamic (*Gibbs entropy*) and information entropy (*both definitions differ only by a constant factor meaning that they are perfectly linear dependent*) that led Shannon to name his metric entropy. The similarity is further highlighted by the fact that originally the Gibbs entropy did not contain the Boltzmann constant. This was added later by Planck and only then entailed the units $\mathrm{J\,K^{-1}}$.

*The probabilistic nature of entropy*

However, the similarities between the two metrics go beyond the mathematical similarity as also the questions they were designed to answer are akin to a certain extent. The Gibbs entropy is applied to answer the question how probable a given microscopic state is on average, given a certain observed macroscopic state. The Shannon entropy describes how probable a given event will occur on average, given an underlying probability distribution. In thermodynamics the microstate is typically a probability of a certain kinetic energy of a single molecule and the macrostate the observed temperature distribution of a cylinder filled with a gas. In the information case the microstate is for instance the probability of a given letter and the macrostate the entire underlying probability distribution of all letters in a chosen language. Both thermodynamic and information entropy are therefore united by their probabilistic nature as they link the likeliness of a microstate with a given macrostate and hence the certainty of our knowledge about a given microstate given knowledge about the macrostate of a system.

It is important to recall that the objective of this thesis is not to pursue the ongoing discussion about whether or not information theory is a child of thermodynamic reasoning but to examine how we can improve hydrological predictions at the catchment scale. The important point here is that the connection between information theory and thermodynamics, established on purpose or not, gives us a set of powerful tools, like the maximization of entropy (Jaynes, 1957), to do statistical inference which can be straightforwardly linked to a deeper physical understanding.

For example, to calculate information theoretic metrics using real-world data it is often necessary to bin the data into similar groups, thereby reducing precision, estimate probabilities and subsequently work with discrete probability distributions. This process is frequently criticized and interpreted as a weakness of information theory as the choice of the bin width can strongly influence the outcome of an investigation. However, from a physical perspective a meaningful choice of the bin width is rather an advantage than a weakness. It forces the researcher to define the term similarity clearly and specific to the research question before doing any statistical inference or modeling. This means, for instance, that the point when a given model cannot be further improved or two observations are indistinguishable needs to be defined a priori. A clear definition of similarity – as opposed to sharing residual model errors – is thereby the key to generalize findings beyond the boundaries of a research environment.

Another concept taken from information theory with a large potential for hydrological research is the field of data compression. As shown in chapter 3 the ideas and tools behind this concept can be used to investigate questions related to model complexity or to the optimal size of a measurement network by identifying and minimize redundancy in data. The methods are thereby founded on an entire research field called algorithmic information theory (Cover and Thomas, 2005) which includes interesting methodical approaches as well as an underlying philosophy which has barely been explored in environmental science (Weijs et al., 2013b). The key findings of chapter 3 highlight, however, the merits of further investigating in this direction also because most environmental models are computer programs and (algorithmic) information theory provides the natural framework to analyze their structure and compare models in a sophisticated way.

*Data compression as tool to identify hydrological similarity*

Another advantage of information theory is that a large part of the statistical concepts used today were developed in times where no or little computer power was available and are thus often based on simplifications which might seem unnecessary from today's point of view. Information theory, on the other hand, evolved hand in hand with the development of modern computers and many of the typical simplifications and requirements for statistical methods (e.g. normally distributed datasets) do not need to be fulfilled for information theoretical analyses. Therefore, the approaches from information theory are from a statistical point of view essentially assumption free, apart from the choice of a bin width (e.g. Cover and Thomas, 2005). This means that it is unnecessary to assume a priori that a data set follows a specific theoretical distribution and work with statistical moments like the variance which are only meaningful if the data is normally distributed. For instance, in chapter 4 I demonstrate that the Jensen-Shannon divergence could be used to distinguish visually dissimilar distributions without any assumptions besides the chosen bin width. The latter is obviously only an advantage if the sample distribution underlying a data set can be estimated in a robust way.

*Information theory a modern mindset to do statistical inference in the computer age.*

Concepts like mutual information to identify non-linear correlations between data sets, the Kullback-Leibler divergence to examine for instance the performance of model ensembles (e.g. Weijs et al., 2013a), or the "distance to the maximum entropy" method to establish a link between Bayesian inference and thermodynamic optimality (Jaynes, 1957) are promising tools for different hydrological research questions which highlight the potential of information theory.

### 5.3.2 *Hydrology beyond closing the mass balance*

Catchment hydrology is - somewhat logically - a field of research with a strong focus on the estimation of quantitative water amounts. This emphasis on the water balance is already implicit to the characteristic research unit, the catchments, a concept founded essentially on mass conservation. It is hence not surprising that a large fraction of the available hydrological catchment models are fundamentally based on a closed water balance while other conservation laws like energy conservation or momentum conservation are often neglected (*together with mass conservation the three build the foundation of describing fluid flow*). The issue with this procedure is that the amount of water stored and released by a catchment alone is a weak proxy for the state of a catchment. As long as we are hence focusing exclusively on the mass balance we are unable to identify an equilibrium state which serves as an attractor where all relevant driving potentials for water flow are close to depletion. This means that fundamentally, a hydrological system like a catchment can only be described exhaustively if both, mass and energy, are considered equally (Zehe et al., 2018).

Many models and concepts in hydrology compensate their lack of an appropriate physical system description by deriving empirical relationships founded on long-term observations of hydrological fluxes. It is the beauty but also the pitfall of modern hydrology that finding these empirical relationships is somewhat straightforward and often works surprisingly well. The stunning performance of these mathematically simple models to mimic the hydrological function of a diverse catchment explains, at least partly, the bias of hydrological research towards statistical concepts like optimization and parameter estimation. Patterns and processes we observe in hydrology are, however, by no means a result of randomness but highly organized and fitting models to data will always mean that we will have limited success to predict future states under instationary conditions.

One way to eventually improve our ability to make predictions at the catchment scale is hence to re-visit the theoretical foundation of our models and statistical concepts (Clark et al., 2016; Dooge, 1986; Kirchner, 2006; Sivapalan, 2003; Zehe et al., 2014). As catchments belong to a category of systems referred to as "organized complexity" this avenue can only be successful if we combine theories which improve the *physical foundation* of hydrology with theories helping us to improve the way we do *statistical inference*.

In this thesis, I chose information theory and thermodynamics and therefore the concept of information and energy as theoretical foundations to conduct research. In combination with established hydrological concepts and models I was able to shed new light on well-known hydrological research questions ranging from the identification of hydrological similar landscape units to the development of hydrological model concepts for the meso-scale. My results indicate that the combination of information, energy and mass could serve as a general scientific scheme for developing a hydrological theory for the catchment scale. This theoretical foundation is one prerequisite if we want to better manage the challenges hydrology is facing in the $21^{th}$ century and in general improve our ability to do predictions in hydrology.

# Part VI

# APPENDIX

# A APPENDIX

## A.1 APPENDIX CHAPTER II

*A 1.1: Subsurface structure and bedrock topography*

Spatial subsurface information of representative hillslopes were obtained from 2d ERT sections collected using a GeoTom (GeoLog) device at seven profiles on two hillslopes in the Colpach catchment. We used a Wenner configuration with electrode spacing of 0.5 m and 25 depth levels: electrode positions were recorded at a sub-centimeter accuracy using a total station providing 3d position information. Application of a robust inversion scheme as implemented in Res2Dinv (Loke, 2003) resulted in the two-layered subsurface resistivity model shown in Fig. 2.6b. The upper 1-3 m are characterized by high resistivity values larger than 1500 $\Omega$ m. This is underlain by a layer of generally lower resistivity values smaller than 1500 $\Omega$ m. In line with the study of Wrede et al. (2015) and in correspondence with the maximum depth of the local auger profiles, we interpreted the transition from high to low resistivity values to reflect the transition zone between bedrock and unconsolidated soil. In consequence, we regard the 1500 $\Omega$ m isoline as being representative for the soil-bedrock interface. For our modeling study we have access to seven ERT profiles within the Colpach area (example see Fig. 2.6b).

*A 1.2: Soil hydraulic properties, infiltrability and dye staining experiments*

Saturated hydraulic conductivity was determined with undisturbed 250 ml ring samples with the KSAT apparatus (UMS GmbH). The apparatus records the falling head of the water supply through a highly sensitive pressure transducer which is used to calculate the flux. The soil water retention curve of the drying branch was measured with the same samples in the HYPROP apparatus (UMS GmbH) and subsequently in the WP4C dew point hygrometer (Decagon Devices Inc.). The HYPROP records total mass and matric head in two depths in the sample over some days when it was exposed to free evaporation (Jackisch, 2015; Peters and Durner, 2008 for further details). For both geological settings we estimated a mean soil retention curve by grouping the observation points of all soil samples (62 and 25 for schist and marl, respectively), and averaging them in steps of 0.05 pF. We then fitted a van Genuchten-Mualem model using a maximum likelihood method to these averaged values (Tab. 2.1 and Fig. 2.7). We used a representative soil water retention curve because the young soils on periglacial slope deposits prevail in the both headwaters exhibit large heterogeneity which cannot be grouped in a simple manner. This is due to a) the general mismatch of the scale of 250 mL undisturbed core samples with the relevant flow paths and b) the high content of gravel and voids, which affect the retention curve especially above field capacity and concerning its scaling with available pore space (Jackisch, 2015; Jackisch et al., 2016). The dye tracer images, Fig. 2.2 b and d, were obtained with high rainfall intensities of 50 mm in 1 h on 1 $m^2$ and the sprinkling water was enriched with 4.0 g $1^{-1}$ Brilliant Blue dye tracer (Jackisch et al., 2016). The aim of these rainfall simulations was to visualize the macropore networks in the topsoil, to gather information on the potential preferential flow paths relevant for infiltration.

### *A 1.3: Physically-based model CATFLOW*

The model CATFLOW has been successfully used and specified in numerous studies (e.g. Wienhöfer and Zehe, 2014; Zehe et al., 2005, 2010). The basic modeling unit is a two-dimensional hillslope. The hillslope profile is discretized by curvilinear orthogonal coordinates in vertical and downslope directions; the third dimension is represented via a variable width of the slope perpendicular to the slope line at each node. Soil water dynamics are simulated based on the Richards equation in the pressure based form and numerically solved using an implicit mass conservative "Picard iteration" (Celia et al., 1990). The model can simulate unsaturated and saturated subsurface flow and hence has no separate groundwater routine. Soil hydraulic functions after van Genuchten-Mualem are commonly used, though several other parameterizations are possible. Overland flow is simulated using the diffusion wave approximation of the Saint-Venant equation and explicit upstreaming. The hillslope module can simulate infiltration excess runoff, saturation excess runoff, re-infiltration of surface runoff, lateral water flow in the subsurface as well as return flow. For catchment modeling several hillslopes can be interconnected by a river network for collecting and routing their runoff contributions, i.e. surface runoff or subsurface flow leaving the hillslope, to the catchment outlet. CATFLOW has no routine to simulate snow or frozen soil.

#### *A 1.3.1 Evaporation controls, root water uptake and vegetation phenology*

Soil evaporation, plant transpiration and evaporation from the interception store is simulated based on the Penman–Monteith equation. Soil moisture dependence of the soil albedo is also accounted for as specified in Zehe and Flühler (2001). Annual cycles of plant pheno-

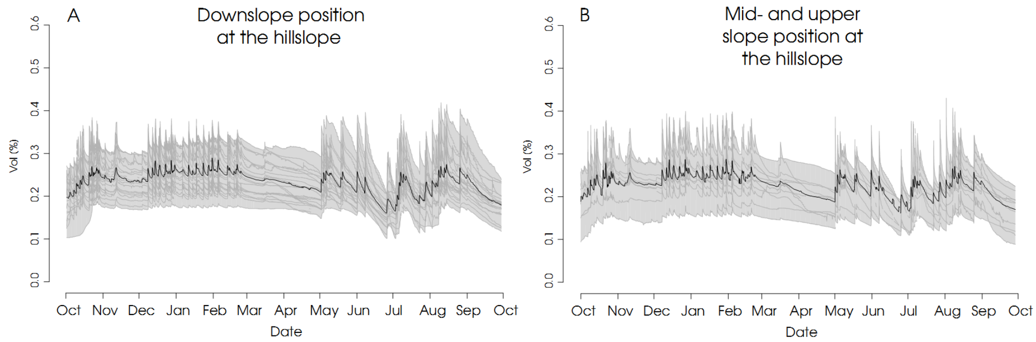


Figure A.1: Soil moisture observations grouped by their landscape position. (A) Soil moisture observations at the hillslope foot and hence close to the river. (B) Soil moisture observations at the upper part of the hillslope.

logical parameters, plant albedo and plant roughness are accounted for in the form of tabulated data (Zehe and Flühler, 2001). Optionally, the impact of local topography on wind speed and on radiation may be considered, if respective data are available. The atmospheric resistance is equal to wind speed in the boundary layer over the squared friction velocity. The former depends on observed wind speed, plant roughness and thus plant height. The friction velocity depends on observed wind speed as well as atmospheric stability, which is represented through six stability classes depending on prevailing global radiation, air temperature and humidity. The canopy resistance is the product of leaf area index and leaf resistance, which in turn depends on stomata and cuticular resistance. The stomata resistance varies around a minimum value, which depends on the Julian day as well as on air temperature, water availability in the root zone, the water vapor saturation deficit and photosynthetic active radiation (Jarvis, 1976). The resulting root water uptake is accounted for as a sink in the Richards equations term using a soil water dependent root extraction function (Feddes et al., 1976), and is specified as a flux per volume, which is extracted uniformly along the entire root depth.

### *A 1.4 Soil moisture observations*

Fig. A.1 shows the soil moisture observations of the Colpach catchment group by their position at the hillslope. This figure highlight, similar to Fig. 2.7 for the soil water retention properties, that the small-scale variability of the prevailing soils make a simple grouping by the landscape position difficult.

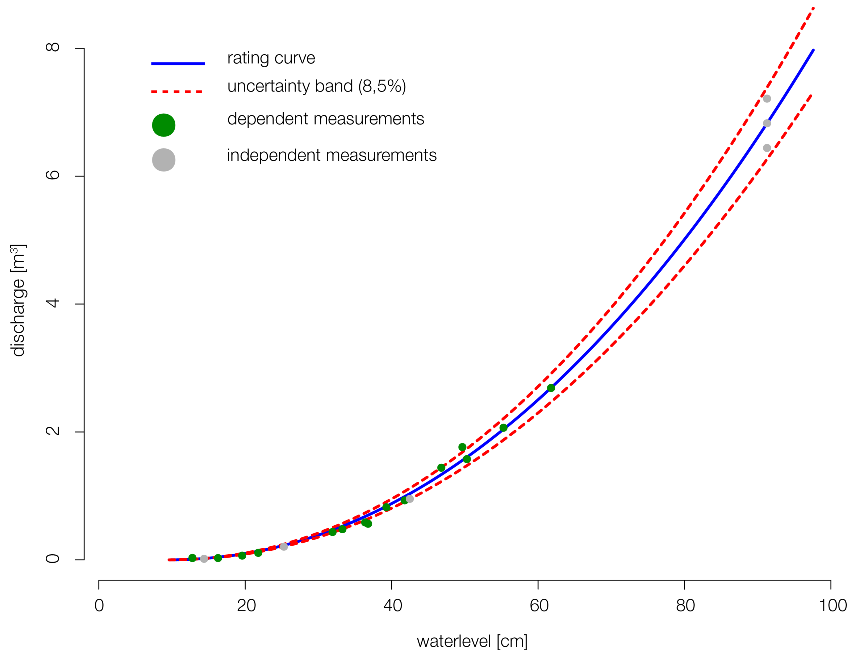


Figure A.2: Rating curve of the Colpach gauge. Green dots which were used to estimate the rating curve, gray dots independent discharge measurements.

## A.2 APPENDIX CHAPTER III

### *A 2.1: Uncertainty of the rating curve*

For the gauge "Colpach" the rating curve was given with:

$$Q = 10.59(h - 0.11)^{2.14} \tag{A.1}$$

where Q is discharge ($m^3\,s^{-1}$) and h is gauge level (m). It was derived by ordinary least square fitting to 15 direct discharge measurements (Fig. A.2 green dots). Using the rating curve for flood frequency analyses would require a validation against an independent set of direct discharge measurements (grey dots). In order to we use it as proxy for the binning width to estimate the pdfs, we calculated its overall uncertainty relative to the total set of direct discharge measurements (green and grey dots) as RMSE with a value of 8.5 % (dashed red line).

### *A 2.2: Uncertainty of the rating curve*

In Fig. A.3 we illustrate the influence of different bin widths when calculating the Shannon entropy of our discharge simulations as function of time. We start as already described in section 3.3.1 with a discharge value of 0.01 mm and then progressively increase the bin width by

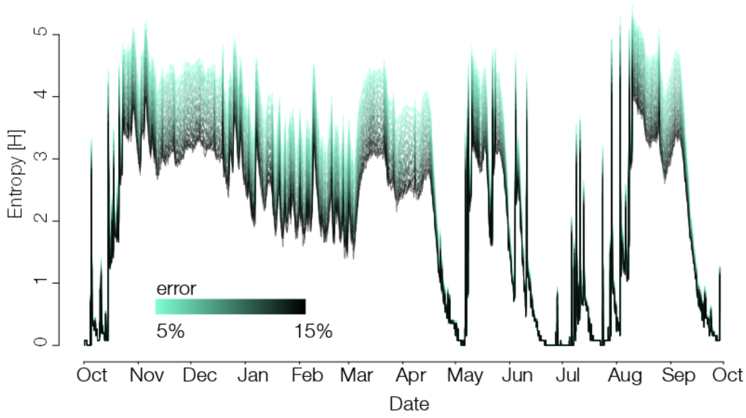


Figure A.3: Influence of the bin width.

factors ranging from 5 % to 15 % in 0.05 % steps. This graph highlights that the absolute value of the Shannon entropy depends strongly on the chosen binning size. However, more important for this study is that the overall pattern of the Shannon entropy in time does not change depending on the chosen bin size.

*A 2.3: Comparison of the NMI*

To illustrate the performance of this metric, Fig. A.4 shows a comparison of normalized mutual information (NMI) to the Pearson correlation and the Euclidean distance for four different synthetic cases:

- linear relationship between X and Y
- difference between two sinusoidal functions with different amplitudes
- quadratic relationship between X and Y
- two independent random variables X and Y

We used equally distant bin widths of 0.05 to estimate the pdf for the calculation of the mutual information in all four cases.

*A 2.4: Shannon entropy of the runoff simulations against the median discharge of the runoff simulations*

Relation between the area-weighted median of the discharge simulation against the Shannon entropy of all discharge simulations for each time step (Fig. A.5). The graph highlights that there is no simple linear relation between discharge height, time of the year and the Shannon entropy.

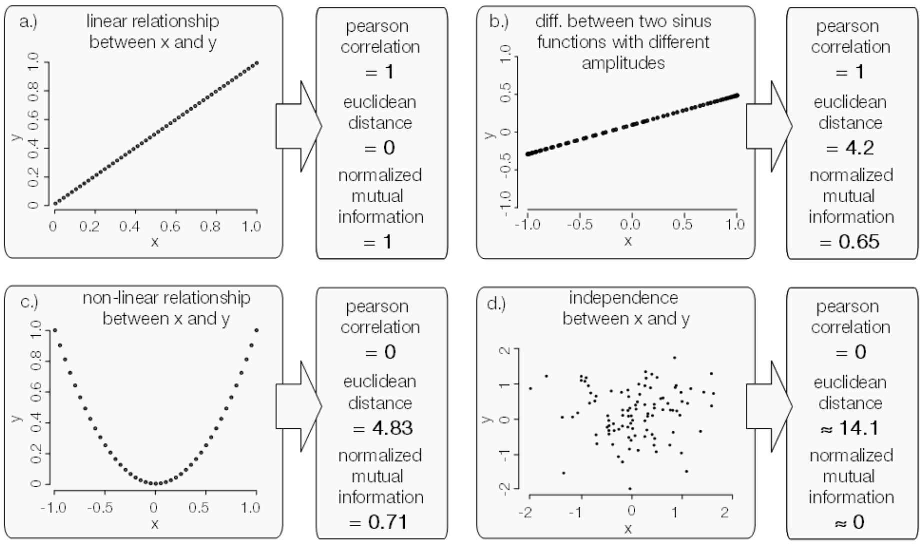


Figure A.4: Difference between the Pearson correlation coefficient, Euclidean distance and the normalized mutual information. Four cases are shown (a) linear relationship, (b) the difference between two sinus functions with different amplitude, (c) a quadratic relationship and (d) two independent variables. The pdf was estimated using an equally distant bin width of 0.05 in all four cases.

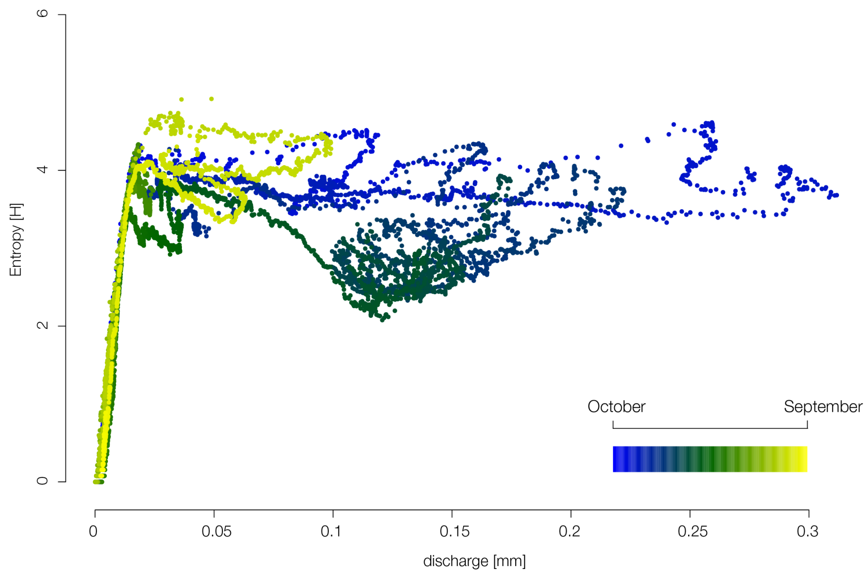


Figure A.5: Shannon entropy of the 105 discharge simulations against the area-weighted median of the discharge simulations. The color key range from blue (winter) over green (autumn / spring) to yellow (summer) and illustrates the time of the year.

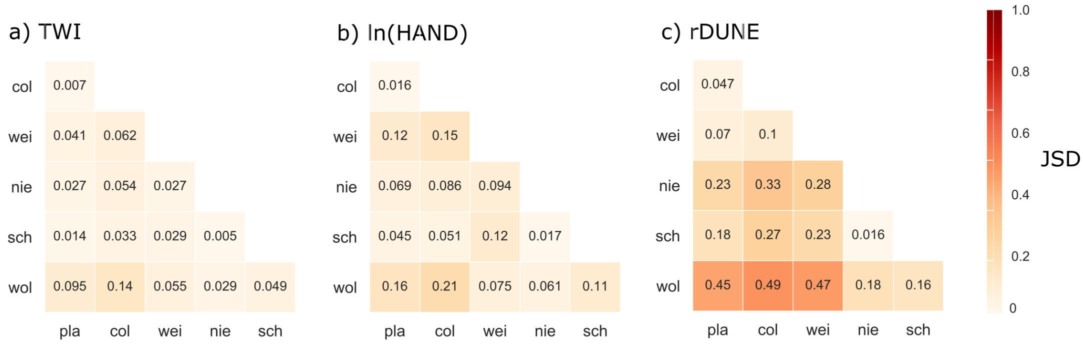


Figure A.6: JSD values for the six research catchments (Schist: Platen (pla), Colpach (col), Weierbach (wei); Marl: Niederpallen (nie), Schwebich (sch), Wollefsbach (wol) ). Panel a JSD of between the TWI distributions, b between the ln(HAND) distributions and c between the rDUNE distributions. A high JSD value indicates a high divergence between the distributions with a maximum of 1. The difference between this figure and Fig. 4.4 is the chosen bin width when we estimated the JSD between the different distributions.

## A.3 APPENDIX CHAPTER IV

### *A 3.1: Influence of different bin widths on the Jensen-Shannon divergence*

In Fig. A.6 we illustrate the influence of a different bin width when calculating the Jensen-Shannon divergence between the TWI, ln(HAND) and rDUNE distributions. Instead of using the largest bin width as described in Sect. 4.2.2 we use the smallest meaningful bin width which is 0.1 for the TWI, 0.03 for ln(HAND) and 0.05 for rDUNE. This figure in comparison to Fig. 4.4 highlights that the overall picture does persists even if we would have chosen the smallest statistical feasible bin width instead of the largest.

Angermann, Lisa et al. (2017). "Form and function in hillslope hydrology: characterization of subsurface flow based on response observations." In: *Hydrology and Earth System Sciences* 21.7, pp. 3727–3748. DOI: 10.5194/hess-21-3727-2017.

Applebaum, David (1996). *Probability and Information*. First. Cambridge University Press, pp. 105–126.

Arnaud, Patrick, Christophe Bouvier, Leonardo Cisneros, and Ramon Dominguez (2002). "Influence of rainfall spatial variability on flood prediction." In: *Journal of Hydrology* 260.1-4, pp. 216–230. DOI: 10.1016/S0022-1694(01)00611-4.

Babcock, K. L. and Roy Overstreet (Oct. 1955). "Thermodynamics of soil moisture: A new application." In: *Soil Science* 80.4, pp. 257–263. DOI: 10.1097/00010694-195510000-00002.

Bachmair, Sophie and Markus Weiler (2011). *Forest Hydrology and Biogeochemistry*. Ed. by Delphis F. Levia, Darryl Carlyle-Moses, and Tadashi Tanaka. Vol. 216. Ecological Studies 1. Dordrecht: Springer Netherlands, p. 455. DOI: 10.1007/978-94-007-1363-5.

Bárdossy, A. (2007). "Calibration of hydrological model parameters for ungauged catchments." In: *Hydrology and Earth System Sciences* 11.2, pp. 703–710. DOI: 10.5194/hess-11-703-2007.

Barling, Rowan D., Ian D. Moore, and Rodger B. Grayson (Apr. 1994). "A quasi-dynamic wetness index for characterizing the spatial distribution of zones of surface saturation and soil water content." In: *Water Resources Research* 30.4, pp. 1029–1044. DOI: 10.1029/93WR03346.

Bear, Jacob (1972). *Dynamics of Fluids in Porous Media*. American Elsevier, New York.

Ben-Naim, Arieh (2008). *A Farewell to Entropy*. World scientific. DOI: 10.1142/6469.

Berghuijs, Wouter R., Murugesu Sivapalan, Ross A. Woods, and Hubert H. G. Savenije (2014). "Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales." In: *Water Resources Research* 50.7, pp. 5638–5661. DOI: 10.1002/2014WR015692. eprint: 2014WR016527 (10.1002).

Bergström, Sten and Aren Forsman (1973). "Development of a conceptual deterministic rainfall-runoff-model." In: *Hydrology Research* 4.3, pp. 147 –170.

Berne, A., R. Uijlenhoet, and P. A. Troch (2005). "Similarity analysis of subsurface flow response of hillslopes with complex geometry." In: *Water Resources Research* 41.9, pp. 1–10. DOI: 10.1029/2004WR003629.

Beven, K. and M. J. Kirkby (1979). *A physically based, variable contributing area model of basin hydrology*. DOI: 10.1080/02626667909491834.

Beven, Keith J. (2000). "Uniqueness of place and process representations in hydrological modelling." English. In: *Hydrology and Earth System Sciences* 4.2, pp. 203–213. DOI: 10.5194/hess-4-203-2000.

Beven, Keith (1989). "Changing ideas in hydrology — The case of physically-based models." In: *Journal of Hydrology* 105.1-2, pp. 157–172. DOI: 10.1016/0022-1694(89)90101-7.

Beven, Keith (2001). "Dalton Lecture: How Far Can We Go In Distributed Hydrological Modelling? Keith Beven Lancaster University." In: *Computer* 5, pp. 1–12.

Beven, Keith (2006a). "A manifesto for the equifinality thesis." In: *Journal of Hydrology* 320.1-2, pp. 18–36. DOI: 10.1016/j.jhydrol.2005.07.007.

Beven, Keith (2006b). *Streamflow Generation Processes: Introduction*. Ed. by K. J. Beven. 1st edtion. University of Lancaster, Lancaster, UK, p. 10.

Beven, Keith and Jim Freer (2001). "A dynamic topmodel." In: *Hydrological Processes* 15.10, pp. 1993–2011. DOI: 10.1002/hyp.252.

Beven, Keith and Peter Germann (1982). "Macropores and water flow in soils." In: *Water Resources Research* 18.5, pp. 1311–1325. DOI: 10.1029/WR018i005p01311.

Beven, Keith and Peter Germann (2013). "Macropores and water flow in soils revisited." In: *Water Resources Research* 49.6, pp. 3071–3092. DOI: 10.1002/wrcr.20156.

Beven, Keith and Peter Young (2013). "A guide to good practice in modeling semantics for authors and referees." In: *Water Resources Research* 49.8, pp. 5092–5098. DOI: 10.1002/wrcr.20393.

Bishop, Joshua M., Michael V. Callaghan, Edwin E. Cey, and Larry R. Bentley (2015). "Measurement and simulation of subsurface tracer migration to tile drains in low permeability, macroporous soil." In: *Water Resources Research* 51.6, pp. 3956–3981. DOI: 10.1002/2014WR016310.

Blöschl, Günther and Murugesu Sivapalan (1995). "Scale issues in hydrological modelling: A review." In: *Hydrological Processes* 9.3-4, pp. 251–290. DOI: 10.1002/hyp.3360090305.

Bos, R. van den, L Hoffmann, J Juilleret, P Matgen, and L Pfister (1996). "Conceptual modelling of individual HRU ' s as a trade-off between bottom-up and top-down modelling , a case study ." In: *Conf. Environmental Modelling and Software. Proc. 3rd Biennal meeting of the international Environmental Modelling and Software Society. Vermont, USA.*

Breuer, Lutz, Klaus Eckhardt, and Hans-Georg Frede (2003). "Plant parameter values for models in temperate climates." In: *Ecological Modelling* 169.2-3, pp. 237–293. DOI: 10.1016/S0304-3800(03)00274-6.

Bronstert, A. et al. (2017). "The Braunsbach Flashflood of Mai 29th, 2016 – Origin, Pathways and Impacts of an Extreme Hydro-Meteorological Event. Part 1: Meteorological and Hydrological Analysis." In: *Hydrologie und Wasserbewirtschaftung* 61, pp. 150–162. DOI: 10.5675/HyWa_2017,3_2.

Bronstert, Axel and Erich J. Plate (1997). "Modelling of runoff generation and soil moisture dynamics for hillslopes and micro-catchments." In: *Journal of Hydrology* 198.1-4, pp. 177–195. DOI: 10.1016/S0022-1694(96)03306-9.

Brunner, Philip and Craig T. Simmons (2012). "HydroGeoSphere: A Fully Integrated, Physically Based Hydrological Model." In: *Ground Water* 50.2, pp. 170–176. DOI: 10.1111/j.1745-6584.2011.00882.x.

Brunsell, N. A. (2010). "A multiscale information theory approach to assess spatial-temporal variability of daily precipitation." In: *Journal of Hydrology* 385.1-4, pp. 165–172. DOI: 10.1016/j.jhydrol.2010.02.016.

Burgess, S. S. et al. (2001). "An improved heat pulse method to measure low and reverse rates of sap flow in woody plants." In: *Tree physiology* 21.9, pp. 589–598. DOI: 10.1093/treephys/21.9.589.

Camporese, M., C. Paniconi, M. Putti, and S. Orlandini (2010). "Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data." In: *Water Resources Research* 46.2. DOI: 10.1029/2008WR007536.

Castiglione, Paolo et al. (2003). "Lateral Water Diffusion in an Artificial Macroporous System: Modeling and Experimental Evidence." In: *Vadose Zone Journal* 2.2, pp. 212–221. DOI: 10.2113/2.2.212.

Celia, Michael A., Efthimios T. Bouloutas, and Rebecca L. Zarba (1990). "A general mass-conservative numerical solution for the unsaturated flow equation." In: *Water Resources Research* 26.7, pp. 1483–1496. DOI: 10.1029/WR026i007p01483.

Chaubey, I., A. S. Cotter, T. A. Costello, and T. S. Soerens (2005). "Effect of DEM data resolution on SWAT output uncertainty." In: *Hydrological Processes* 19.3, pp. 621–628. DOI: 10.1002/hyp.5607.

Clark, Martyn P. and Dmitri Kavetski (2010). "Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes." In: *Water Resources Research* 46.10. DOI: 10.1029/2009WR008894.

Clark, Martyn P. et al. (2016). "Improving the theoretical underpinnings of process-based hydrologic models." In: *Water Resources Research* 52.3, pp. 2350–2365. DOI: 10.1002/2015WR017910.

Coenders-Gerrits, A. M J, L. Hopp, H. H G Savenije, and L. Pfister (2013). "The effect of spatial throughfall patterns on soil moisture patterns at the hillslope scale." In: *Hydrology and Earth System Sciences* 17.5, pp. 1749–1763. DOI: 10.5194/hess-17-1749-2013.

Cover, Thomas M. and Joy A. Thomas (2005). *Elements of Information Theory*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1–748. DOI: 10.1002/047174882X. eprint: ISBN0-471-06259-6.

Crave, A. and C. Gascuel-Odoux (1997). "The Influence of topography on time and space distribution of soil surface water content." In: *Hydrological Processes* 11.2, pp. 203–210. DOI: 10.1002/(SICI)1099-1085(199702)11:2⟨203::AID-HYP432⟩3.0.CO;2-K.

Currie, Iain G. (2003). "Fundamental Mechanics of Fluids." In: *Mechanical Engineering* 1, xiv, 525 p. DOI: 10.1017/CBO9781107415324.004. eprint: arXiv:1011.1669v3.

Darscheid, Paul, Anneli Guthke, and Uwe Ehret (2018). "A Maximum-Entropy Method to Estimate Discrete Distributions from Samples Ensuring Nonzero Probabilities." In: *Entropy* 20.8, p. 601. DOI: 10.3390/e20080601.

Das, Tapash, András Bárdossy, Erwin Zehe, and Yi He (2008). "Comparison of conceptual model performance using different representations of spatial variability." In: *Journal of Hydrology* 356.1-2, pp. 106–118. DOI: 10.1016/j.jhydrol.2008.04.008.

Davies, Paul C. W. (1990). "Why is the physical world so comprehensible?" In: *Complexity, entropy and the physics of information* VIII.1976, pp. 61–71.

Dooge, James C. I. (1973). *Linear Theory*. Washington, D.C.: US Government Printing Office, p. 327. DOI: 10.2307/2228880.

Dooge, James C. I. (1986). "Looking for hydrologic laws." In: *Water Resources Research* 22.9S, 46S–58S. DOI: 10.1029/WR022i09Sp0046S.

Dunne, Thomas and Richard D. Black (1970). "Partial Area Contributions to Storm Runoff in a Small New England Watershed." In: *Water Resources Research* 6.5, pp. 1296–1311. DOI: 10.1029/WR006i005p01296.

Dusek, Jaromir, Tomas Vogel, Michal Dohnal, and Horst H. Gerke (2012). "Combining dual-continuum approach with diffusion wave model to include a preferential flow component in hillslope scale modeling of shallow subsurface runoff." In: *Advances in Water Resources* 44, pp. 113–125. DOI: 10.1016/j.advwatres.2012.05.006.

Ebel, Brian a. and Keith Loague (2006). "Physics-based hydrologic-response simulation: Seeing through the fog of equifinality." In: *Hydrological Processes* 20.13, pp. 2887–2900. DOI: 10.1002/hyp.6388.

Ebel, Brian a., Keith Loague, David R. Montgomery, and William E. Dietrich (2008). "Physics-based continuous simulation of long-term near-surface hydrologic response for the Coos Bay experimental catchment." In: *Water Resources Research* 44.7, pp. 1–23. DOI: 10.1029/2007WR006442.

Ehret, Uwe et al. (2014). "Advancing catchment hydrology to deal with predictions under change." In: *Hydrology and Earth System Sciences* 18.2, pp. 649–671. DOI: 10.5194/hess-18-649-2014.

Emmanuel, I., H. Andrieu, E. Leblois, N. Janey, and O. Payrastre (2015). "Influence of rainfall spatial variability on rainfall-runoff modelling: Benefit of a simulation approach?" In: *Journal of Hydrology* 531, pp. 337–348. DOI: 10.1016/j.jhydrol.2015.04.058.

Endres, D.M. and J.E. Schindelin (July 2003). "A new metric for probability distributions." In: *IEEE Transactions on Information Theory* 49.7, pp. 1858–1860. DOI: 10.1109/TIT.2003.813506.

Faigle, Benjamin, Rainer Helmig, Ivar Aavatsmark, and Bernd Flemisch (2014). "Efficient multiphysics modelling with adaptive grid refinement using a MPFA method." In: *Computational Geosciences* 18.5, pp. 625–636. DOI: 10.1007/s10596-014-9407-1.

Farr, Tom G. et al. (2007). "The Shuttle Radar Topography Mission." In: *Reviews of Geophysics* 45.2, RG2004. DOI: 10.1029/2005RG000183.

Feddes, Reinder A., Piotr Kowalik, Krystina Kolinska-Malinka, and Henryk Zaradny (1976). "Simulation of field water uptake by plants using a soil water dependent root extraction function." In: *Journal of Hydrology* 31.1-2, pp. 13–26. DOI: 10.1016/0022-1694(76)90017-2.

Fenicia, Fabrizio, Dmitri Kavetski, and Hubert H G Savenije (2011). "Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development." In: *Water Resources Research* 47.11, pp. 1–13. DOI: 10.1029/2010WR010174.

Fenicia, Fabrizio, Dmitri Kavetski, Hubert H. G. Savenije, and Laurent Pfister (2016). "From spatially variable streamflow to distributed hydrological models: Analysis of key modeling decisions." In: *Water Resources Research* 52.2, pp. 954–989. DOI: 10.1002/2015WR017398.

Fenicia, Fabrizio et al. (2013). "Catchment properties, function, and conceptual model representation: is there a correspondence?" In: *Hydrological Processes* 28.4, pp. 2451–2467. DOI: 10.1002/hyp.9726.

Feynman, Richard (1963). "Probability." In: *The Feynman Lectures on Physics*.

Flügel, Wolfgang-Albert (1996). "Hydrological Response Units (HRUs) as modeling entities for hydrological river basin simulation and their method- ological potential for modeling complex environmental process systems." In: *Die Erde* 127, pp. 42–62.

Francke, T., A. Güntner, G. Mamede, E. N. Müller, and A. Bronstert (2008). "Automated catena-based discretization of landscapes for the derivation of hydrological modelling units." In: *International Journal of Geographical Information Science* 22.2, pp. 111–132. DOI: 10.1080/13658810701300873.

Freeze, R. Allen and John. A. Cherry (1979). "Groundwater." In: *Progress in Physical Geography*, pp. 1–624.

Freeze, R.Allan and R.L. Harlan (1969). "Blueprint for a physically-based, digitally-simulated hydrologic response model." In: *Journal of Hydrology* 9.3, pp. 237–258. DOI: 10.1016/0022-1694(69)90020-1.

Gao, H., M. Hrachowitz, F. Fenicia, S. Gharari, and H. H. G. Savenije (May 2014). "Testing the realism of a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe, China." In: *Hydrology and Earth System Sciences* 18.5, pp. 1895–1915. DOI: 10.5194/hess-18-1895-2014.

Gao, H. et al. (Feb. 2019). "A simple topography-driven and calibration-free runoff generation module." In: *Hydrology and Earth System Sciences* 23.2, pp. 787–809. DOI: 10.5194/hess-23-787-2019.

Gharari, S., M. Hrachowitz, F. Fenicia, and H. H. G. Savenije (2011). "Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment." In: *Hydrology and Earth System Sciences* 15.11, pp. 3275–3291. DOI: 10.5194/hess-15-3275-2011.

Gong, Wei, Dawen Yang, Hoshin V. Gupta, and Grey Nearing (2014). "Estimating information entropy for hydrological data: One-dimensional case." In: *Water Resources Research* 50.6, pp. 5003–5018. DOI: 10.1002/2014WR015874.

Grabs, T., J. Seibert, K. Bishop, and H. Laudon (2009). "Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model." In: *Journal of Hydrology* 373.1-2, pp. 15–23. DOI: 10.1016/j.jhydrol.2009.03.031.

Graeff, T. et al. (2010). "A quality assessment of Spatial TDR soil moisture measurements in homogenous and heterogeneous media with laboratory experiments." In: *Hydrology and Earth System Sciences* 14.6, pp. 1007–1020. DOI: 10.5194/hess-14-1007-2010.

Graeff, Thomas et al. (2009). "Process identification through rejection of model structures in a mid-mountainous rural catchment: observations of rainfall-runoff response, geophysical conditions and model inter-comparison." In: *Hydrological Processes* 23.5, pp. 702–718. DOI: 10.1002/hyp.7171.

Grayson, Rodger B., Ian D. Moore, and Thomas A. McMahon (1992). "Physically based hydrologic modeling: 2. Is the concept realistic?" In: *Water Resources Research* 28.10, pp. 2659–2666. DOI: 10.1029/92WR01259.

Gupta, Hoshin V., Martyn P. Clark, Jasper a. Vrugt, Gab Abramowitz, and Ming Ye (2012). "Towards a comprehensive assessment of model structural adequacy." In: *Water Resources Research* 48.8, pp. 1–16. DOI: 10.1029/2011WR011044.

Gupta, Hoshin V., Harald Kling, Koray K. Yilmaz, and Guillermo F. Martinez (2009). "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling." In: *Journal of Hydrology* 377.1-2, pp. 80–91. DOI: 10.1016/j.jhydrol.2009.08.003.

Gupta, Hoshin V and Grey S Nearing (2014). "Debates-the future of hydrological sciences: A (common) path forward? Using models and data to learn: A systems theoretic perspective on the future of hydrological science." In: *Water Resources Research* 50.6, pp. 5351–5359. DOI: 10.1002/2013WR015096.

Hack, John T, Fred A Seaton, and Thomas B Nolan (1957). *Studies of Longitudinal Stream Profiles in Virginia and Maryland*. Tech. rep. U.S. Geological Survey Professional Paper, pp. 45–97.

Harman, C. J., M. Sivapalan, and P. Kumar (2009). "Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics." In: *Water Resources Research* 45.9, pp. 1–13. DOI: 10.1029/2008WR007392.

Hassler, Sibylle K, Markus Weiler, and Theresa Blume (2017). "Tree-, stand- and site-specific controls on landscape-scale patterns of transpiration." In: February, pp. 1–18. DOI: 10.5194/hess-2017-47.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009). *The Elements of Statistical Learning*. Second Edi. Springer Series in Statistics. New York, NY: Springer New York, p. 587. DOI: 10.1007/978-0-387-84858-7.

Hazenberg, P. et al. (2016). "Testing the hybrid-3-D hillslope hydrological model in a controlled environment." In: *Water Resources Research* 52.2, pp. 1089–1107. DOI: 10.1002/2015WR018106.

Hergarten, Stefan, G. Winkler, and S. Birk (2014). "Transferring the concept of minimum energy dissipation from river networks to subsurface flow patterns." In: *Hydrology and Earth System Sciences* 18.10, pp. 4277–4288. DOI: 10.5194/hess-18-4277-2014.

Hildebrandt, Anke, Axel Kleidon, and Marcel Bechmann (2016). "A thermodynamic formulation of root water uptake." In: *Hydrology and Earth System Sciences* 20.8, pp. 3441–3454. DOI: 10.5194/hess-20-3441-2016.

Hjerdt, K. N., J. J. McDonnell, J. Seibert, and A. Rodhe (2004). "A new topographic index to quantify downslope controls on local drainage." In: *Water Resources Research* 40.5, pp. 1–6. DOI: 10.1029/2004WR003130.

Hopp, L. and J.J. McDonnell (2009). "Connectivity at the hillslope scale: Identifying interactions between storm size, bedrock permeability, slope angle and soil depth." In: *Journal of Hydrology* 376.3-4, pp. 378–391. DOI: 10.1016/j.jhydrol.2009.07.047.

Horton, R.E (1945). "Erosional Development of Streams and their Drainage Basins "Hydro-Physical Approach to Quantitative Morphology"." In: *Bulletin of the Geological Society of America* 56.March, pp. 275–370.

Howard, Alan D. (Sept. 1990). "Theoretical model of optimal drainage networks." In: *Water Resources Research* 26.9, pp. 2107–2117. DOI: 10.1029/WR026i009p02107.

Hrachowitz, Markus and Martyn P. Clark (2017). "HESS Opinions: The complementary merits of competing modelling philosophies in hydrology." In: *Hydrology and Earth System Sciences* 21.8, pp. 3953–3973. DOI: 10.5194/hess-21-3953-2017.

Hrachowitz, Markus et al. (2013). "A decade of Predictions in Ungauged Basins (PUB)—a review." In: *Hydrological Sciences Journal* 58.6, pp. 1198–1255. DOI: 10.1080/02626667.2013.803183.

Hubbert, M. King (1940). "The theory of ground-water motion." In: *Transactions, American Geophysical Union* 21.2, p. 648. DOI: 10.1029/TR021i002p00648-1.

Hughes, Joseph D. and Jie Liu (Nov. 2008). "MIKE SHE: Software for Integrated Surface Water/Ground Water Modeling." In: *Ground Water* 46.6, pp. 797–802. DOI: 10.1111/j.1745-6584.2008.00500.x.

Huuskonen, Asko, Elena Saltikoff, and Iwan Holleman (2014). "The Operational Weather Radar Network in Europe." In: *Bulletin of the American Meteorological Society* 95.6, pp. 897–907. DOI: 10.1175/BAMS-D-12-00216.1.

Jackisch, Conrad (2015). "Linking structure and functioning of hydrological systems." In: p. 171. DOI: 10.5445/IR/1000051494.

Jackisch, Conrad et al. (2016). "In situ investigation of rapid subsurface flow: Identification of relevant spatial structures beyond heterogeneity." In: *Hydrology and Earth System Sciences Discussions* 2016, pp. 1–32. DOI: 10.5194/hess-2016-190.

Jackisch, Conrad et al. (2017). "Form and function in hillslope hydrology: in situ imaging and characterization of flow relevant structures." In: *Hydrology and Earth System Sciences* 21.7, pp. 3749–3775. DOI: 10.5194/hess-21-3749-2017.

Jakeman, A. J. and G. M. Hornberger (1993). "How much complexity is warranted in a rainfall-runoff model?" In: *Water Resources Research* 29.8, pp. 2637–2649. DOI: 10.1029/93WR00877. arXiv: arXiv:1011.1669v3.

Jarvis, P. G. (1976). "The Interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 273.927, pp. 593–610. DOI: 10.1098/rstb.1976.0035.

Jasiewicz, Jarosław and Markus Metz (2011). "A new GRASS GIS toolkit for Hortonian analysis of drainage networks." In: *Computers & Geosciences* 37.8, pp. 1162–1173. DOI: 10.1016/j.cageo.2011.03.003.

Jaynes, E. T. (1957). "Information Theory and Statistical Mechanics." In: *Physical Review* 106.4, pp. 620–630. DOI: 10.1103/PhysRev.106.620.

Juilleret, J., J. F. Iffly, L. Pfister, and C. Hissler (2011). "Remarkable Pleistocene periglacial slope deposits in Luxembourg (Oesling): pedological implication and geosite potential." In: *Bulletin de la Société des Naturalistes Luxembourgeois* 112.1, pp. 125–130.

Kawachi, Toshihiko, Takeo Maruyama, and Vijay P. Singh (2001). "Rainfall entropy for delineation of water resources zones in Japan." In: *Journal of Hydrology* 246.1-4, pp. 36–44. DOI: 10.1016/S0022-1694(01)00355-9.

Kirchner, James W. (2006). "Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology." In: *Water Resources Research* 42.3. DOI: 10.1029/2005WR004362.

Kirkby, M.J. (1975). *Hydrograph modelling strategies.* Ed. by Chrisholm M and Hagget P. chapter 5. Leeds: University of Leeds, pp. 69–90.

Kirkby, M.J. (1976). "Tests of the random network model, and its application to basin hydrology." In: *Earth Surface Processes* 1.August 1975, pp. 197–212. DOI: 10.1002/esp.3290010302.

Klaus, Julian and Erwin Zehe (2010). "Modelling rapid flow response of a tile-drained field site using a 2D physically based model: assessment of 'equifinal' model setups." In: *Hydrological Processes* 24.12, pp. 1595–1609. DOI: 10.1002/hyp.7687.

Klaus, Julian and Erwin Zehe (2011). "A novel explicit approach to model bromide and pesticide transport in connected soil structures." In: *Hydrology and Earth System Sciences* 15.7, pp. 2127–2144. DOI: 10.5194/hess-15-2127-2011.

Kleidon, Axel (2010). "A basic introduction to the thermodynamics of the Earth system far from equilibrium and maximum entropy production." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1545, pp. 1303–1315. DOI: 10.1098/rstb.2009.0310.

Kleidon, Axel (2016). *Thermodynamic Foundations of the Earth System*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139342742.

Kleidon, Axel and Maik Renner (2013). "A simple explanation for the sensitivity of the hydrologic cycle to surface temperature and solar radiation and its implications for global climate change." In: *Earth System Dynamics* 4.2, pp. 455–465. DOI: 10.5194/esd-4-455-2013.

Kleidon, Axel, Erwin Zehe, Uwe Ehret, and Ulrike Scherer (2013). "Thermodynamics, maximum power, and the dynamics of preferential river flow structures at the continental scale." In: *Hydrology and Earth System Sciences* 17.1, pp. 225–251. DOI: 10.5194/hess-17-225-2013.

Klemeš, Vit (1986). "Operational testing of hydrological simulation models." In: *Hydrological Sciences Journal* 31.1, pp. 13–24. DOI: 10.1080/02626668609491024.

Kondepudi, Dilip and I Prigogine (1998). "From heat engines to dissipative structures." In: DOI: 10.1002/9781118698723.ch19.

Kondepudi, Dilip and Ilya Prigogine (Nov. 2014). *Modern Thermodynamics*. Chichester, UK: John Wiley & Sons, Ltd, pp. 421–450. DOI: 10.1002/9781118698723.

Koutsoyiannis, Demetris (2014). "Entropy: From Thermodynamics to Hydrology." In: *Entropy* 16.3, pp. 1287–1314. DOI: 10.3390/e16031287.

Kullback, S. and R. A. Leibler (1951). "On Information and Sufficiency." In: *The Annals of Mathematical Statistics* 22.1, pp. 79–86. DOI: 10.1214/aoms/1177729694.

Lamy, Edvina, Laurent Lassabatere, Beatrice Bechet, and Herve Andrieu (2009). "Modeling the influence of an artificial macropore in sandy columns on flow and solute transfer." In: *Journal of Hydrology* 376.3-4, pp. 392–402. DOI: 10.1016/j.jhydrol.2009.07.048.

Langbein, W. B. and L. B. Leopold (June 1964). "Quasi-equilibrium states in channel morphology." In: *American Journal of Science* 262.6, pp. 782–794. DOI: 10.2475/ajs.262.6.782.

Lee, H, E Zehe, and M Sivapalan (2007). "Predictions of rainfall-runoff response and soil moisture dynamics in a microscale catchment using the CREW model." In: *Hydrology and Earth System Sciences* 11.2, pp. 819–849. DOI: 10.5194/hess-11-819-2007.

Lehmann, Peter, Christoph Hinz, G McGrath, H. J. Tromp-van Meerveld, and J J McDonnell (2007). "Rainfall threshold for hillslope outflow: an emergent property of flow pathway connectivity." In: *Hydrology and Earth System Sciences* 11.2, pp. 1047–1063. DOI: 10.5194/hess-11-1047-2007.

Leopold, Luna B. and Thomas Maddock (1953). *The Hydraulic Geomtry of Stream Channels and Some Physiographic Implications*. Tech. rep., p. 57. DOI: 10.3133/PP252.

Leopold, Luna Bergere and Walter Basil Langbein (1962). *The concept of entropy in landscape evolution*. English. Tech. rep. Washington, D.C., A1–A20.

Lin, Jianhua (1991). "Divergence measures based on the Shannon entropy." In: *IEEE Transactions on Information Theory* 37.1, pp. 145–151. DOI: 10.1109/18.61115.

Lindley, David (2007). *Uncertainty: Einstein, Heisenberg, Bohr, and the Struggle for The Soul of Science*. New York: Doubleday.

Lindsay J.B. (2014). "The Whitebox Geospatial Analysis Tools project and open-access GIS," in: *Proceedings of the GIS Research UK 22nd Annual Conference* 001, p. 8. DOI: 10.13140/RG.2.1.1010.8962.

Loague, Keith and Joel E. VanderKwaak (2004). "Physics-based hydrologic response simulation: Platinum bridge, 1958 Edsel, or useful tool." In: *Hydrological Processes* 18.15, pp. 2949–2956. DOI: 10.1002/hyp.5737.

Lobligeois, F., V. Andréassian, C. Perrin, P. Tabary, and C. Loumagne (2014). "When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events." In: *Hydrology and Earth System Sciences* 18.2, pp. 575–594. DOI: 10.5194/hess-18-575-2014.

Loke, MH (2003). "Rapid 2D Resistivity & IP Inversion using the least-squares method." In: *Geotomo Software, Manual*.

Loritz, Ralf et al. (2017). "Picturing and modeling catchments by representative hillslopes." In: *Hydrology and Earth System Sciences* 21.2, pp. 1225–1249. DOI: 10.5194/hess-21-1225-2017.

Loritz, Ralf et al. (2018). "On the dynamic nature of hydrological similarity." In: *Hydrology and Earth System Sciences* 22.7, pp. 3663–

3684. DOI: 10.5194/hess-22-3663-2018. arXiv: hess-22-3663-2018 [10.5194].

Lyon, Steve W. and Peter A. Troch (2010). "Development and application of a catchment similarity index for subsurface flow." In: *Water Resources Research* 46.3, pp. 1–13. DOI: 10.1029/2009WR008500.

Majtey, A. P., P. W. Lamberti, and D. P. Prato (2005). "Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states." In: *Physical Review A - Atomic, Molecular, and Optical Physics* 72.5, pp. 1–6. DOI: 10.1103/PhysRevA.72.052310. arXiv: 0508138 [quant-ph].

Martínez-Carreras, N. et al. (2015). "Hydrological connectivity inferred from diatom transport through the riparian-stream system." In: *Hydrology and Earth System Sciences* 19.7, pp. 3133–3151. DOI: 10.5194/hess-19-3133-2015.

Martínez-Carreras, Núria et al. (2012). "The Influence of Sediment Sources and Hydrologic Events on the Nutrient and Metal Content of Fine Grained Sediments (Attert River Basin, Luxembourg)." In: *Water, Air, & Soil Pollution* 223.9, pp. 5685–5705. DOI: 10.1007/s11270-012-1307-1.

Martínez-Carreras, Núria et al. (2016). "Storage controls on the generation of double peak hydrographs in a forested headwater catchment." In: *Journal of Hydrology* 543, pp. 255–269. DOI: 10.1016/j.jhydrol.2016.10.004.

Maurer, Thomas. (1997). "Physikalisch begründete zeitkontinuierliche Modellierung des Wassertransports in kleinen ländlichen Einzugsgebieten." PhD thesis. Karlsruher Institut für Technologie, p. 1.

McDonnell, J. J. et al. (2007). "Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology." en. In: *Water Resources Research* 43.7, n/a–n/a. DOI: 10.1029/2006WR005467.

McGuire, K. J. et al. (2005). "The role of topography on catchment-scale water residence time." In: *Water Resources Research* 41.5, pp. 1–14. DOI: 10.1029/2004WR003657.

Mendoza, Pablo A. et al. (2015). "Are we unnecessarily constraining the agility of complex process-based models?" In: *Water Resources Research* 51.1, pp. 716–728. DOI: 10.1002/2014WR015820.

Menzel, Annette, Gert Jakobi, Rein Ahas, Helfried Scheifinger, and Nicole Estrella (2003). "Variations of the climatological growing season (1951-2000) in Germany compared with other countries." In: *International Journal of Climatology* 23.7, pp. 793–812. DOI: 10.1002/joc.915.

Michaels, George S et al. (1998). "Cluster analysis and data visualization of large-scale gene expression data." In: *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* 3, pp. 42–53.

Montanari, Alberto et al. (Aug. 2013). ""Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific

Decade 2013–2022." In: *Hydrological Sciences Journal* 58.6, pp. 1256–1275. DOI: 10.1080/02626667.2013.809088.

Montgomery, D. R. and W. E. Dietrich (1992). "Channel Initiation and the Problem of Landscape Scale." In: *Science* 255.5046, pp. 826–830. DOI: 10.1126/science.255.5046.826.

Montgomery, David R. and William E. Dietrich (1988). "Where do channels begin?" In: *Nature* 336.6196, pp. 232–234. DOI: 10.1038/336232a0.

Mueller, E. N., A. Güntner, T. Francke, and G. Mamede (2010). "Modelling sediment export, retention and reservoir sedimentation in drylands with the WASA-SED model." In: *Geoscientific Model Development* 3.1, pp. 275–291. DOI: 10.5194/gmd-3-275-2010.

Murtagh, Fionn and Pierre Legendre (2014). "Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?" In: *Journal of Classification* 31.3, pp. 274–295. DOI: 10.1007/s00357-014-9161-z. arXiv: arXiv:1111.6285v2.

Musa, Z. N., I. Popescu, and A. Mynett (2015). "A review of applications of satellite SAR, optical, altimetry and DEM data for surface water modelling, mapping and parameter estimation." In: *Hydrology and Earth System Sciences* 19.9, pp. 3755–3769. DOI: 10.5194/hess-19-3755-2015.

Mutzner, Raphaël, Paolo Tarolli, Giulia Sofia, Marc B. Parlange, and Andrea Rinaldo (2016). "Field study on drainage densities and rescaled width functions in a high-altitude alpine catchment." In: *Hydrological Processes* 30.13, pp. 2138–2152. DOI: 10.1002/hyp.10783.

Nearing, Grey S. et al. (2016). "A philosophical basis for hydrological uncertainty." In: *Hydrological Sciences Journal* 61.9, pp. 1666–1678. DOI: 10.1080/02626667.2016.1183009.

Neteler, Markus, M. Hamish Bowman, Martin Landa, and Markus Metz (2012). "GRASS GIS: A multi-purpose open source GIS." In: *Environmental Modelling & Software* 31, pp. 124–130. DOI: 10.1016/j.envsoft.2011.11.014.

Nicótina, L., E. Alessi Celegon, A. Rinaldo, and M. Marani (2008). "On the impact of rainfall patterns on the hydrologic response." In: *Water Resources Research* 44.12, pp. 1–14. DOI: 10.1029/2007WR006654.

Nieber, John L. and Roy C. Sidle (2010). "How do disconnected macropores in sloping soils facilitate preferential flow?" In: *Hydrological Processes* 24.12, pp. 1582–1594. DOI: 10.1002/hyp.7633.

Nieber, John L. and Glenn S. Warner (1991). "Soil pipe contribution to steady subsurface stormflow." In: *Hydrological Processes* 5.4, pp. 329–344. DOI: 10.1002/hyp.3360050402.

Niu, Guo-Yue et al. (2011). "The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements." In: *Journal of Geophysical Research* 116.D12, p. D12109. DOI: 10.1029/2010JD015139.

Obled, Ch, J. Wendling, and K. Beven (1994). "The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data." In: *Journal of Hydrology* 159.1-4, pp. 305–333. DOI: 10.1016/0022-1694(94)90263-1.

Or, Dani, Peter Lehmann, and Shmuel Assouline (2015). "Natural length scales define the range of applicability of the Richards equation for capillary flows." In: *Water Resources Research* 51.9, pp. 7130–7144. DOI: 10.1002/2015WR017034. eprint: 2014WR016527 (10.1002).

Oudin, Ludovic, Alison Kay, Vazken Andréassian, and Charles Perrin (2010). "Are seemingly physically similar catchments truly hydrologically similar?" In: *Water Resources Research* 46.11, pp. 1–15. DOI: 10.1029/2009WR008887.

Pechlivanidis, Ilias G, Bethanna Jackson, Hilary Mcmillan, and Hoshin V Gupta (2016). "Robust informational entropy-based descriptors of flow in catchment hydrology." In: *Hydrological Sciences Journal* 61.1, pp. 1–18. DOI: 10.1080/02626667.2014.983516.

Pérez, A.J., R. Abrahão, J. Causapé, O.A. Cirpka, and C.M. Bürger (2011). "Simulating the transition of a semi-arid rainfed catchment towards irrigation agriculture." In: *Journal of Hydrology* 409.3-4, pp. 663–681. DOI: 10.1016/j.jhydrol.2011.08.061.

Peters, A and W Durner (2008). "Simplified evaporation method for determining soil hydraulic properties." In: *Journal of Hydrology* 356.1-2, pp. 147–162. DOI: 10.1016/j.jhydrol.2008.04.016.

Pfeiffer, Allison M., Noah J. Finnegan, and Jane K. Willenbring (Mar. 2017). "Sediment supply controls equilibrium channel geometry in gravel rivers." In: *Proceedings of the National Academy of Sciences* 114.13, pp. 3346–3351. DOI: 10.1073/pnas.1612907114.

Pfister, L. et al. (2018). "Contrasting Hydrologic Response in the Cuesta Landscapes of Luxembourg." In: *The Luxembourg Gutland Landscape*. Ed. by A. M. Kooijman, L.H. Cammeraat, and A.C. Seijmonsbergen. Cham: Springer International Publishing, pp. 73–87. DOI: 10.1007/978-3-319-65543-7_4.

Pfister, Laurent, Joel Humbert, and Lucien Hoffmann (2000). "Recent trends in rainfall-runoff characteristics in the Alzette River basin, Luxembourg." In: *Climatic Change* 45.2, pp. 323–337. DOI: 10.1023/A:1005567808533.

Pfister, Laurent, Jean-François Iffly, Lucien Hoffmann, and Joel Humbert (2002). "Use of regionalized stormflow coefficients with a view to hydroclimatological hazard mapping." In: *Hydrological Sciences Journal* 47.3, pp. 479–491. DOI: 10.1080/02626660209492948.

Pfister, Laurent and James W Kirchner (2017). "Debates-Hypothesis testing in hydrology: Theory and practice." In: *Water Resources Research* 53.3, pp. 1792–1798. DOI: 10.1002/2016WR020116.

Pokhrel, Prafulla and Hoshin V. Gupta (2010). "On the use of spatial regularization strategies to improve calibration of distributed watershed models." In: *Water Resources Research* 46.1, pp. 1–17. DOI: 10.1029/2009WR008066.

Pokhrel, Prafulla, Koray K. Yilmaz, and Hoshin V. Gupta (2012). "Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures." In: *Journal of Hydrology* 418-419, pp. 49–60. DOI: 10.1016/j.jhydrol.2008.12.004.

Popper, Karl R. (1959). "The Propensity Interpretation of Probability." In: *The British Journal for the Philosophy of Science* 10.37, pp. 25–42.

Rao, Radhakrishna C (1982). "and Dissimilarity." In: *Theoretical Population Biology* 43, pp. 24–43.

Refsgaard, JC and B Storm (1995). *MIKE SHE in Computer models of watershed hydrology*. Ed. by V P Singh. Water Resources Publications, pp. 809–846.

Refsgaard, J.C., M. Thorsen, J.B. Jensen, S. Kleeschulte, and S. Hansen (1999). "Large scale modelling of groundwater contamination from nitrate leaching." In: *Journal of Hydrology* 221.3-4, pp. 117–140. DOI: 10.1016/S0022-1694(99)00081-5.

Refsgaard, Jens Christian (1997). "Parameterisation, calibration and validation of distributed hydrological models." In: *Journal of Hydrology* 198.1-4, pp. 69–97. DOI: 10.1016/S0022-1694(96)03329-X.

Reggiani, P. and T. H. M. Rientjes (2005). "Flux parameterization in the representative elementary watershed approach: Application to a natural basin." In: *Water Resources Research* 41.4. DOI: 10.1029/2004WR003693.

Reggiani, Paolo, Murugesu Sivapalan, and S. Majid Hassanizadeh (1998). "A unifying framework for watershed thermodynamics: Balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics." In: *Advances in Water Resources* 22.4, pp. 367–398. DOI: 10.1016/S0309-1708(98)00012-8.

Reid, D. E. B., Uldis Silins, Carl Mendoza, and Victor J. Lieffers (2005). "A unified nomenclature for quantification and description of water conducting properties of sapwood xylem based on Darcy's law." In: *Tree Physiology* 25.8, pp. 993–1000. DOI: 10.1093/treephys/25.8.993.

Rennó, Camilo Daleles et al. (2008). "HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia." In: *Remote Sensing of Environment* 112.9, pp. 3469–3481. DOI: 10.1016/j.rse.2008.03.018.

Rinaldo, Andrea et al. (1992). "Minimum energy and fractal structures of drainage networks." In: *Water Resources Research* 28.9, pp. 2183–2195. DOI: 10.1029/92WR00801.

Robinson, Justin S., Murugesu Sivapalan, and John D. Snell (1995). "On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments." In: *Water Resources Research* 31.12, pp. 3089–3101. DOI: 10.1029/95WR01948.

Rodhe, A. and J. Seibert (1999). "Wetland occurrence in relation to topography: a test of topographic indices as moisture indicators." In: *Agricultural and Forest Meteorology* 98-99, pp. 325–340. DOI: 10.1016/S0168-1923(99)00104-5.

Rodríguez-Iturbe, Ignacio et al. (1992). "Energy dissipation, runoff production, and the three-dimensional structure of river basins." In: *Water Resources Research* 28.4, pp. 1095–1103. DOI: 10.1029/91WR03034.

Savenije, Hubert H. G. (2009). "HESS Opinion "The art of hydrology"*." In: *Hydrology and Earth System Sciences* 13.2, pp. 157–161. DOI: 10.5194/hess-13-157-2009.

Savenije, Hubert H. G. (Dec. 2010). "HESS Opinions "Topography driven conceptual modelling (FLEX-Topo)"." In: *Hydrology and Earth System Sciences* 14.12, pp. 2681–2692. DOI: 10.5194/hess-14-2681-2010.

Savenije, Hubert H. G. and Markus Hrachowitz (2017). "HESS Opinions "Catchments as meta-organisms – a new blueprint for hydrological modelling"." In: *Hydrology and Earth System Sciences* 21.2, pp. 1107–1116. DOI: 10.5194/hess-21-1107-2017.

Sawicz, K, T Wagener, M Sivapalan, P A Troch, and G Carrillo (2011). "Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA." In: *Hydrology and Earth System Sciences* 15.9, pp. 2895–2911. DOI: 10.5194/hess-15-2895-2011.

Schaefli, Bettina and Hoshin V Gupta (2007). "Do Nash values have value?" In: *Hydrological Processes* 21.15, pp. 2075–2080. DOI: 10.1002/hyp.6825.

Schierholz, I., D. Schäfer, and O. Kolle (2000). "The Weiherbach data set: An experimental data set for pesticide model testing on the field scale." In: *Agricultural Water Management* 44.1-3, pp. 43–61. DOI: 10.1016/S0378-3774(99)00083-9.

Schneider, E.D. and J.J. Kay (1994). "Life as a manifestation of the second law of thermodynamics." In: *Mathematical and Computer Modelling* 19.6-8, pp. 25–48. DOI: 10.1016/0895-7177(94)90188-0.

Schoorl, J. M., M. P. W. Sonneveld, and A. Veldkamp (2000). "Three-dimensional landscape process modelling: the effect of DEM resolution." In: *Earth Surface Processes and Landforms* 25.9, pp. 1025–1034. DOI: 10.1002/1096-9837(200008)25:9⟨1025::AID-ESP116⟩3.0.CO;2-Z.

Schoups, G., N. C. van de Giesen, and H. H G Savenije (2008). "Model complexity control for hydrologic prediction." In: *Water Resources Research* 44.12, p. 14. DOI: 10.1029/2008WR006836.

Schweizer, Daniel, Philipp Blum, and Christoph Butscher (2017). "Uncertainty assessment in 3-D geological models of increasing complexity." In: *Solid Earth* 8.2, pp. 515–530. DOI: 10.5194/se-8-515-2017.

Scott, David W. (1979). "On Optimal and Data-Based Histograms." In: *Biometrika* 66.3, p. 605. DOI: 10.2307/2335182.

Scudeler, Carlotta et al. (2016). "Multiresponse modeling of an unsaturated zone isotope tracer experiment at the Landscape Evolution Observatory." In: *Hydrology and Earth System Sciences Discussions* May, pp. 1–29. DOI: 10.5194/hess-2016-228.

Seibert, Jan and Jeffrey J. McDonnell (2002). "On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration." In: *Water Resources Research* 38.11, pp. 23–1–23–14. DOI: 10.1029/2001WR000978.

Seibert, Jan and Brian L. McGlynn (2007). "A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models." In: *Water Resources Research* 43.4, pp. 1–8. DOI: 10.1029/2006WR005128.

Seibert, Simon Paul, Uwe Ehret, and Erwin Zehe (2016). "Disentangling timing and amplitude errors in streamflow simulations." In: *Hydrology and Earth System Sciences Discussions* March, pp. 1–37. DOI: 10.5194/hess-2016-145.

Seibert, Simon Paul, Conrad Jackisch, Uwe Ehret, Laurent Pfister, and Erwin Zehe (2017). "Unravelling abiotic and biotic controls on the seasonal water balance using data-driven dimensionless diagnostics." In: *Hydrology and Earth System Sciences* 21.6, pp. 2817–2841. DOI: 10.5194/hess-21-2817-2017.

Shannon, Claude E (1948). "A mathematical theory of communication." In: *Bell System Technical Journal* 27.1, pp. 379–423. DOI: 10.1145/584091.584093. arXiv: 9411012 [chao-dyn].

Šimůnek, Jiří, Martinus Th. van Genuchten, and Miroslav Šejna (2016). "Recent Developments and Applications of the HYDRUS Computer Software Packages." In: *Vadose Zone Journal* 15.7, p. 0. DOI: 10.2136/vzj2016.04.0033.

Simunek, J, M. T. V. Genuchten, and M Sejna (2006). "The HYDRUS Software Package for Simulating the Two-and Three-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media Technical Manual." In: *Version 1.0, PC-Progress, Prague, Czech Republic*. DOI: 10.1016/j.earscirev.2013.09.007.

Singh, Vijay P. (2013). *Entropy Theory and its Application in Environmental and Water Engineering*. Vol. 49. 4. Chichester, UK: John Wiley & Sons, Ltd, pp. 2253–2273. DOI: 10.1002/9781118428306.

Sivapalan, Murugesu (2003). "Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection?" In: *Hydrological Processes* 17.5, pp. 1037–1041. DOI: 10.1002/hyp.5109.

Sivapalan, Murugesu (2005). "Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale." In: *Encyclopedia of Hydrological Sciences*. Vol. -. -. Chichester, UK: John Wiley & Sons, Ltd, pp. –. DOI: 10.1002/0470848944.hsa012.

Sivapalan, Murugesu (Mar. 2018). "From engineering hydrology to Earth system science: Milestones in the transformation of hydrologic science." In: *Hydrology and Earth System Sciences* 22.3, pp. 1665–1693. DOI: 10.5194/hess-22-1665-2018.

Sivapalan, Murugesu, Keith J. Beven, and Eric Woods (1987). "On Hydrologic Similarity." In: 23.12, pp. 2266–2278.

Sivapalan, Murugesu et al. (2003). "IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences." In: *Hydrological Sciences Journal* 48.6, pp. 857–880. DOI: 10.1623/hysj.48.6.857.51421.

Smith, Michael B. et al. (2004). "Runoff response to spatial variability in precipitation: an analysis of observed data." In: *Journal of Hydrology* 298.1-4, pp. 267–286. DOI: 10.1016/j.jhydrol.2004.03.039.

Song, Charles C S (1992). "The Role of Energy Dissipation in Fluid Flows and River Mechanics." In: pp. 447–455. DOI: 10.1007/978-94-011-2430-0_25.

Sørensen, Rasmus, U. Zinko, and J. Seibert (2006). "On the calculation of the topographic wetness index: evaluation of different methods based on field observations." In: *Hydrology and Earth System Sciences* 10.1, pp. 101–112. DOI: 10.5194/hess-10-101-2006. arXiv: hessd/2005-2-1807 [1812-2116].

Sprenger, M., T. H. M. Volkmann, T. Blume, and M. Weiler (June 2015). "Estimating flow and transport parameters in the unsaturated zone with pore water stable isotopes." In: *Hydrology and Earth System Sciences* 19.6, pp. 2617–2635. DOI: 10.5194/hess-19-2617-2015.

Tetzlaff, D, S Uhlenbrook, and P Molnar (2005). "Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data." In: *Hydrology & Earth System Sciences* 9.1/2, pp. 29–41. DOI: 10.5194/hess-9-29-2005.

Thompson, James A., Jay C. Bell, and Charles A. Butler (2001). "Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling." In: *Geoderma* 100.1-2, pp. 67–89. DOI: 10.1016/S0016-7061(00)00081-1.

Tolman, Richard Chace (1938). *The Principles of Statistical Mechanics*. Oxford: The Clarendon Press.

Troch, Peter A., Claudio Paniconi, and E. Emiel van Loon (2003). "Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response." In: *Water Resources Research* 39.11, n/a–n/a. DOI: 10.1029/2002WR001728.

Tromp-Van Meerveld, H. J. and J. J. McDonnell (2006). "Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis." In: *Water Resources Research* 42.August 2005, pp. 1–11. DOI: 10.1029/2004WR003800.

Tucker, G E and R L Bras (2000). "A stochastic approach to modeling the role of rainfall variability in drainage basin evolution." In: *Water Resources Research* 36.7, pp. 1953–1964. DOI: 10.1029/2000wr900065.

Wagener, Thorsten and Hoshin V. Gupta (2005). "Model identification for hydrological forecasting under uncertainty." In: *Stochastic Environmental Research and Risk Assessment* 19.6, pp. 378–387. DOI: 10.1007/s00477-005-0006-5.

Wagener, Thorsten, Murugesu Sivapalan, Peter Troch, and Ross Woods (2007). "Catchment Classification and Hydrologic Similarity." In: *Geography Compass* 1.4, pp. 901–931. DOI: 10.1111/j.1749-8198.2007.00039.x.

Weijs, S. V., N. van de Giesen, and M. B. Parlange (2013a). "Data compression to define information content of hydrological time series." In: *Hydrology and Earth System Sciences* 17.8, pp. 3171–3187. DOI: 10.5194/hess-17-3171-2013.

Weijs, S. V. and Nick van de Giesen (2013). "An information-theoretical perspective on weighted ensemble forecasts." In: *Journal of Hydrology* 498, pp. 177–190. DOI: 10.1016/j.jhydrol.2013.06.033.

Weijs, S. V., Nick van de Giesen, and Marc Parlange (2013b). "HydroZIP: How Hydrological Knowledge can Be Used to Improve Compression of Hydrological Data." In: *Entropy* 15.4, pp. 1289–1310. DOI: 10.3390/e15041289.

Weijs, S. V., G. Schoups, and N. van de Giesen (2010). "Why hydrological predictions should be evaluated using information theory." In: *Hydrology and Earth System Sciences* 14.12, pp. 2545–2558. DOI: 10.5194/hess-14-2545-2010.

Weinberg, Gerald M. (1975). *An Introduction to General Systems Thinking*. New York: John Wiley & Sons, p. 279.

Wienhöfer, J. and E. Zehe (2014). "Predicting subsurface stormflow response of a forested hillslope – the role of connected flow paths." In: *Hydrology and Earth System Sciences* 18.1, pp. 121–138. DOI: 10.5194/hess-18-121-2014.

Wood, Eric F., M. Sivapalan, Keith Beven, and Larry Band (1988). "Effects of spatial variability and scale with implications to hydrologic modeling." In: *Journal of Hydrology* 102.1-4, pp. 29–47. DOI: 10.1016/0022-1694(88)90090-X.

Wrede, Sebastian et al. (2015). "Towards more systematic perceptual model development: a case study using 3 Luxembourgish catchments." In: *Hydrological Processes* 29.12, pp. 2731–2750. DOI: 10.1002/hyp.10393.

Yakirevich, A. et al. (2013). "Augmentation of groundwater monitoring networks using information theory and ensemble modeling with pedotransfer functions." In: *Journal of Hydrology* 501, pp. 13–24. DOI: 10.1016/j.jhydrol.2013.07.032.

Yoshida, Takeo and Peter A. Troch (2016). “Coevolution of volcanic catchments in Japan.” In: *Hydrology and Earth System Sciences* 20.3, pp. 1133–1150. DOI: 10.5194/hess-20-1133-2016.

Zehe, E., H. Elsenbeer, F. Lindenmaier, K. Schulz, and G. Blöschl (2007). “Patterns of predictability in hydrological threshold systems.” In: *Water Resources Research* 43.7. DOI: 10.1029/2006WR005589.

Zehe, Erwin, Rolf Becker, András Bárdossy, and Erich Plate (2005). “Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation.” In: *Journal of Hydrology* 315.1-4, pp. 183–202. DOI: 10.1016/j.jhydrol.2005.03.038.

Zehe, Erwin and Hannes Flühler (2001). “Preferential transport of isoproturon at a plot scale and a field scale tile-drained site.” In: *Journal of Hydrology* 247.1-2, pp. 100–115. DOI: 10.1016/S0022-1694(01)00370-5.

Zehe, Erwin, T. Graeff, M. Morgner, A. Bauer, and A. Bronstert (2010). “Plot and field scale soil moisture dynamics and subsurface wetness control on runoff generation in a headwater in the Ore Mountains.” In: *Hydrology and Earth System Sciences* 14.6, pp. 873–889. DOI: 10.5194/hess-14-873-2010.

Zehe, Erwin, H Lee, and M Sivapalan (2006). “Dynamical process upscaling for deriving catchment scale state variables and constitutive relations for meso-scale process models To cite this version : Dynamical process upscaling for deriving catchment scale state variables and constitutive relations fo.” In:

Zehe, Erwin, Thomas Maurer, Jürgen Ihringer, and Erich Plate (2001). “Modeling water flow and mass transport in a loess catchment.” In: *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* 26.7-8, pp. 487–507. DOI: 10.1016/S1464-1909(01)00041-7.

Zehe, Erwin et al. (2013). “A thermodynamic approach to link self-organization, preferential flow and rainfall–runoff behaviour.” In: *Hydrology and Earth System Sciences* 17.11, pp. 4297–4322. DOI: 10.5194/hess-17-4297-2013.

Zehe, Erwin et al. (2014). “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.” In: *Hydrology and Earth System Sciences* 18.11, pp. 4635–4655. DOI: 10.5194/hess-18-4635-2014.

Zehe, Erwin et al. (2018). “Energy states of soil water – a thermodynamic perspective on storage dynamics and the underlying controls.” In: *Hydrology and Earth System Sciences Discussions* June, pp. 1–33. DOI: 10.5194/hess-2018-346.

Zhang, G. P. and H. H. G. Savenije (2005). “Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach.”

In: *Hydrology and Earth System Sciences* 9.3, pp. 243–261. DOI: 10.5194/hess-9-243-2005.

Zhang, Weihua and David R Montgomery (1994). "Digital elevation model grid size, landscape representation, and hydrologic simulations." In: *Water Resources Research* 30.4, pp. 1019–1028. DOI: 10.1029/93WR03553.

Zhang, Zhilin and Hubert H. G. Savenije (2018). "Thermodynamics of saline and fresh water mixing in estuaries." In: *Earth System Dynamics* 9.1, pp. 241–247. DOI: 10.5194/esd-9-241-2018.

*Ohana* means family.
Family means nobody gets left behind, or forgotten.

— Lilo & Stitch

# ACKNOWLEDGMENTS

This thesis would not have been possible without the inspiration and support of a number of wonderful individuals — my thanks and appreciation to all of them for making this thesis possible.

First and foremost I would like to express my deep gratitude to my supervisor Prof. Dr. Erwin Zehe. Thank you for your guidance, for your general support and for the confidence you had in me and in my work. Erwin you were (are) an excellent mentor always making time for inspiring discussions, to share your expertise, and to answer my countless questions. You provided an environment without pressure in which I could work freely and enjoy science as a true learning process.

This thesis would not have been possible without the help of Sibylle Hassler, Jan Wienhöfer and Uwe Ehret who considerably improved my thesis as well as my journal publications by their proofreading and numerous valuable comments. Furthermore, I would like to thank Malte Neuper, Lucas Reid, Mirko Mälicke, Franziska Villinger and Rik van Pruijssen for their general support, for listing to all my complaints and for the various scientific and non-scientific discussions we had in the last years.

Special thank go to Conrad Jackisch, Axel Kleidon, Hoshin Gupta and Simon Seibert. Your expertise helped me as well as challenged my own thinking. Your drive to learn and to improve motivated me every time to do better science when we had the chance to have a discussion or work together.

I gratefully acknowledge the German Science Foundation (DFG) for partly funding my PhD thesis and the Open Access Publishing Fund of Karlsruhe Institute of Technology for covering the processing charges of my articles.

I would also like to thank my partner Nina and my son Emil. Nina you are a wonderful person and I am very grateful to have you in my life. Life is not always simple and smooth but I am sure with you on my side we will make the best out of it. Last but not least I would like to thank my entire family, particularly my mother. Without you I would have listened to my teachers and would not have had the courage to graduate school and go to university after I failed so many times. I still can hardly read or write a sentence without making plenty of mistakes but fortunately good grades and correct spelling do not make a good scientist.

# OWN PUBLICATIONS

## FIRST AUTHOR; PEER-REVIEWED INTERNATIONAL PUBLICATIONS

**Loritz, R.**, Kleidon, A., Jackisch, C., Westhoff, M., Ehret, U., Gupta, H., and Zehe, E. (2019): A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation. Hydrology and Earth System Sciences, 23(9), 3807–3821. doi:10.5194/hess-23-3807-2019

**Loritz, R.**, Gupta, H., Jackisch, C., Westhoff, M., Kleidon, A., Ehret, U., and Zehe, E. (2018): On the dynamic nature of hydrological similarity. Hydrology and Earth System Sciences, 22(7), 3663-3684. doi: 10.5194/hess-22-3663-2018

**Loritz, R.**, Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J. and Zehe, E. (2017): Picturing and modeling catchments by representative hillslopes, Hydrol. Earth Syst. Sci., 21(2), 1225-1249, doi: 10.5194/hess-21-1225-2017

## CO AUTHOR; PEER-REVIEWED INTERNATIONAL PUBLICATIONS

Zehe, E., **Loritz, R.**, Jackisch, C., Westhoff, M., Kleidon, A., Blume, T., et al. (2019). Energy states of soil water – a thermodynamic perspective on soil water dynamics and storage-controlled streamflow generation in different landscapes. Hydrology and Earth System Sciences, 23(2), 971–987. doi: 10.5194/hess-23-971-2019

Sternagel, A., **Loritz, R.**, Wilcke, W., and Zehe, E. (2019). Simulating preferential soil water flow and tracer transport using the Lagrangian Soil Water and Solute Transport Model. Hydrology and Earth System Sciences Discussions, (April), 1-25. doi: 10.5194/hess-2019-132

## CONFERENCE CONTRIBUTIONS (ORAL)

**Loritz, R.** (2018, Invited) - Is there an optimum? How information theory and thermodynamics can help to identify the needed spatial complexity of hydrological models. Workshop - Improving the theoretical underpinnings of hydrologic models - Sopron, Hungary

**Loritz, R.**, H. Gupta, C. Jackisch, M. Westhoff, A. Kleidon, U. Ehret, and Zehe, E. (2018): On the dynamic nature of hydrological similarity. European Geosciences Union General Assembly 2018, Vienna. Austria

**Loritz, R.**, Neuper, M., Gupta, H., and Zehe, E. (2017): How to handle spatial heterogeneity in hydrological models. European Geosciences Union General Assembly 2017, Vienna. Austria

**Loritz, R.**, Hassler, S., Jackisch, C., and Zehe, E. (2016): Picturing and modelling catchments by representative hillslopes. European Geosciences Union General Assembly 2016, Vienna. Austria

## Conference contributions (poster)

**Loritz, R**., Zehe, E. (2019): Modelkomplexität und räumliche Heterogenität - Wo ist der Zusammenhang. Tag der Hydrologie 2019, Karlsruhe, Germany

Kiese, N., **Loritz, R.**, Allroggen, N. and Zehe, E. (2017): Influence of bedrock topography on the runoff generation under use of ERT data. European Geosciences Union General Assembly 2017. Vienna. Austria

Wienhöfer, J., **R. Loritz** and Zehe, E. (2015): Model-based quantification of hydrological similarity at the hillslope scale. European Geosciences Union General Assembly 2015, Vienna. Austria

**Loritz, R**., Weiler, M. and Seibert, S. (2015): Data mining methods for predicting event runoff coefficients in ungauged basins using static and dynamic characteristics. European Geosciences Union General Assembly 2015, Vienna. Austria

# DECLARATION

*Authors contributions:*

**Chapter 2** Ralf Loritz et al. (2017). Picturing and modeling catchments by representative hillslopes, Hydrol. Earth Syst. Sci., 21(2), 1225-1249, doi:10.5194/hess-21-1225-2017

Ralf Loritz (RL) carried out the hydrological modeling using CATFLOW in consultation with Jan Wienhöfer(JW) and Erwin Zehe (EZ). RL developed and implemented the general approach and analyzed the resulting data by developing R and Python codes. RL wrote the manuscript, whereby EZ, JW and Sibylle Hassler (SH) significantly improved the manuscript. The final manuscript was reviewed by all remaining authors. RL accompanied it through the review process with help of EZ.

**Chapter 3**: Ralf Loritz et al. (2018). On the dynamic nature of hydrological similarity. Hydrology and Earth System Sciences, 22(7), 3663-3684. doi:10.5194/hess-22-3663-2018

RL wrote the paper, carried out the analysis and developed the underlying idea together with EZ. Hoshin Gupta (HG) and Uwe Ehret (UE) contributed to the theoretical framework and helped with the information theory related part. The final results and manuscript was reviewed and interpreted by all authors. RL accompanied it through the review process.

**Chapter 4**: Ralf Loritz et al. (2019). A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation. Hydrology and Earth System Sciences, 23(9), 3807–3821. doi:10.5194/hess-23-3807-2019

RL wrote the paper, carried out the analysis and developed the underlying theoretical framework together with EZ. Axel Kleidon (AK) and Martijn Westhoff (MW) contributed to the theoretical framework and helped with the thermodynamic interpretation. The final results and manuscript was reviewed and interpreted by all authors. RL accompanied it through the review process.

Eidesstattliche Versicherung gemäß §6 Abs. 1 Ziff. 4 der Promotionsordnung des Karlsruher Instituts für Technologie für die Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften:

1. Bei der eingereichten Dissertation zu dem Thema *The role of energy and information in hydrological modeling* handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

*Karlsruhe, 2019*

Ralf Loritz

## COLOPHON

This document was typeset using the typographical look-and-feel classicthesis developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". classicthesis is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of classicthesis usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

Thank you very much for your feedback and contribution.

*Final Version* as of January 22, 2020 (classicthesis v4.6).